

Joint Channel Estimation and Equalization in Massive MIMO Using a Single Pilot Subcarrier

Danilo Lelin Li, Arman Farhang

Abstract—The focus of this letter is on the reduction of the large pilot overhead in orthogonal frequency division multiplexing (OFDM) based massive multiple-input multiple-output (MIMO) systems. We propose a novel joint channel estimation and equalization technique that requires only one pilot subcarrier, reducing the pilot overhead by orders of magnitude. We take advantage of the coherence bandwidth spanning over multiple subcarrier bands. This allows for a band of subcarriers to be equalized with the channel frequency response (CFR) at a single subcarrier. Subsequently, the detected data symbols are considered as virtual pilots, and their CFRs are updated without additional pilot overhead. Thereafter, the remaining channel estimation and equalization can be performed in a sliding manner. With this approach, we use multiple channel estimates to equalize the data at each subcarrier. This allows us to take advantage of frequency diversity and improve the detection performance. Finally, we corroborate the above claims through extensive numerical analysis, showing the superior performance of our proposed technique compared to conventional methods.

Index Terms—OFDM, massive MIMO, channel estimation, linear combining, spatial diversity, spectral efficiency.

I. INTRODUCTION

Massive multiple-input multiple-output (MIMO) has brought substantial improvements to 5G systems and it continues to be a dominant technology in next-generation networks. In particular, spatial diversity and beamforming gains provided by massive MIMO have led to significant improvements in capacity, spectral, and energy efficiency compared to 4G networks [1]. In massive MIMO, multiple users reuse the same time-frequency resources, as their signals can be distinguished based on their channel responses. Therefore, obtaining accurate channel state information (CSI) at each base station (BS) antenna is crucial. Conventional channel estimation methods, e.g. least squares (LS), rely on the transmission of pilot sequences from each user. These methods are favorable due to their low complexity. However, this comes at the expense of spectral efficiency (SE) loss [2].

Hence, extensive research has been conducted to reduce the pilot overhead, see [3]–[5] and the references therein. For instance, (semi) blind techniques emphasize reducing the pilot length and estimating the channel by solving underdetermined systems of equations using iterative algorithms [3]. Moreover, superimposed pilots share the same time-frequency resources for both data and pilot transmission. In these techniques, the channel is estimated by treating data as noise. This can completely remove the pilot overhead, however, at the cost of interference and extra computational load [4], [5]. Most works on massive MIMO consider a narrow band channel. Therefore, utilizing these channel estimation techniques for orthogonal frequency division multiplexing (OFDM) would require pilots

on each subcarrier. Meanwhile, it was shown in [6] that the number of pilot subcarriers needs to be at least equal to the channel length. However, this leads to a large overhead, as the channel length varies from 7% to 25% of the OFDM symbol duration, [7]. To alleviate this issue, the authors in [6] proposed a technique that can reduce the required pilot subcarriers by 80%, however, the technique is limited to sparse channels.

To address the aforementioned limitations of the existing literature, in this letter, we show that only one reference subcarrier is sufficient for both channel estimation and data detection in massive MIMO. It is worth noting that our proposed technique is independent of channel sparsity and length. We take advantage of multiple adjacent subcarriers being within the channel coherence bandwidth. Thus, the transmitted data over a band of subcarriers can be detected using only one reference subcarrier's channel estimate. Thereafter, the detected data symbols are considered as virtual pilots to update their corresponding channel estimates. Each updated channel estimate is then used to equalize a new band of subcarriers within the coherence bandwidth. This way, we can obtain multiple estimates of the transmitted data symbols at each subcarrier and average them. With this approach, we take advantage of the frequency diversity apart from the spatial diversity to achieve an improved performance. By repeating this process in a sliding manner, all the data symbols and the channel frequency responses (CFRs) can be estimated, starting from one reference pilot.

Furthermore, we prove that in the large antenna regime, linear combining can be performed using the channel estimate at any subcarrier. The combined signal at a subcarrier will be scaled by a coefficient, which depends on the channel power delay profile (PDP) and frequency spacing between the subcarrier and the CFR used for combining. However, this coefficient can take very small values leading to noise enhancement. To avoid this issue, we define a *depth* factor in our proposed sliding technique that determines the number of subcarriers within each band to be equalized using the same CFR. This ensures that only large values of the aforementioned coefficient are considered. To evaluate the effectiveness of our proposed technique, we compare its bit error rate (BER) and signal-to-interference and noise ratio (SINR) performance with the conventional techniques. For 15 MHz transmission bandwidth, the existing techniques require over 70 pilot subcarriers, whereas our proposed technique reduces this overhead to only one pilot subcarrier. Moreover, our proposed technique outperforms the linear equalizers by around 2 dB, at high signal-to-noise ratios (SNRs), in terms of the BER performance.

Notations: Matrices, vectors, and scalar quantities are denoted by boldface uppercase, boldface lowercase, and normal letters, respectively. $[\mathbf{A}]_{m,n}$ represents the element on row m and column n of $\mathbf{A}$. $\text{tr}\{\mathbf{A}\}$ represents the trace of $\mathbf{A}$. $\mathbf{I}_M$ and

Danilo Lelin Li and Arman Farhang are with the Department of Electronic and Electrical Engineering, Trinity College Dublin, Dublin 2, Ireland (Email: {lelinli, arman.farhang}@tcd.ie). This publication has emanated from research conducted with the financial support of Science Foundation Ireland under Grant number 19/FFP/7005(T) and 21/US/3757.

