

DEEP LEARNING-BASED BEAM ALIGNMENT IN MMWAVE VEHICULAR NETWORKS

Nitin Jonathan Myers, Yuyang Wang, Nuria González-Prelcic, and Robert W. Heath Jr.

Department of Electrical and Computer Engineering, The University of Texas at Austin.

Email :{nitinjmyers,yuywang,ngprelcic,rheath}@utexas.edu

ABSTRACT

Millimeter wave channels exhibit structure that allows beam alignment with fewer channel measurements than exhaustive beam search. From a compressed sensing (CS) perspective, the received channel measurements are usually obtained by multiplying a CS matrix with a sparse representation of the channel matrix. Due to the constraints imposed by analog frontends, designing CS matrices that efficiently exploit the channel structure is challenging. In this paper, we propose an end-to-end deep learning technique to design a structured CS matrix that is well suited to the underlying channel distribution, leveraging both sparsity and the particular spatial structure that appears in vehicular channels. The channel measurements acquired with the designed CS matrix are then used to predict the best beam for link configuration. Simulation results for vehicular communication channels indicate that our deep learning-based approach achieves better beam alignment than standard CS techniques that use the random phase shift-based design.

Index Terms— Compressed sensing, mmWave, deep learning

1. INTRODUCTION

Beam alignment in millimeter wave (mmWave) radios is a challenging problem due to the use of large antenna arrays and fewer radio frequency chains than antennas [1]. Exhaustive search-based beam alignment results in a substantial overhead at mmWave [2]. A possible approach to reduce this overhead is to use compressed sensing (CS)-based methods that acquire a lower dimensional channel representation [3]. The use of random compressive channel projections together with algorithms that exploit sparsity of mmWave channels can achieve fast and accurate beam alignment [4, 5]. Prior work has shown that structured compressive channel projections may result in better beam alignment than the use of random projections [6].

Convolutional compressed sensing (CCS) is a structured CS technique in which the signal of interest is projected onto fewer circulant shifts of a known signal [7]. We investigate 2D-CCS, i.e., CCS of 2D signals, as our focus is on planar antenna arrays. In 2D-CCS-based beam alignment, the transmitter (TX) applies fewer 2D-circulant shifts of a matrix to its antenna array for the receiver (RX) to acquire compressed channel measurements. Then, the best beam at the TX is estimated from the compressed measurements using optimization algorithms [6]. The matrix used in 2D-CCS, called the base matrix, determines the success of channel sparsity-aware beam alignment. In typical vehicular communication scenarios, channels exhibit structure beyond sparsity in the space of beam directions. For example, some beam directions may be more likely to be optimal than the others [8]. In such a case, the use of

structured random CS matrices can result in better beam alignment than standard designs [9]. In this paper, we use deep learning as a tool to find a 2D-CCS matrix that is well suited to the channel prior.

Prior work in imaging [10–12] and beamforming [13] has considered deep learning-based CS matrix optimization. In this paper, we propose a structured CS matrix optimization framework for the beam alignment problem. Our contributions are as follows. First, we show how 2D-CCS matrices with complex entries can be realized with real valued convolutional layers. Second, we propose a deep learning-based 2D-CCS matrix optimization procedure that accounts for the hardware constraints associated with radio frequency (RF) phase shifters. Last, we interpret the CS matrix optimized with deep learning using properties of the Fourier transform. The structured training in our 2D-CCS-based method results in fewer optimization parameters than the CS matrix optimization approach in [13].

Notation: \mathbf{A} is a matrix, \mathbf{a} is a column vector and a, A denote scalars. \mathbf{A}_R and \mathbf{A}_I are the real and imaginary parts of \mathbf{A} . \bar{a} denotes the complex conjugate of a . $A(k, \ell)$ is the entry of \mathbf{A} in the k^{th} row and the ℓ^{th} column. The ℓ^{th} column of \mathbf{A} is denoted by $\mathbf{A}(:, \ell)$. $|\mathbf{A}|$ is a matrix that contains the magnitude of the entries in \mathbf{A} . The Frobenius norm of \mathbf{A} is $\|\mathbf{A}\|_F$. The inner product of \mathbf{A} and \mathbf{B} is defined as $\langle \mathbf{A}, \mathbf{B} \rangle = \sum_{k, \ell} A(k, \ell) \bar{B}(k, \ell)$. The matrix $[\mathbf{A}; \mathbf{B}]$ is obtained by vertically stacking \mathbf{A} and \mathbf{B} . \mathbf{U}_N is an $N \times N$ unitary discrete Fourier transform (DFT) matrix. $j = \sqrt{-1}$.

