

# **IEICE** **TRANSACTIONS**

## **on Fundamentals of Electronics, Communications and Computer Sciences**

DOI:10.1587/transfun.2024EAL2064

Publicized:2024/10/16

This advance publication article will be replaced by  
the finalized version after proofreading.



A PUBLICATION OF THE ENGINEERING SCIENCES SOCIETY

The Institute of Electronics, Information and Communication Engineers

Kikai-Shinko-Kaikan Bldg., 5-8, Shibakoen 3 chome, Minato-ku, TOKYO, 105-0011 JAPAN

## LETTER

# A novel Distributed Stagewise Orthogonal Matching Pursuit algorithm for mmWave MIMO channel estimation

Ngoc-Son DUONG<sup>†</sup>, Member, Lan-Nhi VU THI<sup>†</sup>, Nonmember, Sinh-Cong LAM<sup>†</sup>, Member, Phuong-Dung CHU THI<sup>†</sup>, and Thai-Mai DINH THI<sup>\*†</sup>, Nonmembers

**SUMMARY** In this letter, we investigate the problem of multiple-input multiple-output (MIMO) mmWave channel estimation in a hybrid analog-digital architecture by exploiting both sparsity and the structure of the channel. To gain noise robustness, we first introduce a method that applies the Stagewise Orthogonal Matching Pursuit (StOMP) algorithm to a distributed setting, where subsystems over sub-carriers share the same support set. To further enhance the accuracy of the estimation, we propose a novel algorithm to estimate the number of paths present in the channel. This technique leverages a modified Silhouette method to determine the exact support for the mmWave sparse system, thereby reducing the ambiguity of the estimate returned by the Distributed StOMP (DStOMP) algorithm. Simulation results demonstrate that our proposed method outperforms the standard OMP method and achieves nearly the same recovery accuracy compared to the Simultaneous OMP method, even without prior knowledge of signal sparsity.

**key words:** mmWave; channel estimation; distributed compressive sensing; OMP; SOMP; StOMP.

## 1. Introduction

The demand for high data rates and low-latency communication will continue to increase for next-generation networks due to the emergence of new types of services such as high-definition communication, immersive experiences, industrial automation, etc. As a result, millimeter-wave (mmWave) or even shorter wavelengths at terahertz frequencies were and will be adopted for use in multiple-input, multiple-output (MIMO) systems. However, raising the frequency also lowers the wave's capacity to propagate. Consequently, this leads to significant path loss at the user equipment (UE) and hence requires advanced methods like directional beamforming to compensate for its disadvantage. To support beamforming, sparse channel estimation algorithms were introduced [2]. For example, they are Orthogonal Matching Pursuit (OMP) [1], Sparse Bayesian Learning (SBL) [9], and Approximate Message Passing Algorithm (AMP) algorithm [6] [10].

Although the benefits of the sparse method have been proven in many previous works, structural understanding of mmWave channels and the benefits of certain methods for that structure still need more work to surpass the current state-of-the-art in mmWave channel estimation. There are two weaknesses that exist in current studies on sparse channel estimation: *i*) the number of paths, i.e., sparsity level, is often assumed to be known in advance, and *ii*) the mmWave

channel in the angular domain is assumed to be  $L$ -sparse, where  $L$  is the number of paths. These two assumptions reduce the practical meaning of sparse estimation methods. Firstly, estimating the number of paths in the channel is not a trivial task. Secondly, the true nature of the mmWave channel is only nearly sparse in the sense that there is a set of non-zero entries with smaller amplitudes concentrated around the dominant (highest) one [5].

In this letter, we propose a novel Distributed Stagewise Orthogonal Matching Pursuit (DStOMP) <sup>†</sup> algorithm for the estimation of the mmWave MIMO channel. Our approach leverages a distributed setting, estimating the channel with multiple sub-carriers rather than a single carrier. It is based by the fact that a multiple measurement vector setting invariably outperforms the single measurement vector setting in terms of noise robustness. The StOMP [1] algorithm was selected as the main algorithm for two practical reasons. First, it does not require prior knowledge of the sparse level. Second, it is particularly suited to the mmWave channel, which exhibits a specific structure. Leveraging the DStOMP and the channel's special structure, we introduce a novel clustering algorithm that is capable of identifying the number of paths and identifying the dominant supports in cases where the channel is considered  $L$ -sparse.

