This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/TCOMM.2024.3362146, IEEE Transactions on Communications

Robust Precoding for HF Skywave Massive MIMO with Slepian Transform

Linfeng Song, *Graduate Student Member, IEEE*, Ding Shi, *Graduate Student Member, IEEE*, Lu Gan, *Senior Member, IEEE*, and Xiqi Gao, *Fellow, IEEE*

***Abstract*—In this paper, we address robust precoding in high-frequency (HF) skywave massive multiple-input multiple-output (MIMO) systems with imperfect channel state information (CSI). We first employ a sparse beam based *a posteriori* channel model and demonstrate that robust precoding can be efficiently solved in the Slepian transform domain with a large number of base station (BS) antennas. Next, we introduce two Slepian transform based robust precoding methods, including a joint approach that leverages inverse fast Fourier transform (IFFT) for reduced complexity with a large number of user terminals (UTs). We then establish a local optimum for the Slepian transform domain robust precoder (STRP) design using the majorization minimization (MM) algorithm, taking advantages of HF skywave massive MIMO channel sparsity and Slepian sequence properties. Further, two distinct designs are presented: separate STRP (SSTRP) and joint STRP (JSTRP). Simulation results confirm the effectiveness of proposed robust precoders, showcasing their excellent ergodic sum-rate performance and low complexity.**

***Index Terms*—Massive MIMO, HF skywave communications, robust precoding, Slepian transform.**

I. Introduction

THE future of wireless networks aims to offer high-data-rate global services, even in remote areas [1]. High frequency (HF) skywave communications, operating in the 3 to 30 MHz range, facilitate beyond-line-of-sight communications through ionospheric refraction [2]. Traditional HF skywave systems, however, are limited by lower data rates in single-input single-output configurations. The adoption of massive multiple-input multiple-output (MIMO) technology, a cornerstone of 5G systems, significantly boosts sum-rate and spectrum efficiency by employing numerous antennas at the base station (BS) to serve multiple UTs [3]–[5]. Integrating massive MIMO into HF skywave communications, as shown in [6], markedly improves spectral and energy efficiency in these systems.

In massive MIMO downlink (DL) transmissions, precoder design is crucial for boosting sum-rate performance. In environments where the ionosphere changes slowly and UTs move at lower speeds, HF skywave massive MIMO channels are quasi-static [7]. This enables leveraging instantaneous DL channel state information (CSI) at the base station (BS) in time-division duplex (TDD) mode. Consequently, precoders such as signal-to-leakage noise ratio (SLNR) [8], minimum mean-squared error (MMSE) [6], and weighted MMSE (WMMSE) [9] are applicable. However, under rapidly changing HF skywave channels, due to a disturbed ionosphere [2] and high UT mobility, perfect instantaneous DL CSI at the BS becomes unattainable. In these scenarios, the channel varies symbol by symbol, impacting precoding in data segments. Precoders designed for slowly-varying statistical CSI or imperfect DL CSI are then utilized [10]–[13]. Despite this, precoders using statistical CSI in quasi-static scenarios [10] show limited efficacy, and the independent and identically distributed (i.i.d.) assumption for channels is inaccurate in spatially correlated HF skywave massive MIMO contexts [11]–[13].

In massive MIMO systems, it's essential to tackle challenges such as channel estimation errors, channel aging, and spatial correlation. A robust precoding technique, based on the *a posteriori* channel model and utilizing the majorization-minimization (MM) algorithm, has been established to effectively reduce multi-user interference and improve sum-rate performance [14]. Advanced strategies have also emerged, including robust precoding through matrix manifold optimization [15] and deep learning based approaches [16]. However, these techniques become complex with an increasing number of base station (BS) antennas, primarily due to their operation in the spatial domain. To counter this complexity, the beam domain robust precoder (BDRP) for HF skywave massive MIMO has been introduced [17], significantly reducing complexity by utilizing the sparsity of beam domain channels. Nonetheless, the non-orthogonality of the beam matrix in BDRP poses a challenge to sum-rate performance.

This paper introduces Slepian transform domain robust precoders for HF skywave massive MIMO systems, addressing the limitations of the BDRP approach. Utilizing the Slepian basis, known for its orthogonality and exceptional spectral concentration [18]–[24], we make several key contributions:

- We develop a sparse beam based *a posteriori* channel model to describe imperfect CSI at the BS, highlighting the crucial relationship between beam domain channels and the Fourier spectrum of spatial domain channels for robust precoder design.
- We formulate the robust precoding problem to maximize

This work was supported by the National Key R&D Program of China under Grant 2018YFB1801103, the Jiangsu Province Basic Research Project under Grant BK20192002, the Fundamental Research Funds for the Central Universities under Grant 2242022k60007, the Key R&D Plan of Jiangsu Province under Grant BE2022067, and the Huawei Cooperation Project. (*Linfeng Song and Ding Shi contributed equally to this work.*) (*Corresponding author: Xiqi Gao.*)

L. F. Song, D. Shi, and X. Q. Gao are with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China, and also with Purple Mountain Laboratories, Nanjing 211111, China (e-mail: songlf@seu.edu.cn; shiding@seu.edu.cn; xqgao@seu.edu.cn).

L. Gan is with the Department of Electronic and Electrical Engineering, Brunel University London, London UB8 3PH, UK. (e-mail: lu.gan@brunel.ac.uk).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/TCOMM.2024.3362146, IEEE Transactions on Communications

the ergodic sum-rate under a total transmit power constraint and transform it into the Slepian transform domain. We demonstrate that Slepian transform domain robust precoding is asymptotically optimal for a sufficiently large number of BS antennas. We propose two precoding approaches based on the Slepian transform and introduce an efficient joint precoding implementation that utilizes a low-dimensional inverse fast Fourier transform (IFFT). The implementation complexities of these algorithms are also thoroughly analyzed.

- To address the computational intensity of ergodic sum rate, we focus on maximizing an ergodic sum-rate upper bound. A local optimum for the Slepian transform domain robust precoder (STRP) design is derived using the MM algorithm. By exploiting the properties of Slepian sequences, we simplify the STRP design into separate (SSTRP) and joint (JSTRP) versions. The computational complexity analysis and simulation results demonstrate that our robust precoder designs maintain low complexity with near-optimal performance.

The remainder of this paper is structured as follows: Section II introduces the channel model for HF skywave massive MIMO systems. Section III investigates robust precoding with Slepian transform, while Section IV details the design of Slepian transform domain robust precoders and their complexity analysis. Simulation results are presented in Section V, and the paper concludes in Section VI.

Notations: Boldface lower case letters and boldface upper case letters represent vectors and matrices, respectively. The operators $(\cdot)^*$, $(\cdot)^{\mathrm{T}}$, and $(\cdot)^{\mathrm{H}}$ denote the conjugate, transpose and conjugate-transpose, respectively. $\mathbf{I}_M$ and $\mathbf{O}_{M\times N}$ represent M-dimensional identity matrix and $M \times N$ matrix with all zeros, respectively. $\mathbf{e}_{M,m}$, $\mathbf{1}_M$ and $\mathbf{0}_M$ indicate the m-th column of $\mathbf{I}_M$, M-dimensional all-one vector and all-zero vector, respectively. $\mathrm{diag}\{\mathbf{a}\}$ and $[\mathbf{A}]_{i,j}$ denote the diagonal matrix with $\mathbf{a}$ along its main diagonal and the (i,j)-th entry of $\mathbf{A}$, respectively. $\otimes$ and $\odot$ represent the Kronecker product operator and the Hadamard product operator, respectively. $\bar{\jmath} = \sqrt{-1}$ denotes the imaginary unit. $|\mathcal{A}|$ denotes the cardinality for set $\mathcal{A}$. $\mathbb{E}\{\cdot\}$, $\langle\cdot\rangle_N$ and $\|\cdot\|$ denote the ensemble expectation, modulo-N operation and the Frobenius norm, respectively. $\mathcal{Z}^+$ represents the set of positive integers and $\mathcal{Z}_N^+ \triangleq \{1, \cdots, N\}$. $\lfloor a \rfloor$ ($\lceil a \rceil$) indicates the maximum (minimum) integer that is not greater (less) than a. $\mathcal{CN}\{\mathbf{a}, \mathbf{A}\}$ denotes the circular symmetric complex Gaussian distribution with mean $\mathbf{a}$ and convariance $\mathbf{A}$.

II. System Model

In this section, we present the system configuration for HF skywave massive MIMO orthogonal frequency division multiplexing (OFDM) systems. We introduce the beam based *a priori* and *a posteriori* channel models. Additionally, we establish the relationship between the sparsity of the beam domain channel and the Fourier spectrum of the spatial domain channel.

A. System Configuration

We examine an HF skywave massive MIMO-OFDM system operating in TDD mode, where a uniform linear array (ULA) with M antennas is deployed at the base station (BS) to serve U single-antenna UTs, as depicted in Fig. 1. Notably, ULA is a practical choice in the HF band, as implementing multiple antennas in the elevation direction is challenging due to the substantial wavelength. Additionally, it's worth mentioning while the antenna array aperture is typically large, it remains relatively small compared to the overall propagation distance.

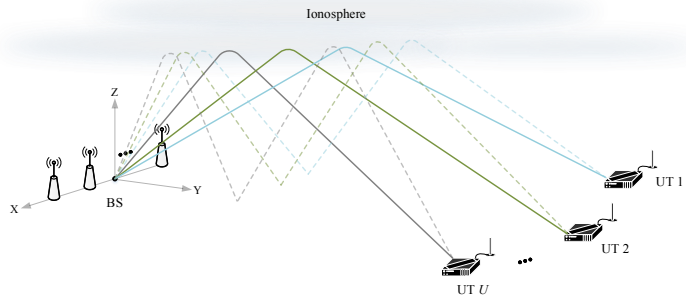

Fig. 1. HF skywave massive MIMO-OFDM system.

The frame structure of the HF skywave massive MIMO-OFDM system is visualized in Fig. 2. Each frame contains N OFDM symbols that are further divided into $\widetilde{N}$ uplink (UL) data symbols, along with one training symbol, and $N-\widetilde{N}-1$ DL data symbols. Moreover, the number of subcarriers, the length of the cyclic prefix (CP), the number of valid subcarriers and subcarrier spacing are denoted by N_{c}, N_{g}, N_{v} and Δ_f, respectively. We define f_{c} as the carrier frequency and f_{o} as the highest system operating frequency since f_{c} varies with different ionosphere condition [2], [25]. The inter-antenna spacing is set as $d = \lambda_{\mathrm{o}}/2$ instead of half wavelength of carrier in traditional terrestrial massive MIMO systems, where $\lambda_{\mathrm{o}} = c/f_{\mathrm{o}}$ is the wavelength corresponding to f_{o} and c is the speed of light.

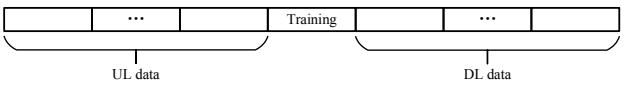

Fig. 2. Frame structure of the HF skywave massive MIMO-OFDM system.

B. Beam Based A Priori Channel Model

In DL transmission, the signal transmitted for UT u at subcarrier k of symbol n is denoted as $\overline{\mathbf{x}}_{u,n,k} \in \mathbb{C}^M$. The corresponding received signal takes the following form [6]

$$y_{u,n,k} = \mathbf{h}_{u,n,k}^{\mathrm{H}} \sum_{u'=1}^{U} \overline{\mathbf{x}}_{u',n,k} + z_{u,n,k}, \tag{1}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/TCOMM.2024.3362146, IEEE Transactions on Communications

where $z_{u,n,k} \sim \mathcal{CN}(0, \sigma_z^2)$ represents the additive white Gaussian noise, and the channel frequency response $\mathbf{h}_{u,n,k}$ is given by

$$\mathbf{h}_{u,n,k} = \sum_{p=1}^{P_u} \eta_{u,p,n,k} \mathbf{v}_k(\Omega_{u,p}) \in \mathbb{C}^M, \tag{2}$$

where $\eta_{u,p,n,k}$ defines the path characteristics, and $\mathbf{v}_k(\Omega)$ is a vector function of the direction, as expressed in the following equations:

$$\eta_{u,p,n,k} \triangleq \beta_{u,p} e^{\bar{\jmath}\phi_{u,p}} e^{\bar{\jmath}2\pi f_{u,p}^{\mathrm{d}} n T_{\mathrm{sym}}} e^{-\bar{\jmath}2\pi k \Delta_f \tau_{u,p}}, \tag{3}$$

$$\mathbf{v}_k(\Omega) \triangleq \left[1, e^{-\bar{\jmath}2\pi(f_c+k\Delta_f)\Delta_\tau 1\Omega}, \cdots, e^{-\bar{\jmath}2\pi(f_c+k\Delta_f)\Delta_\tau(M-1)\Omega}\right]^{\mathrm{T}}. \tag{4}$$

Besides, in (2), P_u denotes the number of paths for UT u, and $\Delta_\tau = d/c$. Parameters $\beta_{u,p}$, $\phi_{u,p}$, $f_{u,p}^{\mathrm{d}}$, $\tau_{u,p}$, and $\Omega_{u,p}$ represent the path gain, initial phase, Doppler frequency, propagation delay, and direction cosine of the p-th path, respectively.

We derive a beam based channel model by uniformly sampling the direction cosine range $[-1, 1)$ into segments $\mathcal{S}_{\widetilde{m}}, \widetilde{m} \in \mathcal{Z}_{\widetilde{M}}^+$, with $\widetilde{M} = \lceil FMf_c/f_o \rceil$ representing the number of samples, determined by the fine factor F. Each segment is defined as $\mathcal{S}_{\widetilde{m}} = [\Omega_{\widetilde{m}}, \Omega_{\widetilde{m}} + 2/\widetilde{M})$, where $\Omega_{\widetilde{m}} = (2\widetilde{m} - 2 - \widetilde{M})/\widetilde{M}$. For large values of $\widetilde{M}$, the channel frequency response $\mathbf{h}_{u,n,k}$ is approximated as

$$\mathbf{h}_{u,n,k} = \sum_{\widetilde{m}=1}^{\widetilde{M}} \sum_{\Omega_{u,p} \in \mathcal{P}_u \cap \mathcal{S}_{\widetilde{m}}} \eta_{u,p,n,k} \mathbf{v}_k(\Omega_{\widetilde{m}}) = \mathbf{V}_k \widetilde{\mathbf{h}}_{u,n,k}, \tag{5}$$

where $\mathcal{P}_u \triangleq \{\Omega_{u,1}, \cdots, \Omega_{u,P_u}\}$ represents the set of direction cosines for all paths to UT u. The beam matrix at subcarrier k is given by $\mathbf{V}_k = \frac{1}{\sqrt{M}}[\mathbf{v}_k(\Omega_1), \cdots, \mathbf{v}_k(\Omega_{\widetilde{M}})] \in \mathbb{C}^{M \times \widetilde{M}}$. The beam domain channel vector $\widetilde{\mathbf{h}}_{u,n,k} \in \mathbb{C}^{\widetilde{M}}$, where each element $[\widetilde{\mathbf{h}}_{u,n,k}]_{\widetilde{m}} = \sqrt{M} \sum_{\Omega_{u,p} \in \mathcal{P}_u \cap \mathcal{S}_{\widetilde{m}}} \eta_{u,p,n,k}, \widetilde{m} \in \mathcal{Z}_{\widetilde{M}}^+$. It should be noted that $\mathbf{V}_k$ is correlated with subcarriers due to the spatial-wideband effect, stemming from the propagation delay across the large-scale antenna array [26]. This correlation leads to (5), forming a beam based *a priori* channel model. It is worth mentioning that the traditional massive MIMO channel model corresponds to specific conditions, such as $f_o = f_c$, $F = 1$, and the absence of the spatial-wideband effect. The beam domain channel power vector for UT u, defined as

$$\begin{aligned}\widetilde{\boldsymbol{\omega}}_u &\triangleq \mathbb{E}\{\widetilde{\mathbf{h}}_{u,n,k} \odot \widetilde{\mathbf{h}}_{u,n,k}^*\} \\ &= M\left[\textstyle\sum_{\Omega_{u,p} \in \mathcal{P}_u \cap \mathcal{S}_1} \beta_{u,p}^2, \cdots, \sum_{\Omega_{u,p} \in \mathcal{P}_u \cap \mathcal{S}_{\widetilde{M}}} \beta_{u,p}^2\right]^{\mathrm{T}},\end{aligned} \tag{6}$$

serving as the statistical CSI, reflecting the aggregated path gains within each segment.