$\mathbf{0}_{M \times N}$ are the $M \times M$ identity and $M \times N$ zero matrices, respectively. The superscripts $(\cdot)^\dagger$, $(\cdot)^H$, $(\cdot)^T$ and $(\cdot)^*$ indicate pseudo inverse, Hermitian, transpose and conjugate operations, respectively. $|\cdot|$, $((\cdot))_M$ and $\mathbb{E}\{\cdot\}$ are the absolute value, modulo- M , and expected value operators, respectively. Finally, $\mathbf{F}_M$ is the normalized M -point discrete Fourier transform (DFT) matrix and $\mathbf{f}_{M,m}$ is the m^{th} column of $\mathbf{F}_M$.

II. SYSTEM MODEL

We consider the uplink (UL) of a single-cell massive MIMO system operating in the time-division duplex (TDD) mode. OFDM with M subcarriers is deployed as the modulation format, with the cyclic prefix (CP) of length M_{CP} . M_{CP} is chosen to be larger than the channel length to avoid intersymbol interference. The BS is equipped with Q antennas and serves K single antenna users. The duration of each frame is assumed to be within the channel coherence time, including N number of OFDM time symbols in the UL. Hence, the channel remains time-invariant within each frame.

The UL transmission is divided into two phases, training/pilot and data transmission. N_p and N_d time slots are allocated to pilot and data transmission, respectively. Thus, a given user k transmits the time-frequency symbols $\mathbf{X}^k = [\mathbf{P}^k, \mathbf{D}^k]$, where $\mathbf{P}^k \in \mathbb{C}^{M \times N_p}$ represents the pilot and $\mathbf{D}^k \in \mathbb{C}^{M \times N_d}$ the data symbols. Hence, the OFDM transmit signal of user k is obtained as $\mathbf{S}^k = \mathbf{A}_{\text{CP}} \mathbf{F}_M^H \mathbf{X}^k$, where $\mathbf{A}_{\text{CP}} = [\mathbf{G}_{\text{CP}}^T, \mathbf{I}_M]^T$ is the CP addition matrix and the $M_{\text{CP}} \times M$ matrix $\mathbf{G}_{\text{CP}}$ includes the last M_{CP} rows of $\mathbf{I}_M$.

The signal is transmitted through the channel and undergoes OFDM demodulation. Thus, the received signal from all the users at a given BS antenna q is obtained as

$$\mathbf{Y}_q = \mathbf{F}_M \mathbf{R}_{\text{CP}} \sum_{k=0}^{K-1} \mathbf{H}_{q,k} \mathbf{S}^k + \mathbf{W}_q = \sum_{k=0}^{K-1} \text{diag}\{\lambda_{q,k}\} \mathbf{X}^k + \mathbf{W}_q, \quad (1)$$

where $\mathbf{R}_{\text{CP}} = [\mathbf{0}_{M \times M_{\text{CP}}}, \mathbf{I}_M]$ is the CP removal matrix, $\mathbf{H}_{q,k}$ denotes the Toeplitz channel matrix realizing the linear convolution, which is formed by the channel impulse response (CIR) of length L between user k and antenna q , i.e., $\mathbf{h}_{q,k} = [h_{q,k}[0], \dots, h_{q,k}[L-1]]^T$, and $\mathbf{W}_q$ includes the complex additive white Gaussian noise (AWGN) in the frequency domain, with the variance σ_w^2 , i.e., $[\mathbf{W}_q]_{m,n} \sim \mathcal{CN}(0, \sigma_w^2)$. We assume the CIR between each user and the BS antennas to be independent and identically distributed (i.i.d.) complex random variables, $\mathbf{h}_{q,k} \sim \mathcal{CN}(\mathbf{0}_{L \times 1}, \mathbf{\Sigma}_k)$ for $q = 0, \dots, Q-1$ and $k = 0, \dots, K-1$. $\mathbf{\Sigma}_k$ is a diagonal matrix with the diagonal elements formed by the PDP of the channel $\rho_k = [\rho_k[0], \dots, \rho_k[L-1]]^T$, where the normalized PDP is considered, i.e., $\sum_{l=0}^{L-1} \rho_k[l] = 1$. As $M_{\text{CP}} \geq L-1$, $\tilde{\mathbf{H}}_{q,k} = \mathbf{R}_{\text{CP}} \mathbf{H}_{q,k} \mathbf{A}_{\text{CP}}$ is a circulant matrix, with the first column formed by the zero-padded CIR, i.e., $\tilde{\mathbf{h}}_{q,k} = [\mathbf{h}_{q,k}^T, \mathbf{0}_{M-L \times 1}^T]^T$. Hence, the channel matrix $\tilde{\mathbf{H}}_{q,k}$ can be diagonalized by the DFT and inverse DFT matrices as $\mathbf{F}_M \tilde{\mathbf{H}}_{q,k} \mathbf{F}_M^H = \text{diag}\{\lambda_{q,k}\}$, where $\lambda_{q,k} = [\lambda_{q,k}[0], \dots, \lambda_{q,k}[M-1]]^T$ represents the CFR with the elements obtained as

$$\lambda_{q,k}[m] = \sqrt{M} \mathbf{f}_{M,m}^T \tilde{\mathbf{h}}_{q,k}. \quad (2)$$