## 2. System model and mmWave channel estimation problem

### 2.1 System model

Consider a mmWave MIMO system, where a single UE is served by a single base station (BS) that employs a hybrid architecture. Both BS and UE are equipped with a uniform linear array (ULA) consisting of  $N_t$  and  $N_r$  equally spaced antennas, respectively. Initially, BS sent a pilot sequence to the UE  $G$  times where the  $g$ -th transmission over  $n$ -th sub-carrier includes known pilot  $\mathbf{x}^{(g)}[n] \in \mathbb{C}^{M_t}$ . The signal at time  $g$  that was sent over the  $n$ -th sub-carrier is expressed as  $\mathbf{F}^{(g)}[n]\mathbf{x}^{(g)}[n]$ , where  $\mathbf{F}^{(g)}[n] \in \mathbb{C}^{N_r \times M_t}$  is the beamforming matrix. We assume that there are  $L$  well-separated paths in the channel. The matrix channel on the  $n$ -th subcarrier can be written as

$$\mathbf{H}[n] = \mathbf{A}_R[n]\mathbf{\Gamma}[n]\mathbf{A}_T[n]^T, \quad (1)$$

<sup>†</sup>Code is available at <https://github.com/SonUET/DStOMP>

<sup>†</sup>The authors are with the Faculty of Electronics and Telecommunication, VNU University of Engineering and Technology, Hanoi 100000, Vietnam. \*Corresponding author.

where,  $\mathbf{\Gamma}[n]$ ,  $\mathbf{A}_T[n]$ , and  $\mathbf{A}_R[n]$  are the complex channel gain matrix, the steering matrix, and the response matrix over  $n$ -th subcarrier, respectively [8]. The received signal over  $n$ -th subcarrier can be written as

$$\mathbf{y}^{(g)}[n] = \mathbf{H}[n]\mathbf{F}^{(g)}[n]\mathbf{x}^{(g)}[n] + \mathbf{w}^{(g)}[n], \quad (2)$$

where,  $\mathbf{w}$  is the complex Gaussian noise with zeros mean and variance  $N_0/2$  per real dimension. Our purpose is to estimate the channel  $\mathbf{H}[n]$  when knowing received signal  $\mathbf{y}$ , beamforming matrix  $\mathbf{F}$  and pilot sequence  $\mathbf{x}$ .

## 2.2 Beamspace channel and sparse channel estimation

To reveal the channel sparsity, we multiply the  $\mathbf{H}$  matrix by two unitary transformation matrices, say  $\mathbf{U}_{R_x}$  and  $\mathbf{U}_{T_x}$ , as defined in [8]. Then, we call the quantity  $\mathbf{U}_{R_x}^H \mathbf{H}[n] \mathbf{U}_{T_x}$  the beamspace channel. Its essence is Dirichlet kernel [5] and can be illustrated by Fig. 1a. In addition to the main lobe, it also has many side lobes spanning the domains of AOA and AOD. Intuitively, the expansion of side lobes forms a cross line over the column and row space of the channel matrix. Assuming the paths do not affect each other, we call the position with the highest peak in the main lobe the main support.

For the  $g$ -th transmission, let  $\mathbf{\Omega}[n] = (\mathbf{U}_{T_x}^H \mathbf{F}[n] \mathbf{x}[n])^\top$  and  $\mathbf{h}[n] = \text{vec}(\mathbf{U}_{R_x}^H \mathbf{H}[n] \mathbf{U}_{T_x})$ , we obtain

$$\mathbf{y}[n] = \mathbf{\Omega}[n] \mathbf{h}[n] + \mathbf{w}[n]. \quad (3)$$

We refer to (3) as the distributed setting where sensing matrix  $\mathbf{\Omega}$  over all sub-carriers shares the same support set. Sparse algorithms such as OMP and its family are the best candidates to find sparse  $\mathbf{h}[n]$  in (3) when knowing  $\mathbf{y}[n]$  and  $\mathbf{\Omega}[n]$ .

