Low-Complexity Beam Selection Algorithms Based on SVD for mmWave Massive MIMO Systems

Jihong Yu‡, Member, IEEE, Jinxing Yang, Shuai Wang‡, Yuting Cai, and Jiahao Liu‡

Abstract—To realize beamspace MIMO with beam selection more efficiently, we propose low-complexity beam selection algorithms based on singular value decomposition (SVD). We first diagonalize the channel matrix by SVD, and select the appropriate beams one by one in a decremental or incremental order to maximize the sum-rate. To reduce the complexity of the proposed algorithms, we update eigenvalues by solving the secular equation instead of SVD from scratch again. Our method can achieve $O(N_r^2 N_t^2)$, which is at most one in $\frac{N_r^2 + N_t^2}{N_r}$ of the state-of-the-art method. Analytical and simulation results demonstrate that our proposed algorithms outperform the state-of-the-art algorithms in terms of the sum-rate and complexity.

Index Terms—Massive MIMO, mmWave communications, beamspace, beam selection, precoding.

I. INTRODUCTION

The rapid development of mobile communications, satellite networks, and Internet of Things (IoT) technologies will generate a large amount of data. To meet the explosive capacity demand, millimeter-wave (mmWave) communication is a promising technology. However, each antenna must be connected to one dedicated radio-frequency (RF) chain, leading to unbearable power consumption and hardware cost in massive multiple-input multiple-output (MIMO) scenarios with a large number of antennas [1], [2], [3].

To reduce the required number of RF chains, a feasible way is to transform the traditional MIMO channel into beamspace MIMO (B-MIMO) channel [1] by employing a designed discrete lens array (DLA). As DLA plays the role of the convex lens, the signal converges at different points of the focal surface, leading to the sparse channel with angle-dependent energy-focused capabilities in B-MIMO [2]. The sparse structure of B-MIMO enables us to only use a subset of the energy-focused beams while keeping the performance loss negligible [3], reducing the required number of RF chains drastically compared with the traditional MIMO. However, it is challenging to select the optimum subset of the beams, which has been proved NP-hard [4]. Note that there exists another way to reduce the number of RF chains named hybrid digital and analog precoding constructed by analog phase shifters [5], but our work focuses on addressing the lower-complexity beam selection problem.

To address the beam selection problem, a method named MM-BS selects the beams with the highest magnitude for each user [3], making multiple users select the same beam. The interference-aware beam selection algorithm named IA-BS is proposed [6], which circumvents this problem and achieves better performance than MM-BS by considering the potential interferences among users. Several algorithms based on other criteria have been proposed in [2], such as maximizing signal-to-interference ratio (SINR) and maximizing capacity. Based on the QR decomposition of the beamspace channel matrix, a beam selection algorithm and associated precoding matrix are designed in [7] (referred to as QRD-BS). Though this method achieves better performance than other algorithms, it suffers from very high computational complexity. To reduce the complexity of QRD-BS, the state-of-the-art algorithm named RQRD-BS is proposed [8], which could reduce complexity without the system performance degraded. However, RQRD-BS still exhibits high computational complexity.

In order to further reduce the complexity, this letter presents a series of novel low-complexity beam selection algorithms based on singular value decomposition (SVD). The complexity is reduced in two aspects:

- First, we choose a fixed number of high-energy beams to get a reduced-dimensional channel instead of applying to all beams as done in prior works [2], [6], [7], [8].
- Second, the appropriate beam is selected one by one by SVD in a decremental or incremental order to maximize sum-rate. These algorithms can avoid repeating SVD from scratch, achieving $O(N_r^2)$ complexity for computing sum-rate criterion, which is better than $O(N_r^2 N_t)$ and $O(N_r^2 + N_r N_t)$ in prior works [7], [8].

Another advantage of our proposed algorithms lies in producing the precoding matrix more simply and effectively than RQRD-BS without additional computation. The simulation results show that our proposed algorithms outperform the prior beam selection works.

