IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

IEEE VEHICULAR TECHNOLOGY SOCIETY SECTION

Received 19 June 2023, accepted 5 July 2023, date of publication 21 July 2023, date of current version 2 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3297884

RESEARCH ARTICLE

Channel Estimation Based on Compressed Sensing for Massive MIMO Systems With Lens Antenna Array

ELHAM SHARIFI[1,2], MAHMOOD MOHASSEL FEGHHI[3], GHANBAR AZARNIA[4], SAJJAD NOURI[3], DUEHEE LEE[5], (Member, IEEE), AND MD. JALIL PIRAN[6], (Senior Member, IEEE)

[1]Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz 51666, Iran
[2]Faculty of Electrical and Computer Engineering, Tarbiat Modares University, Tehran 14115, Iran
[3]Department of Communications Engineering, Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz 51666, Iran
[4]Engineering Faculty of Khoy, Urmia University of Technology, Urmia 57166, Iran
[5]Department of Electrical and Electronics Engineering, Konkuk University, Seoul 27478, South Korea
[6]Department of Computer Science and Engineering, Sejong University, Seoul 05006, South Korea

Corresponding authors: Mahmood Mohassel Feghhi (mohasselfeghhi@tabrizu.ac.ir) and Elham Sharifi (sh.elham95@ms.tabrizu.ac.ir)

This paper was supported by Industry & Energy South Korea, under Grant 20204010600220 and Grant RS-2023-00237035.

ABSTRACT With the emergence of fifth-generation cellular networks (5G), there has been significant interest in multi-input-multi-output (MIMO) systems. MIMO systems aim to achieve several key objectives, including increasing capacity, mitigating the negative effects of multi-path propagation, minimizing interference, and achieving higher data rates. Furthermore, the utilization of millimeter-wave (mmWave) technology and high bandwidth can address traffic congestion and interference challenges, leading to substantial improvements in data rates, spectral efficiency, and overall bandwidth. In the realm of mmWave communications, massive MIMO systems incorporating lens antenna arrays have proven effective in reducing the number of required radio-frequency chains. This paper presents two distinct approaches to address the challenge of massive MIMO channel estimation in mmWave communications. The first approach proposes a compressed sensing (CS) scheme based on convex optimization, which offers accurate and low-complexity channel estimation. The second approach introduces an estimation algorithm based on the greedy method, which provides fast reconstruction, a straightforward geometric interpretation, and a mathematically efficient framework. Extensive simulations demonstrate the superior performance of the proposed algorithms compared to similar methods such as support detection (SD), orthogonal matching pursuit (OMP), and sparsity mask detection (SMD). The proposed algorithms exhibit higher channel estimation accuracy, better recovery quality, and fast convergence rates.

INDEX TERMS Channel estimation, compressed sensing, lens antenna array, massive MIMO, millimeter-wave communication.

I. INTRODUCTION

The fifth-generation (5G) of cellular communications was developed to meet the stringent requirements of applications not supported by the previous generation. Some of these requirements are: data rate, connection density, mobility, energy efficiency, flexibility and latency. One of the key services of 5G networks is to empower the ultra-reliable and low latency communications (URLLC) capacity, which is required to help special degrees of high reliability and low latency end-to-end (E2E) communication. The 3rd generation partnership project (3GPP) specifies the basic URLLC reliability essentials for a single data frame of 32 as 99.9%, and an E2E latency of $<$1ms. Also, in this generation, the data rate

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Quan.

has reached a peak of 20 Gbps and an average of 100 Mbps, it has a mobility of 500 km/h and energy efficiency with a battery life of 10 to 15 years.

New technologies have been introduced to achieve these requirements in the new generation, such as massive multi-input-multi-output (MIMO) and mmWave [1], [2], [3]. Since high-quality audio and video services have developed and the number of home communication devices has increased in recent years [4], [5], the demand for high-speed wireless communications at short distances has increased accordingly. Its large bandwidth and high spectral efficiency make mmWave technology a key technology for 5G communications. The development of massive MIMO technology is driven by its ability to increase data rates significantly, increase communication bandwidth without requiring additional bandwidth or power, and improve spectrum efficiency and energy efficiency [6]. There are a variety of types of phase and frequency distortion that affect wireless channels, including MIMO channels, and the effects of each type depend on the channel and the network fluctuations. Channel estimation determines how the physical channel affects the input string. In Massive MIMO systems for mmWave communications, there are various types of channel estimation methods categorized as follows: 1) Training-based Channel Estimation Methods: These methods employ training information or sequences to estimate the channel. The training sequences are designed using specific algorithms to achieve optimal channel estimation. An example is the Advanced Training-based Method, which enhances channel estimation using lens antenna arrays and advanced techniques [7], [8]. 2) Blind Channel Estimation Methods: a) Self-Recovery Algorithms: These methods estimate the channel using specific computations and algorithms, such as source combination and feedback estimation. b) Correlation-based Algorithms: These methods estimate the channel based on the correlation between transmitted and received signals, utilizing linear or nonlinear correlation-based algorithms [9], [10]. 3) Semi-Blind Channel Estimation Methods or Statistical Information-based Methods: These methods utilize statistical information along with received data to estimate the channel. Statistical properties of signals and the channel, including the statistical distribution, are considered. Examples include maximum likelihood-based algorithms and Kalman filter-based algorithms [11].

These channel estimation methods are used in Massive MIMO systems for mmWave communications, each with its own characteristics and advantages. The selection of a specific method depends on the communication conditions and requirements. For example while channel blind estimators exhibit good performance in transmitting information, they face limitations in channels that experience rapid changes over time. In modern telecommunications, factors such as movement, Doppler effects, and channel variability play crucial roles, presenting significant challenges in terms of hardware and software complexities [12], [13], [14].

The use of lens antenna arrays [15], [16] can provide enhanced beamforming and beam steering capabilities, enabling improved spatial resolution and signal focusing. This, in turn, contributes to more accurate channel estimation. Accurate channel estimation is crucial for various aspects of Massive MIMO systems, including beamforming, precoding, spatial multiplexing, and interference management. Reliable channel state information (CSI) is essential for optimizing system performance, mitigating interference, and achieving high data rates. Furthermore, precise channel estimation can help overcome channel fading effects and improve the overall system robustness and reliability.

By conducting research on channel estimation techniques specifically tailored for lens antenna arrays in Massive MIMO channels, researchers can unlock the technology's potential to provide substantial gains in terms of capacity, coverage, and quality of service. This, in turn, paves the way for advanced wireless communication systems like 5G and beyond, facilitating innovative applications such as the Internet of Things (IoT), augmented reality, and autonomous vehicles.

The channel estimation method based on training sequences in mmWave telecommunications has advantages such as having more detailed information about channel characteristics, such as delay, interference and signal flow, ensuring data transmission without errors and interference, and optimal use of resources such as energy and bandwidth. Therefore, this method can help to estimate the channel of mmWave massive mimo systems. Algorithms for channel estimation and its types have been described in the related works section.

mmWave bands, however, have limited scattering, so only a few dominant beams can be selected. In this way, the number of radio-frequency chains can be reduced significantly [17]. Furthermore, due to the sparsity of the mmWave in the angle domain, selecting a suitable beam can be made using the compressive sensing (CS) methods [18], [19].

II. RELATED WORK

Massive MIMO systems with lens antenna arrays in mmWave communications can significantly reduce the number of radio frequency (RF) chains with beam selection. However, beam selection requires the base station (BS) to obtain accurate information from the beamspace. This is a challenging task, for example, the channel size of the beam space is large while the number of RF chains is limited. In the desired system model, the problem of beamspace channel estimation in mmWave massive MIMO with lens antenna array is investigated. So that an adaptive selective network including a small number of one-bit phase shifters is considered for mmWave massive MIMO systems with lens antenna array [20].

The proposed selective network serves a dual purpose in data transmission and channel estimation. During data transmission, it operates by selecting beams as usual. However, during channel estimation, it acts as a combiner to obtain effective beamspace channel measurements. Therefore, the

beam space channel estimation problem can be formulated as a weak signal recovery problem based on the adaptive selective network [17].

In another study [21], the authors proposed a channel estimation method that incorporates complete knowledge of large-scale fading. They introduced a pilot reuse sequence to eliminate pilot contamination in edge users. The study considered Rayleigh fading for small-scale and an identically independent distributed (i.i.d.) channel model in multi-cell massive MIMO systems. The authors estimated the large-scale fading and analyzed the performance in terms of maximum ratio transmission and zero-forcing precoding as the number of antenna elements and users increased. They further enhanced the channel quality by allocating orthogonal pilot reuse sequences specifically to the edge user group to mitigate pilot contamination.

Moreover, the achievable data rate and channel quality improve by maintaining a constant grouping parameter based on large-scale fading and dividing users into two groups: edge users and central users. This grouping approach optimizes transmission parameters according to the characteristics of each user group, resulting in increased data rates and improved channel quality.

To estimate the sparse channel, various algorithms have been presented, including greedy, statistical, threshold-based, and convex optimization algorithms. Greedy algorithms iteratively approximate the signal and support coefficients until convergence or reaching a certain threshold. Some of these algorithms perform comparably to convex optimization methods, such as norm l_p minimization. One method called orthogonal matching pursuit (OMP) iteratively identifies the signal support and reconstructs the signal using a virtual-inverse approach. OMP identifies columns of the measurement matrix with the highest correlation and repeats this step, calculating correlations with the residual signal to recover the sparse signal. The stopping condition is defined as the number of iterations or the requirement to achieve the observed data. OMP can recover the sparse signal by repeating this step several times [22].

In addition, the authors employ a wideband dictionary and demonstrate that channels across different orthogonal frequency division multiplexing (OFDM) subcarriers share a common support. This insight enables the application of a variant of the simultaneous OMP algorithm, known as generalized simultaneous OMP (GSOMP). GSOMP exploits information from multiple subcarriers to increase the probability of successfully recovering the common support [23].

Another category of algorithms that are used in compressed reception is threshold-based algorithms. The insight that justifies the classification of these algorithms as a different family relies on the approximation of the matrix inverse and the adjoint matrix (measurement). The basic threshold algorithm consists of determining the s-sparse support of the desired signal that must be reconstructed from the observation vector $\boldsymbol{y} = \boldsymbol{Ax}$, in the form of indices corresponding to the absolutely large s component of $\boldsymbol{A} * \boldsymbol{y}$, and then finding a vector with this support in such a way that have the best compatibility with the measurement vector. Algorithms related to this set include iterative hard thresholding (IHT) [24], [25], normalized iterative hard thresholding (NIHT) [26], [27] and hard thresholding pursuit (HTP) [28], [29].

Another type of algorithms are sublinear algorithms. These algorithms reconstruct sparse signals through group testing and are much faster and more efficient compared to greedy or convex optimization algorithms. But they need a special pattern in the observation matrix, and that is that the observation matrix must also be sparse [30]. In addition to the stated methods, there are also statistical methods such as Bayes method, which, unlike other methods, use previous information about signal coefficients. That is, they consider a model for the signal. While in non-statistical methods, the only prior information is the sparsity of the signal. As an example, [31] has proposed a method in which Gaussian-Bernoulli prior information is considered for the s-sparse signal. Also, in the proposed method [32], [33], signal sparsity is modeled using Laplace's hierarchical form as prior information. Although it is reasonable to assume additive Gaussian noise; However, the Gaussian-Bernoulli or Laplace prior information assumption for signal statical moments will not always correspond to reality. Accordingly, [34] and [35] has presented a Bayes-based method for sparse signal reconstruction that uses noise statistics and signal sparsity rate.

Another group of methods is convex optimization techniques, which are very powerful for calculating sparse representations. Convex optimization is a sub-branch of optimization problems in mathematics that deals with the minimization of convex functions on convex sets. In [36] and [37], to reduce the overhead of pilot-based channel estimation for cell-free systems, the authors propose a novel joint channel estimation and data detection (JED) algorithm that exploits (i) channel sparsity and (ii) the fact that the user equipment (UE) transmit signals are taken from a bounded constellation set (e.g., QPSK). JED algorithm uses forward-backward splitting (FBS) [8] to efficiently solve a biconvex optimization problem. Recently, algorithms based on sparsity mask detection (SMD) [38] and support detection (SD) [17] proposed to estimate the beam space channel in mmWave massive MIMO systems.

