FALP: Fast beam alignment in mmWave systems with low-resolution phase shifters

Nitin Jonathan Myers, Student Member, IEEE, Amine Mezghani, Member, IEEE, and Robert W. Heath Jr., Fellow, IEEE.

Abstract—Millimeter wave (mmWave) systems can enable high data rates if the link between the transmitting and receiving radios is configured properly. Fast configuration of mmWave links, however, is challenging due to the use of large antenna arrays and hardware constraints in these systems. The large amount of training overhead incurred by exhaustive search-based beam alignment in typical mmWave systems is one common example. We present a framework for Fast beam Alignment with Low-resolution Phase shifters which we refer to as FALP. FALP designs an efficient set of antenna weight vectors to acquire channel measurements, and allows faster beam alignment when compared to exhaustive scan. The antenna weight vectors in FALP can be realized in ultra-low power phase shifters whose resolution can be as low as one-bit. From a compressed sensing (CS) perspective, the CS matrix designed in FALP satisfies the restricted isometry property and allows CS algorithms to exploit the fast Fourier transform. The proposed framework also establishes a new connection between channel acquisition in phased arrays and magnetic resonance imaging.

Index Terms—Perfect arrays, 2D sparse recovery, one-bit phase shifters, magnetic resonance imaging, mm-Wave

I. INTRODUCTION

Millimeter wave communication, currently used in the IEEE 802.11ad standard [1], can support giga-bit-per-second data rates by exploiting the large amount of bandwidth available at mmWave carrier frequencies [2]. Although existing circuit technology can support communication at mmWave frequencies, link configuration in typical mmWave hardware can be challenging [3]. For example, in a phased antenna array-based mmWave radio that is used in IEEE 802.11ad standard compliant devices [1], the number of radio frequency (RF) chains is smaller than the number of antennas [4]. As a result, phased arrays can only obtain a lower dimensional projection of the multiple-input multiple-output (MIMO) channel. Due to the use of large number of antennas and the lower dimensional nature of channel projections in mmWave phased arrays, standard exhaustive search-based beam alignment can result in a lot of training overhead [3].

Compressed sensing (CS) is a technique that allows reconstructing a sparse signal with fewer measurements when compared to the dimension of the signal [5], [6]. Due to the sparse nature of mmWave channels in an appropriate dictionary, CS is a promising solution for mmWave channel estimation or beam alignment with sub-Nyquist channel measurements [7]. The channel measurements in CS are obtained by projecting the channel onto a lower dimensional subspace using a CS matrix [5]. The channel is then recovered from the lower dimensional projections using optimization techniques that exploit sparsity of the channel [7], [8]. The guarantees on the recovery of sparse signals and the complexity of the reconstruction algorithms, however, depend on the choice of the CS matrix used to obtain these projections. The restricted isometry property (RIP) [9] is one metric that characterizes the efficiency of a CS matrix in recovering sparse signals. Unfortunately, several random CS matrices that are known to satisfy the RIP cannot be realized in the common phased array architecture due to hardware constraints [10]. Prior work has used random IID phase shift-based CS matrices for channel estimation or beam alignment [7], [8]. Although such matrices work well, it is possible that other CS matrices can achieve better beam alignment performance at a lower computational complexity [10], [11]. In this paper, we design CS matrices that are compatible with planar phased arrays, satisfy the RIP, and aid low complexity CS algorithms.

Convolutional compressed sensing (CCS) is a form of structured compressed sensing that is useful in hardware constrained signal acquisition systems like the phased array [12]. In CCS, the sparse signal is projected onto different circulantely shifted versions of a modulation sequence. For a properly designed modulation sequence, CCS can recover the sparse signal using fewer circulant shift-based projections when compared to the dimension of the signal. Most of the work on CCS, however, is limited to the acquisition of sparse vectors and the design of efficient modulation sequences [13]. Prior work has shown that ideal modulation sequences for vector CCS exist up to a certain dimension that depends on the size of the alphabet [12], [14]. For CCS in phased arrays, the size of the alphabet is determined by the resolution of the phase shifters. Due to the finite resolution of phase shifters in practical mmWave radios, it may not be possible to perform efficient vector CCS in large antenna arrays.

In this paper, we propose a novel framework called FALP for 2D convolutional acquisition of mmWave channel matrices in planar phased arrays. The channel measurements in our 2D-CCS framework are obtained by projecting the channel matrix onto 2D-circulantely shifted versions of a hardware compatible modulation matrix. Similar to vector CCS, the performance of 2D-CCS depends on the choice of the modulation matrix. As the number of hardware compatible matrices in phased arrays is exponential in the array dimensions, a brute-force...
approach to find the best modulation matrix is not practically feasible for large arrays. In our prior work, i.e., Swift-Link [10], we used Zadoff-Chu (ZC) sequences for efficient CCS-based channel estimation in planar arrays. For linear phased arrays, the efficiency of CCS-based channel estimation with ZC sequences was studied in [11]. The techniques in [10] and [11] assume a Kronecker product codebook, i.e., the equivalent modulation matrix used in a 2D-CCS setting is an outer product of two ZC sequences. Realizing such ZC-based modulation matrices, however, requires a phase shift resolution that is logarithmic in the number of antennas [11]. As a result, the ZC-based CCS techniques in [10] and [11] cannot be used in large phased arrays with ultra-low resolution phase shifters. This result allows applying FALP to large phased arrays with one-bit phase shifters, and makes it a candidate solution for next generation wireless systems. We summarize our main contributions as follows.

- For analog beamforming with uniform planar arrays, we propose a framework called FALP that realizes partial 2D-DFT CS matrices using partial 2D-circulant shifts of a modulation matrix. We characterize the properties of hardware compatible modulation matrices that preserve sparsity of the mmWave channel.

- We show that perfect arrays [15], [16] can be used as efficient modulation matrices in 2D-CCS. For a given resolution of phase shifters, such arrays exist for a family of array dimensions. For other cases, we derive the sub-optimality gap of CS algorithms when non-ideal modulation matrices are used in FALP.

- We provide a beamspace de-convolution perspective to CS-based beam alignment using ideas from CS in magnetic resonance imaging (MRI) [17]. For ultra-low-complexity beam alignment using a single 2D-fast Fourier transform, we derive lower bounds on the beam alignment probability using results from CS in MRI. Our analysis incorporates the impact of multi-path interference on beam discovery and matches well with simulations.

- With standard CS-based techniques, FALP’s training achieves better beam alignment than the commonly used random phase shift-based training for the same settings. Furthermore, CS algorithms that use the proposed training can exploit the fast Fourier transform and have a lower computational complexity than those that use random phase shift-based training.

We would like to highlight that Swift-Link and FALP solve two independent problems. On the one hand, FALP develops efficient modulation matrices for 2D-CCS in low resolution phased arrays. On the other hand, Swift-Link designs trajectories to perform CS-based beam alignment that is robust to carrier frequency offset (CFO). For simplicity of exposition, we assume perfect frame timing and carrier synchronization. Nevertheless, Swift-Link’s trajectory can be used in FALP for CFO robust beam alignment in low-resolution phased arrays.

**Notation:** $A$ is a matrix, $a$ is a column vector and $a, A$ denote scalars. Using this notation $A^T, \overline{A}$ and $A^*$ represent the transpose, conjugate and conjugate transpose of $A$. We use $\text{diag}(a)$ to denote a diagonal matrix with entries of $a$ on its diagonal. The scalar $a[m]$ denotes the $m^{th}$ element of $a$. The $\ell_2$ norm of $a$ is denoted by $\|a\|_2$. The $k^{th}$ row and the $\ell^{th}$ column of $A$ are denoted by $A(k, :)$ and $A(:, \ell)$. The scalar $A(k, \ell)$ or $A_{k,\ell}$ denotes the entry of $A$ in the $k^{th}$ row and $\ell^{th}$ column. The matrix $|A|$ contains the element-wise magnitude of $A$, i.e., $|A|_{k,\ell} = |A_{k,\ell}|$. The entries of $1$ is denoted by $\|A\|_1$. The inner product of two matrices $A$ and $B$ is defined as $\langle A, B \rangle = \sum_{k,\ell} A(k, \ell) B(k, \ell)$. We use $I$ to denote an all-ones matrix and $I$ to denote the identity matrix. The symbols $\odot$ and $\otimes$ are used for the Hadamard product and 2D circular convolution [18]. The unitary Discrete Fourier Transform (DFT) matrix. The set $\mathbb{Z}_N$ denotes the set of integers $\{0, 1, 2, ..., N-1\}$. We use $e_k$ to represent the $(k + 1)^{th}$ canonical basis vector.

## II. SYSTEM AND CHANNEL MODEL

In this section, we describe a planar phased antenna array system considered in FALP. To explain our framework, we assume a narrowband mmWave system and focus on the transmit beam alignment problem. We extend our algorithm to the wideband setting in Section VIII.

### A. System model

![Fig. 1: Channel acquisition in a phased array system with a uniform planar array of antennas at the transmitter. The receiver is assumed to operate with a fixed beam pattern during channel acquisition for transmit beam alignment.](image)

We consider an analog beamforming system in which the transmitter (TX) is equipped with a uniform planar array (UPA) of antennas as shown in Fig. 1. For ease of notation, we consider an equal number of antennas, i.e., $N$, along each of the azimuth and elevation dimensions of the UPA. Our framework can also be extended to other rectangular arrays using array response vectors of appropriate dimensions in the formulation. The beamforming architecture at the TX uses a single radio frequency (RF) chain as shown in Fig. 1. Each antenna element in the UPA is connected to the RF chain through a digitally controlled phase shifter. By appropriately configuring the phase shifters, the TX can perform directional transmission [4]. In practice, the resolution of the phase shifters is limited to reduce the power consumption in the mmWave phased array [3]. For a $q$-bit phase shifter, we define the set of possible phase shifts as $Q_q = \{ e^{j2\pi k/2^q} : k \in \mathbb{Z}_{2^q} \}$. As each antenna in the UPA is connected to a unique phase shifter, it is possible to configure $N^2$ phase shifters. As a result, the
phase shift matrix applied to the phased array at the TX is constrained to be an element in \( \mathbb{Q}_q^{N \times N} \). The transmit beam alignment problem is to determine a phase shift matrix at the TX that maximizes the SNR at the receiver (RX).