Notations: Matrices and vectors are denoted by boldface uppercase and lowercase letters, respectively. $a_i$ and $A_{ij}$ denote the $i$th element of vector $a$ and the $(ij)^{th}$ element of matrix $A$, respectively. $A_{-j}$ denotes $A$ with its $j^{th}$ row removed. $I_M$ represents an $M \times M$ identity matrix. The superscripts $-1, T, *$ indicate inverse, transpose, and conjugate transpose operators, respectively. $\|a\|$ denotes its Frobenius norm.

II. SYSTEM MODEL

A mmWave MIMO system is considered where the Transmitter (TX) and Receiver (RX) employ $N_t$ transmit and $N_r$ receive antennas respectively, and both of them have $N_r$ RF
chains without loss of generality. For the typical uniform linear array (ULA) with $N$ antennas, the steering vector $\mathbf{a}(\psi, N) = \frac{1}{\sqrt{N}} \sum_{n=I(N)} e^{-j2\pi kn\psi}$, where $\mathcal{I}(N) = \{0, 1, \ldots, N - 1\}$. The MIMO channel matrix $\mathbf{G}$ can be characterized by a geometric channel model \cite{9}

$$
\mathbf{G} = \sqrt{\frac{N_r N_t}{N_c N_f}} \sum_{l=1}^{N_c} \sum_{i=1}^{N_r} \alpha_{i,l} \mathbf{a}(\varphi_{i,l}, N_t) \mathbf{a}(\psi_{i,l}, N_r)^*,
$$

(1)

where $N_c$ is the number of scattering clusters and each cluster is composed of $N_t$ subpaths and the spatial direction is defined as $\varphi_{i,l} = \frac{d}{\lambda}$ and $\psi_{i,l} = \frac{d}{\lambda}$, where $\theta_i$ and $\phi_i$ are azimuth angles of departure and arrival of the TX and RX respectively, $\alpha_{i,l}$ is the complex gain, $\lambda$ is the signal wavelength, and $d$ is the distance between neighboring antenna elements.

The carefully designed DLA can be viewed as a discrete Fourier transform (DFT) matrix \cite{1}, which can transform the conventional MIMO into the B-MIMO. Specifically, the DFT matrix $\mathbf{F}$, which can transform the elements.

$$
\mathbf{F} = \begin{bmatrix}
\cos \left( \frac{2\pi}{N} \right) & \cos \left( \frac{4\pi}{N} \right) & \ldots & \cos \left( \frac{(N-1)2\pi}{N} \right) \\
\sin \left( \frac{2\pi}{N} \right) & \sin \left( \frac{4\pi}{N} \right) & \ldots & \sin \left( \frac{(N-1)2\pi}{N} \right)
\end{bmatrix}
$$

The beamspace channel matrix $\mathbf{H} \in \mathbb{C}^{N_r \times N_c}$ experiences the bigger number of rows $N_t$ than the number of columns $N_c$. We can select $N_c$ highest energy beams to form a new thin beamspace matrix $\mathbf{H} \in \mathbb{C}^{N_c \times N_r}$, where $N_c$ is the size of the candidate beam set. Now we aim to select $N_c$ beams out of $N_r$ beams rather than $N_c$ beams.

We first denote SVD of $\mathbf{H}$ as

$$
\mathbf{H} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^*,
$$

(7)

where $\mathbf{U} \in \mathbb{C}^{N_r \times N_c}$ and $\mathbf{V} \in \mathbb{C}^{N_c \times N_c}$ are complex unitary matrices, and $\mathbf{\Sigma}$ is an $N_r \times N_c$ rectangular diagonal matrix with non-negative real numbers on the diagonal. The received signal vector $\mathbf{y}$ without beam selection can be written as

$$
\mathbf{y} = \mathbf{P}_r^* \mathbf{\Sigma} \mathbf{U}^* \mathbf{F} \mathbf{s} + \mathbf{P}_r^* \mathbf{n},
$$

