This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2023.3343740

Non-Stationary Channel Estimation for Extremely Large-Scale MIMO

Yuhao Chen, *Graduate Student Member, IEEE* and Linglong Dai, *Fellow, IEEE*

***Abstract*—Extremely large-scale multiple-input multiple-output (XL-MIMO) is considered as a key technology for future 6G communications. To realize effective precoding, channel estimation schemes are essential to acquire precise channel state information (CSI), while most existing schemes work relying on the spatial stationary assumption. However, in XL-MIMO systems, the spatial non-stationary effect naturally exists. Such an effect can hardly be recognized by most existing channel estimation schemes, leading to a severe accuracy loss of channel estimation. In order to deal with this problem, we study the spatial non-stationary channel estimation in XL-MIMO systems in this paper. Specifically, the spatial non-stationary channel in an XL-MIMO system is converted to a series of spatial stationary channels by a proposed group time block code (GTBC) based signal extraction scheme. The key idea is to artificially create the time-domain relevance of non-stationary effect, which brings XL-MIMO the ability to recognize such effect in the space domain. Based on the extracted signals, an on-grid GTBC-based polar-domain simultaneous orthogonal matching pursuit (GP-SOMP) algorithm and an off-grid GTBC-based polar-domain simultaneous iterative gridless weighted (GP-SIGW) algorithm are proposed to effectively estimate the non-stationary channel. Then, analyses of the time complexity and performances of the above two algorithms are carried out and the Cramér-Rao lower bound is derived. Finally, numerical results reveal that the proposed algorithms can recognize the spatial non-stationary effect and realize a much more accurate channel estimation than existing schemes.**

***Index Terms*—Spatial non-stationary, XL-MIMO, channel estimation, hybrid precoding.**

I. Introduction

To meet the ever-growing demand on communication capacity in 6G networks, the dimension of wireless signals in both spatial and frequency domains are increasing rapidly [1]. In the spatial domain, massive multiple-input-multiple-output (mMIMO) is evolving to extremely large-scale MIMO (XL-MIMO) from 5G to 6G. With a much larger antenna number, the XL-MIMO embraces 10-fold increase in spectral efficiency [2]. Meanwhile, the operating frequency also climbs from sub-6G to millimeter-wave (mmWave) and terahertz (THz) bands in order to enjoy richer spectral resources [3]. Because of the small wavelength, the size of high-frequency antennas is small accordingly, and several high-gain arrays at high frequency have been designed [4], [5]. As a consequence, the integration of XL-MIMO and high-frequency communications is natural and has been acknowledged as a key technology for future 6G communications.

In high-frequency communications, hybrid precoding architecture is utilized to reduce the high power consumption of the radio-frequency (RF) chain [6], which is the same as the current 5G mmWave massive MIMO [7]. In order to realize efficient hybrid precoding, accurate channel state information (CSI) is necessary. However, in hybrid precoding XL-MIMO, the size of channel matrix is much larger than the number of RF chains, so the signals at each antenna cannot be received by the BS respectively [8], which will lead to an unacceptable pilot overhead in XL-MIMO communicaton systems.

A. Prior Works

To realize a low-overhead channel estimation, several algorithms based on compressive sensing (CS) have been studied by exploiting the channel sparsity of high-frequency communication systems [9]–[23]. Specifically, the angular-domain sparsity of far-field channels was exploited by OMP based algorithms in massive MIMO systems. By conducting Fourier transform, the channel could be transformed into the angular-domain [9]–[11]. These algorithms hold a common on-grid assumption, which states that the angles of departures (AoDs) and the angles of arrivals (AoAs) lie in discrete sampling points in the angular domain. This assumption limited the channel estimation accuracy since most of the AoDs/AoAs might not lie exactly in the sampling points in real systems. To solve these problems, several off-grid algorithms were proposed in [12]–[15] to fine-tune the estimated AoAs and AoDs through optimization methods. Another category of channel estimation technologies focus on the paradigm shift of electromagnetic field structure in XL-MIMO systems. With the increase of antenna number in XL-MIMO systems, many receivers may locate in the near-field region of the electromagnetic field [24]. For example, in a communication system with the carrier frequency being 28 GHz, the Rayleigh distance is up to 187 meters when the array aperture is 1m [25]. In this case, spherical wavefront, which accounted for both the distance and the angle between the scatterer and BS, should be considered in the channel model. As indicated in [16], the spherical wavefront of the near-field channel destroys the angular-domain sparsity widely exploited in [12]–[15]. In order to catch the sparsity in near-field channels, [17] represented the near-field channel sparsely in the polar domain,

This work was supported by the National Natural Science Foundation of China (Grant No. 62031019). The conference version of this paper was presented in part at IEEE VTC'23, Florence, Italy, June 2023. The associate editor coordinating the review of this article and approving it for publication was K. Adachi. *(Corresponding author: Linglong Dai.)*

All authors are with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China, and also with the Beijing National Research Center for Information Science and Technology (BNRist), Beijing 100084, China. (e-mails: chen-yh21@mails.tsinghua.edu.cn; daill@tsinghua.edu.cn).

Authorized licensed use limited to: Tsinghua University. Downloaded on March 16,2024 at 02:20:29 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2023.3343740

which utilized the distance and angular information at the same time. Based on the polar-domain representation, [17] studied the on-grid algorithm called P-SOMP and the off-grid algorithm called P-SIGW to recover the near-field channel accurately, which are also adaptive to far-field channels. To reduce the complexity of the polar-domain dictionary, authors in [18], [19] exploited the spatial information in the near field, based on which the corresponding channel estimation schemes were studied to acquire a higher channel estimation accuracy.Authors in [20], [21] further took the near-field beam split effect in wideband systems into consideration and studied the accurate wideband channel estimation schemes in wideband near-field scenarios.In addition, the hybrid field, where both far-field and near-field exist simultaneously, was considered in [22], [23], and the correponding OMP based channel estimation schemes were studied.

The above channel estimation algorithms all assumed that the channel is spatial stationary, i.e., all BS's antennas are visible to the same scatterers and users. However, as the scale of antenna array increases in XL-MIMO systems, different regions of the array are possibly visible to totally different scatterers or users, which could be described by the visibility regions (VR) [26]. As a result, the channel estimation accuracy will dramatically decrease since the spatial non-stationarity of the channel cannot be captured by traditional algorithms. To handle the spatial non-stationarity in XL-MIMO systems, several schemes have been studied [27]–[33]. Specifically, [27], [28] studied a sub-array-wise channel estimation algorithm, where the antenna array was divided into several sub-arrays first, with each sub-array being spatial stationary. Then, the OMP algorithm was applied to each sub-array and recover the corresponding channel. In order to recognize the VR corresponding to each scatterer more efficiently, [29] studied a YOLO based algorithm to estimate the sub-arrays belonging to each scatterer. In addition, authors in [30], [31] estimated the spatial non-stationary channel effectively based on the expectation maximization and expectation propagation, respectively. Furthermore, [32], [33] captured the specific structure of XL-MIMO such as the common sparsity in the antenna domain and studied the Bayesian based channel estimation algorithms to estimate the spatial non-stationary channel. The above schemes have one thing in common. Though the channel is spatial non-stationary for the whole antenna array, the channel corresponding to a part of the antenna array (i.e., a sub-array) is treated as spatial stationary. This is also verified by the measurement in [34].

A common issue of the above method tailored for spatial non-stationarity is that they are all based on a fully digital precoding structure, where the received signal has the same dimension as the channel. This setting makes it easy to analyze the received signals of different sub-arrays independently. Unfortunately, in practical XL-MIMO systems, the hybrid precoding architecture is widely employed. This fact means that for a certain RF chain, the received signals of all antennas are mixed together, making it difficult to decouple the received signal of a certain sub-array. Therefore, existing schemes cannot be applied in hybrid precoding XL-MIMO systems. As far as we know, the channel estimation considering spatial non-stationarity in hybrid precoding XL-MIMO systems has not been studied in the literature, which is a big challenge in future communications.

B. Our Contributions

To fill in this gap, in this paper, we study the spatial non-stationary channel estimation for both far-field and near-field scenarios in XL-MIMO systems with hybrid precoding structure, which is realized by a group time block code (GTBC) based signal extraction scheme. Specifically, the contribution of this paper can be summarized as follows.

- Firstly, we show that the problem of estimating the channel of the entire antenna array can be converted into several sub-problems of estimating the sub-channel with respect to each sub-array. Though the channel is spatial non-stationary for the whole antenna array, the channel with respect to each sub-array can still be considered as spatial stationary. Thus, for each sub-array, we can apply several CS based channel estimation algorithms to recover the corresponding sub-channel with a low pilot overhead. Based on this, the whole channel can be recovered by combining all the estimated sub-channels.
- Considering the hybrid precoding structure, the received signals are the mixture of all BS's antennas, which makes it hard to extract the received signals corresponding to each sub-array for further process. To overcome this difficulty, we propose a GTBC based signal extraction scheme, which artificially create the relevance in the time domain to acquire the ability to recognize the spatial non-stationarity in the space domain. The proposed scheme consists of the encoding stage and the decoding stage. At the encoding stage, the antennas in a spatial stationary sub-array are packed together as a group. In each time slot, the combining matrix of the antennas in a group is changed as a whole according to the designed GTBC to create the time-domain relevance. This is different from traditional channel estimation algorithms, where the combining matrix is generated in a totally random way in each time slot. Through the above encoding process, the antennas in each group exhibit consistent behaviors, which is the foundation of extracting the signals correponding to each group. At the decoding stage, the original received signal is combined based on the designed GTBC, and the received signal corresponding to each sub-array is extracted for further process.
- Based on the proposed GTBC based signal extraction scheme and the channel estimation algorithms in [17], an on-grid GTBC-based polar-domain simultaneous orthogonal matching pursuit (GP-SOMP) algorithm and an off-grid GTBC-based polar-domain simultaneous iterative gridless weighted (GP-SIGW) algorithm are proposed to estimate the spatial non-stationary channel efficiently, which are appliable for both the near-field scenario and the far-field scenario. In addition, analyses of the performance and the computational complexity of the proposed two algorithms are conducted. Furthermore, the Cramér-Rao lower bound (CRLB) is derived to assess the effectiveness of the proposed algorithms. Finally, numerical

Authorized licensed use limited to: Tsinghua University. Downloaded on March 16,2024 at 02:20:29 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2023.3343740

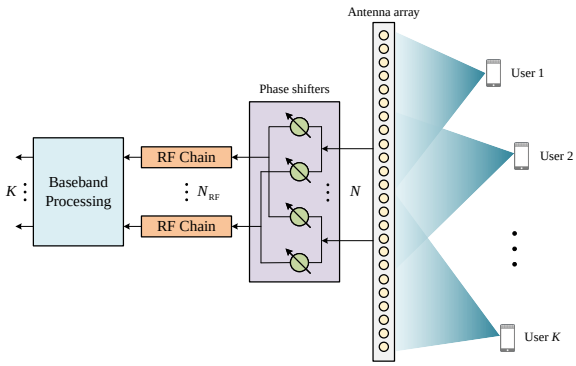


Fig. 1. XL-MIMO system with hybrid precoding.

experiments are conducted to reveal the effectiveness of the proposed algorithms.

C. Organization and Notation

Organization: The remainder of this paper is organized as follows. In Section II, the system model is introduced, where the spherical wave propagation effects and the spatial non-stationary effects in XL-MIMO systems are elaborated. In Section III, the GTBC based signal extraction scheme is proposed. In Section IV, we propose an on-grid algorithm called GP-SOMP and an off-grid algorithm called GP-SIGW, and the performance and computational complexity analysis is elaborated. In Section V, we carry out the numerical experiments, and finally conclusions are drawn in Section VI.

Notation: Vectors and matrices are denoted by lower-case and upper-case boldface letters; $\mathbf{X}[i,j]$ denotes the (i,j)-th element of the matrix $\mathbf{X}$; $\mathbf{X}[i,:]$ and $\mathbf{X}[:,j]$ denote the i-th row and the j-th column of the matrix $\mathbf{X}$; $(\cdot)^T$ and $(\cdot)^H$ are the transpose and conjugate transpose, respectively; $|\cdot|$ is the absolute operator; $\|\cdot\|_F$ is the Frobenius norm operator; $\odot$ is the Hadamard product operator; $\mathrm{Tr}(\cdot)$ is the trace operator; $\mathcal{CN}(\boldsymbol{\mu},\boldsymbol{\Sigma})$ and $\mathcal{U}(a,b)$ are the Gaussian distribution with mean μ and covariance Σ, and the uniform distribution between a and b, respectively.

II. System Model

In this paper, we consider an uplink time division duplexing (TDD) based XL-MIMO communication system, which is illustrated in Fig. 1. A fully-connected hybrid precoding architecture with N_{RF} RF chains and an N-antenna uniform linear array (ULA) is equipped at the base station (BS)[1]. The antenna spacing of the ULA is $d=\frac{c}{2f_c}$, where f_c is the central carrier frequency. M subcarriers serve K single antenna users simultaneously. Since different users utilize orthogonal pilot sequences during the channel estimation stage [35], an arbitrary user is considered in the following sections. Specifically,

[1]It is worth noting that though we consider the fully-connected hybrid precoding architecture here, the proposed schemes are also appliable to partially-connected architectures. In order to apply proposed schemes in such systems, we only need to consider each RF-chain and the antennas connected to it as a smaller fully-connected architecture.