To pave the way toward the derivations in the following sections, we rearrange the received signals into the space-time representation. By stacking the received samples at a given subcarrier m , i.e., the m^{th} row of $\mathbf{Y}_q$ from all the Q receive antennas, we obtain the $Q \times N$ space-time matrix $\bar{\mathbf{Y}}_m$. Thus, the input-output relationship for a given subcarrier m across all the antennas can be represented as

$$\bar{\mathbf{Y}}_m = \bar{\mathbf{\Lambda}}_m \bar{\mathbf{X}}_m + \bar{\mathbf{W}}_m, \quad (3)$$

where the elements of the matrices $\bar{\mathbf{\Lambda}}_m \in \mathbb{C}^{Q \times K}$, $\bar{\mathbf{X}}_m \in \mathbb{C}^{K \times N}$ and $\bar{\mathbf{W}}_m \in \mathbb{C}^{Q \times N}$ are given by $[\bar{\mathbf{\Lambda}}_m]_{q,k} = \lambda_{q,k}[m]$, $[\bar{\mathbf{X}}_m]_{k,n} = [\mathbf{X}^k]_{m,n}$ and $[\bar{\mathbf{W}}_m]_{q,n} = [\mathbf{W}_q]_{m,n}$, respectively.

Finally, $\bar{\mathbf{X}}_m \triangleq [\bar{\mathbf{P}}_m, \bar{\mathbf{D}}_m]$, where $\bar{\mathbf{P}}_m$ and $\bar{\mathbf{D}}_m$ represent the transmitted pilot and data symbols, respectively. Correspondingly, $\bar{\mathbf{Y}}_m \triangleq [\bar{\mathbf{Y}}_m^p, \bar{\mathbf{Y}}_m^d]$ and $\bar{\mathbf{W}}_m \triangleq [\bar{\mathbf{W}}_m^p, \bar{\mathbf{W}}_m^d]$.

III. CONVENTIONAL CHANNEL ESTIMATION AND EQUALIZATION

In this section, the widely used linear pilot-based channel estimation and equalization techniques are described. In particular, we consider LS-based channel estimation, and two linear combining techniques, namely maximum ratio combining (MRC) and minimum mean square error (MMSE) equalization techniques. At the channel sounding stage, each user transmits a pilot sequence with the length $N_p \geq K$ over different time slots on a given subcarrier allocated to pilot symbols [1]. The pilot sequence of length N_p for a given user k at a given subcarrier m , $\bar{\mathbf{p}}_m^k$, is chosen from a pilot book $\bar{\mathbf{P}}_m = [\bar{\mathbf{p}}_m^0, \bar{\mathbf{p}}_m^1, \dots, \bar{\mathbf{p}}_m^{K-1}]^T$ that satisfies the orthogonality property $\bar{\mathbf{P}}_m \bar{\mathbf{P}}_m^H = N_p \mathbf{I}_K$, [8]. We consider a pilot book where each $\bar{\mathbf{p}}_m^k$ is obtained from k cyclic shifts of the Zadoff-Chu (ZC) root sequence of length N_p . Using (3), the received pilots from all the users at subcarrier m can be represented as $\bar{\mathbf{Y}}_m^p = \bar{\mathbf{\Lambda}}_m \bar{\mathbf{P}}_m + \bar{\mathbf{W}}_m^p$. Consequently, the channel response for all the users at a given subcarrier m can be estimated as

$$\hat{\bar{\mathbf{\Lambda}}}_m = \frac{1}{N_p} \bar{\mathbf{Y}}_m^p \bar{\mathbf{P}}_m^H = \bar{\mathbf{\Lambda}}_m + \frac{1}{N_p} \bar{\mathbf{W}}_m^p \bar{\mathbf{P}}_m^H \triangleq \bar{\mathbf{\Lambda}}_m + \tilde{\mathbf{W}}_m, \quad (4)$$

where $\tilde{\mathbf{W}}_m = \frac{1}{N_p} \bar{\mathbf{W}}_m^p \bar{\mathbf{P}}_m^H$.

To estimate the whole CFR, it is sufficient to allocate L equally spaced subcarriers as pilots [6]. Let $\mathcal{I}$ represent the set of subcarrier indices allocated to the pilot, and $\lambda_{q,k}^{\mathcal{I}}$ the vector formed by selecting the L entries of $\lambda_{q,k}$ at the pilot subcarriers, each element of $\lambda_{q,k}^{\mathcal{I}}$ can be estimated from (4), as $\lambda_{q,k}[m] = [\hat{\bar{\mathbf{\Lambda}}}_m]_{q,k}$. From (2), the following relation between CFR and CIR holds

$$\lambda_{q,k}^{\mathcal{I}} = \sqrt{M} \mathbf{F}_M^{\mathcal{I}} \mathbf{h}_{q,k}, \quad (5)$$

where $\mathbf{F}_M^{\mathcal{I}}$ is a $L \times L$ matrix formed by selecting the first L columns of $\mathbf{F}_M$ with the rows indexed by $\mathcal{I}$. Hence, the CIR of user k at a given BS antenna q can be estimated by solving (5) for $\mathbf{h}_{q,k}$, i.e., $\mathbf{h}_{q,k} = \frac{1}{\sqrt{M}} (\mathbf{F}_M^{\mathcal{I}})^{-1} \lambda_{q,k}^{\mathcal{I}}$. Finally, after obtaining the CIR, $\hat{\bar{\mathbf{\Lambda}}}_m$ at each subcarrier can be reconstructed with (2). Therefore, the channel estimation process, (4) and (5), requires a minimum of LK pilots.