C. Beam Based A Posteriori Channel Model

Consider $\widehat{\mathbf{h}}_{u,k,\widetilde{N}+1}$ as the estimated beam domain channel for UT u at subcarrier k on the training symbol $\widetilde{N}+1$. It is assumed that accurate CSI can be estimated [27], [28]. To accommodate channel aging, the channel at subcarrier k of symbol n is modeled using a first-order Gauss-Markov process [29], [30]:

$$\mathbf{h}_{u,n,k} = \mathbf{V}_k \left(\alpha_{u,n} \widehat{\mathbf{h}}_{u,\widetilde{N}+1,k} + \sqrt{1-\alpha_{u,n}^2} \widetilde{\mathbf{h}}_{u,n,k}\right), \tag{7}$$

where $\alpha_{u,n}$ is the temporal correlation coefficient between training symbol $\widetilde{N}+1$ and symbol n, reflecting the channel Doppler spread and encapsulates channel uncertainties. This equation represents an *a posteriori* channel model, considering the imperfections in CSI across various mobile scenarios. If $\alpha_{u,n}$ is close to 1, the channel is quasi-static, whereas a value nearing 0 indicates rapid channel changes.

Owing to the narrow angle spread from the BS to each UT [31], the beam domain channel shows remarkable sparsity, especially as M increases. This leads to $\mathbf{h}_{u,n,k}$ being characterized as a narrowband sequence with a continuous Fourier spectrum band-limit $\mathcal{L}_{u,k} = [\nu_{u,k}^{\mathrm{l}}, \nu_{u,k}^{\mathrm{r}}]$. The limits $\nu_{u,k}^{\mathrm{l}}$ and $\nu_{u,k}^{\mathrm{r}}$ can be approximated as follows [32]

$$\nu_{u,k}^{\mathrm{l}} = -(f_c + k\Delta_f)\Delta_\tau \left(\frac{2}{\widetilde{M}} \max\{\mathcal{B}_u\} - 1\right), \tag{8a}$$

$$\nu_{u,k}^{\mathrm{r}} = -(f_c + k\Delta_f)\Delta_\tau \left(\frac{2}{\widetilde{M}} \min\{\mathcal{B}_u\} - 1\right), \tag{8b}$$

respectively, where $\mathcal{B}_u$ is the index set of non-zero elements in the beam domain channel vector, and B_u indicates the count of non-zero beams.

III. Robust Precoding With Slepian Transform

In this section, our focus is on robust DL precoding for HF massive MIMO systems using the *a posteriori* channel model, based on Slepian transform. First, we outline the problem of robust precoding along with an introduction to Slepian sequences. We then establish that robust precoding can be seamlessly transformed into the Slepian transform domain while maintaining optimality. This holds true when the objective is to maximize the ergodic sum-rate or its upper bound. Finally, we introduce two variants of Slepian transform based precoding. In particular, the joint Slepian transform based approach enables efficient implementation through IFFT, with a detailed discussion on its implementation complexity.

A. Robust Precoding

We consider DL precoding for HF massive MIMO systems at subcarrier k of symbol n, using the *a posteriori* channel model. For simplicity, we drop the indices n and k henceforth. In the context of linear precoding, the signal model simplifies to

$$y_u = \mathbf{h}_u^{\mathrm{H}} \overline{\mathbf{x}} + z_u = \mathbf{h}_u^{\mathrm{H}} \sum_{u'=1}^{U} \mathbf{p}_{u'} x_{u'} + z_u = \mathbf{h}_u^{\mathrm{H}} \mathbf{p}_u x_u + n_u, \tag{9}$$

where $\mathbf{p}_u$ is the precoder for UT u, x_u is the intended data symbol with zero mean and unit variance, and $\overline{\mathbf{x}} \triangleq \sum_{u'=1}^{U} \mathbf{p}_{u'} x_{u'}$ represents the precoded transmit signal. The term $n_u = \mathbf{h}_u^{\mathrm{H}} \sum_{u' \in \mathcal{Z}_U^+ \setminus u} \mathbf{p}_{u'} x_{u'} + z_u$ denotes the aggregate interference-plus-noise, with its covariance given by

$\sum_{u' \in \mathcal{Z}_U^+ \setminus u} \mathbf{p}_{u'}^{\mathrm{H}} \mathbb{E}\left\{\mathbf{h}_u \mathbf{h}_u^{\mathrm{H}}\right\} \mathbf{p}_{u'} + \sigma_{\mathrm{z}}^2$. Assuming that n_u is Gaussian noise, the ergodic rate of UT u can be expressed as [14]:

$$r_u(\mathbf{p}_1, \cdots, \mathbf{p}_U) = \mathbb{E}\left\{\log\left(1 + \frac{\mathbf{p}_u^{\mathrm{H}} \mathbf{h}_u \mathbf{h}_u^{\mathrm{H}} \mathbf{p}_u}{\sigma_{\mathrm{z}}^2 + \sum_{u' \in \mathcal{Z}_U^+ \setminus u} \mathbf{p}_{u'}^{\mathrm{H}} \mathbb{E}\left\{\mathbf{h}_u \mathbf{h}_u^{\mathrm{H}}\right\} \mathbf{p}_{u'}}\right)\right\}, \tag{10}$$

where the expectation is over the channel realizations.

The robust precoding problem, aimed at maximizing the ergodic sum-rate, is formulated as:

$$\begin{aligned} \mathbf{p}_1^{\mathrm{op}}, \cdots, \mathbf{p}_U^{\mathrm{op}} = & \underset{\mathbf{p}_1, \cdots, \mathbf{p}_U \in \mathbb{C}^M}{\arg\max} \sum_{u=1}^{U} r_u(\mathbf{p}_1, \cdots, \mathbf{p}_U) \\ & \text{s.t.} \quad \sum_{u=1}^{U} \mathbf{p}_u^{\mathrm{H}} \mathbf{p}_u \le P, \end{aligned} \tag{11}$$

where $\mathbf{p}_u^{\mathrm{op}}$ is the robust precoder for each UT u, under the total transmit power constraint P. However, this non-convex problem presents significant computational challenges in massive MIMO systems, even though suboptimal solutions can be found using the MM algorithm [33].

To simplify the robust precoding problem, we leverage the characteristics of HF skywave massive MIMO channels, transforming the high-dimensional problem in (11) into a lower-dimensional format using Slepian sequences, introduced in the following subsection.

B. Slepian Sequences

Slepian sequences were developed to maximize energy concentration within the Fourier spectrum interval $[-W, +W]$ ($W < 1/2$) across all finite-length sequences. This is achieved by maximizing [34]

$$\lambda = \frac{\int_{-W}^{+W} |\sum_{n=0}^{M-1} s[n] e^{-\bar{\jmath} 2\pi n \nu}|^2 \mathrm{d}\nu}{\int_{-\frac{1}{2}}^{\frac{1}{2}} |\sum_{n=0}^{M-1} s[n] e^{-\bar{\jmath} 2\pi n \nu}|^2 \mathrm{d}\nu} = \frac{\mathbf{s}^{\mathrm{H}} \mathbf{B} \mathbf{s}}{\mathbf{s}^{\mathrm{H}} \mathbf{s}}, \tag{12}$$

where $\mathbf{s} \triangleq [s[0], \cdots, s[M-1]]^{\mathrm{T}} \in \mathbb{R}^M$ is a sequence of length M and $\mathbf{B} \in \mathbb{R}^{M \times M}$ is the prolate matrix with elements $[\mathbf{B}]_{i,j} \triangleq 2W \mathrm{sinc}(2W(i-j)), i, j \in \mathcal{Z}_M^+$. Eq. (12) represents the Rayleigh quotient. Its maximum value corresponds to the maximum eigenvalue of $\mathbf{B}$, with the optimal sequence being the respective eigenvector. Thus, the eigenvectors of $\mathbf{B}$ are identified as Slepian sequences, denoted by $\mathbf{s}_m, m \in \mathcal{Z}_M^+$. The eigenvalues of $\mathbf{B}$ follow the order $1 > \lambda_1 > \cdots > \lambda_M > 0$. Approximately the first $2MW$ eigenvalues are close to 1, while the rest tend towards 0 [35]. This indicates that the initial $\approx 2MW$ Slepian sequences adequately represent any sequence of length M with energy predominantly concentrated in the interval $[-W, +W]$.

Building on Slepian sequences, modulated Slepian sequences have been developed to characterize sequences with energy primarily concentrated in a shifted interval $\mathcal{M}(\nu) = [\nu - W, \nu + W]$ for $\nu \neq 0$. These sequences are represented by the matrix $\mathbf{D}(\nu) \triangleq [\mathbf{d}_1(\nu), \cdots, \mathbf{d}_K(\nu)] \in \mathbb{C}^{M \times K}$ [36], where $\mathbf{d}_m(\nu) = \boldsymbol{\omega}(\nu) \odot \mathbf{s}_m \in \mathbb{C}^M, m \in \mathcal{Z}_M^+$ and

$$\boldsymbol{\omega}(\nu) = \left[1, e^{\bar{\jmath} 2\pi \nu}, \cdots, e^{\bar{\jmath} 2\pi \nu (M-1)}\right]^{\mathrm{T}} \in \mathbb{C}^M. \tag{13}$$

Upon normalization, $\mathbf{d}_m(\nu), m \in \mathcal{Z}_M^+$ form an orthonormal basis in $\mathbb{C}^M$, termed as the Slepian basis. Considering a sequence $\mathbf{a} = [a[0], \cdots, a[M-1]]^{\mathrm{T}}$ within the majority band limit of $\mathcal{M}(\nu)$, we observe that for $K \approx 2MW$, the approximation $\mathbf{D}(\nu)\mathbf{D}^{\mathrm{H}}(\nu)\mathbf{a} \approx \mathbf{a}$ holds. This implies that the sequence $\mathbf{a}$ can be effectively compressed to $\overline{\mathbf{a}}(\nu) = \mathbf{D}^{\mathrm{H}}(\nu)\mathbf{a}$ and later reconstructed using $\mathbf{D}(\nu)\overline{\mathbf{a}}(\nu)$ with minimal energy loss. Therefore, $\overline{\mathbf{a}}(\nu)$ is thus termed as the Slepian transform of $\mathbf{a}$, leveraged in our robust precoder designs.

C. Problem Formulation of Slepian Transform Based Robust Precoding

Motivated by the beam structured robust precoding introduced in [17], we propose using the Slepian basis as an alternative to the conventional beam matrix. This involves representing the precoder as $\mathbf{p}_u = \mathbf{D}(\nu_u)\mathbf{w}_u$, where $\mathbf{w}_u \in \mathbb{C}^K$ is a low-dimensional vector. By doing so, we simplify the precoder design problem from $\mathbf{p}_u$ to $\mathbf{w}_u$. The ergodic rate of UT u can be rewritten as

$$\overline{r}_u(\mathbf{w}_1, \cdots, \mathbf{w}_U) = \mathbb{E}\left\{\log\left(1 + \rho_u^{-1} \mathbf{w}_u^{\mathrm{H}} \mathbf{G}_u(\nu_u) \overline{\mathbf{h}}_u \overline{\mathbf{h}}_u^{\mathrm{H}} \mathbf{G}_u^{\mathrm{H}}(\nu_u) \mathbf{w}_u\right)\right\}, \tag{14}$$

where $\rho_u \triangleq \sigma_{\mathrm{z}}^2 + \sum_{u' \in \mathcal{Z}_U^+ \setminus u} \mathbf{w}_{u'}^{\mathrm{H}} \boldsymbol{\Theta}_u(\nu_{u'}) \mathbf{w}_{u'}$, and

$$\mathbf{G}_u(\nu_{u'}) \triangleq \mathbf{D}^{\mathrm{H}}(\nu_{u'}) \mathbf{V} \mathbf{N}_u \in \mathbb{C}^{K \times B_u}, \tag{15}$$

$$\boldsymbol{\Theta}_u(\nu_{u'}) \triangleq \mathbf{G}_u(\nu_{u'}) \mathbf{A}_u \mathbf{G}_u^{\mathrm{H}}(\nu_{u'}) \in \mathbb{C}^{K \times K}, \tag{16}$$

$$\mathbf{A}_u \triangleq \alpha_u^2 \underline{\mathbf{h}}_u \underline{\mathbf{h}}_u^{\mathrm{H}} + (1 - \alpha_u^2) \mathrm{diag}\{\overline{\boldsymbol{\omega}}_u\} \in \mathbb{C}^{B_u \times B_u}, \tag{17}$$

where $\overline{\mathbf{h}}_u \triangleq \mathbf{N}_u^{\mathrm{T}} \widetilde{\mathbf{h}}_u \in \mathbb{C}^{B_u}$, $\underline{\mathbf{h}}_u \triangleq \mathbf{N}_u^{\mathrm{T}} \widehat{\mathbf{h}}_u \in \mathbb{C}^{B_u}$, and $\overline{\boldsymbol{\omega}}_u \triangleq \mathbf{N}_u^{\mathrm{T}} \widetilde{\boldsymbol{\omega}}_u \in \mathbb{C}^{B_u}$ represent the reduced-dimensional beam domain instantaneous CSI, estimated CSI and statistical CSI, respectively. The beam selection matrix $\mathbf{N}_u$ is given by $\mathbf{N}_u \triangleq \left[\mathbf{e}_{\widetilde{M}, b_u[1]}, \mathbf{e}_{\widetilde{M}, b_u[1]}, \cdots, \mathbf{e}_{\widetilde{M}, b_u[B_u]}\right] \in \mathbb{R}^{\widetilde{M} \times B_u}$, where $b_u[1] < b_u[2] < \cdots < b_u[B_u]$ are elements in $\mathcal{B}_u$. The constraint $\sum_{u=1}^{U} \mathbf{p}_u^{\mathrm{H}} \mathbf{p}_u \le P$ leads to $\sum_{u=1}^{U} \mathbf{w}_u^{\mathrm{H}} \mathbf{w}_u \le P$ since

$$\sum_{u=1}^{U} \mathbf{p}_u^{\mathrm{H}} \mathbf{p}_u = \sum_{u=1}^{U} \mathbf{w}_u^{\mathrm{H}} \mathbf{D}^{\mathrm{H}}(\nu_u) \mathbf{D}(\nu_u) \mathbf{w}_u = \sum_{u=1}^{U} \mathbf{w}_u^{\mathrm{H}} \mathbf{w}_u. \tag{18}$$

Accordingly, the problem of (11) can be reformulated as

$$\begin{aligned} \mathbf{w}_1^{\mathrm{op}}, \cdots, \mathbf{w}_U^{\mathrm{op}} = & \underset{\mathbf{w}_1, \cdots, \mathbf{w}_U \in \mathbb{C}^K}{\arg\max} \sum_{u=1}^{U} \overline{r}_u(\mathbf{w}_1, \cdots, \mathbf{w}_U) \\ & \text{s.t.} \quad \sum_{u=1}^{U} \mathbf{w}_u^{\mathrm{H}} \mathbf{w}_u \le P. \end{aligned} \tag{19}$$

Since $\mathbf{w}_u$ can be expressed as the Slepian transform of $\mathbf{p}_u$, i.e., $\mathbf{w}_u = \mathbf{D}^{\mathrm{H}}(\nu_u)\mathbf{D}(\nu_u)\mathbf{w}_u = \mathbf{D}^{\mathrm{H}}(\nu_u)\mathbf{p}_u$, we refer to $\mathbf{w}_u^{\mathrm{op}}$ in (19) as the Slepian transform domain robust precoder (STRP).

