Hybrid Beamformer Optimization for Coherent Signal Detection

Saidur R. Pavel †, Yimin D. Zhang †, and Braham Himed ‡

† Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA 19122, USA
‡ Passive RF Systems Branch, Air Force Research Laboratory, WPAFB, OH 45433, USA

***Abstract*—Massive multiple-input multiple-output (MIMO) systems with hybrid beamforming have emerged as a promising technology for future wireless communication and sensing, offering an effective tradeoff between hardware complexity and system performance. However, detecting coherent signals under hybrid beamforming architectures remains challenging, primarily due to the inaccessibility of full-dimensional analog array signals and the need for specific data structures to decorrelate signals. This paper addresses the problem using a subarray-based hybrid beamforming architecture designed for coherent signal detection. In this approach, the antenna array is divided into multiple overlapping subarrays, and the hybrid beamforming matrix is optimized using a mutual information maximization criterion. To further reduce complexity, forward-backward spatial smoothing is exploited, thereby lowering the number of required subarrays for effective implementation.**

Keywords*—*Hybrid beamforming, coherent signal, subarray-based hybrid beamforming, massive MIMO, mutual information.

I. Introduction

Massive multiple-input multiple-output (MIMO) with large antenna arrays is a key enabling technology for next-generation wireless communication and sensing. These systems offer orders-of-magnitude improvements in spectral and energy efficiency, capacity, and reliability through relatively simple signal processing [1]–[7]. Massive MIMO is also essential for millimeter-wave (mmWave) communication and radar systems. The high beamforming gain attainable in massive MIMO ensures sufficient signal-to-noise ratio (SNR) to counteract the severe path loss of mmWave channels, thereby enabling access to high-frequency spectrum and improving spectral efficiency [8]–[10]. In massive MIMO systems, conventional digital beamforming requires one dedicated radio-frequency (RF) chain per antenna, typically consisting of high-resolution analog-to-digital converters (ADCs), mixers, filters, and amplifiers. This architecture incurs significant hardware cost and power consumption, motivating the adoption of hybrid beamforming.

Hybrid beamforming substantially reduces the number of required RF chains by partitioning the beamforming operation into two stages: high-dimensional analog beamforming and low-dimensional digital baseband beamforming [11]–[19]. The analog beamforming stage is typically realized using phase shifters, while the digital stage operates on a reduced signal dimension. Optimizing the analog beamformer in hybrid beamforming systems, also referred to as compressive measurement matrix (CMM) [19], poses significant challenges. Many existing approaches rely on prior knowledge of the direction-of-arrival (DOA) or channel state information (CSI) to optimize the analog beamformer. In practice, however, such information is typically unavailable. This limitation arises because the impinging analog signals at the massive MIMO receiver cannot be directly accessed. Rather, only the digitally processed outputs of the analog beamformer are observable. Consequently, the CMM needs to be optimized without access to the full-dimensional analog signals. To address this problem, information-theoretic methods have been proposed in [11], [17], [20], where the CMM is optimized by maximizing the mutual information between the compressed measurements obtained after analog beamforming and the underlying DOA distribution.

The existing methods for CMM optimization assume uncorrelated signals. In practice, however, coherent or correlated signals frequently arise as a result of effects such as multipath propagation in wireless communication systems and low grazing-angle scattering in radar sensing. In the presence of coherent signals, the covariance matrix of the received data becomes rank-deficient, causing conventional beamforming and DOA estimation algorithms, such as minimum variance distortionless response (MVDR), multiple signal classification (MUSIC), and estimation of signal parameters via rotational invariance techniques (ESPRIT), to fail. To address this issue, various decorrelation techniques, such as spatial smoothing-based methods [21]–[23] and Toeplitz matrix and tensor reconstruction [24]–[28], have been extensively studied. Nevertheless, in the context of hybrid beamforming, effective decorrelation of coherent signals is not adequately explored. In hybrid beamforming systems, the decorrelation process becomes more challenging due to the inaccessibility of the full-dimensional analog signal and the need for specific data structures to decorrelate signals. As we show in this paper, the conventional fully connected hybrid beamforming (FHBF) architecture developed for uncorrelated signals is unable to adequately handle coherent signals.

In addressing such challenges, this paper develops a subarray-based hybrid beamforming (SHBF) scheme to effectively handle coherent signal detection. Although the proposed SHBF scheme is applicable to broader array geometries, this work focuses on a uniform linear array (ULA). In this architecture, a ULA is divided into overlapping subarrays, and analog beamforming is applied independently to each subarray using a shared CMM. This yields parallel compressed measurements from the subarrays that retain the translational invariance property for effective signal decorrelation through

This material is based upon work supported by the Air Force Office of Scientific Research under award number FA9550-23-1-0255. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the United States Air Force.

spatial smoothing. In this approach, the CMM is optimized by maximizing the mutual information between the compressed measurements and the signal DOAs. The compressed baseband measurements from each subarray are then employed to decorrelate the covariance matrix using forward–backward spatial smoothing (FBSS). The resulting decorrelated covariance matrix is used to estimate the spatial spectrum, whose normalized form is treated as the posterior DOA distribution to optimize the CMM in an iterative manner. The proposed approach enables coherent signal detection while requiring significantly fewer RF chains compared to a fully digital scheme.

Notations: Bold lowercase letters are used to denote vectors, while bold uppercase letters denote matrices. In particular, $\boldsymbol{I}_{M\times M}$ and $\mathbf{0}_{M\times N}$ represent the $M\times M$ identity matrix and the $M\times N$ zero matrix, respectively. The operators $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ indicate the transpose and Hermitian transpose of a vector or matrix, respectively. The notation $\mathrm{Diag}(\cdot)$ denotes a diagonal matrix constructed from the elements of a given vector, and $\lceil\cdot\rceil$ denotes the ceiling operation.