A possible way to perform transmit beam alignment is to estimate a reasonable approximation of the channel and use it to configure the phased array. Let \( \mathbf{H} \in \mathbb{C}^{N \times N} \) be the antenna domain channel matrix when a fixed beam pattern is used at the RX. This fixed beam pattern at the RX can correspond to quasi-omnidirectional or directional reception. The RX acquires projections of \( \mathbf{H} \) on different phase shift matrices, in successive training slots. In the \( m \)th training slot, the TX applies the phase shift matrix \( \mathbf{P}[m] \in \mathbb{Q}_q^{N \times N} \) to its phased array, and the RX acquires the channel measurement \( y[m] \).

We use \( M \) to denote the number of channel measurements acquired by the RX. In this paper, we assume perfect frame timing and carrier synchronization. Our assumption is valid in cellular scenarios where synchronization is performed using separate control channels. With the perfect synchronization assumption, the channel measurement obtained by the RX, in the \( m \)th training slot, can be expressed as

\[
y[m] = (\mathbf{H}, \mathbf{P}[m]) + v[m],
\]

where \( v[m] \sim N_c(0, \sigma^2) \) is additive white Gaussian noise. As each channel measurement in \( \mathbb{C}^{N \times N} \) is a scalar, estimating a generic \( N \times N \) channel matrix requires \( M = N^2 \) channel measurements. Exhaustive beam search is one such approach that obtains the projections of \( \mathbf{H} \) on all the \( N^2 \) elements of the 2D-DFT dictionary \([3]\). Such a solution, however, does not scale well with the array dimensions. Prior work has exploited the structure in mmWave channels through CS, to estimate them with \( M < N^2 \) measurements \([7], [8]\). In this paper, we develop a novel set of phase shift matrices for channel acquisition in CS. We prove that a good approximation of mmWave channels can be obtained from \( M = \mathcal{O}(\log N) \) channel measurements that are acquired using the proposed set. We also show that CS algorithms that use the proposed design have a lower complexity than those that use the common random phase shift-based design \([7], [8]\).

### B. Channel model

We consider a geometric-ray-based model for the narrow-band mmWave channel when the RX receives signals with a fixed beam pattern \([4]\). Let \( \gamma_k, \theta_{\ell,k} \) and \( \theta_{\ell,k} \) denote the complex ray gain, elevation angle-of-departure and azimuth angle-of-departure of the \( k \)th ray. At this point, we do not make any assumption on these parameters and they can come from any distribution. We define the beamspace angles \( \omega_{\ell,k} = \sin \theta_{\ell,k} \sin \theta_{\ell,k} \) and \( \omega_{\ell,k} = \cos \theta_{\ell,k} \cos \theta_{\ell,k} \). We define the Vandermonde vector \( \mathbf{a}_N(\Delta) \in \mathbb{C}^{N \times 1} \) as

\[
\mathbf{a}_N(\Delta) = \left[ 1, e^{j\Delta}, e^{j2\Delta}, \ldots, e^{j(N-1)\Delta} \right]^T.
\]

The wireless channel for a half wavelength spaced UPA in the baseband is given by

\[
\mathbf{H} = \sum_{k=1}^{K} \gamma_k \mathbf{a}_N(\omega_{\ell,k}) \mathbf{a}_N^T(\omega_{\ell,k}).
\]

As large antenna arrays are used in typical mmWave settings, the dimension of the channel, i.e., \( N^2 \), can be large in mmWave systems when compared to conventional lower frequency systems.

Channel matrices at mmWave are sparse in a well chosen dictionary, because of the propagation characteristics of the environment \([3]\). For UPAs, the 2D-DFT basis is often chosen for a sparse representation of \( \mathbf{H} \) \([19]\). Let \( \mathbf{X} \in \mathbb{C}^{N \times N} \) denote the inverse 2D-DFT of \( \mathbf{H} \), such that

\[
\mathbf{H} = \mathbf{X} \mathbf{U}_N. \tag{4}
\]

The matrix \( \mathbf{X} \) is called the beamspace channel as it contains the channel coefficients seen when different directional beams are used at the TX \([19]\). The sparsity of the mmWave channel in the angle domain translates to the sparsity of the beamspace channel matrix \( \mathbf{X} \). As the beamspace angles-of-departure (AoD) in the channel may not align exactly with those corresponding to the DFT dictionary, there can be leakage effects in the 2D-DFT representation \([3]\). As a result, the matrix \( \mathbf{X} \) is approximately sparse. In such case, dictionaries that use a finer AoD domain representation can be used for a sparser representation of \( \mathbf{H} \) \([7]\). Using such a dictionary, however, increases the dimensionality of the CS problem. For our analysis, we consider \( \mathbf{X} \) to be perfectly sparse, while our simulation results are for the realistic case where \( \mathbf{X} \) is approximately sparse.

### III. Convolutional CS in planar arrays

In this section, we propose a 2D-CCS framework to acquire matrices that are sparse in the 2D-DFT basis. It is important to note that the proposed framework is not the same as standard CS \([12]\) of the vectorized version of a matrix.

An efficient way to acquire a sparse signal is to sample it using a basis that is maximally incoherent with the sparsity basis \([5]\). For example, it is known that the canonical basis is maximally incoherent with the DFT basis \([5]\). As a result, any signal that is sparse in the frequency domain can be recovered from fewer time domain samples compared to the dimension of the signal \([5]\). Similarly, matrices that have a sparse 2D-DFT representation can be recovered from fewer samples in the canonical representation. Although \( \mathbf{H} \) in \( \mathbf{H} \) has a sparse 2D-DFT representation, it is not possible to sample the entries of \( \mathbf{H} \) using phased arrays \([3]\). For example, acquiring \( \mathbf{H}(k, \ell) = (\mathbf{H}, \mathbf{e}_k \mathbf{e}_\ell^T) \) in a single training slot requires the application of \( \mathbf{e}_k \mathbf{e}_\ell^T \) to the phased array. Unfortunately, such canonical projection matrices that achieve maximum incoherence in sampling mmWave channels cannot be realized in phased arrays as \( \mathbb{Q}_q^{N \times N} \) does not contain the canonical basis elements.

The notion behind considering 2D-convolutional CS comes from the observation that the canonical projection matrices are 2D-circulantly shifted versions of a base matrix, i.e., \( \mathbf{e}_k \mathbf{e}_\ell^T \). In this paper, we show that maximally incoherent sampling can also be performed using 2D-circulantly shifted versions of a properly chosen matrix in \( \mathbb{Q}_q^{N \times N} \). Let \( \mathbf{P} \in \mathbb{Q}_q^{N \times N} \) be the matrix used for 2D-CCS in the phased array. The matrix \( \mathbf{P} \) is defined as a base matrix or a modulation matrix.
use $J \in \mathbb{R}^{N \times N}$ to denote a circulant delay matrix with its first row as $(0,1,0,0,...,0)$. The subsequent rows of $J$ are generated by right circularly shifting the previous row by 1 unit. For an $N \times 1$ vector $x$, the transformed vector $Jx$ is $(x[1], x[2],..., x[N-1], x[0])^T$. Using this notation, we define the $d$ circulant delay matrix as $J_d = J \cdot J \cdot \cdots \cdot J$ ($d$ times). Analogous to the canonical sampling case, a channel measurement is obtained by applying a 2D-circulantly shifted version of the base matrix $P$ to the phased array. In the $m^{th}$ training slot, the TX applies a phase shift matrix $P[m]$, that is an $(r[m], c[m])$ 2D-circulantly shifted version of $P$. The matrix $P[m]$ is obtained by circularly shifting the rows of $P$ by $r[m]$ units and circularly shifting the columns of the resultant matrix by $c[m]$ units, i.e.,

$$P[m] = J_{r[m]}^cP J_{c[m]}.$$ (5)

Note that $r[m] \in \mathbb{I}_N$ and $c[m] \in \mathbb{I}_N$. For a given $P$, it can be observed that a maximum of $N^2$ phase shift matrices can be used to obtain channel measurements with 2D-CCS.