2. BEAM ALIGNMENT WITH CONVOLUTIONAL CS

We consider a narrowband system with an $N \times N$ uniform planar array (UPA) at the TX and a single antenna at the RX. The single antenna assumption at the RX is made for simplicity; an extension of our approach to receivers with multiple antennas will be considered in our future work. The TX is equipped with a q -bit phased antenna array that uses a single RF chain and is mounted on a road side unit (RSU). The RSU serves an RX which is mounted on top of a vehicle, as illustrated in Fig. 1. We use $\mathbf{H} \in \mathbb{C}^{N \times N}$ to denote the channel matrix between the UPA at the TX and the single antenna RX. The matrix representation allows our method to exploit the 2D-structure of the channel. Note that the $(r, c)^{\text{th}}$ entry of \mathbf{H} represents the channel coefficient between the $(r, c)^{\text{th}}$ antenna at the TX and the RX antenna. We use \mathbb{Q}_q to denote the set of possible phase shifts in the analog beamforming network, i.e., $\mathbb{Q}_q = \{e^{j2\pi b/2^q}/N : b \in \{1, 2, \dots, 2^q\}\}$. In the m^{th} beam training slot, the TX applies the phase shift matrix $\mathbf{P}[m] \in \mathbb{Q}_q^{N \times N}$ to its phased array. The channel measurement received by the RX for a unit pilot symbol is then

$$y[m] = \langle \mathbf{H}, \mathbf{P}[m] \rangle + v[m], \quad (1)$$

where $v[m] \in \mathbb{C}$ denotes additive white Gaussian noise (AWGN). We define \mathcal{F} as the standard 2D-DFT codebook for the UPA-based

This research was supported in part by the National Science Foundation under Grant numbers NSF-CNS-1731658 and NSF-CNS-1702800.

TX. The transmit beam alignment problem is to estimate a beam-former $\mathbf{P}_{\text{BF}} \in \mathcal{F}$ that maximizes $|\langle \mathbf{H}, \mathbf{P}_{\text{BF}} \rangle|$.

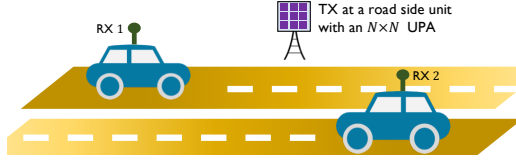


Fig. 1: An illustration of a vehicular communication scenario considered in this paper. The RSU serves vehicles on the two lanes.

A possible beam alignment technique is one that estimates all the N^2 coefficients of \mathbf{H} using scalar projections of the form in (1). Then, beam alignment can be performed using the estimated channel. Estimating the $N \times N$ channel matrix, however, results in a significant overhead as the channel dimension is large in typical mmWave settings. Due to the structure in mmWave vehicular channels, it may be possible to estimate the best beam without explicit channel estimation. For example, prior work has shown that beam alignment can be performed with just $\mathcal{O}(\log N)$ random phase shift-based compressed measurements of a sparse channel [5]. Although random phase shift-based CS performs well, the question is if it is possible to construct a CS matrix that is better matched to the channel prior in vehicular settings.

We now discuss the additional channel structure present in a vehicular communication scenario, such as the one shown in Fig. 1. For an ensemble of receivers within the coverage of an RSU, we show the probability distribution of the best transmit beams in Fig. 2a. The $(i, j)^{\text{th}}$ entry of the beamspace prior in Fig. 2a is the probability that the $(i, j)^{\text{th}}$ 2D-DFT beam, i.e., $\mathbf{U}_N(:, i)\mathbf{U}_N^H(:, j)$, is optimal. In our simulation scenario, the beamspace prior is concentrated over two strips on an $N \times N$ grid of 2D-DFT angles. These strips correspond to directional beams that point along the two lanes shown in Fig. 1. Compressed sensing using random phase shift matrices, however, does not exploit the fact that the beamspace probability is concentrated in a small support set. This is because the beam associated with a random phase shift matrix is quasi-omnidirectional with a high probability; an example of such a beam is shown in Fig. 2b. Sensing with a quasi-omnidirectional pattern is inefficient in vehicular applications, since the vehicle locations are constrained to the road lanes, and many angular directions can be discarded a priori.