II. Signal Model

Consider K narrowband far-field coherent signals $\boldsymbol{s}(t)=[s_1(t),s_2(t),\cdots,s_K(t)]^{\mathrm{T}}\in\mathbb{C}^K$ impinging from DOAs $\boldsymbol{\theta}=[\theta_1,\theta_2,\cdots,\theta_K]^{\mathrm{T}}$ on a massive MIMO system equipped with a uniform linear array of N omnidirectional antennas. By taking the first waveform $s_1(t)$ as a reference, the remaining signals can be expressed as

$$s_k(t)=\alpha_k s_1(t), \tag{1}$$

where α_k is a complex scalar with $\alpha_1=1$. Accordingly, the analog RF signal vector received at time t is represented as

$$\begin{aligned}\boldsymbol{x}^{\mathrm{RF}}(t)&=s_1(t)\sum_{k=1}^{K}\alpha_k\boldsymbol{a}(\theta_k)e^{j\omega_c t}+\boldsymbol{n}^{\mathrm{RF}}(t)\\&=\boldsymbol{A}\boldsymbol{s}(t)e^{j\omega_c t}+\boldsymbol{n}^{\mathrm{RF(t)}},\end{aligned} \tag{2}$$

where $\boldsymbol{a}(\theta_k)=[1,e^{-j\frac{2\pi}{\lambda}d\sin\theta_k},\cdots,e^{-j\frac{2\pi}{\lambda}(N-1)d\sin\theta_k}]^{\mathrm{T}}\in\mathbb{C}^{N\times 1}$ denotes the steering vector of the kth signals, $\boldsymbol{A}=[\boldsymbol{a}(\theta_1),\boldsymbol{a}(\theta_2),\cdots,\boldsymbol{a}(\theta_K)]\in\mathbb{C}^{N\times K}$ is the array manifold matrix, ω_c represents the angular frequency of the carrier, and $\boldsymbol{n}^{\mathrm{RF(t)}}\sim\mathcal{CN}(\mathbf{0},\sigma_n^2\boldsymbol{I})$ is the additive white Gaussian noise (AWGN) vector with noise power σ_n^2.

An FHBF architecture developed for uncorrelated signals is depicted in Fig. 1 [11], [17], [20]. In this structure, $M\ll N$ RF chains are connected to each antenna through analog multipliers which are typically implemented by a series of phase shifters. The output RF signal at the mth RF chain, $y_m^{\mathrm{RF}}(t)$, is obtained by analog multiplication of the input RF signal $\boldsymbol{x}^{\mathrm{RF}}(t)$ and the analog beamforming vector $\boldsymbol{\phi}_m\in\mathbb{C}^{N\times 1}$ as

$$y_m^{\mathrm{RF}}(t)=\sum_{n=1}^{N}\phi_{m,n}x_n^{\mathrm{RF}}(t), \tag{3}$$

where $\phi_{m,n}$ is the nth element of $\boldsymbol{\phi}_m$, denoting the analog beamforming coefficient for the nth antenna and the mth RF chain.

Denoting $\boldsymbol{x}(t)$ as the baseband signal corresponding to $\boldsymbol{x}^{\mathrm{RF}}(t)$, the digital baseband signal obtained from M RF chains can be expressed as

$$\boldsymbol{y}(t)=\boldsymbol{\Phi}\boldsymbol{x}(t)=\boldsymbol{\Phi}\boldsymbol{A}\boldsymbol{s}(t)+\boldsymbol{\Phi}\boldsymbol{n}(t), \tag{4}$$

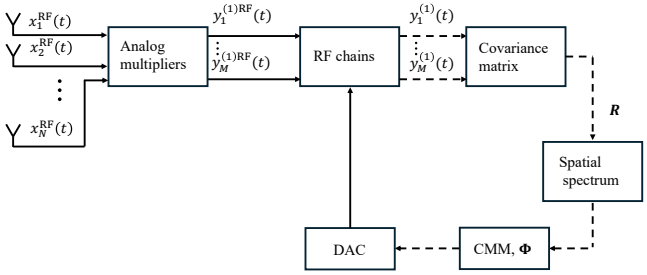


Fig. 1: Fully connected hybrid beamforming structure.

where $\boldsymbol{\Phi}=[\boldsymbol{\phi}_1,\cdots,\boldsymbol{\phi}_M]^{\mathrm{T}}\in\mathbb{C}^{M\times N}$ denotes the analog beamforming matrix subject to row-orthonormal constraint $\boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{H}}=\boldsymbol{I}$. Due to the presence of coherent signals, the source covariance matrix, i.e.,

$$\boldsymbol{R}_s=\mathbb{E}[\boldsymbol{s}(t)\boldsymbol{s}^{\mathrm{H}}(t)]=\sigma_s^2\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{H}}, \tag{5}$$

is rank one, where $\boldsymbol{\alpha}=[\alpha_1,\alpha_2,\cdots,\alpha_K]^{\mathrm{T}}$ and σ_s^2 represents the power of signal $s_1(t)$. Consequently, the noise-free covariance matrix of $\boldsymbol{y}(t)$, given by

$$\boldsymbol{R}_y=\mathbb{E}[\boldsymbol{y}(t)\boldsymbol{y}^{\mathrm{H}}(t)]=\boldsymbol{\Phi}\boldsymbol{A}\boldsymbol{R}_s\boldsymbol{A}^{\mathrm{H}}\boldsymbol{\Phi}^{\mathrm{H}}, \tag{6}$$

is also rank one. Therefore, digital beamforming and DOA estimation methods, such as MVDR, MUSIC, and ESPRIT, cannot be directly applied to the baseband outputs of the analog beamformer. Furthermore, due to the inaccessibility of the full dimensional signal $\boldsymbol{x}(t)$, it is not possible to implement spatial smoothing-based processing in this FHBF architecture.