Now, we show that channel measurements obtained using 2D-CCS can be interpreted as projections of a masked beamspace. A similar conclusion was derived in [10] for channel acquisition using ZC-based phase shift matrices of rank one. Realizing ZC-based matrices, however, may require high resolution phase shifters for channel acquisition [10]. Using (14) and (5), the $m^{th}$ channel measurement is

$$y[m] = \langle H, J_{r[m]}^cP J_{c[m]} \rangle v[m].$$ (6)

Using the property that the 2D-DFT preserves the inner product between any two matrices [18], i.e., $\langle A, B \rangle = \langle U_N A U_N^*, U_N B U_N^* \rangle$, we rewrite (6) as

$$y[m] = \langle U_N H U_N^*, U_N J_{r[m]}^cP J_{c[m]} U_N^* \rangle v[m].$$ (7)

As $J_{c[m]}$ is a circulant matrix, the columns of the DFT matrix serve as its eigenvectors. The matrix $J_{c[m]}$ can be diagonalized as $U_N J_{c[m]} U_N = \Lambda_{c[m]}$; equivalently $J_{c[m]} U_N = U_N \Lambda_{c[m]}$. In this case, $\Lambda_{c[m]}$ is a diagonal matrix that contains the scaled DFT of the first row of $J_{c[m]}$ on its diagonal, i.e., $\Lambda_{c[m]} = \text{diag}(\sqrt{N} u_n e_{c[m]}(\nu))$. Similarly, we have $J_{r[m]} U_N = U_N \Lambda_{r[m]}$. It can be observed that the circulant delay matrix is real, i.e., $J_d = J_d^\dagger$. Putting these observations together with (14), we rewrite (7) as

$$y[m] = \langle X, (J_{r[m]} U_N)^* P J_{c[m]} U_N \rangle v[m]$$ (8)

$$= \langle X, (U_N A_{r[m]} U_N^*)^* P U_N \Lambda_{c[m]} U_N \rangle v[m]$$ (9)

$$= \langle X, \Lambda_{r[m]}^* U_N^* P U_N A_{c[m]} \rangle v[m].$$ (10)

We define the vectors $\bar{r}[m] = U_N e_{r[m]}$ and $\bar{c}[m] = U_N e_{c[m]}$ as the $r[m]$ and $c[m]$ columns of the DFT matrix. Using the identity diag(a)Bdiag(c) = $B \odot (ac^T)$, one of the matrices in the inner product in (10) can be expressed as

$$\Lambda_{r[m]}^* U_N^* P U_N A_{c[m]} = N \text{diag}(\bar{r}[m]) U_N^* P U_N \text{diag}(\bar{c}[m])$$ (11)

$$= (N U_N^* P U_N) \odot (\bar{r}[m] \bar{c}[m]^T).$$ (12)

The channel measurement $y[m]$ can be rewritten using (10) and (12) as

$$y[m] = \langle X, (N U_N^* P U_N) \odot (\bar{r}[m] \bar{c}[m]^T) \rangle + v[m].$$ (13)

We use the identity $\langle A, B \odot C \rangle = \langle A \odot \bar{B}, C \rangle$ to express (13) as

$$y[m] = \langle X \odot (N U_N P U_N), \bar{r}[m] \bar{c}[m]^T \rangle + v[m].$$ (14)

We define the spectral mask [10] of the base matrix $P$ as

$$Z = N U_N P U_N,$$ (15)

and the masked beamspace matrix [10] as $S = X \odot Z$. The spectral mask $Z$ multiplies each entry of the true beamspace $X$ to result in a masked beamspace matrix. For any sparse beamspace matrix $X$, the masked beamspace matrix $S$ is also sparse because $S(k, \ell) = 0$ when $X(k, \ell) = 0$.

We now transform 2D-CCS of mmWave channel matrices into a partial 2D-DFT problem over the masked beamspace. Similar to the definition of $H$ in [4], we define the virtual channel matrix $G = U_N S U_N$ as the 2D-DFT of the masked beamspace. Using (14) and the identity $\langle A, b c^T \rangle = b^* A c$, (14) can be expressed as

$$y[m] = \bar{r}[m] S \bar{c}[m] + v[m]$$ (16)

$$= e_{r[m]}^T U_N S U_N e_{c[m]} + v[m]$$ (17)

$$= \langle G, e_{r[m]} e_{c[m]}^T \rangle + v[m]$$ (18)

$$= G(r[m], c[m]) + v[m].$$ (19)

An interesting observation from (18) is that the $(r[m], c[m])$ entry of the virtual channel matrix $G$ can be sampled by applying a $(r[m], c[m])$ 2D-circulantly shifted version of the base matrix. In contrast, the entries of the true channel matrix $H$ cannot be sampled directly due to hardware constraints. The most intriguing aspect in our 2D-CCS framework is that virtual channel acquisition uses all the $N^2$ antennas at the TX, when compared to true channel acquisition that uses a single antenna. In practical settings where there is a power constraint per antenna, virtual channel samples can have a higher SNR over corresponding true channel samples. For example, it was shown in [10] that ZC-based acquisition of sparse mmWave channels is power efficient over true channel acquisition by a factor of $N^2$.

An important observation from (6) and (17) is that convolutional acquisition of a channel that is sparse in the Fourier basis, results in subsampled Fourier measurements of its masked beamspace. The masked beamspace $S$ can be efficiently recovered from a subsampled version of $G$ using ideas from partial 2D-DFT compressed sensing [13, 20]. Recovering $S = Z \odot X$, however, does not guarantee the reconstruction of the true beamspace, i.e., $X$. For example, if $Z(k, \ell) = 0$ for some $k$ and $\ell$, it is not possible to recover $X(k, \ell)$ as the masked beamspace component $S(k, \ell) = 0$. To avoid such blanking effects in the masked beamspace, well conditioned spectral masks must be designed to recover $X$ from the spectral mask equation, i.e., $S = Z \odot X$.

IV. CS WITH PHASE SHIFTERS: ONE BIT CAN DO IT

In this section, we translate the guarantees on CS-based recovery of the masked beamspace to the true beamspace, and derive conditions on the spectral mask, i.e., $Z$, for efficient
CS. We show that perfect arrays designed in [15] and [16] can be used as efficient spectral masks. Furthermore, we prove that the guarantees and performance of CS algorithms can be made independent of the resolution of phase shifters for several array configurations. For such configurations, FALP can be implemented with low-cost and low-power phased arrays that use just one-bit phase shifters.

For recovery of the sparse masked beamspace matrix \( \mathbf{S} \), the CS matrix that results from acquiring \( M = O(\log N) \) 2D-DFT samples of \( \mathbf{S} \) at random is known to satisfy the restricted isometry property with high probability [13]. It can be observed from [6] that the \( M \) channel measurements in our framework are defined by the base matrix \( \mathbf{P} \) and the set \( \Omega = \{(r[1], c[1]), (r[2], c[2]), \ldots, (r[M], c[M])\} \). We use \( \mathcal{P}_\Omega : \mathbb{C}^{N \times M} \to \mathbb{C}^M \) to denote the projection operator that returns the entries of a matrix at the locations in \( \Omega \). The vector of \( M \) masked beamspace measurements defined by \( \Omega \) can be expressed using [17] as

\[
y = \mathcal{P}_\Omega(\mathbf{U}_N \mathbf{S} \mathbf{U}_N) + \mathbf{v}. \tag{20}
\]

The masked beamspace \( \mathbf{S} \) can be estimated from the channel measurements in (20) using [5]

\[
\hat{\mathbf{S}} = \arg \min \| \mathbf{y} \|_1, \text{s.t.} \| \mathbf{y} - \mathcal{P}_\Omega(\mathbf{U}_N \mathbf{W} \mathbf{U}_N) \|_2 \leq \sqrt{M} \sigma. \tag{21}
\]

The optimization program in (21) “encourages” sparse masked beamspace solutions that are consistent with the observed channel measurements [5].

Now, we derive reconstruction guarantees for compressive beamspace reconstruction via (21) when every coordinate in \( \Omega \) is chosen uniformly at random from \( \mathcal{I}_N \times \mathcal{I}_N \). Let \( \mathbf{X} \) be a solution to the true beamspace. As \( \mathbf{S} = \mathbf{Z} \odot \mathbf{X} \), the estimate \( \hat{\mathbf{X}} \) must satisfy \( \hat{\mathbf{S}} = \mathbf{Z} \odot \hat{\mathbf{X}} \). For the spectral mask \( \mathbf{Z} \), we define \( Z_{\text{max}} = \max_{k,\ell} |\mathbf{Z}(k,\ell)| \) and \( Z_{\text{min}} = \min_{k,\ell} |\mathbf{Z}(k,\ell)| \). We use \( (\mathbf{A})_k \) to denote the \( k \) sparse representation of \( \mathbf{A} \). The matrix \( (\mathbf{A})_k \) is obtained from \( \mathbf{A} \) by retaining the \( k \) largest entries in magnitude and setting the rest to 0.

**Theorem 1.** For a fixed constant \( \gamma \in (0, 1) \), a solution \( \hat{\mathbf{X}} \) such that \( \hat{\mathbf{S}} = \mathbf{Z} \odot \hat{\mathbf{X}} \) satisfies

\[
\| \mathbf{X} - \hat{\mathbf{X}} \|_F \leq C_1 \frac{Z_{\text{max}} \| \mathbf{X} - (\mathbf{X})_k \|_1}{\sqrt{k}Z_{\text{min}}} + C_2 \frac{N \sigma}{Z_{\text{min}}}, \tag{22}
\]

with a probability of at least \( 1 - \gamma \) if \( M \geq Ck \max \{2\log^2(2k) \log(N), \log(\gamma^{-1})\} \). The constants \( C, C_1 \) and \( C_2 \) are independent of all the other parameters.

**Proof.** See Section III-A of [21].

For a given \( \mathbf{X}, M \), and \( \sigma \), it can be observed from Theorem 1 that the reconstruction error, i.e., \( \| \mathbf{X} - \hat{\mathbf{X}} \|_F \), depends on the maximum and minimum values of \( |\mathbf{Z}| \). To achieve a tight upper bound on the reconstruction error, \( Z_{\text{max}} \) must be minimized and \( Z_{\text{min}} \) must be maximized. As \( \| \mathbf{P} \|_F = 1 \) for any \( \mathbf{P} \in \mathbb{Q}_q^{N \times N} \), it can be observed from [15] that the norm of the spectral mask is \( \| \mathbf{Z} \|_F = N \). Therefore, \( Z_{\text{max}} \) and \( Z_{\text{min}} \) satisfy \( Z_{\text{max}} \geq 1 \) and \( Z_{\text{min}} \leq 1 \). It follows from (22) that the best reconstruction guarantee for \( \hat{\mathbf{X}} \) is obtained when \( Z_{\text{max}} = 1 \) and \( Z_{\text{min}} = 1 \), i.e., all the \( N^2 \) entries of \( \mathbf{Z} \) must be unimodular. The unimodular property of the spectral mask \( \mathbf{Z} \) is determined by the base matrix. Finding a base matrix \( \mathbf{P} \in \mathbb{Q}_q^{N \times N} \) such that its 2D-DFT in (15) is unimodular, however, is a difficult non-convex optimization problem. Prior work has considered subsampled convolution using random sequences [13]; the 2D extension of such technique is 2D-CCS using a \( \mathbf{P} \) that is chosen at random from \( \mathbb{Q}_q^{N \times N} \). A random choice for \( \mathbf{P} \), however, may not satisfy the unimodular 2D-DFT property. In this paper, we show that ideal base matrices, i.e., matrices with a unimodular 2D-DFT, exist for several combinations of \( q \) and \( N \).