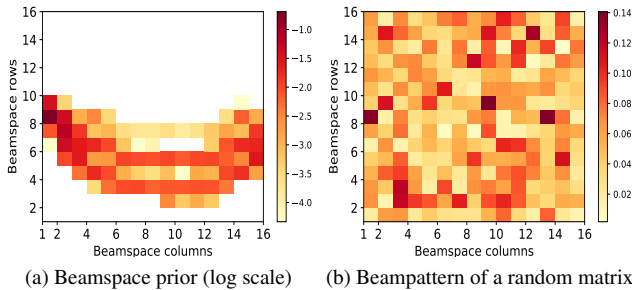


Fig. 2: For the vehicular communication scenario in our simulations, the beamspace prior is non-zero over a small set in the space of 2D-DFT directions. Random phase shift-based CS uses quasi-omnidirectional beams and does not exploit such an information.

2.1. Motivation to use 2D-CCS over standard CS

We describe the measurement model in 2D-CCS when the TX uses $\mathbf{P} \in \mathbb{Q}_q^{N \times N}$ as the base matrix [6]. In 2D-CCS, the TX applies M distinct 2D-circulant shifts of \mathbf{P} to its phased array for the RX

to acquire channel measurements. Consider a scenario where the TX applies an (r, c) 2D-circulant shift of \mathbf{P} to its array. The shifted version is generated by circulantly shifting every column of \mathbf{P} by r units down and every row of the resultant matrix by c units to the right. The noiseless channel measurement in this case is defined as

$$G(r, c) = \sum_{k, \ell} H(k, \ell) \bar{P}((k - r) \bmod N, (\ell - c) \bmod N). \quad (2)$$

For a fixed base matrix \mathbf{P} , the TX can apply a maximum of N^2 distinct 2D-circulant shifts of \mathbf{P} by choosing r and c independently from $\{0, 1, \dots, N - 1\}$. The noiseless measurements acquired by the RX for N^2 2D-circulant shifts are entries of an $N \times N$ matrix \mathbf{G} , where $G(r, c)$ is defined in (2).

The matrix \mathbf{G} in (2) is the 2D-circular cross-correlation between \mathbf{H} and \mathbf{P} , and is expressed as

$$\mathbf{G} = \mathbf{H} \star \mathbf{P}. \quad (3)$$

In 2D-CCS, the TX applies $M \ll N^2$ random 2D-circulant shifts of \mathbf{P} for compressed channel measurements. We use Ω to denote an ordered set of circulant shifts used by the TX. For example, $\Omega = \{(0, 0), (1, 2), (2, 1)\}$ is one possible set for $M = 3$. In this case, the TX applies $(0, 0)$, $(1, 2)$ and $(2, 1)$ circulant shifts of \mathbf{P} in sequence. It can be observed from (2) that the RX acquires $G(0, 0)$, $G(1, 2)$ and $G(2, 1)$. In general, the RX acquires a subsampled version of \mathbf{G} at the ordered locations in Ω with 2D-CCS. We define $\mathcal{P}_\Omega(\mathbf{G})$ as the vector containing the entries of \mathbf{G} at the locations in Ω . The compressed channel measurement vector in 2D-CCS is then

$$\mathbf{y} = \mathcal{P}_\Omega(\mathbf{G}) + \mathbf{v}. \quad (4)$$

The measurements are called convolutional channel measurements as the subsampled cross-correlation operation in (4) can be realized using subsampled convolution. For a well designed base matrix \mathbf{P} , optimization algorithms can estimate the best beam from the 2D-CCS-based measurements in (4) even when $M \ll N^2$ [6].