III. SHBF Architecture for Coherent Signals

In this section, the challenge of applying spatial smoothing within the FHBF architecture is first discussed, which then motivates the development of the proposed SHBF architecture.

A. Limitations of FHBF for Coherent Signals

In this subsection, we first examine the feasibility of applying spatial smoothing in the FHBF structure, developed for signal dimension reduction of uncorrelated signals. Since the full-dimensional received signal vector $\boldsymbol{x}(t)$ is not directly accessible under this architecture, we begin by describing spatial smoothing as if $\boldsymbol{x}(t)$ were available, in order to clarify the conditions required for its implementation. In this case, vector $\boldsymbol{x}(t)$ is partitioned into Q overlapping subarrays of size L. The data corresponding to the qth subarray can be extracted by selecting the appropriate rows from the full array as

$$\boldsymbol{x}^{(q)}(t)=\boldsymbol{H}_x^{(q)}\boldsymbol{x}(t)=\boldsymbol{H}_x^{(q)}\boldsymbol{A}\boldsymbol{s}(t)+\boldsymbol{H}_x^{(q)}\boldsymbol{n}(t), \tag{7}$$

where $\boldsymbol{H}_x^{(q)}\in\{0,1\}^{L\times N}$ is a selection matrix defined as

$$\boldsymbol{H}_x^{(q)}=\left[\mathbf{0}_{L\times(q-1)}\ \boldsymbol{I}_{L\times L}\ \mathbf{0}_{L\times(N-L-q+1)}\right]. \tag{8}$$

The signal component of the covariance matrix for the qth subarray can be expressed as

$$\boldsymbol{R}_x^{(q)}=\mathbb{E}[\boldsymbol{x}^{(q)}(t)\boldsymbol{x}^{(q)^{\mathrm{H}}}(t)]=\sigma_s^2\boldsymbol{H}_x^{(q)}\boldsymbol{A}\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{H}_x{}^{(q)})^{\mathrm{H}}. \tag{9}$$

By averaging over the Q subarrays, the spatially smoothed covariance matrix becomes

$$\boldsymbol{R}_{dx}=\frac{\sigma_s^2}{Q}\sum_{q=1}^{Q}\boldsymbol{H}_x^{(q)}\boldsymbol{A}\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{H}_x{}^{(q)})^{\mathrm{H}}. \tag{10}$$

Now, since the array manifold matrix exhibits a translational invariance property, the array manifold matrix of the qth subarray satisfies

$$\boldsymbol{A}^{(q)} = \boldsymbol{H}_x^{(q)}\boldsymbol{A} = \boldsymbol{A}^{(1)}\boldsymbol{\Gamma}^{q-1}, \tag{11}$$

where $\boldsymbol{\Gamma} = \mathrm{Diag}([e^{-j\frac{2\pi}{\lambda}d\sin\theta_1}, \cdots, e^{-j\frac{2\pi}{\lambda}d\sin\theta_K}])$ and $\boldsymbol{A}^{(1)}$ is the array manifold matrix of the first subarray. This property maintains the Vandermonde structure of the array manifold matrix. Therefore, Eq. (10) can be rewritten as

$$\boldsymbol{R}_{dx} = \frac{\sigma_s^2}{Q}\boldsymbol{A}^{(1)}\left(\sum_{q=1}^{Q}\boldsymbol{\Gamma}^{q-1}\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{H}}(\boldsymbol{\Gamma}^{q-1})^{\mathrm{H}}\right)(\boldsymbol{A}^{(1)})^{\mathrm{H}}, \tag{12}$$

where the term in parentheses corresponds to the source covariance matrix after smoothing. This formulation restores the rank to K provided that $Q \geq K$. Furthermore, the columns of $\boldsymbol{R}_{dx}$ span the same subspace as $\boldsymbol{A}^{(1)}$, since $\boldsymbol{R}_{dx}$ can be factored as $\boldsymbol{A}^{(1)}\Delta(\boldsymbol{A}^{(1)})^{\mathrm{H}}$, where $\Delta = \sum_{q=1}^{Q}\boldsymbol{\Gamma}^{q-1}\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{H}}(\boldsymbol{\Gamma}^{q-1})^{\mathrm{H}}$ is full rank.

After compression, the high-dimensional covariance matrix $\boldsymbol{R}_{dx}$ yields the compressed covariance matrix as

$$\boldsymbol{R}_{dy_f} = \boldsymbol{\Phi}\boldsymbol{R}_{dx}\boldsymbol{\Phi}^{\mathrm{H}}, \tag{13}$$

which retains the same structure as in the uncorrelated case and can therefore be employed for spatial spectrum estimation using digital beamforming. To distinguish between the data obtained from FHBF and SHBF, the subscripts f and s are used in $\boldsymbol{y}$ to denote the FHBF and SHBF cases, respectively.

It is important to note that, in practice, the full-dimensional signal $\boldsymbol{x}(t)$ is not directly accessible. Consequently, spatial smoothing cannot be applied prior to compression, and decorrelation is not feasible in this setting.