(8)

If we choose $\mathbf{P}_r = \mathbf{V}$ and $\mathbf{P}_r = \mathbf{U}$, we can rewrite (8) as

$$
\mathbf{y} = \mathbf{\Sigma} \mathbf{s} + \mathbf{P}_r^* \mathbf{n}.
$$

(9)

Thus $R_s$ with uniform power allocation is given by

$$
R_s = \frac{1}{N_c} \log_2 \left( 1 + \frac{P_s}{\bar{P}_s N_0 \sigma_k^2} \right) \text{bits/s/Hz},
$$

(10)

where $\sigma_k$ is the $(kk)^{th}$ element of $\mathbf{\Sigma}$. From (9), the interference is equal to zero, and we get an equivalent representation of $\mathbf{H}$ as a parallel Gaussian channel \cite{10} for mitigating interference. Meanwhile, (10) implies we can reduce the complexity by updating $\sigma_k$ efficiently \cite{8}.

### A. SSVD-BS: Simplified SVD-Based Algorithm

From (10) the sum-rate $R_s$ can be estimated as

$$
R_s = \frac{1}{N_c} \log_2 \left( 1 + \frac{P_s}{\bar{P}_s N_0 \sigma_k^2} \right) \leq \frac{P_s}{\bar{P}_s N_0 N_c} \sum_{k=1}^{N_c} \sigma_k^2.
$$

(11)

Due to

$$
\sum_{k=1}^{N_c} \sigma_k^2 = \text{Tr} (\mathbf{H}^* \mathbf{H}) = \sum_{i,j} |h_{ij}|^2 = ||\mathbf{H}||^2_F,
$$

we can design a simplified SVD-based algorithm, referred to as SSVD-BS, which directly selects $N_r$ beams of highest energy to maximize the upper bound of $R_s$.

---

1The RX can also replace full digital precoding with beam selection and hybrid precoding.
Algorithm 1 Decremental SVD-Based Algorithm

Input: \( H \)

1: Initialize \( H^{(0)} = H \)
2: for \( i = 0, 1, \ldots, N_c - N_r - 1 \) do
3:   for \( j = 1, 2, \ldots, N_c - i \) do
4:     Remove \( j^{th} \) row from \( H^{(i)} \) to get \( H^{(i) \setminus j} \)
5:     Compute \( H^{(i) \setminus j} = U^{(i) \setminus j} \Sigma^{(i) \setminus j} (V^{(i) \setminus j})^* \)
6:   end for
7: end for
8: \( j = \arg \min_{j \in \{1, \ldots, N_c - i\}} \alpha_j^{(i)} \)
9: Remove \( j^{th} \) row from \( H^{(i)} \) to get \( H^{(i+1)} \)
10: end for

Output: \( H^{(N_c - N_r)} \)

i.e., \( N_r \log_2 \left( 1 + \frac{P_s}{N_c N_r} \sum_{k=1}^{N_r} \sigma_k^2 \right) \). This method is degraded from [11]. In this way, SVD-BS enjoys the lowest computational complexity \( O(N_c N_r) \) at the cost of performance loss compared with the other two algorithms introduced shortly since we only need to scan the channel matrix \( H \) once.

B. DSVD-BS: Decremental SVD-Based Algorithm

We now propose a novel decremental algorithm based on (10) instead of the upper bound of \( R_c \), referred to as DSVD-BS. DSVD-BS consists of \( N_c - N_r \) iterations. We define the channel matrix at the beginning of \( i^{th} \) iteration as \( H^{(i)} \). In an arbitrary iteration \( i \), we need to eliminate a beam (i.e., a row of \( H^{(i)} \)) with the minimum sum-rate loss, as detailed below.