The SMD algorithm first selects beams with higher power using a beam training process between the base station and users, and as a result, by reducing the dimensions of the channel and solving it with classical algorithms, such as least squares (LS), it achieves an estimate of the channel. One of the advantages of this method is reducing the training pilots [34], [38]. However, the number of training pilots required to check all beams is proportional to the number of base station antennas, which is a large number (for example, 256 antennas). This number of pilots is significantly reduced based on the SD-based channel estimation scheme [17].

The basic idea is to decompose the total channel estimation problem into a series of sub-problems, each of which only considers one sparse channel component (a vector containing the information of a specific propagation direction) [39]. First, for each channel component its support is detected (i.e., the index set of non-zero elements in a sparse vector) by exploiting the structural characteristics of mmWave beamspace channel. Then, the influence of this channel component is removed, and the support of the next channel component is detected in a similar method. After the supports of all channel components have been detected, the beamspace channel of large size can be estimated with low pilot overhead.

In [40], the authors proposed the regular scanning support detection (RSSD) algorithm, a channel estimation technique for the 5G mmWave massive MIMO system with 3D metallic plate lens antenna array that saves energy and focus beams. Initially, entire channel estimation obstacles are divided by some alternate obstacles named as channel components. After that, using regular scanning the component support is detected along with fading action. Finally, channel component will be abolished from the entire beamspace channel estimation complication. In [41] and [42], the authors propose separable compressive sampling match pursuit (SCoSamp) algorithm by exploiting the separable structured sparsity of mmWave massive MIMO channel. They present that the angular spreads of mmWave channel give rise to the sparse separable structure, such that the sparsity of mmWave channel can be divided into angle of arrival (AoA) and departure (AoD) domains separately. After designing the precoding and combining matrices under the metric of mutual information, according to the separability structure of the channel, SCoSamp algorithm is proposed.

Finally, it solves the problem by calculating the correlation matrix of the signal and separating the coordinates of the channel elements into two sub-matrices through least squares. In this method, the number of pilots to estimate the channel is equal to the number of elements, and since it performs the estimation twice for each matrix, as a result, it has more computational load and the number of pilots is also high.

In [43], the authors based on the specific triple-structured sparsity in the cascaded beamspace channel, proposed a novel multi-user joint cascaded channel estimation scheme named as MTSCS-CE. The calculational complexity is close to conventional OMP method but pilot overhead can be remarkably reduced under the same normalized mean square error (NMSE) accuracy. Their proposed algorithm in the paper concentrates on on-grid scenarios, while super resolution estimation based on the triple-structured sparsity and the trade-off between complexity and estimation accuracy will be further discussion to overcome the energy leakage problem.

In [44], a multi-user mMIMO system with CSI feedback resource adaptation framework and closed-loop pilot has proposed. The framework includes CSI quality estimation, joint compressive CSI recovery and pilot resource adaptation. From simulation results, the framework proposed in [44] upgraded CSI estimation performance, and improved robustness against dynamic channel sparsity.

A combination of compressed sensing, S-CoSaMP, block iterative support-detection (Block-ISD) and AoD algorithms have given in [45] as a solution to decrease feedback overhead of channel estimation. Simulation results show that this proposed method reduces the pilot feedback overhead by a significant percentage which improves the overall system spectral and energy efficient (SE and EE).

In [46] an estimation framework extended using enhanced Newtonized orthogonal matching pursuit (eNOMP) algorithm is given. The algorithm is applied to extract from the uplink the frequency-independent parameters, which can be apply to expand the downlink training scheme. eNOMP is an iteration-based method that extracts a new component within each iteration through the e-OMP and eNewton steps. At the end of every iteration of the eNOMP algorithm, the corresponding component path will be eliminated from the noisy mixture.

Finally, the number of practical components equals the number of extracted components if each component is accurately estimated. Numerical results display that the proposed framework in [46] can reconstruct the CSI with training overhead and small amount of feedback.

In [47] the authors proposed a CS estimation scheme for frequency division duplexing (FDD) mMIMO systems by proposing a structured sparse adaptive (CoSaMP) algorithm that modifies the coding sampling matching pursuit (CoSaMP) algorithm. Simulation results show that the proposed algorithm in [47] can reduce the pilot overhead and has good performance and better SE in low signal-to-noise ratio (SNR).

Authors in [48] propose a non-orthogonal pilot design with CS-based channel estimation algorithm. By exploiting spatial and temporal joint dispersion in the mMIMO delay domain, the results show that less pilot overhead is achieved for accurate CSI estimation. Authors in [49] proposed a design consisting of scheduling and joint precoding algorithm for multi-cell mMIMO systems to deal with the pilot feedback overhead issue in FDD mMIMO systems. Also, a greedy scheduling algorithm is introduced to further reduce the computational complexity of the reported algorithm in [49]. Simulations confirm that overhead feedback reduction is achieved.

In [50] the intrinsic tensor feature of the full-dimension MIMO (FD-MIMO) channel is explored, to increase CSI estimation. The proposed estimation algorithm in [50] is based on the expectation-maximization framework via tensor as the processing data structure. Moreover, the Cramér-Rao lower bound is employed as a metric for evaluation. Simulation results show an enhancement in CSI estimations, providing flexibility in realizing trade-off points between downlink and uplink throughput using pilot-data superposition strategy.

In the aforementioned algorithms, the number of training pilots required for channel estimation is often high, sometimes equal to the number of antennas at the base station, which can be challenging in mmWave massive MIMO systems. Considering this, our work is based on two key algorithms: OMP and the SD algorithm, which offers the lowest training pilot requirement.

In the OMP algorithm, all channel components are initially considered for estimation, resulting in increased pilot overhead and decreased accuracy. This reduction in accuracy is mainly due to the presence of numerous channel components with very small or almost zero values. Estimating these components introduces errors, as their values are forced to be nonzero according to the minimum square error (MSE) metric. Therefore, it is unnecessary to estimate all components, and a subset of channels is sufficient for accurate estimation. On the other hand, the SD algorithm significantly reduces the support set, but it lacks responsiveness to different channel conditions and paths. It relies on a fixed series of indices, requiring prior knowledge of the number of transmitted paths.

The proposed methods address these limitations by offering distinct advantages. Firstly, they do not rely on knowing the number of signal transmission paths. Secondly, the channel support set consistently includes the most relevant and effective channel components for accurate estimation. This ensures both improved accuracy and reduced training pilot requirements.

III. CONTRIBUTIONS

This paper introduces two different systems models that include conventional channels and channels with a lens antenna array. We then propose an algorithm for channel estimation with an acceptable overhead, reduced complexity, and a low estimation error by utilizing compressive sensing methods. This type of problem has already been considered, but some drawbacks reduce its effectiveness. The SD algorithm, relies on a limited number of signal transmission paths, including one line-of-sight (LoS) path and two non-line-of-sight (NLoS), all of which must be orthogonal and independent. In addition, the algorithm must consider the multiple signal paths that arise due to the multi-path phenomenon of the mmWave in MIMO channels. Additionally, only a limited number of beams are selected based on the sparse channel's non-zero components (support), including the most active beam and seven symmetrically surrounding beams. However, if the number of beams is insufficient to model the channel, it would increase the estimation error.

In addition, we present an algorithm based on the convex optimization method, in which the channel estimation is independent of the number of paths and recognizes the sparsity level of the channel better. Hence, the estimation can be done without knowing the number of paths. The proposed method also has the advantage of being highly accurate and of reduced complexity. Additionally, to speed up the massive MIMO channel recovery and to deal with the sparse nature of mmWave channels, a new algorithm based on a greedy strategy is proposed for channel estimation in massive MIMO systems with a lens antenna array in mmWave communications.

Our main contributions are summarized as follows:

- We propose a new channel estimation algorithm based on compressive sensing, inspired by convex optimization methods for massive MIMO systems without requiring that the number of channel paths is known.
- We propose another low-complexity algorithm based on greedy algorithm for increasing the speed of recovering the components of a massive MIMO channel.
- Through extensive simulation, we demonstrate that the both proposed frameworks reduce estimation error and require fewer pilots for the channel estimation procedure compared with the SD algorithm, OMP and SMD.

The rest of the paper is organized as follows. Section IV introduces the system model and problem formulation. Section V summarizes the general situation and the necessary prerequisites for channel estimation and then presents channel estimation algorithms based on a group lasso with an effective support set and the greedy method, respectively. In Section VI, Section IV presents the simulation results from VI-A and VI-B subsections based on the proposed Group Lasso with Effective Support (GLES) and the Greedy Algorithm with Intelligent Selection Support (GAISS) algorithms, respectively. Finally, the conclusions are drawn in Section VII.

NOTATION:

Lowercase letters show vectors, and boldface uppercase letters show matrices; We illustrated scaler with small normal face and vectors with small bold face; the notations $(\cdot)^{\mathrm{T}}$, $(\cdot)^{*}$,$(\cdot)^{\mathrm{H}}$ and $\|\cdot\|$ represent transpose, conjugate, Hermitian operators and the Frobenius norm, respectively, and the trace of a matrix $\boldsymbol{A}$ is noted by tr($\boldsymbol{A}$).Card (A) shows the cardinality of a set A; $\|\mathrm{a}\|_2$ shows the l_2-norm of the vector a; the amplitude of vector a is noted by |a |, and $\boldsymbol{I}_K$ indicates the $K \times K$ identity matrix.

IV. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a massive MIMO time-division duplexing (TDD) system in mmWave communications. The base station has M antennas and M_{RF} radio frequency chains to serve K single-antenna users. For the channel estimation, we consider the downlink and uplink models. There are two subsections to a massive MIMO channel model in mmWave communications, including traditional model and lens antenna array. The channel model of a massive MIMO system in mmWave communication is presented by two subsections of traditional model and the lens antenna array.

A. TRADITIONAL MILLIMETER-WAVE MASSIVE MIMO

Figure 1(a) shows a massive MIMO system in the traditional mode in a mmWave communication system. The received signal Vector $\boldsymbol{y}^{DL}$ can be expressed as:

$$y^{DL} = \boldsymbol{H}^{H}\boldsymbol{G}s + v, \tag{1}$$

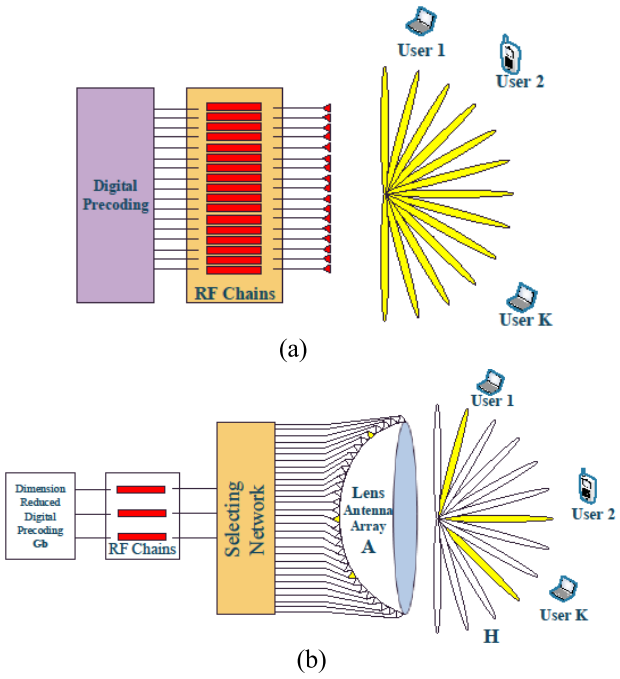

FIGURE 1. a) Architecture of traditional mmWave massive MIMO vs. (b) Architecture of mmWave massive MIMO with a lens antenna array.

where $\boldsymbol{s} = [s_1, s_2, \ldots, s_k]$ is the transmitted signal vector with dimensions $K \times 1$ for all K users. Its normalized power is $E(\boldsymbol{s}\boldsymbol{s}^H) = \boldsymbol{I}_K, \boldsymbol{H}^H \in C^{K\times M}$ shows the downlink matrix of the channel, that $\boldsymbol{H} = [\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_k]$ shows the uplink matrix of the channel based on channel reciprocity, the $M \times 1$ vector h_k is the spatial vector of the channel between the k^{th} user and the BS [51], [52]. The precoding matrix with dimension $M \times K$ is denoted by $\boldsymbol{G} = [\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k]$ such that $tr(\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}}) \leq \omega$, where ω is the downlink transmit power for $k = \{1, 2, \cdots, K\}$, and $\boldsymbol{v}$ is the noise vector drawn from the distribution $CN \sim (0, \delta_{DL}^2)$, where δ_{DL}^2 is the downlink noise power.