An alternative solution is to apply smoothing directly on the compressed measurements $\boldsymbol{y}_f(t)$ by partitioning them into Q subarrays. In this case, the data for the qth subarray is given by

$$\boldsymbol{y}_f^{(q)}(t) = \boldsymbol{H}_y^{(q)}\boldsymbol{\Phi}\boldsymbol{x}(t) = \boldsymbol{H}_y^{(q)}\boldsymbol{\Phi}\boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{H}_y^{(q)}\boldsymbol{\Phi}\boldsymbol{n}(t), \tag{14}$$

where $\boldsymbol{H}_y^{(q)} \in \{0,1\}^{L\times M}$ is the selection matrix for the compressed domain, analogous to the uncompressed case. The signal component of the covariance matrix for $\boldsymbol{y}_f^{(q)}$ can then be expressed as

$$\begin{aligned}\boldsymbol{R}_{y_f}^{(q)} &= \sigma_s^2\boldsymbol{H}_y^{(q)}\boldsymbol{\Phi}\boldsymbol{A}\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{\Phi}^{\mathrm{H}}(\boldsymbol{H}_y{}^{(q)})^{\mathrm{H}} \\ &= \sigma_s^2\boldsymbol{H}_y^{(q)}\boldsymbol{B}\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{H}}\boldsymbol{B}^{\mathrm{H}}(\boldsymbol{H}_y{}^{(q)})^{\mathrm{H}},\end{aligned} \tag{15}$$

where $\boldsymbol{B} = \boldsymbol{\Phi}\boldsymbol{A}$. It is important to note that, unlike the uncompressed case, $\boldsymbol{B}^{(q)}$ does not satisfy the translational-invariance property, i.e.,

$$\boldsymbol{B}^{(q)} = \boldsymbol{H}_y^{(q)}\boldsymbol{B} \neq \boldsymbol{B}^{(1)}\tilde{\boldsymbol{\Gamma}}^{q-1}, \tag{16}$$

for some diagonal matrix $\tilde{\boldsymbol{\Gamma}}$. This destroys the Vandermonde structure of the array manifold and prevents spatial smoothing from restoring the signal subspace rank. As a result, averaging

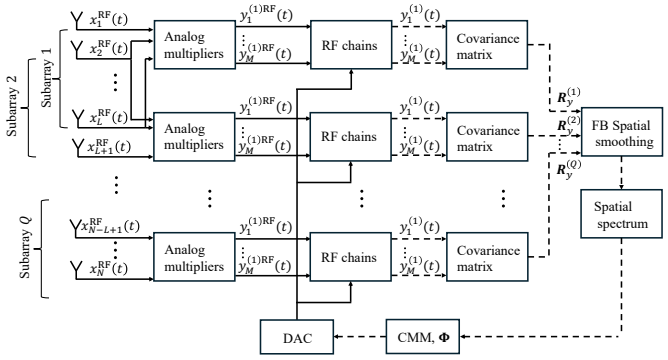


Fig. 2: Subarray-based hybrid beamforming structure for coherent signal processing.

over Q subarrays results

$$\begin{aligned}\boldsymbol{R}_{dy_f} &= \frac{\sigma_s^2}{Q}\sum_{q=1}^{Q}\boldsymbol{H}_y^{(q)}\boldsymbol{\Phi}\boldsymbol{A}\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{\Phi}^{\mathrm{H}}(\boldsymbol{H}_y{}^{(q)})^{\mathrm{H}} \\ &= \frac{\sigma_s^2}{Q}\sum_{q=1}^{Q}\boldsymbol{H}_y^{(q)}\boldsymbol{z}\boldsymbol{z}^{\mathrm{H}}(\boldsymbol{H}_y{}^{(q)})^{\mathrm{H}}.\end{aligned} \tag{17}$$

In this case, the columns of $\boldsymbol{R}_{dy_f}$ are contained in the span of the sliding versions of the same vector $\boldsymbol{z} = \boldsymbol{\Phi}\boldsymbol{A}\boldsymbol{\alpha}$. Hence, $\boldsymbol{R}_{dy}$ cannot span the desired signal subspace.

B. Proposed SHBF architecture

From the previous discussion, it becomes evident that spatial smoothing–based decorrelation techniques cannot be directly applied in the FHBF structure. To address this limitation, we propose an SHBF architecture designed for coherent signal detection, as illustrated in Fig. 2. In this configuration, the N-antenna array is partitioned into Q overlapping subarrays, each of size $L = N - Q + 1$, which are connected to M RF chains. Consequently, the total number of RF chains required in this structure is QM. Due to this subarray-based arrangement, analog beamforming is carried out separately for each subarray using the same CMM $\boldsymbol{\Phi} \in \mathbb{C}^{M\times L}$. Consequently, in the digital baseband domain, the compressed measurement vectors corresponding to all subarrays are obtained simultaneously.

Note that the SHBF configuration and processing described here differ from those in the non-overlapping subarray-based S-HBF structure, which handles uncorrelated signals with reduced complexity [29].

