SCIENCE CHINA
Information Sciences

• **LETTER** •

Beam Alignment for Millimeter Wave Multiuser MIMO Systems Using Sparse-Graph Codes

Long CHENG[1], Guangrong YUE[1*], Pei XIAO[2], Ning WEI[3,4] & Shaoqian Li[1]

[1]*National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu* 611731*, China;*
[2]*5GIC &6GIC, Institute for Communication Systems (ICS), University of Surrey, Guildford GU2 7XH, U.K.;*
[3] *ZTE Corporation, Shenzhen 518057, China;*
[4] *State Key Laboratory of Mobile Network and Mobile Multimedia Technology, Shenzhen 518055, China*

Citation Cheng L, Yue G R, XIAO Pei, et al.
. Sci China Inf Sci, for review

Dear editor,

In order to achieve millimeter wave (mmWave) beam alignment, a class of beam scanning and searching schemes have been extensively studied [1–3]. Recently, to address the problems of the traditional algorithms have a high sample complexity, some adaptive beam scanning approaches utilize the hierarchical beamforming codebook to reduce the training time at the cost of frequent feedback [2]. Then, to eliminate the feedback link, a random beam alignment algorithm is proposed by utilizing the pseudo-random spreading codes [3]. However, it needs a Pseudo-Noise (PN) sequences with sufficient length to ensure the good correlation properties of different beams. Furthermore, in addition to the above disadvantages, most of the existing algorithms require either a separate pilot sequence per user or long beam scanning time when considering mmWave multi-user uplinking systems.

To solve the above problems, a novel class of beam alignment algorithms based on the sparse graph coding theory are proposed in this paper. Firstly, we investigate the uplink mmWave beam training structure. Based on the analysis, the mmWave multi-user beam alignment problem is transformed into the sparse-graph design and detection problem. Secondly, a beam alignment algorithm framework based on sparse-graph coding and decoding is proposed. Furthermore, we derive the theoretical bound to chose the optimal parameters of the designed coding matrix. Finally, two beam alignment algorithms are proposed to detect the beam index in different settings. Simulation results confirm that our beam algorithms outperform the conventional beam training methods.

Proposed Uplink Beam Training Scheme. This paper considers a typical uplink mmWave MU-MIMO system, where the BS communicates with K UEs simultaneously. Suppose that the BS is equipped with $N_{\rm R}$ antennas and $N_{\rm RF}$ RF chains, while the k-th UE has $M_{\rm T}$ antennas and $M_{\rm RF}$ RF chains. Then, the channel associated with the k-th UE can be given by [4]

$$\mathbf{H}_k=\sum_{l=1}^{L_k}\alpha_{l,k}\mathbf{a}_{\rm BS}(\theta_{l,k})\mathbf{a}_{\rm UE}^H(\varphi_{l,k}), \tag{1}$$

where $(\theta_{l,k},\varphi_{l,k},\alpha_{l,k})$ denote the AoA, AoD, complex gain of the l-th path of the k-th UE, respectively. The multipath complex gain can be modeled as Rice fading given by [5, 6]

$$\alpha_{l,k}\sim\sqrt{\rho_{l,k}}\left(\sqrt{\frac{\eta_{l,k}}{\eta_{l,k}+1}}+\sqrt{\frac{1}{\eta_{l,k}+1}}\widetilde{\alpha}_{l,k}\right), \tag{2}$$

where $\sqrt{\rho_{l,k}}$ is the overall multipath complex gain strength, $\eta_{l,k}$ denotes the ratio between the LOS component and the NLOS components, and $\widetilde{\alpha}_{l,k}\sim\mathcal{CN}(0,1)$ denotes the complex Gaussian random variable. Furthermore, the uniform linear array (ULA) is used by the BS and the UEs.

In our scheme, the BS can locally customize its own beamforming codebook, then combining matrix $\mathbf{w}(t)$=$\mathbf{W}_{\rm RF}(t)\mathbf{W}_{\rm BB}(t)$=$\mathbf{F}_{\rm BS}\mathbf{v}(t)$, where $\mathbf{W}_{\rm RF}(t)\in\mathbb{C}^{N_{\rm R}\times N_{\rm RF}}$, $\mathbf{W}_{{\rm BB},k}(t)\in\mathbb{C}^{N_{\rm RF}\times N_{\rm RF}}$, $\mathbf{F}_{\rm BS}\in\mathbb{C}^{N_{\rm R}\times N_{\rm R}}$, and $\mathbf{v}(t)\in\mathbb{C}^{N_{\rm R}\times 1}$are the RF precoder matrix, the digital precoder matrix, the Discrete Fourier Transform (DFT) matrix, and the index vector of the quantized angle, respectively. Specially, the non-zero value in the beam selection vector $\mathbf{v}(t)$ indicates which direction the candidate beam should be formed. Without loss of generality, we set the non-zero value as one. On the UE side, the beamforming matrix $\mathbf{f}_k(t)$=$\mathbf{F}_{{\rm RF},k}(t)\mathbf{F}_{{\rm BB},k}(t)$=$\mathbf{F}_{\rm MS}\boldsymbol{\psi}_k(t)$, where $\mathbf{F}_{{\rm RF},k}(t)\in\mathbb{C}^{M_{\rm T}\times M_{\rm RF}}$, $\mathbf{F}_{{\rm BB},k}(t)\in\mathbb{C}^{M_{\rm RF}\times M_{\rm RF}}$, $\mathbf{F}_{\rm MS}\in\mathbb{C}^{M_{\rm T}\times M_{\rm T}}$, and $\boldsymbol{\psi}_k(t)\in\mathbb{C}^{M_{\rm T}\times 1}$ denote the RF precoder matrix , the digital precoder, the DFT matrix, and the coding vector, respectively. Note that the mmWave channel can be given by a beamspace representation expression.