We explain the efficiency of 2D-CCS in exploiting the beamspace prior using the Fourier transform. We define \mathbf{X} as the 2D-DFT of the channel \mathbf{H} and $\mathbf{Z}[m]$ as the 2D-DFT of the phase shift matrix $\mathbf{P}[m]$. The matrix \mathbf{X} is called the beamspace and the prior associated with the maximizer of $|\mathbf{X}|$ is shown in Fig. 2a. This prior is also the probability distribution of the best beam within the 2D-DFT codebook. To explain the efficiency of 2D-CCS, we consider an example where $\mathbf{P}[1]$ is set to a given base matrix, i.e., $\mathbf{P}[1] = \mathbf{P}$. The first channel measurement is then $\langle \mathbf{H}, \mathbf{P} \rangle$. Due to the unitary nature of the 2D-DFT, the inner product between \mathbf{H} and \mathbf{P} can also be expressed as the inner product of their 2D-DFTs, i.e., $\langle \mathbf{X}, \mathbf{Z}[1] \rangle$. It can be observed from Fig. 2a that the information required for beam alignment is encoded in fewer coefficients of \mathbf{X} , i.e., the entries of \mathbf{X} along the two strips. To extract this information from $\langle \mathbf{X}, \mathbf{Z}[1] \rangle$ in a way that is robust to noise, $\mathbf{Z}[1]$ must have a large amplitude at the locations along the two strips. A base matrix \mathbf{P} whose 2D-DFT satisfies such a property can be used to efficiently acquire features or measurements of the sparse channel.

An interesting property of 2D-CCS is that the matrices used to obtain channel projections have the same 2D-DFT magnitude as the base matrix, i.e., $|\mathbf{Z}[m]| = |\mathbf{Z}[1]| \forall m$. The observation follows from the fact that 2D-circulant shifts over a matrix do not change the magnitude of its 2D-DFT [14]. Using this property, it can be observed that a careful design of the base matrix ensures that all the matrices $\{\mathbf{Z}[m]\}_{m=1}^M$ have a large amplitude at the desired locations. Design a good base matrix, however, is hard due to the non-convex nature of

the optimization problem. In Section 3, we discuss how deep learning can be used to solve the base matrix design problem for efficient 2D-CCS. We assume that the subsampling set Ω in (4) is chosen at random and is fixed throughout optimization. Optimizing the subsampling set is an interesting direction for future work.

3. BASE MATRIX DESIGN WITH DEEP LEARNING

We explain the proposed deep neural network architecture using Fig. 3. The first part of our deep neural network contains a convolutional layer to emulate 2D-CCS-based compressed channel sensing in (4). The weights of the filters in the convolutional layer model the real and imaginary components of the base matrix \mathbf{P} in 2D-CCS. The second part of the network is a cascade of fully connected layers that predicts the best beam from the compressed channel measurements. Through end-to-end learning, our network learns a base matrix \mathbf{P} and the weights of the fully connected layers that maximize the beam alignment probability. In this section, we describe the key components of our network and the training procedure; our implementation is available online on GitHub [15].

3.1. 2D-CCS using real valued convolutional layers

We explain how to implement the circular cross-correlation in (3) using the example of a real matrix $\mathbf{H}_R \star \mathbf{P}_R$. To realize a circulant structure in the correlation, we define a matrix $\mathbf{A}_{R,\text{pad}} = [\mathbf{A}_R, \mathbf{A}_R; \mathbf{A}_R, \mathbf{A}_R]$. Note that $\mathbf{H}_{R,\text{pad}}$ is a $2N \times 2N$ matrix. Now, applying the convolutional filter \mathbf{P}_R over $\mathbf{H}_{R,\text{pad}}$, in a valid-padding mode with a stride of 1, results in an $(N+1) \times (N+1)$ matrix. The 2D-circular cross-correlation $\mathbf{H}_R \star \mathbf{P}_R$ is obtained by simply deleting the last row and the last column of the convolved output.