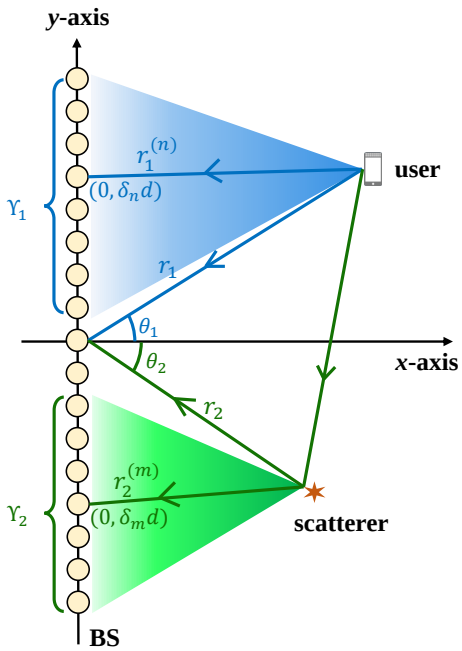


Fig. 2. The near-field non-stationary channel model with two paths.

the transmit pilot in time slot p at the m-th subcarrier are denoted by $s_{m,p}$. Then, the received pilot $\mathbf{y}_{m,p}\in\mathbb{C}^{N_{\mathrm{RF}}\times 1}$ can be represented as

$$\mathbf{y}_{m,p}=\mathbf{A}_p\mathbf{h}_m s_{m,p}+\mathbf{A}_p\mathbf{n}_{m,p}, \tag{1}$$

where $\mathbf{A}_p\in\mathbb{C}^{N_{\mathrm{RF}}\times N}$ is the analog combining matrix satisfying the constant modulus constraint $|\mathbf{A}_p(i,j)|=\frac{1}{\sqrt{N}}$, $\mathbf{n}_{m,p}\in\mathbb{C}^{N\times 1}$ is the complex Gaussian noise following the distribution $\mathcal{CN}\left(\mathbf{0},\sigma^2\mathbf{I}_N\right)$ with σ^2 being the noise power. We define the pilot overhead as P. By assuming that $s_{m,p}=1$ for all $p=1,2,...,P$, the overall received pilots at the m-th subcarrier $\mathbf{y}_m=\left[\mathbf{y}_{m,1}^T,...,\mathbf{y}_{m,P}^T\right]^T\in\mathbb{C}^{PN_{\mathrm{RF}}\times 1}$ can be formulated as

$$\mathbf{y}_m=\mathbf{A}\mathbf{h}_m+\mathbf{n}_m, \tag{2}$$

where $\mathbf{A}=\left[\mathbf{A}_1^T,...,\mathbf{A}_P^T\right]^T\in\mathbb{C}^{PN_{\mathrm{RF}}\times N}$ is the overall combining matrix, with the elements randomly selected from $\mathbb{H}=\frac{1}{\sqrt{N}}\{-1,1\}$, $\mathbf{n}_m=\left[\mathbf{n}_{m,1}^T\mathbf{A}_1^T,...,\mathbf{n}_{m,P}^T\mathbf{A}_P^T\right]^T\in\mathbb{C}^{PN_{\mathrm{RF}}\times 1}$ denotes the effective noise.

In XL-MIMO systems, the channel matrix $\mathbf{h}_m$ differs from traditional massive MIMO in two aspects: *spherical wave propagation* and *spatial non-stationarity*. The two aspects are elaborated in the following two subsections.

A. Spherical Wave Propagation

Rayleigh distance could divide the radiation field of electromagnetic into two regions [24], which is defined as $Z=\frac{2D^2}{\lambda_c}$ in terms of the array aperture D and the wavelength of the central subcarrier λ_c. In most literature, planar wave assumption is applied. In this case, the steering vector accounts for only the angle, and can be represented as a Fourier vector. Therefore, the channel exhibits sparsity in its angular domain by applying a spatial discrete Fourier transform [36]. This sparsity enables

Authorized licensed use limited to: Tsinghua University. Downloaded on March 16,2024 at 02:20:29 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2023.3343740

the application of compressive sensing (CS) based channel estimation schemes that can recover the channel matrix with a reduced pilot overhead [9]–[11], [36]–[38]. However, in XL-MIMO systems, the Rayleigh distance can be up to hundreds of meters. Therefore, the users are inevitably located in the range of near-field regions. To describe the channel more accurately, spherical wave propagation should be considered in XL-MIMO systems.

Specifically, as illustrated in Fig. 2, considering the spherical wavefront, the near-field channel model can be formulated as [39]

$$\tilde{\mathbf{h}}_m = \sqrt{\frac{N}{L}} \sum_{\ell=1}^{L} g_\ell e^{-jk_m r_\ell} \mathbf{b}\left(\theta_\ell, r_\ell\right), \tag{3}$$

with $k_m = \frac{2\pi f_m}{c}$ being the wave number, L being the path number, and g_ℓ being the complex path gain, r_ℓ being the distance, θ_ℓ being the sine of the physical angle of the ℓ-th path, respectively. It is worth noting that $\mathbf{b}\left(\cdot\right)$ is dependent on both the angle θ_ℓ and the distance r_ℓ as

$$\mathbf{b}\left(\theta_\ell, r_\ell\right) = \frac{1}{\sqrt{N}} \left[e^{-jk_c\left(r_\ell^{(1)} - r_\ell\right)}, ..., e^{-jk_c\left(r_\ell^{(N)} - r_\ell\right)} \right]^T, \tag{4}$$

with $k_c = \frac{2\pi}{\lambda_c}$ being the wave number of the central subcarrier and $r_\ell^{(n)}$ being the distance between the l-th scatterer and the n-th BS antenna, which can be formulated as

$$r_\ell^{(n)} = \sqrt{r_\ell^2 - 2r_\ell \delta_n d\theta_\ell + \delta_n^2 d^2}, \tag{5}$$

with $\delta_n = \frac{2n-N-1}{2}, n = 1, 2, ..., N$. The non-linearity of the phase of each element in (5) indicates that there is no sparsity in the angular domain. Fortunately, researchers in [17] considers both the distance and the angle in the near-field channel simultaneously to obtain a polar-domain codebook. This codebook reveals the sparsity of near-field channel, which enables the application of CS based schemes to recover the channel matrix with a low pilot overhead.

B. Spatial Non-stationarity

The large array aperture not only causes the change in wavefront shape, but also brings the spatial non-stationary effect, as illustrated in Fig. 2. Spatial non-stationary effect means that different antennas are visible to different scatterers or users, which is described by the visibility regions (VRs). Taking the VR of each path into account, the channel model should be modified based on (3) as

$$\mathbf{h}_m = \sqrt{\frac{N}{L}} \sum_{\ell=1}^{L} g_\ell e^{-jk_m r_\ell} \mathbf{b}\left(\theta_\ell, r_\ell\right) \odot \mathbf{v}\left(\Upsilon_\ell\right), \tag{6}$$

where $\mathbf{v}\left(\cdot\right)$ denotes the mask which selects the antennas visible to the ℓ-th scatterer, and Υ_ℓ is the indices of antennas that can see[2] the ℓ-th scatterer. Specifically, the n-th element of $\mathbf{v}\left(\Upsilon_\ell\right)$ is defined as

$$\left[\mathbf{v}\left(\Upsilon_\ell\right)\right]_n = \begin{cases} 1, & n \in \Upsilon_\ell \\ 0, & n \notin \Upsilon_\ell \end{cases}. \tag{7}$$

As shown by the measurement in [34], though the entire channel is spatial non-stationary, the sub-channels with respect to a part of the array (i.e., a sub-array) can be treated as spatial stationary. Considering an XL-MIMO divided into N_s smaller sub-arrays, without loss of generality, we assume that $\frac{N}{N_s}$ is an integer. The set $\tilde{\Upsilon}_\ell$ is introduced to index the sub-arrays that are visible to the ℓ-th scatter, which is represented as

$$\tilde{\Upsilon}_\ell = \left\{n_{s,1}, ..., n_{s,L}\right\}, \tag{8}$$

where $1 \leq n_{s,i} \leq N_s$ denotes the index of sub-array. Then, (7) can be further represented as

$$\left[\mathbf{v}\left(\Upsilon_\ell\right)\right]_n = \begin{cases} 1, & \left\lceil \frac{nN_s}{N} \right\rceil \in \tilde{\Upsilon}_\ell \\ 0, & \text{else} \end{cases}. \tag{9}$$

The mask vector $\mathbf{v}\left(\cdot\right)$ brings new challenge to the channel estimation since traditional CS based schemes, with the basis vector being the Fourier vector or Polar-domain vector, all imply the assumption of spatial stationary. The mismatch of the basis vector causes a non-negligible performance loss in XL-MIMO systems.

III. Proposed GTBC based Signal Extraction Scheme

To realize accurate channel estimation in spatial non-stationary XL-MIMO systems, we can divide the entire antenna array into several sub-arrays. We divide the antennas in the same sub-array into a group. Since the size of a sub-array is much smaller than the entire antenna array, the sub-channel can be treated as spatial stationary. In this section, we will propose a group time block code (GTBC) based signal extraction scheme to extract the received signal with respect to each sub-array.

A. Encoding Stage

In this subsection, we introduce the encoding stage based on the GTBC, where the combining matrix of the antennas in a group is changed as a whole in different time slots based on the GTBC. Existing schemes considering spatial non-stationarity all assumed that the received signals have the same dimension as the channel. In this case, the extraction of the received signal with respect to each sub-array is trivial. However, in XL-MIMO hybrid precoding systems, the number of RF chains is far smaller than that of BS antennas, and the received signals of each RF chain is merely the mixture of the received signal of all BS antennas. This makes it challenging to decouple the the received signal with respect to each sub-array. The main reason of the above challenge is that in traditional channel estimation procedure in hybrid precoding systems, for a certain RF chain, the analog combining matrices in different time slots are all randomly generated. This makes the received signals of different time slots independent from each other. The independence in the time domain sacrifices the ability

[2]Here, the word "see" does not necessarily mean the visual blockage. To be more rigorous, the word "see" means that the signals from the ℓ-th scatterer can reach the antenna elements.

Authorized licensed use limited to: Tsinghua University. Downloaded on March 16,2024 at 02:20:29 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2023.3343740

of decouple signals of different antenna groups in the space domain. In order to exploit the received signals of different time slots to acquire the received signal with respect to each sub-array, we need to artificially create the relevance in the time domain based on the GTBC.

We consider the received signal of the i-th RF chain, $1 \leq i \leq N_{\mathrm{RF}}$. In the p-th time slot, the received signal of the m-th subcarrier can be formulated as

$$y_{m,p,i} = \mathbf{A}_p\left[i,:\right]\mathbf{h}_m s_{m,p} + \mathbf{A}_p\left[i,:\right]\mathbf{n}_{m,p}. \tag{10}$$

To further simplify the expression, the indices of the RF chains are omitted, and (10) can be rewritten as

$$y_{m,p} = \mathbf{A}_p\mathbf{h}_m s_{m,p} + \mathbf{A}_p\mathbf{n}_{m,p}. \tag{11}$$

We start from a simple case with the XL-MIMO divided into two sub-arrays. To describe the sub-channels corresponding to the two sub-arrays more clearly, $\mathbf{A}$ and $\mathbf{h}_m$ are divided into two parts and (11) is further represented as

$$\begin{aligned} y_{m,p} &= \begin{bmatrix}\mathbf{A}_{p,1} & \mathbf{A}_{p,2}\end{bmatrix}\begin{bmatrix}\mathbf{h}_{m,1}\\ \mathbf{h}_{m,2}\end{bmatrix} s_{m,p} + \mathbf{A}_p\mathbf{n}_{m,p} \\ &= \mathbf{A}_{p,1}\mathbf{h}_{m,1}s_{m,p} + \mathbf{A}_{p,2}\mathbf{h}_{m,2}s_{m,p} + \mathbf{A}_p\mathbf{n}_{m,p}, \end{aligned} \tag{12}$$

where $\mathbf{A}_{p,j}, \mathbf{h}_{m,j}$ denote the combing matrix of the p-th time slot and the sub-channel corresponding to the j-th sub-array, respectively. The total received signal is composed of three parts: the received signal from the first sub-array, the received signal from the second sub-array and the effective noise. Though the received signal from the two sub-arrays cannot be extracted in only one time slot, the combining matrix corresponding to each sub-array can be easily decoupled. To divide the above two received signals, two time slots, i.e., the p-th and $(p+1)$-th time slots are needed. Traditional channel estimation schemes generate the combining matrix in a total random way, making $\mathbf{A}_{p,j}$ and $\mathbf{A}_{p+1,j}$ independent of each other, so the two parts of received signal cannot be divided.

Instead of a totally random design, we design the combining matrix in every two adjacent time slots based on the GTBC, where the combining matrix of the antennas in one group is changed as a whole. Specifically, $\mathbf{A}_{p+1,j}$ is generated based on $\mathbf{A}_{p,j}$. In the above case, let $\mathbf{A}_{p+1,1} = \mathbf{A}_{p,1}, \mathbf{A}_{p+1,2} = -\mathbf{A}_{p,2}, s_{m,p+1} = s_{m,p}$, and $y_{m,p+1}$ is represented as

$$\begin{aligned} y_{m,p+1} &= \begin{bmatrix}\mathbf{A}_{p+1,1} & \mathbf{A}_{p+1,2}\end{bmatrix}\begin{bmatrix}\mathbf{h}_{m,1}\\ \mathbf{h}_{m,2}\end{bmatrix} s_{m,p} + \mathbf{A}_{p+1}\mathbf{n}_{m,p+1} \\ &= \begin{bmatrix}\mathbf{A}_{p,1} & -\mathbf{A}_{p,2}\end{bmatrix}\begin{bmatrix}\mathbf{h}_{m,1}\\ \mathbf{h}_{m,2}\end{bmatrix} s_{m,p} + \mathbf{A}_{p+1}\mathbf{n}_{m,p+1} \\ &= \mathbf{A}_{p,1}\mathbf{h}_{m,1}s_{m,p} - \mathbf{A}_{p,2}\mathbf{h}_{m,2}s_{m,p} + \mathbf{A}_{p+1}\mathbf{n}_{m,p+1}. \end{aligned} \tag{13}$$

For a more general scenario, where the XL-MIMO is divided into N_s sub-arrays, how to design the GTBC to generate the combining matrices in every N_s adjacent time slots is the main challenge. In the above case, the operation in a certain group can be abstracted as

$$\begin{bmatrix}\mathbf{A}_p\\ \mathbf{A}_{p+1}\end{bmatrix} = \begin{bmatrix}\mathbf{A}_{p,1} & \mathbf{A}_{p,2}\\ \mathbf{A}_{p,1} & \mathbf{A}_{p,2}\end{bmatrix} \odot \underbrace{\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}}_{\mathbf{P}} = \begin{bmatrix}\mathbf{A}_{p,1} & \mathbf{A}_{p,2}\\ \mathbf{A}_{p,1} & -\mathbf{A}_{p,2}\end{bmatrix}, \tag{14}$$

where the matrix $\mathbf{P}$ is called the GTBC matrix in this paper, based on which the combining matrices in every N_s adjacent time slots can be generated by a random basic combining matrix. Then, the design of the GTBC in every N_s adjacent time slots is converted to the design of GTBC matrix $\mathbf{P}$. While designing $\mathbf{P}$, the following factors are considered. 1) $\mathbf{P} \in \mathbb{C}^{N_s \times N_s}$ to provide enough rank to extract the received signals corresponding to all N_s sub-arrays. 2) Each column of $\mathbf{P}$ is orthogonal to each other to eliminate the noise. 3) The phase of elements in $\mathbf{P}$ and $\mathbf{P}^{-1}$ need to be as coarse as possible considering the non-ideal phase shift of the phase shifters at the BS.