## 3. Proposed Methods

### 3.1 Proposed Distributed Stagewise Orthogonal Matching Pursuit algorithm: The Procedure

Stacking the observations of all subcarriers to obtain  $\bar{\mathbf{y}}$ ,  $\bar{\mathbf{h}}$  and  $\bar{\mathbf{\Omega}}$ , we innate the algorithm with counter  $s = 0$ , solution  $\bar{\mathbf{h}}^{(0)} = \mathbf{0}$  and residual as  $\bar{\mathbf{r}}^{(0)} = \bar{\mathbf{y}}$ . It then calculates the correlation between each column in the sensing matrix and the corresponding residual. In the distributed case, we have

$$c_k^{(s)} = \sum_{n=1}^N |\langle \bar{\mathbf{r}}^{(s-1)}[n], \omega_i[n] \rangle| \quad (4)$$

where  $\omega_i[n]$  is the normalized column of  $\mathbf{\Omega}[n]$ . With  $G = 1$  and  $N = 1$ ,  $c_k$  is can be written as

$$\begin{aligned} c_k &= \left[ \left( \mathbf{u}_{T,i}^H \mathbf{F} \mathbf{x} \right)^\top \otimes \mathbf{u}_{R,j} \right]^H \left( \mathbf{U}_T^H \mathbf{F} \mathbf{x} \right)^\top \otimes \mathbf{U}_R \mathbf{r} \\ &= \left[ \left( \mathbf{u}_{T,i}^H \mathbf{F} \mathbf{x} \right)^\top \right]^H \left( \mathbf{U}_T^H \mathbf{F} \mathbf{x} \right)^\top \otimes \mathbf{u}_{R,j}^H \mathbf{U}_R \mathbf{r} \\ &= \mathbf{u}_{T,i}^\top \left[ (\mathbf{F} \mathbf{x})^\top \right]^H (\mathbf{F} \mathbf{x})^\top \left[ (\mathbf{U}_T)^H \right]^\top \otimes \mathbf{u}_{R,j}^H \mathbf{U}_R \mathbf{r} \end{aligned} \quad (5)$$

where  $\mathbf{u}_{T,i}$  and  $\mathbf{u}_{R,j}$  are  $i$ -th and  $j$ -th column of  $\mathbf{U}_{T,i}$  and  $\mathbf{U}_{R,j}$ , respectively and  $i, j \in [N_b]$ ,  $k = i + N_b(j - 1)$  assuming  $N_t = N_r$  and  $N_b$  is number of virtual angles. We next identifies all coordinates with amplitudes exceeding a hard threshold.

$$\mathcal{J}^{(s)} = \{k : c_k^{(s)} > t^{(s)} \sigma^{(s)}\}, \quad (6)$$

where  $\sigma^{(s)}$  is the noise level and  $t^{(s)}$  is the threshold parameter. Since the number of rows of the sensing matrix increases by  $N$  times compared to a single carrier case and  $c_i$  is accumulated by  $N$  sub-carriers, the noise level and the threshold should be scaled by  $N$  and become  $\frac{\|\bar{\mathbf{r}}^{(s)}\|_2}{N\sqrt{N_t G}}$  and  $N\delta^{(s)}$ , respectively, where  $\delta^{(s)}$  is some noise-dependent number. After that, we updates the estimates by merging the previously determined support with the newly subset as

$$\mathcal{I}^{(s)} = \mathcal{I}^{(s-1)} \cup \mathcal{J}^{(s)} \quad (7)$$

After this step, if  $|\mathcal{I}^{(s)}| = |\mathcal{J}^{(s)}|$ , we only need to complete the remaining steps once and terminate the procedure since the algorithm cannot find a new column. Let  $\mathbf{\Omega}_{\mathcal{I}}[n] \in \mathbb{C}^{N_t G \times |\mathcal{I}|}$  be the matrix that includes columns chosen from the support set  $\mathcal{I}$  over the  $n$ -th sub-carrier, we construct the matrix  $\bar{\mathbf{\Omega}}_{\mathcal{I}} \in \mathbb{C}^{N_t G N \times |\mathcal{I}| N}$  as

