

A compressive channel estimation technique robust to synchronization impairments

Nitin Jonathan Myers and Robert W. Heath Jr.
Department of Electrical and Computer Engineering
The University of Texas at Austin
Email: {nitinjmyers, rheath}@utexas.edu

Abstract—Initial access at millimeter wave frequencies is a challenging problem due to hardware non-idealities and low SNR measurements prior to beamforming. Prior work has exploited the observation that mmWave MIMO channels are sparse in the spatial angle domain and has used compressed sensing based algorithms for channel estimation. Most of them, however, ignore hardware impairments like carrier frequency offset and phase noise, and fail to perform well when such impairments are considered. In this paper, we develop a compressive channel estimation algorithm for narrowband mmWave systems, which is robust to such non idealities. We address this problem by constructing a tensor that models both the mmWave channel and CFO, and estimate the tensor while still exploiting the sparsity of the mmWave channel. Simulation results show that under the same settings, our method performs better than comparable algorithms that are robust to phase errors.

Index Terms—Millimeter wave channel estimation, tensor compressed sensing, analog beamforming, channel estimation

I. INTRODUCTION

Millimeter wave (mmWave) communication is a potential candidate for 5G systems due to the enormous amount of spectrum available at mmWave frequencies [1]. Such systems are likely to employ large antenna arrays and use highly directional beamforming to provide sufficient received signal power [2]. Channel estimation at mmWave, however, is challenging as the best beams have to be estimated using low SNR measurements [2]. A naive way to estimate the best beams would be to perform an exhaustive beam search at the transmitter (TX) and receiver (RX) to operate with the beam pair corresponding to the maximum received SNR. This method, however, incurs a lot of training overhead.

Unlike the lower frequency systems, MIMO channels at mmWave are marginally sparse in the spatial angle domain due to clustering in the propagation environment [2]. The sparse nature of mmWave MIMO channels has been exploited to estimate the channel using compressed sensing (CS) [3][4][5]. Most of the existing methods, however, assume an ideal CS measurement model and fail to perform well in the presence of hardware impairments like phase noise, carrier frequency offset (CFO) etc. Although the exhaustive beam search method (which considers just the magnitude of beam space measurements) is robust to such non idealities, it does not make use of the fact that mmWave channels are sparse. Hence, there is a need to develop better channel estimation algorithms that are simultaneously robust to such non-idealities and require fewer

measurements by exploiting the sparsity of mmWave channel. To the best of our knowledge, the only work that addresses this problem is [5], where the optimal beam pair for analog beamforming is estimated using the components of the channel matrix along specially designed measurement matrices, that satisfy hardware constraints. Although it performs better than the conventional CS based algorithms that do not model CFO, it ignores noise and neglects the phase of measurements in its model. Our approach handles both these issues.

In this paper, we propose a compressive algorithm to jointly estimate the CFO and channel for narrowband analog beamforming systems (having a single RF chain) with uniform linear arrays (ULAs) at the base station (BS) and the mobile station (MS). Extension of our method to uniform planar arrays is straightforward. We assume that timing synchronization is already performed and that the channel is sparse per the virtual channel model [2]. Our key contributions are in modeling the MIMO channel and CFO using a third order tensor and compressively estimating the tensor using the available measurements. We highlight the fact that unlike in [5] where just the angle-of-arrivals (AoAs) and angle-of-departures (AoDs) are found, our method estimates the channel matrix and CFO. Further, simulations show that our algorithm performs better than that proposed in [5] for the same setting.

We use the following notation: $\mathbf{A}$ is a matrix, $\mathbf{a}$ is a column vector, $\mathcal{A}$ is a tensor and a, A denote scalars. Using this notation, $\bar{a}$ is the complex conjugate of a , $\mathbf{A}^*$ is the conjugate transpose of $\mathbf{A}$ and $\mathbf{A}^{(i)}$ denotes the i^{th} row of $\mathbf{A}$. We use $[N]$ to denote the set $\{1, 2, 3, \dots, N\}$. The symbols $\odot$ and $\otimes$ are used to denote the outer product [6] and kronecker product respectively. The matrix $\mathbf{U}_N \in \mathbb{C}^{N \times N}$ denotes a DFT matrix of dimension N and is given by $\mathbf{U}_N(k, \ell) = e^{-j \frac{2\pi(k-1)(\ell-1)}{N}}$, for $k, \ell \in [N]$. We define $\mathbf{e}_n \in \mathbb{C}^{M \times 1}$ to be a canonical basis vector with its n^{th} entry as 1.