Fortunately, Hadamard matrix provides a proper solution to all the above problems. We denote n-order Hadamard matrix as $\mathbf{\Lambda}_n$. Hadamard matrix is composed of 1 and -1, and $\frac{1}{\sqrt{n}}\mathbf{\Lambda}_n$ is an orthogonal matrix, which satisfies the above three factors simultaneously. In the above case where the XL-MIMO is divided into two sub-arrays, $\mathbf{P} = \mathbf{\Lambda}_2 = \begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$. The Hadamard matrix with order 2^k can be recursively obtained as

$$\mathbf{\Lambda}_{2^k} = \begin{bmatrix}\mathbf{\Lambda}_{2^{k-1}} & \mathbf{\Lambda}_{2^{k-1}}\\ \mathbf{\Lambda}_{2^{k-1}} & -\mathbf{\Lambda}_{2^{k-1}}\end{bmatrix}. \tag{15}$$

Considering an XL-MIMO system, where the BS is divided into N_s sub-arrays, the adjacent N_s time slots share a common basic combining matrix. We assume that $N_s = 2^k$. To match the dimension of the BS antenna number, the $\mathbf{\Lambda}_{N_s}$ needs to be processed as (16), where $\mathbf{1}_{1\times N/N_s}$ denotes the all-one matrix with dimension $1 \times N/N_s$.

Taking consecutive N_s time slots from the p-th time slot as an example, the combining matrix can be generated based on $\mathbf{A}_p$ and the GTBC matrix by

$$\begin{bmatrix}\mathbf{A}_p\\ \mathbf{A}_{p+1}\\ \vdots\\ \mathbf{A}_{p+N_s-1}\end{bmatrix} = \begin{bmatrix}\mathbf{A}_{p,1} & \mathbf{A}_{p,2} & \cdots & \mathbf{A}_{p,N_s}\\ \mathbf{A}_{p,1} & \mathbf{A}_{p,2} & \cdots & \mathbf{A}_{p,N_s}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{A}_{p,1} & \mathbf{A}_{p,2} & \cdots & \mathbf{A}_{p,N_s}\end{bmatrix} \odot \tilde{\mathbf{\Lambda}}_{N_s}, \tag{17}$$

where $\mathbf{A}_{p,j}$ denotes the $((j-1)\,N/N_s+1)$-th column to the (jN/N_s)-th column of $\mathbf{A}_p$. With the help of Hadamard matrix, the combining matrix can be generated based on the GTBC, which is the foundation to extract the received signals

$$\tilde{\mathbf{\Lambda}}_{N_s} = \begin{bmatrix}\mathbf{1}_{1\times N/N_s}\times\mathbf{\Lambda}_{N_s}[1,1] & \mathbf{1}_{1\times N/N_s}\times\mathbf{\Lambda}_{N_s}[1,2] & \cdots & \mathbf{1}_{1\times N/N_s}\times\mathbf{\Lambda}_{N_s}[1,N_s]\\ \mathbf{1}_{1\times N/N_s}\times\mathbf{\Lambda}_{N_s}[2,1] & \mathbf{1}_{1\times N/N_s}\times\mathbf{\Lambda}_{N_s}[2,2] & \cdots & \mathbf{1}_{1\times N/N_s}\times\mathbf{\Lambda}_{N_s}[2,N_s]\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{1}_{1\times N/N_s}\times\mathbf{\Lambda}_{N_s}[N_s,1] & \mathbf{1}_{1\times N/N_s}\times\mathbf{\Lambda}_{N_s}[N_s,2] & \cdots & \mathbf{1}_{1\times N/N_s}\times\mathbf{\Lambda}_{N_s}[N_s,N_s]\end{bmatrix} \tag{16}$$

Authorized licensed use limited to: Tsinghua University. Downloaded on March 16,2024 at 02:20:29 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2023.3343740

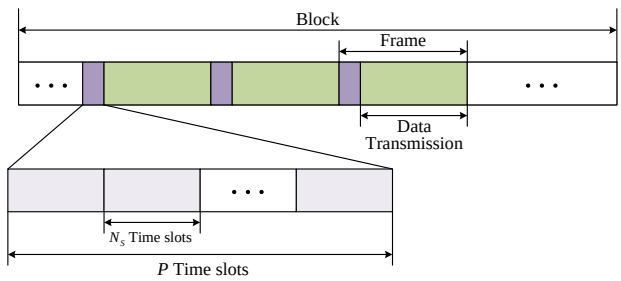


Fig. 3. The frame structure for proposed signal extraction scheme

corresponding to each sub-array.

To employ the above combining matrix design in the communication systems, the channel estimation procedure should be divided according to the number of sub-arrays. As illustrated in Fig. 3, P pilot transmission time slots are divided into $\frac{P}{N_s}$ sub-frames, each having N_s time slots. $\frac{P}{N_s}$ basic combining matrices are firstly generated randomly and independently, which are denoted as $\mathbf{A}_{(i-1)N_s+1}, i = 1, 2, ..., P/N_s$. Then, for each sub-frame, the P combining matrices are generated by (17) for pilot transmission.

B. Decoding Stage

After receiving the combined received signal, the received signal corresponding to each sub-array is extracted at the decoding stage based on the GTBC. Starting from the simple case where the XL-MIMO is divided into two sub-arrays, by combining (12) and (13), we get

$$\begin{cases} \tilde{y}_{m,p,1} = \dfrac{y_{m,p} + y_{m,p+1}}{2} = \mathbf{A}_{p,1}\mathbf{h}_{m,1}s_{m,p} + \tilde{\mathbf{n}}_{m,p,1} \\ \tilde{y}_{m,p,2} = \dfrac{y_{m,p} - y_{m,p+1}}{2} = \mathbf{A}_{p,2}\mathbf{h}_{m,2}s_{m,p} + \tilde{\mathbf{n}}_{m,p,2} \end{cases}, \tag{18}$$

where $\tilde{\mathbf{n}}_{m,p,1} = \frac{\mathbf{A}_p\mathbf{n}_{m,p}+\mathbf{A}_{p+1}\mathbf{n}_{m,p+1}}{2}$, and $\tilde{\mathbf{n}}_{m,p,2} = \frac{\mathbf{A}_p\mathbf{n}_{m,p}-\mathbf{A}_{p+1}\mathbf{n}_{m,p+1}}{2}$. The received signals corresponding to each sub-array are properly extracted by a simple linear combination at the BS.

The above process can be abstracted as

$$\begin{bmatrix} \tilde{y}_{m,p,1} \\ \tilde{y}_{m,p,2} \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}}_{\mathbf{P}^{-1}} \begin{bmatrix} y_{m,p} \\ y_{m,p+1} \end{bmatrix}. \tag{19}$$

For a more general case, after the pilot transmission stage, the BS receives P signals, denoted as $y_{m,j}, j = 1, 2, ..., P$. For the i-th pilot transmission sub-frame, the extracted signals of each sub-array can be formulated as

$$\begin{bmatrix} \tilde{y}_{m,i,1} \\ \tilde{y}_{m,i,2} \\ \vdots \\ \tilde{y}_{m,i,N_s} \end{bmatrix} = \underbrace{\frac{1}{N_s}\mathbf{\Lambda}_{N_s}}_{\mathbf{P}^{-1}} \begin{bmatrix} y_{m,(i-1)N_s+1} \\ y_{m,(i-1)N_s+2} \\ \vdots \\ y_{m,iN_s} \end{bmatrix}. \tag{20}$$

Combining all pilot transmission sub-frames, the received signal corresponding to the k-th sub-array can be denoted as $\tilde{\mathbf{y}}_{m,k} = \left[\tilde{y}_{m,1,k}, \tilde{y}_{m,2,k}, ..., \tilde{y}_{m,P/N_s,k}\right], k = 1, 2, ..., N_s$. The corresponding observation matrix can be represented as

$$\mathbf{A}_{\text{sub},k} = \begin{bmatrix} \mathbf{A}_{1,k} \\ \mathbf{A}_{N_s+1,k} \\ \vdots \\ \mathbf{A}_{P-N_s+1,k} \end{bmatrix}. \tag{21}$$

Then, the received signal corresponding to the k-th sub-array can be represented as

$$\tilde{\mathbf{y}}_{m,k} = \mathbf{A}_{\text{sub},k}\mathbf{h}_{m,k} + \tilde{\mathbf{n}}_{m,k}, \tag{22}$$

and the received pilots corresponding to each sub-array are extracted. The basic idea of the proposed GTBC based signal extraction scheme is similar to the Alamouti Code, which can extract the data from different data streams in one antenna [40]. However, the scheme proposed in [40] require the complete CSI to operate normally, while the proposed scheme can perform well without the CSI and help us recognize the spatial non-stationarity in XL-MIMO systems. In addition, the schemes studied in [40] are hard to extend to systems with a large number of antennas or systems with different number of data streams, while our proposed scheme can work well in XL-MIMO systems and the number of sub-arrays can be adjusted flexibly.

Though the proposed scheme can apply to different N_s, it is helpful to determine the appropriate N_s for a better system performance. When the antenna array is divided into N_s sub-arrays, the effective pilot length decreases and becomes P/N_s. When N_s is large, the proposed scheme has a better ability to recognize the spatial non-stationary effect in the system, but accordingly, the channel estimation accuracy of each sub-array may decrease if N_s is set too large. On the other hand, when N_s is small, the spatial non-stationary effect may not be recognized properly. Therefore, for a given system, the N_s should be neither too small nor too large, and the proper N_s should be determined based on the system parameters, which is discussed in Section V.

IV. Proposed Non-stationary Channel Estimation Schemes

By applying the above GTBC based signal extraction scheme, an on-grid channel estimation algorithm called GTBC-based polar-domain simultaneous orthogonal matching pursuit (GP-SOMP) and an off-grid channel estimation algorithm called GTBC-based polar-domain simultaneous iterative gridless weighted (GP-SIGW) are proposed in this section to effectively recover the non-stationary XL-MIMO channel. In addition, the complexities and performances of the proposed algorithms are analyzed.

A. On-Grid Non-Stationary Channel Estimation

As mentioned in Section II, uplink channel estimation is carried out independently for each user since orthogonal pilot sequences are utilized. For a certain user, the received pilot at the m-th subcarrier corresponding to the k-th sub-array can be represented as

$$\tilde{\mathbf{y}}_{m,k} = \mathbf{A}_{\text{sub},k}\mathbf{W}_{N_s}\mathbf{h}^{\mathcal{P}}_{m,k} + \tilde{\mathbf{n}}_{m,k} = \mathbf{\Psi}_k\mathbf{h}^{\mathcal{P}}_{m,k} + \tilde{\mathbf{n}}_{m,k}, \tag{23}$$

Authorized licensed use limited to: Tsinghua University. Downloaded on March 16,2024 at 02:20:29 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2023.3343740

where $\mathbf{\Psi}_k = \mathbf{A}_{\text{sub},k}\mathbf{W}_{N_s}$ and $\mathbf{W}_{N_s}$ is the polar-domain codebook in [17].

During the combining and signal extraction process, the noise $\tilde{\mathbf{n}}_{m,k}$ is colored noise. Specifically, the noise corresponding to the i-th pilot transmission group can be represented as

$$\tilde{n}_{m,i,k} = \frac{1}{N_s}\mathbf{\Lambda}_{N_s}[k,:]\begin{bmatrix}\mathbf{A}_{(i-1)N_s+1}\mathbf{n}_{m,(i-1)N_s+1}\\ \mathbf{A}_{(i-1)N_s+2}\mathbf{n}_{m,(i-1)N_s+2}\\ \vdots \\ \mathbf{A}_{iN_s}\mathbf{n}_{m,iN_s}\end{bmatrix}. \quad (24)$$

Then, the covariance matrix of the combined noise is represented as

$$\mathbf{C} = \mathbb{E}\left(\tilde{\mathbf{n}}_{m,k}\tilde{\mathbf{n}}_{m,k}^H\right) = \text{blkdiag}\left(\sigma^2\mathbf{\Gamma}_1, \sigma^2\mathbf{\Gamma}_2, ..., \sigma^2\mathbf{\Gamma}_{P/N_s}\right), \quad (25)$$

where $\mathbf{\Gamma}_i = \frac{\sum_{j=1}^{N_s}\mathbf{A}_{(i-1)N_s+j}\mathbf{A}_{(i-1)N_s+j}^H}{N_s^2}$. The covariance matrix in (25) can be decomposed as $\mathbf{C} = \sigma^2\mathbf{D}\mathbf{D}^H$ by Cholesky factorization. The received signal is then whitened as

$$\bar{\mathbf{y}}_{m,k} = \mathbf{D}^{-1}\tilde{\mathbf{y}}_{m,k} = \bar{\mathbf{\Psi}}_k\mathbf{h}_{m,k}^{\mathcal{P}} + \bar{\mathbf{n}}_{m,k}, \quad (26)$$

where $\bar{\mathbf{\Psi}}_k = \mathbf{D}^{-1}\mathbf{\Psi}_k$ and $\bar{\mathbf{n}}_{m,k} = \mathbf{D}^{-1}\tilde{\mathbf{n}}_{m,k}$. Thus, the covariance matrix of $\bar{\mathbf{n}}_{m,k}$ is $\bar{\mathbf{C}} = \mathbf{D}^{-1}\mathbf{C}\mathbf{D}^{-H} = \sigma^2\mathbf{I}_{P/N_s}$ and the noise is properly whitened for further process.

