Distributed Spatially Non-Stationary Channel Estimation for Extremely-Large Antenna Systems

Yanqing Xu[1], Shuai Wang[2,*], Ruihong Jiang[3], Zhou Wang[4]

[1]School of Science and Engineering, the Chinese University of Hong Kong, Shenzhen 518172, China
[2]Information Systems Technology and Design, Singapore University of Technology and Design, 487372, Singapore
[3]State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China
[4]School of Integrated Circuits, Tsinghua University, Beijing 100084, China

Abstract

The purpose of this paper is to develop a distributed channel estimation (CE) algorithm for spatially non-stationary (SNS) channels in extremely large aperture array systems, addressing the issues of high communication cost and computational complexity associated with traditional centralized algorithms. However, SNS channels differ from conventional spatially stationary channels, presenting new challenges such as varying sparsity patterns for different antennas. To overcome these challenges, we propose a novel distributed CE algorithm accompanied by a simple-yet-effective hard thresholding scheme. The proposed algorithm is not only suitable for uniform antenna arrays but also for irregularly deployed antennas. Simulation results demonstrate the advantages of the the proposed algorithm in terms of estimation accuracy, communication cost and computational complexity.

1. Introduction

A massive multiple-input multiple-output (MIMO) system is equipped with a much larger number of antennas, typically tens or even hundreds [1]. This offers several benefits, such as a significant increase in capacity, simplification of scheduling in the frequency domain [1], and the ability to average interference using the large number theorem [2]. Therefore, massive MIMO has become a key technology in 5G wireless communication systems operating in sub-6 GHz bands. To fully utilize the spatial multiplexing gains and the array gains of massive MIMO, accurate channel state information (CSI) is essential in wireless communication systems and is the foundation of many key physical algorithms such as the multi-user detection [3] and beamforming [4–6]. However, CSI has to be estimated using training symbols in practice,

and the complexity scales with the number of transmit antennas.

Traditional channel estimation algorithms are typically based on the centralized baseband processing (CBD) architecture [7–9]. However, this architecture is faced with challenges such as high communication cost and computational complexity. These issues require a powerful and expensive CBP unit, which may not be feasible for large antenna sizes. To address these challenges, the decentralized baseband processing (DPB) architecture has been proposed [10]. In the DPB architecture, antennas are divided into clusters, each equipped with an independent and more affordable baseband processing unit (BBU) (see Fig. 1). A naive approach under the DPB architecture is for each BBU to perform channel estimation based solely on its locally received signal. However, this fully decentralized scheme would result in significant performance loss due to the neglect of inter-cluster correlations. To address this issue, the DPB architecture leverages advanced distributed signal processing (SP) techniques

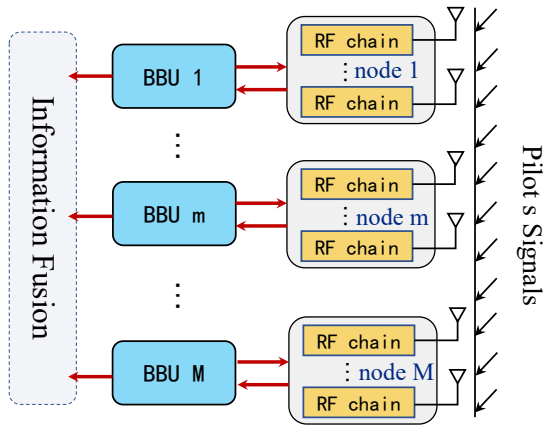*Corresponding author. Email: shuai_wang@sutd.edu.sg

Figure 1. Illustration of the DBP architecture, where the "RF chain" is the radio frequency chain connecting to each antenna.

to achieve promising CE performance while minimizing inter-BBU communication cost and BBU computation complexity. In recent years, researchers have explored a range of distributed and decentralized signal processing algorithms for DBP-based massive MIMO systems, as documented in previous studies [10–16]. However, the majority of these investigations have focused on the uplink equalization or downlink precoding problems, with few studies exploring the design of distributed CE algorithms.

To accommodate a larger number of terminals in the future, an approach that may be adopted is the deployment of thousands of antennas in a specific geographic area such as along the walls of a building or stadium structure [17, 18]. This approach, referred to as extremely large aperture array (ELAA), exhibits channel characteristics that differ from conventional MIMO systems. Non-stationarity of massive MIMO channels is observed in two aspects. Firstly, with a large number of base station (BS) antennas, the distance between the BS array and a scatterer in the propagation channel may become smaller than the Rayleigh distance, rendering the far-field and plane wavefront assumptions invalid. As a result, the wavefront should be modeled as a spherical wavefront to account for drifts in the received power, delays, and directions of incoming/outgoing multipaths along the BS array. Secondly, scatterer clusters dynamically appear and disappear, making some clusters visible to only a part of the BS array, which may cause variations in the number of multipaths on the array axis and enlarge the variations in the received power, delays, and directions of multipaths across different BS antennas[19]. As a result, the channel of ELAA system may be spatially non-stationary and the channels of antennas would have different sparsity patterns, and

the channel estimation in ELAA systems presents a new challenge.