**A. When is the spectral mask unimodular?**

In this section, we show the equivalence between unimodularity of \( \mathbf{Z} \) and the perfect periodic spatial autocorrelation property of its inverse 2D-DFT, i.e., \( \mathbf{P} \). We provide details about base matrices that satisfy the desired autocorrelation properties in Sec. IV-B. For an \( N \times N \) matrix \( \mathbf{A} \), we define its \( N \times N \) spatial autocorrelation matrix \( \mathbf{R}_A \) as [21]

\[
\mathbf{R}_A(x, y) = \sum_{k=0}^{N-1} \sum_{\ell=0}^{N-1} \mathbf{A}(k, \ell) \overline{\mathbf{A}(k + x, \ell + y)}. \tag{23}
\]

Note that the addition of indices in (23) is calculated modulo \( N \), and autocorrelation implies periodic autocorrelation in this paper. We define \( \delta(x, y) \) as the 2D-Dirac delta function, i.e., \( \delta(0, 0) = 1 \) and \( \delta(x, y) = 0 \) whenever \( x \neq 0 \) or \( y \neq 0 \). The duality between perfect spatial autocorrelation and unimodular 2D-DFT properties is explained in Theorem 2.

**Theorem 2.** A base matrix \( \mathbf{P} \) has a unimodular spectral mask, i.e., \( Z_{\text{max}} = 1 \) and \( Z_{\text{min}} = 1 \), if and only if

\[
\mathbf{R}_P(x, y) = \delta(x, y) \quad \forall x \in \mathcal{I}_N \text{ and } y \in \mathcal{I}_N. \tag{24}
\]

**Proof.** See Section III-A of [21].

Note that 2D-CCS obtains projections of the channel matrix over different circulantly shifted versions of the base matrix \( \mathbf{P} \). When \( \mathbf{P} \) satisfies the perfect autocorrelation property, it can be observed that the set of \( N^2 \) 2D-circulantly shifted versions of \( \mathbf{P} \) forms an orthogonal basis of \( \mathbb{C}^{N \times N} \). Due to the unimodularity of the spectral mask, channel acquisition using projections over randomly chosen elements of such a basis set achieves maximum incoherence [5]. As a result, the proposed framework to acquire 2D-DFT sparse signals is efficient when base matrices with perfect autocorrelation are used in 2D-CCS.

**B. Perfect arrays as base matrices**

The problem of designing matrices over small alphabets, i.e., a small \( q \), with perfect spatial autocorrelation has been well investigated in [15] and [16]. Such matrices, called perfect arrays, are known to exist for several array dimensions that extend up to infinity. Although there are several hardware constrained 2D-CS applications, perfect arrays over finite alphabets have not been used in the context of CS to the best of our knowledge.
The construction of perfect arrays over large alphabets, i.e., a large \( q \), can be trivial, but is challenging for a small \( q \). For example, a matrix that is an outer product of two Zadoff-Chu sequences satisfies the conditions in Theorem 2 and is a perfect array. The ZC-based matrix, however, may not be realizable in low resolution phased arrays [10], [11]. Quantizing the phase of the ZC-based matrix using a small \( q \) may distort its perfect autocorrelation property. Therefore, it is necessary to develop new matrices in \( \mathbb{Q}_{q \times N}^{N\times N} \) for efficient 2D-CCS with a small \( q \).

Perfect arrays over \( \mathbb{Q}_{q \times N} \) were constructed for binary and quaternary alphabets, i.e., for \( q = \log_2 2 \) and \( q = \log_2 4 \) in [15] and [16]. In this paper, we consider square arrays, i.e., arrays of size \( N \times N \) for simplicity. We also consider the extreme case of perfect binary arrays, i.e., \( q = 1 \), as such arrays can be implemented in phase shifters of any resolution. For any natural number \( k \), it was shown in [15] that perfect binary arrays over \( \mathbb{Q}_{2^k \times N} \) exist when \( N = 2^k \) or \( N = 3 \cdot 2^k \). Furthermore, a recursive construction to generate such perfect arrays was provided for square arrays and other rectangular configurations. For \( q = 2 \), several perfect arrays for which perfect binary arrays of the same dimension do not exist were proposed in [16]. FALP uses perfect arrays as base matrices for 2D-CCS in planar arrays.

FALP allows the CS algorithm in (21) to achieve the smallest upper bound on the reconstruction error in (22), as \( Z_{\text{max}} = 1 \) and \( Z_{\text{min}} = 1 \) for any perfect array. It can be observed from (22) that the guarantee on the reconstructed channel becomes independent of \( q \), as long as there exists a perfect array in \( \mathbb{Q}_{q \times N}^{N \times N} \). The existence of perfect arrays for a small \( q \) makes FALP a promising solution for channel estimation or beam alignment in low resolution phased arrays.

V. FALP: A FRAMEWORK FOR LOW-COMPLEXITY BEAM ALIGNMENT

We explain how the complexity of CS-based channel estimation can be reduced with FALP by exploiting the partial 2D-DFT nature of the CS problem in (21). Prior work in CS has used the fast Fourier transform (FFT) to accelerate partial 2D-DFT CS problems [13]. In this section, we describe partial 2D-DFT CS and summarize our FFT-based implementation. The definitions introduced in this section will be used to derive the beam alignment probability in Sec. VI.

The masked beamspace matrix \( \mathbf{S} \) can be estimated from \( \mathbf{y} \) in (20) using a standard CS algorithm. We use orthogonal matching pursuit (OMP) [22] for sparse masked beamspace recovery, and discuss about our low-complexity implementation. Nevertheless, similar complexity reduction arguments can be made for other sparse recovery algorithms when they use FALP’s training. Let \( \mathbf{A}_{CS} \in \mathbb{C}^{N \times N} \) be the CS matrix corresponding to the sparse recovery problem in (20). Using (17), the \( m \)th row of \( \mathbf{A}_{CS} \) can be explicitly written as

\[
\mathbf{A}_{CS}(m,:) = (\mathbf{U}_N \mathbf{e}_{c[m]})^T \otimes (\mathbf{U}_N \mathbf{e}_{r[m]})^T.
\]

We define an \( N^2 \times 1 \) vector \( \mathbf{s} = \text{vec}(\mathbf{S}) \) to rewrite (20) in standard form as

\[
\mathbf{y} = \mathbf{A}_{CS} \mathbf{s} + \mathbf{v}.
\]

Each iteration of the OMP to solve for \( \mathbf{s} \) requires computing \( \mathbf{A}_{CS} \mathbf{w} \) and \( \mathbf{A}_{CS}^* \mathbf{d} \) [22]. The variables \( \mathbf{w} \) and \( \mathbf{d} \) represent the sparse signal estimate and the measurement residual in an OMP iteration. The complexity of each OMP iteration for a generic CS matrix is \( \mathcal{O}(MN^2) \), and can be reduced when \( \mathbf{A}_{CS} \) is a partial 2D-DFT CS matrix.

We describe how matrix multiplications in the OMP algorithm corresponding to (20) can be replaced by the 2D-FFT. From (20), it can be observed that the \( m \)th entry of \( \mathbf{A}_{CS} \mathbf{w} \) is \( (\mathbf{A}_{CS} \mathbf{w})_m = (\mathbf{U}_N \mathbf{e}_{c[m]})^T \otimes (\mathbf{U}_N \mathbf{e}_{r[m]})^T \mathbf{w} \). We define a matrix \( \mathbf{W} \) such that \( \mathbf{w} = \text{vec}(\mathbf{W}) \). Using the identity \((\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B}) = \text{vec}(\mathbf{ABC})\), \( (\mathbf{A}_{CS} \mathbf{w})_m \) can be rewritten as \( (\mathbf{A}_{CS} \mathbf{w})_m = e_{c[m]}^T \mathbf{U}_N \mathbf{W} \mathbf{U}_N e_{r[m]} \). Therefore, the vector \( \mathbf{A}_{CS} \mathbf{w} \) is a subsampled version of \( \mathbf{U}_N \mathbf{W} \mathbf{U}_N \) at the locations in \( \Omega \), i.e.,

\[
\mathbf{A}_{CS} \mathbf{w} = \mathbf{P}_\Omega (\mathbf{U}_N \mathbf{W} \mathbf{U}_N).
\]

In CS of approximately sparse signals, the number of non-zero entries in \( \mathbf{w} \) progressively increase with the OMP iteration. Therefore, it is reasonable to compute \( \mathbf{A}_{CS} \mathbf{w} \) using matrix and sparse vector multiplication for the initial OMP iterations. For a sufficiently large OMP iteration, \( \mathbf{A}_{CS} \mathbf{w} \) can be computed using 2D-FFT followed by subsampling.