In the considered scenario, the bandwidth is small compared to the central frequency. In this case, all subcarriers share the same sparse channel support thanks to the same steering vectors [11]. By estimating all subcarriers simultaneously, the channel estimation accuracy can be improved. Combining all subcarriers in (26), we get

$$\bar{\mathbf{Y}}_k = \mathbf{D}^{-1}\tilde{\mathbf{Y}}_k = \bar{\mathbf{\Psi}}_k\mathbf{H}_k^{\mathcal{P}} + \bar{\mathbf{N}}_k, \quad (27)$$

where $\bar{\mathbf{Y}}_k = [\bar{\mathbf{y}}_{1,k}, \bar{\mathbf{y}}_{2,k}, ..., \bar{\mathbf{y}}_{M,k}]$, $\tilde{\mathbf{Y}}_k = [\tilde{\mathbf{y}}_{1,k}, \tilde{\mathbf{y}}_{2,k}, ..., \tilde{\mathbf{y}}_{M,k}]$, $\mathbf{H}_k^{\mathcal{P}} = \left[\mathbf{h}_{1,k}^{\mathcal{P}}, \mathbf{h}_{2,k}^{\mathcal{P}}, ..., \mathbf{h}_{M,k}^{\mathcal{P}}\right]$, and $\bar{\mathbf{N}}_k = [\bar{\mathbf{n}}_{1,k}, \bar{\mathbf{n}}_{2,k}, ..., \bar{\mathbf{n}}_{M,k}]$. The $\bar{\mathbf{Y}}_k$ and $\bar{\mathbf{\Psi}}_k$ are utilized at the BS to estimate the channel [17]. After recovering the sub-channel with respect to each sub-array, a combining operation of the matrices can recover the entire channel matrix. The proposed GP-SOMP algorithm is summarized in **Algorithm 1**.

Algorithm 1 The Proposed GP-SOMP Algorithm

Input: Received signal $\mathbf{Y}$; combining matrix $\mathbf{A}$; path number $\hat{L}$; sub-array number N_s; pilot number P

Output: The estimation of the non-stationary channel $\hat{\mathbf{H}}$

1: **for** $i = 1$ to P/N_s **do**
2: Extract the received signals corresponding to each sub-array at the i-th pilot transmission group by (20)
3: **end for**
4: Combine the results in step 2 to get the received signals corresponding to each sub-array $\tilde{\mathbf{Y}}_k$
5: Construct the polar-domain codebook $\mathbf{W}_{N_s}$ by **Algorithm 1** in [17]
6: Convariance matrix $\mathbf{C} = \text{blkdiag}\left(\mathbf{\Gamma}_1, \mathbf{\Gamma}_2, ..., \mathbf{\Gamma}_{P/N_s}\right)$
7: Calculating the pre-whitening matrix $\mathbf{D}$ by Cholesky factorization $\mathbf{C} = \mathbf{D}\mathbf{D}^H$
8: **for** $k = 1$ to N_s **do**
9: Extract the combining matrix corresponding to the k-th sub-array $\mathbf{A}_{\text{sub},k}$ by (21)
10: Pre-whitening: $\bar{\mathbf{Y}}_k = \mathbf{D}^{-1}\tilde{\mathbf{Y}}_k$, $\bar{\mathbf{\Psi}}_k = \mathbf{D}^{-1}\mathbf{A}_{\text{sub},k}\mathbf{W}_{N_s}$
11: Initialization: residual $\mathbf{R} = \bar{\mathbf{Y}}_k$, support set $\Xi = \{\emptyset\}$
12: **for** $l = 1$ to $\hat{L}$ **do**
13: Correlation matrix: $\mathbf{\Phi} = \bar{\mathbf{\Psi}}_k^H\mathbf{R}$
14: New support: $\varrho^* = \arg\max_\varrho \sum_{m=1}^M \left|\mathbf{\Phi}(\varrho, m)\right|^2$
15: Update support set: $\Xi = \Xi \cup \varrho^*$
16: Update sub-channel: $\hat{\mathbf{H}}_k^{\mathcal{P}}[\varrho,:] = \bar{\mathbf{\Psi}}_k^\dagger[:,\varrho]\bar{\mathbf{Y}}_k$
17: Update residual: $\mathbf{R} = \mathbf{R} - \bar{\mathbf{\Psi}}_k[:,\varrho]\hat{\mathbf{H}}_k^{\mathcal{P}}[\varrho,:]$
18: **end for**
19: $\hat{\mathbf{H}}_k = \mathbf{W}_{N_s}[:,\varrho]\hat{\mathbf{H}}_k^{\mathcal{P}}[\varrho,:]$
20: **end for**
21: $\hat{\mathbf{H}} = \left[\hat{\mathbf{H}}_1^T, \hat{\mathbf{H}}_2^T, ..., \hat{\mathbf{H}}_{N_s}^T\right]^T$

Specifically, in steps 1-4, the received signals with respect to each sub-array are extracted by the scheme in Section III. Then, in steps 5-7, the polar-domain codebook and the pre-whitening matrix are calculated for the following procedure. Next in steps 8-20, the sub-channels corresponding to each sub-array are estimated respectively and then combined to recover the entire channel. In steps 9-11, initialization is carried out and the received signals are pre-whitened. Then, the estimation is carried out for each path. The correlation matrix $\mathbf{\Phi}$ is firstly calculated in step 13. Then, the common support ϱ^* is determined as $\varrho^* = \arg\max_\varrho \sum_{m=1}^M \left|\mathbf{\Phi}(\varrho, m)\right|^2$ in step 14. The support set is then updated in step 15, and in step 16, the path gain of the l-th path component is acquired by orthogonal least square. After this, the residual is then updated. Repeating steps 13-17 $\hat{L}$ times, all path components in the system can be detected. In practical systems, the estimation of path number $\hat{L}$ can be obtained according to the statistical CSI during a long period of time. Then, the sub-channel corresponding to the k-th sub-array is recovered as $\hat{\mathbf{H}}_k = \mathbf{W}_{N_s}[:,\varrho]\hat{\mathbf{H}}_k^{\mathcal{P}}[\varrho,:]$ in step 19. After the sub-channels corresponding to all sub-arrays are estimated, the entire estimated channel are recovered in step 21 by combining all the sub-channels.

The proposed GP-SOMP is capable of distinguishing the received signals corresponding to different sub-arrays. Thus, the mismatch of the steering vectors brought by the spatial non-stationary effect can be properly settled. Moreover, since the number of sub-arrays can be flexibly adjusted, the proposed GP-SOMP algorithm also works well in the spatial stationary scenario[3].

However, the accuracy of the above algorithm is limited for two reasons. On one hand, the on-grid assumption limits the resolution of the proposed GP-SOMP algorithm. On the other hand, since the number of antennas in each sub-array is N_s times smaller than that in the entire array, the sampled number correspondingly decreases, which further decreases the accuracy. Therefore, in the next subsection, we proposed an off-grid GP-SIGW algorithm to further improve the channel estimation accuracy.

[3]The idea of combining different sub-arrays for analysis may increase the channel estimation accuracy [27]. However, in the near field, whether the supports belong to the same scatterer/user is hard to confirm, so such idea is not practical in the considered scenarios.

Authorized licensed use limited to: Tsinghua University. Downloaded on March 16,2024 at 02:20:29 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2023.3343740

B. Off-Grid Non-Stationary Channel Estimation

To reduce the channel estimation error caused by the on-grid sampling, we propose an off-grid GP-SIGW algorithm. In this algorithm, the angles and distances of the paths corresponding to each sub-array are updated according to the maximum likelihood principle, and the path gains are obtained by the least square principle. The proposed GP-SIGW algorithm is illustrated in **Algorithm 2**.

Algorithm 2 The Proposed GP-SIGW Algorithm

Input: Received signal $\mathbf{Y}$; combining matrix $\mathbf{A}$; path number $\hat{L}$; sub-array number N_s; pilot number P; maximum iteration number N_{iter}

Output: The estimation of the non-stationary channel $\hat{\mathbf{H}}$

1: **for** $i = 1$ to P/N_s **do**
2: Extract the received signals corresponding to each sub-array at the i-th pilot transmission group by (20)
3: **end for**
4: Combine the results in step 2 to get the received signals corresponding to each sub-array $\bar{\mathbf{Y}}_k$
5: **for** $k = 1$ to N_s **do**
6: Initialize the distances $\hat{\mathbf{r}}_k^0 = \left[\hat{r}_{k,1}^0, \hat{r}_{k,2}^0, ..., \hat{r}_{k,\hat{L}}^0\right]$ and the angles $\hat{\boldsymbol{\theta}}_k^0 = \left[\hat{\theta}_{k,1}^0, \hat{\theta}_{k,2}^0, ..., \hat{\theta}_{k,\hat{L}}^0\right]$ by **Algorithm 1**.
7: **for** $n = 1$ to N_{iter} **do**
8: Choose the line search step l_1 by Goldstein condition
9: Update the distances by $\frac{1}{\hat{\mathbf{r}}_k^n} = \frac{1}{\hat{\mathbf{r}}_k^{n-1}} - l_1 \nabla_{\frac{1}{\hat{\mathbf{r}}_k}} \mathcal{L}\left(\hat{\mathbf{r}}_k, \hat{\boldsymbol{\theta}}_k^{n-1}\right)\Big|_{\hat{\mathbf{r}}_k = \hat{\mathbf{r}}_k^{n-1}}$ in (33)
10: Choose the line search step l_2 by Goldstein condition
11: Update the angles by $\hat{\boldsymbol{\theta}}_k^n = \hat{\boldsymbol{\theta}}_k^{n-1} - l_2 \nabla_{\hat{\boldsymbol{\theta}}_k} \mathcal{L}\left(\hat{\mathbf{r}}_k^n, \hat{\boldsymbol{\theta}}_k\right)\Big|_{\hat{\boldsymbol{\theta}}_k = \hat{\boldsymbol{\theta}}_k^{n-1}}$ in (34)
12: Update the path gains by $\hat{\mathbf{G}}_k^{\text{opt}} = \hat{\boldsymbol{\Psi}}_k^{\dagger}\left(\hat{\mathbf{r}}_k^n, \hat{\boldsymbol{\theta}}_k^n\right)\bar{\mathbf{Y}}_k$ in (30)
13: **end for**
14: $\hat{\mathbf{H}}_k = \left[\mathbf{b}\left(\hat{\theta}_{1,k}^{N_{\text{iter}}}, \hat{r}_{1,k}^{N_{\text{iter}}}\right), ..., \mathbf{b}\left(\hat{\theta}_{\hat{L},k}^{N_{\text{iter}}}, \hat{r}_{\hat{L},k}^{N_{\text{iter}}}\right)\right]$
15: **end for**
16: $\hat{\mathbf{H}} = \left[\hat{\mathbf{H}}_1^T, \hat{\mathbf{H}}_2^T, ..., \hat{\mathbf{H}}_{N_s}^T\right]^T$

Specifically, the proposed GP-SIGW algorithm is still carried out for each sub-array and then combined to recover the entire channel. In steps 1-4, the received signals with respect to each sub-array are firstly extracted. Then, in steps 5-15, the sub-channels with respect to each sub-array are refined based on **Algorithm 1**. In step 6, the initial results of the distances, angles and complex path gains are obtained as a start point of the iteration. Then, the target of the iteration can be formulated as

$$\min_{\hat{\mathbf{r}}_k, \hat{\boldsymbol{\theta}}_k, \hat{\mathbf{G}}_k} \left\|\hat{\mathbf{Y}}_k - \hat{\boldsymbol{\Psi}}_k\left(\hat{\mathbf{r}}_k, \hat{\boldsymbol{\theta}}_k\right)\hat{\mathbf{G}}_k\right\|_F^2, \tag{28}$$

where $\hat{\boldsymbol{\Psi}}_k\left(\hat{\mathbf{r}}_k, \hat{\boldsymbol{\theta}}_k\right)$ can be represented as $\hat{\boldsymbol{\Psi}}_k\left(\hat{\mathbf{r}}_k, \hat{\boldsymbol{\theta}}_k\right) = \mathbf{D}^{-1}\mathbf{A}_{\text{sub},k}\hat{\mathbf{W}}\left(\hat{\boldsymbol{\theta}}_k, \hat{\mathbf{r}}_k\right)$, with $\hat{\mathbf{W}}\left(\hat{\boldsymbol{\theta}}_k, \hat{\mathbf{r}}_k\right)$ being

$$\hat{\mathbf{W}}\left(\hat{\boldsymbol{\theta}}_k, \hat{\mathbf{r}}_k\right) = \left[\mathbf{b}\left(\hat{\theta}_{1,k}, \hat{r}_{1,k}\right), ..., \mathbf{b}\left(\hat{\theta}_{\hat{L},k}, \hat{r}_{\hat{L},k}\right)\right]. \tag{29}$$

Due to the non-convex property of (28), the alternating minimization scheme is utilized. For fixed $\hat{\mathbf{r}}_k$ and $\hat{\boldsymbol{\theta}}_k$, the optimal solution for $\hat{\mathbf{G}}_k$ can be directly derived as

$$\hat{\mathbf{G}}_k^{\text{opt}} = \hat{\boldsymbol{\Psi}}_k^{\dagger}\left(\hat{\mathbf{r}}_k, \hat{\boldsymbol{\theta}}_k\right)\bar{\mathbf{Y}}_k. \tag{30}$$