Note that this new problem formulation essentially adds the constraint $\mathbf{p}_u \in \mathrm{span}\{\mathbf{D}(\nu_u)\}$ to (11), where $\mathrm{span}\{\cdot\}$ denotes the column space. Although there is an inherent ergodic sum-rate performance loss in this new solution compared to (11), the gap is small due to the sparse nature of beam domain channels in HF skywave massive MIMO. Proper parameterization of $\mathbf{D}(\nu_u)$ further mitigates this loss, as we demonstrate in the following theorem.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/TCOMM.2024.3362146, IEEE Transactions on Communications

Theorem 1. Assume $\mathcal{P}_u \cap \mathcal{P}_{u'} = \emptyset$ for all $u, u' \in \mathcal{Z}_U^+$ and $u \neq u'$, when $\mathcal{L}_u \subseteq \mathcal{M}(\nu_u)$ for all $u \in \mathcal{Z}_U^+$, $K = \lfloor 2MW(1+\epsilon) \rfloor$, $0 < \epsilon < \frac{1}{2W} - 1$, we have

$$\lim_{M\to\infty} \sum_{u=1}^{U} \left(r_u(\mathbf{p}_1^{\mathrm{op}}, \cdots, \mathbf{p}_U^{\mathrm{op}}) - \overline{r}_u(\mathbf{w}_1^{\mathrm{op}}, \cdots, \mathbf{w}_U^{\mathrm{op}})\right) = 0. \quad (20)$$

Proof: See Appendix A.

Theorem 1 indicates that under given conditions, the problem in (11) can be converted to (19) without any loss of optimality.

The expression of $\overline{r}_u(\mathbf{w}_1, \cdots, \mathbf{w}_U)$ in (14) involves expectation, which means it is hard to obtain the optimal solution of the problem (19) with close-form. To tackle this issue, we use the ergodic sum-rate upper bound to obtain an approximate solution for the original problem. According to Jensen's inequality, an upper bound of $\overline{r}_u(\mathbf{w}_1, \cdots, \mathbf{w}_U)$ is given by

$$\begin{aligned} &\overline{r}_u^{\mathrm{ub}}(\mathbf{w}_1, \cdots, \mathbf{w}_U) \\ &= \log\left(1 + \rho_u^{-1} \mathbf{w}_u^{\mathrm{H}} \mathbf{G}_u(\nu_u) \mathbb{E}\{\overline{\mathbf{h}}_u \overline{\mathbf{h}}_u^{\mathrm{H}}\} \mathbf{G}_u^{\mathrm{H}}(\nu_u) \mathbf{w}_u\right) \quad (21) \\ &= \log\left(1 + \rho_u^{-1} \mathbf{w}_u^{\mathrm{H}} \mathbf{\Theta}_u(\nu_{u'}) \mathbf{w}_u\right), \end{aligned}$$

where the ergodic rate upper bound $\overline{r}_u^{\mathrm{ub}}$ is typically tight for single-antenna UTs [15]. This leads to a new optimization problem

$$\begin{aligned} \mathbf{w}_1^{\mathrm{ub,op}}, \cdots, \mathbf{w}_U^{\mathrm{ub,op}} = &\underset{\mathbf{w}_1, \cdots, \mathbf{w}_U \in \mathbb{C}^K}{\arg\max} \sum_{u=1}^{U} \overline{r}_u^{\mathrm{ub}}(\mathbf{w}_1, \cdots, \mathbf{w}_U) \\ &\text{s.t.} \quad \sum_{u=1}^{U} \mathbf{w}_u^{\mathrm{H}} \mathbf{w}_u \leq P. \quad (22) \end{aligned}$$

In accordance with Jensen's inequality and $\mathbb{E}\left\{\mathbf{h}_u \mathbf{h}_u^{\mathrm{H}}\right\} = \mathbf{V}\left(\alpha_u^2 \widehat{\mathbf{h}}_u \widehat{\mathbf{h}}_u^{\mathrm{H}} + (1-\alpha_u^2)\mathrm{diag}\{\widetilde{\boldsymbol{\omega}}_u\}\right)\mathbf{V}^{\mathrm{H}}$, an upper bound of $r_u(\mathbf{p}_1, \cdots, \mathbf{p}_U)$ is

$$r_u^{\mathrm{ub}}(\mathbf{p}_1, \cdots, \mathbf{p}_U) = \log\left(1 + \frac{\mathbf{p}_u^{\mathrm{H}} \mathbf{V}\left(\alpha_u^2 \widehat{\mathbf{h}}_u \widehat{\mathbf{h}}_u^{\mathrm{H}} + (1-\alpha_u^2)\mathrm{diag}\{\widetilde{\boldsymbol{\omega}}_u\}\right)\mathbf{V}^{\mathrm{H}} \mathbf{p}_u}{\sigma_{\mathrm{z}}^2 + \sum_{u' \in \mathcal{Z}_U^+ \setminus u} \mathbf{p}_{u'}^{\mathrm{H}} \mathbf{V}\left(\alpha_u^2 \widehat{\mathbf{h}}_u \widehat{\mathbf{h}}_u^{\mathrm{H}} + (1-\alpha_u^2)\mathrm{diag}\{\widetilde{\boldsymbol{\omega}}_u\}\right)\mathbf{V}^{\mathrm{H}} \mathbf{p}_{u'}}\right). \quad (23)$$

Replacing $r_u(\mathbf{p}_1, \cdots, \mathbf{p}_U)$ in (11) by (23) yields the spatial domain robust precoding problem based on maximizing the ergodic sum-rate upper bound, whose optimal solution is denoted as $\mathbf{p}_u^{\mathrm{ub,op}}, u \in \mathcal{Z}_U^+$. Similarly, one can prove that the following relation holds when the conditions in Theorem 1 are satisfied

$$\lim_{M\to\infty} \sum_{u=1}^{U} \left(r_u^{\mathrm{ub}}\left(\mathbf{p}_1^{\mathrm{ubop}}, \cdots, \mathbf{p}_U^{\mathrm{ubop}}\right) - \overline{r}_u^{\mathrm{ub}}\left(\mathbf{w}_1^{\mathrm{ubop}}, \cdots, \mathbf{w}_U^{\mathrm{ubop}}\right)\right) = 0, \quad (24)$$

which implies that with the objective of maximizing the ergodic sum-rate upper bound, there is only negligible performance loss in converting the spatial domain robust precoding problem to Slepian transform domain.

D. Slepian Transform Based Robust Precoding

By solving the optimization problem (19) or (22), we obtain the STRP, which is denoted as $\mathbf{w}_u, u \in \mathcal{Z}_U^+$ for brevity. By judiciously configuring the parameters for $\mathbf{D}(\nu_u)$, one can ensure the reliable performance of precoding while achieving efficient implementation. Specifically, we develope the following DL precoding approaches.

1) Separate Slepian Transform Based Robust Precoding: Based on Theorem 1, we define the parameters for each UT as follows:

$$\nu_u = \overline{\nu}_u, \quad W \geq \overline{W}, \quad (25)$$

where $\overline{\nu}_u \triangleq (\nu_u^{\mathrm{l}} + \nu_u^{\mathrm{r}})/2$ and $\overline{W} \triangleq \max_{u \in \mathcal{Z}_U^+} (\nu_u^{\mathrm{r}} - \nu_u^{\mathrm{l}})/2$. These settings ensure $\mathcal{L}_u \subseteq \mathcal{M}(\nu_u)$, affirming the optimality of the Slepian transform domain robust precoder (STRP) when $K \geq \lfloor 2MW \rfloor$ and M is sufficiently large. Since $\mathbf{D}(\overline{\nu}_u)$ is designed individually for each UT, we refer to $\mathbf{w}_u, u \in \mathcal{Z}_U^+$ as the separate STRP (SSTRP), denoted by $\overline{\mathbf{w}}_u, u \in \mathcal{Z}_U^+$. In line with this, the precoded signal vector can be written as $\overline{\mathbf{x}} = \sum_{u=1}^{U} \mathbf{D}(\overline{\nu}_u) \overline{\mathbf{w}}_u x_u$.

2) Joint Slepian Transform Based Robust Precoding: When the number of UTs increases, designing $\mathbf{D}(\overline{\nu}_u)$ for each UT separately becomes complex. To address this, we construct L fixed transform matrices $\mathbf{D}(\widetilde{\nu}_l), l \in \mathcal{Z}_L^+$, where M is divisible by L and $\widetilde{\nu}_l, l \in \mathcal{Z}_L^+$ are set as

$$\widetilde{\nu}_l = L^{-1}(l - 0.5) - 0.5. \quad (26)$$

UTs are then grouped into L categories $\mathcal{N}_l, l \in \mathcal{Z}_L^+$, defined as

$$\mathcal{N}_l \triangleq \left\{u \in \mathcal{Z}_U^+ \middle| l_u = l\right\}, \text{ with } N_l = |\mathcal{N}_l|, \quad (27)$$

where l_u is the group index for UT u. Each UT is allocated to the nearest group, determined by

$$l_u \triangleq \underset{l \in \mathcal{Z}_L^+}{\arg\min} |\widetilde{\nu}_l - \overline{\nu}_u| = \lfloor L(\overline{\nu}_u + 0.5) \rfloor + 1. \quad (28)$$

We set $\nu_u = \widetilde{\nu}_{l_u}$, meaning UTs in $\mathcal{N}_l$ reuse $\mathbf{D}(\widetilde{\nu}_l)$ for precoding. To satisfy $\mathcal{L}_u \subseteq \mathcal{M}(\nu_u)$ as per Theorem 1, we set

$$W \geq \overline{W} + \frac{1}{2L}. \quad (29)$$

Alternatively, when W is fixed, L should be at least $\frac{1}{2(W-\overline{W})}$, and ideally as small as possible to reduce complexity. Thus, the number of groups L can be chosen by

$$\min L, \quad \text{s.t. } L \in \mathcal{Z}^+, L \geq \frac{1}{2(W - \overline{W})} \text{ and } L | M, \quad (30)$$

where $L|M$ represents M is divisible by L. This grouping leads to the joint STRP (JSTRP) for $\mathbf{w}_u, u \in \mathcal{Z}_U^+$, denoted as $\widetilde{\mathbf{w}}_u, u \in \mathcal{Z}_U^+$.

The precoded vector $\overline{\mathbf{x}}$ based on JSTRP can be expressed as

$$\overline{\mathbf{x}} = \sum_{u=1}^{U} \mathbf{D}(\widetilde{\nu}_{l_u}) \widetilde{\mathbf{w}}_u x_u = \sum_{l=1}^{L} \mathbf{D}(\widetilde{\nu}_l) \mathbf{c}_l = \mathbf{Q}\mathbf{c}, \quad (31)$$

where $\mathbf{c}_l \triangleq \widetilde{\mathbf{W}}_l \widetilde{\mathbf{N}}_l^{\mathrm{T}} \mathbf{x} \in \mathbb{C}^K, l \in \mathcal{Z}_L^+$ are transform domain precoded vectors, $\widetilde{\mathbf{N}}_l \triangleq \left[\mathbf{e}_{U,u_l[1]}, \mathbf{e}_{U,u_l[2]}, \cdots, \mathbf{e}_{U,u_l[N_l]}\right] \in$

$\mathbb{R}^{U\times N_l}$ is the UT selection matrix for group l with elements $u_l[1] < u_l[2] < \cdots < u_l[N_l]$ in $\mathcal{N}_l$, and

$$\widetilde{\mathbf{W}}_l \triangleq [\widetilde{\mathbf{w}}_1, \widetilde{\mathbf{w}}_2, \cdots, \widetilde{\mathbf{w}}_U]\widetilde{\mathbf{N}}_l \in \mathbb{C}^{K\times N_l}, \tag{32}$$

$$\begin{aligned}\mathbf{Q} &\triangleq [\mathbf{D}(\widetilde{\nu}_1), \mathbf{D}(\widetilde{\nu}_2), \cdots, \mathbf{D}(\widetilde{\nu}_L)] \\ &= [\mathrm{diag}\{\mathbf{d}_1(\widetilde{\nu}_1)\}, \cdots, \mathrm{diag}\{\mathbf{d}_K(\widetilde{\nu}_1)\}](\mathbf{I}_K \otimes \mathbf{\Omega})\mathbf{M}_{L,K}^{\mathrm{T}},\end{aligned} \tag{33}$$

$$\mathbf{\Omega} \triangleq [\boldsymbol{\omega}(0), \boldsymbol{\omega}(1/L), \cdots, \boldsymbol{\omega}((L-1)/L)] = \mathbf{1}_{M/L} \otimes \mathbf{F}_L^{\mathrm{H}}, \tag{34}$$

$$\mathbf{c} \triangleq \left[\mathbf{c}_1^{\mathrm{T}}, \mathbf{c}_2^{\mathrm{T}}, \cdots, \mathbf{c}_L^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{C}^{KL}, \tag{35}$$

where $\mathbf{M}_{L,K} \in \mathbb{R}^{KL\times KL}$ is a permutation matrix with the i-th column being $\mathbf{e}_{KL,K\langle i-1\rangle_L + \lfloor (i-1)/L\rfloor + 1}$, $\mathbf{F}_L$ is L-dimensional DFT matrix with $[\mathbf{F}_L]_{i,j} = e^{-\jmath 2\pi (i-1)(j-1)/L}, i, j \in \mathcal{Z}_L^+$. Consequently, substituting (33) and (34) into (31) yields

$$\begin{aligned}\overline{\mathbf{x}} &= [\mathrm{diag}\{\mathbf{d}_1(\widetilde{\nu}_1)\}, \cdots, \mathrm{diag}\{\mathbf{d}_K(\widetilde{\nu}_1)\}](\mathbf{I}_K \otimes \mathbf{\Omega})\mathbf{M}_{L,K}^{\mathrm{T}}\mathbf{c} \\ &= \sum_{i=1}^{K} \mathbf{d}_i(\widetilde{\nu}_1) \odot \left(\mathbf{1}_{M/L} \otimes \left(\mathbf{F}_L^{\mathrm{H}}\mathbf{m}_i\right)\right),\end{aligned} \tag{36}$$

where $\mathbf{m}_i \triangleq (\mathbf{e}_{K,i} \otimes \mathbf{I}_L)\mathbf{M}_{L,K}^{\mathrm{T}}\mathbf{c} \in \mathbb{C}^L, i \in \mathcal{Z}_K^+$. Thus the computation of $\overline{\mathbf{x}}$ can be efficiently implemented with L-point IFFT, as shown in Fig. 3. Importantly, this precoding method is not limited to JSTRP, it can be applied to all precoders with a joint Slepian transform structure.

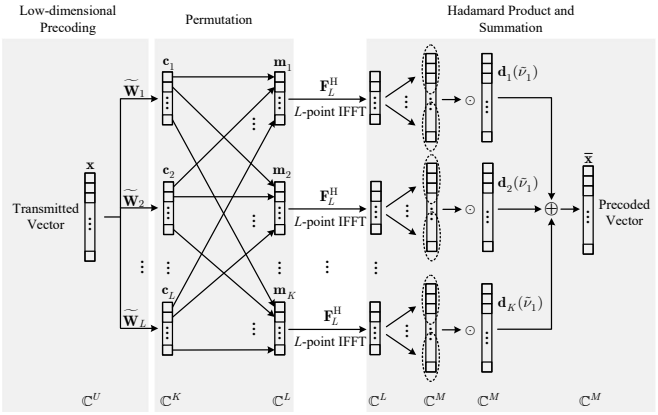

Fig. 3. Efficient implementation of joint Slepian transform based precoding.