Now, we explain a low complexity implementation of \( \mathbf{A}_{CS}^* \mathbf{d} \) in OMP. Using (25), the vector \( \mathbf{A}_{CS}^* \mathbf{d} = \sum_{m=1}^M d[m] \mathbf{A}_{CS}^*(m,:) \) can be expressed as

\[
\mathbf{A}_{CS}^* \mathbf{d} = \sum_{m=1}^M d[m] \text{vec}(\mathbf{U}_N^* e_{c[m]} e_{r[m]}^T \mathbf{U}_N^*).
\]

We define an \( N \times N \) matrix \( \mathbf{D}_{\text{adj}} \) that contains the entries of \( \mathbf{d} \) at \( \Omega = \{(r[m], c[m])\}_{m=1}^M \), and zeros at the rest of the locations, i.e.,

\[
\mathbf{D}_{\text{adj}}(r[m], c[m]) = d[m] \quad \text{for} \ 1 \leq m \leq M, \quad \text{and} \quad \mathbf{D}_{\text{adj}}(k, \ell) = 0 \quad \forall (k, \ell) \notin \Omega.
\]

With the definition of \( \mathbf{D}_{\text{adj}} \), (28) can be rewritten as

\[
\mathbf{A}_{CS}^* \mathbf{d} = \sum_{k, \ell \in \mathcal{I}_N} \mathbf{D}_{\text{adj}}(k, \ell) \text{vec}(\mathbf{U}_N^* e_{k} e_{\ell}^T \mathbf{U}_N^*).
\]

It can be observed from (30) that \( \mathbf{A}_{CS}^* \mathbf{d} = \text{vec}(\mathbf{U}_N^* \mathbf{D}_{\text{adj}} \mathbf{U}_N^*) \). Therefore, \( \mathbf{A}_{CS}^* \mathbf{d} \) can be computed using the 2D-FFT.

We provide a summary of the low-complexity OMP to estimate \( \mathbf{S} \) in Algorithm 1. Note that \( \text{supp}(\mathbf{A}) \) represents the support of a matrix \( \mathbf{A} \), i.e., the matrix \( \mathbf{A} \) is 0 at all locations that are not in \( \text{supp}(\mathbf{A}) \). The masked beamspace estimate \( \hat{\mathbf{S}} \) obtained from Algorithm 1 is transformed to obtain the true beamspace estimate \( \hat{\mathbf{X}} \) by inverting the spectral mask. As FALP uses a perfect array as a base matrix, the spectral mask \( \mathbf{Z} \) is unimodular and the true beamspace can be estimated as \( \hat{\mathbf{X}} = \hat{\mathbf{S}} \otimes \mathbf{Z} \). The channel estimate is then \( \mathbf{H} = \mathbf{U}_N \hat{\mathbf{X}} \mathbf{U}_N \). When \( \alpha = \alpha_0 \) for some constant \( \alpha_0 \) < 1, it can be observed that each iteration in the low-complexity OMP implementation has a complexity of \( \mathcal{O}(N^4) \), while the standard implementation requires a complexity of \( \mathcal{O}(N^6) \).
VI. FALP DEMYSTIFIED WITH CS IN MRI

One of the early applications of CS was magnetic resonance imaging using subsampled measurements. For example, as natural angiogram images are sparse in their canonical representation, CS can reconstruct them from fewer samples of their 2D-DFT \[17\]. In this section, we establish an equivalence between MR imaging and mmWave channel estimation with FALP. Furthermore, we derive lower bounds on the probability of beam alignment using ideas from MRI.

FALP estimates the masked beamspace \( \mathbf{S} \) from random subsamples of \( \mathbf{G} \), the 2D-DFT of \( \mathbf{S} \). For subsampling determined by the set \( \Omega = \{ [r|m], [c|m] \}_{m=1}^M \), it can be observed from \[19\] that FALP acquires samples of the virtual channel \( \mathbf{G} \) at the locations in \( \Omega \). We define a trajectory \[10\] as a curve that sequentially traverses through the coordinates \( \{ [r|m], [c|m] \}_{m=1}^M \) of \( \mathbf{G} \). An example of a trajectory for \( M = 9 \) and \( N = 5 \) is shown in Fig. 2a. In the context of angiogram virtual channel as any circumstantial shift of a base matrix can be applied to the phased array.

A. One and done: Zero filling in MRI meets single step OMP

In this section, we use ideas from zero filling reconstruction in MRI \[17\] to show that a single iteration of Algorithm 1 can provide near optimal beamforming gain for sub-Nyquist sampling of sufficiently sparse channels. Our result is useful as CS-based algorithms that often require iterative optimization can be avoided. Prior work has considered single iteration OMP for CS using random phase shift-based CS matrices \[23\]. The guarantees on the beam alignment probability, however, were not derived in \[23\]. The connection with MRI allows us to study CS-based beam alignment in FALP as a function of subsampling, and derive analytical guarantees on the beam alignment probability under certain conditions. To provide a better illustration and for a tractable analysis, we ignore the measurement noise in the system, i.e., \( \sigma = 0 \).

Now, we show that subsampling the virtual channel \( \mathbf{G} \) results in convolutional distortion in \( \mathbf{U}_N^* \mathbf{D}_\text{adj}^{(0)} \mathbf{U}_N^* \) i.e., the matrix constructed in the first iteration of Algorithm 1. The matrix \( \mathbf{D}_\text{adj}^{(0)} \) contains the entries of \( \mathbf{y} \) in \( \Omega \) and zeros in the rest of the locations. To study the impact of subsampling on the masked beamspace estimate, we define a binary matrix \( \mathbf{N}_\Omega \in \mathbb{C}^{N \times N} \) such that

\[
\mathbf{N}_\Omega(r, c) = \begin{cases} 
1, & \text{if } (r, c) \in \Omega \\
0, & \text{if } (r, c) \notin \Omega 
\end{cases}.
\]

For single step OMP, i.e., Algorithm 1 for \( T_{\text{iter}} = 1 \), it can be observed that FALP uses the coordinate where \( |\mathbf{U}_N^* \mathbf{D}_\text{adj}^{(0)} \mathbf{U}_N^*| \) achieves its maximum for beam alignment. In contrast, beam alignment using exhaustive scan over the 2D-DFT dictionary uses the coordinate where \( |\mathbf{X}| \) achieves its maximum \[23\]. We define a kernel matrix \( \mathbf{K}_\Omega \) as the scaled inverse 2D-DFT of \( \mathbf{N}_\Omega \), i.e., \( \mathbf{K}_\Omega = N \mathbf{U}_N^* \mathbf{N}_\Omega \mathbf{U}_N^*/M \). For \( \sigma = 0 \), it can be observed that \( \mathbf{y} = \mathcal{P}_\Omega(\mathbf{G}) \) and \( \mathbf{D}_\text{adj}^{(0)} = \mathbf{G} \odot \mathbf{N}_\Omega \). For single step OMP, i.e., Algorithm 1, the zero fill of \( \mathbf{F}_\Omega \) can be observed that FALP uses the coordinate where \( |\mathbf{U}_N^* \mathbf{D}_\text{adj}^{(0)} \mathbf{U}_N^*| \) achieves its maximum for beam alignment. In contrast, beam alignment via single step OMP is successful when the coordinate that maximizes \( |\mathbf{U}_N^* \mathbf{D}_\text{adj}^{(0)} \mathbf{U}_N^*| \) also maximizes \( |\mathbf{S}| \).

As element-wise multiplication of two matrices results in 2D-circular convolution of their inverse 2D-DFTs \[18\], we can write

\[
U_N^* D^{(0)}_{\text{adj}} U_N^* = (U_N^* G U_N^*) \odot (U_N^* N_\Omega U_N^*)/N \tag{34}
\]

The MRI analogue of \( U_N^* D^{(0)}_{\text{adj}} U_N^* \) is the zero filling reconstruction-based estimate of the sparse MR image \[17\]. It can be observed from \[35\] that the matrices \( U_N^* D^{(0)}_{\text{adj}} U_N^* \) and \( \mathbf{S} \) differ by a convolutional distortion due to \( \mathbf{K}_\Omega \). For full sampling, i.e., \( M = N^2 \), we have \( \mathbf{N}_\Omega = 1_{N \times N} \), \( \mathbf{K}_\Omega = \mathbf{e}_0 \mathbf{e}_0^T \), and \( U_N^* D^{(0)}_{\text{adj}} U_N^* = \mathbf{S} \). Hence, single step OMP achieves the performance of exhaustive search when \( M = N^2 \). For the subsampled case, i.e., \( M < N^2 \), single step OMP succeeds only if the coordinate that maximizes \( |\mathbf{S} \odot \mathbf{K}_\Omega| \) is the same as the one that maximizes \( |\mathbf{S}| \).
We use ideas from CS in MRI to illustrate the impact of random subsampling on beam alignment. With the definition of 2D circular convolution, (35) can be expressed as [18]

$$\mathbf{S} \otimes \mathbf{K}_\Omega = \sum_{k,\ell} \mathbf{S}(k,\ell) \mathbf{J}^*_k \mathbf{K}_\Omega^{1,\ell}. \quad (36)$$

It can be observed from (36) that the $(k,\ell)$ entry in $\mathbf{S}$ contributes or “spreads” to all the entries in $\mathbf{S} \otimes \mathbf{K}_\Omega$ through the kernel matrix $\mathbf{K}_\Omega$. The matrix $\mathbf{K}_\Omega$ is called the point spread function (PSF) in MRI literature [17]. We define the subsampling ratio in FALP as $\rho = M/N^2$. As $\Omega$ has $M$ coordinates, the binary matrix $\mathbf{N}_\Omega$ has $M$ ones and $N^2 - M$ zeros. Therefore, $\mathbf{K}_\Omega(0,0)$, the scaled DC-component of $\mathbf{N}_\Omega$, is 1. The other entries of $\mathbf{K}_\Omega$ explicitly depend on the elements in the sampled set $\Omega$ unlike $\mathbf{K}_\Omega(0,0)$. As $\Omega$ is sampled at random, $\mathbf{K}_\Omega(r,c)$ can be modelled as a random variable for any $(r,c) \neq (0,0)$ with a variance [17]

$$\xi^2 = 1 - \rho / \rho N^2. \quad (37)$$

Note that the variance of the PSF at the $N^2 - 1$ locations other than $(0,0)$ is exactly the same when $\Omega$ is chosen uniformly at random. The magnitude of the PSF, i.e., $|\mathbf{K}_\Omega|$, is shown in Fig. 2b for $N = 32$ and $\rho = 1/16$. For a two-sparse channel in Fig. 3a, the matrix $\mathbf{S} \otimes \mathbf{K}_\Omega$ is shown in Fig. 3b. It can be observed that $\mathbf{S} \otimes \mathbf{K}_\Omega$ is a distorted version of $\mathbf{S}$, and the distortion is determined by the undersampling factor $\rho$.