Then, denoting $\hat{\boldsymbol{\Psi}}_k\left(\hat{\mathbf{r}}_k, \hat{\boldsymbol{\theta}}_k\right)\hat{\boldsymbol{\Psi}}_k^{\dagger}\left(\hat{\mathbf{r}}_k, \hat{\boldsymbol{\theta}}_k\right)$ as $\mathbf{T}\left(\hat{\mathbf{r}}_k, \hat{\boldsymbol{\theta}}_k\right)$, (28) can be written as

$$\begin{aligned}&\left\|\bar{\mathbf{Y}}_k - \mathbf{T}\left(\hat{\mathbf{r}}_k, \hat{\boldsymbol{\theta}}_k\right)\bar{\mathbf{Y}}_k\right\|_F^2\\ =&\text{Tr}\left\{\bar{\mathbf{Y}}_k^H\left(\mathbf{I} - \mathbf{T}\left(\hat{\mathbf{r}}_k, \hat{\boldsymbol{\theta}}_k\right)\right)^H\left(\mathbf{I} - \mathbf{T}\left(\hat{\mathbf{r}}_k, \hat{\boldsymbol{\theta}}_k\right)\right)\bar{\mathbf{Y}}_k\right\}\\ =&\text{Tr}\left\{\bar{\mathbf{Y}}_k^H\bar{\mathbf{Y}}_k - \bar{\mathbf{Y}}_k^H\mathbf{T}\left(\hat{\mathbf{r}}_k, \hat{\boldsymbol{\theta}}_k\right)\bar{\mathbf{Y}}_k\right\}.\end{aligned} \tag{31}$$

Therefore, the new optimization problem can be represented as

$$\max_{\hat{\mathbf{r}}_k, \hat{\boldsymbol{\theta}}_k} \mathcal{L}\left(\hat{\mathbf{r}}_k, \hat{\boldsymbol{\theta}}_k\right) = \text{Tr}\left\{\bar{\mathbf{Y}}_k^H\mathbf{T}\left(\hat{\mathbf{r}}_k, \hat{\boldsymbol{\theta}}_k\right)\bar{\mathbf{Y}}_k\right\}. \tag{32}$$

$\mathcal{L}\left(\hat{\mathbf{r}}_k, \hat{\boldsymbol{\theta}}_k\right)$ can be optimized by an iterative gradient descent approach. For the distance, since $\frac{1}{r}$ is uniformly sampled when generating $\mathbf{W}_{N_s}$, the gradient descent approach expressed in the inverse-distance domain, i.e., $\frac{1}{\hat{\mathbf{r}}_k} = \left[\frac{1}{\hat{r}_{k,1}}, \frac{1}{\hat{r}_{k,2}}, ..., \frac{1}{\hat{r}_{k,\hat{L}}}\right]$. In the n-th iteration, the inverse of the distances are updated as

$$\frac{1}{\hat{\mathbf{r}}_k^n} = \frac{1}{\hat{\mathbf{r}}_k^{n-1}} - l_1\mathbf{G}_{\frac{1}{\hat{\mathbf{r}}_k}}, \tag{33}$$

where l_1 denotes the step length for updating the inverse of distances and $\mathbf{G}_{\frac{1}{\hat{r}_k}}$ denotes the gradient satisfying $\mathbf{G}_{\frac{1}{\hat{\mathbf{r}}_k}} = \nabla_{\frac{1}{\hat{\mathbf{r}}_k}}\mathcal{L}\left(\hat{\mathbf{r}}_k, \hat{\boldsymbol{\theta}}_k^{n-1}\right)\Big|_{\hat{\mathbf{r}}_k = \hat{\mathbf{r}}_k^{n-1}}$. Similarly, the angles can be updated as

$$\hat{\boldsymbol{\theta}}_k^n = \hat{\boldsymbol{\theta}}_k^{n-1} - l_2\mathbf{G}_{\hat{\theta}_k}, \tag{34}$$

where l_2 denotes the step length for updating the angles and $\mathbf{G}_{\hat{\theta}_k}$ denotes the gradient satisfying $\mathbf{G}_{\hat{\theta}_k} = \nabla_{\hat{\boldsymbol{\theta}}_k}\mathcal{L}\left(\hat{\mathbf{r}}_k^n, \hat{\boldsymbol{\theta}}_k\right)\Big|_{\hat{\boldsymbol{\theta}}_k = \hat{\boldsymbol{\theta}}_k^{n-1}}$. the step lengths are decided by Goldstein backtracking line search, which guarantees fast convergence as well as the increment during every iteration at the same time. The gradient of the objective function $\mathcal{L}\left(\hat{\mathbf{r}}_k, \hat{\boldsymbol{\theta}}_k\right)$ is derived in [17]. Based on the above derivation, the parameters in (28) are updated in steps 8-12. The sub-channel corresponding to the k-th sub-array is thus concatenated based on the refined parameters after N_{iter} rounds of iterations in step 14. Finally, the entire refined channel are recoverd in step 16 by combining all the sub-channels. In section V, the performances of the proposed GP-SOMP algorithm and GP-SIGW algorithm are evaluated the demonstrate the efficiency of proposed algorithms.

C. Complexity and Performance Analysis

Authorized licensed use limited to: Tsinghua University. Downloaded on March 16,2024 at 02:20:29 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2023.3343740

TABLE I
COMPLEXITY COMPARISON

Scheme	Complexity $\mathcal{O}(\cdot)$
GP-SOMP	$\mathcal{O}(\hat{L}PN_{\mathrm{RF}}NSM/N_s)$
P-SOMP	$\mathcal{O}(\hat{L}PN_{\mathrm{RF}}NSM)$
GP-SIGW	$\mathcal{O}(\hat{L}PN_{\mathrm{RF}}NSM/N_s)+$ $\mathcal{O}(N_{\mathrm{iter}}(P^2N_{\mathrm{RF}}^2M/N_s+PN_{\mathrm{RF}}M^2))$
P-SIGW	$\mathcal{O}(\hat{L}PN_{\mathrm{RF}}NSM)+$ $\mathcal{O}(N_{\mathrm{iter}}(P^2N_{\mathrm{RF}}^2M+PN_{\mathrm{RF}}M^2))$

1) Complexity Analysis: For the proposed GP-SOMP algorithm in **Algorithm 1**, the SOMP procedure is the dominant factor of the overall complexity. For a certain path corresponding to a sub-array, the dimensions of $\bar{\mathbf{\Psi}}_k$, $\mathbf{R}$ and $\bar{\mathbf{Y}}_k$ are $PN_{\mathrm{RF}}/N_s \times SN/N_s$, $PN_{\mathrm{RF}}/N_s \times M$ and $PN_{\mathrm{RF}}/N_s \times M$, respectively. Therefore, the computation complexities of step 13, 14, 16, 17 are $\mathcal{O}(PN_{\mathrm{RF}}NSM/N_s^2)$, $\mathcal{O}(NSM/N_s^2)$, $\mathcal{O}(\hat{L}^2PN_{\mathrm{RF}}/N_s+\hat{L}PN_{\mathrm{RF}}M/N_s)$, and $\mathcal{O}(\hat{L}PN_{\mathrm{RF}}M/N_s)$, respectively. Due to the sparse nature of high-frequency communication systems, $\hat{L}$ is generally small, so the complexity of SOMP procedure can be formulated by $\mathcal{O}(PN_{\mathrm{RF}}NSM/N_s^2)$. Then, the overall computational complexity of the iteration part is $\mathcal{O}(\hat{L}PN_{\mathrm{RF}}NSM/N_s)$. As compared in Table I, the computational complexity of the proposed GP-SOMP is N_s times smaller than that of the spatial stationary near-field scheme P-SOMP [17]. This is because the antenna number and pilot number for each sub-array are all N_s times smaller than the entire antenna array.

For the proposed GP-SIGW algorithm in **Algorithm 2**, the computational complexity contains two parts, the first parts is the initialization stage, which has the same computational complexity as GP-SOMP algorithm, i.e., $\mathcal{O}(\hat{L}PN_{\mathrm{RF}}NSM/N_s)$. In the refinement stage, the complexity is mainly determined by updating $\hat{\mathbf{r}}_k$, $\hat{\boldsymbol{\theta}}_k$, and $\hat{\mathbf{G}}_k$. As for updating $\hat{\mathbf{r}}_k$ and $\hat{\boldsymbol{\theta}}_k$, the computational complexities are $\mathcal{O}(P^2N_{\mathrm{RF}}^2M/N_s^2+PN_{\mathrm{RF}}M^2/N_s)$, $\mathcal{O}(\hat{L}^2PN_{\mathrm{RF}}/N_s+\hat{L}P^2N_{\mathrm{RF}}^2/N_s^2)$, $\mathcal{O}(\hat{L}^3+\hat{L}^2PN_{\mathrm{RF}}/N_s)$, and $\mathcal{O}(PN_{\mathrm{RF}}N/N_s)$. As for updating $\hat{\mathbf{G}}_k$, the computational complexity of (30) is $\mathcal{O}(\hat{L}^2PN_{\mathrm{RF}}/N_s+\hat{L}PN_{\mathrm{RF}}M/N_s)$. Since $\hat{L}$ is small, the complexity of the updating the parameters is determined by $\mathcal{O}(P^2N_{\mathrm{RF}}^2M/N_s^2+PN_{\mathrm{RF}}M^2/N_s)$. After the iterations, the complexity of refinement for all sub-arrays is $\mathcal{O}(N_{\mathrm{iter}}(P^2N_{\mathrm{RF}}^2M/N_s+PN_{\mathrm{RF}}M^2))$. Therefore, the complexity is $\mathcal{O}(\hat{L}PN_{\mathrm{RF}}NSM/N_s+N_{\mathrm{iter}}(P^2N_{\mathrm{RF}}^2M/N_s+PN_{\mathrm{RF}}M^2))$. From Table I, the complexity of the proposed GP-SIGW algorithm is slightly smaller than that of the spatial stationary near-field scheme P-SIGW [17]. This is because the antenna number for each sub-array is N_s times smaller than the entire antenna number, which reduce a the computational complexity to a large extent when calculating the inverse of the matrices.

2) Performance Analysis: For the proposed GP-SOMP algorithm, we prove that the physical distances and angles can be accurately recovered with a guaranteed probability. Since the procedure of channel estimation is totally the same for different sub-arrays, the following analysis is based on only one sub-array and the subscript k is omitted. Thus (26) can be reformulated as

$$\mathbf{y}=\mathbf{\Psi}\mathbf{h}^{\mathcal{P}}+\mathbf{n}, \tag{35}$$

where the polar-domain channel has the known sparsity $\left\|\mathbf{h}^{\mathcal{P}}\right\|_0=L$, and $\mathbf{n}\sim\mathcal{N}\left(\mathbf{0},\sigma^2\mathbf{I}\right)$ is the random noise vector. We denote Ξ as the support set of $\mathbf{h}^{\mathcal{P}}$ of size L. Then, the SOMP algorithm can detect the entire support set if and only if

$$\min_{j\in\Xi}\left|\boldsymbol{\psi}_j^H\mathbf{y}\right|>\max_{j\notin\Xi}\left|\boldsymbol{\psi}_j^H\mathbf{y}\right|, \tag{36}$$

with $\boldsymbol{\psi}_j$ being the j-th column of $\mathbf{\Psi}$ [41]. To analyze the bound of the above equation, we consider a special event $\mathcal{E}$ where the random noise is bounded by a constant τ, which can be defined as

$$\mathcal{E}=\left\{\max_{1\le i\le s}\left|\boldsymbol{\psi}_i^H\mathbf{n}\right|<\tau\right\}, \tag{37}$$

with s being the size of the polar-domain transform matrix $\mathbf{W}_{N_s}$. We denote μ as

$$\mu\triangleq\max_{i\neq j}\left|\boldsymbol{\psi}_i^H\boldsymbol{\psi}_j\right|. \tag{38}$$

Then, under the event $\mathcal{E}$,

$$\begin{aligned}\min_{j\in\Xi}\left|\boldsymbol{\psi}_j^H\mathbf{y}\right|&=\min_{j\in\Xi}\left|h_j+\boldsymbol{\psi}_j^H\mathbf{n}+\sum_{i\in\Xi\setminus\{j\}}h_i\boldsymbol{\psi}_j^H\boldsymbol{\psi}_i\right|\\&>\left|h_{\min}\right|-\tau-(L-1)\mu\left|h_{\max}\right|,\end{aligned} \tag{39}$$

where h_i denotes the i-th element in $\mathbf{h}^{\mathcal{P}}$, $|h_{\min}|=\min_{i\in\Xi}|h_i|$, and $|h_{\max}|=\max_{i\in\Xi}|h_i|$, respectively. Similarly,

$$\begin{aligned}\max_{j\notin\Xi}\left|\boldsymbol{\psi}_j^H\mathbf{y}\right|&=\max_{j\notin\Xi}\left|\boldsymbol{\psi}_j^H\mathbf{n}+\sum_{i\in\Xi}h_i\boldsymbol{\psi}_j^H\boldsymbol{\psi}_i\right|\\&\le\max_{j\notin\Xi}\left|\boldsymbol{\psi}_j^H\mathbf{n}\right|+\max_{j\notin\Xi}\sum_{i\in\Xi}\left|h_i\boldsymbol{\psi}_j^H\boldsymbol{\psi}_i\right|\\&<\tau+L\mu\left|h_{\max}\right|.\end{aligned} \tag{40}$$

Substituting (39), (40) into (36), the condition can be represented as

$$\left|h_{\min}\right|-(2L-1)\mu\left|h_{\max}\right|\ge2\tau. \tag{41}$$

As for event $\mathcal{E}$, the probability can be represented as

$$\Pr\{\mathcal{E}\}=\Pr\left\{\max_{1\le i\le s}\left|\boldsymbol{\psi}_i^H\mathbf{n}\right|<\tau\right\}\ge\prod_{i=1}^{s}\Pr\left\{\left|\boldsymbol{\psi}_i^H\mathbf{n}\right|\le\tau\right\} \tag{42}$$

For the i-th column in $\mathbf{\Psi}$, each random variable $\boldsymbol{\psi}_i^H\mathbf{n}$ satisfies $\mathcal{N}\left(0,\sigma^2\right)$, so the probability is

$$\Pr\left\{\left|\boldsymbol{\psi}_i^H\mathbf{n}\right|\le\tau\right\}=1-2Q\left(\frac{\tau}{\sigma}\right), \tag{43}$$

with $Q(x)=(1/\sqrt{2\pi})\int_x^\infty e^{-z^2/2}dz$ being the Gaussian tail probability, which has the bound

$$Q(x)\le\frac{1}{x\sqrt{2\pi}}e^{-x^2/2}. \tag{44}$$

Authorized licensed use limited to: Tsinghua University. Downloaded on March 16,2024 at 02:20:29 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2023.3343740

Then, the probability for each random variable becomes

$$\Pr\left\{\left|\boldsymbol{\psi}_i^H \mathbf{n}\right| \leq \tau\right\} \geq 1-\sqrt{\frac{2}{\pi}}\frac{\sigma}{\tau}e^{-\tau^2/2\sigma^2}, \tag{45}$$

so, the probability of event $\mathcal{E}$ can be represented as

$$\Pr\{\mathcal{E}\} \geq \left(1-\sqrt{\frac{2}{\pi}}\frac{\sigma}{\tau}e^{-\tau^2/2\sigma^2}\right)^{s}. \tag{46}$$

Then, for all N_s sub-arrays, the probability of recovering all the paths can be represented as

$$\Pr\{\mathcal{E}\} \geq \left(1-\sqrt{\frac{2}{\pi}}\frac{\sigma}{\tau}e^{-\tau^2/2\sigma^2}\right)^{sN_s}. \tag{47}$$

Thus, the performance of the proposed GP-SOMP algorithm can be stated as follows. For a certain sub-array, we assume the physical distances and angles correponding to all paths locate on the polar-domain samples. When (41) holds, the proposed GP-SOMP algorithm can accurately estimate all the paths with a probability exceeding (47).