E. Implementation Complexity Analysis

We next analyze the implementation complexities of proposed precoding approaches in terms of required complex multiplications. The implementation complexity with the SSTRP comes from the computation of $\overline{\mathbf{x}} = \sum_{u=1}^{U} \mathbf{D}(\overline{\nu}_u)\overline{\mathbf{w}}_u x_u$, which requires $(M+1)KU$ complex multiplications. For JSTRP, the utilization of IFFT can significantly reduce the implementation complexity. Note that $\mathbf{m}_i$ is extracted from the permutation of $\mathbf{c}$ and $\mathbf{1}_{M/L} \otimes \left(\mathbf{F}_L^{\mathrm{H}}\mathbf{m}_i\right)$ represents stacking with M/L copies of $\mathbf{F}_L^{\mathrm{H}}\mathbf{m}_i$, which contribute no computational complexity. And the complex multiplications required for the computation of $\mathbf{c}$, the IFFT of $\mathbf{m}_i$ and the Hadamard product of $\mathbf{d}_i(\widetilde{\nu}_1) \odot \left(\mathbf{1}_{M/L} \otimes \left(\mathbf{F}_L^{\mathrm{H}}\mathbf{m}_i\right)\right)$ are $\sum_{l=1}^{L} KN_l = KU$, $0.5L\log L$ and M, respectively. Consequently, the implementation complexity with JSTRP is $KU + \sum_{i=1}^{K}(0.5L\log L + M) = K(M + U + 0.5L\log L)$.

For comparison, the implementation complexity with the SDRP [14] and the BDRP [17] are respectively MU and $UB + M + \widetilde{M} + S(1+\log S)$, where $B \triangleq \frac{1}{U}\sum_{u=1}^{U} B_u$ and $S \geq M + \widetilde{M} - 1$. When M and U are both large, the implementation complexity of the JSTRP is the lowest among all robust precoders.

IV. Design Of Slepian Transform Domain Robust Precoders

In this section, we derive a STRP design by maximizing the ergodic sum-rate upper bound within the framework of MM algorithm. Further, we prove that STRP design can be simplified by omitting partial inter-UT interferences. Consequently, the SSTRP design and the JSTRP design are developed, respectively. Finally, the design complexities of different robust precoders are analyzed.

A. Slepian Transform Domain Robust Precoder Design With MM Algorithm

For the sake of making the STRP design easier to implement, we focus on the solution of problem (22) and derive a local optimum for it within the framework of MM algorithm [37]. Denote $\mathbf{w}_u^{(i)}$ as the i-th iteration of STRP $\mathbf{w}_u$. Define a minorizing function f to minorize $\sum_{u=1}^{U}\overline{r}_u^{\mathrm{ub}}$ at the point $\left(\mathbf{w}_1^{(i)}, \cdots, \mathbf{w}_U^{(i)}\right)$ with

$$f\left(\mathbf{w}_1, \cdots, \mathbf{w}_U \middle| \mathbf{w}_1^{(i)}, \cdots, \mathbf{w}_U^{(i)}\right) \leq \sum_{u=1}^{U}\overline{r}_u^{\mathrm{ub}}(\mathbf{w}_1, \cdots, \mathbf{w}_U), \tag{37a}$$

$$f\left(\mathbf{w}_1^{(i)}, \cdots, \mathbf{w}_U^{(i)} \middle| \mathbf{w}_1^{(i)}, \cdots, \mathbf{w}_U^{(i)}\right) = \sum_{u=1}^{U}\overline{r}_u^{\mathrm{ub}}\left(\mathbf{w}_1^{(i)}, \cdots, \mathbf{w}_U^{(i)}\right), \tag{37b}$$

which yields

$$\left.\frac{\partial f}{\partial \mathbf{w}_u}\right|_{\mathbf{w}_u = \mathbf{w}_u^{(i)}} = \sum_{u'=1}^{U}\left.\frac{\partial \overline{r}_{u'}^{\mathrm{ub}}}{\partial \mathbf{w}_u}\right|_{\mathbf{w}_u = \mathbf{w}_u^{(i)}}, \quad \text{for } u \in \mathcal{Z}_U^+. \tag{38}$$

The MM algorithm relies on finding a minorizing function for the objective function at each point of the optimization process. Once the minorizing function f is obtained, it is used to replace the objective function, and the algorithm proceeds by maximizing f to obtain an iterative solution. Denote $\left(\mathbf{w}_1^{(i+1)}, \cdots, \mathbf{w}_U^{(i+1)}\right)$ as the maximizer of f, then

$$\sum_{u=1}^{U}\overline{r}_u^{\mathrm{ub}}\left(\mathbf{w}_1^{(i+1)}, \cdots, \mathbf{w}_U^{(i+1)}\right) \geq \sum_{u=1}^{U}\overline{r}_u^{\mathrm{ub}}\left(\mathbf{w}_1^{(i)}, \cdots, \mathbf{w}_U^{(i)}\right). \tag{39}$$

(38) and (39) guarantee the generated sequence can converge to the locally optimal solution of maximizing $\sum_{u=1}^{U}\overline{r}_u^{\mathrm{ub}}$.

Theorem 2. A minorizing function for $\sum_{u=1}^{U}\overline{r}_u^{\mathrm{ub}}$ at the point $\left(\mathbf{w}_1^{(i)}, \cdots, \mathbf{w}_U^{(i)}\right)$ is given by

$$\begin{aligned}f(\mathbf{w}_1, \cdots, \mathbf{w}_U | \mathbf{w}_1^{(i)}, \cdots, \mathbf{w}_U^{(i)}) &= \sum_{u=1}^{U} a_u^{(i)} + \sum_{u=1}^{U}\left(\mathbf{q}_u^{(i)}\right)^{\mathrm{H}}\mathbf{w}_u \\ &+ \sum_{u=1}^{U}\mathbf{w}_u^{\mathrm{H}}\mathbf{q}_u^{(i)} - \sum_{u=1}^{U}\sum_{u'=1}^{U}\mathbf{w}_u^{\mathrm{H}}\mathbf{T}_{u',u}^{(i)}\mathbf{w}_u,\end{aligned} \tag{40}$$

where

$$\begin{aligned} a_u^{(i)} =& \log\Big(1+\big(\rho_u^{(i)}\big)^{-1}\big(\mathbf{w}_u^{(i)}\big)^{\mathrm{H}}\mathbf{\Theta}_u(\nu_u)\mathbf{w}_u^{(i)}\Big) \\ &- \sigma_{\mathrm{z}}^2\big(\rho_u^{(i)}\big)^{-1} + \sigma_{\mathrm{z}}^2\big(\rho_u^{(i)}+\big(\mathbf{w}_u^{(i)}\big)^{\mathrm{H}}\mathbf{\Theta}_u(\nu_u)\mathbf{w}_u^{(i)}\big)^{-1} \\ &- \big(\rho_u^{(i)}\big)^{-1}\big(\mathbf{w}_u^{(i)}\big)^{\mathrm{H}}\mathbf{\Theta}_u(\nu_u)\mathbf{w}_u^{(i)}, \end{aligned} \tag{41a}$$

$$\rho_u^{(i)} = \sigma_{\mathrm{z}}^2 + \textstyle\sum_{u'\in\mathcal{Z}_U^+\setminus u}\big(\mathbf{w}_{u'}^{(i)}\big)^{\mathrm{H}}\mathbf{\Theta}_u(\nu_{u'})\mathbf{w}_{u'}^{(i)}, \tag{41b}$$

$$\mathbf{q}_u^{(i)} = \big(\rho_u^{(i)}\big)^{-1}\mathbf{\Theta}_u(\nu_u)\mathbf{w}_u^{(i)}, \tag{41c}$$

$$\mathbf{T}_{u,u'}^{(i)} = \Big(\big(\rho_u^{(i)}\big)^{-1} - \Big(\rho_u^{(i)}+\rho_u^{(i)}\big(\mathbf{w}_u^{(i)}\big)^{\mathrm{H}}\mathbf{q}_u^{(i)}\Big)^{-1}\Big)\mathbf{\Theta}_u(\nu_{u'}), \tag{41d}$$

Proof: See Appendix B.

Based on the minorizing function f, the optimal solution in the $(i+1)$-th iteration can be obtained by solving the following optimization problem

$$\begin{aligned} \mathbf{w}_1^{(i+1)},\cdots,\mathbf{w}_U^{(i+1)} =& \underset{\mathbf{w}_1,\cdots,\mathbf{w}_U\in\mathbb{C}^K}{\arg\max} \\ \sum_{u=1}^{U}\Big(&\big(\mathbf{q}_u^{(i)}\big)^{\mathrm{H}}\mathbf{w}_u + \mathbf{w}_u^{\mathrm{H}}\mathbf{q}_u^{(i)} - \mathbf{w}_u^{\mathrm{H}}\sum_{u'=1}^{U}\mathbf{T}_{u',u}^{(i)}\mathbf{w}_u\Big) \\ &\text{s.t.} \quad \sum_{u=1}^{U}\mathbf{w}_u^{\mathrm{H}}\mathbf{w}_u \le P. \end{aligned} \tag{42}$$

The problem in (42) is a concave quadratic optimization problem, which can be solved by the Lagrange multiplier methods [38]. Construct the Lagrangian function

$$\begin{aligned} \mathcal{L} =& \sum_{u=1}^{U}\Big(\big(\mathbf{q}_u^{(i)}\big)^{\mathrm{H}}\mathbf{w}_u + \mathbf{w}_u^{\mathrm{H}}\mathbf{q}_u^{(i)} - \mathbf{w}_u^{\mathrm{H}}\sum_{u'=1}^{U}\mathbf{T}_{u',u}^{(i)}\mathbf{w}_u\Big) \\ &- \mu\Big(\sum_{u=1}^{U}\mathbf{w}_u^{\mathrm{H}}\mathbf{w}_u - P\Big), \end{aligned} \tag{43}$$

where $\mu \ge 0$ represents the Lagrange multiplier. According to the first-order optimal conditions, the iteration equation of the precoder is given by

$$\mathbf{w}_u^{(i+1)} = \Big(\sum_{u'=1}^{U}\mathbf{T}_{u',u}^{(i)} + \mu^{\mathrm{op}}\mathbf{I}_K\Big)^{-1}\mathbf{q}_u^{(i)}, \tag{44}$$

where $\mu^{\mathrm{op}} \ge 0$ is the optimal Lagrange multiplier. Note that $\sum_{u=1}^{U}\mathbf{w}_u^{\mathrm{H}}\mathbf{w}_u$ is a monotonically decreasing function of μ. Thus, if $\mu^{\mathrm{op}} = 0$ and $\sum_{u=1}^{U}\mathbf{w}_u^{\mathrm{H}}\mathbf{w}_u \le P$, the optimal solution is $\mathbf{w}_u^{(i+1)} = \big(\sum_{u'=1}^{U}\mathbf{T}_{u',u}^{(i)}\big)^{-1}\mathbf{q}_u^{(i)}$, Otherwise, μ^{op} can be obtained by using a bisection method. Eq. (44) provides an iterative process to obtain general STRP. Note that the computation of $\rho_u^{(i)}$ and $\mathbf{w}_u^{(i+1)}$ requires summation over all UTs, which can be further simplified by the following theorem.

Theorem 3. When $\mathcal{L}_u \cap \mathcal{M}(\nu_{u'}) = \emptyset, u, u' \in \mathcal{Z}_U^+$, $K = \lfloor 2MW(1-\epsilon)\rfloor$, $0<\epsilon<1$, we have

$$\lim_{M\to\infty}\mathbf{G}_u(\nu_{u'}) = \mathbf{O}_{K\times B_u}, \tag{45}$$

Proof: See Appendix C.

Theorem 3 implies that when $K < \lfloor 2MW\rfloor$, the negligibly small interference from UTs that are not overlapped with UT u can be omitted to simplify the computation of $\rho_u^{(i)}$ and $\mathbf{w}_u^{(i+1)}$. However, $K < \lfloor 2MW\rfloor$ conflicts with the condition in Theorem 1, which means a portion of optimality will be sacrificed. Nevertheless, it can be known from the simulation that the loss of optimality is negligible when K is slightly smaller than $\lfloor 2MW\rfloor$. To this end, we introduce a threshold λ^{th} which is less than and close to 1, and set

$$K = \max_{K'\in\mathcal{Z}_M^+,\,\lambda_{K'}>\lambda^{\mathrm{th}}} K'. \tag{46}$$

Therefore, for $\mathcal{M}(\nu_{u'})\cap\mathcal{L}_u = \emptyset$, we have $\mathbf{G}_u(\nu_{u'}) \approx \mathbf{O}_{K\times B_u}$ and

$$\mathbf{\Theta}_u(\nu_{u'}) = \mathbf{G}_u(\nu_{u'})\mathbf{A}_u\mathbf{G}_u^{\mathrm{H}}(\nu_{u'}) \approx \mathbf{O}_{K\times K}, \tag{47}$$

thus partial summation terms in (41b) and (44) can be ignored for simplification. Concretely, the designs of SSTRP and JSTRP in III-D are developed as the following two subsections.

B. Separate Slepian Transform Domian Robust Precoder Design

For SSTRP design, according to Throrem 3, $\rho_u^{(i)}$ in (41b) and $\mathbf{w}_u^{(i+1)}$ in (44) can be respectively simplified as

$$\overline{\rho}_u^{(i)} = \sigma_{\mathrm{z}}^2 + \textstyle\sum_{u'\in\overline{\mathcal{G}}_u}\big(\overline{\mathbf{w}}_{u'}^{(i)}\big)^{\mathrm{H}}\mathbf{\Theta}_u(\overline{\nu}_{u'})\overline{\mathbf{w}}_{u'}^{(i)}, \tag{48}$$

$$\overline{\mathbf{w}}_u^{(i+1)} = \Big(\textstyle\sum_{u'\in\overline{\mathcal{U}}_u}\overline{\zeta}_{u'}^{(i)}\mathbf{\Theta}_{u'}(\overline{\nu}_u) + \mu^{\mathrm{op}}\mathbf{I}_K\Big)^{-1}\overline{\mathbf{q}}_u^{(i)}, \tag{49}$$

where

$$\overline{\zeta}_u^{(i)} = \big(\overline{\rho}_u^{(i)}\big)^{-1} - \Big(\overline{\rho}_u^{(i)}+\overline{\rho}_u^{(i)}\big(\overline{\mathbf{w}}_u^{(i)}\big)^{\mathrm{H}}\overline{\mathbf{q}}_u^{(i)}\Big)^{-1}, \tag{50}$$

$$\overline{\mathbf{q}}_u^{(i)} = \big(\overline{\rho}_u^{(i)}\big)^{-1}\mathbf{\Theta}_u(\overline{\nu}_u)\overline{\mathbf{w}}_u^{(i)}, \tag{51}$$

and the interference UT sets are respectively defined as

$$\overline{\mathcal{G}}_u \triangleq \big\{u'\in\mathcal{Z}_U^+\big|\mathcal{M}(\overline{\nu}_{u'})\cap\mathcal{L}_u\neq\emptyset, u'\neq u\big\}, \tag{52a}$$

$$\overline{\mathcal{U}}_u \triangleq \big\{u'\in\mathcal{Z}_U^+\big|\mathcal{M}(\overline{\nu}_u)\cap\mathcal{L}_{u'}\neq\emptyset\big\}, \tag{52b}$$

with $\overline{G}_u = |\overline{\mathcal{G}}_u|$ and $\overline{U}_u = |\overline{\mathcal{U}}_u|$.

