

UC Irvine

UC Irvine Previously Published Works

Title

Channel Acquisition for Massive MIMO-OFDM with Adjustable Phase Shift Pilots

Permalink

<https://escholarship.org/uc/item/4f3020j2>

Journal

IEEE Transactions on Signal Processing, 64(6)

ISSN

1053-587X

Authors

You, Li
Gao, Xiqi
Swindlehurst, A Lee
et al.

Publication Date

2016-03-15

DOI

10.1109/tsp.2015.2502550

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Channel Acquisition for Massive MIMO-OFDM with Adjustable Phase Shift Pilots

Li You, *Student Member, IEEE*, Xiqi Gao, *Fellow, IEEE*, A. Lee Swindlehurst, *Fellow, IEEE*, and Wen Zhong

Abstract—We propose adjustable phase shift pilots (APSPs) for channel acquisition in wideband massive multiple-input multiple-output (MIMO) systems employing orthogonal frequency division multiplexing (OFDM) to reduce the pilot overhead. Based on a physically motivated channel model, we first establish a relationship between channel space-frequency correlations and the channel power angle-delay spectrum in the massive antenna array regime, which reveals the channel sparsity in massive MIMO-OFDM. With this channel model, we then investigate channel acquisition, including channel estimation and channel prediction, for massive MIMO-OFDM with APSPs. We show that channel acquisition performance in terms of sum mean square error can be minimized if the user terminals' channel power distributions in the angle-delay domain can be made non-overlapping with proper phase shift scheduling. A simplified pilot phase shift scheduling algorithm is developed based on this optimal channel acquisition condition. The performance of APSPs is investigated for both one symbol and multiple symbol data models. Simulations demonstrate that the proposed APSP approach can provide substantial performance gains in terms of achievable spectral efficiency over the conventional phase shift orthogonal pilot approach in typical mobility scenarios.

Index Terms—Adjustable phase shift pilots, massive MIMO-OFDM, channel estimation, channel prediction, channel acquisition, pilot phase shift scheduling.

I. INTRODUCTION

FORTHCOMING 5G cellular wireless systems are expected to support 1000 times faster data rates than the currently deployed 4G long-term evolution (LTE) system. To achieve the high data rates required by 5G, many technologies have been proposed [1]–[3]. Among them, massive multiple-input multiple-output (MIMO) systems, which deploy unprecedented numbers of antennas at the base stations (BSs) to simultaneously serve a relatively large number of user terminals (UTs), are believed to be one of the key candidate technologies for 5G [4]–[6].

Copyright © 2015 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

This work was supported in part by the National Natural Science Foundation of China under Grants 61471113, 61320106003, and 61201171, the China High-Tech 863 Plan under Grants 2015AA01A701 and 2014AA01A704, the National Science and Technology Major Project of China under Grant 2014ZX03003006-003, and the Program for Jiangsu Innovation Team. The work of L. You was supported in part by the China Scholarship Council (CSC). This work was presented in part at the IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Stockholm, Sweden, 2015.

L. You, X. Q. Gao, and W. Zhong are with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: liyou@seu.edu.cn; xqgao@seu.edu.cn; wzhong@seu.edu.cn).

A. L. Swindlehurst is with the Center for Pervasive Communications and Computing (CPCC), University of California, Irvine, CA 92697 USA (e-mail: swindle@uci.edu).

Orthogonal frequency division multiplexing (OFDM) is a multi-carrier modulation technology suited for high data rate wideband wireless transmission [7], [8]. Due to its robustness to channel frequency selectivity and relatively efficient implementation, OFDM combined with massive MIMO is a promising technique for wideband massive MIMO transmission [4]. As in conventional MIMO-OFDM, the performance of massive MIMO-OFDM is highly dependant on the quality of the channel acquisition. Pilot design and channel acquisition for massive MIMO-OFDM is of great practical importance.

Optimal pilot design and channel acquisition for conventional MIMO-OFDM has been extensively investigated in the literature. The most common approach is to estimate the channel response in the delay domain, and optimal pilots sent from different transmit antennas are typically assumed to satisfy the phase shift orthogonality condition in both the single-user case [9]–[11] and the multi-user case [12]. Note that such phase shift orthogonal pilots (PSOPs) have been adopted in LTE [13]. When channel spatial correlations are taken into account, optimal pilot design has been investigated for both the single-user case [14] and multi-user case [15]. Although these orthogonal pilot approaches can eliminate pilot interference in the same cell, they do not take into account the pilot overhead issue, which is thought to be one of the limiting factors for throughput in massive MIMO-OFDM [4]. When such approaches are directly adopted in time-division duplex (TDD) massive MIMO-OFDM, the corresponding pilot overhead is proportional to the sum of the number of UT antennas, and would be prohibitively large as the number of UTs becomes large. This becomes the system bottleneck, especially in high mobility scenarios where pilots must be transmitted more frequently. Therefore, a pilot approach that takes the pilot overhead issue into account is of importance for massive MIMO-OFDM systems.

In this paper, we propose adjustable phase shift pilots (APSPs) for massive MIMO-OFDM to reduce the pilot overhead. For APSPs, one sequence along with different adjustable phase shifted versions of itself in the frequency domain are adopted as pilots for different UTs. The proposed APSPs are different from conventional PSOPs [9], [10], [12], in which phase shifts for different pilots are fixed, and phase shift differences between different pilots are no less than the maximum channel delay (divided by the system sampling duration) of all the UTs. Since in our approach the phase shifts for different pilots are adjustable, more pilots are available compared with conventional PSOPs, which leads to significantly reduced pilot overhead.

The proposed APSPs exploit the following two channel

properties: First, wireless channels are sparse in many typical propagation scenarios; most channel power is concentrated in a finite region of delays and/or angles due to limited scattering [16]–[19]. Such channel sparsity can be resolved in the angle domain in massive MIMO due to the relatively large antenna array apertures, which has been observed in recent massive MIMO channel measurement results [20], [21]. Second, channel sparsity patterns, i.e., channel power distributions in the angle-delay domain, for different UTs are usually different.¹ For APSPs, when the phase shifts for pilots employed by different UTs are properly scheduled according to the above channel properties, channel acquisition can be achieved simultaneously in an almost interference-free manner as with conventional PSOPs. There has recently been increased research interest on utilizing channel sparsity for channel acquisition in massive MIMO. For instance, a time-frequency training scheme [25] and a distributed Bayesian channel estimation scheme [24] were proposed for massive MIMO-OFDM by exploiting the channel sparsity. As the approaches in [24] and [25] focus on channel acquisition for a single UT, the corresponding pilot overhead would still grow linearly with the number of UTs. Channel sparsity has also been exploited to mitigate pilot contamination in multi-cell massive MIMO [26], [27]. Note that compressive sensing has been applied to sparse channel acquisition in some recent works (see, e.g., [19], [22], [23], [28] and references therein), in which the corresponding pilot signals are usually assumed to be randomly generated. However, it is usually quite difficult to implement random pilot signals in practical systems [29]. For example, adopting large dimensional random pilot signals in the massive MIMO-OFDM systems considered here requires huge storage space and high complexity channel acquisition algorithms. In addition, a low peak-to-average power ratio (PAPR) for randomly generated pilot signals usually cannot be guaranteed. These drawbacks can be mitigated via proper design of the deterministic sensing matrices (see, e.g., [30], [31] and references therein).

The main contributions of this paper are summarized as follows:

- Based on a physically motivated channel model, we establish a relationship between the space-frequency domain channel covariance matrix (SFCCM) and the channel power angle-delay spectrum for massive MIMO-OFDM. We show that when the number of BS antennas is sufficiently large, the eigenvectors of the SFCCMs for different UTs tend to be equal, while the eigenvalues depend on the respective channel power angle-delay spectra, which reveals the channel sparsity in the angle-delay domain. Then we propose the angle-delay domain channel response matrix (ADCRM) and the corresponding angle-delay domain channel power matrix (ADCPM), which can model the massive MIMO-OFDM channel sparsity

¹There has been recent work that considers channels with a sparse common support [22], [23]. However, for massive MIMO channels, the common support assumption might not hold due to the increased angle resolution [22], [24]. Thus, in this work we assume that the channel sparsity patterns of different UTs are different (but not necessarily totally different), although the proposed APSP approach can also be applied to the common support cases.

in the angle-delay domain, and are convenient for further analyses.

- With the presented channel model, we propose APSP-based channel acquisition (APSP-CA) for massive MIMO-OFDM in TDD mode. For APSPs, equivalent channels for different UTs will experience corresponding cyclic shifts in the delay domain. Using this property, we show that the sum mean square error (MSE) of channel estimation (MSE-CE) can be minimized if the UTs' channel power distributions in the angle-delay domain can be made non-overlapping with proper pilot phase shift scheduling. Taking the time-varying nature of the channel into account, we further investigate channel prediction during the data segment using the received pilot signals. We show that the sum MSE of channel prediction (MSE-CP) can also be minimized if the UTs' channel power distributions in the angle-delay domain can be made non-overlapping with proper pilot phase shift scheduling, which coincides with the optimal channel estimation condition. A simplified pilot phase shift scheduling algorithm is developed based on this optimal channel acquisition condition. The proposed APSP-CA approach is investigated for cases involving both one symbol and multiple consecutive symbols.
- The proposed APSP-CA is evaluated in several typical propagation scenarios, and significant performance gains in terms of achievable spectral efficiency over the conventional PSOP-based channel acquisition (PSOP-CA) are demonstrated, especially in high mobility scenarios.

Portions of this work previously appeared in the conference paper [32].

A. Notations

We adopt the following notation throughout the paper. We use $\bar{j} = \sqrt{-1}$ to denote the imaginary unit. $\lfloor x \rfloor$ ($\lceil x \rceil$) denotes the largest (smallest) integer not greater (smaller) than x . $\langle \cdot \rangle_N$ denotes the modulo- N operation. $\delta(\cdot)$ denotes the delta function. Upper (lower) case boldface letters denote matrices (column vectors). The notation \triangleq is used for definitions. Notations \sim and \propto represent “distributed as” and “proportional to”, respectively. We adopt \mathbf{I}_N to denote the $N \times N$ dimensional identity matrix, and $\mathbf{I}_{N \times G}$ to denote the matrix composed of the first G ($\leq N$) columns of \mathbf{I}_N . We adopt $\mathbf{0}$ to denote the all-zero vector or matrix. The superscripts $(\cdot)^H$, $(\cdot)^T$, and $(\cdot)^*$ denote the conjugate-transpose, transpose, and conjugate operations, respectively. The operator $\text{diag}\{\mathbf{x}\}$ denotes the diagonal matrix with \mathbf{x} along its main diagonal. We employ $[\mathbf{a}]_i$, $[\mathbf{A}]_{i,j}$ and $[\mathbf{A}]_{i,:}$ to denote the i th element of the vector \mathbf{a} , the (i, j) th element of the matrix \mathbf{A} and the i th row of the matrix \mathbf{A} , respectively, where the element indices start with 0. $\mathbb{C}^{M \times N}$ ($\mathbb{R}^{M \times N}$) denotes the $M \times N$ dimensional complex (real) vector space. $\mathbb{E}\{\cdot\}$ denotes the expectation operation. $\mathcal{CN}(\mathbf{a}, \mathbf{B})$ denotes the circular symmetric complex Gaussian distribution with mean \mathbf{a} and covariance \mathbf{B} . \otimes and \odot denote the Kronecker product and Hadamard product, respectively. $\text{vec}\{\cdot\}$ represents the vectorization operation. \mathbf{F}_N denotes the N -dimensional unitary discrete Fourier transform

(DFT) matrix. $\mathbf{F}_{N \times G}$ denotes the matrix composed of the first $G (\leq N)$ columns of \mathbf{F}_N . $\mathbf{f}_{N,q}$ denotes the q th column of the matrix $\sqrt{N}\mathbf{F}_N$. We further define the permutation matrix $\mathbf{\Pi}_N^n \triangleq \begin{bmatrix} \mathbf{0} & \mathbf{I}_{N-(n)} \\ \mathbf{I}_{(n)} & \mathbf{0} \end{bmatrix}$. The notation \setminus denotes the set subtraction operation.

B. Outline

The rest of the paper is organized as follows. In Section II, we investigate the sparse nature of the massive MIMO-OFDM channel model. In Section III, we propose APSP-CA over one OFDM symbol in massive MIMO-OFDM, including channel estimation and prediction. We investigate the multiple consecutive pilot symbol case in Section IV. Simulation results are presented in Section V, and conclusions are given in Section VI.

II. MASSIVE MIMO-OFDM CHANNEL MODEL

In this section, we propose a physically motivated massive MIMO-OFDM channel model, and investigate the inherent channel sparsity property. We consider a single-cell TDD wideband massive MIMO wireless system which consists of one BS equipped with M antennas and K single-antenna UTs. We denote the UT set as $\mathcal{K} = \{0, 1, \dots, K-1\}$ where $k \in \mathcal{K}$ represents the UT index. We assume that the channels of different UTs are statistically independent. We assume that the BS is equipped with a one-dimensional uniform linear array (ULA),² with antennas separated by one-half wavelength. Then the BS array response vector corresponding to the incidence angle θ with respect to the perpendicular to the array is given by [17]

$$\mathbf{v}_{M,\theta} = \begin{bmatrix} 1 & \exp(-j\pi \sin(\theta)) & \dots \\ \dots & \exp(-j\pi(M-1)\sin(\theta)) \end{bmatrix}^T \in \mathbb{C}^{M \times 1}. \quad (1)$$

We assume that the signals seen at the BS are constrained to lie in the angle interval $\mathcal{A} = [-\pi/2, \pi/2]$, which can be achieved through the use of directional antennas at the BS, and thus no signal is received at the BS for incidence angles $\theta \notin \mathcal{A}$ [33].

We consider OFDM modulation with N_c subcarriers, performed via the N_c -point inverse DFT operation, appended with a guard interval (a.k.a. cyclic prefix) of length $N_g (\leq N_c)$ samples. We employ $T_{\text{sym}} = (N_c + N_g)T_s$ and $T_c = N_c T_s$ to denote the OFDM symbol duration with and without the guard interval, respectively, where T_s is the system sampling duration [13]. We assume that the guard interval length $T_g = N_g T_s$ is longer than the maximum channel delay of all the UTs [34], [35].