* Corresponding author (email: yuegr@uestc.edu.cn)

Thus, the received signal can be rewritten as

$$\begin{aligned}\mathbf{r}(t) &= \mathbf{w}^H(t)\sum_{k=1}^{K}\mathbf{H}_k\mathbf{f}_k(t)+\mathbf{N}(t)\\ &= \mathbf{v}^H(t)\mathbf{F}_{\mathrm{BS}}^H\sum_{k=1}^{K}\mathbf{F}_{\mathrm{BS}}\widehat{\mathbf{H}}_k\mathbf{F}_{\mathrm{MS}}^H\mathbf{F}_{\mathrm{MS}}\boldsymbol{\psi}_k(t)+\mathbf{N}(t),\\ &= \mathbf{v}^H(t)\sum_{k=1}^{K}\widehat{\mathbf{H}}_k\boldsymbol{\psi}_k(t)+\mathbf{N}(t)\end{aligned} \tag{3}$$

where $\widehat{\mathbf{H}}_k$ is the virtual angle domain index. Furthermore, considering that the BS can separate the received signal and the non-zero value in the selection vector $\mathbf{v}(t)$ is one, the received signal of the i-th RF chain can be written as

$$r_i(t)=\boldsymbol{\psi}^T(t)\widehat{\mathbf{h}}+n(t) \tag{4}$$

where $\widehat{\mathbf{h}}=[\widehat{\mathbf{h}}_1^T,\widehat{\mathbf{h}}_2^T,\cdots\widehat{\mathbf{h}}_K^T]^T\in\mathbb{C}^{KM_T\times 1}$, $\widehat{\mathbf{h}}_k$ denotes the selected row of $\widehat{\mathbf{H}}_k$, and then $\boldsymbol{\psi}^T(t) = [\boldsymbol{\psi}_1^T(t),\boldsymbol{\psi}_2^T(t)\cdots\boldsymbol{\psi}_K^T(t)\in\mathbb{C}^{1\times KM_T}$.

Furthermore, by collecting T received pilot signals at the BS, we obtain

$$\mathbf{r}_i=\boldsymbol{\psi}\widehat{\mathbf{h}}+\mathbf{N}, \tag{5}$$

where $\boldsymbol{\psi}=[\boldsymbol{\psi}^T(1),\boldsymbol{\psi}^T(2),\cdots\boldsymbol{\psi}^T(T)]\in\mathbb{C}^{T\times KM_T}$. Note that the nonzero elements in vector $\boldsymbol{\psi}_k^T(t)$ denote the selected beam indexes, therefore $\boldsymbol{\psi}$ is the measurement matrix. In particular, since we process the signals of each receiving link separately, we denote $\mathbf{r}_i=\mathbf{r}$ for simplicity whenever no ambiguity arises.

Proposed Beam Alignment Scheme Using Sparse-Graph Codes. For the purposes of later analysis, we first define the $T = NM$. Then, we divide the measurement matrix $\boldsymbol{\psi}\in\mathbb{C}^{NM\times KM_T}$ into two parts, sparse coding matrix $\mathbf{G}\in\mathbb{C}^{M\times M_T}$ and bin detection matrix $\mathbf{S}\in\mathbb{C}^{N\times M_T}$. Thus, the measurement matrix $\boldsymbol{\psi}$ is given by

$$\boldsymbol{\psi}=\mathbf{G}\odot\mathbf{S}, \tag{6}$$

where $\odot$ denotes the *row-tensor* operator. Mathematically, we rewrite the measurement matrix $\boldsymbol{\psi}$ as follows

$$\boldsymbol{\psi}=[\mathbf{G}_1\otimes\mathbf{S}_1\cdots\mathbf{G}_{M_{\mathrm{T}}}\otimes\mathbf{S}_{M_{\mathrm{T}}}], \tag{7}$$

where $\mathbf{G}_i$ and $\mathbf{S}_i$ represent the i column of the matrix $\mathbf{G}$ and the matrix $\mathbf{S}$, respectively, and $\otimes$ denotes the Kronecker product.

Inspired by the traditional LDPC coding method [7], we construct a regular bipartite graph $\Gamma^{M_{\mathrm{T}}}(R,b)$ with M_{T} left nodes and R right nodes for the sparse coding matrix, where each left node is connected to b right nodes at random [8]. Here we devide the observation detections into three categories: **zero-ton**, **single-ton**, and **multi-ton**. Specially, the detailed descriptions can be found in Appendix A.