Since the full-dimensional signal is not accessible, decorrelation must be performed in the baseband domain. Unlike the FHBF case, where compressed measurements are partitioned into subarrays, the SHBF architecture naturally provides subarray measurements corresponding to the actual physical subarrays. The compressed measurement vector associated with the qth subarray is expressed as

$$\boldsymbol{y}_s^{(q)}(t) = \boldsymbol{\Phi}\boldsymbol{x}^{(q)}(t) = \boldsymbol{\Phi}\boldsymbol{A}^{(q)}\boldsymbol{s}(t) + \boldsymbol{\Phi}\boldsymbol{n}^{(q)}(t), \tag{18}$$

where $\boldsymbol{n}^{(q)}(t)$ is the noise term associated with the qth subarray. Since the array manifold matrix exhibits the translational invariance property, Eq. (18) can be rewritten as

$$\boldsymbol{y}_s^{(q)}(t) = \boldsymbol{\Phi}\boldsymbol{A}^{(1)}\boldsymbol{\Gamma}^{q-1}\boldsymbol{s}(t) + \boldsymbol{\Phi}\boldsymbol{n}^{(q)}(t), \tag{19}$$

with the signal component of the covariance matrix being

$$\boldsymbol{R}_{y_s}^{(q)} = \sigma_s^2 \boldsymbol{\Phi} \boldsymbol{A}^{(1)} \boldsymbol{\Gamma}^{q-1} \boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{H}} \boldsymbol{\Gamma}^{q-1\mathrm{H}} (\boldsymbol{A}^{(1)})^{\mathrm{H}} \boldsymbol{\Phi}^{\mathrm{H}}. \tag{20}$$

By averaging over Q subarrays, the smoothed covariance matrix is obtained as as

$$\boldsymbol{R}_{dy_s} = \frac{\sigma_s^2}{Q} \boldsymbol{\Phi} \boldsymbol{A}^{(1)} \left(\sum_{q=1}^{Q} \boldsymbol{\Gamma}^{q-1} \boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{H}} \boldsymbol{\Gamma}^{q-1\mathrm{H}} \right) (\boldsymbol{A}^{(1)})^{\mathrm{H}} \boldsymbol{\Phi}^{\mathrm{H}}, \tag{21}$$

which is identical to the expression in Eq. (13). This demonstrates that the proposed SHBF architecture achieves effective decorrelation through spatial smoothing in the baseband, without requiring access to the full-dimensional received signal.

Note that the rank of the decorrelated covariance matrix $\boldsymbol{R}_{dy_s}$ is K, provided that $K \leq Q$. Therefore, to detecet K coherent sources, at least KM RF chains are required. To further reduce the number of required subarrays, backward smoothing is applied by utilizing flipped and conjugated subarray measurements. With the forward–backward smoothing procedure, the rank of the decorrelated covariance matrix becomes $\min(K, 2Q)$, which equals K when $K \leq 2Q$. Therefore, detecting K coherent signals requires at least $Q = \lceil \frac{K}{2} \rceil$ subarrays, corresponding to $\lceil \frac{K}{2} \rceil M$ RF chains.

C. Optimization of subarray CMM

This subsection describes the optimization of the subarray CMM $\boldsymbol{\Phi} \in \mathbb{C}^{M \times L}$ based on the mutual information maximization criterion as developed in [11], [17]. Since the true DOA distribution is unknown, a uniform initial prior is assumed for the DOA θ. The probability density function (pdf) of the compressed measurement for the qth subarray is given as

$$f(\boldsymbol{y}^{(q)}) = \int_{\theta \in \Theta} f(\boldsymbol{y}^{(q)}|\theta) f(\theta) \mathrm{d}\theta, \tag{22}$$

where $f(\theta)$ denotes the pdf of θ, and Θ represents the angular observation region. Discretizing the pdf $f(\theta)$ into K angular bins with equal width $\Delta\bar{\theta}$ leads to approximated probability mass function (pmf) p_k, given as

$$p_k = f(\bar{\theta}_k) \Delta\bar{\theta}, \tag{23}$$

with $\sum_{k \in \mathcal{K}} p_k = 1$, where $\bar{\theta}_k$ denotes the nominal DOA of the kth angular bin and $\mathcal{K} = \{1, 2, \cdots, K\}$ represents the index set of the angular bins. Using this discretization, the pdf of the compressed measurement $\boldsymbol{y}^{(q)}(t)$ for the qth subarray from Eq. (22) can be approximated as

$$f(\boldsymbol{y}^{(q)}) \approx \sum_{k \in \mathcal{K}} p_k f(\boldsymbol{y}^{(q)}|\bar{\theta}_k), \tag{24}$$

where $f(\boldsymbol{y}^{(q)}|\bar{\theta}_k)$ denotes the conditional distribution of $\boldsymbol{y}^{(q)}$ given a particular DOA $\bar{\theta}_k$. The compressed measurement vector $\boldsymbol{y}^{(q)}(t)$ corresponding to a signal impinging from the kth angular bin with nominal DOA $\bar{\theta}_k$ is assumed to be zero-mean complex Gaussian and can be expressed as

$$\boldsymbol{y}^{(q)}|_{\theta=\bar{\theta}_k} = \boldsymbol{\Phi} \left[\boldsymbol{a}^{(q)}(\bar{\theta}_k) s(t) + \boldsymbol{n}^{(q)}(t) \right], \tag{25}$$

where $\boldsymbol{a}^{(q)}(\bar{\theta}_k)$ denotes the steering vector associated with the qth subarray with conditional pdf [17]

$$f(\boldsymbol{y}^{(q)}|\bar{\theta}_k) = \frac{1}{\pi^M |\boldsymbol{C}^{(q)}_{\boldsymbol{y}\boldsymbol{y}|\bar{\theta}_k}|} e^{-(\boldsymbol{y}^{(q)})^{\mathrm{H}} (\boldsymbol{C}^{(q)}_{\boldsymbol{y}\boldsymbol{y}|\bar{\theta}_k})^{-1} \boldsymbol{y}^{(q)}}, \tag{26}$$

with

$$\boldsymbol{C}^{(q)}_{\boldsymbol{y}\boldsymbol{y}|\bar{\theta}_k} = \boldsymbol{\Phi} \boldsymbol{E}_k^{(q)} \boldsymbol{\Phi}^{\mathrm{H}} \tag{27}$$

representing the covariance matrix of the compressed measurements for the qth subarray given the DOA $\bar{\theta}_k$, where

$$\boldsymbol{E}_k^{(q)} = \sigma_s^2 \boldsymbol{a}^{(q)}(\bar{\theta}_k) (\boldsymbol{a}^{(q)}(\bar{\theta}_k))^{\mathrm{H}} + \sigma_n^2 \boldsymbol{I}. \tag{28}$$