We assume that the channels remain constant during one OFDM symbol, and evolve from symbol to symbol. We denote the uplink (UL) channel gain between the antenna of the k th UT and the m th antenna of the BS over OFDM symbol ℓ and subcarrier n as $[\mathbf{g}_{k,\ell,n}]_m$. Using a physical channel modeling

approach (see, e.g., [17], [36]–[39]), the channel response vector $\mathbf{g}_{k,\ell,n} \in \mathbb{C}^{M \times 1}$ can be described as

$$\begin{aligned} \mathbf{g}_{k,\ell,n} &= \sum_{q=0}^{N_g-1} \int_{-\infty}^{\infty} \int_{-\pi/2}^{\pi/2} \mathbf{v}_{M,\theta} \cdot \exp\left(-j2\pi \frac{n}{T_c} \tau\right) \\ &\quad \cdot \exp(j2\pi \nu \ell T_{\text{sym}}) \cdot g_k(\theta, \tau, \nu) \cdot \delta(\tau - qT_s) d\theta d\nu \\ &= \sum_{q=0}^{N_g-1} \int_{-\infty}^{\infty} \int_{-\pi/2}^{\pi/2} \mathbf{v}_{M,\theta} \cdot \exp\left(-j2\pi \frac{n}{N_c} q\right) \\ &\quad \cdot \exp(j2\pi \nu \ell T_{\text{sym}}) \cdot g_k(\theta, qT_s, \nu) d\theta d\nu \end{aligned} \quad (2)$$

where $\mathbf{v}_{M,\theta}$ is given in (1), $g_k(\theta, \tau, \nu)$ is the complex-valued joint angle-delay-Doppler channel gain function of UT k corresponding to the incidence angle θ , delay τ , and Doppler frequency ν . Note that the number of significant channel taps in the delay domain is usually limited, and smaller than N_g ; i.e., $|g_k(\theta, qT_s, \nu)|$ is approximately 0 for most q . Since the locations of the significant channel taps in the delay domain are usually different for different UTs, we adopt (2) in this paper to obtain a general channel representation applicable for all the UTs.

We write the k th UT's channel at OFDM symbol ℓ over all subcarriers as

$$\mathbf{G}_{k,\ell} = [\mathbf{g}_{k,\ell,0} \quad \mathbf{g}_{k,\ell,1} \quad \dots \quad \mathbf{g}_{k,\ell,N_c-1}] \in \mathbb{C}^{M \times N_c} \quad (3)$$

which will be referred to as the space-frequency domain channel response matrix (SFCRM). From (2), it is not hard to show that

$$\text{vec}\{\mathbf{G}_{k,\ell}\} = \sum_{q=0}^{N_g-1} \int_{-\infty}^{\infty} \int_{-\pi/2}^{\pi/2} [\mathbf{f}_{N_c,q} \otimes \mathbf{v}_{M,\theta}] \cdot \exp(j2\pi \nu \ell T_{\text{sym}}) \cdot g_k(\theta, qT_s, \nu) d\theta d\nu. \quad (4)$$

We assume that channels with different incidence angles, delays, and/or Doppler frequencies are uncorrelated [17], [38], [39]. We also assume that the temporal correlations and joint space-frequency domain correlations of the channels can be separated [35], [38], i.e.,

$$\begin{aligned} &\mathbb{E}\{g_k(\theta, \tau, \nu) g_k^*(\theta', \tau', \nu')\} \\ &= \mathbf{S}_k^{\text{ADD}}(\theta, \tau, \nu) \cdot \delta(\theta - \theta') \delta(\tau - \tau') \delta(\nu - \nu') \\ &= \mathbf{S}_k^{\text{AD}}(\theta, \tau) \cdot \mathbf{S}_k^{\text{Dop}}(\nu) \cdot \delta(\theta - \theta') \delta(\tau - \tau') \delta(\nu - \nu') \end{aligned} \quad (5)$$

where $\mathbf{S}_k^{\text{ADD}}(\theta, \tau, \nu)$, $\mathbf{S}_k^{\text{AD}}(\theta, \tau)$, and $\mathbf{S}_k^{\text{Dop}}(\nu)$ represent the power angle-delay-Doppler spectrum, power angle-delay spectrum, and power Doppler spectrum of UT k , respectively [17], [40].

From (4) and (5), we can obtain the following channel statistical property (see Appendix A for the derivations)

$$\mathbb{E}\{\text{vec}\{\mathbf{G}_{k,\ell+\Delta_\ell}\} \text{vec}^H\{\mathbf{G}_{k,\ell}\}\} = \varrho_k(\Delta_\ell) \cdot \mathbf{R}_k \quad (6)$$

where $\varrho_k(\Delta_\ell)$ is the channel temporal correlation function (TCF) given by

$$\varrho_k(\Delta_\ell) \triangleq \int_{-\infty}^{\infty} \exp(j2\pi \nu \Delta_\ell T_{\text{sym}}) \cdot \mathbf{S}_k^{\text{Dop}}(\nu) d\nu \quad (7)$$

²We adopt the ULA model in this paper for clarity, although our work can be readily extended to more general antenna array models using the techniques in [33].

and \mathbf{R}_k is the space-frequency domain channel covariance matrix (SFCCM) given by

$$\mathbf{R}_k \triangleq \sum_{q=0}^{N_g-1} \int_{-\pi/2}^{\pi/2} [\mathbf{f}_{N_c, q} \otimes \mathbf{v}_{M, \theta}] [\mathbf{f}_{N_c, q} \otimes \mathbf{v}_{M, \theta}]^H \cdot \mathbf{S}_k^{\text{AD}}(\theta, qT_s) d\theta \in \mathbb{C}^{MN_c \times MN_c}. \quad (8)$$

In this work, we consider the widely accepted Clarke-Jakes channel power Doppler spectrum,³ with the corresponding channel TCF given by [40], [41]

$$\varrho_k(\Delta_\ell) = J_0(2\pi\nu_k T_{\text{sym}} \Delta_\ell) \quad (9)$$

where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind, and ν_k is the Doppler frequency of UT k . Note that the Clarke-Jakes power Doppler spectrum is an even function, i.e., $\varrho_k(\Delta_\ell) = \varrho_k(-\Delta_\ell)$, and satisfies $\varrho_k(0) = 1$. Also, we assume that according to the law of large numbers, the channel elements exhibit a joint Gaussian distribution, i.e., $\text{vec}\{\mathbf{G}_{k, \ell}\} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_k)$.

Before proceeding, we investigate in the following proposition a property of the large dimensional SFCCM, and present a relationship between the SFCCM and the power angle-delay spectrum for massive MIMO-OFDM channels.

Proposition 1: Define $\mathbf{V}_M \in \mathbb{C}^{M \times M}$ as $[\mathbf{V}_M]_{i, j} \triangleq \frac{1}{\sqrt{M}} \exp\left(-j2\pi \frac{i(j-M/2)}{M}\right)$, and $\mathbf{\Omega}_k \in \mathbb{R}^{M \times N_g}$ as

$$[\mathbf{\Omega}_k]_{i, j} \triangleq MN_c(\theta_{i+1} - \theta_i) \cdot \mathbf{S}_k^{\text{AD}}(\theta_i, \tau_j) \quad (10)$$

where $\theta_m \triangleq \arcsin(2m/M - 1)$, and $\tau_n \triangleq nT_s$. Then when the number of antennas $M \rightarrow \infty$, the SFCCM \mathbf{R}_k tends to $(\mathbf{F}_{N_c \times N_g} \otimes \mathbf{V}_M) \text{diag}\{\text{vec}\{\mathbf{\Omega}_k\}\} (\mathbf{F}_{N_c \times N_g} \otimes \mathbf{V}_M)^H$ in the sense that, for fixed non-negative integers i and j ,

$$\lim_{M \rightarrow \infty} \left[\mathbf{R}_k - (\mathbf{F}_{N_c \times N_g} \otimes \mathbf{V}_M) \text{diag}\{\text{vec}\{\mathbf{\Omega}_k\}\} \cdot (\mathbf{F}_{N_c \times N_g} \otimes \mathbf{V}_M)^H \right]_{i, j} = 0. \quad (11)$$

Proof: See Appendix B. ■

The relationship between the space-frequency domain channel joint correlation property and the channel power distribution in the angle-delay domain for massive MIMO-OFDM is established in Proposition 1. Specifically, for massive MIMO-OFDM channels in the asymptotically large array regime, the eigenvectors of the SFCCMs for different UTs tend to be the same, which shows that massive MIMO-OFDM channels can be asymptotically decorrelated by the fixed space-frequency domain statistical eigendirections, while the eigenvalues depend on the corresponding channel power angle-delay spectra.

Proposition 1 indicates that, for massive MIMO-OFDM channels, when the number of BS antennas M is sufficiently large, the SFCCM can be well approximated by

$$\mathbf{R}_k \approx (\mathbf{F}_{N_c \times N_g} \otimes \mathbf{V}_M) \text{diag}\{\text{vec}\{\mathbf{\Omega}_k\}\}$$

³Although the waves impinging on the BS are assumed to be sparsely distributed in the angle domain due to limited scattering around the BS (typically mounted at an elevated position), the waves departing the mobile UTs are usually uniformly distributed in angle of departure. Thus the Clarke-Jakes spectrum is suitable to model the time variation of the channel [40], [41].

$$\cdot (\mathbf{F}_{N_c \times N_g} \otimes \mathbf{V}_M)^H. \quad (12)$$

It is worth noting that the approximation in (12) is consistent with existing results in the literature. For frequency-selective single-input single-output channels, (12) agrees with the results in [35], [42]. For frequency-flat massive MIMO channels, the approximation given in (12) has been shown to be accurate enough for a practical number of antennas, which usually ranges from 64 to 512 [27], [33], [43], [44], and a detailed numerical example can be found in [27]. Since the SFCCM model given in (12) is a good approximation to the more complex physical channel model in (8) when the number of BS antennas is sufficiently large, we will thus exclusively use the simplified SFCCM model in (12) in the rest of the paper.

Realistic wireless channels are usually not wide-sense stationary [17], i.e., \mathbf{R}_k varies as time evolves, although with a relatively large time scale.⁴ In practice, acquisition of the large dimensional \mathbf{R}_k is rather difficult and resource-intensive for massive MIMO-OFDM. However, when we shift our focus from the space-frequency domain to the angle-delay domain, the problem can be significantly simplified. Motivated by the eigenvalue decomposition of the SFCCM given in (12), we decompose the SFCCM as follows

$$\mathbf{G}_{k, \ell} = \mathbf{V}_M \mathbf{H}_{k, \ell} \mathbf{F}_{N_c \times N_g}^T \quad (13)$$

where

$$\mathbf{H}_{k, \ell} = \mathbf{V}_M^H \mathbf{G}_{k, \ell} \mathbf{F}_{N_c \times N_g}^* \in \mathbb{C}^{M \times N_g} \quad (14)$$

is referred to as the angle-delay domain channel response matrix (ADCRM) of UT k at OFDM symbol ℓ . In the following proposition, we derive a statistical property of the ADCRM.

Proposition 2: For massive MIMO-OFDM channels, when the number of antennas $M \rightarrow \infty$, elements of the ADCRM $\mathbf{H}_{k, \ell}$ satisfy

$$\begin{aligned} & \mathbb{E} \left\{ [\mathbf{H}_{k, \ell + \Delta_\ell}]_{i, j} [\mathbf{H}_{k, \ell'}^*]_{i', j'} \right\} \\ & = \varrho_k(\Delta_\ell) \delta(i - i') \delta(j - j') \cdot [\mathbf{\Omega}_k]_{i, j} \end{aligned} \quad (15)$$

where $\mathbf{\Omega}_k$ is given in (10).

Proof: See Appendix C. ■

Proposition 2 shows that, for massive MIMO-OFDM channels, different elements of the ADCRM $\mathbf{H}_{k, \ell}$ are approximately mutually statistically uncorrelated, which lends the ADCRM in (14) its physical interpretation. Specifically, different elements of the ADCRM correspond to the channel gains for different incidence angles and delays, which can be resolved in massive MIMO-OFDM with a sufficiently large antenna array aperture. Note that $[\mathbf{\Omega}_k]_{i, j}$ corresponds to the average power of $[\mathbf{H}_k]_{i, j}$, and can describe the sparsity of the wireless channels in the angle-delay domain. Hereafter we will refer to $\mathbf{\Omega}_k$ as the angle-delay domain channel power matrix (ADCPM) of UT k . The dimension of the ADCPM $\mathbf{\Omega}_k$ is much smaller than that of the SFCCM \mathbf{R}_k , and most elements in $\mathbf{\Omega}_k$ are approximately zero due to the channel sparsity. In addition, $\mathbf{\Omega}_k$ is composed of the variances of

⁴The degree of channel stationarity depends on the propagation scenarios. In typical scenarios, the channel statistics vary on the order of seconds [45], while the OFDM symbol length is usually on the order of millisecond [46].

independent angle-delay domain channel elements, and thus can be estimated in an element-wise manner. Therefore, in practice there will be enough resources for one to obtain an estimate of $\mathbf{\Omega}_k$ with guaranteed accuracy. In the rest of the paper, we will assume that the ADCPMs of all the UTs are known by the BS.

Before we conclude this section, we define the extended ADCRM as follows

$$\begin{aligned}\bar{\mathbf{H}}_{k,\ell,(N_c)} &\triangleq \mathbf{H}_{k,\ell} \mathbf{I}_{N_c \times N_g}^T \\ &= [\mathbf{H}_{k,\ell} \quad \mathbf{0}_{M \times (N_c - N_g)}] \in \mathbb{C}^{M \times N_c}.\end{aligned}\quad (16)$$

Similarly, the extended ADCPM, which corresponds to the power distribution of the extended ADCRM $\bar{\mathbf{H}}_{k,\ell,(N_c)}$, is defined as

$$\begin{aligned}\bar{\mathbf{\Omega}}_{k,(N_c)} &\triangleq \mathbf{\Omega}_k \mathbf{I}_{N_c \times N_g}^T \\ &= [\mathbf{\Omega}_k \quad \mathbf{0}_{M \times (N_c - N_g)}] \in \mathbb{R}^{M \times N_c}.\end{aligned}\quad (17)$$

Such definitions will be employed to simplify the analyses in the following sections.