In this paper, we aim to develop distributed channel estimation (CE) algorithms for spatially non-stationary (SNS) channels in extremely large aperture array systems to overcome the challenges of high communication cost and computational complexity of the traditional centralized algorithm. On the other hand, different to the conventional spatially stationary channels, the SNS channel brings new challenges, e.g., the sparsity patterns for different antennas may be different. To resolve these challenges, we proposed a novel distributed CE algorithm, together with a simple-yet-effective hard thresholding scheme. The proposed algorithm is not only applicable to the uniform antenna array, but also suitable to the irregularly deployed antennas. Simulation results indicate that our proposed distributed CE algorithm can substantially enhance the accuracy of channel estimation in spatially non-stationary channels.

2. System Model and Problem Description

2.1. Spatially Non-Stationary Channel

In this work, we consider the channel estimation in a massive MIMO system where a base station (BS) equipped with N_R antennas communicates with a single-antenna user through N_C subcarriers. While we will primarily focus on the uplink channels, the concepts presented in this report can also be extended to the downlink with appropriate pilot/training designs. Specifically, in the uplink, the received signal at the BS via the k-th subcarrier is given by

$$\mathbf{y}(k) = \mathbf{h}(k) + \mathbf{w}(k), k = 1, ..., N_C. \tag{1}$$

where $\mathbf{y}(k), \mathbf{w}(k) \in \mathbb{C}^{N_R}$ are the received signal and noise at the receiver, respectively, and $\mathbf{h}(k) \in \mathbb{C}^{N_R}$ denotes the channel vector between BS and user. By concatenating all the received signal, the received signal at the receiver can be written as the matrix representation as follows: where $\mathbf{y}(k), \mathbf{w}(k) \in \mathbb{C}^{N_R}$ are the received signal and noise at the receiver, respectively, and $\mathbf{h}(k) \in \mathbb{C}^{N_R}$ denotes the channel vector between BS and user. By concatenating all the received signal, the received signal at the receiver can be written as the matrix representation as follows:

$$\mathbf{Y} = \mathbf{H} + \mathbf{W} \tag{2}$$
$$= \widetilde{\mathbf{H}}\mathbf{F}_{N_C}^H + \mathbf{W}. \tag{3}$$

where $\mathbf{W} \in \mathbb{C}^{N_R \times N_C}$ is the additive white Gaussian noise, and the antenna-and-frequency-domain channel matrix $\mathbf{H}$ can be expressed by an antenna-and-delay-domain channel matrix $\widetilde{\mathbf{H}} \in \mathbb{C}^{N_R \times N_C}$ as follows

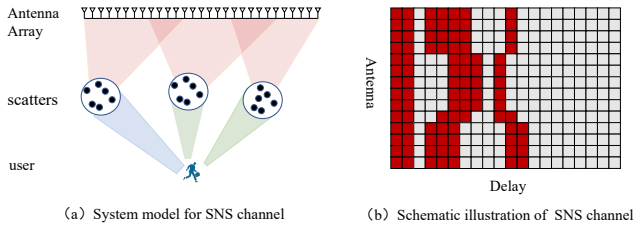$$\widetilde{\mathbf{H}} = \mathbf{H}\mathbf{F}_{N_C} \tag{4}$$

Figure 2. Illustration of the spatially non–stationary wireless channel, where (a) shows that only part of the antenna elements can see the scatters due to the large size of antenna array, and (b) implies that the delay–domain sparsities for different antennas are different.

where $\mathbf{F}_{N_R} \in \mathbb{C}^{N_R \times N_R}$ and $\mathbf{F}_{N_C} \in \mathbb{C}^{N_C \times N_C}$ are the antenna-domain and delay-domain DFT matrices respectively. The DFT matrices are defined as

$$[\mathbf{F}_{N_C}]_{mn} = \frac{1}{\sqrt{N_C}} e^{-j \frac{2\pi}{N_C} mn}, \; m, n = 1, ..., N_C, \quad (5)$$

where $j = \sqrt{-1}$.

Massive MIMO systems feature a large antenna array, causing the channel parameters, including power, delay spread, angle spread, and the number of clusters, to vary over the wide-scale array. As a result, the wireless channel becomes spatially non-stationary, particularly when the number of antennas is significantly large. In such channels, some clusters are only visible to a portion of the antenna array [3], causing the delay-domain channel sparse structures to differ among the antennas, as illustrated in Fig. 2.

2.2. Signal Model Based on DBP Architecture

Under the DBP architecture, the antennas are divided into multiple non-overlapping antenna clusters, and each antenna cluster connects to a dedicated BBU to handle its received signal. Each antenna cluster and its BBU are viewed as a local node. In this work, we consider the star network, where a central node coordinates the local nodes to performs the distributed channel estimation to approach the centralized scheme, and we assume that the local nodes can only communicate with the central node, as shown in Fig. 1.