Using the channel estimates $\hat{\bar{\mathbf{\Lambda}}}_m$, and deploying a linear combining technique such as MRC, the transmit data symbols for all the users can be estimated as

$$\hat{\mathbf{X}}_m^{\text{MRC}} = \mathbf{\Gamma}_m^{-1} \hat{\bar{\mathbf{\Lambda}}}_m^H \bar{\mathbf{Y}}_m^d = \mathbf{\Gamma}_m^{-1} \hat{\bar{\mathbf{\Lambda}}}_m^H (\bar{\mathbf{\Lambda}}_m \bar{\mathbf{D}}_m + \bar{\mathbf{W}}_m^d), \quad (6)$$

where $\mathbf{\Gamma}_m$ is a $K \times K$ diagonal matrix, with the k^{th} diagonal element given by $\frac{1}{Q} \sum_{q=0}^{Q-1} |\lambda_{q,k}[m]|^2$, considered solely to normalize the amplitude of the equalizer output. In the literature, the effect of channel estimation noise is often neglected and the diagonal elements of $\mathbf{\Gamma}_m$ are formed from the diagonal elements of $\bar{\mathbf{\Lambda}}_m^H \bar{\mathbf{\Lambda}}_m$, i.e., the norm squared of the channel vector for each user at a given subcarrier m . However, in the presence of channel estimation errors, using (4), the normalization factors on the elements of $\bar{\mathbf{\Lambda}}_m^H \bar{\mathbf{\Lambda}}_m$ are obtained as the diagonal elements of

$$\hat{\mathbf{\Lambda}}_m^H \hat{\mathbf{\Lambda}}_m = \hat{\mathbf{\Lambda}}_m^H \bar{\mathbf{\Lambda}}_m + \bar{\mathbf{\Lambda}}_m^H \tilde{\mathbf{W}}_m + \frac{1}{N_p^2} \bar{\mathbf{P}}_m (\bar{\mathbf{W}}_m^p)^H \bar{\mathbf{W}}_m^p \bar{\mathbf{P}}_m^H. \quad (7)$$

The term $\bar{\mathbf{\Lambda}}_m^H \tilde{\mathbf{W}}_m$ tends to zero in the asymptotic regime, as $Q \rightarrow \infty$, since the channel gain and noise are independent. However, the same is not true for $(\bar{\mathbf{W}}_m^p)^H \bar{\mathbf{W}}_m^p$. As $\bar{\mathbf{W}}_m^p$ represents the AWGN, in the asymptotic regime, we have

$$\left(\bar{\mathbf{\Lambda}}_m^H \tilde{\mathbf{W}}_m + \frac{1}{N_p^2} \bar{\mathbf{P}}_m (\bar{\mathbf{W}}_m^p)^H \bar{\mathbf{W}}_m^p \bar{\mathbf{P}}_m^H \right) \rightarrow \frac{Q\sigma_w^2}{N_p^2} \bar{\mathbf{P}}_m \bar{\mathbf{P}}_m^H. \quad (8)$$

Hence, taking a similar approach to [9], the channel estimation noise can be mitigated by subtracting (8) from the normalization factor, i.e., $\mathbf{\Gamma}_m$ is formed from the diagonal elements of $(\hat{\mathbf{\Lambda}}_m^H \hat{\mathbf{\Lambda}}_m - \frac{Q\sigma_w^2}{N_p^2} \bar{\mathbf{P}}_m \bar{\mathbf{P}}_m^H)$.

Due to the channel hardening effect [1], in the asymptotic regime, as $Q \rightarrow \infty$, we have

$$\frac{\bar{\mathbf{\Lambda}}_{m,k}^H \bar{\mathbf{\Lambda}}_{m,k'}}{Q} \rightarrow \begin{cases} \mathbb{E}\{|\bar{\lambda}_{m,k}[q]|^2\} = 1 & \text{if } k = k' \\ 0 & \text{otherwise} \end{cases}, \quad (9)$$

where $\bar{\mathbf{\Lambda}}_{m,k}$, of length Q , represents the k^{th} column of $\bar{\mathbf{\Lambda}}_m$. Therefore, as $Q \rightarrow \infty$, $\mathbf{\Gamma}_m^{-1} \hat{\mathbf{\Lambda}}_m^H \bar{\mathbf{\Lambda}}_m \rightarrow \mathbf{I}_K$.