III. CHANNEL ACQUISITION WITH APSPS OVER ONE SYMBOL

Based on the sparse massive MIMO-OFDM channel model presented in the previous section, we propose APSP-CA for massive MIMO-OFDM, including channel estimation and prediction. In this section, we first investigate the case where the APSPs are sent over one OFDM symbol, while the multiple symbol case will be investigated in the next section.

A. APSPs over One Symbol

We assume that all the UTs are synchronized. During the UL pilot segment, namely, the ℓ th OFDM symbol of each frame, all the UTs transmit the scheduled pilots simultaneously, and the space-frequency domain signal received at the BS can be represented as

$$\mathbf{Y}_\ell = \sum_{k'=0}^{K-1} \mathbf{G}_{k',\ell} \mathbf{X}_{k'} + \mathbf{Z}_\ell \in \mathbb{C}^{M \times N_c} \quad (18)$$

where $[\mathbf{Y}_\ell]_{i,j}$ denotes the received pilot signal at the i th antenna over the j th subcarrier, $\mathbf{G}_{k,\ell}$ is the SFCRM defined in (3), $\mathbf{X}_k = \text{diag}\{\mathbf{x}_k\} \in \mathbb{C}^{N_c \times N_c}$ denotes the frequency domain pilot signal sent from the k th UT, \mathbf{Z}_ℓ is the additive white Gaussian noise (AWGN) matrix during the UL pilot segment with elements identically and independently distributed (i.i.d.) as $\mathcal{CN}(0, \sigma_{\text{ztr}})$, and σ_{ztr} is the noise power.

The proposed APSP over one OFDM symbol for a given UT k is given by

$$\mathbf{X}_k \triangleq \sqrt{\sigma_{\text{xtr}}} \underbrace{\text{diag}\{\mathbf{f}_{N_c, \phi_k}\}}_{\triangleq \mathbf{D}_{\phi_k}} \mathbf{X}, \quad \phi_k = 0, 1, \dots, N_c - 1 \quad (19)$$

where $\mathbf{X} = \text{diag}\{\mathbf{x}\} \in \mathbb{C}^{N_c \times N_c}$ satisfying $\mathbf{X}\mathbf{X}^H = \mathbf{I}_{N_c}$ is the basic pilot matrix shared by all UTs in the same cell, and σ_{xtr} is the pilot signal transmit power. The APSP signal given in (19) can be seen as a phase shifted version of $\sqrt{\sigma_{\text{xtr}}}\mathbf{X}$ with phase shift ϕ_k in the frequency domain. Note that the proposed APSP has the same PAPR as that of \mathbf{X} in the

time domain, thus existing low PAPR sequence designs can be easily incorporated into our approach. In addition, as the basic pilot matrix \mathbf{X} can be predetermined, only \mathbf{X} and the pilot phase shift indices rather than the entire pilot matrices are required to be stored, and the required storage space can be significantly reduced.

From (19), it can be readily obtained that, for $\forall k, k' \in \mathcal{K}$,

$$\mathbf{X}_{k'} \mathbf{X}_k^H = \sigma_{\text{xtr}} \mathbf{D}_{\phi_{k'} - \phi_k} \quad (20)$$

which indicates that cross correlations of the proposed APSPs for different UTs depend only on the associated phase shift difference. It is worth noting that, for conventional PSOPs, the phase shift differences for different pilots are set to satisfy the orthogonality condition $|\phi_{k'} - \phi_k| \geq N_g \quad \forall k' \neq k$. However, for our APSPs, the phase shifts for different pilots are adjustable, and pilots for different UTs can even share the same phase shift, which leads to more available pilots, and thus pilot overhead can be significantly reduced.

B. Channel Estimation with APSPs

In this section we investigate channel estimation during the pilot segment under the minimum MSE (MMSE) criterion using the proposed APSPs. Direct MMSE estimation of the SFCRM $\mathbf{G}_{k,\ell}$ requires information about the large dimensional SFCCM \mathbf{R}_k and a large dimensional matrix inversion, which is difficult to implement in practice. However, with the sparse massive MIMO-OFDM channel model presented above, when we shift our focus from the space-frequency domain to the angle-delay domain, channel estimation can be greatly simplified. The BS can first estimate the ADCRM to obtain $\bar{\mathbf{H}}_{k,\ell}$, then the SFCRM estimates can be readily obtained as $\hat{\mathbf{G}}_{k,\ell} = \mathbf{V}_M \bar{\mathbf{H}}_{k,\ell} \mathbf{F}_{N_c \times N_g}^T$ via exploiting the unitary equivalence between the angle-delay domain channels and the space-frequency domain channels given in (13), while the same MSE-CE performance can be maintained. In the following, we focus on estimation of the ADCRM $\mathbf{H}_{k,\ell}$ under the MMSE criterion.

Recalling (13), the received pilot signal at the BS in (18) can be rewritten as

$$\mathbf{Y}_\ell = \sum_{k'=0}^{K-1} \mathbf{V}_M \mathbf{H}_{k',\ell} \mathbf{F}_{N_c \times N_g}^T \mathbf{X}_{k'} + \mathbf{Z}_\ell. \quad (21)$$

After decorrelation and power normalization of \mathbf{Y}_ℓ , the BS can obtain an observation of the UL channel $\mathbf{H}_{k,\ell}$, given by (22) shown at the top of the next page, where (a) follows from (20). Using the unitary transformation property, it can be readily shown that the pilot noise term in (22) exhibits a Gaussian distribution with i.i.d. elements distributed as $\mathcal{CN}(0, \sigma_{\text{ztr}}/\sigma_{\text{xtr}})$, and (22) can be simplified as

$$\mathbf{Y}_{k,\ell} = \mathbf{H}_{k,\ell} + \sum_{k' \neq k} \mathbf{H}_{k',\ell}^{\phi_{k'} - \phi_k} + \frac{1}{\sqrt{\rho_{\text{tr}}}} \mathbf{Z}_{\text{iid}} \quad (23)$$

where $\rho_{\text{tr}} \triangleq \sigma_{\text{xtr}}/\sigma_{\text{ztr}}$ is the signal-to-noise ratio (SNR) during the pilot segment, and $\mathbf{Z}_{\text{iid}} \in \mathbb{C}^{M \times N_g}$ is the normalized AWGN matrix with i.i.d. elements distributed as $\mathcal{CN}(0, 1)$.

$$\begin{aligned}
\mathbf{Y}_{k,\ell} &= \frac{1}{\sigma_{\text{xtr}}} \mathbf{V}_M^H \mathbf{Y}_\ell \mathbf{X}_k^H \mathbf{F}_{N_c \times N_g}^* \\
&= \frac{1}{\sigma_{\text{xtr}}} \sum_{k'=0}^{K-1} \mathbf{H}_{k',\ell} \mathbf{F}_{N_c \times N_g}^T \mathbf{X}_{k'} \mathbf{X}_k^H \mathbf{F}_{N_c \times N_g}^* + \frac{1}{\sigma_{\text{xtr}}} \mathbf{V}_M^H \mathbf{Z}_\ell \mathbf{X}_k^H \mathbf{F}_{N_c \times N_g}^* \\
&\stackrel{(a)}{=} \underbrace{\mathbf{H}_{k,\ell} + \sum_{k' \neq k} \mathbf{H}_{k',\ell} \mathbf{F}_{N_c \times N_g}^T \mathbf{D}_{\phi_{k'} - \phi_k} \mathbf{F}_{N_c \times N_g}^*}_{\text{pilot interference} \triangleq \sum_{k' \neq k} \mathbf{H}_{k',\ell}^{\phi_{k'} - \phi_k}} + \underbrace{\frac{1}{\sigma_{\text{xtr}}} \mathbf{V}_M^H \mathbf{Z}_\ell \mathbf{X}_k^H \mathbf{F}_{N_c \times N_g}^*}_{\text{pilot noise}} \quad (22)
\end{aligned}$$

Note that the pilot interference term $\mathbf{H}_{k',\ell}^{\phi_{k'} - \phi_k}$ defined in (22) satisfies

$$\begin{aligned}
\mathbf{H}_{k',\ell}^{\phi_{k'} - \phi_k} &= \mathbf{H}_{k',\ell} \mathbf{F}_{N_c \times N_g}^T \mathbf{D}_{\phi_{k'} - \phi_k} \mathbf{F}_{N_c \times N_g}^* \\
&= \mathbf{H}_{k',\ell} \mathbf{I}_{N_c \times N_g}^T \mathbf{F}_{N_c}^T \mathbf{D}_{\phi_{k'} - \phi_k} \mathbf{F}_{N_c}^* \mathbf{I}_{N_c \times N_g} \\
&\stackrel{(a)}{=} \bar{\mathbf{H}}_{k',\ell,(N_c)} \mathbf{F}_{N_c}^T \mathbf{D}_{\phi_{k'} - \phi_k} \mathbf{F}_{N_c}^* \mathbf{I}_{N_c \times N_g} \\
&\stackrel{(b)}{=} \bar{\mathbf{H}}_{k',\ell,(N_c)} \mathbf{\Pi}_{N_c}^{\phi_{k'} - \phi_k} \mathbf{I}_{N_c \times N_g} \quad (24)
\end{aligned}$$

where (a) follows from (16), and (b) follows from the permutation matrix definition given in Section I-A. Thus, the pilot interference term $\mathbf{H}_{k',\ell}^{\phi_{k'} - \phi_k}$ in (23) is a column truncated version of the extended ADCRM $\bar{\mathbf{H}}_{k',\ell,(N_c)}$ with a cyclic column shift, where the shift factor depends on the corresponding pilot phase shift difference $\phi_{k'} - \phi_k$. Thus elements of $\mathbf{H}_{k',\ell}^{\phi_{k'} - \phi_k}$ can be readily obtained as

$$\left[\mathbf{H}_{k',\ell}^{\phi_{k'} - \phi_k} \right]_{i,j} = \begin{cases} \left[\mathbf{H}_{k',\ell} \right]_{i, (j - (\phi_{k'} - \phi_k))_{N_c}}, & \langle j - (\phi_{k'} - \phi_k) \rangle_{N_c} \leq N_g - 1 \\ 0, & \text{else.} \end{cases} \quad (25)$$

Recalling Proposition 2, elements of the ADCRM $\mathbf{H}_{k',\ell}$ are statistically uncorrelated. Consequently, elements of the pilot interference term $\mathbf{H}_{k',\ell}^{\phi_{k'} - \phi_k}$, a column truncated copy of $\mathbf{H}_{k',\ell}$ with cyclic column shift, are also statistically uncorrelated. Thus, using the same methodology as in the previous section, the corresponding power matrix of the pilot interference term $\mathbf{H}_{k',\ell}^{\phi_{k'} - \phi_k}$ can be defined as

$$\begin{aligned}
\mathbf{\Omega}_{k'}^{\phi_{k'} - \phi_k} &\triangleq \mathbb{E} \left\{ \mathbf{H}_{k',\ell}^{\phi_{k'} - \phi_k} \odot \left(\mathbf{H}_{k',\ell}^{\phi_{k'} - \phi_k} \right)^* \right\} \\
&= \bar{\mathbf{\Omega}}_{k',(N_c)} \mathbf{\Pi}_{N_c}^{\phi_{k'} - \phi_k} \mathbf{I}_{N_c \times N_g} \quad (26)
\end{aligned}$$

which is a column truncated version of the extended ADCPM $\bar{\mathbf{\Omega}}_{k',(N_c)}$ defined in (17) with cyclic column shift $\phi_{k'} - \phi_k$.

With the channel observation $\mathbf{Y}_{k,\ell}$ in (23), and the fact that the angle-delay domain channel elements are uncorrelated as derived in Proposition 2, the MMSE estimate $\hat{\mathbf{H}}_{k,\ell}$ can be obtained in an element-wise manner as follows [47]

$$\left[\hat{\mathbf{H}}_{k,\ell} \right]_{i,j} = \frac{[\mathbf{\Omega}_k]_{i,j}}{\sum_{k'=0}^{K-1} [\mathbf{\Omega}_{k'}^{\phi_{k'} - \phi_k}]_{i,j} + \frac{1}{\rho_{\text{tr}}}} [\mathbf{Y}_{k,\ell}]_{i,j}. \quad (27)$$

Let $\tilde{\mathbf{H}}_{k,\ell} = \mathbf{H}_{k,\ell} - \hat{\mathbf{H}}_{k,\ell}$ be the angle-delay domain channel estimation error of the k th UT, then the corresponding MSE-CE can be obtained as

$$\epsilon_k^{\text{CE}} \triangleq \sum_{i=0}^{M-1} \sum_{j=0}^{N_g-1} \mathbb{E} \left\{ \left| \left[\tilde{\mathbf{H}}_{k,\ell} \right]_{i,j} \right|^2 \right\}$$

$$\begin{aligned}
&\stackrel{(a)}{=} \sum_{i=0}^{M-1} \sum_{j=0}^{N_g-1} \mathbb{E} \left\{ \left| \left[\mathbf{H}_{k,\ell} \right]_{i,j} \right|^2 - \left| \left[\hat{\mathbf{H}}_{k,\ell} \right]_{i,j} \right|^2 \right\} \\
&= \sum_{i=0}^{M-1} \sum_{j=0}^{N_g-1} \left\{ [\mathbf{\Omega}_k]_{i,j} - \frac{[\mathbf{\Omega}_k]_{i,j}^2}{\sum_{k'=0}^{K-1} [\mathbf{\Omega}_{k'}^{\phi_{k'} - \phi_k}]_{i,j} + \frac{1}{\rho_{\text{tr}}}} \right\} \quad (28)
\end{aligned}$$

where (a) follows from the orthogonality principle of MMSE estimation [47].

Before we proceed, we define the sum MSE-CE of all the UTs as

$$\epsilon^{\text{CE}} \triangleq \sum_{k=0}^{K-1} \epsilon_k^{\text{CE}}. \quad (29)$$

Due to the incurred pilot interference, performance of the APSP-based channel estimation might deteriorate. However, we will show in the following proposition that such effects can be eliminated with proper phase shift scheduling for different pilots.