Our network obtains the complex valued circular cross-correlation by computing the real and imaginary components of \mathbf{G} in (3). As $\mathbf{P} = \mathbf{P}_R + j\mathbf{P}_I$, the real and imaginary components of \mathbf{G} are

$$\mathbf{G}_R = \mathbf{H}_R \star \mathbf{P}_R + \mathbf{H}_I \star \mathbf{P}_I, \text{ and} \quad (5)$$

$$\mathbf{G}_I = \mathbf{H}_I \star \mathbf{P}_R - \mathbf{H}_R \star \mathbf{P}_I. \quad (6)$$

To realize (5) and (6), a $2N \times 4N \times 2$ real valued tensor is first constructed from the complex channel \mathbf{H} . The first slice of this tensor is the $2N \times 4N$ matrix $[\mathbf{H}_{R,\text{pad}}, \mathbf{H}_{I,\text{pad}}]$ and the second slice is $[\mathbf{H}_{I,\text{pad}}, -\mathbf{H}_{R,\text{pad}}]$, as shown in Fig. 3. Then, an $N \times N \times 2$ convolutional filter with \mathbf{P}_R and \mathbf{P}_I as the filter weights, is applied over the input channel tensor. The filtering is performed in a valid-padding mode with a stride of 1 to result in an $(N+1) \times (3N+1)$ matrix. Finally, \mathbf{G}_R and \mathbf{G}_I are $N \times N$ matrices which begin at the $(0, 0)$ and $(2N, 0)$ coordinates of the $(N+1) \times (3N+1)$ matrix.

The channel measurements in (4) can be split into real and imaginary components as $\mathbf{y}_R = \mathcal{P}_\Omega(\mathbf{G}_R) + \mathbf{v}_R$ and $\mathbf{y}_I = \mathcal{P}_\Omega(\mathbf{G}_I) + \mathbf{v}_I$. It is important to note that the subsampling set in (4), i.e., Ω , is same for both the real and the imaginary components of \mathbf{y} . We implement such a subsampling by using an identical dropout technique for both \mathbf{G}_R and \mathbf{G}_I . Our dropout procedure discards $N^2 - M$ entries of \mathbf{G}_R and \mathbf{G}_I at the same locations which are chosen at random. The subsampled features, i.e., $\mathcal{P}_\Omega(\mathbf{G}_R)$ and $\mathcal{P}_\Omega(\mathbf{G}_I)$, are perturbed by AWGN to emulate the compressed sensing model in (4).

3.2. End-to-end learning for beam prediction

The subsampled channel feature vector of dimension $2M \times 1$ is fed into a beam prediction network that is a cascade of fully connected layers. ReLU activation is used at all the fully connected layers and

a softmax is used at the output layer. In this paper, beam alignment is performed using a 2D-DFT codebook which has N^2 elements. Each codebook element is modeled by a class and the output of the neural network is an $N^2 \times 1$ vector that contains the predicted class probabilities.

The network is trained using the restructured channels, i.e., a collection of $2N \times 4N \times 2$ tensors defined in Section 3.1, and the best beam class indices corresponding to the channels. The cross entropy between the predicted class probabilities and the one-hot encoded vectors corresponding to the best beam index is minimized for training. It is important to note that the convolutional filters \mathbf{P}_R and \mathbf{P}_I are part of the neural network and are optimized during training. After successful training, the fully connected layers predict the best beam direction from the $2M \times 1$ real-valued feature vector which represents the compressed channel measurements in (4).

3.3. Weight quantization and implementation in hardware

The filter weights \mathbf{P}_R and \mathbf{P}_I obtained after training may not be hardware compatible, i.e., the corresponding base matrix $\mathbf{P} = \mathbf{P}_R + j\mathbf{P}_I$ may not belong to $\mathbb{Q}_q^{N \times N}$. To this end, we use the weight quantization and retraining approach in [16]. We define $\text{phase}_q(w)$ as the phase of w rounded to a nearest integer multiple of $2\pi/2^q$. The q -bit phase quantized version of \mathbf{P} is defined as $\mathbf{P}_{\text{quant}} = \exp(j\text{phase}_q(\mathbf{P}))/N$. The optimized filter weights \mathbf{P}_R and \mathbf{P}_I are replaced by the real and imaginary parts of $\mathbf{P}_{\text{quant}}$. Quantizing the filter weights in the convolutional layer, however, can reduce the accuracy of the trained neural network. To compensate for the reduction in accuracy, the fully connected layers following the convolutional layers are retrained. It is important to note that the convolutional layer containing $\mathbf{P}_{\text{quant}}$ is not updated during retraining.