In practical systems, the number of BS antennas is limited, and the off-diagonal elements of $\hat{\mathbf{\Lambda}}_m^H \bar{\mathbf{\Lambda}}_m$, in (6), lead to a significant amount of interference. Thus, MMSE combining,

$$\hat{\mathbf{X}}_m^{\text{MMSE}} = \mathbf{\Phi}_m^{\text{MMSE}} \bar{\mathbf{Y}}_m^d, \quad (10)$$

where $\mathbf{\Phi}_m^{\text{MMSE}} \triangleq (\hat{\mathbf{\Lambda}}_m^H \hat{\mathbf{\Lambda}}_m - \frac{Q\sigma_w^2}{N_p^2} \bar{\mathbf{P}}_m \bar{\mathbf{P}}_m^H + \sigma_w^2 \mathbf{I}_K)^{-1} \hat{\mathbf{\Lambda}}_m^H$, provides an improved performance compared to MRC.

IV. PROPOSED CHANNEL ESTIMATION AND COMBINING

The linear combining techniques presented in the previous section take advantage of the channel hardening effect and spatial diversity gains in massive MIMO. However, it requires knowledge of the whole CFR, and therefore, a minimum of LK time-frequency slots are allocated to the pilots for channel estimation. In practical systems, L can take large values, especially as the bandwidth increases. Hence, in this section, we study the possibility of reducing the pilot overhead to a single pilot subcarrier. Therefore, only K time-frequency slots would be necessary for pilot allocation, reducing the training overhead by a factor of L .

In current standards, the subcarrier spacing Δf is chosen to be much shorter than the coherence bandwidth [10]. Hence, the channel can be considered flat across the frequency band of multiple adjacent subcarriers. This suggests that the channel

at a single subcarrier can be considered for the equalization of its adjacent subcarriers. Moreover, the correlation between the CFR of adjacent subcarriers can be obtained from the following approximation [11]

$$|\alpha_{\Delta m,k}| \approx \sqrt{1 - \left(\frac{\Delta f \Delta m}{F_c} \right)^2}, \quad (11)$$

where we define $\alpha_{\Delta m,k} \triangleq \mathbb{E}\{\bar{\lambda}_{m,k}[q] \bar{\lambda}_{((m+\Delta m)_M,k)}^*[q]\}$, $F_c = \frac{1}{\ell_\tau}$ is the coherence bandwidth and ℓ_τ is the maximum delay spread of the channel.

Taking this into account, we consider a reference pilot in one subcarrier and utilize its channel estimate for equalization of the neighboring subcarriers. The detected data symbols will be considered as virtual pilots, updating their channel estimates and allowing for the equalization of data at their adjacent subcarriers. We repeat this process, in a sliding manner, until all the data and the CFR in the UL frame are estimated. Hence, the data $\bar{\mathbf{X}}_m$, at any given subcarrier m , can be equalized using (10) and (11) with the CFR at subcarrier $m-1$ as

$$\hat{\mathbf{X}}_m^{\text{MMSE}} = |\Psi_1|^{-1} \mathbf{\Phi}_{m-1}^{\text{MMSE}} \bar{\mathbf{Y}}_m, \quad (12)$$

where $\Psi_{\Delta m}$ is a diagonal matrix with the diagonal elements formed by the vector $[\alpha_{\Delta m,0}, \dots, \alpha_{\Delta m,K-1}]$. In (12), we consider the approximation $(\hat{\mathbf{\Lambda}}_m)^H \hat{\mathbf{\Lambda}}_m \approx (\hat{\mathbf{\Lambda}}_{m-1})^H \hat{\mathbf{\Lambda}}_{m-1}$, and the phase of $\alpha_{1,k}$ to be negligible.

After obtaining $\hat{\mathbf{X}}_m^{\text{MMSE}}$, we perform hard decision by quantizing each symbol to their nearest quadrature amplitude modulated (QAM) constellation point, resulting in $\hat{\mathbf{X}}_m^{\text{HD}}$. With $\hat{\mathbf{X}}_m^{\text{HD}}$, the CFR estimate at subcarrier m can be updated as

$$\hat{\mathbf{\Lambda}}_m = \bar{\mathbf{Y}}_m (\hat{\mathbf{X}}_m^{\text{HD}})^\dagger. \quad (13)$$

This channel estimate is then used for equalization of the following subcarrier, $m+1$, using (12). In particular, since we now consider $\hat{\mathbf{X}}_m$ as the pilot, the noise mitigation term in (8) should be corrected accordingly. That is, the term $\bar{\mathbf{P}}_m^H$ in $\mathbf{\Gamma}_m$ and $\mathbf{\Phi}_m^{\text{MMSE}}$ should be substituted by $(\hat{\mathbf{X}}_m^{\text{HD}})^\dagger$. Then the procedure is repeated in a sliding manner until the whole frame is equalized. If the detected symbols $\hat{\mathbf{X}}_m^{\text{HD}}$ are erroneous, (13) provides imperfect channel estimates, and hence, the sliding technique propagates the error to further subcarriers. Therefore, as we will show in the later part of this section, we propose the concept of depth that significantly improves the accuracy of $\hat{\mathbf{X}}_m^{\text{HD}}$, alleviating the error propagation issue.

It is worth noting that $\bar{\mathbf{X}}_m$ can be a rank-deficient matrix. Specifically, if the rank of $\bar{\mathbf{X}}_m$ is lower than K , (13) fails to retrieve $\hat{\mathbf{\Lambda}}_m$. To solve this issue, we calculate $\alpha_{\Delta m,k}$ for any value of Δm . This way, $\hat{\mathbf{X}}_m^{\text{MMSE}}$ at any subcarrier can be detected with the same reference pilot using (12).