Proposition 3: The sum MSE-CE ϵ^{CE} is lower bounded by

$$\epsilon^{\text{CE}} \geq \varepsilon^{\text{CE}} = \sum_{k=0}^{K-1} \sum_{i=0}^{M-1} \sum_{j=0}^{N_g-1} \left\{ [\mathbf{\Omega}_k]_{i,j} - \frac{[\mathbf{\Omega}_k]_{i,j}^2}{[\mathbf{\Omega}_k]_{i,j} + \frac{1}{\rho_{\text{tr}}}} \right\} \quad (30)$$

and the lower bound can be achieved under the condition that, for $\forall k, k' \in \mathcal{K}$ and $k \neq k'$,

$$\left(\bar{\mathbf{\Omega}}_{k,(N_c)} \mathbf{\Pi}_{N_c}^{\phi_k} \right) \odot \left(\bar{\mathbf{\Omega}}_{k',(N_c)} \mathbf{\Pi}_{N_c}^{\phi_{k'}} \right) = \mathbf{0}. \quad (31)$$

Proof: See Appendix D. ■

Proposition 3 shows that with the proposed APSPs, the sum MSE-CE can be minimized when phase shifts for different pilots are properly scheduled according to the condition given in (31). The interpretation is very intuitive. With frequency domain phase shifted pilots, equivalent channels will exhibit corresponding cyclic shifts in the delay domain, as seen from (24). If the equivalent channel power distributions in the angle-delay domain for different UTs can be made non-overlapping after pilot phase shift scheduling, the pilot interference effect can be eliminated, and the sum MSE-CE can be minimized.

Wireless channels are approximately sparse in the angle-delay domain in many practical propagation scenarios, and typically only a few elements of the ADCPM $\mathbf{\Omega}_k$ are dominant in massive MIMO-OFDM. When such channel sparsity is properly taken into account, the equivalent angle-delay domain channels for different UTs are almost non-overlapping with

high probability, assuming proper pilot phase shifts. This suggests the feasibility of the proposed APSPs for massive MIMO-OFDM.

Note that performance of the proposed APSP approach is related to the channel sparsity level. For the case where channels of different UTs have a sparse common support with s ($\leq N_g$) representing the number of the columns containing non-zero elements in the ADCPM [22], [23], the maximum number of UTs that can be served without pilot interference is $\lfloor N_c/s \rfloor$. However, for practical wireless channels, most of the channel elements in the angle-delay domain are close to zero, and the condition in (31) usually cannot be satisfied exactly, which will lead to degradation of the channel acquisition performance. In such cases, it is clear that the more sparse the channels are, the better performance can be achieved by the proposed APSP approach.

Before we conclude this subsection, we remark here that several existing pilot approaches satisfy the optimal condition given in Proposition 3. For the case where channel sparsity property is not known, it is reasonable to assume that all the angle-delay domain channel elements are identically distributed, i.e., all the ADCPM elements are equal, in which case the optimal condition in (31) can be achieved when $|\phi_k - \phi_{k'}| \geq N_g$ for $\forall k \neq k'$, i.e., the extended channels in the delay domain for different UTs are totally separated, which coincides with the conventional PSOPs [12]. For frequency-flat massive MIMO channels, i.e., $N_c = 1$, the condition in (31) can be achieved when $\Omega_k \odot \Omega_{k'} = \mathbf{0}$ for $\forall k \neq k'$, i.e., different UTs can share the same pilot when the respective channels have non-overlapping support in the angle domain, which coincides with previous works such as [33], [43]. In our work, the proposed APSPs exploit the joint angle-delay domain channel sparsity in massive MIMO-OFDM, and are more efficient and general from the pilot overhead point of view.

C. Channel Prediction with APSPs

In the previous subsection, we investigated channel estimation during the pilot segment. Directly employing the pilot segment channel estimates in the data segment might not always be appropriate [48], especially in high mobility scenarios, which are the main focus of the APSPs. In this subsection, we investigate channel prediction during the data segment based on the received pilot signals, using the proposed APSPs.

For frame-based massive MIMO-OFDM transmission, the BS utilizes the received signals during the pilot segment to acquire the channels in the current frame. If the pilot segment channel estimate $\hat{\mathbf{H}}_{k,\ell}$ is directly employed as the estimate of the channel $\mathbf{H}_{k,\ell+\Delta_\ell}$ during the data segment, the corresponding sum MSE-CE for a given delay Δ_ℓ between the pilot symbol and data symbol can be written as

$$\begin{aligned} \epsilon^{\text{CE}}(\Delta_\ell) &= \sum_{k=0}^{K-1} \sum_{i=0}^{M-1} \sum_{j=0}^{N_g-1} \mathbb{E} \left\{ \left| \left[\mathbf{H}_{k,\ell+\Delta_\ell} - \hat{\mathbf{H}}_{k,\ell} \right]_{i,j} \right|^2 \right\} \\ &= \sum_{k=0}^{K-1} \sum_{i=0}^{M-1} \sum_{j=0}^{N_g-1} \left\{ [\Omega_k]_{i,j} + [1 - 2\varrho_k(\Delta_\ell)] \right. \end{aligned}$$

$$\left. \cdot \frac{[\Omega_k]_{i,j}^2}{\sum_{k'=0}^{K-1} \left[\Omega_{k'}^{\phi_{k'} - \phi_k} \right]_{i,j} + \frac{1}{\rho_{\text{tr}}}} \right\}. \quad (32)$$

In high mobility scenarios, the channel TCF satisfies $\varrho_k(\Delta_\ell) \rightarrow 0$ for relatively large delay $|\Delta_\ell|$. When $\varrho_k(\Delta_\ell) < 1/2$, i.e., $1 - 2\varrho_k(\Delta_\ell) > 0$, it can be observed from (32) that the sum MSE-CE expression $\epsilon^{\text{CE}}(\Delta_\ell)$ is even larger than the sum channel power $\sum_{k=0}^{K-1} \sum_{i=0}^{M-1} \sum_{j=0}^{N_g-1} [\Omega_k]_{i,j}$, and channel estimation performance cannot be guaranteed, which motivates the need for channel prediction.

For channel prediction, the BS utilizes the received pilot signals as well as the channel TCF to get estimates of the channels during the data segment. Under the MMSE criterion, with the angle-delay domain channel property of massive MIMO-OFDM given in Proposition 2, it is not hard to show that an estimate of the ADCRM $\mathbf{H}_{k,\ell+\Delta_\ell}$ based on $\mathbf{Y}_{k,\ell}$ can be obtained in an element-wise manner as follows

$$\begin{aligned} & \left[\hat{\mathbf{H}}_{k,\ell+\Delta_\ell} \right]_{i,j} \\ &= \varrho_k(\Delta_\ell) \frac{[\Omega_k]_{i,j}}{\sum_{k'=0}^{K-1} \left[\Omega_{k'}^{\phi_{k'} - \phi_k} \right]_{i,j} + \frac{1}{\rho_{\text{tr}}}} [\mathbf{Y}_{k,\ell}]_{i,j}. \quad (33) \end{aligned}$$

Recalling the pilot segment channel estimate in (27), it can be seen that

$$\hat{\mathbf{H}}_{k,\ell+\Delta_\ell} = \varrho_k(\Delta_\ell) \hat{\mathbf{H}}_{k,\ell} \quad (34)$$

which indicates that optimal channel estimates during the data segment can be easily obtained via prediction with initial channel estimates obtained during the pilot segment, and the complexity of channel prediction in massive MIMO-OFDM can be further reduced. Similar to (29), the sum MSE-CP for a given delay Δ_ℓ between the data symbol and pilot symbol can be defined as

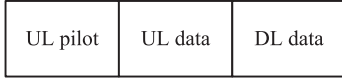
$$\begin{aligned} \epsilon^{\text{CP}}(\Delta_\ell) &\triangleq \sum_{k=0}^{K-1} \sum_{i=0}^{M-1} \sum_{j=0}^{N_g-1} \mathbb{E} \left\{ \left| \left[\mathbf{H}_{k,\ell+\Delta_\ell} - \hat{\mathbf{H}}_{k,\ell+\Delta_\ell} \right]_{i,j} \right|^2 \right\} \\ &= \sum_{k=0}^{K-1} \sum_{i=0}^{M-1} \sum_{j=0}^{N_g-1} \left\{ [\Omega_k]_{i,j} - \varrho_k^2(\Delta_\ell) \right. \\ & \quad \left. \cdot \frac{[\Omega_k]_{i,j}^2}{\sum_{k'=0}^{K-1} \left[\Omega_{k'}^{\phi_{k'} - \phi_k} \right]_{i,j} + \frac{1}{\rho_{\text{tr}}}} \right\}. \quad (35) \end{aligned}$$

From (35), it can be seen that pilot interference will affect channel prediction performance similar to the channel estimation case. However, we will show in the following proposition that such effects can still be eliminated with proper pilot phase shift scheduling.

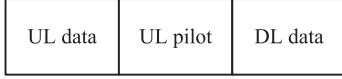
Proposition 4: The sum MSE-CP $\epsilon^{\text{CP}}(\Delta_\ell) \forall \Delta_\ell$ is lower bounded by

$$\begin{aligned} \epsilon^{\text{CP}}(\Delta_\ell) &\geq \varepsilon^{\text{CP}}(\Delta_\ell) \\ &= \sum_{k=0}^{K-1} \sum_{i=0}^{M-1} \sum_{j=0}^{N_g-1} \left\{ [\Omega_k]_{i,j} - \varrho_k^2(\Delta_\ell) \frac{[\Omega_k]_{i,j}^2}{[\Omega_k]_{i,j} + \frac{1}{\rho_{\text{tr}}}} \right\} \quad (36) \end{aligned}$$

and the lower bound can be achieved under the condition that,



(a) Type-A frame structure



(b) Type-B frame structure

Fig. 1. Frame structures for TDD transmission.

for $\forall k, k' \in \mathcal{K}$ and $k \neq k'$,

$$\left(\bar{\Omega}_{k,(N_c)} \mathbf{\Pi}_{N_c}^{\phi_k}\right) \odot \left(\bar{\Omega}_{k',(N_c)} \mathbf{\Pi}_{N_c}^{\phi_{k'}}\right) = \mathbf{0}. \quad (37)$$

Proof: The proof is similar to that of Proposition 3, and is omitted for brevity. ■

D. Frame Structure

There exist two typical frame structures for TDD massive MIMO transmission [49]. One type of frame structure (which will be referred to as type-A) begins with the UL pilot segment, followed by the UL and downlink (DL) data segments, as shown in Fig. 1(a). In the second type (which will be referred to as type-B), the UL pilot segment is placed between the UL data segment and DL data segment, as shown in Fig. 1(b). For the proposed APSP approach, the delay between the tail-end symbols of the data segment and the pilot segment will be longer than the PSOP approach due to the reduced pilot segment length. In addition, the APSP approach focuses on high mobility scenarios where channels vary relatively quickly. Thus the type-B frame structure is well-suited for the proposed APSP approach.

E. Pilot Phase Shift Scheduling

In the previous subsections, we investigated channel estimation and prediction for massive MIMO-OFDM with APSPs, and obtained the optimal pilot phase shift scheduling condition applicable to both channel estimation and prediction. However, such an optimal condition cannot always be met in practice, but pilot phase shift scheduling can still be beneficial. Several scheduling criteria can be adopted. For example, if we schedule the pilot phase shifts based on the MMSE-CE criterion, the problem can be formulated as

$$\arg \min_{\{\phi_k : k \in \mathcal{K}\}} \epsilon^{\text{CE}} \quad (38)$$

where ϵ^{CE} is defined in (29). Such a scheduling problem is combinatorial, and optimal solutions must be found through an exhaustive search. Note that the optimal phase shift scheduling conditions for channel estimation and prediction are the same, thus solution of the problem (38) can also be expected to perform well under the MMSE-CP criterion.

Algorithm 1 Pilot Phase Shift Scheduling Algorithm

Input: The UT set \mathcal{K} and the corresponding ADCPMs $\{\Omega_k : k \in \mathcal{K}\}$; the preset threshold γ

Output: Pilot phase shift pattern $\{\phi_k : k \in \mathcal{K}\}$

- 1: Initialization: $\phi_0 = 0$, scheduled UT set $\mathcal{K}^{\text{sch}} = \{0\}$, unscheduled UT set $\mathcal{K}^{\text{un}} = \mathcal{K} \setminus \mathcal{K}^{\text{sch}}$
- 2: **for** $k \in \mathcal{K}^{\text{un}}$ **do**
- 3: Search for a phase shift ϕ that satisfies $\xi\left(\bar{\Omega}_{k,(N_c)} \mathbf{\Pi}_{N_c}^{\phi}, \sum_{k' \in \mathcal{K}^{\text{sch}}} \bar{\Omega}_{k',(N_c)} \mathbf{\Pi}_{N_c}^{\phi_{k'}}\right) \leq \gamma$
- 4: If ϕ cannot be found in step 3, then $\phi = \arg \min_x \xi\left(\bar{\Omega}_{k,(N_c)} \mathbf{\Pi}_{N_c}^x, \sum_{k' \in \mathcal{K}^{\text{sch}}} \bar{\Omega}_{k',(N_c)} \mathbf{\Pi}_{N_c}^{\phi_{k'}}\right)$
- 5: Update $\phi_k = \phi$, $\mathcal{K}^{\text{sch}} \leftarrow \mathcal{K}^{\text{sch}} \cup \{k\}$, $\mathcal{K}^{\text{un}} \leftarrow \mathcal{K}^{\text{un}} \setminus \{k\}$
- 6: **end for**

Motivated by the optimal condition for channel estimation and prediction obtained in previous subsections, a simplified pilot phase shift scheduling algorithm can be developed. We first define the following function that measures the degree of overlap between two real matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{M \times N}$ as follows

$$\xi(\mathbf{A}, \mathbf{B}) \triangleq \frac{\sum_{i,j} [\mathbf{A} \odot \mathbf{B}]_{i,j}}{\sqrt{\sum_{i,j} [\mathbf{A}]_{i,j}^2} \cdot \sqrt{\sum_{i,j} [\mathbf{B}]_{i,j}^2}}. \quad (39)$$

From the Cauchy-Schwarz inequality, it is obvious that the overlapping degree function in (39) satisfies $0 \leq \xi(\mathbf{A}, \mathbf{B}) \leq 1$. When \mathbf{A} is a scaled version of \mathbf{B} , $\xi(\mathbf{A}, \mathbf{B}) = 1$. When the locations of non-zero elements in \mathbf{A} and \mathbf{B} are non-overlapping, $\xi(\mathbf{A}, \mathbf{B}) = 0$. In our algorithm, we preset a threshold to balance the tradeoff between the algorithm complexity and channel acquisition performance. Specifically, we schedule the pilot phase shifts for different UTs to make the overlap function between the ADCPMs for different UTs smaller than the preset threshold γ . Intuitively, the smaller the threshold γ , the better the channel acquisition performance will be, although with a higher algorithm complexity. The description of the proposed algorithm is summarized in Algorithm 1.