The number of RF chains in the massive MIMO systems in the traditional mode is the same as the number of antennas, see Figure 1 (a) [53], [54], e.g., $M_{RF} = M = 256$. In this paper, we assume the Saleh-Valenzuela model for the channel [55], [56], so the channel vector of k^{th} user ($\boldsymbol{h}_k$) can be expressed as:

$$\boldsymbol{h}_k = \sqrt{\frac{M}{L_k + 1}} \sum_{i=0}^{L_k} \gamma_k^{(i)} f\left(\partial_k^{(i)}\right) = \sqrt{\frac{M}{L_k + 1}} \sum_{i=0}^{L_k} c_{k,i}, \quad (2)$$

where $c_{k,0} = \gamma_k^{(0)} f\left(\partial_k^{(0)}\right)$ is the LoS component of $\boldsymbol{h}_k$ where $\gamma_k^{(0)}$ denotes the complex gain and $\partial_{\mathrm{k}}^{(0)}$ represents the spatial direction, $c_{k,i} = \gamma_k^{(i)} f\left(\partial_k^{(i)}\right)$ represents the i^{th} NLoS component of $\boldsymbol{h}_k$ for $1 \leq i \leq L_k$, and L_k is the number of NLoS components, $\partial_{\mathrm{k}}^{(\mathrm{i})}$ denoting the spatial direction, and $f(\partial)$ is the array steering vector with dimension $M \times 1$. For the uniform linear array (ULA) in typical mode, the array steering vector is:

$$\boldsymbol{f}(\partial) = \frac{1}{\sqrt{M}} [e^{-j2\pi\partial m}]_{m\epsilon I(M)}, \quad (3)$$

where $\boldsymbol{I}(M) = \{M - \frac{p-1}{2}, p = 0, 1, \ldots, M-1\}$ shows a set of centered indices around zero, symmetrically. The spatial direction obtains from $\partial = \frac{d}{\lambda} sin\theta$, where θ specifies the physical direction in the interval $[-\frac{\pi}{2} : \frac{\pi}{2}]$, the signal wavelength denotes λ, and the antenna spacing denotes d, which generally is $d = \lambda/2$ in mm-Wave communications.

B. MASSIVE MIMO WITH THE LENS ANTENNA ARRAY IN MILLIMETER-WAVE COMMUNICATION

The lens array can concentrate the signals from different directions (beams) on different antennas [57], [58]. Therefore, the spatial channel is converted into the beamspace channel, as shown in Figure 1 (b). Mathematically, the lens array antenna is the orthogonal beamforming vectors of M predefined beams in angular space that is as follows. Therefore, the system model of the lens antenna array is defined as:

$$\boldsymbol{A} = [f(\frac{1}{M}(\frac{3-M}{2}), \frac{1}{M}(\frac{5-M}{2}), \ldots, \frac{1}{M}(\frac{M-1}{2}))]^H. \quad (4)$$

where the received downlink signal vector is $\tilde{y}^{DL}$.

$$\tilde{\boldsymbol{y}}^{DL} = \boldsymbol{H}^H \boldsymbol{A}^H \boldsymbol{G}\boldsymbol{s} + \boldsymbol{n} = \tilde{\boldsymbol{H}}^H \boldsymbol{G}\boldsymbol{s} + \boldsymbol{n}, \quad (5)$$

In equation (5), $\tilde{H}$ represents the beamspace channel as:

$$\tilde{\boldsymbol{H}} = [\tilde{\boldsymbol{h}}_1, \tilde{\boldsymbol{h}}_2, \ldots, \tilde{\boldsymbol{h}}_K] = \boldsymbol{A}\boldsymbol{H} = [\boldsymbol{A}\boldsymbol{h}_1, \boldsymbol{A}\boldsymbol{h}_2, \ldots, \boldsymbol{A}\boldsymbol{h}_K]. \quad (6)$$

We select only a small number of dominant beams using a lens antenna array to reduce the MIMO dimension according to [45]:

$$\tilde{\boldsymbol{y}}^{DL} = \tilde{\boldsymbol{H}}_b^H \boldsymbol{G}_b s + n. \quad (7)$$

The dimension-reduced beamspace channel is shown by $\tilde{\boldsymbol{H}}_b = \tilde{\boldsymbol{H}}(j, :)_{j\epsilon k}$, where k represents the set of selected beams as illustrated in Figure 2 and $\boldsymbol{G}_b$ denotes the dimension-reduced digital precoding matrix with size $k \times K$ [52].

The interference-aware (IA) beam selection algorithm can be described as follows: (a) First, the algorithm identifies two types of users: interference-users (IUs) and non-interference-users (NIUs). This classification is based on the level of interference they experience in the system. (b) Second, the algorithm aims to find the best beam that is not shared with other users. This beam is considered the foremost unshared beam and is selected as the optimal choice. In summary, the IA beam selection algorithm identifies users experiencing interference and non-interference, and then selects the beam that offers the least interference by prioritizing unshared beams. This approach helps optimize the overall system performance and minimize interference-related issues [52].

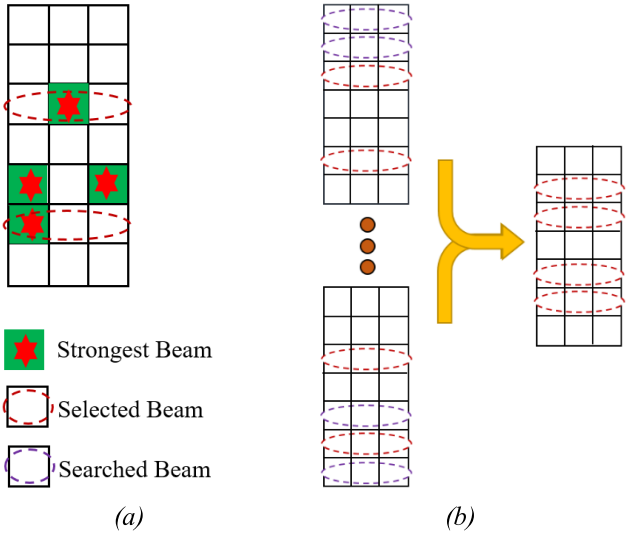

FIGURE 2. The interference-aware (IA) beam selection algorithm: (a) First: identify interference-users (IUs) and non-interference-users (NIUs); (b) Second: look for the foremost unshared beam.

The benefits of the lens antenna array are the low hardware cost/power consumption and the reduction of the number of RF chains. The smallest number of required RF chains is $M_{RF} = K$.

V. BEAMSPACE CHANNEL ESTIMATION

The users require to transmit Q samples to the BS as the known pilot sequences. These instants must be in block-wise form, with each block having a size of $K \times K$ [59]. The strategy that we have considered in this article is as follows: Q instants are divided into M blocks and each block consists of K instants, i.e., $Q = MK$. For the mth block, we define $\boldsymbol{\varphi}_m$ of size $K \times K$ as the pilot matrix, which contains K mutually orthogonal pilot sequences transmitted by K users over K instants [59]. Due to the normalization of the pilot power to the unit, the following relations are established:

$$\boldsymbol{\varphi}_m \boldsymbol{\varphi}_m^H = \boldsymbol{I}_k, \boldsymbol{\varphi}_m^H \boldsymbol{\varphi}_m = \boldsymbol{I}_k$$

As a result, the uplink received signal matrix is due to reciprocity in TDD systems and in the mth block, it is equal to:

$$\tilde{\boldsymbol{Y}}_m^{UP} = \boldsymbol{U}\boldsymbol{H}\boldsymbol{\varphi}_m + \boldsymbol{N}_m = \tilde{\boldsymbol{H}}\boldsymbol{\varphi}_m + \boldsymbol{N}_m \,, m = 1, 2, \ldots, M \quad (8)$$

During the pilot transmission in the mth block, the BS should employ a combiner $\boldsymbol{W}_m$ of size $K \times N$ to combine the received uplink signal matrix $\tilde{\boldsymbol{Y}}_m^{UP}$ (8). As a result, we have:

$$\boldsymbol{R}_m = \boldsymbol{W}_m \tilde{\boldsymbol{Y}}_m^{UP} = \boldsymbol{W}_m \tilde{\boldsymbol{H}} \boldsymbol{\varphi}_m + \boldsymbol{W}_m \boldsymbol{N}_m \quad (9)$$

where $\boldsymbol{N}_m$ is the $N \times K$ noise matrix in the mth block.

The $K \times K$ measurement matrix $\boldsymbol{Z}_m$ of the beamspace channel $\tilde{\boldsymbol{H}}$ achieved by multiplying the known pilot matrix $\boldsymbol{\varphi}_m^H$ on the right side of (9):

$$\boldsymbol{Z}_m = \boldsymbol{R}_m \boldsymbol{\varphi}_m^H = \boldsymbol{W}_m \tilde{\boldsymbol{H}} + \boldsymbol{N}_m^{eff} \quad (10)$$

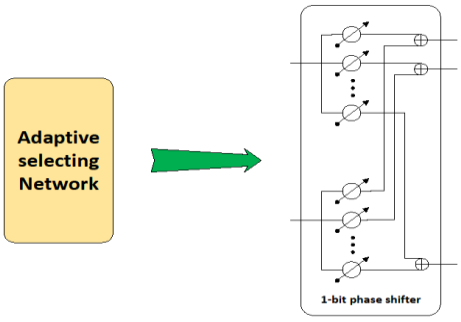

FIGURE 3. Adaptive selecting network for mmWave massive MIMO with lens antenna array.

In the above equation $\boldsymbol{N}_m^{eff} = \boldsymbol{W}_m \boldsymbol{N}_m \boldsymbol{\varphi}_m^H$ is the effective noise matrix. Finally, the M matrix of the $\tilde{\boldsymbol{H}}$ channel measurement block is equal to:

$$\boldsymbol{Z}_K = \boldsymbol{W}\tilde{\boldsymbol{h}}_K + \tilde{\boldsymbol{n}}_k, \quad (11)$$

where $\boldsymbol{Z}_K$ denotes the measurement matrix and, $\tilde{\boldsymbol{n}}_k$ shows the effective noise matrix.

In the system model considered for massive MIMO systems in mmWave communication with a lens antenna array, the selecting network is a one-bit phase shifter that replaces the selecting network block in Figure 1 (b). The adaptive selecting network is configured for beam selection during the data transmission (Figure 3). One of the benefits of this network is a significant reduction in energy consumption and cost [60], [61]. This network is used as a $\boldsymbol{W}_m$-combiner to combine the uplink signals and is designed to minimize cross-correlation [17]. The number of dominant scatters in the mmWave prorogation environments is limited because $\tilde{\boldsymbol{h}}_K$ is a sparse vector. Therefore, by utilizing the adaptive selecting network, it is guarantee that $\boldsymbol{Z}_K$ has complete information of $\tilde{\boldsymbol{h}}_K$ even if $Q < N$. Then, (11) can be formulated as a typical sparse signal recovery problem.

To guarantee the success of the CS reconstruction algorithms, the $\boldsymbol{W}$ matrix must satisfy the Restricted Isometry Property (RIP) condition. For matrix $\boldsymbol{W}$, the smallest $\Omega \geq 0$ represents the *sth* restricted isometry constant $\Omega_S = \Omega_S(\boldsymbol{W})$ such that (10) is met for every s-sparse vectors $\tilde{\boldsymbol{h}}_K$.

$$(1 - \Omega) \left\| \tilde{\boldsymbol{h}}_K \right\|_2^2 \leq \left\| \boldsymbol{W}\tilde{\boldsymbol{h}}_K \right\|_2^2 \leq (1 + \Omega) \left\| \tilde{\boldsymbol{h}}_K \right\|_2^2. \quad (12)$$

The Bernoulli random matrix is selected as the $\boldsymbol{W}$ matrix, where the columns are normalized to the unit norm.

A. CHANNEL ESTIMATION BASED ON THE PROPOSED GLES ALGORITHM

In order to solve (11), we can use greedy algorithms, such as the OMP or the CoSaMP. Due to the limited transmit power of the users, it is generally expected that the channel will be evaluated with a low SNR.

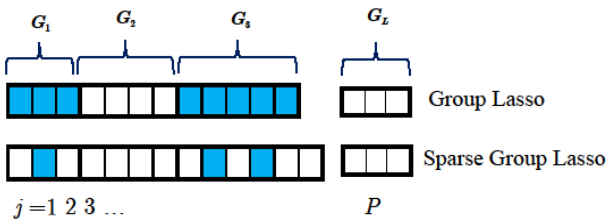

FIGURE 4. Group lasso method of a selection of blocks.