Furthermore, the function of the matrix S is to effectively distinguish the types of right nodes. Then, if the receiver gets the oracle information about the types of the right nodes, similar to the message passing algorithm [7], a peeling decoder can be applied to peel off all the **single-ton** in the bipartite graph. Furthermore, by iteratively repeating the process of peeling off, all the edges can be removed from the graph. Finally, we carry out a probability analysis of the proposed peeling-decoder, over a random selection from the regular graph ensemble $\Re^K(F,m)$. In this ensemble, the m detections is divided into d stages, with each left node connected to one right node per stage randomly. The set F is defined as $F=\{f_1\cdots f_d\}$, where the number of right nodes in stage i is f_i. In particular, $f_i=\mu K+O(1)$, for all i and some redundancy parameter μ. Furthermore, by using the density evolution analysis in [8], our proposed peeling decoder can recover K sparse $\widehat{\mathbf{h}}$ with a probability given as follows:

Theorem 1. If the proposed peeling decoder over a random graph from the ensemble $\Re^K(F,m)$, satisfies the stages of $d\geqslant 3$ and $f_i=\mu K+O(1)$ with the constant μ being chosen from Table B1 (Appendix B), then it can successfully recover K non-zero entries with probability $1-O(1/m)$.

The proof is given in Appendix C.

Thus, based on the same sparse coding matrix $\mathbf{G}$, we proposed two beam alignment algorithms by designing different detection matrices $\mathbf{S}$. One is for the noiseless case in Appendix D and another is for the noisy environment in Appendix E.

Experiment. The experimental results of our proposed algorithms can be seen in Appendix F.

Conclusion. This paper proposed several beam alignment algorithm designs for mmWave multiuser MIMO systems using sparse-graph codes. By utilizing the *peeling off* method and the density evolution analysis method, the proposed algorithm can approach the theoretical bound under the noiseless setting and obtain the reduction of pilot overhead under the noise setting.

Acknowledgements This work was partly supported by the National Key Research and Development Program of China (No. 2020YFB1805001), in part by the National Natural Science Foundation of China (No. 61831004), in part by ZTE Corporation, and in part by the China Scholarship Council. The work of Pei Xiao was supported by the U.K. Engineering and Physical Sciences Research Council under Grant EP/P03456X/1.

Supporting information Appendix A-F. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Liu C, Li M, Hanly S V, et al. Millimeter-wave small cells: base station discovery, beam alignment, and system design challenges. IEEE Wireless Communications, 2018, 25: 40-46

2 Noh S, Zoltowski M D, Love D J. Multi-resolution codebook and adaptive beamforming sequence design for millimeter wave beam alignment. IEEE Transactions on Wireless Communications, 2017, 16: 5689-5701

3 Chiu S, Ronquillo N, Javidi T. Active learning and CSI acquisition for mmWave initial alignment. IEEE Journal on Selected Areas in Communications, 2019, 37: 2474-2489

4 Alkhateeb A, Ayach O E, Leus Gand, et al. Channel estimation and hybrid precoding for millimeter wave cellular systems, IEEE Journal of Selected Topics in Signal Processing, 2014, 8: 831-846

5 Song X, Haghighatshoar S, Caire G. Efficient beam alignment for millimeter wave single-carrier systems with hybrid MIMO transceivers. IEEE Transactions on Wireless Communications, 2019, 18: 1518-1533

6 P. Lioliou, M. Viberg and M. Matthaiou. Bayesian approach to channel estimation for AF MIMO relaying systems. IEEE Journal on Selected Areas in Communications, 2012, 30: 1440-1451

7 Richardson T J, Urbanke R L. The capacity of low-density parity-check codes under message-passing decoding. IEEE Transactions on Information Theory, 2001, 47: 599-618

8 Luby M G, Mitzenmacher M, Shokrollahi M A, et al. Efficient erasure correcting codes. IEEE Transactions on Information Theory, 2001, 47: 569-58

SCIENCE CHINA
Information Sciences

• **Supplementary File** •

Beam Alignment for Millimeter Wave Multiuser MIMO Systems Using Sparse-Graph Codes

Long CHENG[1], Guangrong YUE[1*], Pei XIAO[2], Ning WEI[3,4] & Shaoqian Li[1]

[1]*National Key Laboratory of Science and Technology on Communications,*
University of Electronic Science and Technology of China, Chengdu 611731*, China;*
[2]*5GIC &6GIC, Institute for Communication Systems (ICS),*
University of Surrey, Guildford GU2 7XH, U.K.;
[3] *ZTE Corporation, Shenzhen 518057, China;*
[4] *State Key Laboratory of Mobile Network and Mobile Multimedia Technology, Shenzhen 518055, China*

Appendix A Three Classifications of Observation Nodes

According to the distribution of multipath direction of mmWave channel, we can classify the observation detections based on its edge degree as follows

1. **zero-ton:** A right node is a zero-ton if it is not connected to any non-zero entry of the angle domain channel $\widehat{\mathbf{h}}_k$.
2. **single-ton:** A right node is a single-ton if it is connected to only one non-zero entry of the angle domain channel $\widehat{\mathbf{h}}_k$.
3. **multi-ton:** A right node is a multi-ton if it is connected to more than one non-zero entry of the angle domain channel $\widehat{\mathbf{h}}_k$.

Appendix B The Minimum Threshold Value of η of the Number of Stage d

Table B1 The Minimum Threshold Value of η of the Number of Stage d

d	3	4	5	6	7	8
μ	0.407	0.323	0.285	0.261	0.245	0.233
$d\mu$	1.221	1.292	1.425	1.566	1.715	1.864

Appendix C Proof of Theorem 1

Define Y as the total number of edges that are not decoded over a random graph from the ensemble $\Re^K(F, m)$, we have

$$\begin{aligned} &\mathbb{E}(Y) < 2Kdp_i, \\ &P(|Y - \mathbb{E}(Y)| > Kd\varsigma) < e^{-\delta\varsigma^2 K^{1/(4i+1)}}, \end{aligned} \tag{C1}$$

where δ, ς denote the variable parameters, ς is the arbitrary parameter, and p_i is the probability of the event that an edge exists after the i-th peeling-off operation.