For \( j = 1, \ldots, N_c - i \), we remove the \( j^{th} \) row of \( H^{(i)} \) to get the matrix \( H^{(i) \setminus j} \). We then compute its SVD \( H^{(i) \setminus j} = U^{(i) \setminus j} \Sigma^{(i) \setminus j} (V^{(i) \setminus j})^* \) and the sum-rate

\[
\alpha_j^{(i)} = \sum_{u=1}^{N_r} \log_2 \left( 1 + \frac{P_s}{N_0 N_r} \left( \sigma_j^{(i) \setminus u} \right)^2 \right), \tag{12}
\]

where \( \sigma_j^{(i) \setminus u} \) denotes the \((u)_j^{th}\) element of \( \Sigma_j^{(i)} \). We can eliminate the beam whose contribution to the sum-rate is the least, which implies less interference. This means that we remove \( j^{th} \) row from \( H^{(i)} \) to obtain \( H^{(i+1)} \), where \( j = \arg \max_{j \in \{1, \ldots, N_c - i\}} \alpha_j^{(i)} \). Algorithm 1 states this process.

C. ISVD-BS: Incremental SVD-Based Algorithm

We further design the incremental SVD-based algorithm named ISVD-BS with \( N_r \) iterations. In iteration \( i \), in contrast to DSVD-BS, ISVD-BS selects the beam (i.e., a row of \( H^{(i)} \)) whose contribution to the sum-rate is the highest.

Consider an arbitrary iteration \( i \), let \( H^{(i) \setminus j} \) be the matrix formed by the beams that were selected at the end of the \((i - 1)^{th}\) iteration. For \( j = 1, \ldots, N_c - i \), appending the \( j^{th} \) row of \( H^{(i)} \) to \( H^{(i) \setminus j} \) yields a new one denoted by \( H^{(i) \setminus j} \), i.e.,

\[
H^{(i) \setminus j} = \begin{pmatrix} H^{(i)} \\ h_j^{(i)} \end{pmatrix}, \tag{13}
\]

where \( h_j^{(i)} \) is the \( j^{th} \) row of \( H^{(i)} \). We can compute SVD

\[
H^{(i) \setminus j} = U_j^{(i) \setminus j} \Sigma_j^{(i) \setminus j} (V_j^{(i) \setminus j})^* \text{ and the sum-rate}
\]

\[
\eta_j^{(i)} = \sum_{u=1}^{N_r} \log_2 \left( 1 + \frac{P_s}{N_0 N_r} \left( \sigma_j^{(i) \setminus u} \right)^2 \right), \tag{14}
\]

where \( N_r' \) is the number of positive singular value of \( \Sigma_j^{(i)} \), and \( \sigma_j^{(i) \setminus u} \) denotes the \((u)_j^{th}\) element of \( \Sigma_j^{(i)} \). Here we can select the beam which improves the sum-rate most, i.e., \( j = \arg \max_{j \in \{1, \ldots, N_c - i\}} \eta_j^{(i)} \).

Then we append the \( j^{th} \) row of \( H^{(i)} \) to \( H^{(i) \setminus j} \) and remove \( j^{th} \) row of \( H^{(i)} \) to get \( H^{(i+1)} \). This process is stated in Algorithm 2.

D. Complexity Reduction

For DSVD-BS, we first omit the superscript \((i)\) for readability. We denote the diagonal elements of \( \Sigma_2 \) and \( \Sigma_2^{\setminus j} \) by \( D = \text{diag}\{d_1, d_2, \cdots, d_{N_r}\} \) and \( D' = \text{diag}\{d_1', d_2', \cdots, d_{N_r}'\} \), respectively, where \( d_1 \geq d_2 \geq \cdots \geq d_{N_r} \) and \( d_1' \geq d_2' \geq \cdots \geq d_{N_r}' \). Since \( H^{(i) \setminus j} H^{(i) \setminus j} = H^{(i)} H = h_j h_j^* \), our idea is to make use of the prior knowledge \( \{d_1, \cdots, d_{N_r}\} \) to avoid computing \( \{d_1', \cdots, d_{N_r}'\} \) from scratch again for reducing the complexity of step 5 in DSVD-BS. Based on eigenvalue decomposition (EVD) of \( H^{(i) \setminus j} H^{(i) \setminus j} \) and \( H^{(i)} H \), we can get

\[
V_{-j} D' V_{-j}^* = V D V^* - h_j h_j^*. \tag{15}
\]