$$\bar{\mathbf{\Omega}}_{\mathcal{I}} = \text{blkdiag}(\mathbf{\Omega}_{\mathcal{I}}[1], \mathbf{\Omega}_{\mathcal{I}}[2], \dots, \mathbf{\Omega}_{\mathcal{I}}[N]) \quad (8)$$

to estimate the channel gains for all sub-carrier

$$\bar{\mathbf{h}}_{\mathcal{I}^{(s)}}^{(s)} = \left( \bar{\mathbf{\Omega}}_{\mathcal{I}^{(s)}}^H \bar{\mathbf{\Omega}}_{\mathcal{I}^{(s)}} \right)^{-1} \bar{\mathbf{\Omega}}_{\mathcal{I}^{(s)}}^H \bar{\mathbf{y}} \quad (9)$$

Next, we subtract the effect of them from the received signal as

$$\bar{\mathbf{r}}^{(s)} = \bar{\mathbf{y}} - \bar{\mathbf{\Omega}} \bar{\mathbf{h}}^{(s)} \quad (10)$$

The procedure is terminated when number of stage reach to  $S$ . Finally, we obtain  $\bar{\mathbf{h}}_{\mathcal{I}^{(s)}}$  as an estimate.

### 3.2 Proposed algorithm for estimating number of path and locating main supports

At the end of the above procedure, we obtain a set of supports whose number of elements is usually greater than the number of paths present in the channel. This is both a strength and a weakness of the StOMP algorithm compared to OMP. The essence of OMP family algorithms is to determine the support set based on the correlation between the columns of the sensing matrix and the received signal. They must ensure the following events occur with high probability [7].

$$E = \{ \max_{1 \leq i \leq m} |\omega_i^H[n] \mathbf{w}[n]| < \tau \} \quad (11)$$

where  $\tau := \sigma \sqrt{2(1 + \alpha) \log m}$  for some constant  $\alpha > 0$ ,  $m$  is the length of  $\omega_i$  and  $\sigma$  is the noise standard deviation. Under the influence of noise  $\sigma$ , OMP cannot guarantee selecting the correct support with only one column selected at each

---

**Algorithm 1:** Modified Silhouette method for the DStOMP algorithm
 

---

**Input:** Supports with coordinates:  $\mathbf{X}$   
**Initialization:** Number of clusters one want to test:  $K$ ,  
 Beamwidth of the main lobe (assuming  $G_t = G_r$ ):  $r$ .

```

for  $k = 2 \rightarrow K$  do
  1. Perform k-means clustering:  $[\mathbf{i}, \mathbf{C}] = \text{kmeans}(\mathbf{X}, k)$ ;
  for  $j = 1 \rightarrow k$  do
    2. Get the data points in the  $j$ -th cluster:
        $\mathbf{p} = \mathbf{X}(j == \mathbf{i}, :)$ ;
    3. Calculate the distance of each point in the cluster to
       the centroid:  $\mathbf{d} = \|\mathbf{p} - \mathbf{C}(j, :)\|_2^2$ ;
    if  $\max(\mathbf{d}) > r$  then
      4. Increase the number of clusters by 1:  $k = k + 1$ ;
      5. Perform k-means clustering again with the new
          $k$ :  $[\mathbf{i}, \mathbf{C}] = \text{kmeans}(\mathbf{X}, k)$ ;
      break;
    end
  end
  6. Calculate silhouette values:  $\mathbf{s} = \text{silhouette}(\mathbf{X}, \mathbf{i})$ ;
  7. Calculate average silhouette value:  $\mathbf{c}(k - 1) = \text{mean}(\mathbf{s})$ ;
end
8. Find the optimal number of clusters:  $[\cdot, \hat{k}] = \max(\mathbf{c})$ ;
9. Retain supports that satisfy (12) and (13);
Output: Estimated channel
  
```

---

step. In contrast, StOMP selects multiple columns in each step based on a hard threshold and ensures that those sets of columns contain the correct support with a higher probability compared to OMP. In this way, it, however, adds a vagueness to the support set identification.