The CMM is optimized using a gradient ascent strategy to maximize the mutual information $I(\boldsymbol{y}^{(q)}; \theta)$ between the compressed measurements and the signal DOAs, i.e.,

$$\boldsymbol{\Phi} \leftarrow \boldsymbol{\Phi} + \gamma \nabla_{\boldsymbol{\Phi}} I(\boldsymbol{y}^{(q)}; \theta), \tag{29}$$

where $\gamma > 0$ is a step size and $\nabla_{\boldsymbol{\Phi}} I(\boldsymbol{y}^{(q)}; \theta)$ denotes the gradient of the mutual information with respect to $\boldsymbol{\Phi}$, which is obtained using the approximated pmf as [17]

$$\nabla_{\boldsymbol{\Phi}} I(\boldsymbol{y}^{(q)}; \theta) \approx \frac{\sum\limits_{k \in \mathcal{K}} \sigma_n^2 p_k \frac{\boldsymbol{D}_k^{-1}}{|\boldsymbol{D}_k|} \boldsymbol{\Phi} \boldsymbol{E}_k^{(q)}}{\sum\limits_{k \in \mathcal{K}} p_k |\boldsymbol{D}_k|^{-1}} - \sum_{k \in \mathcal{K}} \sigma_n^2 p_k \boldsymbol{D}_k^{-1} \boldsymbol{\Phi} \boldsymbol{E}_k^{(q)} \tag{30}$$

with $\boldsymbol{D}_k = \sigma_n^{-2} \boldsymbol{\Phi} \boldsymbol{E}_k^{(q)} \boldsymbol{\Phi}^{\mathrm{H}}$.

In this paper, $\boldsymbol{\Phi}$ is optimized for a single subarray, and the same optimized $\boldsymbol{\Phi}$ is applied to all identically partitioned subarrays, ensuring computational efficiency. Alternatively, $\boldsymbol{\Phi}$ could be optimized separately for each subarray and averaged for analog beamforming. This approach reduces random variations but incurs a higher computational cost.

D. Iterative update of the DOA prior distribution

Since the prior distribution of the DOAs is unknown, $\boldsymbol{\Phi}$ is initially optimized by assuming a uniform pdf of θ, based on Eqs. (29) and (30). The spatially smoothed covariance matrix $\boldsymbol{R}_{dy_s}$ is then employed to estimate the spatial spectrum using MVDR, given as

$$P_i(\theta) = \frac{1}{L} \frac{(\boldsymbol{a}^{(1)}(\theta))^{\mathrm{H}} \boldsymbol{\Phi}_i^{\mathrm{H}} \boldsymbol{\Phi}_i \boldsymbol{a}^{(1)}(\theta)}{(\boldsymbol{a}^{(1)}(\theta))^{\mathrm{H}} \boldsymbol{\Phi}_i^{\mathrm{H}} \hat{\boldsymbol{R}}_{dy_s,i}^{-1} \boldsymbol{\Phi}_i \boldsymbol{a}^{(1)}(\theta)}, \tag{31}$$

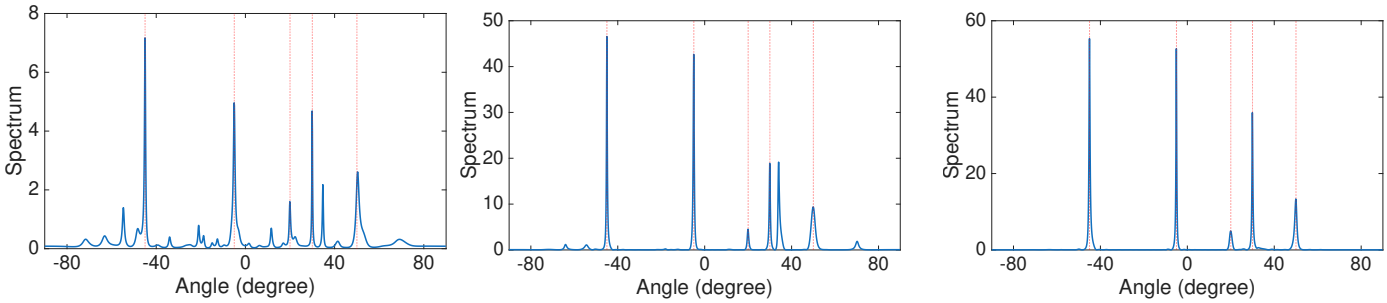
where $\boldsymbol{\Phi}_i$ and $\hat{\boldsymbol{R}}_{dy_s,i}$ represents the CMM and decorrelated covariance matrix for the ith iteration, respectively, $\boldsymbol{a}^{(1)}(\theta)$ is the steering vector for the first subarray. The normalized spectrum is then used to update the prior distribution of the DOAs as [20]

$$p_{k,i+1} \leftarrow \frac{P_i(\theta_k)}{\sum_{k \in \mathcal{K}} P_i(\theta_k)}, \tag{32}$$

which serves as the updated prior in the subsequent iteration for the optimization of $\boldsymbol{\Phi}$.