II. PRELIMINARIES ON TENSORS

In this section, we provide necessary definitions from tensor algebra that will aid in understanding the subsequent sections. We limit our discussion to tensors of order 3, as it suffices to model our problem.

A tensor is a multidimensional array, which is essentially an extension of vectors and matrices. For example, a matrix $\mathbf{X} \in \mathbb{C}^{N_1 \times N_2}$ is a tensor of order 2 and dimension $N_1 N_2$. Similarly $\mathcal{A} \in \mathbb{C}^{N_1 \times N_2 \times N_3}$ is a tensor of order 3 and has

dimension $N_1 N_2 N_3$. For two tensors $\mathcal{A}, \mathcal{B}$, their inner product is defined as

$$\langle \mathcal{A}, \mathcal{B} \rangle = \sum_{k \in [N_3], j \in [N_2], i \in [N_1]} \mathcal{A}(i, j, k) \bar{\mathcal{B}}(i, j, k), \quad (1)$$

and the norm of a $\mathcal{A}$ is given by $\sqrt{\langle \mathcal{A}, \mathcal{A} \rangle}$. The ℓ_1 - norm of $\mathcal{A}$ is given by

$$\|\mathcal{A}\|_{\ell_1} = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \sum_{k=1}^{N_3} |\mathcal{A}(i, j, k)|. \quad (2)$$

The mode 3 unfolding of $\mathcal{A}$ is a matrix denoted by $\mathcal{A}_{(3)} \in \mathbb{C}^{N_3 \times N_1 N_2}$ and is given by $\mathcal{A}_{(3)} =$

$$[\text{vec}(\mathcal{A}(:, :, 1)), \text{vec}(\mathcal{A}(:, :, 2)), \dots, \text{vec}(\mathcal{A}(:, :, N_3))]^T. \quad (3)$$

III. SYSTEM MODEL

In this paper, we focus on initial access in narrowband mmWave systems with analog beamforming at the TX and RX. Consider a point-to-point link with ULAs of N_t and N_r antennas at the BS (TX) and MS (RX) respectively. The antenna arrays at the TX and RX are connected to their corresponding RF chain through a network of digitally controlled phase shifters. For the n^{th} measurement, let $\mathbf{f}_n$ and $\mathbf{w}_n$ be the unit norm beam training vectors applied to the phase shifters at the TX and RX respectively, with $\sqrt{N_t} \mathbf{f}_n \in \mathbb{Q}^{N_t}$, $\sqrt{N_r} \mathbf{w}_n \in \mathbb{Q}^{N_r}$. As the phase shifters are digitally controlled, we have $\mathbb{Q} = \{e^{j\theta_1}, e^{j\theta_2}, \dots, e^{j\theta_q}\}$, where $\theta_i = \frac{2\pi i}{q}$ for a q level phase quantization. For a sequence of M measurements, let $\mathbf{r}[n]$ be the received symbol and $\left(\mathbf{y}[n] = \frac{\mathbf{s}^*[n] \mathbf{r}[n]}{\|\mathbf{s}[n]\|^2}\right)$ be the measurement corresponding to the transmitted symbol $\mathbf{s}[n]$, with $\|\mathbf{s}[n]\| = \sqrt{\rho}$, $\forall n \in [M]$. The n^{th} received symbol and measurement are given by

$$\begin{aligned} \mathbf{r}[n] &= \mathbf{w}_n^* \mathbf{H} \mathbf{f}_n e^{j(\omega_e n + \phi_n)} \mathbf{s}[n] + \tilde{\mathbf{v}}[n] \\ \mathbf{y}[n] &= \mathbf{w}_n^* \mathbf{H} \mathbf{f}_n e^{j(\omega_e n + \phi_n)} + \mathbf{v}[n], \quad \forall n \in [M] \end{aligned} \quad (4)$$