Now, we explain how the trained network can be used for beam alignment in mmWave phased arrays. First, the TX applies M circulant shifts of $\mathbf{P}_{\text{quant}}$ to its phased array, according to the subsampling set Ω . For each phase shift matrix applied at the TX, the RX acquires a channel measurement defined by (1). Finally, the RX obtains a subsampled version of $\mathbf{G} = \mathbf{H} \star \mathbf{P}_{\text{quant}}$ according to the model in (4). The received measurements are then reshaped into a vector $[\mathbf{y}_R; \mathbf{y}_I]$ which is fed into the trained fully connected layers for beam prediction. The RX feedbacks the index of this beam and the TX applies the corresponding 2D-DFT codebook element.

4. SIMULATIONS

We consider a vehicular communication scenario with an RSU at a height of 5 m. The TX at the RSU is equipped with a 16×16 UPA with 3-bit phase shifters, i.e., $N = 16$ and $q = 3$. The carrier frequency is set as 28 GHz and the operating bandwidth is 100 MHz. In our simulation scenario, vehicles move along two lanes that are at a distance of 4 m and 7 m from the foot of the RSU. We use Wireless Insite [17], a ray tracing simulator, to obtain channel matrices between the TX and the receivers which are mounted on top of the vehicles. The beamspace prior corresponding to this scenario is shown in Fig. 2a. Note that our dataset contains both line-of-sight (LoS) and non-LoS (NLoS) channels. The NLoS scenario occurs when the link between the RSU and a vehicle is blocked by tall trucks.

The convolutional layer in our network consists of a single filter of dimensions $16 \times 16 \times 2$. This filter is used to optimize the base matrix in 2D-CCS. To emulate M number of 2D-CCS-based channel measurements, the convolved output is subsampled to a $2M \times 1$ feature vector. This vector is then fed into a cascade of four fully connected layers with output dimensions of 80, 256, 512 and 256.

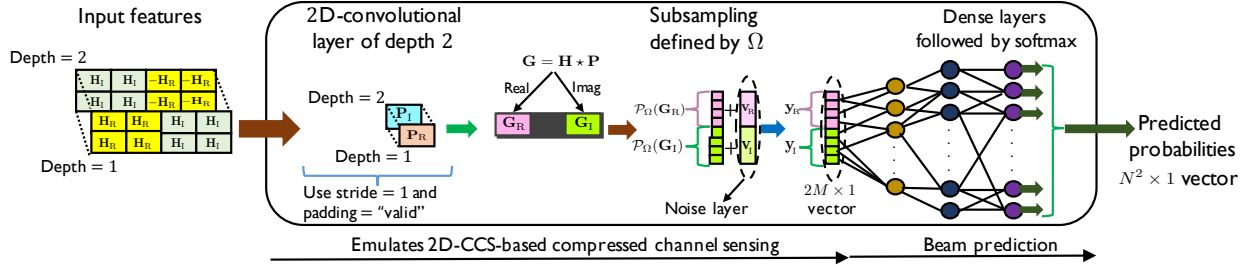
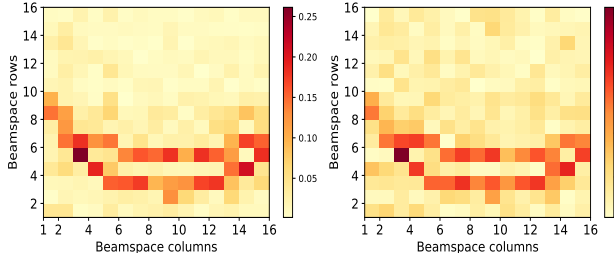


Fig. 3: Channel measurements in 2D-CCS are realized using convolutional filters \mathbf{P}_R and \mathbf{P}_I . Using end-to-end learning, the base matrix in 2D-CCS, i.e., $\mathbf{P} = \mathbf{P}_R + j\mathbf{P}_I$, is optimized to maximize the probability of beam alignment using the compressed channel measurements.