Furthermore, to derive the expression of p_i , we define the edge degree distribution in the ensemble as

$$\begin{aligned} \lambda(\alpha) = \textstyle\sum_{i=1}^{\infty} \lambda_i \alpha^{i-1}, \\ \rho(\alpha) = \textstyle\sum_{i=1}^{\infty} \rho_i \alpha^{i-1}, \end{aligned} \tag{C2}$$

where λ_i and ρ_i are the probability that the edge connected to the left (resp. right) node with degree i. For the proposed *beam-and-detection* procedure, $\lambda(\alpha) = \alpha^{d-1}$. Furthermore, considering that the degree of a right node follows the binomial distribution $B(1/\eta K, K)$, we have

$$\begin{aligned} \rho_i &= i\eta P(a \text{ right node has edge degree i}), \\ &\approx \frac{(1/\eta)^{i-1} e^{-1/\eta}}{(i-1)!}. \end{aligned} \tag{C3}$$

Thus, the edge degree distribution polynomial $\rho(\alpha)$ is given by

$$\rho(\alpha) = e^{(-(1-\alpha)/\eta)}, \tag{C4}$$

* Corresponding author (email: yuegr@uestc.edu.cn)

Then, under the cycle-free assumption, the probability p_i in (C1) for the ensemble $\Re^K(F,m)$ can be expressed as

$$p_{i+1} = (1 - e^{-\frac{p_i}{\eta}})^{d-1}, \tag{C5}$$

where p_1=1, and the choice of η is given to guarantee $p_{i+1} < p_i$ e.g., see Table B1.

According to Eq. (C1), we can draw the conclusion that with high probability, the proposed peeling-off decoder captures all but an small fraction of the variable nodes. However, in a real application scenario, we need to ensure all the variable nodes to be recovered. Therefore, based on the above conclusions, we next study how to complete all the decoding work.

Without loss of generality, in this appendix, we consider a set of left non-zero nodes P in the random graph $\Re^K(F,m)$, and the corresponding right neighborhood $N_i(P)$ of the i-th subset of the right nodes. Note that the proposed peeling-off decoder fails when there are no single-ton nodes in the $N_i(P)$ for $i = 1, \cdots d$. Furthermore, a sufficient condition for the above hypothesis is that the average degree of all the nodes in the neighborhood of P (i.e. $|N_i(P)|$ for all the i) is less than 2, such as $|N_i(P)| > |P|/2$. Specifically, if the average degree of the right nodes is less than 2, the single-ton must occur. Then, considering that $f_i = \eta K + O(1)$, we will discuss the probability of the opposite event $\Im$, i.e., $\max\{|N_i(P)|\}_{i=1}^d \leqslant |P|/2$ as follows

$$\begin{aligned}
\Pr(\aleph) &< \prod_{i=1}^{d} \left(\frac{|P|}{2f_i}\right)^{|P|} \binom{f_i}{|P|/2} \\
&\approx \left(\frac{|P|}{2\eta K}\right) \binom{\eta K}{|P|/2}^{d} \\
&< \left(\frac{|P|}{2\eta K}\right) \left(\frac{2\eta K e}{|P|}\right)^{d|P|/2} \\
&= \left(\frac{|P|\, e}{2\eta K}\right)^{d|P|/2}.
\end{aligned} \tag{C6}$$

Furthermore, by utilizing a union bound, the probability of event $\aleph_s$ that there exist some sets of the variable nodes that follow the rules, i.e., $\max\{|N_i(P)|\}_{i=1}^d \leqslant |P|/2$, we have

$$\begin{aligned}
\Pr(\aleph_s) &< \Pr(\aleph) \binom{K}{|P|} \\
&< \left(\frac{|P|\, e}{2\eta K}\right)^{d|P|/2} \left(\frac{Ke}{|P|}\right)^{|P|} \\
&< \left[\left(\frac{|P|}{\eta K}\right)^{d-2} \left(\frac{e}{2}\right)^{d} \left(\frac{e}{\eta}\right)^{2}\right]^{|P|/2} \\
&\overset{(a)}{<} O\left((|P|/m)^{|P|/2}\right)
\end{aligned} \tag{C7}$$

where (a) comes from the fact that $d \geqslant 3$ and $m = O(\eta K)$. Then, according to $|P| = O(K)$, we get

$$\Pr(\aleph_s) < O(1/m), \tag{C8}$$

Finally, based on Eq. (C1), Eq. (C5) and Eq. (C8), **Theorem 1** can be proved.