Let \( z_j = V^* h_j \), we have

\[
h_j h_j^* = V z_j z_j^* V^*. \tag{16}
\]

Substituting (16) into (15) yields

\[
V_{-j} D' V_{-j}^* = V (D - z_j z_j^*) V^*. \tag{17}
\]

Since \( D - z_j z_j^* \) is a Hermitian matrix, its EVD is \( D - z_j z_j^* = Q S Q^* \). The EVD is unique since the eigenvalues are distinct. Therefore, we have

\[
V_{-j} = V Q, \tag{18}
\]

\[
S = D'. \tag{19}
\]
According to [12], the eigenvalue of \( D - z_j z_j^* \) is the root \( d' = d_1', \ldots, d_{N_r}' \) of the secular function
\[
f(d') = 1 - \left( \frac{|z_1|^2}{d_1' - d'} + \cdots + \frac{|z_{N_r}|^2}{d_{N_r}' - d'} \right),
\]
where \( z_i \) is the \( i \)-th element of \( z_j \). The eigenvector \( q_i \) corresponding to each \( d'_i \) is
\[
q_i = \frac{(D - d'_i I)^{-1} z_j}{\| (D - d'_i I)^{-1} z_j \|}.
\]
The function \( f(d') \) is a monotonically decreasing function between its poles because
\[
f(d') = -\left( \frac{|z_1|^2}{(d_1 - d')^2} + \cdots + \frac{|z_{N_r}|^2}{(d_{N_r} - d')^2} \right) < 0.
\]
Thus, the eigenvalue \( d'_i \) satisfy the following interlacing property:
\[
d_{i+1} \leq d'_i \leq d_i, \quad i = 1, 2, \ldots, N_r
\]
with \( d_{N_r+1} = d_{N_r} - |z_{j}|^2 \).

Similarly, for ISVD-BS, we have \( H^*_j H_j = H^* H + h_i h_i^* \). The eigenvalue \( d_i \) of \( H^*_j H_j \) is the root of the secular equation
\[
g(d) = 1 + \left( \frac{|z_1|^2}{d_1 - d} + \cdots + \frac{|z_{N_r}|^2}{d_{N_r} - d} \right),
\]
and satisfy
\[
d_i \leq d_i \leq d_{i-1}, \quad i = 1, 2, \ldots, N_r
\]
where \( d_0 = d_1 + |z_j|^2 \).

To find the roots of (20) and (25), we can apply bisection or the numerical algorithms in [13], i.e., we can compute \( \{d_1', \ldots, d_{N_r}'\} \) and \( \{d_1, \ldots, d_{N_r}\} \) from \( \{d_1', \ldots, d_{N_r}'\} \) with \( O(N_r^2) \) computational complexity instead of \( O(N_c N_r^2) \), which is the complexity of step 5 in algorithm 1 and 2. Meanwhile, we can get the eigenvector corresponding to eigenvalue by (21) with \( O(N_r^2) \) computational complexity. Our proposed algorithms can be extended to multi-user MIMO by block diagonal (BD) precoding algorithm and please refer to [14] about BD for more details.

### E. Computational Complexity Analysis

Since we can get \( D' \) based on \( D \) directly as analyzed in Section III-D rather than performing SVD, step 5 in Algorithm 1 and 2 can be replaced:

- For DSVD-BS, we can first compute the SVD \( H^{(0)} = U^{(0)} \Sigma^{(0)} (V^{(0)})^* \). Then for the outer iterations \( i = 0, 1, \ldots, N_c - N_t - 1 \) of DSVD-BS, the computational complexity is \( O((N_c - i)N_r^2) \) since the step 5 needs \( O(N_r^2) \) computational complexity. Therefore, the total computation complexity of DSVD-BS is \( O\left( \sum_{i=0}^{N_c-N_t-1} (N_c - i)N_r^2 \right) = O\left( \frac{N_c^2 N_r^2}{2} \right) \).