Note that $\mathbf{\Theta}_u(\overline{\nu}_{u'})$ varies symbol by symbol due to the temporal correlation coefficient α_u varies as OFDM symbol changes. However, the computation of $\mathbf{\Theta}_u(\overline{\nu}_{u'})$ for each symbol can be avoided. Using (16) and (17), $\mathbf{\Theta}_u(\overline{\nu}_{u'})$ can be expressed as

$$\mathbf{\Theta}_u(\overline{\nu}_{u'}) = \alpha_u^2\underline{\mathbf{\Theta}}_u(\overline{\nu}_{u'}) + (1-\alpha_u^2)\overline{\mathbf{\Theta}}_u(\overline{\nu}_{u'}), \tag{53}$$

where

$$\underline{\mathbf{\Theta}}_u(\nu) \triangleq \mathbf{G}_u(\nu)\underline{\mathbf{h}}_u\underline{\mathbf{h}}_u^{\mathrm{H}}\mathbf{G}_u^{\mathrm{H}}(\nu), \tag{54a}$$

$$\overline{\mathbf{\Theta}}_u(\nu) \triangleq \mathbf{G}_u(\nu)\mathrm{diag}\{\overline{\boldsymbol{\omega}}_u\}\mathbf{G}_u^{\mathrm{H}}(\nu). \tag{54b}$$

According to (8a) and (8b), $\mathcal{L}_u$ varies as statistical CSI $\widetilde{\boldsymbol{\omega}}_u$ changes. Since $\widetilde{\boldsymbol{\omega}}_u$ changes slowly compared to the instantaneous CSI [28], the update of $\mathbf{D}(\overline{\nu}_u)$ and $\mathbf{G}(\overline{\nu}_u)$ is also slow. Therefore, the computation of $\underline{\mathbf{\Theta}}_u(\overline{\nu}_{u'})$ is required for each slot since the estimated CSI $\underline{\mathbf{h}}_u$ is obtained in each slot's training symbol. On the other hand, the computation of $\overline{\mathbf{\Theta}}_u(\overline{\nu}_{u'})$ is only needed when $\widetilde{\boldsymbol{\omega}}_u$ changes. Hence, α_u in (53) can be adjusted to achieve the variation of $\mathbf{\Theta}_u(\overline{\nu}_{u'})$ with symbols, where $\underline{\mathbf{\Theta}}_u(\overline{\nu}_{u'})$ and $\overline{\mathbf{\Theta}}_u(\overline{\nu}_{u'})$ can be pre-calculated.

The design procedure of SSTRP is summarized in Algorithm 1.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/TCOMM.2024.3362146, IEEE Transactions on Communications

Algorithm 1: SSTRP design

Initialize $\overline{\mathbf{w}}_u^{(i)}, u \in \mathcal{Z}_U^+$ which satisfies $\sum_{u=1}^U \left(\overline{\mathbf{w}}_u^{(i)}\right)^{\mathrm{H}} \overline{\mathbf{w}}_u^{(i)} \leq P$. Set $i = 0$.
repeat
  Compute $\overline{\rho}_u^{(i)}, \overline{\mathbf{q}}_u^{(i)}, \overline{\zeta}_u^{(i)}, u \in \mathcal{Z}_U^+$ by (48), (51) and (50), respectively.
  Update $\overline{\mathbf{w}}_u^{(i+1)}, u \in \mathcal{Z}_U^+$ by (49) and set $i = i + 1$.
until $\left|\sum_{u=1}^U \overline{r}_u^{\mathrm{ub},(i)} - \sum_{u=1}^U \overline{r}_u^{\mathrm{ub},(i-1)}\right| \leq v_1$ (v_1 is a predefined target)
Obtain $\overline{\mathbf{w}}_u = \overline{\mathbf{w}}_u^{(i)}, u \in \mathcal{Z}_U^+$.

Algorithm 2: JSTRP design

Initialize $\widetilde{\mathbf{w}}_u^{(i)}, u \in \mathcal{Z}_U^+$ which satisfies $\sum_{u=1}^U \left(\widetilde{\mathbf{w}}_u^{(i)}\right)^{\mathrm{H}} \widetilde{\mathbf{w}}_u^{(i)} \leq P$. Set $i = 0$.
repeat
  Compute $\widetilde{\rho}_u^{(i)}, \widetilde{\mathbf{q}}_u^{(i)}, \widetilde{\zeta}_u^{(i)}, u \in \mathcal{Z}_U^+$ by (55), (59) and (57), respectively.
  Compute $\widetilde{\mathbf{E}}_l^{(i)}, l \in \mathcal{Z}_L^+$ by (58).
  Update $\widetilde{\mathbf{w}}_u^{(i+1)}, u \in \mathcal{Z}_U^+$ by (56) and set $i = i + 1$.
until $\left|\sum_{u=1}^U \overline{r}_u^{\mathrm{ub},(i)} - \sum_{u=1}^U \overline{r}_u^{\mathrm{ub},(i-1)}\right| \leq v_2$ (v_2 is a predefined target)
Obtain $\widetilde{\mathbf{w}}_u = \widetilde{\mathbf{w}}_u^{(i)}, u \in \mathcal{Z}_U^+$.

C. Joint Slepian Transform Domain Robust Precoder Design

According to Theorem 3, for JSTRP design, $\mathcal{M}(\widetilde{\nu}_l) \cap \mathcal{L}_u = \emptyset, l \in \mathcal{Z}_L^+, u \in \mathcal{Z}_U^+$ leads to $\mathbf{\Theta}_u(\widetilde{\nu}_l) \approx \mathbf{O}_{K\times K}$. Then $\rho_u^{(i)}$ in (41b) and $\mathbf{w}_u^{(i+1)}$ in (44) can be simplified as

$$\widetilde{\rho}_u^{(i)} = \sigma_{\mathrm{z}}^2 + \textstyle\sum_{u' \in \widetilde{\mathcal{G}}_u} \left(\widetilde{\mathbf{w}}_{u'}^{(i)}\right)^{\mathrm{H}} \mathbf{\Theta}_u\left(\widetilde{\nu}_{l_{u'}}\right) \widetilde{\mathbf{w}}_{u'}^{(i)}, \tag{55}$$

$$\widetilde{\mathbf{w}}_u^{(i+1)} = \widetilde{\mathbf{E}}_{l_u}^{(i)} \widetilde{\mathbf{q}}_u^{(i)}, \tag{56}$$

respectively, where

$$\widetilde{\zeta}_u^{(i)} = \left(\widetilde{\rho}_u^{(i)}\right)^{-1} - \left(\widetilde{\rho}_u^{(i)} + \widetilde{\rho}_u^{(i)} \left(\widetilde{\mathbf{w}}_u^{(i)}\right)^{\mathrm{H}} \widetilde{\mathbf{q}}_u^{(i)}\right)^{-1}, \tag{57}$$

$$\widetilde{\mathbf{E}}_l^{(i)} = \left(\textstyle\sum_{u \in \widetilde{\mathcal{U}}_l} \widetilde{\zeta}_u^{(i)} \mathbf{\Theta}_u(\widetilde{\nu}_l) + \mu^{\mathrm{op}} \mathbf{I}_K\right)^{-1}, \tag{58}$$

$$\widetilde{\mathbf{q}}_u^{(i)} = \left(\widetilde{\rho}_u^{(i)}\right)^{-1} \mathbf{\Theta}_u(\widetilde{\nu}_{l_u}) \widetilde{\mathbf{w}}_u^{(i)}, \tag{59}$$

and the interference UT sets are respectively defined as

$$\widetilde{\mathcal{G}}_u \triangleq \left\{u' \in \mathcal{Z}_U^+ \middle| \mathcal{M}\left(\widetilde{\nu}_{l_{u'}}\right) \cap \mathcal{L}_u \neq \emptyset, u' \neq u\right\}, \tag{60a}$$

$$\widetilde{\mathcal{U}}_l \triangleq \left\{u \in \mathcal{Z}_U^+ \middle| \mathcal{M}(\widetilde{\nu}_l) \cap \mathcal{L}_u \neq \emptyset\right\}, \tag{60b}$$

with $\widetilde{G}_u = |\widetilde{\mathcal{G}}_u|$ and $\widetilde{U}_l = |\widetilde{\mathcal{U}}_l|$. It is worth noting that in each iteration, the matrix inversion term (58) needs to be applied L times instead of U times in (49), which is cost effective when U is significantly larger than L.

Similar to (53), $\mathbf{\Theta}_u(\widetilde{\nu}_l)$ can be rewritten as

$$\mathbf{\Theta}_u(\widetilde{\nu}_l) = \alpha_u^2 \underline{\mathbf{\Theta}}_u(\widetilde{\nu}_l) + (1 - \alpha_u^2) \overline{\mathbf{\Theta}}_u(\widetilde{\nu}_l). \tag{61}$$

Since $\mathbf{D}(\widetilde{\nu}_l)$ and the beam matrix $\mathbf{V}$ are fixed, $\mathbf{G}(\widetilde{\nu}_l) = \mathbf{D}^{\mathrm{H}}(\widetilde{\nu}_l)\mathbf{V}$ can be computed offline. Consequently, $\underline{\mathbf{\Theta}}_u(\widetilde{\nu}_l)$ and $\overline{\mathbf{\Theta}}_u(\widetilde{\nu}_l)$ can be pre-calculated before designing the precoder as well. Additionally, the number of $\mathbf{\Theta}_u(\widetilde{\nu}_l)$ required for the precoder design is $\sum_{l=1}^L \widetilde{U}_l$, whereas that of $\mathbf{\Theta}_u(\overline{\nu}_{u'})$ in SSTRP design is $\sum_{u=1}^U \overline{U}_u$, which implies that the STRP design in this subsection enjoys lower storage requiremnets when the number of UTs gets huge.

Algorithm 2 summarizes the JSTRP design.

D. Design Complexity Analysis

In this subsection, we evaluate the design complexities of proposed precoders. It is worth noting that all robust precoders feature a bisection step that only needs a small number of iterations and can be disregarded in complexity analysis. Therefore, we simplify the analysis by representing the design complexity as the number of complex multiplications required for each iteration. We define $\overline{G} \triangleq \frac{1}{U}\sum_{u=1}^U \overline{G}_u$ and $\widetilde{G} \triangleq \frac{1}{U}\sum_{u=1}^U \widetilde{G}_u$ for the purpose of analysis.

Note that for SSTRP, $\mathbf{\Theta}_u(\overline{\nu}_{u'})$ can be pre-calculated. Therefore, the complexities of computing $\overline{\rho}_u^{(i)}$, $\overline{\mathbf{q}}_u^{(i)}$ and $\overline{\zeta}_u^{(i)}$ are respectively $K(K+1)\overline{G}U$, K^2U and KU, and the complexity of updating $\overline{\mathbf{w}}_u^{(i+1)}$ is $(0.5K^3 + 2.5K^2 + K)U$ [39]. The overall design complexity of the SSTRP is thus given by $(0.5K^3 + (3.5 + \overline{G})K^2 + (\overline{G} + 1)K)U$. In a similar way, the design complexity of the JSTRP is $((\widetilde{G} + 2)K^2 + (\widetilde{G} + 1)K)U + \left(0.5K^3 + 1.5K^2\right)L$.

For comparison, the design complexities of the SDRP [14] and the BDRP [17] are $((U+1)M^2+UM)U+0.5M^3+1.5M^2$ and $(0.5B^3+(2.5+U)B^2+UB)U$, respectively. When M is large, the design complexities of proposed precoders are much lower than that of the SDRP. Moreover, when U is large, the design complexities of proposed STRPs can be lower than that of the BDRP.

V. Simulation Results

In this section, we provide simulation results to demonstrate advantages of the proposed precoders for HF skywave massive MIMO systems.

In our simulations, we use the parameters listed in Table I for the HF skywave massive MIMO-OFDM system. Specifically, the BS is located at $(34°\mathrm{N}, 118°\mathrm{E})$. UTs are distributed at a distance of 2000 km from the BS, with the azimuth angles distributed in $[-70°, 70°]$. Realistic HF skywave channel parameters are generated using the commercial ray-tracing software Proplab-Pro version 3.1 [40]. This includes the signal strength, azimuth (elevation) angles of departure (AoD), and propagation distance for each path, from which we compute the path gain, direction cosine, and propagation delay, respectively. Additionally, we generate random initial phases and Doppler shifts for each path. Using (2), we get the spatial domain channel, while the beam domain channel and statistical CSI are obtained using the method proposed in [28] with $\widetilde{M} = 2M$.

The performances of various DL robust precoders are compared by computing the ergodic sum-rates versus the system configurations and parameters, where the ergodic sum-rate is obtained by averaging channel realizations among DL data symbols and valid subcarriers. For proposed STRPs, we adopt

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/TCOMM.2024.3362146, IEEE Transactions on Communications

TABLE I
SIMULATION PARAMETERS

Parameter	Value
Carrier frequency f_c	16 MHz
Number of antennas at BS M	128 / 256 / 512
Subcarriers spacing Δ_f	250 Hz
Antenna spacing at BS d	9 m
Number of subcarriers N_c	2048
Number of valid subcarriers N_v	1536
Frame structure $(N, \widetilde{N})$	$(14, 6)$
Number of UTs U	64

the threshold $\lambda^{\text{th}} = 0.99$. Specifically, the following DL robust precoders are compared:

- **SDRP:** Spatial domain robust precoder [14].
- **BDRP:** Beam domain robust precoder [17].
- **SSTRP:** Separate Slepian transform domain robust precoder, designed in Algorithm 1.
- **JSTRP:** Joint Slepian transform domain robust precoder, designed in Algorithm 2.

Fig. 4 provides an overview of the ergodic sum-rates for different precoders. Fig. 4(a) specifically demonstrates the robustness of our proposed precoders under varying channel Doppler spreads. These spreads are influenced by UT speeds and ionospheric-induced Doppler effects, with maximum values set to 60 km/h (200 km/h) and 3 Hz (5 Hz), respectively. SSTP (JSTP) stands for the separate (joint) Slepian transform domain precoder using genie-aided perfect instantaneous CSI, which is designed in terms of the criterion of sum-rate maximization within the framwork of MM algorithm as well. Additionally, we employ an MMSE precoder, designed based on channel estimates during training symbols and applied in DL data symbols. The figure reveals that, compared to SSTP and JSTP, the performances of our proposed precoders experience a moderate decline at higher Doppler spreads. This result is expected because perfect instantaneous CSI becomes challenging to obtain when the channel exhibits significant temporal variability. Nonetheless, both SSTRP and JSTRP consistently outperform the MMSE precoder, particularly when the channel undergoes rapid variations. This observation underscores the robustness of our proposed precoders. Further performance comparison with other robust precoders is presented in Fig. 4(b). Notably, JSTRP's sum-rate performance closely aligns with SSTRP, approaching that of SDRP and outperforming BDRP in regions with high transmit power.

In Fig. 5, we depict the relationship between sum-rate performance and system parameters with $M = 256$. The maximum UT speed and ionospheric-induced maximum Doppler spread are fixed at 60 km/h and 3 Hz, respectively, for all subsequent simulations. From the Figure, it's evident that the sum-rate performance of our proposed precoders improves as W increases, surpassing the BDRP and closely approaching the SDRP's performance. Additionally, for JSTRP, performance improves with larger L values but lags behind SSTRP when $L < U$. However, the differences become less significant when $W \geq 0.04$ and $L \geq 32$, as this fulfills conditions in (25) and (29). Consequently, to strike a balance between precoding performance and complexity, choosing $W = 0.04$ and $L = 32$ is a suitable option in this scenario.

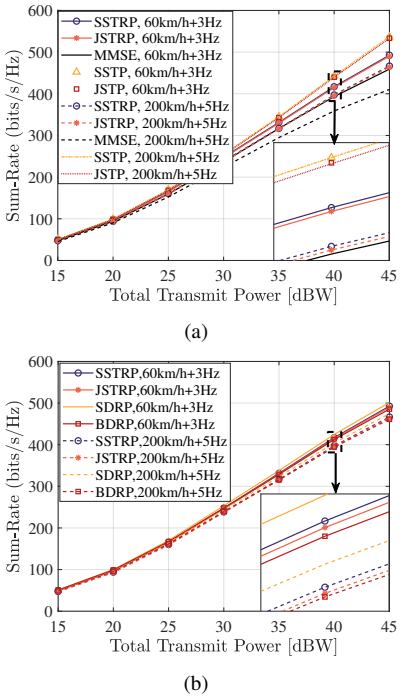

Fig. 4. Ergodic sum-rates of various precoders with $(M, W, L) = (256, 0.04, 32)$ under different channel Doppler spreads. (a) Performance of robustness for SSTRP and JSTRP; (b) Performance comparison with SDRP and BDRP.