Therefore, the support function of the noisy $\tilde{\boldsymbol{h}}_K$, which is calculated by greedy methods, is inaccurate and leads to degraded performance.

In this paper, we exploit the beamspace channel structural characteristics for channel estimation purposes. In this regard, we propose the GLES algorithm, which is based on the group least absolute shrinkage and selection operator (LASSO). It is an algorithm from the convex optimization family with the possibility of the proper support selection [62]. This algorithm can estimate the channel more accurately, and it works better than the other existing methods, such as SD, OMP [63], [64], and SMD [38], especially in low SNR settings.

The GLES algorithm does not require prior knowledge of the number of sparse blocks and the size of each block. Hence, as the channel is sparse and the effective channel components exist only around the strongest one, the group lasso method can be used to appropriately estimate the desired channel [65]. As a result of the lens antenna array features, the proposed method operates on the proper block size detection, which is natural when the strongest channel elements are larger than the other blocks.

Figure 4 shows that the dominant block is active while the remaining blocks are set to zero. The advantage of this method is that the number of paths and the sparsity level of the beamspace channel does not have to be known.

Based on the group LASSO formulation, the estimation problem is formulated as follows. It first divides the channel vector into several blocks as shown in (13).

$$\tilde{\boldsymbol{h}}_K = \left(\tilde{\boldsymbol{h}}_{K1}, \tilde{\boldsymbol{h}}_{K2}, \ldots, \tilde{\boldsymbol{h}}_{KL_k}\right). \tag{13}$$

Using the equation below, the group lasso solves the channel estimation.

$$\hat{\tilde{\boldsymbol{h}}}_K = \text{argmin}_{\tilde{\boldsymbol{h}}_K}\left(\frac{1}{2} * \left\|\boldsymbol{Z}_K - \boldsymbol{W}\tilde{\boldsymbol{h}}_K\right\|_2^2 + \alpha \sum_{i=1}^{c} \left\|\tilde{\boldsymbol{h}}_K\right\|_2\right), \tag{14}$$

where α is the penalty parameter and, $\hat{\tilde{\boldsymbol{h}}}_K$ is the sparse estimation of $\tilde{\boldsymbol{h}}_K$, which minimizes the distance between $\boldsymbol{Z}_K$ and $\boldsymbol{W}\tilde{\boldsymbol{h}}_K$ [39], [66].

The group lasso solves the block basis pursuit problem using Lagrangian relaxation. An alternative direction multiplier method is used in one solution. Due to the uncertainty of the prediction, it does not work correctly. By using the ConVeX (CVX) optimization tool, we reduce the computational complexity of the sparsity recovery algorithms and improve the accuracy of the estimation [66].

The proposed algorithm works as follows. First, it solves the group lasso problem using the CVX in order to obtain the raw channel estimation $\tilde{\boldsymbol{h}}_K$. The index set of $\tilde{\boldsymbol{h}}_K$ is then sorted according to the descending order of the elements of $\tilde{\boldsymbol{h}}_{K1}$, which is shown as the set $\boldsymbol{S}k$. Now the first element of $\boldsymbol{S}k$, is taken as the index of the first main peak of the channel, and the $\boldsymbol{V}1$ elements around it are then selected as the first effective components of the channel, and they are represented with the set $\boldsymbol{T}1$. Actually, $\boldsymbol{T}1$ is the first partial support. In the next step, $\boldsymbol{T}1$ is removed from Sk. Again, as illustrated previously, the first element of Sk and the $\boldsymbol{V}2$ elements around it are selected as $\boldsymbol{T}2$, the second partial support. In this manner, the set $\boldsymbol{T}3$ is obtained. Finally, $\boldsymbol{T}1$, $\boldsymbol{T}2$, and $\boldsymbol{T}3$ are defined as effective support sets, and the other channel components are set to zero. The estimation method is summarized as Algorithm 1.

B. CHANNEL ESTIMATION BASED ON THE PROPOSED GAISS ALGORITHM

The recovery speed of optimization-based methods is insufficient for large-scale problems. In terms of measurement error, the proposed GLES method is robust; however, its computational complexity makes it unsuitable for large-scale problems, such as a large number of antennas. The GLES method also requires an adjustment for α in each SNR. In this section, we present an algorithm based on greedy methods.

This method has the advantages of fast reconstruction, low mathematical complexity, and simple geometric interpretations. Using this algorithm, which is based on a specific rule, one or more columns of the compressing matrix are identified in each iteration as candidates for active columns. The location of the non-zero components is determined first in each iteration of the sparse signal coefficients. These coefficients are then estimated. This section describes the proposed method in low SNRs using the forward-backward (FB) algorithm [67], [68], which is classified as a greedy algorithm. In the high SNR cases, we used a combination of the SD and OMP algorithms, except that the support set was selected intelligently to obtain the most accurate channel estimation.

The proposed algorithm consists of two parts, a) the channel estimation at high SNRs, and b) the channel estimation at low SNRs. Algorithm 2 describes the SD method. This algorithm first finds the position n_i^* of the strongest element of $\tilde{c}_{k,i}$ (the i-th channel component of $\tilde{\boldsymbol{h}}_K$ in the beamspace) by utilizing the low mutual coherence property of $\bar{\boldsymbol{W}}$.

Then supp($\tilde{c}_{k,i}$) utilizing the structural characteristic of mmWave beamspace channel as mentioned in step 9. After that, the non-zero elements of $\tilde{c}_{k,i}$ are estimated by LS algorithm and the effect of this channel component is removed. This process is repeated for all the channel paths. As mentioned at the beginning of the article, we use the

Algorithm 1 Proposed GLES-Based Channel Estimation Algorithm

Input : $\bar{Z}_K, \bar{W}$
Output: Estimated beamspace channel for all users
Measurement vector : $\bar{Z}_K$ in $1 \le k \le K$ (11);
Combining matrix: $\bar{W}$ in (11);
parameters: α, *L-k*, $V1$, $V2$, $V3$
for $k = 1$ to $k \le K$ **do**
  $\arg\min_{h_k} (\frac{1}{2} * \left\| \bar{Z}_K - \bar{W}\tilde{h}_K \right\|_2^2 + \alpha \sum_{i=1}^{L-k} \left\| \tilde{h}_K \right\|_2)$
  $Sk =$ *the index set that ordered by decreasing magnitude of entries of* $\tilde{h}_K$
  $s =$ *the first element of* Sk
  $T1 = \{s - V1/2, \ldots, s + (V1 + 2)/2\}$
  $Sk = Sk \backslash T1$
  $s =$ *the first element of* Sk
  $T2 = mod_N\{s - V2/2, \ldots, s + (V2 + 2)/2\}$
  $Sk = Sk \backslash T2$
  $s =$ *the first element of* Sk
  $T3 = \{s - V_3/2, \ldots, s + (V_3 + 2)/2\}$
  Set the entries of h_k to zero outside of $T1 \cup T2 \cup T3$
end for

strengths of this algorithm for our proposed work and fix its weaknesses and then propose the GAISS algorithm.

C. CHANNEL ESTIMATION FOR MASSIVE MIMO SYSTEMS WITH A LENS ANTENNA ARRAY IN MILLIMETER-WAVE COMMUNICATIONS AT HIGH SNRs

As previously mentioned, the purpose is to recover the channel, which is $\tilde{h}_K$, for the k^{th} user in (11). SD has the advantage of relatively good accuracy at high SNRs, as there is a small amount of noise, and channel recovery is easier. Figure 5 illustrates the performance of this algorithm. A channel has three paths, including one LoS path and two NLoS paths. This beam path has the most power, and it produces the biggest peak. The components around this peak, which are formed by the antennas around the most prominent peak, have a significant impact on the estimation of the channel. This figure shows that one of the disadvantages of the SD algorithm is that it does not detect the complete number of effective beams since it only calculates eight components around each peak by default.

The figure also shows that the components around the highest peak, which are comparable with two of the other peaks, were not estimated. As a result, these un/estimated components increase the error.

Therefore, we need to identify the number of components to estimate the channel correctly. By contrast, expanding the support set, the non-zero components, decreases the estimation accuracy in this method since it estimates some additional components incorrectly.

Due to the fact that the two NLoS paths have less power, the effective beams around them are negligible, which results in the NMSE increasing instead of decreasing.

On this basis, we estimated the NLOS paths based on the SD method, and for the LoS path estimation, which has more effective components around itself, we propose an improved strategy. In the proposed method, the two smaller peaks with the supports S_2 and S_3 are estimated by the SD algorithm (Algorithm 2) [17].

Algorithm 2 SD-Based Channel Estimation Algorithm

Input : $\bar{Z}_K, \bar{W}$
Measurement vector : $\bar{Z}_K$ in (11);
Combining matrix:$\bar{W}$ in (11);
All components of the channel: $L_k + 1$
The number of elements stored for each component: V
Initialization: $\tilde{c}_{k,i}^e = 0_{N\times 1}$ for $0 \le i \le L_k$, $\bar{Z}_k^{(0)} = \bar{Z}_K$
for $0 \le i \le L_k$
  Find the position of the strongest element of $\tilde{c}_{k,i}$ as: $n_i^* = \arg\left| \bar{w}_n^H \bar{Z}_k^{(i)} \right|$, $\bar{w}_n$ is the n-th column of $\bar{W}$;
  Find $\text{supp}\left(\tilde{c}_{k,i}\right) = \text{mod}_N \left\{n_i^* - \frac{V}{2}, \ldots, n_i^* + \frac{V}{2}\right\}$;
  LS estimation of the non-zero elements of $\tilde{c}_{k,i}$ as: $f_i = (\bar{W}_i^H \bar{W}_i)^{-1} \bar{W}_i^H \bar{Z}_k^{(i)}$, $\bar{W}_i = \bar{W}(:, b)_{b\in \text{supp}(\tilde{c}_{k,i})}$;
  From the estimated $\tilde{c}_{k,i}^e$ as $\tilde{c}_{k,i}^e\left(\text{supp}(\tilde{c}_{k,i})\right) = f_i$;
  Remove the influence of $\tilde{c}_{k,i}$ by $\bar{Z}_k^{(i+1)} = \bar{Z}_k^{(i)} - \bar{W}\tilde{c}_{k,i}^e$
  $i = i + 1$;
end for
$S_T = \bigcup_{0\le i\le L_k} \text{supp}\,(\tilde{c}_i)$;
$f_T = (\bar{W}_T^H \bar{W}_T)^{-1} \bar{W}_T^H \bar{Z}_k$, $\bar{W}_T = \bar{W}(:, b)_{b\in S_T}$;
$\tilde{h}_k^e = 0_{N\times 1}$, $\tilde{h}_k^e\,(S_T) = f_T$;
Output: $\tilde{h}_k^e \forall k$

$$\tilde{h}_K\,(S_2, S_3) = \hat{h}_K, \tag{15}$$

where $\tilde{h}_K(S_2, S_3)$ shows the estimated components of the SD algorithm's support of two small peaks. We then subtract the estimated $\tilde{h}_K\,(S_2, S_3)$ from the channel measurement vector, $\bar{Z}_K$.

Then we have:

$$\bar{\boldsymbol{Z}}_K = \bar{\boldsymbol{Z}}_K - \bar{\boldsymbol{W}}\tilde{\boldsymbol{h}}_K\left(\boldsymbol{S}_2, \boldsymbol{S}_3\right). \tag{16}$$

Now, the problem is to reconstruct the remaining components of the channel from $\bar{\boldsymbol{Z}}_K$. The following algorithm estimates the $\boldsymbol{S}_1$ support for the highest peak. The proposed algorithm is similar to the OMP method with the exception that the $\boldsymbol{S}_1$ support in each iteration was selected intelligently to obtain the most appropriate estimate for the highest peak. For this aim, we define a boundary as $\varepsilon \left\|\bar{\boldsymbol{Z}}_K\right\|_2$, where ε is adjusted according to the desired accuracy. According to the defined boundary, the algorithm is repeated until the norm of the residual vector $\boldsymbol{r}$ becomes smaller than this boundary and is consequently the most appropriate support $\boldsymbol{S}_1$ to be selected. In addition, we also set another control parameter, s_{msx}, to terminate the algorithm if the support size exceeds it, regardless of the upper bound. Therefore, the proposed algorithm will perform the following by selecting the initial values of $\boldsymbol{r}^{(0)} = \bar{\boldsymbol{Z}}_K$ and $\tilde{\boldsymbol{h}}_K^{(0)} = 0$ until the control conditions allow.