Appendix D Proposed Noiseless Beam Alignment Algorithm

In this appendix, a noiseless beam alignment algorithm is proposed for mmWave MU-MIMO systems. Considering that in the noiseless case, the mmWave uplink multi-user channel only retains K LOS paths, then we design the bin detection matrix $\mathbf{S}$ as

$$\mathbf{S} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & e^{j\frac{2\pi}{KM_T}} & \cdots & e^{j\frac{2\pi}{KM_T}(KM_T-1)} \end{bmatrix}. \tag{D1}$$

Therefore, the b-th left node is modulated by the b column of the matrix $\mathbf{S}$. Furthermore, each right node is connected to a bin measurement vector $\mathbf{r}^b = [r^b[1], r^b[2]]^T$. Then, in order to effectively identify whether each measured right node is a zero-ton, single-ton or multi-ton, we design the identification method as follows

1. **zero-ton identification:** A measured right node is a zero-ton if all the measurements are zero, such as

$$r^i[1]{=}r^i[2]{=}0. \tag{D2}$$

2. **single-ton identification:** The single-ton identification is utilized to obtain the index of the estimated beam and the magnitude as follows

$$\begin{aligned}
&\widetilde{b} = \tfrac{\angle r^i[2]/r^i[1]}{2\pi/KM_T}, \\
&\widetilde{\mathbf{h}}[\widetilde{b}] = r^i[1],
\end{aligned} \tag{D3}$$

where the measured node can be judged as a single-ton only if the estimated index $\widetilde{b}$ is an integer.

3. **multi-ton identification:** A measured right node is a multi-ton if the measured vector complies with the following rules

$$\begin{aligned}
&\widetilde{i} = \tfrac{\angle r^i[2]/r^i[1]}{2\pi/KM_T} \neq \text{integer}, \\
&\left|r^i[2]\right| \neq \left|r^i[1]\right|.
\end{aligned} \tag{D4}$$

Finally, combined with the *peeling off* method, the proposed noiseless beam alignment algorithm is proposed. And then, the detail steps are listed in the Algorithm D1.

Algorithm D1 Noiseless Beam Alignment Algorithm Using Sparse-Graph Codes

Require: Received signal $\mathbf{r}$, measurement matrix $\boldsymbol{\psi}$, sparse coding matrix $\mathbf{G}$ and bin detection matrix $\mathbf{S}$, number of users K and number of iterations L.
Ensure: The estimated index of beam channel $\widetilde{b}$ and magnitude $\widetilde{\mathbf{h}}$.
  for iteration $1 < l < L$ **do**
    for stage $1 < i < d$ **do**
      for bin $1 < j < f_i$ **do**
        if $\|\mathbf{r}_{i,j}[1]\| = 0$ **then**
          $\mathbf{r}_{i,j}$ is a zero-ton bin vector.
        else
          identify if $\mathbf{r}_{i,j}$ is a single-ton bin vector and get the estimated index-value pair $\left\langle \widetilde{b}, \widetilde{\mathbf{h}}[\widetilde{b}] \right\rangle$.
          if single-ton = 'true' **then**
            peeling off: $\mathbf{r}^{l+1} = \mathbf{r}^{l} - \widetilde{\mathbf{h}}[\widetilde{b}] \begin{pmatrix} 1 \\ e^{j\frac{2\pi\widetilde{b}}{KM_T}} \end{pmatrix}$
          else
            bin vector $\mathbf{r}_{i,j}$ is a **multi-ton** bin vector.
          end if
        end if
      end for
    end for
  end for

$$\underbrace{\left[\begin{array}{cccccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ \hdashline 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ \hdashline 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \end{array}\right]}_{\mathbf{G}} \Rightarrow \underbrace{\left[\begin{array}{cccccc} \mathbf{D}_1[1] & 0 & 0 & \mathbf{D}_1[2] & 0 & 0 \\ 0 & \mathbf{D}_1[1] & 0 & 0 & \mathbf{D}_1[2] & 0 \\ 0 & 0 & \mathbf{D}_1[1] & 0 & 0 & \mathbf{D}_1[2] \\ \hdashline \mathbf{D}_2[1] & 0 & \mathbf{D}_2[2] & 0 & \mathbf{D}_2[3] & 0 \\ 0 & \mathbf{D}_2[1] & 0 & \mathbf{D}_2[2] & 0 & \mathbf{D}_2[3] \\ \hdashline 0 & \mathbf{D}_3[1] & 0 & \mathbf{D}_3[2] & 0 & \mathbf{D}_3[3] \\ \mathbf{D}_3[1] & 0 & \mathbf{D}_3[2] & 0 & \mathbf{D}_3[3] & 0 \end{array}\right]}_{\boldsymbol{\Psi}}$$

Figure E1 An illustration of the sparse codes with its associated measurement matrix, while the coding matrix has three stages with $f_1 = 3$, $f_2 = 2$ and $f_3 = 2$. Note that $\mathbf{D}_i[x]$ denotes the x-th column of the bin matrix $\mathbf{D}_i$.

Appendix E Proposed Robust Beam Alignment Algorithm Using Sparse-Graph Codes

In this appendix, a robust beam alignment algorithm is proposed to enhance robustness in noisy environments. In a real multi-user scenario, we only need to acquire the LOS path for each user [1]. Then we can utilize the same regular graph $\Gamma^{M_T}(R,b)$ as mentioned in previous section to ensure the success probability of beam detection, because it provides a theoretical bound for beam scanning. However, in the noisy scenario, the data received is contaminated and cannot be simply indexed by calculation, we thus propose a new bin detection matrix $\mathbf{D}$ in the sequel.