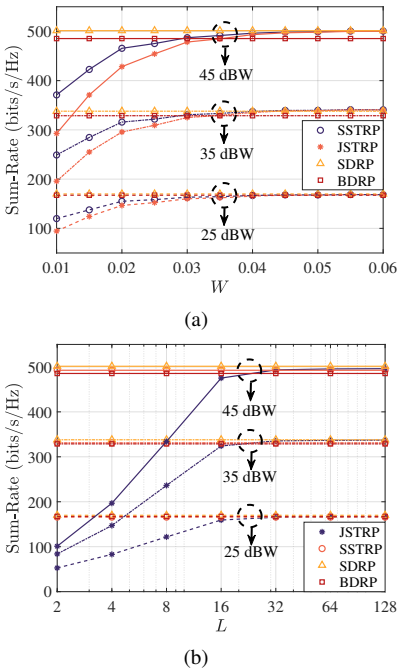

Fig. 5. Ergodic sum-rate comparison with (a) different W ($L = 32$) (b) different L ($W = 0.04$).

Fig. 6 compares the ergodic sum-rate performances of different precoders with different number of BS antennas. From the figure, the proposed precoders achieve a near-optimal sum-rate performance, and the sum-rate gap compared with the SDRP increases as the transmit power grows due to the

impact of inter-UT interference. Further, the gap between proposed precoders and the SDRP gets smaller as M increases since the channel vectors of UTs become nearly orthogonal when M is sufficiently large, in which case the inter-UT interference becomes relatively small, leading to high ergodic sum-rate performance. Moreover, when M tends to infinity, the ergodic sum-rate gap tends to vanish, which is confirmed by Theorem 1.

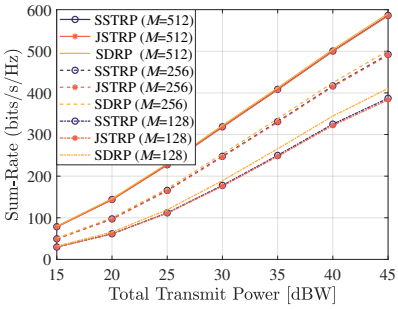

Fig. 6. Ergodic sum-rates of SSTRP, JSTRP, and SDRP with different number of BS antennas and $(W, L) = (0.04, 32)$.

Further, the convergence performance of the proposed robust precoder design algorithms is investigated. Fig. 7 illustrates the ergodic sum-rates of Algorithm1 and Algorithm 2 at each iteration, where the total transmit power respectively equals 15 dBW, 25 dBW, and 35 dBW for $M = 256$. The initial values of precoders are generated randomly. It is observed from the figure that both Algorithm 1 and Algorithm 2 exhibit fast convergence. Specifically, the proposed algorithms converge in only 2 to 4 iterations when the total transmit power is 15 dBW or 25 dBW, while about 10 iterations are required when the total transmit power is 35 dBW. Thus, the effectiveness of the proposed Slepian transform based robust precoding for HF skywave massive MIMO systems is confirmed.

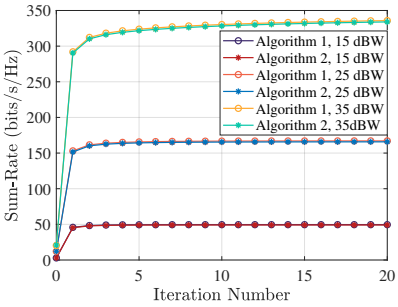

Fig. 7. Convergence of Algorithm 1 and Algorithm 2 with random initial values and $(W, L) = (0.04, 32)$.

Finally, to show the feasibility for practical applications, the computational complexities of SSTRP, JSTRP, SDRP and BDRP are plotted with different number of BS antennas and UTs under typical parameter settings in Fig. 8. The complexities in Fig. 8 refer to the computational complexity for each valid subcarrier in one OFDM symbol. For design complexity, Fig. 8(a), Fig. 8(b) and Fig. 8(c) exhibit that the advantage of the proposed precoders is rather apparent with various typical parameter settings, especially for the JSTRP. For implementation complexity in Fig. 8(d), Fig. 8(e) and Fig. 8(f), the SSTRP has no superiority, while the complexity advantage of the JSTRP is highlighted when the UTs number is sufficiently large. Therefore, for situations with few UTs, the SSTRP can be chosen. For scenarios with a large number of UTs, the JSTRP is preferred.

VI. Conclusions

In this paper, we have investigated the DL robust precoding for HF massive MIMO with Slepian transform. An *a posteriori* channel model based on beams is introduced, and the relationship between the sparse beams and the Fourier spectrum of the spatial domain channel is established. Using this channel model, it is proved that the robust precoding problem of maximizing ergodic sum-rate or its upper bound can be solved in Slepian transform domain. Moreover, we develope two Slepian transform based robust precoding approaches, and propose an efficient implementation of joint Slepian transform based precoding. Additionally, a local optimum for the STRP design is derived within the framework of the MM algorithm. After simplifying the STRP design by using features of modulated Slepian sequences, the SSTRP design and the JSTRP design are proposed, respectively. Simulation results demonstrate that the proposed robust precoders can achieve near-optimal performance with low complexity.

Appendix A
Proof of Theorem 1

According to [17] we know that when $\mathcal{P}_u \cap \mathcal{P}_{u'} = \emptyset$ for all $u, u' \in \mathcal{Z}_U^+$ and $u \neq u'$,

$$\lim_{M\to\infty} (\mathbf{p}_u^{\mathrm{op}} - \mathbf{V}\mathbf{N}_u\mathbf{b}_u^{\mathrm{op}}) = \mathbf{0}_M, \tag{62}$$

where $\mathbf{b}_u^{\mathrm{op}}$ is the optimal beam domain robust precoder (BDRP). (62) demonstrates that when $M \to \infty$, $\mathbf{p}_u^{\mathrm{op}}$ can be modelled as a beam structured vector, and shares the same set of beams with $\mathbf{h}_u$, i.e., $\mathcal{B}_u$, then the support set of $P_u(\nu)$ is $\mathcal{L}_u$, where $P_u(\nu)$ represents the discrete time Fourier transform (DTFT) of sequence $\mathbf{p}_u^{\mathrm{op}}$. According to the Parseval's theorem,

$$\lim_{M\to\infty} \left|\mathbf{d}_i^{\mathrm{H}}(\nu_u)\mathbf{p}_u^{\mathrm{op}}\right| = \lim_{M\to\infty} \left|\int_{\mathcal{L}_u} D_i^*(\nu_u, \nu)P_u(\nu)\mathrm{d}\nu\right|, \tag{63}$$

where $D_i(\nu_u, \nu)$ is the DTFT of sequence $\mathbf{d}_i(\nu_u)$. For $\mathcal{L}_u \subseteq \mathcal{M}(\nu_u)$,

$$\begin{aligned}&\left|\int_{\mathcal{L}_u} D_i^*(\nu_u, \nu)P_u(\nu)\mathrm{d}\nu\right|^2 \overset{\text{(a)}}{\leq} \\ &\left(\int_{\mathcal{M}(\nu_u)} |D_i(\nu_u, \nu)|^2\mathrm{d}\nu\right)\left(\int_{\mathcal{L}_u} |P_u(\nu)|^2\mathrm{d}\nu\right) \overset{\text{(b)}}{\leq} \lambda_i P,\end{aligned} \tag{64}$$

where (a) follows from the Cauchy-Schwarz inequality [41] and $\int_{\mathcal{L}_u} |D_i(\nu_u, \nu)|^2\mathrm{d}\nu \leq \int_{\mathcal{M}(\nu_u)} |D_i(\nu_u, \nu)|^2\mathrm{d}\nu$, (b) follows from

$$\lambda_i = \frac{\mathbf{s}_i^{\mathrm{H}}\mathbf{B}\mathbf{s}_i}{\mathbf{s}_i^{\mathrm{H}}\mathbf{s}_i} = \frac{\mathbf{d}_i^{\mathrm{H}}(\nu_u)\overline{\mathbf{B}}(\nu_u)\mathbf{d}_i(\nu_u)}{\mathbf{d}_i^{\mathrm{H}}(\nu_u)\mathbf{d}_i(\nu_u)} = \int_{\mathcal{M}(\nu)} |D_i(\nu_u, \nu)|^2\mathrm{d}\nu, \tag{65}$$

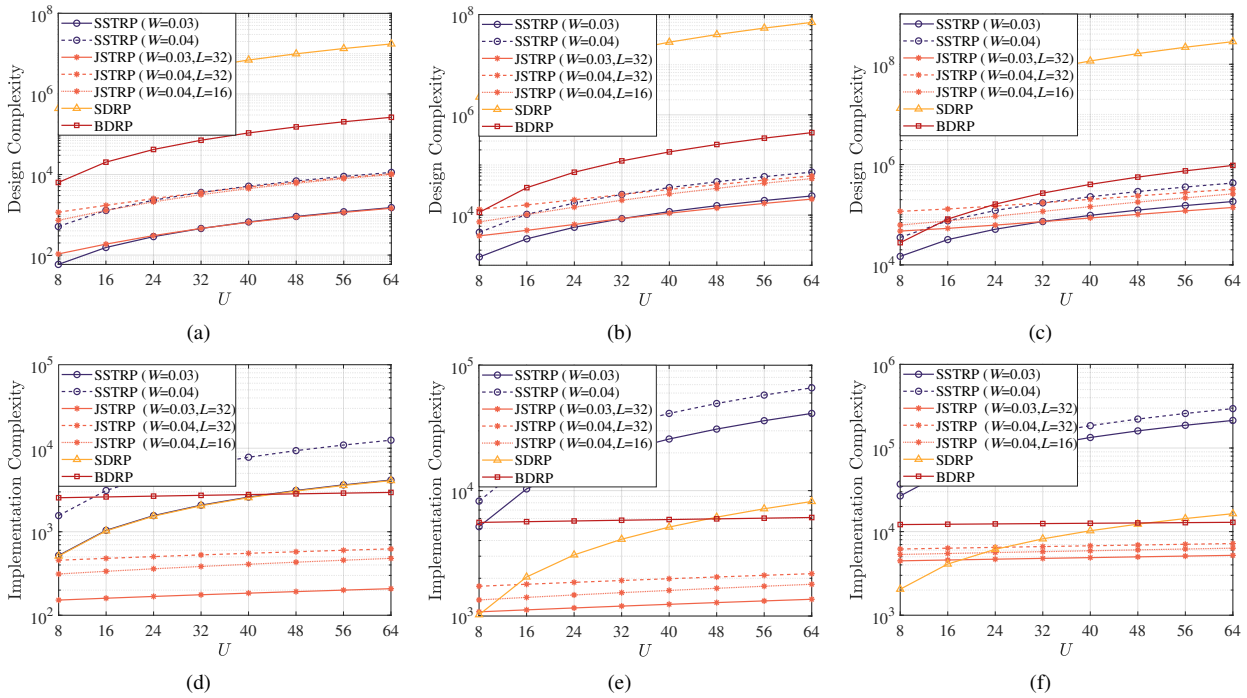

Fig. 8. Complexity comparison of SSTRP, JSTRP, SDRP and BDRP versus the number of UTs. (a) Design complexity with $M = 64$; (b) Design complexity with $M = 128$; (c) Design complexity with $M = 256$; (d) Implementation complexity with $M = 64$; (e) Implementation complexity with $M = 128$; (f) Implementation complexity with $M = 256$.

and power constraint $\left(\mathbf{p}_u^{\mathrm{op}}\right)^{\mathrm{H}}\mathbf{p}_u^{\mathrm{op}} \leq \sum_{u=1}^{U}\left(\mathbf{p}_u^{\mathrm{op}}\right)^{\mathrm{H}}\mathbf{p}_u^{\mathrm{op}} \leq P$, where $\overline{\mathbf{B}}(\nu_u) \triangleq \left(\boldsymbol{\omega}(\nu_u)\boldsymbol{\omega}^{\mathrm{H}}(\nu_u)\right) \odot \mathbf{B}$. According to [18] and [36, Lemma 4.2], when $i > K = \lfloor 2MW(1+\epsilon) \rfloor$, $0 < \epsilon < \frac{1}{2W} - 1$ and $M \geq M_{W,\epsilon}$, we have the following inequality

$$\left| \int_{\mathcal{M}(\overline{\nu})} D_i^*(\overline{\nu}_u, \nu) P_u(\nu) \mathrm{d}\nu \right| \leq \sqrt{\lambda_i P} \leq \sqrt{P e^{-C_{W,\epsilon} M}}, \quad (66)$$

where $C_{W,\epsilon}$ and $M_{W,\epsilon}$ are positive constants depend on W and ϵ. According to $\|\mathbf{d}_i(\nu_u)\| = 1$, (63) and (66), we have

$$\begin{aligned} \left\|\mathbf{p}_u^{\mathrm{op}} - \mathbf{D}(\nu_u)\mathbf{D}^{\mathrm{H}}(\nu_u)\mathbf{p}_u^{\mathrm{op}}\right\| &= \left\| \sum_{i=K+1}^{M} \mathbf{d}_i(\nu_u)\mathbf{d}_i^{\mathrm{H}}(\nu_u)\mathbf{p}_u^{\mathrm{op}} \right\| \\ &\leq \sum_{i=K+1}^{M} \left| \mathbf{d}_i^{\mathrm{H}}(\nu_u)\mathbf{p}_u^{\mathrm{op}} \right|, \quad (67) \end{aligned}$$

$$\begin{aligned} \lim_{M\to\infty} M \sum_{i=K+1}^{M} \left| \mathbf{d}_i^{\mathrm{H}}(\nu_u)\mathbf{p}_u^{\mathrm{op}} \right| &\leq \lim_{M\to\infty} \sqrt{P} M (M-K) e^{-C_{W,\epsilon} M/2} \\ &= 0. \quad (68) \end{aligned}$$

Then we have

$$\lim_{M\to\infty} M \left\| \mathbf{p}_u^{\mathrm{op}} - \mathbf{p}_u^{\diamond} \right\| = 0, \quad (69)$$

where $\mathbf{p}_u^{\diamond} = \mathbf{D}(\nu_u)\mathbf{w}_u^{\diamond}$, $\mathbf{w}_u^{\diamond} = \mathbf{D}^{\mathrm{H}}(\nu_u)\mathbf{p}_u^{\mathrm{op}}$.