We assume that each antenna cluster consists of N_r antennas and $N_r = \frac{N_R}{M}$, where M is the number of clusters. Then, the received antenna-and-frequency-domain signal at the m-th node is given by

$$\mathbf{Y}_m = \mathbf{H}_m + \mathbf{W}_m, m \in \mathcal{M}, \quad (6)$$

where $\mathcal{M} \triangleq \{1, ..., M\}$, $\mathbf{H}_m \in \mathbb{C}^{N_r \times N_C}$ and $\mathbf{W}_m$ are the antenna-and-frequency-domain channel and noise at node m, respectively. We assume that each column of

$\mathbf{W}_m$ follows $\mathcal{CN}(0, \sigma_w^2 \mathbf{I}_{N_r})$ and is independent with other columns. The goal of distributed CE is to recover $\mathbf{H}_m$ from the received signals, $\mathbf{Y}_m$s' from all nodes.

3. Baseline Schemes

Before presenting the proposed algorithm, we will first revisit the MMSE scheme and the diagonal MMSE based centralized schemes.

MMSE Scheme: The MMSE solution can be obtained by solving the following optimization problem

$$\min_{\mathbf{U}} \; \mathbb{E}\left[\|\mathbf{U}\left(\mathbf{F}_{N_C}^H \otimes \mathbf{F}_{N_R}^H\right)\mathbf{y} - \left(\mathbf{F}_{N_C}^H \otimes \mathbf{F}_{N_R}^H\right)\mathbf{h}\|_F^2\right], \quad (7)$$

where $\mathbf{y} = \text{vec}(\mathbf{Y})$, $\mathbf{h} = \text{vec}(\mathbf{H})$, $\mathbf{U} \in \mathbb{C}^{N_R N_C \times N_R N_C}$ is the optimized matrix to minimize the mean square error, and $\|\cdot\|_F$ denotes the Frobenius norm. Define $\widetilde{\mathbf{y}}$ and $\widetilde{\mathbf{h}}$ as follows:

$$\widetilde{\mathbf{y}} = \left(\mathbf{F}_{N_C}^H \otimes \mathbf{F}_{N_R}^H\right)\mathbf{y}, \quad (8a)$$

$$\widetilde{\mathbf{h}} = \left(\mathbf{F}_{N_C}^H \otimes \mathbf{F}_{N_R}^H\right)\mathbf{h} \quad (8b)$$

Note that (7) is a quadratic optimization problem, thus the optimal $\mathbf{U}$ can be obtained by checking the first-order derivation of the objective function, which is given by

$$\mathbf{U} = \mathbf{R}_{\widetilde{\mathbf{y}}\widetilde{\mathbf{h}}}\mathbf{R}_{\widetilde{\mathbf{y}}}^{-1}, \quad (9a)$$

$$= \mathbf{R}_{\widetilde{\mathbf{h}}}\left(\mathbf{R}_{\widetilde{\mathbf{h}}} + \sigma_w^2 \mathbf{I}\right)^{-1}, \quad (9b)$$

where (9b) is due to the assumption that $\widetilde{\mathbf{h}}$ and $\widetilde{\mathbf{w}}$ are independent. Therefore, the MMSE channel estimator is given by

$$\widehat{\mathbf{H}} = \mathbf{F}_{N_R}\text{mat}\left(\mathbf{U}\left(\mathbf{F}_{N_C}^H \otimes \mathbf{F}_{N_R}^H\right)\mathbf{y}\right)\mathbf{F}_{N_C} \quad (10)$$

where $\text{mat}(\mathbf{y})$ is the inverse operation of matrix vectorization to reshape vector $\mathbf{y}$ to a matrix which has the same dimension as $\mathbf{Y}$.

Diagonal MMSE Scheme : In this scheme, the optimization problem becomes

$$\min_{\substack{\mathbf{U} \\ \text{is diagonal}}} \; \mathbb{E}\left[\|\mathbf{U}\left(\mathbf{F}_{N_C}^H \otimes \mathbf{F}_{N_R}^H\right)\mathbf{y} - \left(\mathbf{F}_{N_C}^H \otimes \mathbf{F}_{N_R}^H\right)\mathbf{h}\|_F^2\right], \quad (11)$$

Different from problem (7), $\mathbf{U} \in \mathbb{C}^{N_R N_C \times N_R N_C}$ is a diagonal matrix. By optimizing problem (11), we arrive at the following solution

$$\mathbf{U} = \text{diag}\left(\left\{\frac{[\mathbf{R}_{\widetilde{\mathbf{h}}}]_{\ell\ell}}{[\mathbf{R}_{\widetilde{\mathbf{h}}}]_{\ell\ell} + \sigma_w^2}\right\}_{\ell=1}^{N_C N_R}\right), \quad (12)$$

where $[\mathbf{R}_{\widetilde{\mathbf{h}}}]_{\ell\ell}$ is the ℓ-th element of $\mathbf{R}_{\widetilde{\mathbf{h}}}$. Then, the channel estimate can be obtained in the same way as the MMSE scheme.