Proposition 1: In the asymptotic regime, it is possible to perform MRC with the channel estimates on a single subcarrier. In other words, the data at any given subcarrier m can be equalized with the CFR at any other subcarrier m' .

Proof We detect the data $\bar{\mathbf{X}}_m$, at any given subcarrier m , by performing MRC with the channel estimates at subcarrier m'

$$\hat{\mathbf{X}}_m^{\text{MRC}} = \mathbf{\Gamma}_{m'}^{-1} \mathbf{\Lambda}_{m'}^H \mathbf{Y}_m^d. \quad (14)$$

Considering (9) in the asymptotic regime, as $Q \rightarrow \infty$,

$$\frac{\bar{\lambda}_{m',k'}^H \bar{\lambda}_{m,k}}{Q} \rightarrow \begin{cases} \alpha_{\Delta m,k} & \text{if } k = k' \\ 0 & \text{otherwise} \end{cases}, \quad (15)$$

where $\Delta m = m - m'$. Thus, in (14), $\Gamma_{m'}^{-1} \Lambda_{m'}^H \Lambda_m \rightarrow \Psi_{\Delta m}$.

To obtain the exact value of $\alpha_{\Delta m,k}$, we substitute the CFRs with the expression given in (2). This yields a quadratic form, with the expectation defined as [12]

$$\mathbb{E}\{\tilde{\mathbf{h}}_{q,k}^H (M \mathbf{f}_{M,m}^* \mathbf{f}_{M,((m+\Delta m))}^T) \tilde{\mathbf{h}}_{q,k}\} = \text{tr}\{\mathbf{A} \tilde{\Sigma}_k\} + \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu}, \quad (16)$$

where $\boldsymbol{\mu}$ and $\tilde{\Sigma}_k$ represent the mean and covariance matrices of $\tilde{\mathbf{h}}_{q,k}$, respectively, and $\mathbf{A} = M \mathbf{f}_{M,m}^* \mathbf{f}_{M,((m+\Delta m))}^T$. With the parameters presented in section II, it is given that the channel is zero mean, i.e., $\boldsymbol{\mu} = \mathbf{0}_{M \times 1}$, and $\tilde{\Sigma}_k$ is a diagonal matrix with the diagonal elements formed by the vector $\tilde{\rho}_k = [\rho_k^T, \mathbf{0}_{M-L \times 1}^T]^T$. Hence, (16) reduces to

$$\alpha_{\Delta m,k} = \text{tr}\{\mathbf{A} \tilde{\Sigma}_k\} = \sqrt{M} \mathbf{f}_{M,((\Delta m))}^T \tilde{\rho}_k. \quad (17)$$

Based on this result, as long as the PDP is known, equalization can be performed using (14) followed by scaling the MRC output for each subcarrier by $\frac{1}{\alpha_{\Delta m,k}}$. ■

While *Proposition 1* is sufficient in the asymptotic regime, in practical systems, the aforementioned scaling by $\frac{1}{\alpha_{\Delta m,k}}$ may lead to noise enhancement and, consequently, a performance loss. For this reason, we rely on our proposed sliding equalization technique as $\alpha_{\Delta m,k}$ is larger for the smaller values of $|\Delta m|$. When $\bar{\mathbf{X}}_m$ is rank-deficient, $\hat{\Lambda}_m$ cannot be estimated. Thus, we resort to utilizing the previously estimated channels on the closest subcarrier. Furthermore, if the PDP is known at the receiver, the scaling term $\alpha_{\Delta m,k}$ can be calculated from (17), instead of using the approximation from (11).

When the coherence bandwidth is much larger than the subcarrier spacing, the correlation $\alpha_{\Delta m,k}$ takes large values for multiple values of Δm . This sparks the idea of taking advantage of the frequency diversity in addition to the spatial diversity. To this end, improved SINR performance can be achieved by performing equalization for each subcarrier multiple times, each with the channel of a different subcarrier within the coherence bandwidth. This improvement substantially reduces the error propagation issue from (13).

While the channels from all subcarriers could be considered for the averaging, the channels with a small correlation would bring minimal improvement to the output. Since, the value of $\alpha_{\Delta m,k}$ decreases as we increase the value of $|\Delta m|$, only channels of subcarriers within a range are considered. Hence, we call the maximum range considered as the depth, D .

Ideally, the depth would consider all adjacent subcarriers, at higher and lower indices. However, since the channels are updated in a sliding manner, only the channel on one side of the subcarriers is available at the time of equalization. Therefore, the final output of our proposed technique is achieved with two steps. We perform the first step of the proposed technique by sliding from lower to upper subcarrier indices until the whole frame is equalized. For the second step, the procedure is realized again, however, now sliding from higher to lower subcarrier indices. It is worth noting that both steps can be