To overcome this phenomenon, we introduce a modified Silhouette method [4] to estimate the number of path and retain those are belong to the main lobe. We add to conventional Silhouette method three constraints. The first constraint bound the radius of the main lobe in each angular direction. By this way, the K-means algorithm, which is often used in conjunction with the Silhouette method, will ignore cases where the distance from the cluster center to a data point is greater than the width of the main lobe. Using Silhouette method with the first constraint, we obtain the number of clusters and choose the largest value in each cluster as their main support. Let  $\mathbf{X} = \{\mathbf{X}_k\}$  be the estimated support set after allying the first condition with each support  $\mathbf{X}_k = [x_i, y_j, z_k]$ ,  $k = i + N_b(j - 1)$ , where  $x_i$ ,  $y_j$ , and  $z_k$  are the AOD, AOA and amplitude of that AOD-AOA pair, respectively. To exclude a support that does not belong to the main lobe, we look for supports that satisfy two following conditions

$$\mathbf{X}_k = \left\{ k \mid \max_{1 \leq i \leq N_b} z_k \text{ for each } j \right\} \quad (12)$$

$$\mathbf{X}_k = \{k \mid z_k > \eta\} \quad (13)$$

where  $\eta$  is small positive constant. The second constraint concerns the physical agreement of the estimates returned by the StOMP algorithm. If multiple supports have the same AOA, the algorithm will only retain the largest one. The third constraint concerns magnitude of support whose are

in arbitrary position. If a support have magnitude smaller than a small threshold, the algorithm will exclude it. The modified Silhouette method is given by Algorithm 1.

#### 4. Simulation and Discussion

We consider a wideband mmWave MIMO OFDM system, where  $N_t = N_r = 16$ . The downlink channel includes one LOS path and one NLOS path. It is modeled as [8] for simulation with 5000 realizations. In this work, our proposed method is compared with the standard OMP and Simultaneous OMP (SOMP)[3]. We employ successful recovery probability (SRP) and mean square error (MSE) as two indicators to assess the estimation performance as

$$\text{SRP} = \frac{\text{number of successful trial}}{\text{total number of trial}}, \quad (14)$$

$$\text{MSE} = \frac{\|\mathbf{H} - \hat{\mathbf{H}}\|_F^2}{\|\mathbf{H}\|_F^2}. \quad (15)$$

##### 4.1 On the visual simulation results

The visual simulation results of the proposed methods are given in Fig. 1. Fig 1b shows the channel reconstruction ability when using only the DStOMP algorithm at a high SNR level. In this case, undesirable supports degrade the overall performance of the DStOMP algorithm due to (9), especially in terms of channel amplitude. To preserve the main support's amplitude, LS should only be performed on a single column of the sensing matrix. When it comes to StOMP, the amplitude of the main support is shared with different side supports. As a result, the channel amplitude of the estimated highest peak is far different from the correct amplitude. Fig. 1c shows the estimated beamspace channel when using both the DStOMP algorithm and the modified Silhouette method. Thanks to the ability to determine the number of clusters using the modified Silhouette method and physical rules, undesirable supports are eliminated, and the final estimate retains only main supports.

##### 4.2 Performance of proposed method under different condition of $N$

This simulation evaluates the SRP and MSE of the three methods when changing the number of  $N$  while holding the number of  $G$ . The simulation results are given in Fig. 2. We observe that the two methods using multiple carriers have a higher SRP than the standard OMP which uses only a single carrier. It is worth noting that SOMP outperforms our proposed method in the low SNR range (-10 to 0 dB). However, it should be remembered that SOMP is a method that requires prior knowledge of signal sparsity, which is an informational advantage compared to the proposed method. It can be said that the gap between the two methods is created when Algorithm 1 is wrong in estimating the number of paths that exist in the channel. As shown in Fig. 2, we can write  $\lim_{N \rightarrow \infty} \text{SRP} = 1$  since the noise present in  $\mathbf{y}$  is white noise.