Remark 1. In the MMSE scheme, both delay-domain and antenna-domain correlations are taken into consideration through the frequency-to-delay and antenna-to-angle IDFT operations, respectively. In addition, by designing the matrix $\mathbf{U}$, correlations among all channel elements are accounted for. On the other hand, the diagonal MMSE scheme only considers delay-domain and antenna-domain correlations, resulting in inferior performance compared to the MMSE scheme. However, the diagonal MMSE scheme has significantly lower computational complexity than the full MMSE scheme. Specifically, the computational complexity order of the diagonal MMSE scheme and the MMSE scheme are given by $\mathcal{O}(N_R N_C)$ and $\mathcal{O}(N_R^3 N_C^3)$, respectively.

Inspired by the aforementioned schemes, we will introduce baseline schemes for the nonstationary scenario. Our proposed approach considers delay-domain and antenna-domain correlations without necessitating antenna-to-angle IDFT operations. We will begin by outlining the centralized scheme, followed by the fully distributed scheme.

Centralized Scheme : In the centralized scheme, we solve the following optimization problem:

$$\min_{\mathbf{U}} \ \mathbb{E}\left[\left\| \mathbf{U}\left(\mathbf{F}_{N_C}^H \otimes \mathbf{I}_{N_R}\right)\mathbf{y} - \left(\mathbf{F}_{N_C}^H \otimes \mathbf{I}_{N_R}\right)\mathbf{h} \right\|_F^2 \right], \quad (13)$$

where $\mathbf{y} = \text{vec}(\mathbf{Y})$, $\mathbf{h} = \text{vec}(\mathbf{H})$, $\mathbf{I}_{N_R}$ is an identity matrix of size N_R, $\mathbf{U} \in \mathbb{C}^{N_R N_C \times N_R N_C}$ is a block diagonal matrix which is optimized to minimize the mean square error. More explicitly, $\mathbf{U}$ has the following structure:

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{U}_{N_C} \end{bmatrix}, \quad (14)$$

where $\mathbf{U}_\ell \in \mathbb{C}^{N_R \times N_R}, \ell = 1, ..., N_C$.

Remark 2. Compared with problem (11) for the centralized scheme in stationary scenario, there are two modifications in problem (13). First, the antenna-to-angle-domain IDFT is removed. Therefore, the algorithm is applicable to the scenario where the antennas are not regularly deployed (e.g., uniform linear array (ULA) and uniform planar array (UPA). Second, $\mathbf{U} = \text{blkdiag}\left(\mathbf{U}_1, \mathbf{U}_2, \cdots, \mathbf{U}_{N_C}\right)$ is a block-diagonal matrix where $\mathbf{U}_\ell, \ell = 1, \cdots N_C$, is optimized to estimate the ℓ-th path by taking the correlation of all antennas into consideration.

Then, define $\widetilde{\mathbf{Y}}$ and $\widetilde{\mathbf{H}}$ as follows:

$$\widetilde{\mathbf{Y}} = \mathbf{Y}\mathbf{F}_{N_C}^H, \ \widetilde{\mathbf{H}} = \mathbf{H}\mathbf{F}_{N_C}^H. \quad (15)$$

By exploiting the block structure of $\mathbf{U}$, problem (13) can be equivalently rewritten as

$$\min_{\mathbf{U}} \ \mathbb{E}\left[\sum_{\ell=1}^{N_C} \|\mathbf{U}_\ell \widetilde{\mathbf{y}}_\ell - \widetilde{\mathbf{h}}_\ell\|_F^2 \right], \quad (16)$$

where $\widetilde{\mathbf{y}}_\ell$ and $\widetilde{\mathbf{h}}_\ell$ are the ℓ-th column (corresponds to the ℓ-th path) of $\widetilde{\mathbf{Y}}$ and $\widetilde{\mathbf{H}}$, respectively. By using the decoupling structure of (16), each $\mathbf{U}_\ell$ can be optimized individually and the optimal $\mathbf{U}_\ell \in \mathbb{C}^{N_R \times N_R}$ is given by

$$\mathbf{U}_\ell = \mathbf{R}_{\widetilde{\mathbf{y}}_\ell \widetilde{\mathbf{h}}_\ell} \mathbf{R}_{\widetilde{\mathbf{y}}_\ell}^{-1}, \ell = 1, ..., N_C. \quad (17)$$

Then, the channel estimate can be obtained accordingly. The computational complexity order is given by $\mathcal{O}(N_C N_R^3)$. Next, let us describe the details of the fully decentralized scheme.

Fully Decentralized (FD) Scheme : In FD scheme, the received signal of cluster i is given by

$$\mathbf{Y}_m = \mathbf{H}_m + \mathbf{W}_m, m = 1, ..., M, \quad (18)$$

where M is number of antenna clusters and we assume each cluster consists of $N_r = N_R/M$ antennas. Each cluster uses their own received signal to estimate the local channel. In particular, cluster i needs to solving the following optimization problem:

$$\min_{\mathbf{U}_m} \ \mathbb{E}\left[\|\mathbf{U}_m\left(\mathbf{F}_{N_C}^H \otimes \mathbf{I}_{N_r}\right)\mathbf{y}_m - \left(\mathbf{F}_{N_C}^H \otimes \mathbf{I}_{N_r}\right)\mathbf{h}_m\|_F^2 \right], \quad (19)$$