The difference between the ergodic sum-rates under the precoder $\mathbf{p}_u^{\mathrm{op}}, u \in \mathcal{Z}_U^+$ and $\mathbf{p}_u^{\diamond}, u \in \mathcal{Z}_U^+$ is given by

$$\begin{aligned} &\sum_{u=1}^{U} \left(r_u(\mathbf{p}_1^{\mathrm{op}}, \cdots, \mathbf{p}_U^{\mathrm{op}}) - r_u(\mathbf{p}_1^{\diamond}, \cdots, \mathbf{p}_U^{\diamond}) \right) = \\ &\sum_{u=1}^{U} \left(\log \delta_u + \mathbb{E}\{\log \overline{\delta}_u\} \right), \end{aligned} \quad (70)$$

where

$$\delta_u \triangleq \frac{\sigma_{\mathrm{z}}^2 + \sum_{u' \in \mathcal{Z}_U^+ \setminus u} \mathbb{E}\{(\mathbf{p}_{u'}^{\diamond})^{\mathrm{H}}\mathbf{h}_u\mathbf{h}_u^{\mathrm{H}}\mathbf{p}_{u'}^{\diamond}\}}{\sigma_{\mathrm{z}}^2 + \sum_{u' \in \mathcal{Z}_U^+ \setminus u} \mathbb{E}\{(\mathbf{p}_{u'}^{\mathrm{op}})^{\mathrm{H}}\mathbf{h}_u\mathbf{h}_u^{\mathrm{H}}\mathbf{p}_{u'}^{\mathrm{op}}\}}, \quad (71a)$$

$$\overline{\delta}_u \triangleq \frac{\sigma_{\mathrm{z}}^2 + \sum_{u' \in \mathcal{Z}_U^+ \setminus u} \mathbb{E}\{(\mathbf{p}_{u'}^{\mathrm{op}})^{\mathrm{H}}\mathbf{h}_u\mathbf{h}_u^{\mathrm{H}}\mathbf{p}_{u'}^{\mathrm{op}}\} + (\mathbf{p}_u^{\mathrm{op}})^{\mathrm{H}}\mathbf{h}_u\mathbf{h}_u^{\mathrm{H}}\mathbf{p}_u^{\mathrm{op}}}{\sigma_{\mathrm{z}}^2 + \sum_{u' \in \mathcal{Z}_U^+ \setminus u} \mathbb{E}\{(\mathbf{p}_{u'}^{\diamond})^{\mathrm{H}}\mathbf{h}_u\mathbf{h}_u^{\mathrm{H}}\mathbf{p}_{u'}^{\diamond}\} + (\mathbf{p}_u^{\diamond})^{\mathrm{H}}\mathbf{h}_u\mathbf{h}_u^{\mathrm{H}}\mathbf{p}_u^{\diamond}}. \quad (71b)$$

Obviously, $\delta_u > 0$, then we have

$$1 - \delta_u^{-1} \leq \log \delta_u \leq \delta_u - 1. \quad (72)$$

According to $\sum_{u' \in \mathcal{Z}_U^+ \setminus u} \mathbb{E}\{(\mathbf{p}_{u'}^{\mathrm{op}})^{\mathrm{H}}\mathbf{h}_u\mathbf{h}_u^{\mathrm{H}}\mathbf{p}_{u'}^{\mathrm{op}}\} \geq 0$ and absolute value inequalities, we have

$$\begin{aligned} |\delta_u - 1| &= \frac{\left| \sum_{u' \in \mathcal{Z}_U^+ \setminus u} \mathbb{E}\{(\mathbf{p}_{u'}^{\diamond})^{\mathrm{H}}\mathbf{h}_u\mathbf{h}_u^{\mathrm{H}}\mathbf{p}_{u'}^{\diamond} - (\mathbf{p}_{u'}^{\mathrm{op}})^{\mathrm{H}}\mathbf{h}_u\mathbf{h}_u^{\mathrm{H}}\mathbf{p}_{u'}^{\mathrm{op}}\} \right|}{\sigma_{\mathrm{z}}^2 + \sum_{u' \in \mathcal{Z}_U^+ \setminus u} \mathbb{E}\{(\mathbf{p}_{u'}^{\mathrm{op}})^{\mathrm{H}}\mathbf{h}_u\mathbf{h}_u^{\mathrm{H}}\mathbf{p}_{u'}^{\mathrm{op}}\}} \\ &\leq \sigma_{\mathrm{z}}^{-2} \sum_{u' \in \mathcal{Z}_U^+ \setminus u} \mathbb{E}\left\{ \left| |\mathbf{h}_u^{\mathrm{H}}\mathbf{p}_{u'}^{\diamond}|^2 - |\mathbf{h}_u^{\mathrm{H}}\mathbf{p}_{u'}^{\mathrm{op}}|^2 \right| \right\}. \quad (73) \end{aligned}$$

Moreover,

$$\begin{aligned} &\left| |\mathbf{h}_u^{\mathrm{H}}\mathbf{p}_{u'}^{\diamond}|^2 - |\mathbf{h}_u^{\mathrm{H}}\mathbf{p}_{u'}^{\mathrm{op}}|^2 \right| \\ &= \left(|\mathbf{h}_u^{\mathrm{H}}\mathbf{p}_{u'}^{\diamond}| + |\mathbf{h}_u^{\mathrm{H}}\mathbf{p}_{u'}^{\mathrm{op}}| \right) \left| |\mathbf{h}_u^{\mathrm{H}}\mathbf{p}_{u'}^{\diamond}| - |\mathbf{h}_u^{\mathrm{H}}\mathbf{p}_{u'}^{\mathrm{op}}| \right| \\ &\overset{(a)}{\leq} \left(|\mathbf{h}_u^{\mathrm{H}}\mathbf{p}_{u'}^{\diamond}| + |\mathbf{h}_u^{\mathrm{H}}\mathbf{p}_{u'}^{\mathrm{op}}| \right) |\mathbf{h}_u^{\mathrm{H}}(\mathbf{p}_{u'}^{\diamond} - \mathbf{p}_{u'}^{\mathrm{op}})| \\ &\overset{(b)}{\leq} \|\mathbf{h}_u\|^2 \left(\|\mathbf{p}_{u'}^{\diamond}\| + \|\mathbf{p}_{u'}^{\mathrm{op}}\| \right) \|\mathbf{p}_{u'}^{\diamond} - \mathbf{p}_{u'}^{\mathrm{op}}\| \\ &\overset{(c)}{\leq} 2\sqrt{P} E_u M \|\mathbf{p}_{u'}^{\diamond} - \mathbf{p}_{u'}^{\mathrm{op}}\|, \end{aligned} \quad (74)$$

where (a) follows from the fact that $\left| |x| - |y| \right| \leq |x - y|$, (b) follows from the compatibility of the Frobenius norm, (c)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI10.1109/TCOMM.2024.3362146, IEEE Transactions on Communications

follows from power constraints for precoders and $\|\mathbf{h}_u\|^2 \leq ME_u$, where $E_u \triangleq \left(\sum_{p=1}^{P_u} \beta_{u,p}\right)^2$. Substituting (74) into (73) yields

$$|\delta_u - 1| \leq 2\sigma_{\mathrm{z}}^{-2}\sqrt{P}E_u \sum_{u' \in \mathcal{Z}_U^+ \setminus u} M\|\mathbf{p}_{u'}^{\diamond} - \mathbf{p}_{u'}^{\mathrm{op}}\|. \tag{75}$$

Similarly,

$$|1-\delta_u^{-1}| \leq 2\sigma_{\mathrm{z}}^{-2}\sqrt{P}E_u \sum_{u' \in \mathcal{Z}_U^+ \setminus u} M\|\mathbf{p}_{u'}^{\diamond} - \mathbf{p}_{u'}^{\mathrm{op}}\|. \tag{76}$$

According to (69), $\lim_{M\to\infty} 2\sigma_{\mathrm{z}}^{-2}\sqrt{P}E_u \sum_{u' \in \mathcal{Z}_U^+ \setminus u} M\|\mathbf{p}_{u'}^{\diamond} - \mathbf{p}_{u'}^{\mathrm{op}}\| = 0$. Combining (72), (75), (76) and the squeeze theorem,

$$\lim_{M\to\infty} \log \delta_u = \lim_{M\to\infty} |\delta_u - 1| = \lim_{M\to\infty} |1 - \delta_u^{-1}| = 0. \tag{77}$$

In a similar manner, we have $\lim_{M\to\infty} \log \overline{\delta}_u = 0$. Therefore, according to (70),

$$\lim_{M\to\infty} \sum_{u=1}^{U} \left(r_u(\mathbf{p}_1^{\mathrm{op}}, \cdots, \mathbf{p}_U^{\mathrm{op}}) - r_u(\mathbf{p}_1^{\diamond}, \cdots, \mathbf{p}_U^{\diamond})\right) = 0. \tag{78}$$

Since $\mathbf{p}_u^{\mathrm{op}}, u \in \mathcal{Z}_U^+$ and $\mathbf{w}_u^{\mathrm{op}}, u \in \mathcal{Z}_U^+$ are the optimizers of (11) and (19), respectively, then

$$\begin{aligned} \sum_{u=1}^{U} r_u(\mathbf{p}_1^{\mathrm{op}}, \cdots, \mathbf{p}_U^{\mathrm{op}}) &\geq \sum_{u=1}^{U} \overline{r}_u(\mathbf{w}_1^{\mathrm{op}}, \cdots, \mathbf{w}_U^{\mathrm{op}}) \\ &\geq \sum_{u=1}^{U} \overline{r}_u(\mathbf{w}_1^{\diamond}, \cdots, \mathbf{w}_U^{\diamond}), \end{aligned} \tag{79}$$

Combining (78), (79) and the relationship $\overline{r}_u(\mathbf{w}_1^{\diamond}, \cdots, \mathbf{w}_U^{\diamond}) = r_u(\mathbf{p}_1^{\diamond}, \cdots, \mathbf{p}_U^{\diamond})$, we have

$$\lim_{M\to\infty} \sum_{u=1}^{U} (r_u(\mathbf{p}_1^{\mathrm{op}}, \cdots, \mathbf{p}_U^{\mathrm{op}}) - \overline{r}_u(\mathbf{w}_1^{\mathrm{op}}, \cdots, \mathbf{w}_U^{\mathrm{op}})) = 0. \tag{80}$$

This concludes the proof.

Appendix B
Proof of Theorem 2

Define $\gamma_u \triangleq \left(1 + \rho_u^{-1}\mathbf{w}_u^{\mathrm{H}}\mathbf{\Theta}_u(\nu_u)\mathbf{w}_u\right)^{-1}$, then $\overline{r}_u^{\mathrm{ub}} = -\log \gamma_u$ is convex over γ_u. According to the first-order condition, in the i-th iteration we have

$$\begin{aligned} -\log \gamma_u &\geq -\log \gamma_u^{(i)} - \left(\gamma_u^{(i)}\right)^{-1}\left(\gamma_u - \gamma_u^{(i)}\right) \\ &= 1 - \log \gamma_u^{(i)} - \left(\gamma_u^{(i)}\right)^{-1}\gamma_u, \end{aligned} \tag{81}$$

where $\gamma_u^{(i)} \triangleq \left(1 + \left(\rho_u^{(i)}\right)^{-1}\left(\mathbf{w}_u^{(i)}\right)^{\mathrm{H}}\mathbf{\Theta}_u(\nu_u)\left(\mathbf{w}_u^{(i)}\right)\right)^{-1}$. Let

$$\begin{aligned} \epsilon_u =& 1 - \frac{\gamma_u^{(i)}}{\rho_u^{(i)}}\left(\left(\mathbf{w}_u^{(i)}\right)^{\mathrm{H}}\mathbf{\Theta}_u(\nu_u)\mathbf{w}_u + \mathbf{w}_u^{\mathrm{H}}\mathbf{\Theta}_u(\nu_u)\mathbf{w}_u^{(i)}\right) \\ &+ \left(\frac{\gamma_u^{(i)}}{\rho_u^{(i)}}\right)^2 \frac{\rho_u}{\gamma_u}\left(\mathbf{w}_u^{(i)}\right)^{\mathrm{H}}\mathbf{\Theta}_u(\nu_u)\mathbf{w}_u^{(i)} \\ \overset{(a)}{=}& \rho_u\gamma_u^{-1}\left(\left(\rho_u^{(i)}\right)^{-1}\gamma_u^{(i)}\mathbf{w}_u^{(i)} - \rho_u^{-1}\gamma_u\mathbf{w}_u\right)^{\mathrm{H}}\mathbf{\Theta}_u(\nu_u) \\ &\times \left(\left(\rho_u^{(i)}\right)^{-1}\gamma_u^{(i)}\mathbf{w}_u^{(i)} - \rho_u^{-1}\gamma_u\mathbf{w}_u\right) + \gamma_u \\ \overset{(b)}{\geq}& \gamma_u, \end{aligned} \tag{82}$$

where (a) follows from $\rho_u^{-1}\gamma_u\mathbf{w}_u^{\mathrm{H}}\mathbf{\Theta}_u(\nu_u)\mathbf{w}_u + \gamma_u = 1$, (b) follows from the fact that $\mathbf{\Theta}_u(\nu_u)$ is semi-positive and the equality holds when $\mathbf{w}_u = \mathbf{w}_u^{(i)}$. Therefore, for any $\mathbf{w}_u^{(i)}$, $\left(\gamma_u^{(i)}\right)^{-1}\epsilon_u \geq \left(\gamma_u^{(i)}\right)^{-1}\gamma_u$. Combine with (81) we have

$$-\log \gamma_u \geq 1 - \log \gamma_u^{(i)} - \left(\gamma_u^{(i)}\right)^{-1}\epsilon_u. \tag{83}$$

Construct

$$\begin{aligned} f =& \sum_{u=1}^{U}\left(1 - \log \gamma_u^{(i)} - \left(\gamma_u^{(i)}\right)^{-1}\epsilon_u\right) \\ \overset{(a)}{=}& \sum_{u=1}^{U} a_u^{(i)} + \sum_{u=1}^{U}\left(\mathbf{q}_u^{(i)}\right)^{\mathrm{H}}\mathbf{w}_u + \sum_{u=1}^{U}\mathbf{w}_u^{\mathrm{H}}\mathbf{q}_u^{(i)} \\ &- \sum_{u=1}^{U}\sum_{u'=1}^{U}\mathbf{w}_u^{\mathrm{H}}\mathbf{T}_{u',u}^{(i)}\mathbf{w}_u, \end{aligned} \tag{84}$$

where (a) follows from $\sum_{u=1}^{U}\sum_{u'=1}^{U}\mathbf{w}_{u'}^{\mathrm{H}}\mathbf{T}_{u,u'}^{(i)}\mathbf{w}_{u'} = \sum_{u=1}^{U}\sum_{u'=1}^{U}\mathbf{w}_u^{\mathrm{H}}\mathbf{T}_{u',u}^{(i)}\mathbf{w}_u$, and

$$\begin{aligned} a_u^{(i)} =& 1 - \log \gamma_u^{(i)} - \left(\gamma_u^{(i)}\right)^{-1} + \sigma_{\mathrm{z}}^2\left(\rho_u^{(i)}\right)^{-1}\left(\gamma_u^{(i)} - 1\right) \\ =& \log\left(1 + \left(\rho_u^{(i)}\right)^{-1}\left(\mathbf{w}_u^{(i)}\right)^{\mathrm{H}}\mathbf{\Theta}_u(\nu_u)\mathbf{w}_u^{(i)}\right) \\ &- \sigma_{\mathrm{z}}^2\left(\rho_u^{(i)}\right)^{-1} + \sigma_{\mathrm{z}}^2\left(\rho_u^{(i)} + \left(\mathbf{w}_u^{(i)}\right)^{\mathrm{H}}\mathbf{\Theta}_u(\nu_u)\mathbf{w}_u^{(i)}\right)^{-1} \\ &- \left(\rho_u^{(i)}\right)^{-1}\left(\mathbf{w}_u^{(i)}\right)^{\mathrm{H}}\mathbf{\Theta}_u(\nu_u)\mathbf{w}_u^{(i)}, \end{aligned} \tag{85a}$$

$$\mathbf{q}_u^{(i)} = \left(\rho_u^{(i)}\right)^{-1}\mathbf{\Theta}_u(\nu_u)\mathbf{w}_u^{(i)}, \tag{85b}$$

$$\begin{aligned} \mathbf{T}_{u,u'}^{(i)} =& \left(\left(\rho_u^{(i)}\right)^{-2}\gamma_u^{(i)}\left(\mathbf{w}_u^{(i)}\right)^{\mathrm{H}}\mathbf{\Theta}_u(\nu_u)\mathbf{w}_u^{(i)}\right)\mathbf{\Theta}_u(\nu_{u'}) \\ =& \left(\left(\rho_u^{(i)}\right)^{-1} - \left(\rho_u^{(i)} + \rho_u^{(i)}\left(\mathbf{w}_u^{(i)}\right)^{\mathrm{H}}\mathbf{q}_u^{(i)}\right)^{-1}\right)\mathbf{\Theta}_u(\nu_{u'}). \end{aligned} \tag{85c}$$

According to (82) and (83) we have

$$f = \sum_{u=1}^{U}\left(1 - \log\gamma_u^{(i)} - \left(\gamma_u^{(i)}\right)^{-1}\epsilon_u\right) \leq -\sum_{u=1}^{U}\log\gamma_u = \sum_{u=1}^{U}\overline{r}_u^{\mathrm{ub}}, \tag{86a}$$

$$f|_{\mathbf{w}_u = \mathbf{w}_u^{(i)}, u\in\mathcal{Z}_U^+} = \sum_{u=1}^{U}\overline{r}_u^{\mathrm{ub}}\big|_{\mathbf{w}_u = \mathbf{w}_u^{(i)}, u\in\mathcal{Z}_U^+}, \tag{86b}$$

respectively. Hence, f is a minorizing function for $\sum_{u=1}^{U}\overline{r}_u^{\mathrm{ub}}$ at the point $\left(\mathbf{w}_1^{(i)}, \cdots, \mathbf{w}_U^{(i)}\right)$. This concludes the proof.