First, we obtain the correlation vector between the two variables [41], [42] $\boldsymbol{r}$ and $\bar{\boldsymbol{W}}$ as:

$$\mathrm{Corr}(n) = \bar{\boldsymbol{W}}^T \boldsymbol{r}^{(n-1)} / \left\|\bar{\boldsymbol{W}}^T\right\|_2. \tag{17}$$

Consider the variable $l^{(n)}$ as a column index of $\bar{\boldsymbol{W}}$ which has the greatest correlation with $\boldsymbol{r}^{(n-1)}$ in (17), which can be rewritten as:

$$l^{(\mathrm{n})} = \arg\left(\max\left(\mathrm{Corr}\left(n\right)\right)\right). \tag{18}$$

Now, select the support set as follows.

$$\boldsymbol{S}_1^{(n)} = \left\{l^{(n)} - [\frac{v-1}{2}]\right\} : \left\{l^{(n)} + [\frac{v-1}{2}]\right\}. \tag{19}$$

According to the selected support, the channel is estimated by the least squares method [34], [35], [36], [37], [38] as:

$$\tilde{\boldsymbol{h}}_K^{(n)} = \mathrm{zeros}\left(1, n\right). \tag{20}$$

$$\boldsymbol{f}_T = \left(\left(\bar{\boldsymbol{W}}_{S_1^{(n)}}^T \bar{\boldsymbol{W}}_{S_1^{(n)}}\right)^{-1} \bar{\boldsymbol{W}}_{S_1^{(n)}}^T\right) * \bar{\boldsymbol{Z}}_K. \tag{21}$$

$$\tilde{\boldsymbol{h}}_K(\boldsymbol{S}_1^{(\mathrm{n})}) = \boldsymbol{f}_T \tag{22}$$

Now, the residual vector is updated as follows:

$$\boldsymbol{r}^{(n)} = \acute{\bar{\boldsymbol{Z}}}_K - \bar{\boldsymbol{W}}^T \acute{\tilde{\boldsymbol{h}}}_K, \tag{23}$$

Until the inequality $\left\|r^{(n)}\right\|_2 > \varepsilon \left\|\bar{\boldsymbol{Z}}_K\right\|_2$ is met, the value of parameter v is updated as $v = v + 2$ and the steps are repeated from the beginning.

We form the support set, collectively the two previous support for the NLoS paths and the new $\boldsymbol{S}_1$ support for the LoS path.

$$\boldsymbol{S}_{total} = \left(\boldsymbol{S}_2, \boldsymbol{S}_3\right) \cup \boldsymbol{S}_1^{(n)}. \tag{24}$$

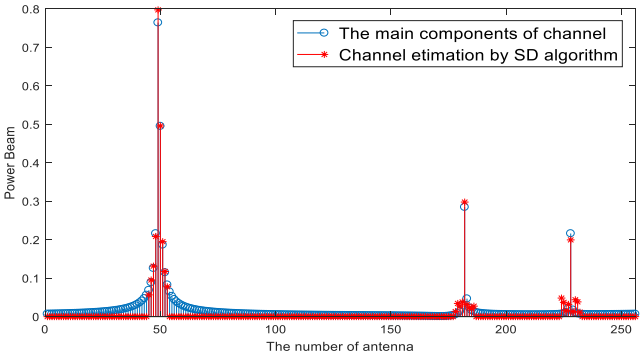

FIGURE 5. Demonstration of the performance of an SD algorithm in the channel components reconstruction at SNR = 30 for the k^{th} user.

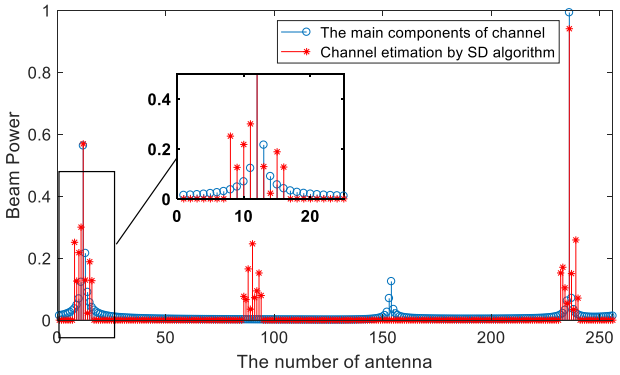

FIGURE 6. Channel reconstruction performance of the SD algorithm at *SNR = 0 dB*.

Finally, the estimated components are as follows.

$$\tilde{\boldsymbol{h}}_K^{(n)} = \mathrm{zeros}\left(1, n\right). \tag{25}$$

$$\boldsymbol{f}_T = \left(\left(\bar{\boldsymbol{W}}_{S\_total}^T \bar{\boldsymbol{W}}_{S\_total}\right)^{-1} \bar{\boldsymbol{W}}_{S\_total}^T\right) * \bar{\boldsymbol{Z}}_K \tag{26}$$

$$\tilde{\boldsymbol{h}}_K^{(n)}(\boldsymbol{S}\_total) = \boldsymbol{f}_T. \tag{27}$$

D. CHANNEL ESTIMATION FOR MASSIVE MIMO SYSTEMS WITH A LENS ANTENNA ARRAY IN MILLIMETER-WAVE COMMUNICATIONS AT LOW SNRs

The noise is noticeable when the SNR is low, and the channel recovery is difficult. Figure 6 shows that the SD algorithm estimates the channel inaccurately as well as the position of the transmitting antenna for the third peak and sometimes for the second peak, resulting in errors. The SD algorithm is used to estimate the highest peak for low SNRs. We propose the FB algorithm for the two small peaks, described in more detail below. FB is one of the greedy methods, which is a two-stage iterative algorithm. As the first step, the forward step estimates the length of the support set.

In the second step, which is the same length as the backward step, the components that have the least impact on the estimated support set are removed. The forward step is

longer than the backward step. In this way, the length of the support set increases with each iteration. Both steps are repeated until the residual power of the observation vector reaches a threshold. Unlike other similar algorithms, the FB algorithm does not need to be aware of the channel sparsity level. The CS converts the signal into a compressed and small-sized signal. Observations are made using the matrix Measure $\bar{\boldsymbol{W}}$. The algorithm is similar to OMP, SP, and SMD algorithms [69], [70]. The performance of this algorithm is as follows. We consider the highest peak with the support that is estimated by the SD algorithm as an acceptable estimate in the proposed method,

$$\tilde{\boldsymbol{h}}_K(\boldsymbol{S}_1) = \hat{\boldsymbol{h}}_K, \tag{28}$$

where $\tilde{\boldsymbol{h}}_K(S_1)$ shows the components of the highest peak estimated by the SD algorithm. We then subtract the estimated $\tilde{\boldsymbol{h}}_K(\boldsymbol{S}_1)$ from the channel measurement vector, $\bar{\boldsymbol{Z}}_K$, so we have:

$$\bar{\boldsymbol{Z}}_K = \bar{\boldsymbol{Z}}_K - \bar{\boldsymbol{W}}\tilde{\boldsymbol{h}}_K(\boldsymbol{S}_1). \tag{29}$$

We need to recover the channel peaks around the smaller of $\bar{\boldsymbol{Z}}_K$. The following algorithm estimates these supports.

The measurement vector $\bar{\boldsymbol{Z}}_K$ and the measurement matrix $\bar{\boldsymbol{W}}$ are applied as the input of the FB algorithm, which results in the output $\tilde{\boldsymbol{h}}_{\mathrm{FB},K}$. We now choose variables l_1 and l_2, as the corresponding indices of the two bigger components of $\tilde{\boldsymbol{h}}_{FB,K}$ and we then define the support set $\boldsymbol{S}_{2\_3}$ as follows:

$$\begin{aligned}\boldsymbol{S}_{2-3} = &\left\{l_1 - \left[\frac{v_1-1}{2}\right] : l_1 + \left[\frac{v_1-1}{2}\right]\right\} \\ &\cup \left\{l_2 - \left[\frac{v_1-1}{2}\right] : l_2 + \left[\frac{v_1-1}{2}\right]\right\}\end{aligned} \tag{30}$$

A community consisting of this estimated support and the support of the larger peak is selected as the final support. The channel under this support is estimated as follows using the least squares algorithm.

$$\boldsymbol{S}_{total} = \boldsymbol{S}_{2-3} \cup \boldsymbol{S}_1, \tag{31}$$

$$\tilde{\boldsymbol{h}}_K^{(n)} = zeros(1, n), \tag{32}$$

$$\boldsymbol{f}_T = \left(\left(\bar{\boldsymbol{W}}_{S\_total}^T \bar{\boldsymbol{W}}_{S\_total}\right)^{-1} \bar{\boldsymbol{W}}_{S\_total}^T\right) * \bar{\boldsymbol{Z}}_K, \tag{33}$$

$$\tilde{\boldsymbol{h}}_K^{(n)}(\boldsymbol{S}_{total}) = \boldsymbol{f}_T. \tag{34}$$

Algorithm 3 Proposed GAISS Channel Estimation Algorithm for High SNR Conditions

Input : $\bar{\boldsymbol{Z}}_K, \bar{\boldsymbol{W}}$
Output: Estimated beamspace channel for all users
Measurement vector : $\bar{\boldsymbol{Z}}_K$ in $1 \le k \le K$ (11);
Combining matrix:$\bar{\boldsymbol{W}}$ in (11);
parameters: ε, s_{msx}, $r^{(0)} = \bar{\boldsymbol{Z}}_K$, $\tilde{\boldsymbol{h}}_K^{(0)} = 0$, v
for each user $k = 1 to\ k \le K$
  Compute $\tilde{\boldsymbol{h}}_K(\boldsymbol{S}_2, \boldsymbol{S}_3) = \hat{h}_K by\ SD\_basedAlgorithm$
  Update residual signal: $\bar{\boldsymbol{Z}}_K = \bar{\boldsymbol{Z}}_K - \bar{\boldsymbol{W}}\tilde{\boldsymbol{h}}_K(S_2, S_3)$
  Compute the correlation:
  $\mathrm{Corr}(n) = \bar{\boldsymbol{W}}^T \boldsymbol{r}^{(n-1)} / \left\|\bar{\boldsymbol{W}}^T\right\|_2$
  Find the index of the maximum correlation:
  $l^{(n)} = \arg(\max(\mathrm{Corr}(n)))$.
  Select the support set as:
  $S_1^{(n)} = \left\{l^{(n)} - [\frac{v-1}{2}]\right\} : \left\{l^{(n)} + [\frac{v-1}{2}]\right\}$
  Repeat
    $\boldsymbol{f}_T = \left(\left(\bar{\boldsymbol{W}}_{S_1^{(n)}}^T \bar{\boldsymbol{W}}_{S_1^{(n)}}\right)^{-1} \bar{\boldsymbol{W}}_{S_1^{(n)}}^T\right) * \bar{\boldsymbol{Z}}_K\, \tilde{\boldsymbol{h}}_K(S_1^{(n)})^{(n)} = \boldsymbol{f}_T.$
    Update residual vector: $r^{(n)} = \bar{\boldsymbol{Z}}_K - \bar{\boldsymbol{W}}^T \tilde{\boldsymbol{h}}_K,$
    Increase the range size: $v = v + 2$
    Update $\boldsymbol{S}_1^{(n)} = \left\{l^{(n)} - [\frac{v-1}{2}]\right\} : \left\{l^{(n)} + [\frac{v-1}{2}]\right\}.$
  Until $\left\|\boldsymbol{r}^{(n)}\right\|_2 > \varepsilon \left\|\bar{\boldsymbol{Z}}_K\right\|_2$ **or** $\left|\boldsymbol{S}_1^{(n)}\right| < s_{max}$;
  Set $\boldsymbol{S}_{total} = (\boldsymbol{S}_2, \boldsymbol{S}_3) \cup \boldsymbol{S}_1^{(n)}$.
  Update the beamspace estimate for the total set: $\tilde{\boldsymbol{h}}_K^{(n)}(\boldsymbol{S}\_total) = \boldsymbol{f}_T$
end for

VI. SIMULATION RESULTS AND DISCUSSION

A. PROPOSED GLES ALGORITHM

In this section, we evaluate the performance of the proposed algorithms. We consider a massive MIMO system in the mmWave communication, where the BS equips a lens antenna array and has $N = 256$ antennas and $N_{\mathrm{RF}} = 16$ RF chains to serve $K = 16$ users. The number of resolvable paths for the k^{th} user in the spatial channel is set to be one LoS component and $L_k = 2$ NLoS components; while $\beta_k^{(0)} \sim \mathrm{CN}(0, 1)$, $\beta_k^{(i)} \sim \mathrm{CN}\left(0, 10^{-.5}\right)$ for $i = 1, 2$.