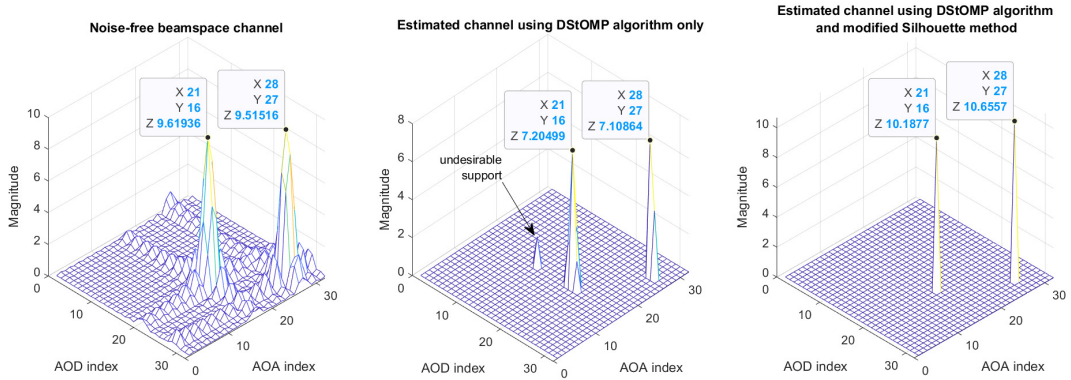


Fig. 1 Visualization of beamspace channel with different estimation schemes

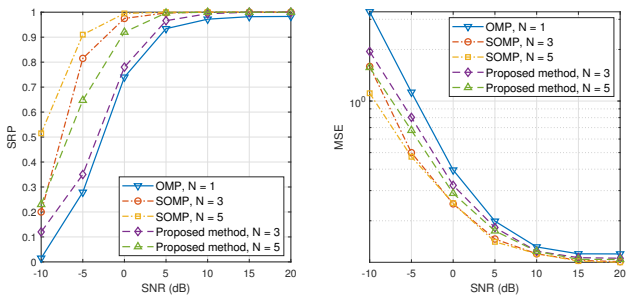


Fig. 2 SRP and MSE performance of all method with  $G = 5$

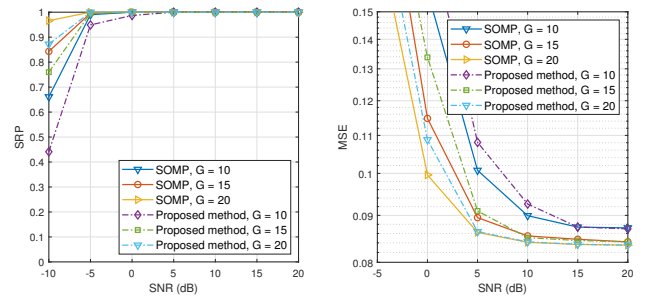


Fig. 3 SRP and MSE performance of all method with  $N = 5$

Increasing  $N$  only makes the power of noise in Eq. (5) smaller and smaller. At SNR ranges from -10 to 10, the MSEs of the methods show differences simply because their SRPs are different. In this case, lower MSEs correspond to higher SRP values. However, at SNR ranges from 15 to 20, the multi-carrier methods have near the same MSE. This is caused by the fact that an atom on one sub-carrier is essentially an alternate, rotated form of an atom with the same index on another sub-carrier. Then, increasing  $N$  does not decrease MSE because the projection of  $\mathbf{y}[n]$  onto subspace spanned by support set with different  $N$  is essentially the same.