IV. CHANNEL ACQUISITION WITH APSPS OVER MULTIPLE SYMBOLS

In the previous section, we investigated channel acquisition for massive MIMO-OFDM with the proposed APSPs over one OFDM symbol. Sometimes pilots over one symbol might be not sufficient to accommodate a large number of UTs. In this section, we extend the use of APSPs to the case of multiple consecutive OFDM symbols.

We assume that the pilots are sent over Q consecutive OFDM symbols starting with the ℓ th symbol in each frame. In practice, the pilot segment length Q is usually short, and we adopt the widely accepted assumption that the channels remain constant during the pilot segment [10]–[12]. Then the received signals by the BS during the pilot segment can be written as

$$\mathbf{Y}_{\ell,(Q)} = \sum_{k'=0}^{K-1} \mathbf{G}_{k',\ell} \mathbf{X}_{k',(Q)} + \mathbf{Z}_{\ell,(Q)}$$

$$\begin{aligned}
&= \sum_{k'=0}^{K-1} \mathbf{V}_M \mathbf{H}_{k',\ell} \mathbf{F}_{N_c \times N_g}^T \mathbf{X}_{k',(Q)} \\
&\quad + \mathbf{Z}_{\ell,(Q)} \in \mathbb{C}^{M \times N_c Q} \quad (40)
\end{aligned}$$

where $\mathbf{Y}_{\ell,(Q)} \triangleq [\mathbf{Y}_\ell \ \mathbf{Y}_{\ell+1} \ \dots \ \mathbf{Y}_{\ell+Q-1}]$, $\mathbf{Y}_\ell \in \mathbb{C}^{M \times N_c}$ represents the received pilot signal at the BS during the ℓ th symbol, $\mathbf{X}_{k,(Q)} \triangleq [\mathbf{X}_{k,0} \ \mathbf{X}_{k,1} \ \dots \ \mathbf{X}_{k,Q-1}]$ represents the pilot signals and $\mathbf{X}_{k,q} = \text{diag}\{\mathbf{x}_{k,q}\} \in \mathbb{C}^{N_c \times N_c}$ represents the signal sent from the k th UT during the q th symbol of the pilot segment, $\mathbf{Z}_{\ell,(Q)}$ is AWGN with i.i.d. elements distributed as $\mathcal{CN}(0, \sigma_{\text{ztr}})$ and σ_{ztr} is the noise power.

Recalling (19), the maximum adjustable phase shift for different pilots over one OFDM symbol is $N_c - 1$. For the Q pilot symbol case, the maximum adjustable pilot phase shift can be extended to $QN_c - 1$. By exploiting the modulo operation, we construct the APSPs over multiple OFDM symbols as follows

$$\begin{aligned}
\mathbf{X}_{k,(Q)} \triangleq \sqrt{Q} [\mathbf{U}]_{\langle \phi_k \rangle_Q, :} \otimes \mathbf{X}_{\lfloor \phi_k / Q \rfloor}, \\
\phi_k = 0, 1, \dots, QN_c - 1 \quad (41)
\end{aligned}$$

where \mathbf{U} is an arbitrary $Q \times Q$ dimensional unitary matrix, and $\mathbf{X}_{\lfloor \phi_k / Q \rfloor}$ is the APSP signal over one symbol defined in (19). Then it can be obtained that, for $\forall k, k' \in \mathcal{K}$,

$$\begin{aligned}
&\mathbf{X}_{k',(Q)} (\mathbf{X}_{k,(Q)})^H \\
&= Q \left([\mathbf{U}]_{\langle \phi_{k'} \rangle_Q, :} \otimes \mathbf{X}_{\lfloor \phi_{k'} / Q \rfloor} \right) \\
&\quad \cdot \left([\mathbf{U}]_{\langle \phi_k \rangle_Q, :} \otimes \mathbf{X}_{\lfloor \phi_k / Q \rfloor} \right)^H \\
&\stackrel{(a)}{=} Q \left([\mathbf{U}]_{\langle \phi_{k'} \rangle_Q, :} [\mathbf{U}]_{\langle \phi_k \rangle_Q, :}^H \right) \otimes \left(\mathbf{X}_{\lfloor \phi_{k'} / Q \rfloor} \mathbf{X}_{\lfloor \phi_k / Q \rfloor}^H \right) \\
&\stackrel{(b)}{=} \sigma_{\text{xtr}} Q \delta \left(\langle \phi_{k'} \rangle_Q - \langle \phi_k \rangle_Q \right) \cdot \mathbf{D}_{\lfloor \phi_{k'} / Q \rfloor - \lfloor \phi_k / Q \rfloor} \quad (42)
\end{aligned}$$

where (a) follows from the Kronecker product identities $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD})$ and $(\mathbf{A} \otimes \mathbf{B})^H = \mathbf{A}^H \otimes \mathbf{B}^H$ [50], and (b) follows from (20). This shows that the available phase shifts for the Q symbol case are divided into Q groups for the proposed APSPs in (41), and the group index depends on the residue of the pilot phase shift ϕ with respect to the pilot segment length Q . Pilot interference can only affect the UTs using APSPs with phase shifts in the same group. For example, if $\langle \phi_{k'} \rangle_Q = \langle \phi_k \rangle_Q$, then phase shifts $\phi_{k'}$ and ϕ_k are within the same group, and the corresponding channel acquisition of UTs k' and k might be mutually affected.

Given the APSP correlation property over multiple symbols in (42), the channel estimation and prediction operations can be performed similarly to the single-symbol case investigated in the previous section, and we will briefly discuss such issues below.

After decorrelation and power normalization with $\mathbf{Y}_{\ell,(Q)}$ given in (40), the BS can obtain an observation of the pilot segment ADCRM $\mathbf{H}_{k,\ell}$ as

$$\begin{aligned}
&\mathbf{Y}_{k,\ell,(Q)} \\
&= \frac{1}{\sigma_{\text{xtr}} Q} \mathbf{V}_M^H \mathbf{Y}_{\ell,(Q)} \mathbf{X}_{k,(Q)}^H \mathbf{F}_{N_c \times N_g}^* \\
&= \frac{1}{\sigma_{\text{xtr}} Q} \sum_{k'=0}^{K-1} \mathbf{H}_{k',\ell} \mathbf{F}_{N_c \times N_g}^T \mathbf{X}_{k',(Q)} \mathbf{X}_{k,(Q)}^H \mathbf{F}_{N_c \times N_g}^*
\end{aligned}$$

$$\begin{aligned}
&+ \frac{1}{\sigma_{\text{xtr}} Q} \mathbf{V}_M^H \mathbf{Z}_{\ell,(Q)} \mathbf{X}_{k,(Q)}^H \mathbf{F}_{N_c \times N_g}^* \\
&\stackrel{(a)}{=} \sum_{k'=0}^{K-1} \delta \left(\langle \phi_{k'} \rangle_Q - \langle \phi_k \rangle_Q \right) \cdot \mathbf{H}_{k',\ell} \mathbf{F}_{N_c \times N_g}^T \\
&\quad \cdot \mathbf{D}_{\lfloor \phi_{k'} / Q \rfloor - \lfloor \phi_k / Q \rfloor} \mathbf{F}_{N_c \times N_g}^* + \frac{1}{\sqrt{\rho_{\text{tr}} Q}} \mathbf{Z}_{\text{iid}} \\
&\stackrel{(b)}{=} \underbrace{\mathbf{H}_{k,\ell} + \sum_{k' \neq k} \delta \left(\langle \phi_{k'} \rangle_Q - \langle \phi_k \rangle_Q \right) \cdot \mathbf{H}_{k',\ell}^{\lfloor \phi_{k'} / Q \rfloor - \lfloor \phi_k / Q \rfloor}}_{\text{pilot interference}} \\
&\quad + \underbrace{\frac{1}{\sqrt{\rho_{\text{tr}} Q}} \mathbf{Z}_{\text{iid}}}_{\text{pilot noise}} \quad (43)
\end{aligned}$$

where (a) follows from (42), $\rho_{\text{tr}} \triangleq \sigma_{\text{xtr}} / \sigma_{\text{ztr}}$ is the pilot segment SNR, \mathbf{Z}_{iid} is the normalized AWGN matrix with i.i.d. elements distributed as $\mathcal{CN}(0, 1)$, and (b) follows from (24).

With the channel observation $\mathbf{Y}_{k,\ell,(Q)}$ in (43), the MMSE estimate of the ADCRM $\mathbf{H}_{k,\ell}$ can be readily obtained in an element-wise manner as (44) shown at the top of the next page, and the corresponding sum MSE-CE is given by (45) shown at the top of the next page. In addition, prediction of the ADCRM $\mathbf{H}_{k,\ell+\Delta_\ell}$ based on $\mathbf{Y}_{k,\ell,(Q)}$ can be performed as (46) shown at the top of the next page, and the corresponding sum MSE-CP with a given delay Δ_ℓ is given by (47) shown at the top of the next page.

Based on the above sum MSE-CE and MSE-CP expressions for the multiple symbol APSP case, we can readily obtain the following proposition.

Proposition 5: The sum MSE-CE $\epsilon_{(Q)}^{\text{CE}}$ is lower bounded by

$$\epsilon_{(Q)}^{\text{CE}} \geq \epsilon_{(Q)}^{\text{CE}} = \sum_{k=0}^{K-1} \sum_{i=0}^{M-1} \sum_{j=0}^{N_g-1} \left\{ [\boldsymbol{\Omega}_k]_{i,j} - \frac{[\boldsymbol{\Omega}_k]_{i,j}^2}{[\boldsymbol{\Omega}_k]_{i,j} + \frac{1}{\rho_{\text{tr}} Q}} \right\} \quad (48)$$

and the sum MSE-CP $\epsilon_{(Q)}^{\text{CP}}(\Delta_\ell)$ for $\forall \Delta_\ell$ is lower bounded by

$$\begin{aligned}
\epsilon_{(Q)}^{\text{CP}}(\Delta_\ell) &\geq \epsilon_{(Q)}^{\text{CP}}(\Delta_\ell) \\
&= \sum_{k=0}^{K-1} \sum_{i=0}^{M-1} \sum_{j=0}^{N_g-1} \left\{ [\boldsymbol{\Omega}_k]_{i,j} - \frac{\varrho_k^2(\Delta_\ell) [\boldsymbol{\Omega}_k]_{i,j}^2}{[\boldsymbol{\Omega}_k]_{i,j} + \frac{1}{\rho_{\text{tr}} Q}} \right\}. \quad (49)
\end{aligned}$$

Both the lower bounds in (48) and (49) can be achieved under the condition that, for $\forall k, k' \in \mathcal{K}$ and $k \neq k'$,

$$\begin{aligned}
&\left(\bar{\boldsymbol{\Omega}}_{k,(N_c)} \boldsymbol{\Pi}_{N_c}^{\lfloor \phi_k / Q \rfloor} \right) \odot \left(\bar{\boldsymbol{\Omega}}_{k',(N_c)} \boldsymbol{\Pi}_{N_c}^{\lfloor \phi_{k'} / Q \rfloor} \right) = \mathbf{0}, \\
&\quad \text{when } \langle \phi_k \rangle_Q = \langle \phi_{k'} \rangle_Q. \quad (50)
\end{aligned}$$

Proof: The proof is similar to that of Proposition 3, and is omitted for brevity. ■

Proposition 5 extends the single-symbol APSP case in the previous section to the multiple symbol case. Actually, when $Q = 1$, Proposition 5 reduces to the results in Proposition 3 and Proposition 4. The interpretation of Proposition 5 is straightforward. For multiple symbol APSPs, different pilot phase shifts are divided into several groups, and pilot interference only affects the UTs using the phase shifts within the same group. If pilot interference can be eliminated through proper phase shift scheduling in all the groups, then optimal channel estimation and prediction performance can

$$\left[\hat{\mathbf{H}}_{k,\ell}\right]_{i,j} = \frac{[\boldsymbol{\Omega}_k]_{i,j}}{\sum_{k'=0}^{K-1} \delta(\langle\phi_{k'}\rangle_Q - \langle\phi_k\rangle_Q) \left[\boldsymbol{\Omega}_{k'}^{\lfloor\phi_{k'}/Q\rfloor - \lfloor\phi_k/Q\rfloor}\right]_{i,j} + \frac{1}{\rho_{\text{tr}}Q}} \left[\mathbf{Y}_{k,\ell,(Q)}\right]_{i,j} \quad (44)$$

$$\epsilon_{(Q)}^{\text{CE}} = \sum_{k=0}^{K-1} \sum_{i=0}^{M-1} \sum_{j=0}^{N_g-1} \left\{ \left[\boldsymbol{\Omega}_k \right]_{i,j} - \frac{[\boldsymbol{\Omega}_k]_{i,j}^2}{\sum_{k'=0}^{K-1} \delta(\langle\phi_{k'}\rangle_Q - \langle\phi_k\rangle_Q) \left[\boldsymbol{\Omega}_{k'}^{\lfloor\phi_{k'}/Q\rfloor - \lfloor\phi_k/Q\rfloor}\right]_{i,j} + \frac{1}{\rho_{\text{tr}}Q}} \right\} \quad (45)$$

$$\left[\hat{\mathbf{H}}_{k,\ell+\Delta_\ell}\right]_{i,j} = \frac{\varrho_k(\Delta_\ell) [\boldsymbol{\Omega}_k]_{i,j}}{\sum_{k'=0}^{K-1} \delta(\langle\phi_{k'}\rangle_Q - \langle\phi_k\rangle_Q) \left[\boldsymbol{\Omega}_{k'}^{\lfloor\phi_{k'}/Q\rfloor - \lfloor\phi_k/Q\rfloor}\right]_{i,j} + \frac{1}{\rho_{\text{tr}}Q}} \left[\mathbf{Y}_{k,\ell,(Q)}\right]_{i,j} \quad (46)$$

$$\epsilon_{(Q)}^{\text{CP}}(\Delta_\ell) = \sum_{k=0}^{K-1} \sum_{i=0}^{M-1} \sum_{j=0}^{N_g-1} \left\{ \left[\boldsymbol{\Omega}_k \right]_{i,j} - \frac{\varrho_k^2(\Delta_\ell) [\boldsymbol{\Omega}_k]_{i,j}^2}{\sum_{k'=0}^{K-1} \delta(\langle\phi_{k'}\rangle_Q - \langle\phi_k\rangle_Q) \left[\boldsymbol{\Omega}_{k'}^{\lfloor\phi_{k'}/Q\rfloor - \lfloor\phi_k/Q\rfloor}\right]_{i,j} + \frac{1}{\rho_{\text{tr}}Q}} \right\} \quad (47)$$

be achieved. When the optimal pilot phase shift scheduling condition in Proposition 5 cannot be met, a straightforward extension of the pilot phase shift scheduling algorithm in the previous section can be applied. Specifically, the UT set can be divided into Q groups, and pilot phase shift scheduling can be performed within each UT group using Algorithm 1. The tradeoff between channel acquisition performance and algorithm complexity can still be balanced with the preset threshold to determine the degree of allowable channel overlap.