- Similarly, for the outer iteration \( i = 0, 1, \ldots, N_r - 1 \) of ISVD-BS, the computational complexity is \( O((N_c - i)N_r^2) \). Therefore, the total computation complexity of ISVD-BS is \( O\left( \sum_{i=0}^{N_r-1} (N_c - i)N_r^2 \right) = O\left( N_r^3 N_c \right) \).

From the above discussion, when \( N_c \geq 2N_r \), the complexity of DSVD-BS is higher than ISVD-BS roughly and vice versa. The complexity of QRD-BS [7] and RQRD-BS [8] algorithms are \( O(N_c^2 N_r^3) \) and \( O(N_r N_c^3 + N_r^2 N_c^3) \) respectively, which are much larger than DSVD-BS in three aspects:

- \( N_r \) is much larger than \( N_c \) for both algorithms.
- Compared with QRD-BS, the degree of the term \( N_t \) in the complexity of DSVD-BS is 3. However, the degree of \( N_c \) in the complexity of QRD-BS is 2.
- Compared with RQRD-BS, ours doesn’t have the term \( O(N_r N_c^3) \). Since \( N_c < N_r \), the complexity of DSVD-BS is at most one in \( \frac{N_r+N_c^2}{N_r} \) of RQRD-BS.

### IV. Simulation Results

In this section, we conduct simulations to evaluate the performance of the proposed algorithms compared with the prior works [8] and fully digital SVD-based precoding at TX (referred to as FD-SVD). For the channel model (1), we set \( N_t = 128, N_r = 8, N_c = 4 \) and \( N_c = 3, \alpha_{dL} \sim CN(0, 0.02) \) for the line-of-sight component, otherwise \( \alpha_{dL} \sim CN(0, 1) \) for non-line-of-sight components. The SNR is set to 25 dB. The \( \varphi_{i,j} \) and \( \psi_{i,j} \) are sampled from the uniform distribution within \( [-\pi, \pi] \). Since ISVD-BS enjoys lower complexity to achieve the sum-rate almost the same as DSVD-BS when \( N_c = 3N_t \), here we only show the performance of ISVD-BS for simplicity. The simulation results are averaged over 1000 channel realizations.

Fig. 1 depicts the sum-rate performance versus the SNR of different beam selection algorithms, including our proposed SSVD-BS and ISVD-BS, FD-SVD, and RQRD-BS [8]. SSVD-BS that has the lowest complexity \( O(N_r N_c) \) can achieve a very close performance to ISVD-BS and outperform RQRD-BS when the SNR is low. Thus, in the stringent computational complexity scenario, SSVD-BS is attractive. We can observe that our ISVD-BS achieve the highest sum-rate for all beam selection algorithms. The lowest-complexity SSVD-BS experiences performance loss compared with ISVD-BS, and the performance gap becomes more pronounced as the SNR increases. Note that the complexity of ISVD-BS is much lower than RQRD-BS.

In Fig. 2, we plot the sum-rate performance versus the number of transmit antennas \( N_t \). Our ISVD-BS and SSVD-BS achieve much better performance compared with RQRD-BS when \( N_t = 20 \). And the impact of \( N_t \) on the sum-rate of

![Fig. 1. Sum-rate performance versus the SNR.](image-url)
ISVD-BS is stronger than SSVD-BS, which illustrates the superiority of ISVD-BS. Meanwhile, ISVD-BS can achieve near performance to FD-SVD when $N_t$ is small. Therefore, ISVD-BS can be deployed in a scenario where a high sum-rate is required with a limited number of antennas.