Since only a few beams in specific directions are sent each time for beam scanning, we can also design the bin detection matrix $\mathbf{D}$ to be sparse, that is, the beams in different directions have different weights each time. Specifically, in each stage of the sparse coding matrix $\mathbf{G}$, we design the corresponding special bin matrix $\mathbf{D}_i \in \mathbb{C}^{N\times M_i}$, where M_i is the number of non-zero entries in each stage of the matrix $\mathbf{G}$. In particular, considering the presence of noise, each entry of the matrix $\mathbf{D}_i$ is randomly selected from a unit circle [2]. Furthermore, an example is shown in Fig. E1 to illustrate the relationship between measurement matrix $\boldsymbol{\psi}$ and sparse coding matrix $\mathbf{G}$.

Based on the special structure of measurement matrix, we propose a robust detection scheme for each bin. Firstly, we simple check if the received vector is zero-ton as follows

$$\|\mathbf{r}_{i,j}[1]\| \leqslant (\delta^2_{\min}\varepsilon_1 + \sigma^2_{\min}), \tag{E1}$$

where $\delta^2_{\min}$ denotes the minimum signal power, $\sigma^2_{\min}$ is the minimum noise power, and ε_1 is the **zero-ton** detection threshold. In particular, it seems that the values of the minimum signal power $\delta^2_{\min}$ and the minimum noise power $\sigma^2_{\min}$ influence the final decision. However, since the minimum signal power and minimum noise power have been determined as a standard when sending signals, we use tuning factor ε_1 to adjust them.

we assume the received bin vector is a single-ton and then estimate the index pair $\left\langle \widetilde{b}, \widetilde{\mathbf{h}}[\widetilde{b}] \right\rangle$. Specifically, for bin j of stage i, a maximum likelihood (ML) method is utilized to obtain the possible coefficient for the k-th column of the bin matrix $\mathbf{D}_i$ as follows

$$\alpha_k = \frac{\mathbf{D}_i^H[k]\mathbf{r}_{i,j}}{\left\|\mathbf{D}_i[k]\right\|^2}. \tag{E2}$$

Then, the estimated $\widetilde{k}$ is the one which minimizes the residual, i.e.,

$$\widetilde{k} = \arg\min_{1<k<M_i} \|\mathbf{r}_{i,j} - \alpha_k \mathbf{D}_i[k]\|. \tag{E3}$$

Algorithm E1 Robust Beam Alignment Algorithm Using Sparse-Graph Codes

Require: Received signal $\mathbf{r}$, measurement matrix $\boldsymbol{\psi}$, sparse coding matrix $\mathbf{G}$ and bin detection matrix $\mathbf{D}_i$ for all stages, number of users K ,number of iterations L, error threshold ε_1, initialization parameter $K_{\rm ue} = 0$ and error threshold ε_2
Ensure: The estimated index of beam channel $\widetilde{b}$ and magnitude $\widetilde{\mathbf{h}}$.
 for iteration $1 < l < L$ **do**
  for stage $1 < i < d$ **do**
   for bin $1 < j < f_i$ **do**
    if Eq. (E1) is real **then**
     $\mathbf{r}_{i,j}$ is a zero-ton bin vector.
    else
     for index $1 < k < M_i$ **do**
      if $K_{ue} \leqslant K$ **then**
       get the estimated index-value pair $\left\langle \widetilde{k}, \widetilde{\alpha}_{\widetilde{k}} \right\rangle$ by using Eq. (E2), Eq. (E3).
       identify if the received bin vector is a single-ton by using Eq. (E4).
       if single-ton = 'true' **then**
        obtain $\widetilde{b}$ by utilizing the estimated $\widetilde{k}$ and the coding pattern in the j-th bin of the i-th stage.
        $\widetilde{\mathbf{h}}[\widetilde{b}] = \widetilde{\alpha}_{\widetilde{k}}$.
        peeling off: $\mathbf{r}^{l+1} = \mathbf{r}^{l} - \widetilde{\mathbf{h}}[\widetilde{b}]\mathbf{D}_i[\widetilde{k}]$.
        $K_{ue} = K_{ue} + 1$;
       else
        bin vector $\mathbf{r}_{i,j}$ is a multi-ton bin vector.
       end if
      else
       break.
      end if
     end for
    end if
   end for
  end for
 end for

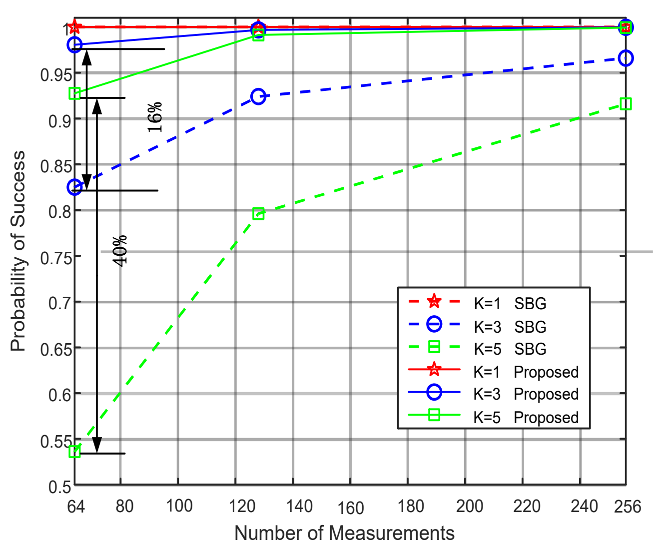

Figure F1 Comparison of the probability of success among the proposed method and the fast beam alignment method versus different measurement cost.