For the beam alignment problem, a reasonable network is one that is invariant to the scaling of the channel. The bias at all the layers is forced to zero to ensure that the network satisfies the scale invariance property. We use about 20,000 channels to train the network and 5,000 other channels for testing. Every channel in the training set is scaled so that $\|\mathbf{H}\|_F = N$. For the test set, however, we use a common scaling for all the channels such that $\mathbb{E}[\|\mathbf{H}\|_F^2] = N^2$. Here, $\mathbb{E}[\cdot]$ denotes the average across the test channels. The channels are then restructured according to the procedure in Section 3.1.



(a) Beam pattern before quantization (b) Beam pattern after quantization

Fig. 4: The plot shows the 2D-DFT magnitude of the base matrices optimized with our procedure for $M = 10$. Weight quantization slightly perturbs the beam pattern. The beams in optimized 2D-CCS radiate power along the directions that are more likely to be optimal.

We explain how our network is trained and provide insights into the optimized base matrix. For M channel measurements, our method first samples M distinct 2D-integer coordinates from an $N \times N$ grid at random and constructs the subsampling set Ω . The $N \times N$ integer grid corresponds to the support of \mathbf{G} in (3). Then, the proposed deep neural network is trained for 30 epochs with the restructured channels and the associated class labels from the training set. A good base matrix is one whose beam pattern has a large magnitude along the directions that are more likely to be optimal. A discrete version of this beam pattern is the 2D-DFT of the base matrix. For the scenario corresponding to Fig. 2a, the 2D-DFT of the optimized base matrix derived after training is shown in Fig. 4a. A 3-bit phase quantization is performed over the optimized base matrix to obtain $\mathbf{P}_{\text{quant}}$; the 2D-DFT of $\mathbf{P}_{\text{quant}}$ is shown in Fig. 4b. It can be observed that the beam pattern corresponding to $\mathbf{P}_{\text{quant}}$ is well matched to the beamspace prior in Fig. 2a. As quantization can lower the beam prediction accuracy of the trained network, the four fully connected layers in our network are retrained for 30 epochs.

We evaluate the proposed deep learning-based 2D-CCS technique in terms of the achievable rate over the test dataset. The $2M \times 1$ compressed channel measurement vector after subsampled convolution with $\mathbf{P}_{\text{quant}}$ is perturbed by AWGN of variance $\sigma^2/2$. The SNR observed at the RX when the TX uses a quasi-omnidirectional pattern is defined as $1/\sigma^2$. As we focus on beam alignment using the 2D-DFT codebook, the SNR after beamforming can be expressed in

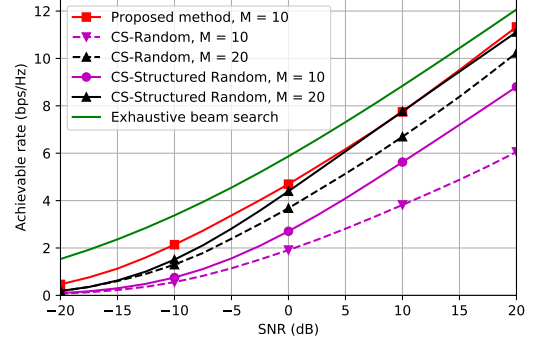


Fig. 5: Deep learning-based beam alignment with an optimized base matrix performs better than standard CS with a random phase shift-based design. The proposed approach can reduce the training overhead by 4 \times when compared to standard CS.

terms of the 2D-DFT of the channel, i.e., \mathbf{X} . The SNR at the RX when the TX applies the $(i, j)^{\text{th}}$ element of the 2D-DFT codebook is $\text{SNR}_{\text{BF}} = |\mathbf{X}(i, j)|^2/\sigma^2$ and the achievable rate is $\log_2(1 + \text{SNR}_{\text{BF}})$. For every channel, the exhaustive search-based approach finds the optimal beam index $(i_{\text{opt}}, j_{\text{opt}})$ where $|\mathbf{X}|$ achieves its maximum. For the proposed deep learning-based approach, the index corresponding to the predicted beam is used to compute the rate. We compare our method with CS-based beam alignment using random phase shifts and structured random phase shifts similar to [9].