where $\mathbf{y}_m = \text{vec}(\mathbf{Y}_m)$, $\mathbf{h}_m = \text{vec}(\mathbf{H}_m)$, $\mathbf{I}_{N_r}$ is an identity matrix of size N_r, and $\mathbf{U}_m = \text{blkdiag}(\mathbf{U}_{m1}, \mathbf{U}_{m2}, \cdots, \mathbf{U}_{mN_C}) \in \mathbb{C}^{N_r N_C \times N_r N_C}$ is a block diagonal matrix. $\mathbf{U}_{m\ell} \in \mathbb{C}^{N_r \times N_r}, m = 1, ..., M, \ell = 1, \cdots N_C$, is optimized to estimate the ℓ-th path with the antennas in cluster m. Same as the centralized scheme, cluster i solves

$$\min_{\mathbf{U}_m} \ \mathbb{E}\left[\sum_{\ell=1}^{N_C} \|\mathbf{U}_{m\ell}\widetilde{\mathbf{y}}_{m\ell} - \widetilde{\mathbf{h}}_{m\ell}\|_F^2 \right], \quad (20)$$

where $\widetilde{\mathbf{y}}_{m\ell}$ and $\widetilde{\mathbf{h}}_{m\ell}$ are the ℓ-th column of $\widetilde{\mathbf{Y}}_m$ and $\widetilde{\mathbf{H}}_m$, respectively. Similar to the centralized scheme, the optimal $\mathbf{U}_{m\ell}$ is given by

$$\mathbf{U}_{m\ell} = \mathbf{R}_{\widetilde{\mathbf{y}}_{m\ell}\widetilde{\mathbf{h}}_{m\ell}} \mathbf{R}_{\widetilde{\mathbf{y}}_{m\ell}}^{-1}, m = 1, ..., M, \ell = 1, ..., N_C. \quad (21)$$

Then, each cluster can estimate its local channel accordingly. The computational complexity order is given by $\mathcal{O}(M N_C N_r^3)$.

To show the efficacy of the baseline schemes, the associated NMSE performance is shown in Fig. 3. In the simulation, the channels are generated by the âĂIJ3GPP-38.901-UMa-NLOSâĂİ model in âĂIJQuaDRiGaâĂİ with 256 antennas, 1024 subcarriers, and 20 MHz
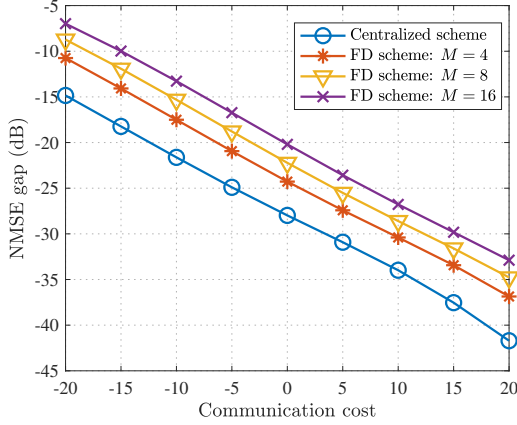
Figure 3. NMSE performance of the baseline schemes in non–stationary scenario.

bandwidth[20]. The antennas are uniformly and linearly deploy. From this figure, one can see that the NMSE decreases as the increase of SNR and the centralized scheme can outperform the FD scheme, and the NMSE gap increases as the number of the clusters. This observation calls for urgent designs of efficient distributed CE algorithms to approach the centralized scheme while having both low communication cost and low computational complexity.

4. Proposed Distributed CE Algorithm

In this subsection, we will introduce the aggregation based distributed scheme in the nonstationary scenario. Take the STAR network for example. Let $\mathbf{D}_m \in \mathbb{R}^{N_R \times N_C}$ be a local windowing matrix in cluster i, $i = 1, \ldots, M$. Each cluster i obtains a hard windowed signal

$$\bar{\mathbf{Y}}_m = \mathbf{D}_m \odot \mathbf{Y}_m \mathbf{F}_{N_C}^H, \quad m = 1, \ldots, M, \tag{22}$$

where the hard-thresholding matrix $\mathbf{D}_m$ with $[\mathbf{D}_m]_{mn} \in \{0, 1\}$, $m = 1, \ldots, N_R$, $n = 1, \ldots, N_C$ is to decide which elements in $\widetilde{\mathbf{Y}}_m$ should be preserved. Remind that in the stationary scenario, the hard-windowing is taken by exploring the column-sparsity structure. However, in the nonstationary cases, some scatters may only be seen by partial antennas. Consequently, the strictly stationarity of delay-domain channel maynot hold. In other words, in some column of delay-domain channel, there may exist partial sparsity. Due to the above reasons, the hard windowing in nonstationary scenario is completed in two steps:

1) Inter-column sparsity-aware hard thresholding: The column-wise hard thresholding operation is carried out to get a column sparse signal. So, $[\mathbf{D}_m]_m$ is designed

by

$$[\mathbf{D}_m]_n = \begin{cases} \mathbf{1}_{N_r}, & \text{if, } \|[\widetilde{\mathbf{Y}}_m]_m\|_2^2 \geq A N_r \sigma_w^2, \\ \mathbf{0}_{N_r}, & \text{otherwise,} \end{cases} \tag{23}$$

where σ_w^2 is the noise power and A is the tunable threshold for hard windowing, $[\mathbf{D}_m]_n$ denotes the n-th column of $\mathbf{D}_m$, $\mathbf{1}_{N_R}$ and $\mathbf{0}_{N_R}$ denote all-one and all-zero column vectors of size N_R respectively.