![Distortion of PSF](image)

(a) A two-sparse $\mathbf{S}$ (b) $\mathbf{S} \otimes \mathbf{K}_\Omega$ for $\mathbf{K}_\Omega$ in Fig. 2b

Fig. 3: The distortion in $\mathbf{S} \otimes \mathbf{K}_\Omega$ depends on the subsampling ratio $\rho$. From Fig. 3a and Fig. 3b, it can be observed that single step OMP can find the best direction for $\rho = 1/16$ and $N = 32$.

### B. Is single step OMP good enough for beam alignment?

Now, we derive a lower bound on the probability of successful beam alignment via single step OMP as a function of the subsampling ratio $\rho$. For simplicity of analysis, we consider a 2-path channel such that the beamspace angle of departure of each path is aligned with the 2D-DFT dictionary. In this case, the beamspace channel $\mathbf{X}$ is 2-sparse and the masked beamspace channel $\mathbf{S}$ in FALP is also 2-sparse as $\mathbf{P}$ is unimodular. Without loss of generality, we consider $\mathbf{S}(0,0) = 1$, and $\mathbf{S}(r_o,c_o) = a$ such that $|a| < 1$ and $(r_o,c_o)$ is some point other than $(0,0)$. The remaining $N^2 - 2$ entries of $\mathbf{S}$ are equal to 0 as seen in Fig. 3a. For the setting considered in our analysis, beam alignment is successful when $|\mathbf{S} \otimes \mathbf{K}_\Omega(0,0)|$ is the largest entry in $\mathbf{S} \otimes \mathbf{K}_\Omega$.

The statistics of the kernel matrix $\mathbf{K}_\Omega$ can be used to determine the entries in $\mathbf{S} \otimes \mathbf{K}_\Omega$. Prior work in MRI [17] and partial 2D-DFT CS [24] has modelled $\mathbf{K}_\Omega(r,c)$ as $\mathcal{N}_C(0,\xi^2)$ for any $(r,c) \neq (0,0)$. The random variables $\mathbf{K}_\Omega(r_1,c_1)$ and $\mathbf{K}_\Omega(r_2,c_2)$ can be dependent for $(r_1,c_1) \neq (0,0)$ and $(r_2,c_2) \neq (0,0)$. For instance, as $\mathbf{N}_\Omega$ is a real matrix, its inverse 2D-DFT must be conjugate symmetric [18], i.e., $\mathbf{K}_\Omega(N-r,N-c) = \mathbf{K}_\Omega(r,c)$ for any $(r,c) \in \mathbb{I}_N \times \mathbb{I}_N$.

B. Is single step OMP good enough for beam alignment?

Now, we derive a lower bound on the probability of successful beam alignment for the 2-sparse channel. The probability of beam alignment via single step OMP can be expressed as

$$p = \Pr((r,c) \cap |\mathbf{S} \otimes \mathbf{K}_\Omega| > |\mathbf{S} \otimes \mathbf{K}_\Omega(0,0)|). \quad (38)$$

We define $Q_1(\alpha, \beta)$ as the first order Marcum-Q function with parameters $\alpha$ and $\beta$ [25]. A lower bound on the beam alignment probability in (38) is derived in Theorem 3.

**Theorem 3.** For a 2-sparse beamspace channel, the probability of successful beam alignment using single step OMP can be lower bounded as

$$p \geq 1 - Q_1(0, N \sqrt{2\rho/(1-\rho)}) - \frac{1}{1+2|a|^2} \exp \left( \frac{-N^2 \rho}{(1+2|a|^2)(1-\rho)} \right). \quad (39)$$

**Proof.** See Section X-B
Fig. [4]. The phase transition plots indicate that single step OMP can perform beam alignment with sub-Nyquist channel measurements. It can also be observed that the number of channel measurements required for successful beam alignment increase with the strength of the second best path, i.e., [\alpha]. A larger number of channel measurements reduces the variance of the entries in \( K_\alpha \), i.e., \( \xi^2 \), and mitigates the beamspace interference that arises due to sub-Nyquist sampling.

Fig. 4: For a 2-sparse channel, the plot shows the beam alignment probability as a function of the strength of second best path and the number of channel measurements. Here, the undersampling ratio \( \rho = M/1024 \). The strength of the second best path is 20 log_{10}(|\alpha|).

From a signal recovery perspective, zero filling reconstruction results in noisy MR images for \( \rho < 1 \). For beam alignment, however, the noisy beamspace matrix for \( \rho < 1 \) can be good enough. For instance, the best direction for transmit beam alignment can still be determined from the noisy beamspace matrix in Fig. [3]. The connection between beam alignment using FALP and MRI opens an interesting research direction in robust compressive beamforming. It can be observed that the CS matrix in FALP is determined by the trajectory for a given base matrix. Therefore, k-space trajectories in MRI can be used to design CS matrices in mmWave systems that achieve robustness to non-idealities like CFO, phase noise, and frame synchronization errors.

VII. BEAM ALIGNMENT WITH FALP

For a given channel estimate \( \hat{H} \), beam alignment can be performed with a matrix \( F \in \mathbb{Q}_q^{N\times N} \) such that \(|\langle \hat{H}, F \rangle|\) is maximum. In practice, searching for such a matrix in \( \mathbb{Q}_q^{N\times N} \) is not feasible for large antenna arrays. As a compromise, we seek a suboptimal procedure that achieves a reasonable beamforming gain using \( \hat{H} \).

The first step of our procedure finds an \( F \in \mathbb{Q}_q^{N\times N} \) that maximizes \(|\langle \hat{H}, F \rangle|\). Note that the phased array implementation requires \(|F_{i,j}| = 1/N\) for every \( i \) and \( j \). By the dual norm inequality [26], we have \(|\langle \hat{H}, F \rangle| \leq \max(|F|)\|\hat{H}\|_1\). Therefore, \(|\langle \hat{H}, F \rangle| \leq \|\hat{H}\|_1/N\). Furthermore, the upper bound in the dual norm inequality is achieved by an \( F_{\text{opt}}(\beta) \) such that \( F_{\text{opt}}(\beta)_{i,j} = 1/N \) and \( \phi_F(\hat{H}, F_{\text{opt}}(\beta)) = \beta + \phi_H(\hat{H}_{i,j}) \) for any \( \beta \in (\pi, \pi] \). The scalar \( \beta \) corresponds to the global phase in \( F_{\text{opt}}(\beta) \). As the angles in \( F_{\text{opt}}(\beta) \) may not be integer multiples of \( 2\pi/2^q \), \( F_{\text{opt}}(\beta) \) may not be directly realized in \( q \)-bit phased arrays. In such case, a \( q \)-bit phase quantized version of \( F_{\text{opt}}(\beta) \) can be used in the phased array, for an appropriate choice of global phase \( \beta \).

The second step of our procedure finds the best \( \beta \) that minimizes phase errors due to \( q \)-bit phase quantization of \( F_{\text{opt}}(\beta) \). This step is important in low resolution phased arrays [27], [28]. Let \( Q_q(\beta) \) denote the \( q \)-bit phase quantized version of \( F_{\text{opt}}(\beta) \). Note that \( Q_q(\cdot) \) performs element-wise phase quantization. The global phase term \( \beta_{\text{est}} \) that minimizes the phase quantization error can be expressed as

\[
\beta_{\text{est}} = \arg\min_{\beta \in (0, 2\pi/2^q)} \|Q_q(F_{\text{opt}}(\beta)) - F_{\text{opt}}(\beta)\|_F. \tag{40}\]

To solve for \( \beta_{\text{est}} \) in (40), we define a phase set \( B \) that contains \( K_B \) uniformly spaced values in \((0, 2\pi/2^q)\). The optimization in (40) is performed using line search over the elements in \( B \) for a sufficiently large \( K_B \). The matrix used for transmit beamforming is then \( F_{BA} = Q_q(F_{\text{opt}}(\beta_{\text{est}})) \).

Now, we explain how FALP can be used for beam alignment in wideband mmWave systems using a Galay sequence-based frame structure [1]. For an elaborate description of the wideband extension, we refer the reader to [29]. We consider an \( L \) tap wideband channel \( \{H[l]\}_{l=0}^{L-1} \), where \( H[l] \in \mathbb{C}^{N \times N} \). For each phase shift configuration in \( \{P[m]\}_{n=1}^{M} \), a Galay complementary sequence of length \( 2N_s \) is transmitted. The use of guard interval prevents inter-frame interference and allow sufficient time to configure the phase shifters [8]. The RX uses the perfect autocorrelation property of complementary Galay sequences to obtain the channel impulse response (CIR) for each phase shift configuration. Specifically, the CIR obtained by using the \( m \)-th phase shift configuration is a noisy version of \( \{\langle H[l], P[m]\rangle\}_{l=0}^{L-1} \). Using several spatial channel projections, it is possible to reconstruct the wideband channel. We define \( Y_{\text{blk}} \in \mathbb{C}^{M \times L} \) as a matrix that contains noisy wideband channel projections. The noise in \( Y_{\text{blk}} \) is modelled using \( V_s \in \mathbb{C}^{M \times L}; Y_{\text{blk}}(m, \ell) = \) then

\[
Y_{\text{blk}}(m, \ell) = \|\langle H[l], P[m]\rangle + V_s(m, \ell)\|_2. \tag{41}\]

As a spreading gain of \( 2N_s \) is achieved at the output of the RX correlator, it can be observed that the entries in \( V_s \) are independent and identically distributed as \( N_c(0, \sigma^2/(2N_s)) \).