where $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ is the channel matrix; $\tilde{\mathbf{v}} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_{M \times M})$, $\mathbf{v} \sim \mathcal{CN}(0, \frac{\sigma^2}{\rho} \mathbf{I}_{M \times M})$; $\phi_n | \phi_{n-1} \sim \mathcal{N}(\phi_{n-1}, \tau^2)$ is considered to be a Wiener phase noise process [7] (with $\phi_0 = 0$); $\omega_e = 2\pi f_e T$, with T as the symbol duration and f_e (Hz) as the carrier frequency offset.

Consider a propagation environment with N_{cl} clusters and N_{ray}^n rays in the n^{th} cluster. For the m^{th} path of the n^{th} cluster, let $\gamma_{n,m}$ denote the complex gain and $\theta_{n,m}^r$ ($\theta_{n,m}^t$) denote the AoA (AoD). Let λ be the carrier wavelength and d be the antenna spacing in the ULAs at the BS and MS. With $\omega_{n,m}^r := \frac{2\pi d}{\lambda} \sin(\theta_{n,m}^r)$, $\omega_{n,m}^t := \frac{2\pi d}{\lambda} \sin(\theta_{n,m}^t)$ and the following definition

$$\mathbf{a}_N(\theta) = [1 e^{j\theta} e^{j2\theta} \dots e^{j(N-1)\theta}]^T, \quad (5)$$

the MIMO channel matrix $\mathbf{H}$, in baseband is given by

$$\mathbf{H} = \frac{1}{\sqrt{N_{cl}}} \sum_{n=1}^{N_{cl}} \frac{1}{\sqrt{N_{ray}^n}} \sum_{m=1}^{N_{ray}^n} \gamma_{n,m} \mathbf{a}_{N_r}(\omega_{n,m}^r) \mathbf{a}_{N_t}^*(\omega_{n,m}^t). \quad (6)$$

At mmWave carrier frequencies, $\mathbf{H}$ in (6) is marginally sparse matrix in the spatial DFT basis [2]. The channel matrix would be exactly sparse if the spatial frequencies align exactly on the grid; when it is not the case, a grid of finer resolution can be chosen to increase compressibility at the expense of higher dimensionality. We assume that the spatial frequencies (of the form (ω_x, ω_y)) of $\mathbf{H}$ come from a discrete set, i.e., $\omega_x \in \left\{0, \frac{2\pi}{N_t}, \frac{4\pi}{N_t}, \dots, \frac{2\pi(N_t-1)}{N_t}\right\}$, $\omega_y \in \left\{0, \frac{2\pi}{N_r}, \frac{4\pi}{N_r}, \dots, \frac{2\pi(N_r-1)}{N_r}\right\}$ for our analysis; off grid extensions can be made using [3][8]. We test our algorithm using off-grid parameters in the simulations. The channel matrix in (4) can be expressed as

$$\mathbf{H} = \sum_{i=1}^{N_r} \sum_{j=1}^{N_t} \alpha_{ij} \mathbf{a}_{N_r} \left(\frac{2\pi i}{N_r} \right) \odot \mathbf{a}_{N_t} \left(\frac{2\pi j}{N_t} \right), \quad (7)$$

With the assumption that $\mathbf{H}$ is sparse in the spatial DFT basis, $\mathbf{C} := \{\alpha_{ij}\}_{i,j=1}^{N_r, N_t}$ is now a sparse matrix, of sparsity K in $N_r N_t$ dimension.