For the proposed GP-SIGW algorithm, the performance is mainly determined by the convergence. The objective function $\left\|\hat{\mathbf{Y}}_k-\hat{\boldsymbol{\Psi}}_k\left(\hat{\mathbf{r}}_k,\hat{\boldsymbol{\theta}}_k\right)\hat{\mathbf{G}}_k\right\|_F^2$ is positive, and thus has a lower bound related to the pilot length and the noise. In each iteration, since Goldstein condition is considered when selecting the step length, the objective function will not increase and the step length will not be too small. In addition, the $\hat{\mathbf{G}}_k^{\text{opt}}$ is the optimal solution in each iteration. Therefore, the alternating minimization scheme in the proposed GP-SIGW algorithm will converge. The convergence is verified in Section V.

D. Cramér-Rao Lower Bound

As a theorerical bound of MSE, Cramér-Rao Lower Bound (CRLB) can assess the effectiveness of the proposed channel estimation algorithms. We first derive the CRLB of estimating the sub-channel of the k-th sub-array at the m-th subcarrier. For brevity, the channel estimation problem can be reformulated as

$$\mathbf{y}_{m,k}=\mathbf{A}_k\mathbf{h}_{m,k}+\mathbf{n}_{m,k}, \tag{48}$$

where $\mathbf{y}_{m,k}\in\mathbb{C}^{\frac{N_{\text{RF}}P}{N_s}\times 1}$ denotes the whitened received signals, $\mathbf{A}_k\in\mathbb{C}^{\frac{N_{\text{RF}}P}{N_s}\times\frac{N}{N_s}}$ denotes the combining matrices corresponding to the k-th sub-array, $\mathbf{h}_{m,k}\in\mathbb{C}^{\frac{N}{N_s}\times 1}$ denotes the sub-channel corresponding to the k-th sub-array, and $\mathbf{n}_{m,k}\in\mathbb{C}^{\frac{N_{\text{RF}}P}{N_s}\times 1}$ denotes the whitened noise. During the signal extraction process, the noise power becomes $\frac{1}{N_s}$ of the original noise. Since the noise of the entire system follows the distribution $\mathcal{CN}(\mathbf{0},\sigma^2\mathbf{I}_N)$, so the noise here follows $\mathbf{n}_{m,k}\sim\mathcal{CN}(\mathbf{0},\frac{\sigma^2}{N_s}\mathbf{I}_{\frac{N_{\text{RF}}P}{N_s}})$. In our proposed algorithms, the elements in the combining matrix $\mathbf{A}_k$ are all real, while the elements in $\mathbf{y}_{m,k}$, $\mathbf{h}_{m,k}$, and $\mathbf{n}_{m,k}$ are all complex, so (48) can be split into two parts: the real part and the imaginary part, which can be represented as

$$\begin{aligned}\mathbf{y}_{m,k}^{\mathcal{R}}&=\mathbf{A}_k\mathbf{h}_{m,k}^{\mathcal{R}}+\mathbf{n}_{m,k}^{\mathcal{R}},\\ \mathbf{y}_{m,k}^{\mathcal{I}}&=\mathbf{A}_k\mathbf{h}_{m,k}^{\mathcal{I}}+\mathbf{n}_{m,k}^{\mathcal{I}},\end{aligned} \tag{49}$$

where $\mathbf{y}_{m,k}^{\mathcal{R}}=\text{Re}\{\mathbf{y}_{m,k}\}$, $\mathbf{y}_{m,k}^{\mathcal{I}}=\text{Im}\{\mathbf{y}_{m,k}\}$, $\mathbf{h}_{m,k}^{\mathcal{R}}=\text{Re}\{\mathbf{h}_{m,k}\}$, $\mathbf{h}_{m,k}^{\mathcal{I}}=\text{Im}\{\mathbf{h}_{m,k}\}$, $\mathbf{n}_{m,k}^{\mathcal{R}}=\text{Re}\{\mathbf{n}_{m,k}\}$, $\mathbf{n}_{m,k}^{\mathcal{I}}=\text{Im}\{\mathbf{n}_{m,k}\}$. The result of the estimated channel are denoted as $\hat{\mathbf{h}}_{m,k}=\hat{\mathbf{h}}_{m,k}^{\mathcal{R}}+\hat{\mathbf{h}}_{m,k}^{\mathcal{I}}$, so the CRLB of the unbiased estimator $\hat{\mathbf{h}}_{m,k}$ can be formulated as

$$\begin{aligned}\text{CRLB}_{m,k}&=\mathbb{E}\left\{\left\|\hat{\mathbf{h}}_{m,k}-\mathbf{h}_{m,k}\right\|^2\right\}\\ &=\text{CRLB}_{m,k}^{\mathcal{R}}+\text{CRLB}_{m,k}^{\mathcal{I}}\\ &=\mathbb{E}\left\{\left\|\hat{\mathbf{h}}_{m,k}^{\mathcal{R}}-\mathbf{h}_{m,k}^{\mathcal{R}}\right\|^2\right\}+\mathbb{E}\left\{\left\|\hat{\mathbf{h}}_{m,k}^{\mathcal{I}}-\mathbf{h}_{m,k}^{\mathcal{I}}\right\|^2\right\}.\end{aligned} \tag{50}$$

We first consider the real part of (50). The conditional probability density function of $\mathbf{y}_{m,k}^{\mathcal{R}}$ with the given $\mathbf{h}_{m,k}^{\mathcal{R}}$ can be derived as

$$\begin{aligned}&p_{\mathbf{y}_{m,k}^{\mathcal{R}}|\mathbf{h}_{m,k}^{\mathcal{R}}}\left(\mathbf{y}_{m,k}^{\mathcal{R}};\mathbf{h}_{m,k}^{\mathcal{R}}\right)\\ &=\frac{1}{\left(2\pi\frac{\sigma^2}{N_s}\right)^{\frac{N_{\text{RF}}P}{2N_s}}}\exp\left\{-\frac{N_s}{2\sigma^2}\left\|\mathbf{y}_{m,k}^{\mathcal{R}}-\mathbf{A}_k\mathbf{h}_{m,k}^{\mathcal{R}}\right\|^2\right\}.\end{aligned} \tag{51}$$

Then, the Fisher information matrix of the real part can be derived as

$$\begin{aligned}&\left[\mathbf{J}_{\mathcal{R}}\right]_{\iota,\ell}\\ &\triangleq-\mathbb{E}\left\{\frac{p_{\mathbf{y}_{m,k}^{\mathcal{R}}|\mathbf{h}_{m,k}^{\mathcal{R}}}\left(\mathbf{y}_{m,k}^{\mathcal{R}};\mathbf{h}_{m,k}^{\mathcal{R}}\right)}{\partial h_{m,k,\iota}^{\mathcal{R}}\partial h_{m,k,\ell}^{\mathcal{R}}}\right\}=\frac{N_s}{\sigma^2}\left[\mathbf{A}_k^H\mathbf{A}_k\right]_{\iota,\ell},\end{aligned} \tag{52}$$

with $\partial h_{m,k,\iota}^{\mathcal{R}}$, $\partial h_{m,k,\ell}^{\mathcal{R}}$ being the ι-th and the ℓ-th element in $\mathbf{h}_{m,k}^{\mathcal{R}}$. Thus, the CRLB of the real part is

$$\begin{aligned}\text{CRLB}_{m,k}^{\mathcal{R}}&=\mathbb{E}\left\{\left\|\hat{\mathbf{h}}_{m,k}^{\mathcal{R}}-\mathbf{h}_{m,k}^{\mathcal{R}}\right\|^2\right\}\\ &\geq\text{Tr}\left\{\mathbf{J}_{\mathcal{R}}^{-1}\right\}=\frac{\sigma^2}{N_s}\text{Tr}\left\{\left(\mathbf{A}_k^H\mathbf{A}_k\right)^{-1}\right\}.\end{aligned} \tag{53}$$

Since $\text{Tr}\left\{\left(\mathbf{A}_k^H\mathbf{A}_k\right)^{-1}\right\}$ satisfies

$$\begin{aligned}\text{Tr}\left\{\left(\mathbf{A}_k^H\mathbf{A}_k\right)^{-1}\right\}&=\left(\sum_{i=1}^{N/N_s}\lambda_i^{-1}\right)\\ &\overset{(a)}{\geq}\frac{N}{N_s}\left(\frac{N}{N_s}/\sum_{i=1}^{N/N_s}\lambda_i\right)=\frac{N^2}{N_s^2\text{Tr}\left\{\mathbf{A}_k^H\mathbf{A}_k\right\}},\end{aligned} \tag{54}$$

where $\{\lambda_i\}_{i=1}^{N/N_s}$ denotes the eigenvalues of $\mathbf{A}_k^H\mathbf{A}_k$. For (a), the equality holds when $\lambda_1=\lambda_2=\cdots=\lambda_{N/N_s}$, i.e., the columns of $\mathbf{A}_k$ are orthogonal. In this case, $\mathbf{A}_k^H\mathbf{A}_k$ has identical diagonals equal to $\frac{N_{\text{RF}P}}{N_s}$ and $\text{Tr}\left\{\mathbf{A}_k^H\mathbf{A}_k\right\}=\frac{N_{\text{RF}}PN}{N_s^2}$. The CRLB of the real part can be finally obtained as

$$\text{CRLB}_{m,k}^{\mathcal{R}}=\sigma^2\frac{N}{N_sN_{\text{RF}}P}. \tag{55}$$

Since the imaginary part has the same form as the real part, the overall CRLB of the k-th sub-array at the m-th subcarrier is

$$\text{CRLB}_{m,k}=\text{CRLB}_{m,k}^{\mathcal{R}}+\text{CRLB}_{m,k}^{\mathcal{I}}=2\sigma^2\frac{N}{N_sN_{\text{RF}}P}. \tag{56}$$

Authorized licensed use limited to: Tsinghua University. Downloaded on March 16,2024 at 02:20:29 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2023.3343740

TABLE II
SIMULATION PARAMETERS

BS antenna number N	512
Subcarrier number M	256
User number K	4
RF chain number N_{RF}	4
Central carrier frequency f_c	100 GHz
Bandwidth B	100 MHz
The distribution of θ	$\mathcal{U}\left(-\frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2}\right)$
Path number L for each user	3

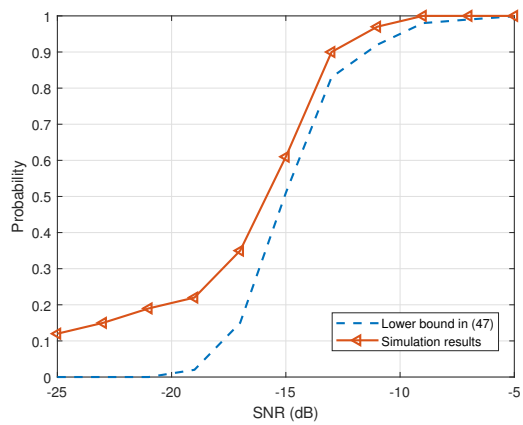


Fig. 4. The probability of recovering all paths by the proposed GP-SOMP scheme.

For the entire antenna array, the CRLB at the m-th subcarrier can then be derived as

$$\mathrm{CRLB} = \sum_{m=1}^{M}\sum_{k=1}^{N_s}\mathrm{CRLB}_{m,k} = 2\sigma^2\frac{NM}{N_{\mathrm{RF}}P}. \tag{57}$$

With the CRLB, the performance of the proposed algorithms can be evaluated more comprehensively, which is elaborated in Section V.

V. SIMULATION RESULTS

In this section, numerical simulations are carried out to evaluate the performance of the proposed non-stationary channel estimation schemes. We utilize the normalized mean square error (NMSE) to evaluate the performance. Specifically, the NMSE is defined as $\mathrm{NMSE} = \mathbb{E}\left(\frac{\|\mathbf{H}-\hat{\mathbf{H}}\|_2^2}{\|\mathbf{H}\|_2^2}\right)$ is utilized to evaluate the performance. In the simulation, we consider a multi-user XL-MIMO OFDM system, and the system parameters for simulations are elaborated in Table II. For each path, the VR is chosen randomly from all sub-arrays.

Fig. 4 provides the verification of the lower bound of the probability of recovering all paths by the proposed GP-SOMP algorithm. For the first path, the sine of the physical angle $\theta_1 = -0.8652$, the distance $r_1 = 17.16$ m, and the first path is visible to the first and the second sub-arrays. For the second path, the sine of the physical angle $\theta_2 = -0.5$, the distance $r_2 = 7.31$ m, and the second path is visible to the third sub-array. For the third path, the sine of the physical angle $\theta_3 = 0.7031$, the distance $r_3 = 11.51$ m, and the third path is visible to the third and the fourth sub-arrays. It is illustrated that the trends of the probability of recovering all paths correctly are consistent with the lower bound of Eq. (47). When the SNR is high, the lower bound in Eq. (47) is tight.