In FALP, a single tap of the wideband channel is used to determine the best beam. Nevertheless, CS-based wideband channel estimation can also be performed at the expense of a higher complexity [7], [6]. The tap used to perform beam alignment is given by \( \ell_o = \arg\max_{\ell} \|\sum_{i}\|Y_{\text{blk}}(:, \ell)\|_2^2 \). The channel measurements considered in FALP are compressive spatial projections of \( H[\ell_o] \), i.e., \( y = Y_{\text{blk}}(:, \ell_o) \). We use \( H[\ell_{\text{opt}}] \) to denote the channel tap that has the maximum energy of the \( L \) taps. In practice, \( H[\ell_o] \) can be different from \( H[\ell_{\text{opt}}] \) as \( \ell_o \) is determined from lower-dimensional spatial projections of the \( L \) channel taps. The matrix \( H[\ell_o] \) obtained using \( y \) can be considered as an equivalent narrowband channel estimate that is used for beam alignment. Note that our approach ignores the correlation between channel taps as it performs beam alignment based on a single tap. Developing better beam
alignment strategies that account for such correlations is an interesting direction for future work.

VIII. Simulations

In this section, we consider a practical wideband mmWave setting, and study CS-based beam alignment in FALP in terms of the achievable rate and computational complexity. We consider an analog beamforming system in Fig. 1 where the TX is equipped with a UPA of size $32 \times 32$, i.e., $N = 32$. The RX acquires channel measurements using a fixed quasi-omnidirectional pattern throughout beam training. We consider a mmWave carrier frequency of 28 GHz and an operating bandwidth of 100 MHz, which corresponds to a symbol duration of 10 ns. The mmWave channel in our simulations was derived from the QuaDriga channel simulator for the 3GPP 38.901 UMi-NLOS scenario [30]. The height of the TX and the RX were 5 m and 2 m in our simulation setup. For a uniform distribution of receivers within a distance of 100 m from the TX, the omnidirectional RMS delay spread was found to be less than 165 ns in more than 90% of the channel realizations. Considering the leakage effects due to pulse shaping, the wideband channel is modelled using $L = 64$ taps corresponding to a duration of 640 ns. The transmit power at the TX is assumed to be 20 dBm.

Now, we explain channel acquisition using different CS-based training solutions. The channel measurements for CS-based beam alignment are acquired using Golay complementary sequences, where each sequence is of length $N_s = 64$. For our simulation settings, it can be observed that the duration of the guard interval that follows a Golay pair is 630 ns. The guard interval is sufficient enough as phase shifters with a settling time of about 30 ns at 28 GHz have been reported in [51].

We compare FALP’s training with the commonly used random IID phase shift-based training and random 2D-CCS-based training, i.e., 2D-CCS using a random base matrix. For the IID phase shift training, each of the phase shift matrices in the frame, i.e., $\{P[m]_{n=1}^{M}\}$, are chosen uniformly at random from $\mathbb{Q}_q^{N \times N}$. In random 2D-CCS, $\{P[m]_{n=1}^{M}\}$ are chosen as random 2D-circular shifts of a base matrix that is chosen at random from $\mathbb{Q}_q^{N \times N}$. In FALP, the matrices $\{P[m]_{n=1}^{M}\}$ are chosen as random 2D-circularly shifted versions of a perfect array.

Using the procedure in Sec. VII the equivalent narrowband channel that corresponds to the tap with “maximum” energy is estimated for beam alignment using CS. The phase shift matrix for transmit beam alignment, i.e., $F_{BA}$, is derived from the equivalent narrowband channel estimate using the procedure in Sec. VII. In this case, the effective single-input single-output (SISO) channel seen after beam alignment is $\{\langle H[\ell], F_{BA}\rangle\}_{\ell=0}^{L-1}$. The training solutions and algorithms are evaluated in terms of the achievable rate of the effective SISO channel. The simulation results we report are the averages over 100 channel realizations taken at each TX-RX separation.

For FALP’s perfect array-based training, CS-based beam alignment is performed using Algorithm I for $T_{iter} = 1$ and $T_{iter} = 200$. The stopping threshold, i.e., $\epsilon$ in Algorithm I, is set as $\sigma_{\sqrt{M}/2N_s}$. For IID phase shift-based training and random 2D-CCS-based training, the standard OMP algorithm $\|\|_1$ is evaluated for $T_{iter} = 1$ and $T_{iter} = 200$ with the same $\epsilon$. The same 2D-DFT dictionary was used for CS using each of the three training solutions for a fair comparison. The single step OMP algorithm that uses $T_{iter} = 1$ is compared with exhaustive beam search using the 2D-DFT dictionary. It can be observed from Fig. 5 and Fig. 6 that single step OMP achieves performance comparable to exhaustive beam search, for sub-Nyquist sampling. Note that the performance evaluation in Fig. 5 and Fig. 6 considered a phased array architecture with 1-bit phase shifters. The number of candidate global phase values in $B$ was chosen as $K_B = 6$. OMP with $T_{iter} = 200$ obtains a better approximation of the sparse channel when compared to single step OMP at the expense of higher computational complexity. As a result, OMP with $T_{iter} = 200$ achieves better beam alignment performance than exhaustive scan over the 2D-DFT dictionary. The loss in achievable rate when compared to the perfect CSI case is due to noise in the channel measurements and leakage effects.
in the beamspace representation.

As seen in Fig. [7] the achievable rate increases with the resolution of the phase shifters. An interesting observation from Fig. [7] is that FALP’s one-bit training achieves better beam alignment than the random training solutions for any resolution of phase shifters. Such property follows from the resolution independent guarantee in Theorem 1 for CS using FALP. From an implementation perspective, CS using FALP’s training has a lower computational complexity than CS using random phase shift-based training. In Fig. 8 we compare the execution time of single step OMP using different training solutions. The low complexity nature and guaranteed beam alignment or channel estimation using FALP motivate its use for real time mmWave radios.

IX. CONCLUSIONS AND FUTURE WORK

In this paper, we have proposed FALP, a framework for compressive beam alignment or channel estimation using a perfect array-based codebook. The existence of perfect arrays over small alphabets allows efficient compressed sensing in low resolution phased arrays through FALP. We have derived guarantees on channel reconstruction from sub-Nyquist sampling using FALP. In addition, we have also derived a lower bound on the beam alignment probability using our approach by establishing the equivalence between CS in FALP and CS in MRI.

FALP establishes a new platform to translate CS ideas from MRI to channel estimation or beam alignment in mmWave systems. For example, deterministic k-space trajectories in MRI can be used in our framework instead of random subsampling trajectories. We have shown how image recovery techniques that are as simple as zero-filling reconstruction can be used for rapid and low complexity beam alignment. In our future work, we will derive insights from MRI to develop sub-sampling trajectories that achieve robustness to hardware impairments.

X. PROOF OF THEOREMS

A. Proof of Theorem 1

For the conditions in Theorem 1 the reconstruction error in the masked beamspace obtained using the ℓ1-norm optimization program in (21) can be bounded as

$$
\|S - \hat{S}\|_F \leq C_1 \frac{\|S - (S)_k\|_1}{\sqrt{k}} + C_2 N\sigma.
$$

(42)

The upper bound on the reconstruction error in (42) follows from Theorems 1 and 3 of [13]. Using the spectral mask equation, i.e., $S = X \odot Z$, and (42), we translate the guarantee on $S$ to the true beamspace estimate, i.e., $\hat{X}$.

We first obtain an upper bound on the ℓ1 approximation error of the masked beamspace in (42). We define $\Gamma \subseteq \mathcal{I}_N \times \mathcal{I}_N$ and its complement as $\Gamma^c$. The cardinality of $\Gamma$ is denoted by $|\Gamma|$. With the definition of $N_{\Omega}$ in (33), we express the sparse approximation error $\|S - (S)_k\|_1$ as

$$
\|S - (S)_k\|_1 = \min_{\Gamma, |\Gamma|=k} \|S - (S)_{\Omega}\|_1
$$

(43)

$$
= \min_{\Gamma, |\Gamma|=k} \sum_{(\ell,m) \in \Gamma} |S(\ell,m)|.
$$

(44)
From the spectral mask equation, we have $S(\ell, m) = X(\ell, m)Z(\ell, m)$. As a result, $|S(\ell, m)| \leq Z_{\text{max}}|X(\ell, m)|$. The $\ell_1$ approximation error in (44) is upper bounded as

$$\|S - (S)\|_1 \leq Z_{\text{max}} \min_{\ell, m, \ell', m'} \sum_{(\ell, m) \in \Gamma'} |X(\ell, m)|.$$  (45)

Using the definition of the $\ell_1$ error in a $k$ sparse approximation of $X$ in (45), we have $|S - (S)\|_1 \leq Z_{\text{max}} \|X - (X)\|_1$.

Now, we derive a lower bound on the error in the true beamspace estimate $\hat{S}$. The squared error in $\hat{S}$ is $\|S - \hat{S}\|_F^2 = \sum_{\ell, m} |Z(\ell, m)|^2 (X(\ell, m) - \hat{X}(\ell, m))^2$. By definition, $|Z(\ell, m)| \geq Z_{\text{min}}$ for every $\ell$ and $m$. Therefore, the error in the beamspace estimate lower bounded as $\|S - \hat{S}\|_F \geq Z_{\text{min}} \|X - \hat{X}\|_F$. The result in Theorem 1 follows by using $\|S - (S)\|_1 \leq Z_{\text{max}} \|X - (X)\|_1$ and $\|S - \hat{S}\|_F \geq Z_{\text{min}} \|X - \hat{X}\|_F$ in (42).