2) Intra-column sparsity-aware hard thresholding: The element-wise hard thresholding operation is implemented on the obtained column-sparse signal and $[\mathbf{D}_m]_{mn}$ is designed by

$$[\mathbf{D}_m]_{mn} = \begin{cases} 1, & \text{if, } \|[\widetilde{\mathbf{Y}}_m]_{mn}\|_2^2 \geq A \sigma_w^2, \\ 0, & \text{otherwise,} \end{cases} \tag{24}$$

With hard windowing the removed signal is given by

$$\bar{\mathbf{Y}}_m^{(c)} = \mathbf{Y}_m \mathbf{F}_{N_C}^H - \bar{\mathbf{Y}}_m \in \mathbb{C}^{N_r \times N_C}, m = 1, \ldots, M, \tag{25}$$

which can be used to estimate the corresponding channel.

After hard thresholding, local node m sends $\bar{\mathbf{Y}}_m$ to the central node. Then, the central node merges the received (sparse) signals by $\bar{\mathbf{Y}} = [\bar{\mathbf{Y}}_1^H, \bar{\mathbf{Y}}_2^H, \ldots, \bar{\mathbf{Y}}_M^H]^H$. With $\bar{\mathbf{Y}}$, the central node solves the following optimization problem

$$\min_{\mathbf{U}} \mathbb{E}\left[\|\mathbf{U}\bar{\mathbf{y}} - \bar{\mathbf{h}}\|_F^2\right], \tag{26}$$

The optimal $\mathbf{U}^*$ is obtained in the same way as (16). Then the central node gets the estimated antenna-delay-domain channel $\widehat{\widetilde{\mathbf{H}}} = \left[\widehat{\widetilde{\mathbf{H}}}_1^H, \widehat{\widetilde{\mathbf{H}}}_2^H, \ldots, \widehat{\widetilde{\mathbf{H}}}_M^H\right]^H$ by

$$\widehat{\widetilde{\mathbf{H}}} = \text{mat}\left(\mathbf{U}^* \bar{\mathbf{y}}\right), \tag{27}$$

where $\bar{\mathbf{y}} = \text{vec}(\bar{\mathbf{Y}})$. Then the central node sends $\widehat{\widetilde{\mathbf{H}}}_m$, $m = 1, \ldots, M$ to each local node.

Then each local node uses $\bar{\mathbf{Y}}_m^{(c)}$ to estimate the corresponding channels by

$$\min_{\mathbf{U}^{(c)}} \mathbb{E}\left[\|\mathbf{U}^{(c)} \bar{\mathbf{y}}^{(c)} - \bar{\mathbf{h}}^{(c)}\|_F^2\right], \tag{28}$$

and the channel estimate is given by

$$\widehat{\widetilde{\mathbf{H}}}^{(\mathbf{c})} = \text{mat}\left(\mathbf{U}^{(c)} \bar{\mathbf{y}}^{(c)}\right), \tag{29}$$

and then the antenna-delay local estimate is given by

$$\widehat{\widetilde{\mathbf{G}}}_m = \widehat{\widetilde{\mathbf{H}}} + \widehat{\widetilde{\mathbf{H}}}^{(\mathbf{c})}, m = 1, \ldots, M. \tag{30}$$

Finally, each local node performs delay-to-frequency-domain DFT operations to obtain the local channel estimate.

$$\widehat{\mathbf{G}}_m = \widehat{\widetilde{\mathbf{G}}}_i \mathbf{F}_{N_C}, m = 1, \ldots, M. \tag{31}$$

Table 1. Computational complexity comparison of the baseline schemes and aggregation based scheme in the non–stationary scenario where $\bar{N}_C$ is preserved columns after hard windowing.

MMSE scheme	$2N_R N_C^2 + N_C^3 N_R^3/4$
Centralized scheme	$2N_R N_C^2 + N_C N_R^3/4$
FD scheme	$2N_R N_C^2 + \frac{1}{4M^2} N_C N_R^3$
Agg. scheme (Central)	$\underbrace{\bar{N}_C N_R^3/4}_{(26)}$
Agg. scheme (Local)	$\underbrace{2N_R N_C^2}_{(22),(31)} + \underbrace{\frac{1}{4M^2}(N_C - \bar{N}_C)N_R^3}_{(28)}$

Table 2. Summary of the channel generating parameters

parameter	value
Number of antennas at BS	256
Polarization of the antennas	$45°/-45°$ dual-polarized
Antenna spacing	$\lambda_f/2$ (half wavelength)
Number of antennas at user	1
Bandwidth	20MHz
Number of subcarriers	1024
Center frequency	3.5GHz
3D-position of BS	$[0, 0, 25]$ (in meter)
3D-position of user	$[100, 20, 1.5]$ (in meter)

Remark 3. The proposed distributed CE algorithm offers a remarkable advantage of achieving a flexible tradeoff between estimation accuracy and inter-BBU communication cost. This can be achieved by selecting different hard thresholding parameters A in (23) and (24), while also incorporating centralized and FD algorithms as special cases. For instance, by setting $A = 0$, all the local information is uploaded to the central node, making it equivalent to the centralized algorithm. On the other hand, for a sufficiently large value of A, no information is uploaded to the central node, and each local node independently estimates its channel with its local received signal. Thus, the algorithm degrades to the FD algorithm. This adaptability of the proposed algorithm adds to its efficacy in various scenarios.