Appendix C
Proof of Theorem 3

The (i,j)-th element of $\mathbf{G}_u(\nu_{u'})$ can be expressed as

$$[\mathbf{G}_u(\nu_{u'})]_{i,j} = \frac{1}{\sqrt{M}}\mathbf{d}_i^{\mathrm{H}}(\nu_{u'})\mathbf{v}(\Omega_{\widetilde{m}_j}),\ i \in \mathcal{Z}_K^+, j \in \mathcal{Z}_{B_u}^+, \tag{87}$$

where $\widetilde{m}_j \in \mathcal{B}_u$ is the beam index which is related to j. Similar to the proof of Theorem 1, as $M \to \infty$, the support set of $V(\Omega_{\widetilde{m}_j}, \nu)$ contains in $\mathcal{L}_u$, where $V(\Omega_{\widetilde{m}_j}, \nu)$ is the DTFT of sequence $\frac{1}{\sqrt{M}}\mathbf{v}(\Omega_{\widetilde{m}_j})$, and

$$\lim_{M\to\infty}\left|\frac{1}{\sqrt{M}}\mathbf{d}_i^{\mathrm{H}}(\nu_{u'})\mathbf{v}(\Omega_{\widetilde{m}_j})\right| = \lim_{M\to\infty}\left|\int_{\mathcal{L}_u} D_i^*(\nu_{u'},\nu)V(\Omega_{\widetilde{m}_j},\nu)\mathrm{d}\nu\right|. \tag{88}$$

When $\mathcal{L}_u \cap \mathcal{M}(\nu_{u'}) = \emptyset$,

$$\left|\int_{\mathcal{L}_u} D_i^*(\nu_{u'},\nu)V(\Omega_{\widetilde{m}_j},\nu)\mathrm{d}\nu\right|^2$$

$$\leq \left(\int_{\mathcal{L}_u} |D_i(\nu_{u'},\nu)|^2 \mathrm{d}\nu\right)\left(\int_{\mathcal{L}_u} |V(\Omega_{\widetilde{m}_j},\nu)|^2 \mathrm{d}\nu\right)$$
$$\leq \left(1-\int_{\mathcal{M}(\nu_{u'})} |D_i(\nu_{u'},\nu)|^2 \mathrm{d}\nu\right)\left(\int_{-\frac{1}{2}}^{\frac{1}{2}} |V(\Omega_{\widetilde{m}_j},\nu)|^2 \mathrm{d}\nu\right)$$
$$\overset{(a)}{=} 1-\lambda_i, \tag{89}$$

where (a) follows from the fact that $\int_{-\frac{1}{2}}^{\frac{1}{2}} |V(\Omega_{\widetilde{m}_j},\nu)|^2 \mathrm{d}\nu = \frac{1}{M}\mathbf{v}^{\mathrm{H}}(\Omega_{\widetilde{m}_j})\mathbf{v}(\Omega_{\widetilde{m}_j}) = 1$. According to [18] and [36, Lemma 4.1], when $i \leq K = \lfloor 2MW(1-\epsilon) \rfloor$, $0 < \epsilon < 1$ and $M \geq \overline{M}_{W,\epsilon}$, we have the following inequality

$$\left|\int_{\mathcal{L}_u} D_i^*(\nu_{u'},\nu) V(\Omega_{\widetilde{m}_j},\nu)\mathrm{d}\nu\right| \leq \sqrt{1-\lambda_i} \leq e^{-\overline{C}_{W,\epsilon}M/2}, \tag{90}$$

where $\overline{C}_{W,\epsilon}$ and $\overline{M}_{W,\epsilon}$ are positive constants depend on W and ϵ. Therefore, for $i \leq K$,

$$\lim_{M\to\infty}[\mathbf{G}_u(\nu_{u'})]_{i,j} = \lim_{M\to\infty}\left|\int_{\mathcal{L}_u} D_i^*(\nu_{u'},\nu)V(\Omega_{\widetilde{m}_j},\nu)\mathrm{d}\nu\right| = \lim_{M\to\infty} e^{-\overline{C}_{W,\epsilon}M/2} = 0 \tag{91}$$

Then $\lim_{M\to\infty}\mathbf{G}_u(\nu_{u'}) = \mathbf{O}_{K\times B_u}$, This concludes the proof.

REFERENCES

[1] C. D. Alwis, A. Kalla, Q.-V. Pham, P. Kumar, K. Dev, W.-J. Hwang, and M. Liyanage, "Survey on 6G Frontiers: Trends, Applications, Requirements, Technologies and Future Research," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 836–886, Apr. 2021.

[2] E. E. Johnson, E. Koski, and W. N. Furman, *Third-Generation and Wideband HF Radio Communications*. Norwood, MA: Artech House, 2013.

[3] J. Zhang, J. Zhang, D. W. K. Ng, S. Jin, and B. Ai, "Improving sum-rate of cell-free massive MIMO with expanded compute-and-forward," *IEEE Trans. Signal Process.*, vol. 70, pp. 202–215, Nov. 2022.

[4] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[5] C. Sun, X. Q. Gao, S. Jin, M. Matthaiou, Z. Ding, and C. Xiao, "Beam division multiple access transmission for massive MIMO communications," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2170–2184, Jun. 2015.

[6] X. L. Yu, A.-A. Lu, X. Q. Gao, G. Y. Li, G. Ding, and C.-X. Wang, "HF skywave massive MIMO communication," *IEEE Trans. Wireless Commun.*, vol. 21, no. 4, pp. 2769–2785, Oct. 2021.

[7] J. Mastrangelo, J. Lemmon, L. Vogler, J. Hoffmeyer, L. Pratt, and C. Behm, "A new wideband high frequency channel simulation system," *IEEE Trans. Commun.*, vol. 45, no. 1, pp. 26–34, Jan. 1997.

[8] M. Sadek, A. Tarighat, and A. H. Sayed, "A leakage-based precoding scheme for downlink multi-user MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1711–1721, May 2007.

[9] S. S. Christensen, R. Agarwal, E. De Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.

[10] X. Gan, C. Zhong, C. Huang, and Z. Zhang, "RIS-assisted multi-user MISO communications exploiting statistical CSI," *IEEE Trans. Commun.*, vol. 69, no. 10, pp. 6781–6792, Oct. 2021.

[11] J. Park and B. Clerckx, "Multi-user linear precoding for multi-polarized massive MIMO system under imperfect CSIT," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2532–2547, May 2015.

[12] T. X. Tran and K. C. Teh, "Spectral and energy efficiency analysis for SLNR precoding in massive MIMO systems with imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 4017–4027, Jun. 2018.

[13] S. Qiu, D. Chen, D. Qu, K. Luo, and T. Jiang, "Downlink precoding with mixed statistical and imperfect instantaneous CSI for massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3028–3041, Apr. 2018.

[14] A.-A. Lu, X. Q. Gao, W. Zhong, C. Xiao, and X. Meng, "Robust transmission for massive MIMO downlink with imperfect CSI," *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5362–5376, Aug. 2019.

[15] C. Wang, A.-A. Lu, X. Q. Gao, and Z. Ding, "Robust precoding for 3D massive MIMO configuration with matrix manifold optimization," *IEEE Trans. Wireless Commun.*, vol. 21, no. 5, pp. 3423–3437, May 2022.

[16] J. Shi, W. Wang, X. Yi, X. Q. Gao, and G. Y. Li, "Deep learning-based robust precoding for massive MIMO," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7429–7443, Nov. 2021.

[17] X. L. Yu, X. Q. Gao, A.-A. Lu, J. Zhang, W. Hebing, and G. Y. Li, "Robust precoding for HF skywave massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 22, no. 10, pp. 6691–6705, Oct. 2023.

[18] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty—V: The discrete case," *Bell Syst. Tech. J.*, vol. 57, no. 5, pp. 1371–1430, May-Jun. 1978.

[19] H. A. A. Saleh, A. F. Molisch, T. Zemen, S. D. Blostein, and N. B. Mehta, "Receive antenna selection for time-varying channels using discrete prolate spheroidal sequences," *IEEE Trans. Wireless Commun.*, vol. 11, no. 7, pp. 2616–2627, Jul 2012.

[20] O. Longoria-Gandara and R. Parra-Michel, "Estimation of correlated MIMO channels using partial channel state information and DPSS," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3711–3719, Nov. 2011.

[21] J. K. Tugnait and S. He, "Multiuser/MIMO doubly selective fading channel estimation using superimposed training and Slepian sequences," *IEEE Trans. Veh. Technol.*, vol. 59, no. 3, pp. 1341–1354, Mar. 2010.

[22] T. Zemen and C. Mecklenbrauker, "Time-variant channel estimation using discrete prolate spheroidal sequences," *IEEE Trans. Signal Process.*, vol. 53, no. 9, pp. 3597–3607, Sep. 2005.

[23] G. Taubock, F. Hlawatsch, D. Eiwen, and H. Rauhut, "Compressive estimation of doubly selective channels in multicarrier systems: Leakage effects and sparsity-enhancing processing," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 255–271, Apr. 2010.

[24] S. Sreepada and S. Kalyani, "Channel estimation in OFDM systems with virtual subcarriers using DPSS," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2462–2465, Dec. 2016.

[25] Z. Qin, J. Wang, J. Chen, G. Ding, Y.-D. Yao, X. Ji, and X. Chen, "Link quality analysis based channel selection in high-frequency asynchronous automatic link establishment: A matrix completion approach," *IEEE Syst. J.*, vol. 12, no. 2, pp. 1957–1968, Jun. 2018.

[26] B. Wang, F. Gao, S. Jin, H. Lin, and G. Y. Li, "Spatial-wideband effect in massive MIMO systems," in *Proc. 23rd Asia-Pacific Conf. Commun.*, Perth, WA, Dec. 2017, pp. 1–6.

[27] J. Gao, M. Hu, C. Zhong, G. Y. Li, and Z. Zhang, "An attention-aided deep learning framework for massive MIMO channel estimation," *IEEE Trans. Wireless Commun.*, vol. 21, no. 3, pp. 1823–1835, Mar. 2022.

[28] D. Shi, L. Song, W. Zhou, X. Q. Gao, C.-X. Wang, and G. Y. Li, "Channel acquisition for HF skywave massive MIMO-OFDM communications," *IEEE Trans. Wireless Commun.*, vol. 22, no. 6, pp. 4074–4089, Jun. 2023.

[29] K. Mamat and W. Santipach, "On optimizing feedback interval for temporally correlated MIMO channels with transmit beamforming and finite-rate feedback," *IEEE Trans. Commun.*, vol. 66, no. 8, pp. 3407–3419, Aug. 2018.

[30] L. Lian, A. Liu, and V. K. N. Lau, "Exploiting dynamic sparsity for downlink FDD-massive MIMO channel tracking," *IEEE Trans. Signal Process.*, vol. 67, no. 8, pp. 2007–2021, Apr. 2019.

[31] A. L. Moustakas, G. C. Alexandropoulos, A. Polydoros, I. Kaddas, and I. Dagres, "Impact of Imperfect Channel Estimation in HF OFDM-MIMO Communications," in *2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Istanbul, Turkey, 2019, pp. 1–6.

[32] L. Song, D. Shi, L. Gan, and X. Q. Gao, "Slepian transform based detectors for HF skywave massive MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 72, no. 10, pp. 13 106–13 119, Oct. 2023.

[33] F. Vaida, "Parameter convergence for EM and MM algorithms," *Statistica Sinica*, pp. 831–840, 2005.

[34] D. B. Percival, A. T. Walden *et al.*, *Spectral Analysis for Physical Applications*. Cambridge University Press, 1993.

[35] S. Karnik, Z. Zhu, M. B. Wakin, J. Romberg, and M. A. Davenport, "The fast Slepian transform," *Appl. Comput. Harmon. Anal.*, vol. 46, no. 3, pp. 624–652, 2019.

[36] M. A. Davenport and M. B. Wakin, "Compressive sensing of analog signals using discrete prolate spheroidal sequences," *Appl. Comput. Harmon. Anal.*, vol. 33, no. 3, pp. 438–472, Feb. 2012.

[37] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 794–816, Feb. 2017.

[38] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[39] R. Hunger, *Floating Point Operations in Matrix-vector Calculus*. Associate Institute for Signal Processing, Tech. Rep., 2007.
[40] *Proplab-pro version 3*, Solar Terrestrial Dispatch. [Online]. Available: https://shop.spacew.com/index.php/product/proplab-pro-hf-radio-propagation-laboratory/
[41] J. M. Steele, *The Cauchy-Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities*. Cambridge University Press, 2004.

Linfeng Song (Graduate Student Member, IEEE) received the B.E. degree in radio wave propagation and antenna from Wuhan University, Wuhan, China, in 2020. He received the M.E. degree with the National Mobile Communications Research Laboratory, Southeast University, Nanjing, China, in 2023. His research interests include massive MIMO communications and HF communications.

Ding Shi (Graduate Student Member, IEEE) received the B.E. degree in information science and engineering from Southeast University, Nanjing, China, in 2018. He is currently pursuing the Ph.D. degree with the National Mobile Communications Research Laboratory, Southeast University, Nanjing, China. From November 2022 to November 2023, he was a visiting student at the Division of Information Science and Engineering, KTH Royal Institute of Technology, Stockholm, Sweden. His research interests include massive MIMO communications, HF communications, and signal processing for wireless communications.

Lu Gan (Senior Member, IEEE) received her B. Eng and M. Eng. degrees from Southeast University, China and the Ph.D degree from Nanyang Technological University, Singapore in 1998, 2000 and 2004 respectively. She is currently a senior lecturer in Brunel University, London. Before she joined Brunel University in 2008, she has been on the faculties with The University of Newcastle (2004-2006), Australia and University of Liverpool, UK (2006-2007). From 2003 to 2004, she was a research associate in Centre for Signal Processing, Nanyang Technological University. Dr. Gan's research interests include fundamental signal processing theories and their applications in image/video coding and processing, non-destructive terahertz and ultrasound imaging, machine learning and wireless communications etc. She is a senior member of IEEE and a regular reviewer for many IEEE journals.

Xiqi Gao (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Southeast University, Nanjing, China, in 1997.

He joined the Department of Radio Engineering, Southeast University, in April 1992. Since May 2001, he has been a professor of information systems and communications. From September 1999 to August 2000, he was a visiting scholar at Massachusetts Institute of Technology, Cambridge, and Boston University, Boston, MA. From August 2007 to July 2008, he visited the Darmstadt University of Technology, Darmstadt, Germany, as a Humboldt scholar. His current research interests include broadband multicarrier communications, massive MIMO wireless communications, satellite communications, optical wireless communications, information theory and signal processing for wireless communications. From 2007 to 2012, he served as an Editor for the IEEE Transactions on Wireless Communications. From 2009 to 2013, he served as an Associate Editor for the IEEE Transactions on Signal Processing. From 2015 to 2017, he served as an Editor for the IEEE Transactions on Communications.

Dr. Gao received the Science and Technology Awards of the State Education Ministry of China in 1998, 2006 and 2009, the National Technological Invention Award of China in 2011, the Science and Technology Award of Jiangsu Province of China in 2014, and the 2011 IEEE Communications Society Stephen O. Rice Prize Paper Award in the field of communications theory.