The achievable rate plot in Fig. 5 indicates that our deep learning-based approach results in reasonable beam alignment with just $M = 10$ channel measurements. Standard CS with a random phase shift-based design, however, required about $M = 40$ channel measurements to achieve comparable performance. CS with structured random phase shifts needs about $M = 20$ measurements. It can be observed from Fig. 4b that the beams corresponding to the proposed training, i.e., $\{\mathbf{Z}[m]\}_{m=1}^M$, focus power in a small set of directions that are more likely to be optimal. Such beams result in a higher SNR in the compressed channel measurements which translates to better beam alignment performance.

5. CONCLUSIONS

We proposed a novel approach for compressive beam alignment with millimeter wave phased arrays. Our method is based on a structured compressed sensing technique called 2D-convolutional compressed sensing. Any CS matrix in 2D-CCS can be parameterized by a base matrix and a subsampling set. We showed how deep learning can be used as a tool to optimize the base matrix in 2D-CCS. The optimized base matrix achieved superior beam alignment than the common random phase shift-based design and other known structured random designs for a reduced beamtraining overhead.

6. REFERENCES

- [1] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, 2016.
- [2] T. Nitsche, C. Cordeiro, A. B. Flores, E. W. Knightly, E. Perahia, and J. C. Widmer, "IEEE 802.11 ad: directional 60 GHz communication for multi-gigabit-per-second wi-fi," *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 132–141, 2014.
- [3] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, 2008.
- [4] Z. Marzi, D. Ramasamy, and U. Madhow, "Compressive channel estimation and tracking for large arrays in mm-wave picocells," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 514–527, 2016.
- [5] J. Rodríguez-Fernández, N. González-Prelcic, K. Venugopal, and R. W. Heath, "Frequency-domain compressive channel estimation for frequency-selective hybrid millimeter wave MIMO systems," *IEEE Trans. on Wireless Commun.*, vol. 17, no. 5, pp. 2946–2960, 2018.
- [6] N. J. Myers, A. Mezghani, and R. W. Heath Jr, "FALP: fast beam alignment in mmWave systems with low-resolution phase shifters," *IEEE Trans. Commun.*, 2019.
- [7] K. Li, L. Gan, and C. Ling, "Convolutional compressed sensing using deterministic sequences," *IEEE Trans. Signal Process.*, vol. 61, no. 3, pp. 740–752, 2013.
- [8] Y. Wang, A. Klautau, M. Ribero, A. C. Soong, and R. W. Heath, "MmWave vehicular beam selection with situational awareness using machine learning," *IEEE Access*, vol. 7, pp. 87 479–87 493, 2019.
- [9] A. Ali, N. González-Prelcic, and R. W. Heath, "Millimeter wave beam-selection using out-of-band spatial information," *IEEE Trans. on Wireless Commun.*, vol. 17, no. 2, pp. 1038–1052, 2017.
- [10] X. Lu, W. Dong, P. Wang, G. Shi, and X. Xie, "Convcsnet: A convolutional compressive sensing framework based on deep learning," *arXiv preprint arXiv:1801.10342*, 2018.
- [11] E. Zisselman, A. Adler, and M. Elad, "Compressed learning for image classification: A deep neural network approach," in *Elsevier Handbook of Numerical Analysis*, 2018, vol. 19, pp. 3–17.
- [12] I. A. Huijben, B. S. Veeling, K. Janse, M. Mischi, and R. J. van Sloun, "Learning sub-sampling and signal recovery with applications in ultrasound imaging," *arXiv preprint arXiv:1908.05764*, 2019.
- [13] X. Li and A. Alkhateeb, "Deep learning for direct hybrid precoding in millimeter wave massive MIMO systems," *arXiv preprint arXiv:1905.13212*, 2019.
- [14] A. C. Kak and M. Slaney, *Principles of computerized tomographic imaging*. IEEE press, 1998.
- [15] N. J. Myers and Y. Wang, "Deep learning-based beam alignment for structured channels," <https://github.com/nitinjmyers>, 2019.
- [16] S. Han, H. Mao, and W. J. Dally, "Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding," *arXiv preprint arXiv:1510.00149*, 2015.
- [17] "Remcom wireless insite," <http://www.remcom.com/wireless-insite>.