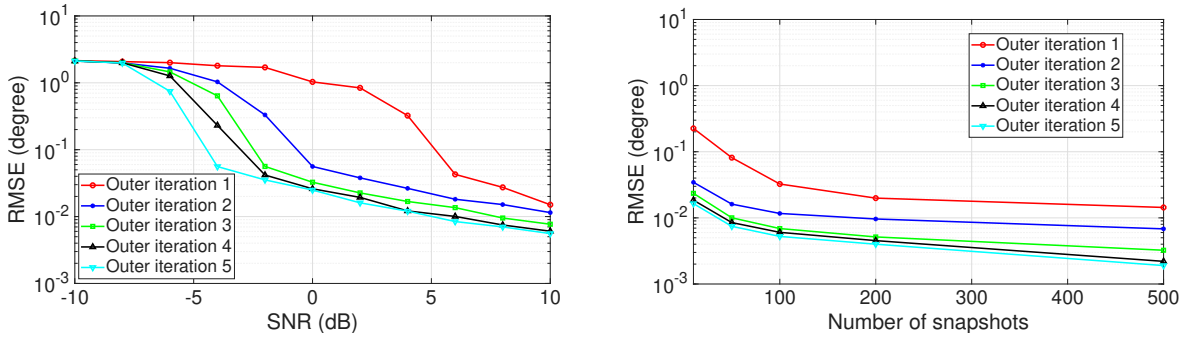
IV. Simulation results

We consider a massive MIMO system where the ULA is equipped with $N = 50$ receive antennas with half-wavelength spacing. $K = 5$ coherent signals are assumed. In applying the FBSS strategy, a minimum of $Q = \lceil \frac{K}{2} \rceil = 3$ subarrays and $M = K + 1 = 6$ RF chains per subarray are required. The hybrid beamforming structure employs a total of $QM = 18$ RF chains. Although higher than the 6 RF chains required for FHBF when all signals are uncorrelated, it handles coherent signals using far fewer RF chains than the 50 required for full digital beamforming. The subarray size in this structure is $L =$



(a) Outer iteration 1 (b) Outer iteration 2 (c) Outer iteration 3

Fig. 3: Estimated DOA spectrum.



(a) RMSE versus input SNR (b) RMSE versus number of snapshots

Fig. 4: Comparison of the RMSE performance.

$N-Q+1=48$, resulting in a CMM of dimension $\mathbf{\Phi}$ to be $M\times L=6\times 48$. The matrix $\mathbf{\Phi}$ is applied to each subarray to obtain the low-dimensional baseband data $\boldsymbol{y}^{(q)}(t)\in\mathbb{C}^{M\times 1}$ from the high-dimensional received data $\boldsymbol{x}^{(q)}(t)\in\mathbb{C}^{L\times 1}$ corresponding to the qth subarray.

For the optimization of $\mathbf{\Phi}$, the pdf is discretized over an angular grid with a resolution of 0.01°, yielding a total of $1,8001$ grid points in the approximate pmf. Assuming no prior knowledge about the signal arrivals, the pdf is initialized with a uniform distribution and subsequently updated according to (32). In each outer iteration, 30 inner iterations are carried out to update $\mathbf{\Phi}$ by maximizing the mutual information using a gradient ascent approach with a step size of $\gamma=0.001$.

We first consider 5 coherent signals with DOAs of -45°, -5°, 20°, 30°, and 50°. The magnitude of the complex scaling factors is randomly drawn from a uniform distribution over $[0.5,2]$, while the phase is selected uniformly between 30° and 60°. The input SNR of the reference signal is set to 20 dB and the number of snapshots is 100. Fig. 3 depicts the estimated DOA spectrum for 1 to 3 outer iterations. It is observed that, as the number of outer iterations increases, the spectrum becomes progressively cleaner. In particular, several spurious peaks present in the initial outer iteration disappear by the third outer iteration. Hence, the proposed subarray-based hybrid beamforming, combined with the information-theoretic iterative optimization, successfully detects coherent signals.

To quantitatively evaluate the DOA estimation performance, we further compute the root-mean-squared error (RMSE) results, defined as

$$\text{RMSE}=\sqrt{\sum_{i=1}^{D}\sum_{k=1}^{K}\frac{1}{DK}\left(\theta_k-\hat{\theta}_{k,i}\right)^2}, \tag{33}$$

where D is the number of Monte Carlo trials and $\hat{\theta}_{k,i}$ is the estimated DOA of the kth signal at the ith trial.

Fig. 4(a) shows the RMSE values with respect to the input SNR using 100 independent Monte Carlo trials. The number of outer iterations is varied from 1 to 5, while the number of inner iterations is fixed at 30. For RMSE computation, 2 coherent signals with DOAs at 60° and 25° are considered using 100 snapshots, and the input SNR varies between -10 dB and 10 dB. As observed in Fig. 4(a), the RMSE decreases as the input SNR increases, and the performance improves with more outer iterations and gradually converges. Fig. 4(b) depicts the RMSE versus the number of snapshots at an input SNR of 10 dB. Consistent with Fig. 4(a), the performance improves with additional outer iterations, achieving a low RMSE after 5 outer iterations, which demonstrates superior estimation accuracy.