Furthermore, a *check* procedure is proposed to determine if it is a single-ton, according to the following criterion

$$\frac{1}{M_i} \left\| \mathbf{r}_{i,j} - \widetilde{\alpha}_{\widetilde{k}} \mathbf{D}_i[\widetilde{k}] \right\| \leqslant (\delta^2_{\min}\varepsilon_2 + \sigma^2_{\min}), \tag{E4}$$

where $\delta^2_{\min}$ denotes the minimum single power, $\sigma^2_{\min}$ is the minimum noise power, and similar to the zero-ton detection threshold ε_1 in Eq. (E1), ε_2 is the single-ton detection threshold.

Finally, the estimated $\widetilde{k}$ can be mapped to the real index $\widetilde{b}$ and the detailed steps are listed in the Algorithm E1.

Appendix F Simulation Results

In this section, we consider a mmWave Massive MIMO system with hybrid precoding architecture at both the UEs and the BS. In particular, the BS utilizes a ULA with $N_R = 64$ and $N_{\rm RF} = 16$ chains to support $K \leqslant 16$ UEs and the UEs utilizes a ULA with $M_{\rm T} = 16$ and $M_{\rm RF} = 4$ chains. For the geometric mmWave channel with noise, as mentioned above, the path loss of the NLOS paths is always higher than the LOS path, thus we consider the channel of each user is a Rician fading channel, e.g.,

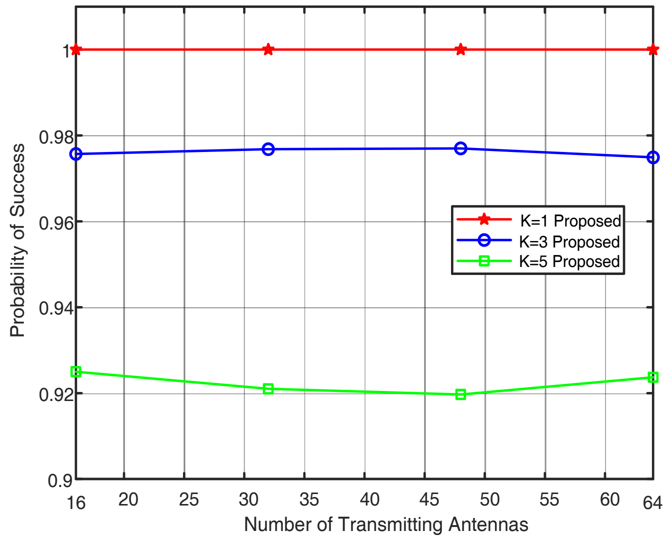

Figure F2 Probability of success of the proposed method versus different number of transmitting antennas, where the number of measurements is 64.

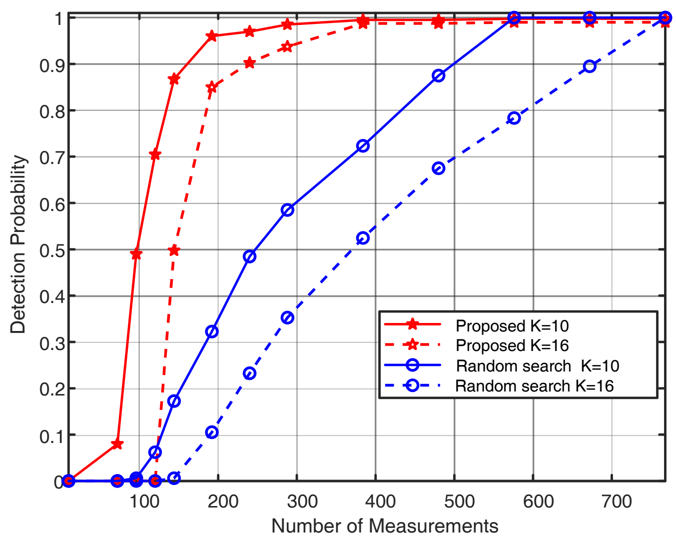

Figure F3 Comparison of the detection probability obtained by the proposed algorithm, the random search algorithm in the case of full beam alignment as a function of measurement cost, where SNR = 5dB, and $N_c = 32$ in the random search algorithm .

$L_k = 4, K_{factor} = 20dB$ for $1 < k < K$ [1]. Then, we utilize a regular graph $\Re_d^K(F, m)$ to construct the measurement matrix, where it has $d \geqslant 3$ stages. In Algorithm D1 and Algorithm E1, we take L = 4 to strike a better tradeoff between the performance and complexity order. In particular, for the threshold in Algorithm E1, we set $\varepsilon_1 = 0.04$ and $\varepsilon_2 = 0.15$, so that the algorithm we proposed can achieve the best performance for different scenarios.

In the following simulations, we evaluate the performance of our proposed scheme according to three viewpoints: i) under the noiseless setting, we show the superiority of our proposed noiseless beam alignment algorithm over recent proposed algorithm [3] through the probability of success, such as

$$\frac{\left\|\widetilde{\mathbf{h}} - \widehat{\mathbf{h}}\right\|_2^2}{\left\|\widehat{\mathbf{h}}\right\|_2^2} \leqslant 10^{-6}, \tag{F1}$$

where a trial is considered successful if the above equation is satisfied; ii) under the noise setting, we compare the detection probability of the LOS path with recent proposed random search algorithm, where a success means if the strongest component of each UE coincides with the central of the strongest scatterer cluster; iii) in order to examine the estimation performance of our algorithm, we compare it with the classical compressed sensing algorithm [1], which not only compare the measurement cost, estimation accuracy, but also includes the calculation time and complexity.