$\psi_k^{(0)}$ and $\psi_k^{(i)}$ follow the independent and identically distributed (i.i.d.) uniform distribution within $[-0{:}5;\ 0{:}5]$. Finally, the uplink and downlink SNR are defined as σ_{UP}^2 and ρ/σ_{UP}^2, respectively. The simulation setups are summarized in Table 1.

In Fig. 7, the NMSE performance of the proposed algorithm is compared with the SD-based, the OMP-based and the SMD-based channel estimation algorithms for the massive MIMO systems in mmWave communication. For the SD-based channel estimation algorithm, we consider the number of high-ower elements $V = 8$.

For the OMP-based channel estimation algorithm, the sparsity level of the beamspace channel is equal to $V(LK+1) =$

Algorithm 4 Proposed GAISS Channel Estimation Algorithm for Low SNR Conditions

Input : $\bar{\boldsymbol{Z}}_K, \bar{\boldsymbol{W}}$
Output: Estimated beamspace channel for all users
Measurement vector : $\bar{\boldsymbol{Z}}_K$ in $1 \leq k \leq K$ (11);
Combining matrix: $\bar{\boldsymbol{W}}$ in (11);
parameters: ε, s_{msx}, $\boldsymbol{r}^{(0)} = \bar{\boldsymbol{Z}}_K, \tilde{\boldsymbol{h}}_K^{(0)} = 0, v$
for each user k from $k = 1$to$k \leq K$
  $\tilde{\boldsymbol{h}}_K(\boldsymbol{S}_2, \boldsymbol{S}_3) = \hat{h}_K$ bySD_based Algorithm
  Update $\bar{\boldsymbol{Z}}_K = \bar{\boldsymbol{Z}}_K - \bar{\boldsymbol{W}}\tilde{\boldsymbol{h}}_K(S_1)$
  Solve $\bar{\boldsymbol{Z}}_K$ in step 2, using Forward-backward algorithm
  Obtain the output of step $3 = \tilde{\boldsymbol{h}}_{FB,K}$
  Set $l_{1,2} = \arg(\max\left(\tilde{\boldsymbol{h}}_{FB,K}\right)_{1,2})$
  $\boldsymbol{S}_{2-3} = \left\{l_1 - \left[\frac{v_1-1}{2}\right] : l_1 + \left[\frac{v_1-1}{2}\right]\right\} \cup \left\{l_2 - \left[\frac{v_1-1}{2}\right] : l_2 + \left[\frac{v_1-1}{2}\right]\right\}$
  $\boldsymbol{S}_{total} = \boldsymbol{S}_{2-3} \cup \boldsymbol{S}_1$,
  $\boldsymbol{f}_T = \left(\left(\bar{\boldsymbol{W}}_{S\_total}^T \bar{\boldsymbol{W}}_{S\_total}\right)^{-1} \bar{\boldsymbol{W}}_{S\_total}^T\right) * \bar{\boldsymbol{Z}}_K$,
  $\tilde{\boldsymbol{h}}_K^{(n)}(\boldsymbol{S}_{total}) = \boldsymbol{f}_T$.
end for

TABLE 1. The simulation setups.

The name / algorithm parameters	Proposed GLES algorithm	Proposed GAISS algorithm
number of beams (transmit antennas)	256	256
number of users (receive antennas)	16	16
RF chain	16	16
number of paths per user	3	3
Pilot transmission	96	96
wavelength of carrier (λ)	1	1
antenna spacing	0.5	0.5
number of SNRs	9	9
Elapsed time for 1 iteration in all of SNRs	0.208 seconds	463.19 seconds

24. This section assumes that the SD and OMP-based channel estimation and the proposed GLES-based channel estimation

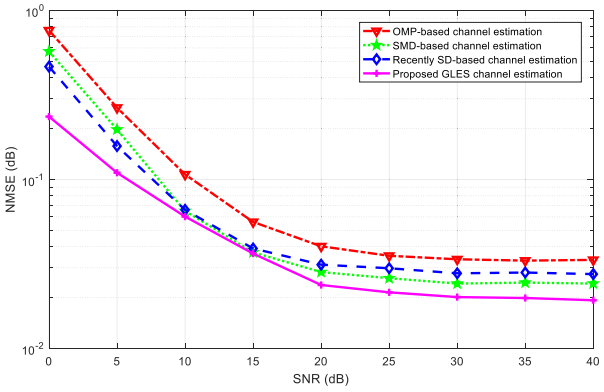

FIGURE 7. NMSE performance comparison among the OMP-based, the SD-based, the SMD-based and the proposed GLES-based channel estimation methods.

send $Q = 96$ instants for $M = 6$ blocks and the SMD-based channel estimation scheme operates $Q = N = 256$ instants for pilot transmission. The proposed GLES-based channel estimation algorithm reduces the estimated error significantly, as shown in Figure 7. Choosing the support set correctly allows us to estimate the main peaks of the channel as well as other effective components of the channel as defined in Algorithm 1, steps 8 to 16. Nevertheless, the error rate is much lower at low SNRs. This can be seen in Fig. 7; at SNR = 0dB, the NMSE of the SD, SMD, OMP and proposed GLES channel estimation algorithm is respectively 0:46, 0.57, 0.75 and 0:26, which is a 43% decline. This shows the superiority of convex optimization algorithms and higher accuracy in recovering signals in sparse environment, step 7 in algorithm1.

In Figure 8, the performance of the proposed GLES-based and SD-based channel estimation algorithms are shown. The proposed algorithm detects the efficient channel components more accurately than the SD-algorithm. As can be seen in the figure, the proposed algorithm detects the effective support for each transmission path (both LOS and NLOS) separately, the coordinates of the most effective beam in each path (the beam with the maximum power) and then the rest of the non-zero components of the channel. This causes the coordinates of the non-zero components to be recognized correctly and thus leads to a decrease in the NMSE value.

Therefore, with the proper block selection, effective support determination, and convex optimization algorithms, the proposed GLES-based channel estimation algorithm is highly accurate compared to greedy methods. Although greedy algorithms require less computation time, they do not always work well, especially in multi-path environments with high correlation. Instead, convex optimization methods are more accurate but slower to compute.

According to Figure 9, the proposed GLES-based channel estimation algorithm determines the sparsity level of the channel by finding the appropriate CVX values for each SNR. In fact, α is the penalty parameter in equation (14) a kind of compromise between the sparsity channel level and the estimation accuracy. In such a way that it applied imposed

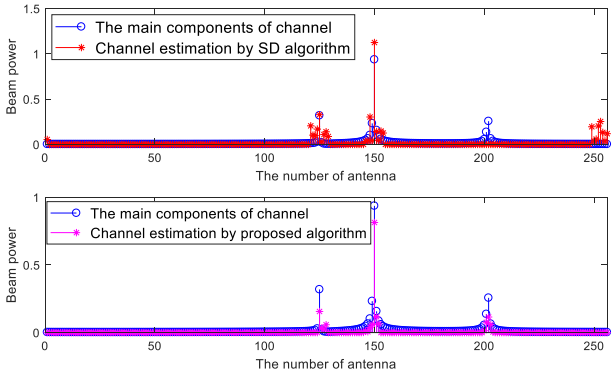

FIGURE 8. The main components of mmWave massive MIMO channel comparison between SD (top) and proposed GLES-based (low) channel estimation by the k^{th} users.

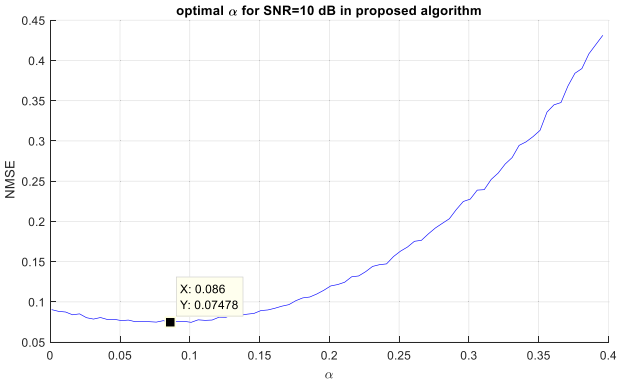

FIGURE 9. The NMSE graph vs. α for SNR=10 dB.

a regularization penalty of the sum of the L2 norm on groups that guaranteed that few groups were selected. But if a group is selected, so are all the predictors in it. The NMSE graph is plotted versus α for SNR = 10dB, as shown in Figure 9. The best α for the proposed algorithm is about 0:086 for SNR =10dB.

Thus, the best α for the other SNRs is obtained using the CVX and is placed in (11). The optimum α for the other SNRs is obtained according to Table 2.

B. PROPOSED GAISS ALGORITHM

In this section, we examine the performance of the second proposed algorithm. Because this algorithm is based on greedy methods with the possibility of smart support selection, it can achieve the desired accuracy error in a reasonable amount of time. The NMSE value in SNR = 0dB has decreased from NMSE = 0:46 for SD-based channel estimation to NMSE = 0:36 for the proposed GAISS-based channel estimation algorithm, as shown in Figure 10.

The simulation results are consistent with algorithms 3 and 4. Considering that high and low SNRs are separated, it is clear that the proposed algorithm works well in all SNRs. For high SNRs, considering that the SD-based algorithm correctly estimates the NLOS paths, for these two smaller peaks,

TABLE 2. Optimum α for different SNRs.

SNR (dB)	0	5	10	15	20	25	30	35	40
α	45	17.455	0.086	2	0.086	0.086	0.086	0.086	0.086

the SD-based algorithm is used according to equation (15). But this algorithm does not have a good estimation accuracy for the LOS path, and we also know that the components around the largest peak of the LOS path have a significant effect on the channel estimation accuracy. Therefore, the GIASS-based proposed algorithm was fully described according to the relationships (16) to (27) for high SNR by intelligently determining the support set to estimate the LOS path and also defining the threshold limit to prevent the estimation of additional components that lead to an increase in NMSE. How the algorithm works is summarized in Algorithm 3.

For low SNRs, according to Figure 6, it was observed that unlike the previous case, the SD-based algorithm works successfully for the LOS path, but it needs to be modified for the two NLOS paths. Therefore, Algorithm 4 was presented specifically for low SNRs, so that it estimates the LOS path according to Algorithm 2 and equation (28), then for two NLOS paths, the algorithm uses the FB method due to its forward and backward property. It is fully described in equations (29) to (34). How the algorithm works is given in Algorithm 4. The proposed GAISS method has the advantage of not being dependent on the sparsity level of the channel, and it can select the support set intelligently, resulting in a better performance than existing algorithms.

On the other hand, the proposed method performs very well under low SNR conditions. The downlink SNR is low due to the limited transmission power of the users, so the proposed GAISS algorithm performs better in these conditions. By considering how many training instants are sent at SNR = 25 dB, Figure 11 compares the previously mentioned algorithms based on NMSE performance. For example, to achieve NMSE= 0:35, the number of Qs required in the OMP and the SD algorithms are $Q = 180$ and $Q = 120$, respectively, and it reaches $Q = 100$ in the proposed GAISS algorithm. As a result, GAISS has an overhead reduction of 44% compared to OMP and 16% to SD. Therefore, the GAISS algorithm has a satisfactory performance with a low pilot overhead.

In summary, we first proposed an algorithm that has a better performance in terms of accuracy with emphasis on having less information about the channel. Hence, the estimation can be done without knowing the number of paths The cost of this better performance is its high time complexity. The NMSE of the SD, SMD, OMP and proposed GLES channel estimation algorithm is respectively 0:46, 0.57, 0.75 and 0:26, which is a 43% decline. On the other hand, for the applications where this computational complexity is not acceptable or

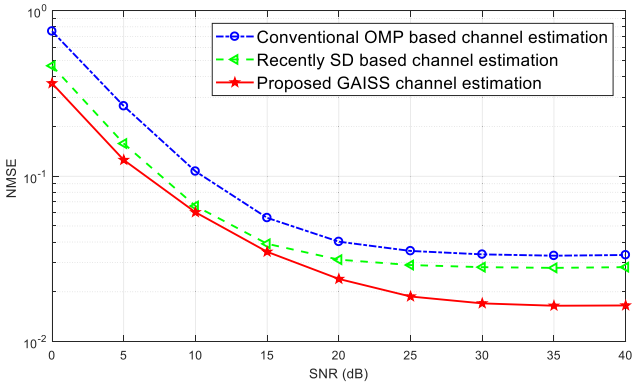

FIGURE 10. NMSE performance comparison among the OMP-based, the SD-based, and the proposed GAISS- based channel estimation methods.