Fig. 3 demonstrates the sum-rate of different algorithms versus normalized mean square error (NMSE) of the channel estimation. As shown in Fig. 3, when the NMSE is larger, the sum-rate of all algorithms decreases. When NMSE $\geq -22.5$dB, the performance of the proposed ISVD-BS exceeds that of FD-SVD. The sum-rate of ISVD-BS exceeds that of the RQRD-SVD and SSVD-BS algorithms, so ISVD-BS is more robust for imperfect CSI than the other algorithms.

Fig. 4 plots energy efficiency of different beam selection algorithms versus the number of RF chains of RX. The energy efficiency is defined as [6]. The proposed ISVD-BS algorithm achieves the highest energy efficiency, followed by RQRD-BS, SSVD-BS, and FD-SVD. The energy efficiency of ISVD-BS is more than ten times that of FD-SVD, which fully demonstrates the superiority of the beam selection system over the fully digital precoding system. When $N_r$ increases, the gap between ISVD-BS and the RQRD-BS is widening.

V. CONCLUSION

In the letter, we proposed three low-complexity beam selection algorithms, namely SSVD-BS, DSVD-BS, and ISVD-BS. All three algorithms can achieve a better sum-rate performance than the fully digital system with much higher energy efficiency. In particular, SSVD-BS has the lowest computation complexity $O(N_t N_r)$ and comparable sum-rate. DSVD-BS and ISVD-BS can outperform the state-of-the-art work named RQRD-BS in terms of the sum-rate with much lower computation complexity.

REFERENCES

[1] Y. Zeng and R. Zhang, “Millimeter wave MIMO with lens antenna array: A new path division multiplexing paradigm,” *IEEE Trans. Commun.*, vol. 64, no. 4, pp. 1557–1571, Apr. 2016.
[2] P. V. Amadori and C. Masouros, “Low RF-complexity millimeter-wave beamspace-MIMO systems by beam selection,” *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2212–2223, Jun. 2015.
[3] A. Sayeed and J. Brady, “Beamspace MIMO for high-dimensional multilayer communication at millimeter-wave frequencies,” in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2013, pp. 3679–3684.
[4] H. Liu, X. Yuan, and Y. J. Zhang, “Statistical beamforming for FDD downlink massive MIMO via spatial information extraction and beam selection,” *IEEE Trans. Wireless Commun.*, vol. 19, no. 7, pp. 4617–4631, Jul. 2020.
[5] F. Sohrabi and W. Yu, “Hybrid digital and analog beamforming design for large-scale antenna arrays,” *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 501–513, Apr. 2016.
[6] X. Gao, L. Dai, Z. Chen, Z. Wang, and Z. Zhang, “Near-optimal beam selection for beamspace mmWave massive MIMO systems,” *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 1054–1057, May 2016.
[7] R. Pal, K. V. Srinivas, and A. K. Chaitanya, “A beam selection algorithm for millimeter-wave multi-user MIMO systems,” *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 852–855, Apr. 2018.
[8] Q. Zhang, X. Li, B. Wu, L. Cheng, and Y. Gao, “On the complexity reduction of beam selection algorithms for beamspace MIMO systems,” *IEEE Wireless Commun. Lett.*, vol. 10, no. 7, pp. 1439–1443, Jul. 2021.
[9] X. Li, J. Fang, H. Li, and P. Wang, “Millimeter wave channel estimation via exploiting joint sparse and low-rank structures,” *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 1123–1133, Feb. 2018.
[10] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[11] S. Vishwanath, N. Jindal, and A. Goldsmith, “Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels,” *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.
[12] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD, USA: JHU Press, 2013.
[13] A. Melman, “A numerical comparison of methods for solving secular equations,” *J. Comput. Appl. Math.*, vol. 86, no. 1, pp. 237–249, Nov. 1997.
[14] L.-U. Choi and R. D. Murch, “A transmit preprocessing technique for multilayer MIMO systems using a decomposition approach,” *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 20–24, Jan. 2004.