Appendix F.1 Noiseless Case

Fig. F1 depicts the probability of success versus different number of measurements of the proposed scheme in comparison with the fast beam alignment algorithm introduced in [3]. As shown in Fig. F1, even the number of measurements is very small (32),

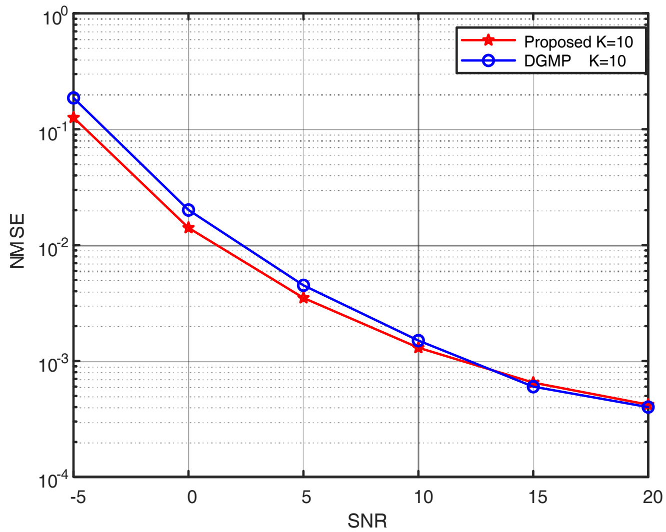

Figure F4 Comparison of the NMSE obtained by the proposed algorithm, the DGMP algorithm as a function of SNR, where $K = 10$. In particular, the number of measurements for the two algorithms is 256 and 96, respectively.

Table F1 Average Run Time of Respective Algorithms: $K = 10$, SNR = 20dB

ALG	**NMSE**	**Average Run Time(s)**
DGMP	4.0013e-4	27.785
Proposed	4.2131e-4	0.565

our proposed noiseless algorithm still achieves a high probability of success (over 90%). Furthermore, compared with the beam alignment algorithm [3], our algorithm provides considerable performance gains up to 16% ($K = 3$) and 40% ($K = 5$) in the success rate. This benefit is due to the fact that our proposed algorithm takes advantage of the *peeling off* method and the density evolution to guarantee the bound achieving performance, even though the fast beam alignment method utilizes a similar precoding structure.

Furthermore, in Fig. F2, the probability of success as a function of the number of transmitting antennas is presented. One can observe that the performance of our proposed scheme remains constant with the growth of M_{T}. This phenomenon implies that the proposed scheme has a measurement independent of M_{T}.

Appendix F.2 Noise Case

To compare with the state-of-the-art beam alignment algorithm, we utilize the random search algorithm, which is modelled as a best case of [4]. Specially, at each time t, the random search algorithm randomly selects a multidirectional beamforming vector w_t and utilizes the PN sequences to ensure the orthogonality of different beams. Then, under the noise setting, the detection probability of the proposed robust beam alignment scheme is compared with the recent random search scheme for $K = 10$ and $K = 16$. In particular, we define a success if the location of the strongest component of the AoAs/AoDs can be detected. As observed in Fig. F3, the proposed algorithm outperforms the random search scheme v.s. different number of measurements. The performance improvement is due to that even though the random search method also sends multiple beams, it requires a longer sequence to maintain the orthogonality of multiple beams. In particular, there is a marked improvement when the measure cost from 96 to 120. This particular phenomenon may be due to the fact that the redundancy parameter μ for constructing this measurement exactly meets the conditions of **Theorem 1**.

Moreover, different from the conventional beam alignment algorithms such as the random search method, our proposed algorithm can extract information such as channel complex gain according to Eq. (E2), so the estimated channel can be reconstructed. Thus, in order to make a comparison with the traditional CS-based algorithm, Fig. F4 examines the performance of the proposed algorithm and the distributed grid matching pursuit (DGMP) algorithm [1] at different SNRs in the case of full beam alignment. Compared with the traditional CS-based algorithm, the proposed beam alignment algorithm can achieve the similar estimated performance at the cost of longer pilot sequences. However, as shown in Table F2, the DGMP algorithm incurs a prohibitive computational complexity due to the large matrix operation, especially the matrix inversion. Furthermore, the simulation results above show that our algorithm achieves a good trade-off between sampling complexity and computational complexity.

References

1 Gao Z, Hu C, Dai L, et al. Channel estimation for millimeter-wave massive MIMO with hybrid precoding over frequency-selective Fading Channels. IEEE Communications Letters, 2016, 20: 1259-1262

2 Zhou Z, Fang J, Yang L, et al. Low-rank tensor decomposition-aided channel estimation for millimeter wave MIMO-OFDM systems. IEEE Journal on Selected Areas in Communications, 2017, 35: 1524-1538

3 Li X, Fang J, Duan H, et al. Fast beam alignment for millimeter wave communications: a sparse encoding and phaseless decoding approach. IEEE Transactions on Signal Processing, 2019, 67: 4402-4417

4 Abari O, Hassanieh H, Rodriguez M, et al. Millimeter wave communications: from point-to-point links to agile network connections. Proc. 15th ACM Workshop Hot Topics Networks, New York, USA, 2016, 169-175