5. Communication Cost and Computational Complexity

In this section, we evaluate the communication cost and computational complexity of the proposed algorithm.

5.1. Communication Cost

The communication cost is measured by the number of real values. Based on the star network architecture, the total number of exchanged real values for the centralized schemeis given by

$$N_c = \frac{4(M-1)N_R N_C}{M},\qquad (32)$$

Similarly, the communication cost of the proposed distributed algorithm is given by

$$C_{dce} = \frac{\sum_{i=1}^{M-1} \bar{N}_{C_i}^{ul}(16N_r + \log_2(N_C)) + \sum_{i=1}^{M-1}(16N_r + \log_2(N_C))}{32N_R N_C(M-1)/M},\qquad (33)$$

5.2. Computational Complexity

The computational complexity is measured by the real multiplications. The computational complexity of the baseline schemes and the proposed algorithm are summarized in Table 1. As it can be seen that the computational complexity of the aggregation based scheme is lower than that of centralized scheme due to the sparsity after hard windowing. For example, with $N_R = 256$, $N_C = 816$, $\bar{N}_C = \frac{1}{10}N_C$ and $M = 16$, the required computational complexity of the aggregation based distributed scheme is only 23.27% of the centralized scheme and the central node will take 39.08% of the total computation. while for $M = 16$, the overall computation decreases to 18.47% of the centralized scheme but more computation (with a percentage of 49.23%) is moved to the central node.

6. Numerical Simulations

In this section, the efficacy of the proposed distributed channel estimation algorithms are numerically evaluated. In the simulations, the channel is generated by the "3GPP-38.901-UMa-NLOS" model in "QuaDRiGa" [20] and the key parameters are summarized in Table 2. The channel covariance is assumed to be known which is approximated by averaging over L (set as 10 in the following simulations) channel realizations. The channel estimation accuracy is evaluated by the normalized mean square error (NMSE), which is given by NMSE $= \|\mathbf{H} - \widehat{\mathbf{H}}\|_F^2/\|\mathbf{H}\|_F^2$.

The NMSE performance versus communication cost of the proposed distributed CE algorithm are simulated in Fig. 4 and Fig. 5. The y-axis represents the NMSE gap between the proposed algorithm and the centralized algorithm. The x-axis is the communication cost computed by $\frac{C_c}{C_{dce}}$ by setting different hard windowing parameter η. Specifically, for SNR = 20dB,and SNR = -20dB, η is set as $[10^8, 25, 15, 10, 7, 5, 3, 1, 0.5, 0.2]$, and $[10^8, 0.9, 0.7, 0.5, 0.3, 0.1, 0.08, 0.06, 0.04, 0.02]$, respectively.
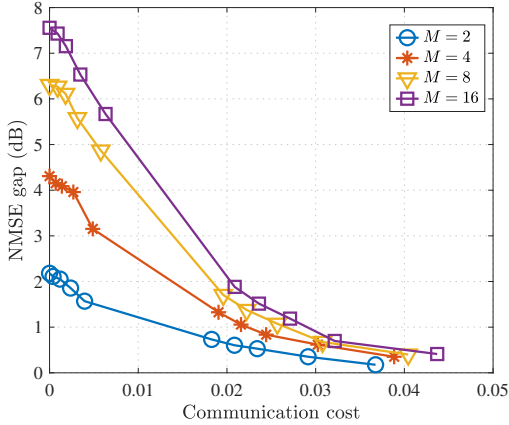
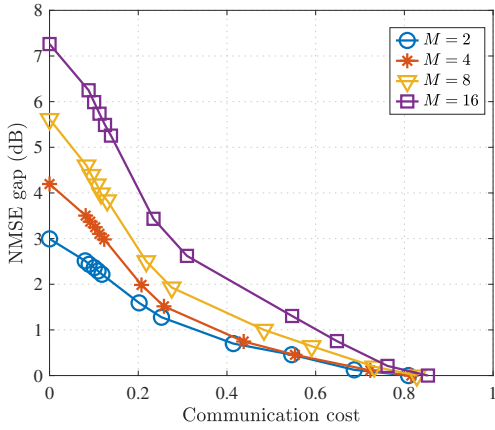**Figure 4.** NMSE performance of the baseline scheme in non–stationary scenario.

Figure 5. NMSE performance of the baseline scheme in non–stationary scenario.