Algorithm 1 Proposed sliding technique with reference pilot at subcarrier index i .

```

1: Initialize: Estimate channel at reference subcarrier (Pilot),
   obtaining  $\hat{\Lambda}_i$  using (4).
2: for  $\xi = [-1, 1]$  do
3:    $m' = i$  {Define the anchor for closest available CFR}
4:   for  $m_{\text{dir}} = 1$  to  $M - 1$  do
5:      $m = ((i + \xi m_{\text{dir}}))_M$ 
6:     Initialize  $\hat{\mathbf{X}}_{m,\xi} = \mathbf{0}_{K \times N}$ 
7:      $c = 0$  {Define total channels used for equalization}
8:     for  $\Delta m = 1$  to  $D$  do
9:       if  $\hat{\Lambda}_{((m-\xi\Delta m))_M}$  was estimated then
10:         $\hat{\mathbf{X}}_{m,\xi} = \hat{\mathbf{X}}_{m,\xi} + \Psi_{\xi\Delta m}^{-1} \Phi_{((m-\xi\Delta m))_M}^{\text{MMSE}} \bar{\mathbf{Y}}_m$ 
11:         $c = c + 1$ 
12:      end if
13:    end for
14:    if  $c == 0$  then
15:       $\hat{\mathbf{X}}_{m,\xi} = \Psi_{m-m'}^{-1} \Phi_{m'}^{\text{MMSE}} \bar{\mathbf{Y}}_m$ 
16:    else
17:       $\hat{\mathbf{X}}_{m,\xi} = \frac{\hat{\mathbf{X}}_{m,\xi}}{c}$ 
18:    end if
19:     $\hat{\mathbf{X}}_{m,\xi}^{\text{HD}}$  is obtained from hard decision of  $\hat{\mathbf{X}}_{m,\xi}$ 
20:    if  $\hat{\mathbf{X}}_{m,\xi}^{\text{HD}}$  is invertible then
21:       $\hat{\Lambda}_m = \bar{\mathbf{Y}}_m (\hat{\mathbf{X}}_m^{\text{HD}})^\dagger$  and update  $m' = m$ 
22:    else
23:       $\hat{\Lambda}_m$  is not estimated and  $m'$  is not updated
24:    end if
25:  end for
26: end for
27: Obtain  $\hat{\mathbf{X}}_m$  by averaging the two sets of result

```

realized in parallel as the steps are independent of each other. The output of both steps are then averaged to obtain the final output of our proposed technique. We distinguish each step with the direction variable $\xi \in [-1, 1]$ representing steps one and two, respectively. Each step can be represented as

$$\hat{\mathbf{X}}_{m,\xi} = \frac{1}{D} \sum_{\Delta m=1}^D \Psi_{\xi\Delta m}^{-1} \Phi_{m-\xi\Delta m}^{\text{MMSE}} \bar{\mathbf{Y}}_m. \quad (18)$$

The final output is then obtained as $\hat{\mathbf{X}}_m = \frac{1}{2}(\hat{\mathbf{X}}_{m,-1} + \hat{\mathbf{X}}_{m,1})$. Our proposed technique is summarized in Algorithm 1.

In the following section, we numerically evaluate the performance of our proposed technique. We compare our solution with the conventional LS-based channel estimation and MMSE combining. We show that our proposed channel estimation and equalization method, using only K pilots, outperforms the conventional method, which requires LK pilots.

V. NUMERICAL RESULTS

In this section, we evaluate the efficacy of our proposed technique and compare it with the conventional LS-based channel estimation and MMSE equalization. In our simulations, we consider the UL transmission of a single frame with 16-QAM and OFDM modulation, $M = 1024$ subcarriers and $N = 14$ time symbols. The BS serves $K = 7$ users, where $N_p = 7$ time slots are allocated to each user for channel estimation and $N_d = 7$ for data transmission. We use the

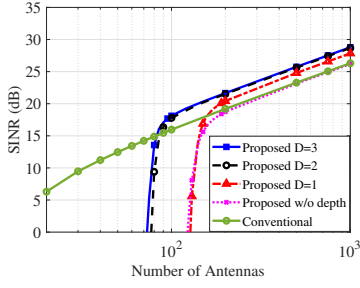


Fig. 1. SINR performance as a function of the number of BS antennas.

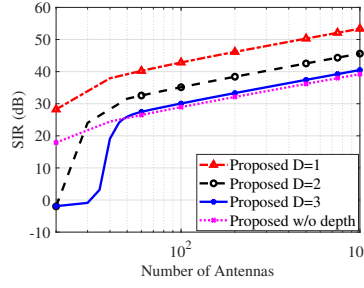


Fig. 2. SIR performance as a function of the number of BS antennas.

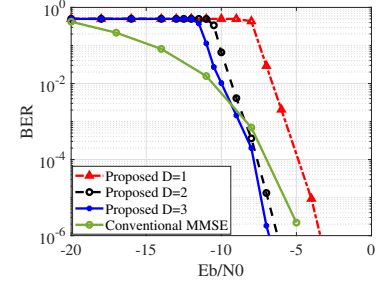


Fig. 3. BER performance for our proposed technique and MMSE combining.

Extended Typical Urban model (ETU) channel model, [13]. ETU channel model has the smallest coherence bandwidth, i.e., $F_c \approx 200$ kHz, among the 3GPP channel models [13]. Since our proposed method benefits from larger coherence bandwidths, we chose this model to show the efficacy of our proposed technique. Considering subcarrier spacing of $\Delta f = 15$ kHz and transmission bandwidth of 15 MHz, the channel length is $L = 77$. Hence, for the conventional LS-based channel estimators, a total of $LN_p = 539$ time-frequency resources are allocated for pilot transmission while our proposed technique requires only $N_p = 7$ time-frequency resources. Perfect knowledge of the PDP for each user at the BS is assumed. These results were obtained through the ensemble average of 5000 random trials.