IV. MODELING CFO AND THE CHANNEL USING TENSORS

The notion behind modeling the channel and CFO in a tensor comes from the intuition to consider the spatial frequencies corresponding to AoAs, AoDs and the CFO (ω_e) along three different dimensions. It may be observed from (4) that $\mathbf{y}[n]$ is a noisy version of the inner product between an unknown matrix $(\mathbf{H} e^{j\Omega_n})$ with a known measurement matrix $\mathbf{w}_n \mathbf{f}_n^*$, where $\Omega_n = \omega_e n + \phi_n$. We model a collection of M such matrices using a tensor $\chi \in \mathbb{C}^{N_r \times N_t \times M}$ such that the k^{th} frontal slab [6] of χ is given by $\chi(:, :, k) = \mathbf{H} e^{j\Omega_k}$. Hence

$$\chi = \mathbf{H} \odot \mathbf{e}_\Omega, \quad (8)$$

where $\mathbf{e}_\Omega = (e^{j\Omega_1}, e^{j\Omega_2}, \dots, e^{j\Omega_M})^T$. We expand $\mathbf{e}_\Omega \in \mathbb{C}^{M \times 1}$ in the M dimensional discrete fourier basis as

$$\mathbf{e}_\Omega = \sum_{k=1}^M \beta_k \mathbf{a}_M \left(\frac{2\pi k}{M} \right), \quad (9)$$

with $\mathbf{a}_M(\cdot)$ defined according to (5). Unlike $\mathbf{C}$, $\mathbf{p} \triangleq (\beta_1, \beta_2, \dots, \beta_M)^T$ is not assumed to be exactly sparse because it is unrealistic to assume that CFO (ω_e) lies on the grid. Moreover phase noise distorts the spectrum corresponding to CFO. From (7), (8) and (9), the tensor χ can be expressed as

$$\chi = \sum_{i,j,k} \mathcal{G}(i, j, k) \mathbf{a}_{N_r} \left(\frac{2\pi i}{N_r} \right) \odot \mathbf{a}_{N_t} \left(\frac{2\pi j}{N_t} \right) \odot \mathbf{a}_M \left(\frac{2\pi k}{M} \right), \quad (10)$$

where $\mathcal{G} \in \mathbb{C}^{N_r \times N_t \times M}$ is another tensor, with $\mathcal{G}(i, j, k) = \alpha_{ij} \beta_k$. Using the tensor notation, the measurements in (4) can now be expressed as

$$\begin{aligned} \mathbf{y}[n] &= \langle \chi, \mathbf{w}_n \odot \bar{\mathbf{f}}_n \odot \mathbf{e}_n \rangle + \mathbf{v}[n] \\ &= \langle \mathcal{G}, \mathcal{M}_n \rangle + \mathbf{v}[n], \end{aligned} \quad (11)$$

where $\mathcal{M}_n = \mathbf{U}_{N_r} \mathbf{w}_n \odot \mathbf{U}_{N_t} \bar{\mathbf{f}}_n \odot \mathbf{U}_M \mathbf{e}_n$. We have thus modelled the channel and CFO using a tensor χ , and their joint estimation is now a tensor estimation problem.

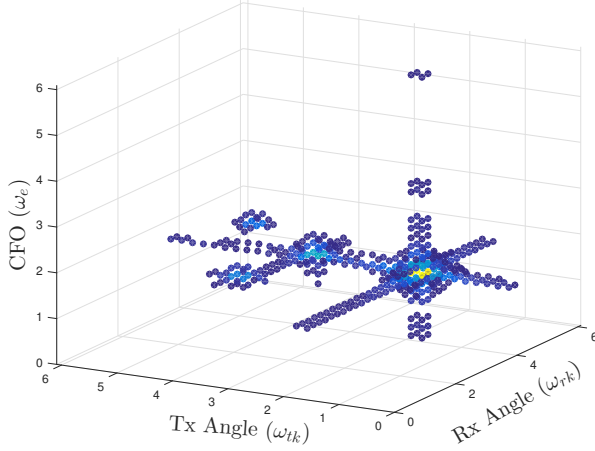


Fig. 1. Magnitude plot of the third order tensor $\mathcal{G}$, for $N_r = N_t = 32$, $N_{cl} = 4$ clusters, CFO $\omega_e = \frac{2\pi}{3}$, $\tau = 0.1$. For illustrative purposes as well as simulations, the on grid assumption on ω_{rk}, ω_{tk} is waived.