From Fig. 4 and 5, we have the following observations: i) the proposed algorithm can realize flexible tradeoff between NMSE performance and communication cost; ii) the NMSE performance increases as the increase of communication cost as more information from the other nodes is used for estimating local channels; and iii) The required communication cost increases as the increase of SNR. This is due to the fact that in the high SNR cases, more information should be preserved after hard windowing to guarantee the estimation accuracy.

7. Conclusions

In this paper, our focus was on developing distributed channel estimation algorithms for SNS channels, which present a new challenge in wireless communication systems. To address this challenge, we proposed a novel distributed CE algorithm that can be applied even when the antennas are irregularly deployed.

This scheme was designed to take into account the spatial non-stationarity of the channel, which can cause significant degradation in the performance of traditional channel estimation algorithms. To evaluate the effectiveness of our proposed scheme, we compared its performance to that of existing algorithms in both stationary and non-stationary scenarios. Remarkably, even when the antennas are regularly deployed, our proposed scheme demonstrated superior performance compared to existing algorithms in the stationary scenario. These findings suggest that our proposed algorithm can significantly improve the accuracy of channel estimation in SNS channels, even in scenarios where the antennas are irregularly deployed.

8. Acknowledgements

References

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

[2] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE signal processing magazine*, vol. 30, no. 1, pp. 40–60, 2012.

[3] D. Cai, J. Wen, P. Fan, Y. Xu, and L. Yu, "Active user and data detection for uplink grant-free NOMA systems," *China Commun.*, vol. 17, no. 11, pp. 12–28, 2020.

[4] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming," *IEEE Signal Process Mag.*, vol. 27, no. 3, pp. 62–75, May 2010.

[5] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.

[6] Y. Xu, C. Shen, Z. Ding, X. Sun, S. Yan, G. Zhu, and Z. Zhong, "Joint beamforming and power-splitting control in downlink cooperative SWIPT NOMA systems," *IEEE Trans. Signal Process.*, vol. 65, no. 18, pp. 4874–4886, Sep. 2017.

[7] Y. Takano, H.-J. Su, M. Juntti, and T. Matsumoto, "A conditional $\ell 1$ regularized MMSE channel estimation technique for IBI channels," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6720–6734, Oct. 2018.

[8] E. Björnson and B. Ottersten, "A framework for training-based estimation in arbitrarily correlated rician MIMO channels with rician disturbance," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1807–1820, Mar. 2010.

[9] D. Cai, P. Fan, Q. Zou, Y. Xu, Z. Ding, and Z. Liu, "Active device detection and performance analysis of massive non-orthogonal transmissions in cellular internet of things," *Sci. China Inf. Sci.*, vol. 65, no. 8, pp. 1–18, 2022.

[10] K. Li, R. R. Sharan, Y. Chen, T. Goldstein, J. R. Cavallaro, and C. Studer, "Decentralized baseband processing for massive MU-MIMO systems," *IEEE J. Emerging Sel. Top. Circuits Syst.*, vol. 7, no. 4, pp. 491–507, Dec. 2017.

[11] C. Jeon, K. Li, J. R. Cavallaro, and C. Studer, "Decentralized equalization with feedforward architectures for massive MU-MIMO," *IEEE Trans. Signal Process.*, vol. 67, no. 17, pp. 4418–4432, Sept. 2019.

[12] ——, "On the achievable rates of decentralized equalization in massive MU-MIMO systems," in *IEEE Int. Symp. Inf. Theory*, Jun. 2017, pp. 1102–1106.

[13] Z. Zhang, H. Li, Y. Dong, X. Wang, and X. Dai, "Decentralized signal detection via expectation propagation algorithm for uplink massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 11 233–11 240, Oct. 2020.

[14] A. Amiri, C. N. ManchÃşn, and E. de Carvalho, "Uncoordinated and decentralized processing in extra-large MIMO arrays," *IEEE Wireless Commun. Lett.*, vol. 11, no. 1, pp. 81–85, Jan. 2022.

[15] V. Croisfelt, T. Abrão, A. Amiri, E. de Carvalho, and P. Popovski, "Decentralized design of fast iterative receivers for massive MIMO with spatial non-stationarities," in *55rd Asilomar Conf. Signals, Syst., Comput.*, 2021, pp. 1242–1249.

[16] A. Kulkarni, M. A. Ouameur, and D. Massicotte, "Hardware topologies for decentralized large-scale MIMO detection using newton method," *IEEE Trans. Circuits Syst. I Regul. Pap.*, vol. 68, no. 9, pp. 3732–3745, Sept. 2021.

[17] À. O. Martínez, E. De Carvalho, and J. Ø. Nielsen, "Towards very large aperture massive MIMO: A measurement based study," in *2014 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2014, pp. 281–286.

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

[19] S. Hou, Y. Wang, T. Zeng, and S. Wu, "Sparse channel estimation for spatial non-stationary massive MIMO channels," *IEEE Commun. Lett.*, vol. 24, no. 3, pp. 681–684, 2020.

[20] J. Stephan, R. Leszek, B. Kai, and T. Lars, "QuaDRiGa: A 3-D multi-cell channel model with time evolution for enabling virtual field trials," *IEEE Trans. Antennas Propagat.*, vol. 62, no. 6, pp. 3242–3256, 2014.