**B. Proof of Theorem 2**

From (38), the probability that $|S \circ K_0|$ achieves maximum at $(0, 0)$ can be expressed as

$$p = 1 - \Pr \left( (r, c, 0, 0) \in \{(r, c, 0, 0), (0, 0, r, c) \} \right).$$  (46)

Note that $S$ is non-zero only at the locations $(0, 0)$ and $(r_o, c_o)$. In this section, we derive closed form expressions for $\Pr((S \circ K_0)_{0, 0} \leq |(S \circ K_0)_{r, c}|)$ and $\Pr((S \circ K_0)_{0, 0} \leq |(S \circ K_0)_{r, c}|)$ for some $(r_1, c_1) \notin \{(0, 0), (r_o, c_o)\}$. We then derive a lower bound on $p$.

The matrix $S$ is 1 at $(0, 0)$, $|a|$ at $(r_o, c_o)$, and 0 at the remaining $N^2 - 2$ locations. Although $|S|$ is maximum at $(0, 0)$ for $|a| < 1$, it is possible that $|S \circ K_0|$ achieves maximum at $(r_o, c_o)$. We define $p_1 = \Pr((S \circ K_0)_{0, 0} \leq |(S \circ K_0)_{r, c}|)$. Using (46), $p_1$ can be expressed as

$$p_1 = \Pr((1 + |a|^* \leq |a + x|).$$  (47)

The inequality $|1 + |a|^* \leq |a + x|$, is equivalent to $1 + |a|^2|\xi|^2 + 2Re(\xi^*ax) - |a|^2|\xi|^2 + 2Re(\xi^*ax)$ can be simplified to $(1 - |a|^2)|\xi|^2 \geq 1 - |a|^2$. For any $|a| < 1$, $p_1$ is given by

$$p_1 = \Pr(|\xi|^2 \geq 1).$$  (48)

As $\chi \sim \chi^2_c(0, \xi^2)$ for a large $N$, $2|\xi|^2$ follows the central $\chi^2$-distribution of order 2. Therefore, $p_1$ in (49) can be expressed in terms of the Marcum Q-function $[29]$ as

$$p_1 = Q_{1}(0, \sqrt{2}|\xi|).$$  (50)

An interesting observation from (50) is that $p_1$ is independent of the strength of the second best path, i.e., $|a|$.

Now, we derive the probability that $|(S \circ K_0)_{0, 0}| \leq |(S \circ K_0)_{r, c}|$ for some $(r_1, c_1)$ such that $S(r_1, c_1) = 0$. We define $p_2 = \Pr(|(S \circ K_0)_{0, 0}| \leq |(S \circ K_0)_{r, c}|)$. Using (56), $p_2$ can be expressed as

$$p_2 = \Pr(|1 + |a|^* \leq |b + aw|).$$  (51)

We assume that the variables $x$, $b$ and $w$ are independent. With this assumption, it can be observed that $b + aw \sim \mathcal{N}_c(0, 1 + |a|^2\xi^2)$ and $1 + ax^* \sim \mathcal{N}_c(1, |a|^2\xi^2)$. We use $\chi^2_c$ and $\chi^2_{NC}$ to denote the central and non-central chi-squared random variables of degree 2 [25]. The non-centrality parameter of $\chi^2_{NC}$ is set as $\lambda_{NC}^2 = 2|a|^2\xi^2$. Using these definitions, it can be shown that $|b + aw|^2 \sim \chi^2_1(1 + |a|^2\xi^2)$ and $1 + |a|^2 \sim |a|^2\xi^2\chi^2_{NC}/2$. We use $f(t)$ to denote the probability density of $\chi^2_{NC}$ at $t$. The probability in (51) is then

$$p_2 = \Pr \left( \chi^2_c \geq \frac{|a|^2}{1 + |a|^2\xi^2} \right)$$

$$= \int_{0}^{\infty} \Pr \left( \chi^2_c \geq \frac{|a|^2}{1 + |a|^2\xi^2} \right) f(t) dt.$$  (52)

We use the complementary cumulative distribution function of $\chi^2_c$ to express (53) as

$$p_2 = \int_{0}^{\infty} \exp \left( \frac{-|a|^2t}{2(1 + |a|^2\xi^2)} \right) f(t) dt.$$  (54)

It can be observed from (54) that $p_2$ is the moment generating function of $\chi^2_{NC}$ [25] evaluated at $-|a|^2/(2 + 2|a|^2)$. The probability in (54) is then

$$p_2 = 1 + |a|^2/2 \exp \left( \frac{-1}{1 + 2|a|^2\xi^2} \right).$$  (55)

As a sanity check, it can be observed that $p_2 = p_1$ for $a = 0$.

We use $p_1$ and $p_2$ in (50) and (55), to derive a lower bound on $p$. The probability that $|S \circ K_0|$ does not achieve its maximum at $(0, 0)$ can be upper bounded using a union bound as

$$1 - p \leq \sum_{(r, c) \notin \{(0, 0)\}} \Pr((S \circ K_0)_{0, 0} \leq |(S \circ K_0)_{r, c}|).$$  (56)

The right hand side in (56) comprises of two distinct terms that are $p_1 = \Pr((S \circ K_0)_{0, 0} \leq |(S \circ K_0)_{r, c}|)$, and $p_2 = \Pr((S \circ K_0)_{0, 0} \leq |(S \circ K_0)_{r, c}|)$ for any $(r_1, c_1) \notin \{(0, 0), (r_o, c_o)\}$. As there are $N^2 - 2$ terms of the second kind, we can write

$$1 - p \leq p_1 + (N^2 - 2)p_2.$$  (57)

The result in Theorem 2 follows by substituting (50) and (55) in (57).
[8] 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.

[9] E. J. Candès, “The restricted isometry property and its implications for compressed sensing,” Comptes rendus mathematique, vol. 346, no. 9-10, pp. 589–592, 2008.

[10] N. J. Myers, A. Mezghani, and R. W. Heath, “Swift-Link: A compressive beam alignment algorithm for practical mmWave radios,” IEEE Trans. on Signal Process., vol. 67, no. 4, pp. 1104–1119, 2019.

[11] C.-R. Tsai and A.-Y. Wu, “Structured random compressed channel sensing for millimeter-wave large-scale antenna systems,” IEEE Trans. on Signal Process., vol. 66, no. 19, pp. 5096–5110, 2018.

[12] 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.

[13] F. Krahmer and H. Rauhut, “Structured random measurements in signal processing,” GAMM-Mitteilungen, vol. 37, no. 2, pp. 217–238, 2014.

[14] H. D. Luke, H. D. Schotten, and H. Hadinejad-Mahram, “Binary and quardiphrase sequences with optimal autocorrelation properties: A survey,” IEEE Trans. on Inform. Theory, vol. 49, no. 12, pp. 3271–3282, 2003.

[15] J. Jedwab, C. Mitchell, F. Piper, and P. Wild, “Perfect binary arrays and difference sets,” Discrete Mathematics, vol. 125, no. 1-3, pp. 241–254, 1994.

[16] K. Arasu and W. de Launey, “Two-dimensional perfect quaternary arrays,” IEEE Trans. on Inform. Theory, vol. 47, no. 4, pp. 1482–1493, 2001.

[17] M. Lustig, D. Donoho, and J. M. Pauly, “Sparse MRI: The application of compressed sensing for rapid MR imaging,” Magnetic Resonance in Medicine, vol. 58, no. 6, pp. 1182–1195, 2007.

[18] A. C. Kak and M. Slaney, Principles of computerized tomographic imaging. IEEE press, 1998.

[19] J. Brady, N. Behdad, and A. M. Sayeed, “Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis, and measurements,” IEEE Trans. on Antennas and Propagation, vol. 61, no. 7, pp. 3814–3827, 2013.

[20] F. Ong, S. Pawar, and K. Ramchandran, “Fast and efficient sparse 2D discrete Fourier transform using sparse-graph codes,” arXiv preprint arXiv:1509.05849, 2015.

[21] H. D. Luke, “Sequences and arrays with perfect periodic correlation,” IEEE Trans. on Aero. and Elec. Sys., vol. 24, no. 3, pp. 287–294, 1988.

[22] J. A. Tropp and A. C. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” IEEE Trans. on Inform. Theory, vol. 53, no. 12, pp. 4655–4666, 2007.

[23] 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, 2018.

[24] M. Abramowitz and I. A. Stegun, Handbook of mathematical functions: with formulas, graphs, and mathematical tables. Courier Corporation, 1965, vol. 55.

[25] S. Boyd and L. Vandenberghe, Convex optimization. Cambridge university press, 2004.

[26] K. Ballüder and M. R. Taghizadeh, “Optimized phase quantization for diffractive elements by use of a bias phase,” Optics letters, vol. 24, no. 23, pp. 1756–1758, 1999.

[27] Z. Wang, M. Li, Q. Liu, and A. L. Swindlehurst, “Hybrid precoder and combiner design with low-resolution phase shifters in mmwave MIMO systems,” IEEE J. Sel. Topics Signal Process., vol. 12, no. 2, pp. 256–269, 2018.

[28] N. J. Myers, A. Mezghani, and R. W. Heath, “Spatial Zadoff-Chu modulation for rapid beam alignment in mmwave phased arrays,” in Proc. of the IEEE Global Proc. on Telegraph. Conf. (GLOBECOM), 2018.

[29] M. Jaekel, L. Raschkowski, K. Börner, and L. Thiele, “QuaDRiGa: A 3-D multi-cell channel model with time evolution for enabling virtual field trials,” IEEE Trans. on Antennas and Propag., vol. 62, no. 6, pp. 3242–3256, 2014.

[30] M. E. Leinonen, G. Destino, O. Kursu, M. Sonkki, and A. Pääsiniemi, “28 GHz wireless backhaul transceiver characterization and radio link budget,” ETRI Journal, vol. 40, no. 1, 2018.