#### 4.3 Performance of proposed method under different condition of $G$

In this simulation, we evaluate the performance of the proposed method when varying  $G$  while keeping  $N$  fixed. The results of the two indicators are shown in Fig. 3. Similar to the assessment in Sec. 4.2, increasing  $G$  also increases performance in terms of SRP. This occurs because  $G$  approaches infinity means transmit signals in all directions, i.e.,  $\lim_{G \rightarrow \infty} [(\mathbf{F}\mathbf{x})^T]^H (\mathbf{F}\mathbf{x})^T = \mathbf{I}$ . This makes coefficients  $c_k$  in (5) tend to form an identical version of the original channel but amplified by a constant. Finally, it increases the chance to select the exact main support. It is noteworthy

that as  $G$  increases, the performance in terms of MSE decreases. This happens because we estimate the channel coefficients through LS. With the same noise static, a system with a larger  $G$ , i.e., with more measurements, will give a LS estimate with lower error. As  $G$  increases to infinity, the MSE is expected to not decrease to infinity since the reconstructed channel is composed of quantized channel parameters. In the range of high SNR and with the same  $G$ , the MSE of the proposed method and that of SOMP are identical because they have the same channel coefficient estimation mechanism, i.e., projecting the received signal onto subspace spanned by the same support set.

## 5. Conclusion

In this study, we have presented a novel approach for MIMO mmWave channel estimation in a hybrid analog-digital architecture. In contrast to earlier work, our approach provides a practical solution with a distributed StOMP algorithm. To reduce the ambiguity of the DStOMP algorithm, we introduced a novel algorithm that estimates the number of paths present in the channel using a modified Silhouette method. Simulation results have shown that our proposed method, without prior knowledge of signal sparsity, can overcome the standard OMP method and achieve nearly the same performance as the SOMP method.

## References

- [1] D. L. Donoho, Y. Tsaig, I. Drori and J. L. Starck, "Sparse Solution of Underdetermined Systems of Linear Equations by Stagewise Orthogonal Matching Pursuit," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1094-1121, 2012.
- [2] T. Nishimura, Y. Ogawa, T. Ohgane, and J. Hagiwara, "Radio techniques incorporating sparse modeling," *IEICE Trans. Fundamentals*, vol. E104.A, no. 3, pp. 591-603, 2021.
- [3] J. A. Tropp, A. C. Gilbert and M. J. Strauss, "Simultaneous sparse approximation via greedy pursuit," *Proc. - ICASSP IEEE Int. Conf. Acoust. Speech Signal Process.*, Philadelphia, USA, pp. v/721-v/724 Vol. 5, 2005.
- [4] L. Kaufman, P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*, John Wiley & Sons, USA, 1990.
- [5] C. K. Anjinappa, Y. Zhou, Y. Yapici, D. Baron and I. Guvenc, "Channel Estimation in mmWave Hybrid MIMO System via Off-Grid Dirichlet Kernels," *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, USA, pp. 1-6, 2019.
- [6] D. L. Donoho, A. Maleki and A. Montanari, "Message passing algorithms for compressed sensing: I. motivation and construction," *Proc. IEEE Int. Theory Workshop (ITW)*, Cairo, Egypt, pp. 1-5, 2010.
- [7] Z. Ben-Haim, Y. C. Eldar and M. Elad, "Coherence-Based Performance Guarantees for Estimating a Sparse Vector Under Random Noise," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5030-5043, 2010.
- [8] A. Shahmansoori, G. Garcia, G. Destino, G. Seco-Granados and H. Wymeersch, "Position and Orientation Estimation Through Millimeter-Wave MIMO in 5G Systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1822-1835, 2018.
- [9] N. S. Duong, Q. T. Nguyen, K. H. Ngo and T. M. Dinh-Thi, "Sparse Bayesian Learning with Atom Refinement for mmWave MIMO Channel Estimation," *Proc. IEEE Stat. Signal Process. Workshop (SSP)*, Hanoi, Vietnam, pp. 155-159, 2023.
- [10] C. Ruan, Z. Zhang, H. Jiang, H. Zhang, J. Dang and L. Wu, "Simplified Learned Approximate Message Passing Network for Beamspace Channel Estimation in mmWave Massive MIMO Systems," *IEEE Trans. Wireless Commun.*, vol. 23, no. 5, pp. 5142-5156, 2024.