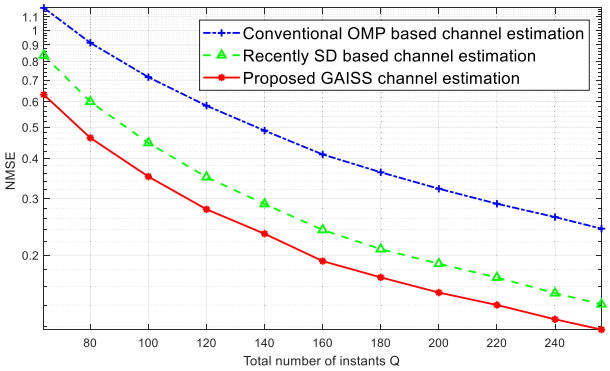

FIGURE 11. NMSE performance comparison against the total number of instants Q for the pilot transmission.

answerable, we proposed a second algorithm that has reduced this computational complexity. So that this amount of time increase in the second method is very small compared to the previous methods. That is, the calculation time for the second proposed method is equal to 0.208 seconds per iteration, while the best processing time in the existing methods belongs to the SD method and is equal to 0.187 seconds. It can be seen that this computational increase is very small compared to the sign- ficant improvement of the NMSE metric. Finally, to compare the time complexity, we have compared both proposed methods for one iteration in all of SNRs and listed them in the table 1.

Lastly, Figure 12 shows a comparison of the two proposed algorithms with the SD, SMD and OMP algorithms. As shown in the figure, due to the ability of convex optimization algorithms to recover sparse noisy signals, the proposed GLES algorithm outperforms the greedy GAISS algorithm under low SNR conditions. The greedy GAISS algorithm, however, performs better than the SD, SMD and OMP under low-SNR conditions and is comparable to the GLES optimization algorithm. However, the GAISS algorithm performs better than the GLES optimization algorithm at high SNRs because it selects support sets intelligently. In addition, it does not need to adjust for each SNR in order to achieve its

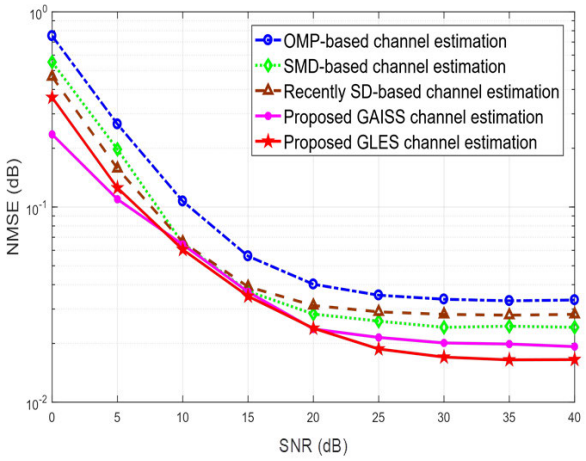

FIGURE 12. NMSE performance comparison among OMP-based, SMD-based, SD-based and two proposed GAISS and GLES-based channel estimation methods.

corresponding sparsity coefficient in the GLES algorithm. Note that Algorithms 3 and 4 are used separately during the simulation process due to the signal-to-noise problem. Each algorithm is equipped with a threshold that allows choosing the appropriate approach based on the user's signal strength and the surrounding environment. This provides flexibility for the user to choose and use any of the algorithms based on the signal they send.

VII. CONCLUSION

This paper proposed two channel estimation techniques for massive MIMO systems with lens antenna arrays in mmWave communications for single-antenna users. We used the training-based method, due to its low complexity. Aside from that, the sparsity of the channel encouraged the use of compressed sensing. Two new algorithms based on the optimization method and greedy strategy have been proposed for channel estimation in massive MIMO systems with a lens antenna array in mmWave communications. The simulation results have showed that the proposed optimization-based GLES algorithm is accurate, especially under low SNR conditions.

Also, this algorithm does not require prior knowledge of the number of sparse blocks and the size of each block. This means that the estimation is independent of the number of paths, so we need less information about the channel and thus it helps to simplify the problem a lot. But it has a slow convergence rate and a high computational complexity. To achieve high accuracy, high convergence speed, and low computational complexity, the greedy GAISS algorithm has been proposed. The performance of the proposed algorithms was compared to that of other existing methods. In addition to reducing estimation error, the proposed methods need fewer pilots for the channel estimation procedure.

A channel estimation algorithm was investigated in this article for systems with a single antenna user. It will be possible in the future to investigate channel estimation in

these systems for users with multiple antennas. One of the challenges that can be addressed in the future is reducing the number of pilots by designing appropriate precoding and training sequences. Future activities could include examining these systems in cellular communications and increasing the number of NOMA users. Also, the use of large Intelligence Surface (LIS or RIS) which has received much attention due to its simple and inexpensive structure. Last but not least, these systems can be estimated using the proposed estimation methods as well as group lasso methods.

REFERENCES

[1] S. Alraih, I. Shayea, M. Behjati, R. Nordin, N. F. Abdullah, A. Abu-Samah, and D. Nandi, "Revolution or evolution? Technical requirements and considerations towards 6G mobile communications," *Sensors*, vol. 22, no. 3, p. 762, Jan. 2022.

[2] M. U. A. Siddiqui, F. Qamar, F. Ahmed, Q. N. Nguyen, and R. Hassan, "Interference management in 5G and beyond network: Requirements, challenges and future directions," *IEEE Access*, vol. 9, pp. 68932–68965, 2021.

[3] R. Ali, Y. B. Zikria, A. K. Bashir, S. Garg, and H. S. Kim, "URLLC for 5G and beyond: Requirements, enabling incumbent technologies and network intelligence," *IEEE Access*, vol. 9, pp. 67064–67095, 2021.

[4] A. A. Raja, H. Pervaiz, S. A. Hassan, S. Garg, M. S. Hossain, and M. J. Piran, "Coverage analysis of mmWave and THz-enabled aerial and terrestrial heterogeneous networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 11, pp. 22478–22491, Nov. 2022.

[5] J. M. Sheeba and S. Deepa, "Beamforming techniques for millimeter wave communications—A survey," in *Proc. Int. Conf. Emerg. Current Trends Comput. Expert Technol.*, Chennai, India, 2020, pp. 1563–1573.

[6] N. Unnisa and M. Tatineni, "Adaptive deep learning strategy with red deer algorithm for sparse channel estimation and hybrid precoding in millimeter wave massive MIMO-OFDM systems," *Wireless Pers. Commun.*, vol. 122, no. 4, pp. 3019–3051, Feb. 2022.

[7] B. Zheng, C. You, W. Mei, and R. Zhang, "A survey on channel estimation and practical passive beamforming design for intelligent reflecting surface aided wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 2, pp. 1035–1071, 2nd Quart., 2022.

[8] A. M. Elbir and S. Coleri, "Federated learning for channel estimation in conventional and RIS-assisted massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 4255–4268, Jun. 2022.

[9] P. Pasangi, M. Atashbar, and M. M. Feghhi, "Blind downlink channel estimation of multi-user multi-cell massive MIMO system in presence of the pilot contamination," *AEU—Int. J. Electron. Commun.*, vol. 117, Apr. 2020, Art. no. 153099.

[10] P. Pasangi, M. Atashbar, and M. M. Feghhi, "Blind downlink channel estimation for TDD-based multiuser massive MIMO in the presence of nonlinear HPA," *ETRI J.*, vol. 41, no. 4, pp. 426–436, Aug. 2019.

[11] M. Naraghi-Pour, M. Rashid, and C. Vargas-Rosales, "Semi-blind channel estimation and data detection for time-varying massive MIMO system," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2021, pp. 1–6.

[12] Z. Yang, H. Shen, W. Xu, and C. Zhao, "Jamming suppression for uplink massive MIMO systems: A semi-blind receiver design," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Aug. 2022, pp. 678–683.

[13] A. Afshar, V. T. Vakili, and S. Daei, "Active user detection and channel estimation for spatial-based random access in crowded massive MIMO systems via blind super-resolution," *IEEE Signal Process. Lett.*, vol. 29, pp. 1072–1076, 2022.

[14] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang, *MIMO-OFDM Wireless Communications With MATLAB*. Hoboken, NJ, USA: Wiley, 2010.

[15] J. Yang, Y. Zeng, S. Jin, C.-K. Wen, and P. Xu, "Communication and localization with extremely large lens antenna array," *IEEE Trans. Wireless Commun.*, vol. 20, no. 5, pp. 3031–3048, May 2021.

[16] N. Shalavi, M. Atashbar, and M. M. Feghhi, "Downlink channel estimation of FDD based massive MIMO using spatial partial-common sparsity modeling," *Phys. Commun.*, vol. 42, Oct. 2020, Art. no. 101138.

[17] X. Gao, L. Dai, S. Han, and X. Wang, "Reliable beamspace channel estimation for millimeter-wave massive MIMO systems with lens antenna array," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6010–6021, Sep. 2017.

[18] G. Azarnia, M. A. Tinati, and T. Y. Rezaii, "Generic cooperative and distributed algorithm for recovery of signals with the same sparsity profile in wireless sensor networks: A non-convex approach," *J. Supercomput.*, vol. 75, no. 5, pp. 2315–2340, May 2019.

[19] G. Azarnia, M. A. Tinati, and T. Y. Rezaii, "Cooperative and distributed algorithm for compressed sensing recovery in WSNs," *IET Signal Process.*, vol. 12, no. 3, pp. 346–357, May 2018.

[20] L. Afeef, G. Mumcu, and H. Arslan, "Energy and spectral-efficient lens antenna subarray design in mmWave MIMO systems," *IEEE Access*, vol. 10, pp. 75176–75185, 2022.

[21] Q. Abdullah, N. Abdullah, A. Salh, L. Audah, N. Farah, A. Ugurenver, and A. Saif, "Pilot contamination elimination for channel estimation with complete knowledge of large-scale fading in downlink massive MIMO systems," 2021, *arXiv:2106.13507*.

[22] S. Khera and S. Singh, "Estimation of channel for millimeter-wave hybrid massive MIMO systems using orthogonal matching pursuit (OMP)," *J. Phys., Conf. Ser.*, vol. 2327, no. 1, 2022, Art. no. 012040.

[23] K. Dovelos, M. Matthaiou, H. Q. Ngo, and B. Bellalta, "Channel estimation and hybrid combining for wideband terahertz massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1604–1620, Jun. 2021.

[24] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Appl. Comput. Harmon. Anal.*, vol. 27, no. 3, pp. 265–274, Nov. 2009.

[25] J. S. Donato and H. W. Levinson, "Structured iterative hard thresholding with on- and off-grid applications," *Linear Algebra Appl.*, vol. 638, pp. 46–79, Apr. 2022.

[26] R. Wang and X. Peng, "Normalized iterative hard thresholding in high dimensional logistic regression with metagenomics data," *Int. J. Intell. Technol. Appl. Statist.*, vol. 15, no. 1, pp. 1–12, 2022.

[27] T. Blumensath and M. E. Davies, "Normalized iterative hard thresholding: Guaranteed stability and performance," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 298–309, Apr. 2010.

[28] S. Foucart, "Hard thresholding pursuit: An algorithm for compressive sensing," *SIAM J. Numer. Anal.*, vol. 49, no. 6, pp. 2543–2563, Jan. 2011.

[29] J.-F. Cai, J. Li, X. Lu, and J. You, "Sparse signal recovery from phaseless measurements via hard thresholding pursuit," *Appl. Comput. Harmon. Anal.*, vol. 56, pp. 367–390, Jan. 2022.

[30] Y. Erlich, N. Shental, A. Amir, and O. Zuk, "Compressed sensing approach for high throughput carrier screen," in *Proc. 47th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Sep. 2009, pp. 539–544.

[31] P. Schniter, L. C. Potter, and J. Ziniel, "Fast Bayesian matching pursuit," in *Proc. Inf. Theory Appl. Workshop*, Jan. 2008, pp. 326–333.

[32] L. Qin, Y. Cao, X. Shao, Y. Luo, X. Rao, Y. Yi, and G. Lei, "A deep heterogeneous optimization framework for Bayesian compressive sensing," *Comput. Commun.*, vol. 178, pp. 74–82, Oct. 2021.