V. NUMERICAL RESULTS

In this section, we present numerical simulations to evaluate the performance of the proposed APSP-CA in massive MIMO-OFDM. The major OFDM parameters, which are based on 3GPP LTE [46], are summarized in Table I. The massive MIMO-OFDM system considered is assumed to be equipped with a 128-antenna ULA at the BS with half wavelength antenna spacing. The number of UTs is set to $K = 42$ as in [4].

We consider channels with 20 taps in the delay domain, which exhibit an exponential power delay profile [18], [51]

$$\mathbf{S}_k^{\text{del}}(\tau) \propto \exp(-\tau/\varsigma_k), \text{ for } \tau \in [0, N_g T_s] \quad (51)$$

where ς_k denotes the channel delay spread of UT k . We assume that transmissions from all the UTs are synchronized [13], [18]. The q th channel tap of UT k is assumed to exhibit a Laplacian power angle spectrum [18], [33], [51]

$$\mathbf{S}_{k,q}^{\text{ang}}(\theta) \propto \exp\left(-\sqrt{2}|\theta - \theta_{k,q}|/\varphi_{k,q}\right), \\ \text{for } \theta \in \mathcal{A} = [-\pi/2, \pi/2] \quad (52)$$

where $\theta_{k,q}$ and $\varphi_{k,q}$ represent the corresponding mean angle of arrival (AoA) and angle spread for the given channel tap, respectively. We assume that the UTs are uniformly distributed in a 120° sector, and the mean AoA $\theta_{k,q}$ is uniformly

TABLE I
OFDM SYSTEM PARAMETERS

Parameter	Value
System bandwidth	20 MHz
Sampling duration T_s	32.6 ns
Subcarrier spacing	15 kHz
Subcarrier number N_c	2048
Guard interval N_g	144
Symbol length T_{sym}	71.4 μ s

distributed in the angle interval $[-\pi/3, \pi/3]$ in radians. We do not consider large scale fading in the simulations, and channels are normalized as $\sum_{i,j} [\boldsymbol{\Omega}_k]_{i,j} = MN_c$ for $\forall k$. We consider channel propagation under several typical mobility scenarios including suburban (SU), urban macro (UMa), and urban micro (UMi). The primary statistical channel parameters under these scenarios are based on the WINNER II channel model [18], [38], and are summarized in Table II. We assume that all UTs exhibit the same Doppler, delay, and angle spread in the simulations.

With the above settings, we compare the performance of the proposed APSP-CA approach with that of the conventional PSOP-CA approach, which serves as the benchmark for comparison of channel acquisition performance. For the conventional PSOP-CA, the required pilot segment length is $Q = \lceil K/(N_c/N_g) \rceil = 3$ OFDM symbols [4]. For the proposed APSP-CA, the pilot segment length can be set to $Q = 1$ or 2. We adopt Algorithm 1 to schedule the pilot phase shifts in the simulations, and the overlap threshold in the algorithm is set as $\gamma = 10^{-4}$. Although this algorithm is suboptimal in general compared with exhaustive search,

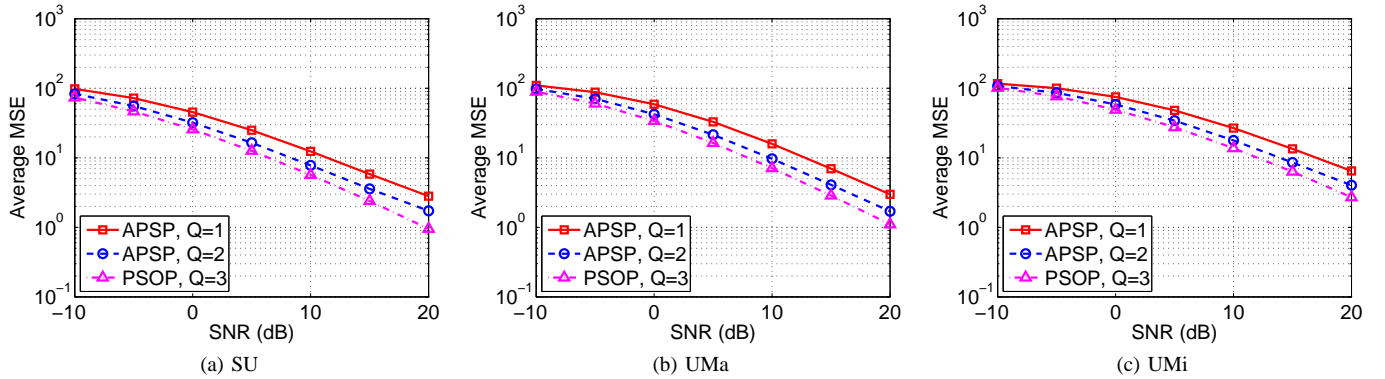


Fig. 2. Comparison of the pilot segment MSE-CE performance of APSPs and PSOPs. Results are shown versus the pilot SNR in several typical scenarios.

TABLE II
STATISTICAL CHANNEL PARAMETERS IN TYPICAL SCENARIOS

Scenario	Doppler νT_{sym} (Velocity)	Delay spread ζ	Angle spread φ
Suburban (SU)	31×10^{-3} (250 km/h)	$0.77 \mu\text{s}$	2°
Urban macro (UMa)	14×10^{-3} (100 km/h)	$1.85 \mu\text{s}$	2°
Urban micro (UMi)	6.6×10^{-3} (50 km/h)	$0.62 \mu\text{s}$	10°

substantial performance gains over the conventional PSOP-CA in terms of achievable spectral efficiency can still be achieved with relatively little computational cost.

In Fig. 2, the pilot segment MSE-CE performance⁵ obtained by the proposed APSPs (with $Q = 1$ and 2) are compared with those for conventional PSOPs ($Q = 3$) under several typical propagation scenarios. It can be observed that, in all the considered scenarios, the MSE-CE performance with APSPs approaches the performance obtained with PSOPs, while the pilot overhead is reduced by 66.7% ($Q = 1$) and 33.3% ($Q = 2$), respectively.

In Fig. 3, we compare the channel acquisition performance during the data segment in terms of MSE versus the delay Δ_ℓ between the data symbol and pilot segment. Both the APSP-CA ($Q = 1$) and PSOP-CA ($Q = 3$) are evaluated. Also, for APSPs, both the channel estimation and prediction MSE performance are calculated. It can be observed that the MSE-CP performance obtained with APSPs approaches that for PSOPs, with the pilot overhead reduced by 66.7%. In addition, with APSPs, channel prediction outperforms channel estimation in all the evaluated scenarios. Note that the channel acquisition performance in terms of both MSE-CE and MSE-CP grows almost linearly with delay, and thus the channel acquisition performance can be improved when combined

⁵All the simulated MSE results are normalized by the number of subcarriers N_c and the number of UTs K .

with the type-B frame structure, as shown in the following simulation results.

At the end of this section, we compare the achievable spectral efficiency of the proposed APSP and the conventional PSOP approaches.⁶ We assume that the frame length equals $500 \mu\text{s}$ as in [4], which is equal to the length of 7 OFDM symbols [46], and that UL and DL data transmission each occupies half of the data segment length. For the conventional PSOP-CA approach, channel estimation and the type-A frame structure in Fig. 1(a) are adopted. For the proposed APSP-CA approach, both APSPs ($Q = 1$) and channel prediction are adopted, and both type-A and type-B frame structures are considered. A MMSE receiver and precoder are employed for both UL and DL data transmissions, and the SNR is assumed to be equal to the pilot SNR. In Fig. 4, the achieved spectral efficiency⁷ of the APSP-CA and PSOP-CA approaches are depicted. It can be observed that the proposed APSP-CA approach shows substantial performance gain in terms of the achievable spectral efficiency over the conventional PSOP-CA approach, especially in the high mobility regime where pilot overhead dominates and the high SNR regime where pilot interference dominates. Specifically, in the high mobility SU scenario (250 km/h) with an SNR of 10 dB, the proposed APSPs can provide about 69% in average spectral efficiency gains over the conventional PSOPs. In addition, the type-B frame structure can provide a gain of about 64% over the type-A frame structure when APSPs are adopted.

VI. CONCLUSION

In this paper, we proposed a channel acquisition approach with adjustable phase shift pilots (APSPs) for massive MIMO-OFDM to reduce the pilot overhead. We first investigated the channel sparsity in massive MIMO-OFDM based on a physically motivated channel model. With this channel model, we investigated channel estimation and prediction for massive

⁶Note that the achievable spectral efficiency can reflect the tradeoff between the transmission performance and pilot overhead. Intuitively, reducing the pilot overhead decreases the channel acquisition quality (which leads to degradation of the achievable spectral efficiency), but also increases the length of the data segments (which leads to increased achievable spectral efficiency).

⁷The achievable UL rate is evaluated using the classical worst case approach as in [52], and the achievable DL rate is evaluated using the approach in [53]. The OFDM guard interval overhead is taken into account.

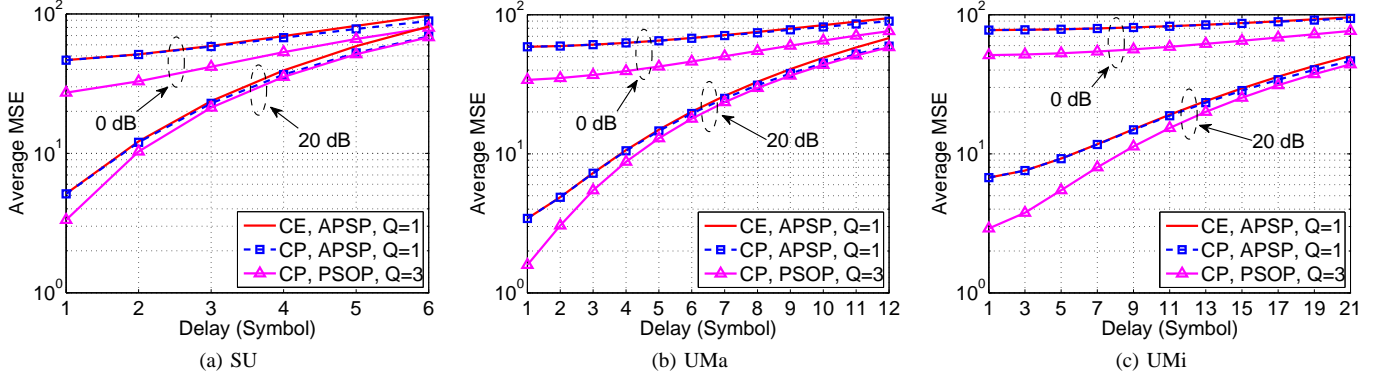


Fig. 3. Comparison of the channel acquisition performance during the data segment between APSPs and PSOPs. For APSPs, the performance of both channel estimation and prediction are depicted. Results are shown versus the channel acquisition delay in several typical scenarios.

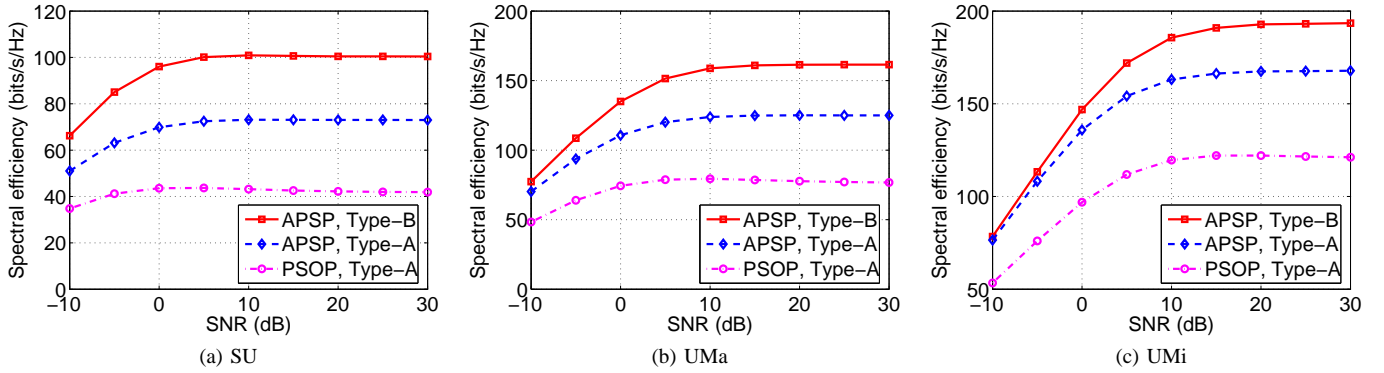


Fig. 4. Comparison of the achievable spectral efficiency between the APSP-CA and PSOP-CA approaches. For the APSP-CA approach, both type-A and type-B frame structures are considered. Results are shown versus the SNR in several typical scenarios.

MIMO-OFDM with APSPs, and provided an optimal pilot phase shift scheduling condition applicable to both channel estimation and prediction. We further developed a simplified pilot phase shift scheduling algorithm based on this optimal channel acquisition condition with APSPs. The proposed APSP-CA implemented over both one and multiple symbols were investigated. Significant performance gains in terms of achievable spectral efficiency were observed for the proposed APSP-CA approach over the conventional PSOP-CA approach in several typical mobility scenarios.

APPENDIX A DERIVATION OF (6)

The derivation of (6) is detailed in (53), shown at the top of the next page, where (a) follows from (5), and (b) follows from the definition of the delta function.