V. Conclusion

This paper proposed a subarray-based hybrid beamforming architecture for coherent signal estimation. Because the full-dimensional analog signal is inaccessible, decorrelation is performed in the digital baseband domain after compressed measurements. Unlike the existing FHBF method, which only handles uncorrelated signals, the proposed subarray design maintains the translational invariance property of array manifold in the compression, thereby facilitating effective decorrelation

using spatial smoothing on post-compression data. Moreover, FBSS is employed to further reduce the required number of RF chains by roughly half, compared with the forward-only smoothing counterpart. Simulation results demonstrated the effectiveness of the proposed approach in achieving reliable coherent signal decorrelation and accurate DOA estimation.

REFERENCES

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

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

[3] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Top. Signal Process.*, vol. 8, no. 5, pp. 742–758, 2014.

[4] F. Jiang, J. Chen, A. L. Swindlehurst, and J. A. López-Salcedo, "Massive MIMO for wireless sensing with a coherent multiple access channel," *IEEE Trans. Signal Process.*, vol. 63, no. 12, pp. 3005–3017, 2015.

[5] Y. Gu and Y. D. Zhang, "Information-theoretic pilot design for downlink channel estimation in FDD massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 67, no. 9, pp. 2334–2346, 2019.

[6] S. Fortunati, L. Sanguinetti, F. Gini, M. S. Greco, and B. Himed, "Massive MIMO radar for target detection," *IEEE Trans. Signal Process.*, vol. 68, pp. 859–871, 2020.

[7] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Top. Signal Process.*, vol. 8, no. 5, pp. 831–846, 2014.

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

[9] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. M. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 122–130, 2014.

[10] A. F. Molisch, V. V. Ratnam, S. Han, Z. Li, S. L. H. Nguyen, L. Li, and K. Haneda, "Hybrid beamforming for massive MIMO: A survey," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 134–141, 2017.

[11] Y. Gu, Y. D. Zhang, and N. A. Goodman, "Optimized compressive sensing-based direction-of-arrival estimation in massive MIMO," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, New Orleans, LA, March 2017, pp. 3181–3185.

[12] X. Wu, D. Liu, and F. Yin, "Hybrid beamforming for multi-user massive mimo systems," *IEEE Trans. Commun.*, vol. 66, no. 9, pp. 3879–3891, 2018.

[13] M. Guo, Y. D. Zhang, and T. Chen, "DOA estimation using compressed sparse array," *IEEE Trans. Signal Process.*, vol. 66, no. 15, pp. 4133–4146, 2018.

[14] I. Ahmed, H. Khammari, A. Shahid, A. Musa, K. S. Kim, E. De Poorter, and I. Moerman, "A survey on hybrid beamforming techniques in 5G: Architecture and system model perspectives," *IEEE Commun. Surveys & Tutorials*, vol. 20, no. 4, pp. 3060–3097, 2018.

[15] T. Lin, J. Cong, Y. Zhu, J. Zhang, and K. B. Letaief, "Hybrid beamforming for millimeter wave systems using the MMSE criterion," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3693–3708, 2019.

[16] J. Zhang, X. Yu, and K. B. Letaief, "Hybrid beamforming for 5G and beyond millimeter-wave systems: A holistic view," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 77–91, 2019.

[17] Y. Gu and Y. D. Zhang, "Compressive sampling optimization for user signal parameter estimation in massive MIMO systems," *Digital Signal Process.*, vol. 94, pp. 105–113, 2019.

[18] S. R. Pavel, Y. D. Zhang, M. S. Greco, and F. Gini, "Deep learning-based compressive sampling optimization in massive MIMO systems," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Rhodes Island, Greece, June 2023, pp. 1–5.

[19] S. R. Pavel and Y. D. Zhang, "Optimization of the compressive measurement matrix in a massive MIMO system exploiting LSTM networks," *Algorithms*, vol. 16, no. 261, pp. 1–16, 2023.

[20] Y. D. Zhang, "Iterative learning for optimized compressive measurements in massive MIMO systems," in *Proc. IEEE Radar Conf.*, New York, NY, 2022, pp. 1–5.

[21] T.-J. Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-of-arrival estimation of coherent signals," *IEEE Trans. Acousti., Speech, Signal Process.*, vol. 33, no. 4, pp. 806–811, 1985.

[22] S. U. Pillai and B. H. Kwon, "Forward/backward spatial smoothing techniques for coherent signal identification," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 1, pp. 8–15, Jan. 1989.

[23] S. R. Pavel and Y. D. Zhang, "2D DOA estimation of coherent signals exploiting moving uniform rectangular array," *IEEE Signal Process. Lett.*, vol. 33, pp. 346–350, 2025.

[24] F.-M. Han and X.-D. Zhang, "An ESPRIT-like algorithm for coherent DOA estimation," *IEEE Antennas Wirel. Propag. Lett.*, vol. 4, pp. 443–446, 2005.

[25] Y.-H. Choi, "ESPRIT-based coherent source localization with forward and backward vectors," *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 6416–6420, Dec.

[26] S. R. Pavel and Y. D. Zhang, "Direction-of-arrival estimation of mixed coherent and uncorrelated signals," *IEEE Signal Process. Lett.*, vol. 31, pp. 2180–2184, 2024.

[27] S. R. Pavel, Y. D. Zhang, and B. Himed, "Structured decorrelation of covariance matrix for DOA estimation of coherent signals," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, Palermo, Italy, Sept. 2025, pp. 2087–2091.

[28] S. R. Pavel, Y. D. Zhang, and S. Sun, "2D DOA estimation of coherent signals exploiting forward-backward covariance tensor," *IEEE Trans. Signal Process.*, 2026 (in press).

[29] S. R. Pavel, Y. D. Zhang, and B. K. Chalise, "Massive MIMO system partitioning for efficient hybrid beamformer optimization," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Hyderabad, India, April 2025, pp. 1–5.