[33] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Bayesian compressive sensing using Laplace priors," *IEEE Trans. Image Process.*, vol. 19, no. 1, pp. 53–63, Jan. 2010.

[34] M. Masood and T. Y. Al-Naffouri, "Sparse reconstruction using distribution agnostic Bayesian matching pursuit," *IEEE Trans. Signal Process.*, vol. 61, no. 21, pp. 5298–5309, Nov. 2013.

[35] G. Azarnia, "Distribution agnostic Bayesian compressive sensing with incremental support estimation," *Multidimensional Syst. Signal Process.*, vol. 33, no. 2, pp. 327–340, Jun. 2022.

[36] H. Song, X. You, C. Zhang, O. Tirkkonen, and C. Studer, "Minimizing pilot overhead in cell-free massive MIMO systems via joint estimation and detection," in *Proc. IEEE 21st Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, May 2020, pp. 1–5.

[37] H. Song, T. Goldstein, X. You, C. Zhang, O. Tirkkonen, and C. Studer, "Joint channel estimation and data detection in cell-free massive MU-MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 4068–4084, Jun. 2022.

[38] J. Hogan and A. Sayeed, "Beam selection for performance-complexity optimization in high-dimensional MIMO systems," in *Proc. Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2016, pp. 337–342.

[39] S. Chen, J. Zhang, E. Bjornsont, Ö. T. Demir, and B. Ai, "Sparse large-scale fading decoding in cell-free massive MIMO systems," in *Proc. IEEE 23rd Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Mar. 2022, pp. 1–5.

[40] P. Paul and M. M. Mowla, "3D metallic plate lens antenna based beamspace channel estimation technique for 5G mmWave massive MIMO systems," *Int. J. Wireless Mobile Netw.*, vol. 13, no. 1, pp. 1–16, Feb. 2021.

[41] T. Jiang, M. Song, X. Zhao, and X. Liu, "Channel estimation for millimeter wave massive MIMO systems using separable compressive sensing," *IEEE Access*, vol. 9, pp. 49738–49749, 2021.
[42] Z. Albataineh, K. Hayajneh, H. B. Salameh, C. Dang, and A. Dagmseh, "Robust massive MIMO channel estimation for 5G networks using compressive sensing technique," *AEU-Int. J. Electron. Commun.*, vol. 120, Jun. 2020, Art. no. 153197.
[43] X. Shi, J. Wang, and J. Song, "Triple-structured compressive sensing-based channel estimation for RIS-aided MU-MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 12, pp. 11095–11109, Dec. 2022.
[44] A. Liu, F. Zhu, and V. K. N. Lau, "Closed-loop autonomous pilot and compressive CSIT feedback resource adaptation in multi-user FDD massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 65, no. 1, pp. 173–183, Jan. 2017.
[45] I. Khan, M. Zafar, M. Jan, J. Lloret, M. Basheri, and D. Singh, "Spectral and energy efficient low-overhead uplink and downlink channel estimation for 5G massive MIMO systems," *Entropy*, vol. 20, no. 2, p. 92, Jan. 2018.
[46] Y. Han, Q. Liu, C.-K. Wen, S. Jin, and K.-K. Wong, "FDD massive MIMO based on efficient downlink channel reconstruction," *IEEE Trans. Commun.*, vol. 67, no. 6, pp. 4020–4034, Jun. 2019.
[47] O. A. Saraereh, I. Khan, Q. Alsafasfeh, S. Alemaishat, and S. Kim, "Low-complexity channel estimation in 5G massive MIMO-OFDM systems," *Symmetry*, vol. 11, no. 5, p. 713, May 2019.
[48] A. Waseem, A. Naveed, S. Ali, M. Arshad, H. Anis, and I. M. Qureshi, "Compressive sensing based channel estimation for massive MIMO communication systems," *Wireless Commun. Mobile Comput.*, vol. 2019, pp. 1–15, May 2019.
[49] B. Han, Z. Jiang, L. Liang, P. Chen, F. Yang, and Q. Bi, "Joint precoding and scheduling algorithm for massive MIMO in FDD multi-cell network," *Wireless Netw.*, vol. 25, no. 1, pp. 75–85, Jan. 2019.
[50] Z. Zhou, L. Liu, and J. Zhang, "FD-MIMO via pilot-data superposition: Tensor-based DOA estimation and system performance," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 931–946, Sep. 2019.
[51] R. S. Sohal, V. Grewal, J. Kaur, and M. L. Singh, "Deep learning based analog beamforming design for millimetre wave massive MIMO system," *Wireless Pers. Commun.*, vol. 126, no. 1, pp. 701–717, Sep. 2022.
[52] 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.
[53] X. Wei, C. Hu, and L. Dai, "Deep learning for beamspace channel estimation in millimeter-wave massive MIMO systems," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 182–193, Jan. 2021.
[54] M. A. L. Sarker, M. F. Kader, and D. S. Han, "Rate-loss mitigation for a millimeter-wave beamspace MIMO lens antenna array system using a hybrid beam selection scheme," *IEEE Syst. J.*, vol. 14, no. 3, pp. 3582–3585, Sep. 2020.
[55] Z. Xiao, P. Xia, and X.-G. Xia, "Channel estimation and hybrid precoding for millimeter-wave MIMO systems: A low-complexity overall solution," *IEEE Access*, vol. 5, pp. 16100–16110, 2017.
[56] A. Salh, L. Audah, Q. Abdullah, Ö. Aydogdu, M. A. Alhartomi, S. H. Alsamhi, F. A. Almalki, and N. S. M. Shah, "Low computational complexity for optimizing energy efficiency in mm-wave hybrid precoding system for 5G," *IEEE Access*, vol. 10, pp. 4714–4727, 2022.
[57] C. Ouyang, H. Xu, X. Zang, and H. Yang, "Lens antenna arrays-assisted mmWave MU-MIMO uplink transmission: Joint beam selection and phase-only beamforming design," 2022, *arXiv:2207.09636*.
[58] S.-H. Park, B. Kim, D. K. Kim, L. Dai, K.-K. Wong, and C.-B. Chae, "Beam squint in ultra-wideband mmWave systems: RF lens array vs. phase-shifter-based array," *IEEE Wireless Commun.*, early access, May 9, 2022, doi: 10.1109/MWC.007.2100530.
[59] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[60] B. Wang, L. Dai, Z. Wang, N. Ge, and S. Zhou, "Spectrum and energy-efficient beamspace MIMO-NOMA for millimeter-wave communications using lens antenna array," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2370–2382, Oct. 2017.
[61] X. Gao, L. Dai, S. Han, and R. W. Heath Jr., "Energy-efficient hybrid analog and digital precoding for mmWave MIMO systems with large antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 998–1009, Apr. 2016.
[62] L. Bottou, F. E. Curtis, and J. Nocedal, "Optimization methods for large-scale machine learning," *SIAM Rev.*, vol. 60, no. 2, pp. 223–311, Jan. 2018.
[63] S. Park and R. W. Heath Jr., "Spatial channel covariance estimation for the hybrid MIMO architecture: A compressive sensing-based approach," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8047–8062, Dec. 2018.
[64] E. Khordad, I. B. Collings, S. V. Hanly, and G. Caire, "Compressive sensing-based beam alignment schemes for time-varying millimeter-wave channels," *IEEE Trans. Wireless Commun.*, vol. 22, no. 3, pp. 1604–1617, Mar. 2023.
[65] Y. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2012.
[66] N. Simon, J. Friedman, T. Hastie, and R. Tibshirani, "A sparse-group lasso," *J. Comput. Graph. Statist.*, vol. 22, no. 2, pp. 231–245, May 2012.
[67] M. Pan, P. Liu, S. Liu, W. Qi, Y. Huang, X. You, X. Jia, and X. Li, "Efficient joint DOA and TOA estimation for indoor positioning with 5G picocell base stations," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–19, 2022.
[68] A. Bazzi, D. T. M. Slock, and L. Meilhac, "On spatio-frequential smoothing for joint angles and times of arrival estimation of multipaths," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2016, pp. 3311–3315.
[69] N. B. Karahanoglu and H. Erdogan, "Compressed sensing signal recovery via forward–backward pursuit," *Digit. Signal Process.*, vol. 23, no. 5, pp. 1539–1548, Sep. 2013.
[70] L. Li, M. Zhou, Y. Zhu, Y. Dai, and X. Liang, "Satellite microvibration measurement based on distributed compressed sensing," *Measurement*, vol. 203, Nov. 2022, Art. no. 112031.

ELHAM SHARIFI received the B.E. degree in electrical engineering from the University of Urmia, Iran, in 2015, and the M.S. degree in communication engineering from the University of Tabriz, Iran, in 2019. She is currently pursuing the Ph.D. degree with the Department of Electronics Engineering, Tarbiat Modares University, Tehran, Iran. Her research interests include optical wireless communications, deep neural networks, and 5G, 6G, massive MIMO, IRS, and satellite communications with an emphasis on signal detection and pre-coding.

MAHMOOD MOHASSEL FEGHHI received the B.S. and M.S. degrees (Hons.) in electrical engineering from the Iran University of Science and Technology, Tehran, Iran, in 2006 and 2009, respectively, and the Ph.D. degree in electrical engineering from the College of Engineering, University of Tehran, Tehran, in 2015. From 2007 to 2016, he was a Senior Design Engineer in communication systems design with several communications industries and Inc. Since 2016, he has been with the Faculty of Electrical and Computer Engineering, University of Tabriz, Iran, where he is currently an Associate Professor and the Director-in-Charge of the *Journal of Advanced Signal Processing*. His current research interests include information theory, wireless communication networks, signal processing, machine learning, and optimization.

GHANBAR AZARNIA received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering (communication) from the University of Tabriz, Tabriz, Iran, in 2011, 2013, and 2017, respectively. He is currently an Assistant Professor with the Engineering Faculty of Khoy, Urmia University of Technology, Urmia, Iran. His current research interests include compressive sensing, statistical signal processing, biomedical signal processing, adaptive filtering, wireless sensor networks, and distributed processing.

SAJJAD NOURI received the B.S. degree in communication engineering from Urmia University, Urmia, Iran, and the M.S. degree in communication systems engineering from the University of Tabriz, Tabriz, Iran, in 2013 and 2016, respectively. He is currently a Sales Engineering Expert with the manufacturing industry of circuit breakers and disconnectors. His current research interests include massive MIMO systems, mmWave communications, energy harvesting wireless communications, simultaneous wireless information and power transfer (SWIPT), wireless Internet of Things (IoT), smart cities, and medium and high voltage switchgears.

DUEHEE LEE (Member, IEEE) received the B.S. degree in electronic and electrical engineering from the Pohang University of Science and Technology, Pohang, South Korea, in 2004, and the M.S. and Ph.D. degrees in electrical and computer engineering from The University of Texas at Austin, Austin, TX, USA, in 2009 and 2015, respectively. He is currently an Associate Professor with the Department of Electrical and Electronics Engineering, Konkuk University.

MD. JALIL PIRAN (Senior Member, IEEE) received the Ph.D. degree in electronics and information engineering from Kyung Hee University, South Korea, in 2016.

He was a Postdoctoral Fellow with the Networking Laboratory, Kyung Hee University. He is currently an Assistant Professor with the Department of Computer Science and Engineering, Sejong University, Seoul, South Korea. He has published a substantial number of technical papers in well-known international journals and conferences in the field of intelligent information and communication technology (IICT), specifically in the fields of machine learning, data science, wireless communications and networking, 5G/6G, the Internet of Things (IoT), and cyber security.

Prof. Piran received the IAAM "Scientist Medal of the Year 2017" for notable and outstanding research in new-age technology and innovation, Stockholm, Sweden. In 2017, he was recognized as an "Outstanding Emerging Researcher" by the Iranian Ministry of Science, Technology, and Research. The Ph.D. dissertation has been selected as the "Dissertation of the Year 2016" by the Iranian Academic Center for Education, Culture, and Research in the Engineering Group. Moreover, he was the Chair of the "5G and Beyond Communications" Session at the 2022 IEEE International Conference on Communications (ICC). He serves as an Editor for IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, *Journal of Engineering Applications of Artificial Intelligence* (Elsevier), *Journal of Physical Communication* (Elsevier), and *Journal of Computer Communication* (Elsevier). In addition, he serves as a reviewer for almost all top journals and is a member of several conferences as well. In the worldwide community, he is an active Delegate from South Korea to the Moving Picture Experts Group (MPEG).