APPENDIX B PROOF OF PROPOSITION 1

We start by defining some auxiliary variables to simplify the derivations. We define $n_d \triangleq \lfloor d/M \rfloor$ and $m_d \triangleq \langle d \rangle_M$ for an arbitrary non-negative integer d . Note that the element indices start from 0 in this paper. Then we can readily obtain that for a matrix $\mathbf{\Omega}_k \in \mathbb{R}^{M \times N_g}$, the d th element of $\text{vec}\{\mathbf{\Omega}_k\}$ equals the (m_d, n_d) th element of $\mathbf{\Omega}_k$, i.e., $[\text{vec}\{\mathbf{\Omega}_k\}]_d = [\mathbf{\Omega}_k]_{m_d, n_d}$.

We can also obtain that for matrices $\mathbf{F} \in \mathbb{C}^{N_c \times N_g}$ and $\mathbf{V} \in \mathbb{C}^{M \times M}$, $[\mathbf{F} \otimes \mathbf{V}]_{i,j} = [\mathbf{F}]_{n_i, n_j} [\mathbf{V}]_{m_i, m_j}$ from the definition of the Kronecker product. With the above definitions and related properties, the proof can be obtained as follows:

$$\begin{aligned}
& \lim_{M \rightarrow \infty} \left[\mathbf{R}_k - (\mathbf{F}_{N_c \times N_g} \otimes \mathbf{V}_M) \text{diag}\{\text{vec}\{\mathbf{\Omega}_k\}\} \right. \\
& \quad \left. \cdot (\mathbf{F}_{N_c \times N_g} \otimes \mathbf{V}_M)^H \right]_{i,j} \\
&= \lim_{M \rightarrow \infty} [\mathbf{R}_k]_{i,j} - \lim_{M \rightarrow \infty} \sum_{d=0}^{MN_g-1} [\text{vec}\{\mathbf{\Omega}_k\}]_d \\
& \quad \cdot [\mathbf{F}_{N_c \times N_g} \otimes \mathbf{V}_M]_{i,d} [\mathbf{F}_{N_c \times N_g} \otimes \mathbf{V}_M]_{j,d}^* \\
& \stackrel{(a)}{=} \lim_{M \rightarrow \infty} [\mathbf{R}_k]_{i,j} - \lim_{M \rightarrow \infty} \sum_{n_d=0}^{N_g-1} \sum_{m_d=0}^{M-1} [\mathbf{\Omega}_k]_{m_d, n_d} \\
& \quad \cdot [\mathbf{F}_{N_c \times N_g}]_{n_i, n_d} [\mathbf{F}_{N_c \times N_g}]_{n_j, n_d}^* [\mathbf{V}_M]_{m_i, m_d} [\mathbf{V}_M]_{m_j, m_d}^* \\
& \stackrel{(b)}{=} \lim_{M \rightarrow \infty} [\mathbf{R}_k]_{i,j} - \lim_{M \rightarrow \infty} \frac{1}{MN_c} \sum_{n_d=0}^{N_g-1} \sum_{m_d=0}^{M-1} MN_c \\
& \quad \cdot (\theta_{m_d+1} - \theta_{m_d}) \mathbf{S}_k^{\text{AD}}(\theta_{m_d}, \tau_{n_d}) \\
& \quad \cdot \exp\left(-j2\pi \frac{(n_i - n_j) n_d}{N_c}\right)
\end{aligned}$$

$$\begin{aligned}
& \mathbb{E} \left\{ \text{vec} \{ \mathbf{G}_{k, \ell + \Delta_\ell} \} \text{vec}^H \{ \mathbf{G}_{k, \ell} \} \right\} \\
&= \mathbb{E} \left\{ \left[\sum_{q=0}^{N_g-1} \int_{-\infty}^{\infty} \int_{-\pi/2}^{\pi/2} [\mathbf{f}_{N_c, q} \otimes \mathbf{v}_{M, \theta}] \cdot \exp(\bar{j}2\pi\nu(\ell + \Delta_\ell) T_{\text{sym}}) \cdot g_k(\theta, qT_s, \nu) \, d\theta d\nu \right] \right. \\
&\quad \cdot \left. \left[\sum_{q'=0}^{N_g-1} \int_{-\infty}^{\infty} \int_{-\pi/2}^{\pi/2} [\mathbf{f}_{N_c, q'} \otimes \mathbf{v}_{M, \theta'}]^H \cdot \exp(-\bar{j}2\pi\nu' \ell T_{\text{sym}}) \cdot g_k(\theta', q'T_s, \nu') \, d\theta' d\nu' \right] \right\} \\
&= \sum_{q=0}^{N_g-1} \sum_{q'=0}^{N_g-1} \int_{-\infty}^{\infty} \int_{-\pi/2}^{\pi/2} \int_{-\infty}^{\infty} \int_{-\pi/2}^{\pi/2} [\mathbf{f}_{N_c, q} \otimes \mathbf{v}_{M, \theta}] [\mathbf{f}_{N_c, q'} \otimes \mathbf{v}_{M, \theta'}]^H \cdot \exp(\bar{j}2\pi\nu(\ell + \Delta_\ell) T_{\text{sym}}) \cdot \exp(-\bar{j}2\pi\nu' \ell T_{\text{sym}}) \\
&\quad \cdot \mathbb{E} \{ g_k(\theta, qT_s, \nu) g_k(\theta', q'T_s, \nu') \} \, d\theta d\nu d\theta' d\nu' \\
&\stackrel{(a)}{=} \sum_{q=0}^{N_g-1} \sum_{q'=0}^{N_g-1} \int_{-\infty}^{\infty} \int_{-\pi/2}^{\pi/2} \int_{-\infty}^{\infty} \int_{-\pi/2}^{\pi/2} [\mathbf{f}_{N_c, q} \otimes \mathbf{v}_{M, \theta}] [\mathbf{f}_{N_c, q'} \otimes \mathbf{v}_{M, \theta'}]^H \cdot \exp(\bar{j}2\pi\nu(\ell + \Delta_\ell) T_{\text{sym}}) \cdot \exp(-\bar{j}2\pi\nu' \ell T_{\text{sym}}) \\
&\quad \cdot \mathbf{S}_k^{\text{AD}}(\theta, qT_s) \cdot \mathbf{S}_k^{\text{Dop}}(\nu) \cdot \delta(\theta - \theta') \delta(q - q') \delta(\nu - \nu') \, d\theta d\nu d\theta' d\nu' \\
&\stackrel{(b)}{=} \sum_{q=0}^{N_g-1} \int_{-\infty}^{\infty} \int_{-\pi/2}^{\pi/2} [\mathbf{f}_{N_c, q} \otimes \mathbf{v}_{M, \theta}] [\mathbf{f}_{N_c, q} \otimes \mathbf{v}_{M, \theta}]^H \cdot \exp(\bar{j}2\pi\nu \Delta_\ell T_{\text{sym}}) \cdot \mathbf{S}_k^{\text{AD}}(\theta, qT_s) \cdot \mathbf{S}_k^{\text{Dop}}(\nu) \, d\theta d\nu \\
&= \underbrace{\int_{-\infty}^{\infty} \exp(\bar{j}2\pi\nu \Delta_\ell T_{\text{sym}}) \cdot \mathbf{S}_k^{\text{Dop}}(\nu) \, d\nu}_{\varrho_k(\Delta_\ell)} \cdot \underbrace{\sum_{q=0}^{N_g-1} \int_{-\pi/2}^{\pi/2} [\mathbf{f}_{N_c, q} \otimes \mathbf{v}_{M, \theta}] [\mathbf{f}_{N_c, q} \otimes \mathbf{v}_{M, \theta}]^H \cdot \mathbf{S}_k^{\text{AD}}(\theta, qT_s) \, d\theta}_{\mathbf{R}_k} \tag{53}
\end{aligned}$$

$$\begin{aligned}
& \cdot \exp\left(-\bar{j}2\pi \frac{(m_i - m_j)(m_d - M/2)}{M}\right) \\
&\stackrel{(c)}{=} \sum_{q=0}^{N_g-1} \int_{-\pi/2}^{\pi/2} [\mathbf{f}_{N_c, q} \otimes \mathbf{v}_{M, \theta}]_i [\mathbf{f}_{N_c, q}(q) \otimes \mathbf{v}_{M, \theta}]_j^* \\
&\quad - \sum_{r=0}^{N_g-1} \int_{-\pi/2}^{\pi/2} \exp\left(-\bar{j}2\pi \frac{(n_i - n_j)r}{N_c}\right) \\
&\quad \cdot \exp(-\bar{j}\pi(m_i - m_j) \sin(\theta)) \cdot \mathbf{S}_k^{\text{AD}}(\theta, \tau_r) \, d\theta \\
&= 0 \tag{54}
\end{aligned}$$

$$\begin{aligned}
& \cdot \mathbf{S}_k^{\text{AD}}(\theta, qT_s) \, d\theta - \lim_{M \rightarrow \infty} \sum_{n_d=0}^{N_g-1} \sum_{m_d=0}^{M-1} (\theta_{m_d+1} - \theta_{m_d}) \\
&\quad \cdot \mathbf{S}_k^{\text{AD}}(\theta_{m_d}, \tau_{n_d}) \exp\left(-\bar{j}2\pi \frac{(n_i - n_j)n_d}{N_c}\right) \\
&\quad \cdot \exp(-\bar{j}\pi(m_i - m_j) \sin(\theta_{m_d})) \\
&\stackrel{(d)}{=} \sum_{q=0}^{N_g-1} \int_{-\pi/2}^{\pi/2} [\mathbf{f}_{N_c, q}]_{n_i} [\mathbf{f}_{N_c, q}]_{n_j}^* [\mathbf{v}_{M, \theta}]_{m_i} [\mathbf{v}_{M, \theta}]_{m_j}^* \\
&\quad \cdot \mathbf{S}_k^{\text{AD}}(\theta, \tau_q) \, d\theta - \sum_{r=0}^{N_g-1} \int_{\theta_0}^{\theta_M} \exp\left(-\bar{j}2\pi \frac{(n_i - n_j)r}{N_c}\right) \\
&\quad \cdot \exp(-\bar{j}\pi(m_i - m_j) \sin(\theta)) \cdot \mathbf{S}_k^{\text{AD}}(\theta, \tau_r) \, d\theta \\
&\stackrel{(e)}{=} \sum_{q=0}^{N_g-1} \int_{-\pi/2}^{\pi/2} \exp\left(-\bar{j}2\pi \frac{(n_i - n_j)q}{N_c}\right) \\
&\quad \cdot \exp(-\bar{j}\pi(m_i - m_j) \sin(\theta)) \cdot \mathbf{S}_k^{\text{AD}}(\theta, \tau_q) \, d\theta
\end{aligned}$$

where (a) follows from the definition of Kronecker product and the definitions of m_d and n_d , (b) follows from (10) and the definitions of $\mathbf{F}_{N_c \times N_g}$ and \mathbf{V}_M , (c) follows from (8) and the definitions of τ_n and θ_m , (d) follows from the definition of the Kronecker product, and (e) follows from (1) and the fact that $\theta_0 = -\pi/2$ and $\theta_M = \pi/2$.

Before concluding the proof, we also have to show that both of the limits in the first equation of (54) exist and are finite. For this purpose, as can be seen from (e) of (54), we only need to show that

$$\begin{aligned}
& \left| \sum_{q=0}^{N_g-1} \int_{-\pi/2}^{\pi/2} \exp\left(-\bar{j}2\pi \frac{(n_i - n_j)q}{N_c}\right) \right. \\
&\quad \cdot \exp(-\bar{j}\pi(m_i - m_j) \sin(\theta)) \cdot \mathbf{S}_k^{\text{AD}}(\theta, \tau_q) \, d\theta \left. \right| \\
&\stackrel{(a)}{\leq} \sum_{q=0}^{N_g-1} \left| \int_{-\pi/2}^{\pi/2} \exp\left(-\bar{j}2\pi \frac{(n_i - n_j)q}{N_c}\right) \right.
\end{aligned}$$

$$\begin{aligned}
& \cdot \exp(-\bar{j}\pi(m_i - m_j) \sin(\theta)) \cdot \mathbf{S}_k^{\text{AD}}(\theta, \tau_q) \, d\theta \Big| \\
& \stackrel{(b)}{\leq} \sum_{q=0}^{N_g-1} \int_{-\pi/2}^{\pi/2} \left| \exp\left(-\bar{j}2\pi \frac{(n_i - n_j)q}{N_c}\right) \right. \\
& \quad \cdot \exp(-\bar{j}\pi(m_i - m_j) \sin(\theta)) \cdot \mathbf{S}_k^{\text{AD}}(\theta, \tau_q) \Big| \, d\theta \\
& = \sum_{q=0}^{N_g-1} \int_{-\pi/2}^{\pi/2} |\mathbf{S}_k^{\text{AD}}(\theta, \tau_q)| \, d\theta \\
& \stackrel{(c)}{<} \infty
\end{aligned} \tag{55}$$

where (a) follows from the triangle inequality $\left| \sum_{q=0}^{N-1} a_q \right| \leq \sum_{q=0}^{N-1} |a_q|$, (b) follows from the integral property $\left| \int_a^b f(x) \, dx \right| \leq \int_a^b |f(x)| \, dx$, and (c) follows from the fact that the power angle-delay spectrum function $\mathbf{S}_k^{\text{AD}}(\theta, \tau)$, which represents the channel power in the angle-delay domain, is bounded. This concludes the proof.

APPENDIX C PROOF OF PROPOSITION 2

To show (15), it suffices to show that

$$\begin{aligned}
& \mathbb{E} \left\{ \text{vec} \{ \mathbf{H}_{k, \ell + \Delta_\ell} \} \text{vec}^H \{ \mathbf{H}_{k, \ell} \} \right\} \\
& = \varrho_k(\Delta_\ell) \cdot \text{diag} \{ \text{vec} \{ \mathbf{\Omega}_k \} \}.
\end{aligned} \tag{56}$$

From the definition of $\mathbf{H}_{k, \ell}$ given in (14), we can obtain

$$\begin{aligned}
\text{vec} \{ \mathbf{H}_{k, \ell} \} & = (\mathbf{F}_{N_c \times N_g}^H \otimes \mathbf{V}_M^H) \text{vec} \{ \mathbf{G}_{k, \ell} \} \\
& = (\mathbf{F}_{N_c \times N_g} \otimes \mathbf{V}_M)^H \text{vec} \{ \mathbf{G}_{k, \ell} \}
\end{aligned} \tag{57}$$

via employing the Kronecker product identities $\text{vec} \{ \mathbf{ABC} \} = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec} \{ \mathbf{B} \}$ and $\mathbf{A}^H \otimes \mathbf{B}^H = (\mathbf{A} \otimes \mathbf{B})^H$ [50].