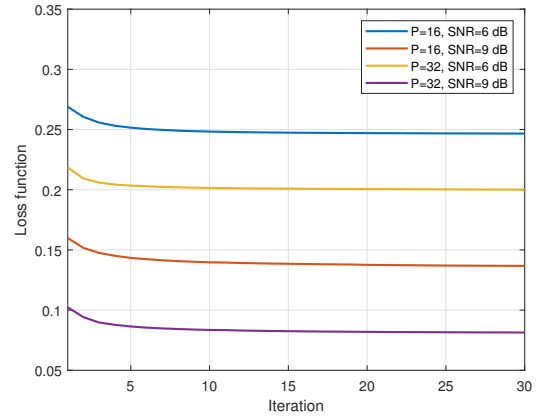


Fig. 5. Objective function with respect to the number of iterations.

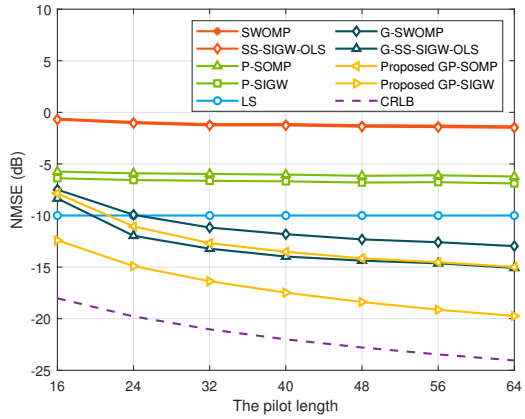
In Fig. 5, in order to verify the convergence of the proposed GP-SIGW algorithm, we test the objective function with respect to the number of iterations under different system configurations. It is demonstrated that under different system parameters, the objective functions all decrease monotonically over iteration, which verifies the convergence analysis in Section IV. In addition, Fig. 5 helps us determine how many iterations are needed to achieve accurate channel estimation. In our simulation, the iteration number is set to 20.

Fig. 6 demonstrates the NMSE performance against the pilot length P. The NMSE performance of the proposed on-grid GP-SOMP algorithm and the off-grid GP-SIGW algorithm are compared with that of the existing schemes, including the on-grid SWOMP algorithm for far-field in [9], the off-grid SS-SIGW-OLS algorithm for far-field in [15], the on-grid P-SOMP algorithm for near-field and the off-grid P-SIGW algorithm for near-field in [17], and the least square algorithm. In addition, the SWOMP algorithm and the SS-SIGW-OLS algorithm are combine with the proposed GTBC based signal extraction scheme for comparison, which are named as G-SWOMP algorithm and G-SS-SIGW-OLS algorithm. The CRLB derived above is also provided to verify the effectiveness of the proposed algorithms. The SNR of the system is set as 10 dB. The number of sub-arrays N_s is set as 4. In Fig. 6(a), the distances between the users/scatterers and the BS satisfy $\mathcal{U}\left(5\mathrm{m}, 10\mathrm{m}\right)$. In Fig. 6(b), the distances between the users/scatterers and the BS satisfy $\mathcal{U}\left(400\mathrm{m}, 450\mathrm{m}\right)$, which is around the Rayleigh distance. The pilot length is increasing from 16 to 64 [4]. Accordingly, the compressive ratio $\frac{PN_{\mathrm{RF}}}{N}$ increases from $\frac{1}{8}$ to $\frac{1}{2}$. Fig. 6 reveals that the NMSE performance of all the above schemes improves with

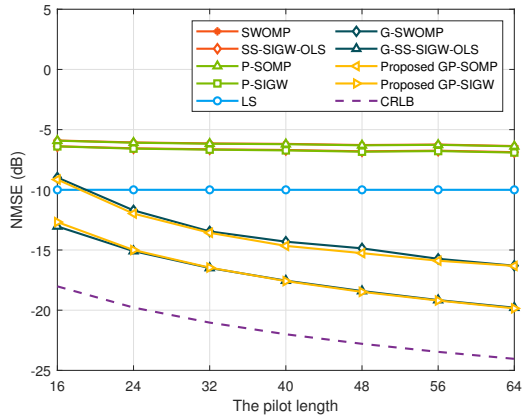
[4]Here, the pilot length seems large compared with the number of users. This is because we consider the hybrid precoding structure, where the number of radio frequency chains N_{RF} is much smaller than that of antennas N.

Authorized licensed use limited to: Tsinghua University. Downloaded on March 16,2024 at 02:20:29 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2023.3343740

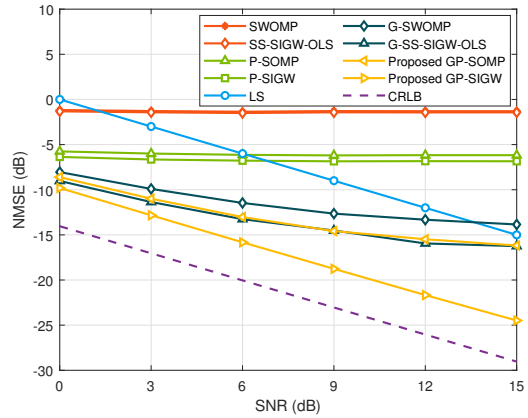


(a)

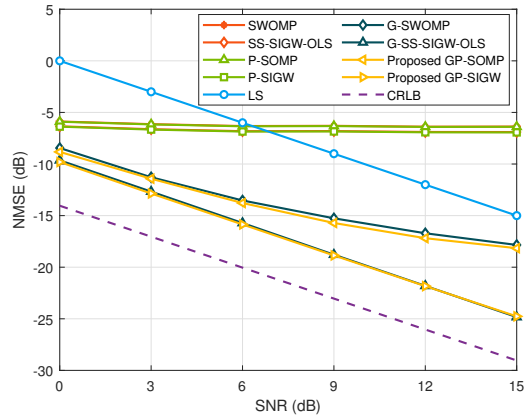


(b)

Fig. 6. The NMSE performance v.s. pilot length. The distances between the BS and the users or the scatterers are randomly chosen from (a) $\mathcal{U}\left(5\mathrm{m}, 10\mathrm{m}\right)$, (b) $\mathcal{U}\left(400\mathrm{m}, 450\mathrm{m}\right)$.



(a)



(b)

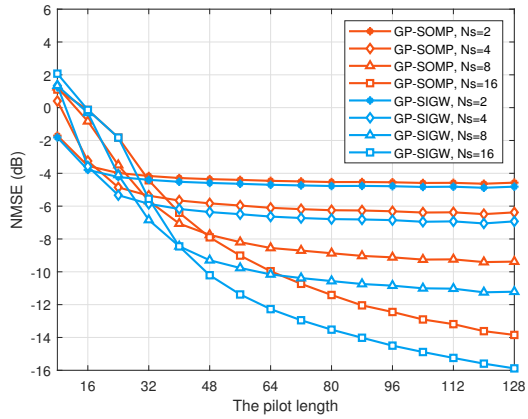
Fig. 7. The NMSE performance vs. the SNR. The distances between the BS and the users or the scatters are randomly chosen from (a) $\mathcal{U}\left(5\mathrm{m}, 10\mathrm{m}\right)$, (b) $\mathcal{U}\left(400\mathrm{m}, 450\mathrm{m}\right)$.

the increment of the pilot length. In Fig. 6(a), users are located in the near-field, the proposed GP-SOMP algorithm and GP-SIGW algorithm outperform other algorithm assuming spatial stationary. When the pilot length is 64, for example, the gap between the proposed GP-SIGW algorithm and the existing algorithms achieves nearly 15 dB, which indicates that the proposed algorithm can capture the feature of the spatial non-stationary and accurately recover the channel. The G-SWOMP algorithm and the G-SS-SIGW-OLS algorithm can obtain better performance compared to the stationary algorithm, but since the near-field propagation is neglected, the proposed algorithms also outperform them. In addition, in Fig. 6(b), the users are located in the far-field, the NMSE performances of the angular-domain algorithms and the polar-domain algorithms are nearly the same. However, since they have not considered the spatial non-stationary effect, the four spatial non-stationary algorithms still outperform existing algorithm. For both cases, due to the existence of spatial non-stationary effect, even the simple LS scheme can outperform all existing schemes. Fig. 6 demonstrates that the proposed GP-SOMP and GP-SIGW algorithms can catch the spatial non-stationary effect and accurately recover the non-stationary channel with a low pilot overhead for both the near-field and far-field scenarios. Furthermore, the proposed algorithms can approach the CRLB, which helps verify the effectiveness of the proposd algorithm.

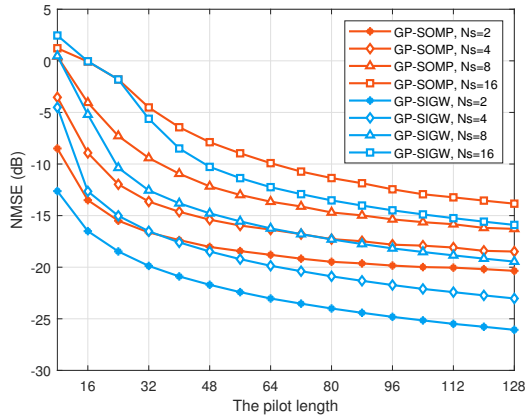
Then, Fig. 7 compares the NMSE performance against SNR, where the pilot length is 64, and the compressive ratio is $\frac{PN_{\mathrm{RF}}}{N} = \frac{1}{2}$. The number of sub-arrays N_s is set as 4. In Fig. 7(a), the distance between the users/scatterers and the BS satisfy $\mathcal{U}\left(5\mathrm{m}, 10\mathrm{m}\right)$, which is corresponding to the near-field scenario. In Fig. 7(b), the distances between the users/scatterers and the BS satisfy $\mathcal{U}\left(400\mathrm{m}, 450\mathrm{m}\right)$, which is corresponding to the far-field scenario. It is shown in Fig. 7(a) that the proposed algorithms outperform existing spatial stationary algorithms at all considered SNR. For far-field scenario in Fig. 7(b), the NMSE performances of the angular domain-algorithms and the polar-domain algorithms are nearly the same, and the NMSE performances of the G-SWOMP algorithm and the G-SS-SIGW-OLS algorithm are similar to the proposed algorithms since the users/scatterers are in the far field. The spatial non-stationary algorithms still outperform existing algorithms. For both cases, the proposed algorithms can approach the CRLB, which reveals the effec-

Authorized licensed use limited to: Tsinghua University. Downloaded on March 16,2024 at 02:20:29 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2023.3343740



(a)



(b)

Fig. 8. The NMSE performance comparison among different sub-array numbers against pilot length when the channel is (a) non-staionary, (b) stationary.

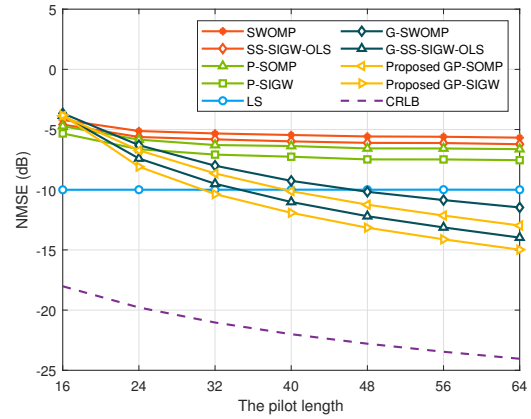


Fig. 9. The NMSE performance v.s. pilot length with $f_c = 3$ GHz.

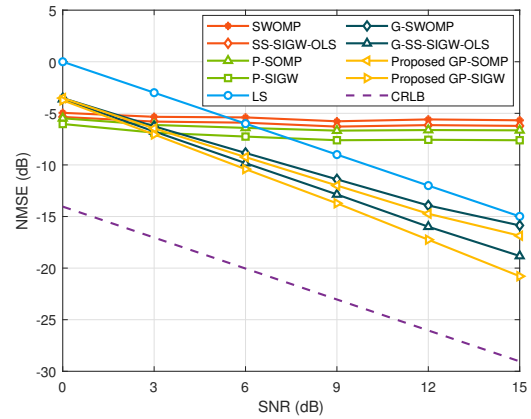


Fig. 10. The NMSE performance v.s. the SNR with $f_c = 3$ GHz.

tiveness of the proposd algorithm.

In addition, we evaluate the impact of the number of sub-arrays to the NMSE performance, as illustrated in Fig. 8. The SNR is set as 10 dB. The distances between the users/scatterers and the BS satisfy $\mathcal{U}(5\text{m}, 400\text{m})$. The pilot length is increasing from 8 to 128. Accordingly, the compressive ratio $\frac{PN_{\text{RF}}}{N}$ increases from $\frac{1}{16}$ to 1. The number of sub-arrays N_s is changed from 2 to 16. In Fig. 8(a), the minimum dimensional of a spatial stationary sub-channel is set to $\frac{N}{16} = 32$. When the pilot length is sufficient, i.e., $P = 128$, the NMSE performance of $N_s = 16$ is the best, because it can best capture the features of spatial non-stationarity. However, when the pilot length is not sufficient, i.e., $P \leq 40$, the NMSE performance of $N_s = 16$ is not the best. This is because the number of effective pilots for each sub-array is inversely propotional to N_s. When the pilot length is small, the effective pilots are not sufficient. On the other hand, in Fig. 8(b) the channel is set as spatial stationary. In this case, the larger N_s is, the lower SNR each sub-array has. Accordingly, the channel estimation accuracy decreases. Therefore, the number of sub-arrays N_s should be neither too small nor too large. If it is too small, the spatial non-stationary effect in the system cannot be recognized properly, which leads to a poor channel estimation accuracy. If it is too large, the effective pilot length is too small. In addition, the SNR of each sub-array also decreases a lot, which also leads to the decrease in the channel estimation accuracy. In practical communication systems, the number of sub-arrays should be selected based on the characteristic of the channel, which can be obtained based on the statistical CSI over a long period of time.