In Fig. 1, we evaluate the SINR performance of our proposed technique versus the number of BS antennas for $D = 1, 2, 3$, and without implementation of *depth*. When *depth* is not implemented, we only consider one sliding direction where one estimate of the data symbol at each subcarrier is obtained, i.e., the output is obtained from $\hat{\mathbf{X}}_m = \Psi_1^{-1} \Phi_{m-1}^{\text{MMSE}} \mathbf{Y}_m$. An input SNR of 0 dB was considered. The results in Fig. 1 show that our proposed technique can effectively average out noise and multiuser interference as the number of BS antennas grows large. Furthermore, our proposed technique, without the implementation of *depth*, can achieve very close performance to that of the conventional MMSE combiner. This is while for $D = 1, 2$ and 3, it leads to the SINR performance gain of around 2 dB compared to the conventional MMSE combining. The SINR performance of our proposed technique becomes linear only after deploying a minimum number of BS antennas. This is due to error propagation at the channel estimation stage in (13). This highlights the benefit of increasing *depth* as it greatly improves the SINR performance, especially when considering fewer BS antennas or lower input SNR. Furthermore, the frequency diversity predominantly averages out the channel estimation noise in (4). However, as it is shown in Fig. 2, in the absence of noise, increasing *depth* adversely affects the signal-to-interference ratio (SIR). This is due to the approximation in (12). The conventional LS channel estimation and MMSE combining lead to infinite SIR in the absence of noise and thus, cannot be included in the results of Fig. 2.

Finally, in Fig. 3, we analyze the BER performance of our proposed technique when $Q = 200$ BS antennas are deployed. The proposed technique with $D = 3$ leads to 2 dB performance gain at large E_b/N_0 s compared to the conventional method. These results show that increasing D from 1 to 3 provides

almost 4 dB performance gain.

VI. CONCLUSION

In this letter, we proposed a sliding joint channel estimation and equalization technique that significantly reduces the pilot overhead in massive MIMO. With the CFR estimates on a single reference subcarrier, we showed that the data symbols on the adjacent subcarriers can be detected with linear combining. The detected data symbols are then utilized as virtual pilots to update their corresponding channel estimates. The updated CFRs are then used for equalization of the remaining data symbols, in a sliding manner. With this approach, we obtain multiple estimates of the transmitted data symbols at each subcarrier that are averaged to provide additional frequency diversity to the spatial diversity gains of massive MIMO. Through extensive numerical analysis, we showed that our proposed technique achieves improved SINR and BER performance compared to the conventional linear combiners.

REFERENCES

- [1] H. Q. Ngo and E. G. Larsson, "No downlink pilots are needed in TDD massive MIMO," *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 2921–2935, 2017.
- [2] M. A. Albreem, M. Juntti, and S. Shahabuddin, "Massive MIMO detection techniques: A survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3109–3132, 2019.
- [3] E. Nayebi and B. D. Rao, "Semi-blind channel estimation for multiuser massive MIMO systems," *IEEE Transactions on Signal Processing*, vol. 66, no. 2, pp. 540–553, 2018.
- [4] D. Verenzuela, E. Björnson, X. Wang, M. Arnold, and S. ten Brink, "Massive-MIMO iterative channel estimation and decoding (MICE) in the uplink," *IEEE Trans. on Commun.*, vol. 68, no. 2, pp. 854–870, 2020.
- [5] Y. Huang, J. Zhang, H. Wang, Z. Zhao, and H. Yin, "Spatial multiplexing superimposed pilots for multicell multiuser MIMO uplink systems," *IEEE Systems Journal*, vol. 16, no. 2, pp. 3151–3162, 2022.
- [6] M. Masood, L. H. Afify, and T. Y. Al-Naffouri, "Efficient coordinated recovery of sparse channels in massive MIMO," *IEEE Transactions on Signal Processing*, vol. 63, no. 1, pp. 104–118, 2015.
- [7] 3GPP, "NR; physical channels and modulation," *3rd Generation Partnership Project (3GPP), Technical Specification (TS) 38.211*, vol. 9, 2018.
- [8] E. Björnson, E. G. Larsson, and M. Debbah, "Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?" *IEEE Trans. Wirel. Commun.*, vol. 15, no. 2, pp. 1293–1308, 2016.
- [9] H. Hosseiny, A. Farhang, and B. Farhang-Boroujeny, "FBMC receiver design and analysis for medium and large scale antenna systems," *IEEE Trans. on Veh. Technol.*, vol. 71, no. 3, pp. 3044–3057, 2022.
- [10] T. ETSI, "138 912, 5g," *Study on New Radio (NR) access technology (3GPP TR 38.912 version 14.0.0 Release 14)*, vol. 14, 2018.
- [11] F. Hlawatsch and G. Matz, *Wireless communications over rapidly time-varying channels*. Academic press, 2011.
- [12] A. M. Mathai and S. B. Provost, *Quadratic forms in random variables: theory and applications*. Dekker, 1992.
- [13] 3GPP, "Evolved universal terrestrial radio access (E-UTRA); base station (BS) radio transmission and reception," *3rd Generation Partnership Project (3GPP), TS 36.104 V15.3.0*, 2018.