Then it can be shown that

$$\begin{aligned}
& \mathbb{E} \left\{ \text{vec} \{ \mathbf{H}_{k, \ell + \Delta_\ell} \} \text{vec}^H \{ \mathbf{H}_{k, \ell} \} \right\} \\
& \stackrel{(a)}{=} (\mathbf{F}_{N_c \times N_g} \otimes \mathbf{V}_M)^H \mathbb{E} \left\{ \text{vec} \{ \mathbf{G}_{k, \ell + \Delta_\ell} \} \text{vec}^H \{ \mathbf{G}_{k, \ell} \} \right\} \\
& \quad \cdot (\mathbf{F}_{N_c \times N_g} \otimes \mathbf{V}_M) \\
& \stackrel{(b)}{=} \varrho_k(\Delta_\ell) \cdot (\mathbf{F}_{N_c \times N_g} \otimes \mathbf{V}_M)^H \mathbf{R}_k (\mathbf{F}_{N_c \times N_g} \otimes \mathbf{V}_M) \\
& \stackrel{(c)}{=} \varrho_k(\Delta_\ell) \cdot \text{diag} \{ \text{vec} \{ \mathbf{\Omega}_k \} \}
\end{aligned} \tag{58}$$

where (a) follows from the fact that $\mathbf{F}_{N_c \times N_g}$ and \mathbf{V}_M are both deterministic matrices, (b) follows from (6), and (c) follows from Proposition 1. This concludes the proof.

APPENDIX D PROOF OF PROPOSITION 3

Due to the fact that the elements of $\mathbf{\Omega}_{k'}^{\phi_{k'} - \phi_k}$ are non-negative, we can obtain

$$\epsilon_k^{\text{CE}} = \sum_{k=0}^{K-1} \sum_{i=0}^{M-1} \sum_{j=0}^{N_g-1} \left\{ [\mathbf{\Omega}_k]_{i,j} \right.$$

$$\left. - \frac{[\mathbf{\Omega}_k]_{i,j}^2}{[\mathbf{\Omega}_k]_{i,j} + \sum_{k' \neq k} [\mathbf{\Omega}_{k'}^{\phi_{k'} - \phi_k}]_{i,j} + \frac{1}{\rho_{\text{tr}}}} \right\} \\
\geq \sum_{k=0}^{K-1} \sum_{i=0}^{M-1} \sum_{j=0}^{N_g-1} \left\{ [\mathbf{\Omega}_k]_{i,j} - \frac{[\mathbf{\Omega}_k]_{i,j}^2}{[\mathbf{\Omega}_k]_{i,j} + \frac{1}{\rho_{\text{tr}}}} \right\} = \epsilon^{\text{CE}}. \tag{59}$$

Furthermore, when the condition $(\bar{\mathbf{\Omega}}_{k, (N_c)} \mathbf{\Pi}_{N_c}^{\phi_k}) \odot (\bar{\mathbf{\Omega}}_{k', (N_c)} \mathbf{\Pi}_{N_c}^{\phi_{k'}}) = \mathbf{0}$ is satisfied, then with the same column permutation and column truncation, multiplications of the corresponding elements still equal zero, i.e.,

$$\begin{aligned}
& (\bar{\mathbf{\Omega}}_{k, (N_c)} \mathbf{\Pi}_{N_c}^{\phi_k} \mathbf{\Pi}_{N_c}^{-\phi_k} \mathbf{I}_{N_c \times N_g}) \\
& \odot (\bar{\mathbf{\Omega}}_{k', (N_c)} \mathbf{\Pi}_{N_c}^{\phi_{k'}} \mathbf{\Pi}_{N_c}^{-\phi_{k'}} \mathbf{I}_{N_c \times N_g}) = \mathbf{0}.
\end{aligned} \tag{60}$$

Recalling the definition in (26) and exploiting the permutation matrix property that $\mathbf{\Pi}_N^a \mathbf{\Pi}_N^b = \mathbf{\Pi}_N^{a+b}$, the condition in (60) is equivalent to

$$\mathbf{\Omega}_k \odot \mathbf{\Omega}_{k'}^{\phi_{k'} - \phi_k} = \mathbf{0}. \tag{61}$$

Then for $\forall i, j$,

$$\begin{aligned}
& [\mathbf{\Omega}_k]_{i,j}^2 \left\{ [\mathbf{\Omega}_k]_{i,j} + \sum_{k' \neq k} [\mathbf{\Omega}_{k'}^{\phi_{k'} - \phi_k}]_{i,j} + \frac{1}{\rho_{\text{tr}}} \right\} \\
& = [\mathbf{\Omega}_k]_{i,j}^2 \left\{ [\mathbf{\Omega}_k]_{i,j} + \frac{1}{\rho_{\text{tr}}} \right\}
\end{aligned} \tag{62}$$

which leads to

$$\frac{[\mathbf{\Omega}_k]_{i,j}^2}{[\mathbf{\Omega}_k]_{i,j} + \sum_{k' \neq k} [\mathbf{\Omega}_{k'}^{\phi_{k'} - \phi_k}]_{i,j} + \frac{1}{\rho_{\text{tr}}}} = \frac{[\mathbf{\Omega}_k]_{i,j}^2}{[\mathbf{\Omega}_k]_{i,j} + \frac{1}{\rho_{\text{tr}}}}. \tag{63}$$

Substituting (63) into (28), the MSE-CE expression ϵ_k^{CE} reduces to

$$\epsilon_k^{\text{CE}} = \sum_{i=0}^{M-1} \sum_{j=0}^{N_g-1} \left\{ [\mathbf{\Omega}_k]_{i,j} - \frac{[\mathbf{\Omega}_k]_{i,j}^2}{[\mathbf{\Omega}_k]_{i,j} + \frac{1}{\rho_{\text{tr}}}} \right\}. \tag{64}$$

Then the minimum in (30) can be achieved. This concludes the proof.

REFERENCES

- [1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.
- [2] F. Boccardi, R. W. Heath, Jr., A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.
- [3] C.-X. Wang, F. Haider, X. Q. Gao, X.-H. You, Y. Yang, D. Yuan, H. M. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 122–130, Feb. 2014.
- [4] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [5] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
- [6] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

- [7] L. J. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Commun.*, vol. 33, no. 7, pp. 665–675, Jul. 1985.
- [8] G. L. Stüber, J. R. Barry, S. W. McLaughlin, Y. G. Li, M. A. Ingram, and T. G. Pratt, "Broadband MIMO-OFDM wireless communications," *Proc. IEEE*, vol. 92, no. 2, pp. 271–294, Feb. 2004.
- [9] Y. G. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 67–75, Jan. 2002.
- [10] I. Barhum, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Trans. Signal Process.*, vol. 51, no. 6, pp. 1615–1624, Jun. 2003.
- [11] H. Minn and N. Al-Dahir, "Optimal training signals for MIMO OFDM channel estimation," *IEEE Trans. Wireless Commun.*, vol. 5, no. 5, pp. 1158–1168, May 2006.
- [12] Y. Chi, A. Gomaa, N. Al-Dahir, and A. R. Calderbank, "Training signal design and tradeoffs for spectrally-efficient multi-user MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 7, pp. 2234–2245, Jul. 2011.
- [13] E. Dahlman, S. Parkvall, and J. Sköld, *4G LTE/LTE-Advanced for Mobile Broadband*. Burlington, MA, USA: Academic Press, 2011.
- [14] H. D. Tuan, H. H. Kha, H. H. Nguyen, and V.-J. Luong, "Optimized training sequences for spatially correlated MIMO-OFDM," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2768–2778, Sep. 2010.
- [15] N. N. Tran, H. H. Nguyen, H. D. Tuan, and D. E. Dodds, "Training signal designs for spatially correlated multi-user multi-input multi-output with orthogonal frequency-division multiplexing systems," *IET Commun.*, vol. 6, no. 16, pp. 2630–2638, Nov. 2012.
- [16] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. New York, NY, USA: Cambridge Univ. Press, 2005.
- [17] B. Clerckx and C. Oestges, *MIMO Wireless Networks: Channels, Techniques and Standards for Multi-Antenna, Multi-User and Multi-Cell Systems*, 2nd ed. Oxford, U.K.: Academic Press, 2013.
- [18] IST-4-027756 WINNER II D1.1.2 V1.2, "WINNER II Channel Models," Tech. Rep., Feb. 2008.
- [19] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proc. IEEE*, vol. 98, no. 6, pp. 1058–1076, Jun. 2010.
- [20] S. Payami and F. Tufvesson, "Channel measurements and analysis for very large array systems at 2.6 GHz," in *Proc. EuCAP*, Prague, Czech Republic, 2012, pp. 433–437.
- [21] X. Gao, F. Tufvesson, and O. Edfors, "Massive MIMO channels – Measurements and models," in *Proc. Annu. ASILOMAR*, Pacific Grove, CA, USA, 2013, pp. 280–284.
- [22] Y. Barbotin, A. Hormati, S. Rangan, and M. Vetterli, "Estimation of sparse MIMO channels with common support," *IEEE Trans. Commun.*, vol. 60, no. 12, pp. 3705–3716, Dec. 2012.
- [23] X. Rao and V. K. N. Lau, "Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3261–3271, Jun. 2014.
- [24] M. Masood, L. H. Afify, and T. Y. Al-Naffouri, "Efficient coordinated recovery of sparse channels in massive MIMO," *IEEE Trans. Signal Process.*, vol. 63, no. 1, pp. 104–118, Jan. 2015.
- [25] L. Dai, Z. Wang, and Z. Yang, "Spectrally efficient time-frequency training OFDM for mobile large-scale MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 251–263, Feb. 2013.
- [26] Z. Chen and C. Yang, "Pilot decontamination in massive MIMO systems: Exploiting channel sparsity with pilot assignment," in *Proc. IEEE GlobalSIP*, Atlanta, GA, USA, 2014, pp. 637–641.
- [27] C.-K. Wen, S. Jin, K.-K. Wong, J.-C. Chen, and P. Ting, "Channel estimation for massive MIMO using Gaussian-mixture Bayesian learning," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1356–1368, Mar. 2015.
- [28] C. R. Berger, Z. Wang, J. Huang, and S. Zhou, "Application of compressive sensing to sparse channel estimation," *IEEE Commun. Mag.*, vol. 48, no. 11, pp. 164–174, Nov. 2010.
- [29] E. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Prob.*, vol. 23, no. 3, pp. 969–985, Apr. 2007.
- [30] R. Calderbank, S. Howard, and S. Jafarpour, "Construction of a large class of deterministic sensing matrices that satisfy a statistical isometry property," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 358–374, Apr. 2010.
- [31] T. Strohmer, "Measure what should be measured: Progress and challenges in compressive sensing," *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 887–893, Dec. 2012.
- [32] L. You, X. Q. Gao, A. L. Swindlehurst, and W. Zhong, "Adjustable phase shift pilots for sparse massive MIMO-OFDM channels," in *Proc. IEEE SPAWC*, Stockholm, Sweden, 2015, pp. 206–210.
- [33] L. You, X. Q. Gao, X.-G. Xia, N. Ma, and Y. Peng, "Pilot reuse for massive MIMO transmission over spatially correlated Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3352–3366, Jun. 2015.
- [34] O. Edfors, M. Sandell, J.-J. van de Beek, S. K. Wilson, and P. O. Börjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 931–939, Jul. 1998.
- [35] Y. G. Li, L. J. Cimini, Jr., and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 902–915, Jul. 1998.
- [36] K. Liu, V. Raghavan, and A. M. Sayeed, "Capacity scaling and spectral efficiency in wide-band correlated MIMO channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2504–2526, Oct. 2003.
- [37] G. Barriac and U. Madhoo, "Characterizing outage rates for space-time communication over wideband channels," *IEEE Trans. Commun.*, vol. 52, no. 12, pp. 2198–2208, Dec. 2004.
- [38] G. Auer, "3D MIMO-OFDM channel estimation," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 972–985, Apr. 2012.
- [39] B. H. Fleury, "First- and second-order characterization of direction dispersion and space selectivity in the radio channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 2027–2044, Sep. 2000.
- [40] M. Pätzold, *Mobile Radio Channels*, 2nd ed. Chichester, U.K.: Wiley, 2012.
- [41] W. C. Jakes, Ed., *Microwave Mobile Communications*. New York, NY, USA: IEEE Press, 1994.
- [42] J.-J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Börjesson, "On channel estimation in OFDM systems," in *Proc. IEEE VTC*, Chicago, IL, USA, 1995, pp. 305–309.
- [43] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.
- [44] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing—The large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441–6463, Oct. 2013.
- [45] A. Liu and V. K. N. Lau, "Two-stage subspace constrained precoding in massive MIMO cellular systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3271–3279, Jun. 2015.
- [46] 3GPP TS 36.211 V12.4.0, "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation (Release 12)," Tech. Rep., Dec. 2014.
- [47] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Upper Saddle River, NJ, USA: Prentice Hall, 2000.
- [48] K. T. Truong and R. W. Heath, Jr., "Effects of channel aging in massive MIMO systems," *J. Commun. Netw.*, vol. 15, no. 4, pp. 338–351, Aug. 2013.
- [49] E. Björnson, M. Matthaiou, and M. Debbah, "Massive MIMO with non-ideal arbitrary arrays: Hardware scaling laws and circuit-aware design," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4353–4368, Aug. 2015.
- [50] G. A. F. Seber, *A Matrix Handbook for Statisticians*. Hoboken, NJ, USA: Wiley, 2008.
- [51] K. I. Pedersen, P. E. Mogensen, and B. H. Fleury, "A stochastic model of the temporal and azimuthal dispersion seen at the base station in outdoor propagation environments," *IEEE Trans. Veh. Technol.*, vol. 49, no. 2, pp. 437–447, Mar. 2000.
- [52] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.
- [53] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD Systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.