In order to assess the scalability of our proposed scheme, we also compare the proposed scheme with existing schemes with a sub-6 GHz setup. Specifically, we consider a system with central frequency $f_c = 3$ GHz, bandwidth $B = 40$ MHz, BS antenna number $N = 512$, subcarrier number $M = 256$, user number $K = 4$, and RF chain number $N_{\text{RF}} = 4$. Due to the lower frequency, the path number L is set to 10, greater than the path number under 100 GHz. In the considered system, the antenna array size is about 25 m. In Fig. 9, the NMSE performance against the pilot length P is illustrated. The SNR of the system is set as 10 dB. Since the number of paths is larger than that in systems with a higher frequency, the necessary pilot length increases accordingly, so the performance gap between the proposed schemes and the CRLB is larger compared to the high-frequency scenario. However, when the pilot length is larger

Authorized licensed use limited to: Tsinghua University. Downloaded on March 16,2024 at 02:20:29 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2023.3343740

than 32, the proposed GP-SIGW algorithm still outperforms existing schemes. When the pilot length is larger than 40, the proposed GP-SOMP algorithm outperforms existing schemes. Thanks to the ability to recognize the spatial non-stationarity in systems, the proposed schemes can still realize more accurate channel estimation compared to existing schemes when the pilot length is sufficient in sub-6 GHz systems.

Fig. 10 compares the NMSE performance against SNR. The pilot length is set as 64, and the compressive ratio is $\frac{1}{2}$. Similarly, the performance gap between the proposed schemes and the CRLB is larger compared to the high-frequency scenario. Nevertheless, it is illustrated that the proposed algorithms can still outperform existing schemes in sub-6 GHz systems when the SNR is larger than 3 dB, which verifies the scalability of the proposed algorithms to sub-6 GHz systems.

In conclusion, the proposed algorithms can estimate the non-stationary channel accurately in all considered scenarios compared to existing algorithms. The reason for the gap between the NMSE of proposed algorithms and CRLB is that the number of antenna elements on each sub-array is small, leading to a relatively coarse codebook. To further improve the channel estimation accuracy, finer codebook can be designed in future works.

VI. Conclusion

In this paper, we have investigated the spatial non-stationary channel estimation problem in XL-MIMO systems with hybrid precoding. The problem of estimating the channel of the entire antenna array is firstly converted into several sub-problems of estimating the sub-channels of each sub-array. Since the sub-channels with respect to each sub-array can be regarded as spatially stationary, traditional CS-based channel estimation algorithms can be utilized to recover the sub-channels. In order to decouple the received signal of each sub-array, a GTBC based signal extraction scheme was proposed, which artificially creates the relevance in the time domain to enable recognition of the spatial non-stationarity in the space domain. This scheme consists of the encoding stage and the decoding stage. At the encoding stage, the combining matrix of a sub-array was changed as a whole according to the designed GTBC. At the decoding stage, the original received signal was combined based on the GTBC to extract the received signal for further process. Based on the above signal extraction scheme, an on-grid algorithm called GP-SOMP and an off-grid algorithm called GP-SIGW were proposed to estimate the spatial non-stationary XL-MIMO channels. Analyses of the complexity and performance of the proposed two algorithms were also conducted and the CRLB is derived. Simulation results revealed that the proposed algorithms can recognize the spatially non-stationary effect and achieve much better NMSE performance in both the far-field scenarios and the near-field scenarios. For future works, the proposed schemes can be extended to the reconfigurable intelligent surface (RIS)-assisted communications.

References

[1] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, May 2013.

[2] T. S. Rappaport, Y. Xing, O. Kanhere, S. Ju, A. Madanayake, S. Mandal, A. Alkhateeb, and G. C. Trichopoulos, "Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond," *IEEE Access*, vol. 7, pp. 78 729–78 757, Jun. 2019.

[3] H. Elayan, O. Amin, B. Shihada, R. M. Shubair, and M.-S. Alouini, "Terahertz band: The last piece of RF spectrum puzzle for communication systems," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 1–32, Nov. 2019.

[4] Q. Yuan, Z.-C. Hao, K.-K. Fan, and G. Q. Luo, "A compact W-band substrate-integrated cavity array antenna using high-order resonating modes," *IEEE Trans. Antennas Propag.*, vol. 66, no. 12, pp. 7400–7405, Dec. 2018.

[5] Y.-W. Wu, Z.-C. Hao, and Z.-W. Miao, "A planar W-band large-scale high-gain substrate-integrated waveguide slot array," *IEEE Trans. Antennas Propag.*, vol. 68, no. 8, pp. 6429–6434, Aug. 2020.

[6] B. Ning, Z. Tian, Z. Chen, C. Han, J. Yuan, and S. Li, "Prospective beamforming technologies for ultra-massive MIMO in terahertz communications: A tutorial," *arXiv preprint arXiv:2107.03032*, 2021.

[7] L. Dai, B. Wang, M. Peng, and S. Chen, "Hybrid precoding-based millimeter-wave massive MIMO-NOMA with simultaneous wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 1, pp. 131–141, Jan. 2018.

[8] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Tran. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Jan. 2014.

[9] J. Rodríguez-Fernández, N. González-Prelcic, K. Venugopal, and R. W. Heath, "Frequency-domain compressive channel estimation for frequency-selective hybrid millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2946–2960, Mar. 2018.

[10] J. Lee, G.-T. Gil, and Y. H. Lee, "Channel estimation via orthogonal matching pursuit for hybrid MIMO systems in millimeter wave communications," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2370–2386, Apr. 2016.

[11] Z. Gao, C. Hu, L. Dai, and Z. Wang, "Channel estimation for millimeter-wave massive MIMO with hybrid precoding over frequency-selective fading channels," *IEEE Commun. Lett.*, vol. 20, no. 6, pp. 1259–1262, Apr. 2016.

[12] J. Rodríguez-Fernández, N. González-Prelcic, and R. W. Heath, "A compressive sensing-maximum likelihood approach for off-grid wideband channel estimation at mmWave," in *Proc. IEEE 7th Int. Workshop Comput. Adv. Multi-Sensor Adapt. Process. (CAMSAP)*, 2017, pp. 1–5.

[13] Z. Zhou, J. Fang, L. Yang, H. Li, Z. Chen, and R. S. Blum, "Low-rank tensor decomposition-aided channel estimation for millimeter wave MIMO-OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1524–1538, Apr. 2017.

[14] C. Hu, L. Dai, T. Mir, Z. Gao, and J. Fang, "Super-resolution channel estimation for mmWave massive MIMO with hybrid precoding," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8954–8958, Jun. 2018.

[15] N. González-Prelcic, H. Xie, J. Palacios, and T. Shimizu, "Wideband channel tracking and hybrid precoding for mmWave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 4, pp. 2161–2174, Oct. 2020.

[16] M. Cui, Z. Wu, Y. Lu, X. Wei, and L. Dai, "Near-field MIMO communications for 6g: Fundamentals, challenges, potentials, and future directions," *IEEE Commun. Mag.*, p. 41, Jan. 2023.

[17] M. Cui and L. Dai, "Channel estimation for extremely large-scale MIMO: Far-field or near-field?" *IEEE Trans. Commun.*, vol. 70, no. 4, pp. 2663–2677, Apr. 2022.

[18] X. Zhang, H. Zhang, and Y. C. Eldar, "Near-field sparse channel representation and estimation in 6G wireless communications," *IEEE Trans. Commun.*, 2023.

[19] S. Tarboush, A. Ali, and T. Y. Al-Naffouri, "Compressive estimation of near field channels for ultra massive-MIMO wideband THz systems," in *IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 2023, pp. 1–5.

[20] M. Cui and L. Dai, "Near-field wideband channel estimation for extremely large-scale MIMO," *Sci. China Inf. Sci.*, vol. 66, no. 7, p. 172303, Jun. 2023.

[21] A. M. Elbir, K. V. Mishra, and S. Chatzinotas, "NBA-OMP: Near-field beam-split-aware orthogonal matching pursuit for wideband THz channel estimation," in *IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 2023, pp. 1–5.

[22] W. Yang, M. Li, and Q. Liu, "A practical channel estimation strategy for XL-MIMO communication systems," *IEEE Commun. Lett.*, Apr. 2023.

Authorized licensed use limited to: Tsinghua University. Downloaded on March 16,2024 at 02:20:29 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2023.3343740

[23] Z. Hu, C. Chen, Y. Jin, L. Zhou, and Q. Wei, "Hybrid-field channel estimation for extremely large-scale massive MIMO system," *IEEE Commun. Lett.*, vol. 27, no. 1, pp. 303–307, Jan. 2022.

[24] K. T. Selvan and R. Janaswamy, "Fraunhofer and fresnel distances: Unified derivation for aperture antennas," *IEEE Antennas Propag. Mag.*, vol. 59, no. 4, pp. 12–15, Jun. 2017.

[25] A. Pizzo, L. Sanguinetti, and T. L. Marzetta, "Fourier plane-wave series expansion for holographic MIMO communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 6890–6905, Mar. 2022.

[26] E. De Carvalho, A. Ali, A. Amiri, M. Angjelichinoski, and R. W. Heath, "Non-stationarities in extra-large-scale massive MIMO," *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 74–80, Aug. 2020.

[27] Y. Han, S. Jin, C.-K. Wen, and X. Ma, "Channel estimation for extremely large-scale massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 633–637, Jan. 2020.

[28] L. Qiao, A. Liao, Z. Gao, and H. Wang, "Indoor massive IoT access relying on millimeter-wave extra-large-scale MIMO," in *IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2023, pp. 1–6.

[29] Y. Han, M. Li, S. Jin, C.-K. Wen, and X. Ma, "Deep learning-based FDD non-stationary massive MIMO downlink channel reconstruction," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 9, pp. 1980–1993, Jun. 2020.

[30] H. Iimori, T. Takahashi, K. Ishibashi, G. T. F. de Abreu, D. González, and O. Gonsa, "Joint activity and channel estimation for extra-large MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7253–7270, Sep. 2022.

[31] H. Wang, A. Kosasih, C.-K. Wen, S. Jin, and W. Hardjawana, "Expectation propagation detector for extra-large scale massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 2036–2051, Mar. 2020.

[32] X. Cheng, K. Xu, J. Sun, and S. Li, "Adaptive grouping sparse bayesian learning for channel estimation in non-stationary uplink massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4184–4198, Jun. 2019.

[33] Y. Zhu, H. Guo, and V. K. N. Lau, "Bayesian channel estimation in multi-user massive MIMO with extremely large antenna array," *IEEE Trans. Signal Process.*, vol. 69, pp. 5463–5478, Sep. 2021.

[34] Z. Yuan, J. Zhang, Y. Ji, G. F. Pedersen, and W. Fan, "Spatial non-stationary near-field channel modeling and validation for massive MIMO systems," *IEEE Trans. Antennas Propag.*, vol. 71, no. 1, pp. 921–933, Jan. 2023.

[35] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[36] C. Huang, L. Liu, C. Yuen, and S. Sun, "Iterative channel estimation using LSE and sparse message passing for mmWave MIMO systems," *IEEE Trans. Signal Process.*, vol. 67, no. 1, pp. 245–259, Nov. 2019.

[37] X. Gao, L. Dai, S. Zhou, A. M. Sayeed, and L. Hanzo, "Wideband beamspace channel estimation for millimeter-wave MIMO systems relying on lens antenna arrays," *IEEE Trans. Signal Process.*, vol. 67, no. 18, pp. 4809–4824, Jul. 2019.

[38] X. Wei, C. Hu, and L. Dai, "Deep learning for beamspace channel estimation in millimeter-wave massive MIMO systems," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 182–193, Sep. 2021.

[39] J. Sherman, "Properties of focused apertures in the fresnel region," *IRE Trans. Antennas Propag.*, vol. 10, no. 4, pp. 399–408, Jul. 1962.

[40] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[41] Z. Ben-Haim, Y. C. Eldar, and M. Elad, "Coherence-based performance guarantees for estimating a sparse vector under random noise," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5030–5043, Jun. 2010.

Yuhao Chen (Graduate Student Member, IEEE) received the B.E. degree in electronic engineering from Tsinghua University, Beijing, China, in 2021, where he is currently pursuing the M.S. degree in electronic engineering. His research interests include massive MIMO, reconfigurable intelligent surface (RIS), and near-field communcations. He received the IEEE ICC Outstanding Demo Award in 2022 and the National Scholarship in 2023.

Linglong Dai (Fellow, IEEE) received the B.S. degree from Zhejiang University, Hangzhou, China, in 2003, the M.S. degree (with the highest honor) from the China Academy of Telecommunications Technology, Beijing, China, in 2006, and the Ph.D. degree (with the highest honor) from Tsinghua University, Beijing, China, in 2011. From 2011 to 2013, he was a Postdoctoral Research Fellow with the Department of Electronic Engineering, Tsinghua University, where he was an Assistant Professor from 2013 to 2016, an Associate Professor from 2016 to 2022, and has been a Professor since 2022. His current research interests include massive MIMO, reconfigurable intelligent surface (RIS), millimeter-wave and Terahertz communications, near-field communications, machine learning for wireless communications, and electromagnetic information theory.

He has coauthored the book *MmWave Massive MIMO: A Paradigm for 5G* (Academic Press, 2016). He has authored or coauthored over 100 IEEE journal papers and over 60 IEEE conference papers. He also holds over 20 granted patents. He has received five IEEE Best Paper Awards at the IEEE ICC 2013, the IEEE ICC 2014, the IEEE ICC 2017, the IEEE VTC 2017-Fall, the IEEE ICC 2018, and the IEEE GLOBECOM 2023. He has also received the Tsinghua University Outstanding Ph.D. Graduate Award in 2011, the Beijing Excellent Doctoral Dissertation Award in 2012, the China National Excellent Doctoral Dissertation Nomination Award in 2013, the URSI Young Scientist Award in 2014, the IEEE Transactions on Broadcasting Best Paper Award in 2015, the Electronics Letters Best Paper Award in 2016, the National Natural Science Foundation of China for Outstanding Young Scholars in 2017, the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2017, the IEEE ComSoc Asia-Pacific Outstanding Paper Award in 2018, the China Communications Best Paper Award in 2019, the IEEE Access Best Multimedia Award in 2020, the IEEE Communications Society Leonard G. Abraham Prize in 2020, the IEEE ComSoc Stephen O. Rice Prize in 2022, the IEEE ICC Outstanding Demo Award in 2022, and the National Science Foundation for Distinguished Young Scholars in 2023. He was listed as a Highly Cited Researcher by Clarivate Analytics from 2020 to 2023. He was elevated as an IEEE Fellow in 2022.

Authorized licensed use limited to: Tsinghua University. Downloaded on March 16,2024 at 02:20:29 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.