Although χ is a tensor of size $N_r N_t M$, it is just an outer product of a rank K matrix with a vector and hence tensor rank $(\chi) \leq K$ [6]. Further, χ is a tensor compressible in the 3D DFT basis, due to sparsity of $\mathbf{C}$ and marginal sparsity of $\mathbf{p}$ (as $\mathcal{G}_{ijk} = \alpha_{ij} \beta_k$), as shown in Fig.1. Hence, low rank recovery [9] or CS based [10] algorithms for tensors can be used to reconstruct χ . We use compressed sensing because the sparsity basis of χ is known and CS exploits this fact unlike low rank tensor recovery algorithms.

V. COMPRESSED SENSING BASED RECOVERY

Similar to the vector case in which a sparse vector can be recovered using projections onto a carefully chosen lower dimensional subspace, a higher dimensional sparse tensor can also be recovered from a lower dimensional one, under certain constraints. An extensive theory on CS for sparse tensor recovery can be found in [10]. The recovery of $\mathcal{G}$ in (11) using the standard ℓ_1 minimization algorithm can be given as,

$$\text{minimize } \|\mathcal{F}\|_{\ell_1} \quad (12)$$

$$\text{s.t. } \sum_{n=1}^M |\mathbf{y}[n] - \langle \mathcal{F}, \mathcal{M}_n \rangle|^2 \leq \sigma_N^2.$$

The simplest way to solve (12) would be to use Kronecker Compressed Sensing (KCS), i.e, to vectorize the tensor and then apply the well known techniques of vector compressed sensing to this problem. Though KCS can recover the sparse tensor, it has huge complexity and the structural information conveyed by the tensor may be lost due to vectorization. We use orthogonal matching pursuit (OMP), a greedy algorithm for sparse tensor recovery. For a given set of measurement tensors $\{\mathcal{M}_n\}_{n=1}^M$, we define a linear operator $\mathcal{P}: \mathbb{C}^{N_1 \times N_2 \times M} \rightarrow$

$\mathbb{C}^M$ and an adjoint $\mathcal{P}^*: \mathbb{C}^M \rightarrow \mathbb{C}^{N_1 \times N_2 \times M}$ as

$$\mathcal{P}(\mathcal{G}) = (\langle \mathcal{G}, \mathcal{M}_1 \rangle, \langle \mathcal{G}, \mathcal{M}_2 \rangle, \dots, \langle \mathcal{G}, \mathcal{M}_n \rangle)^T, \quad (13)$$

$$\mathcal{P}^*(\mathbf{y}) = \sum_{n=1}^M \mathbf{y}[n] \mathcal{M}_n. \quad (14)$$

From (11), (13), we have $\mathbf{y} = \mathcal{P}(\mathcal{G}) + \mathbf{v}$.

A. Tensor estimation algorithm

Let $\epsilon := \mathbb{E}[\mathbf{v}^* \mathbf{v}]$ be the stopping threshold and N_{iter} be the maximum allowed iterations for the OMP based recovery of $\mathcal{G}$.

Input: $\mathbf{y}, \mathcal{P}, \epsilon, \mathcal{F}^0 = \mathbf{0}, \mathcal{D}^0 = \phi, n = 1;$
while $\|\mathbf{y} - \mathcal{P}(\mathcal{F}^{(n-1)})\|^2 > \epsilon$ and $n \leq N_{\text{iter}}$
do
 $\Upsilon^{(n)} = \text{argmax} |\mathcal{P}^*(\mathbf{y} - \mathcal{P}(\mathcal{F}^{(n-1)}))|$
 $\mathcal{D}^{(n)} = \Upsilon^{(n)} \cup \mathcal{D}^{(n-1)}$
 $\mathcal{F}^{(n)} = \text{argmin}_{\mathcal{V}: \text{supp}(\mathcal{V}) = \mathcal{D}^{(n)}} \|\mathbf{y} - \mathcal{P}(\mathcal{V})\|_{\ell_2}$
end
Result: $\hat{\mathcal{G}} = \mathcal{F}^{(n)}.$

Algorithm 1: OMP to recover $\mathcal{G}$

It may be noted that $\text{argmax} |\mathcal{A}|$ returns a tuple corresponding to the location where the tensor takes its maximum (in magnitude). In a nutshell, the OMP algorithm iteratively estimates $\mathcal{G}$, with $\mathcal{F}^{(n)}$ being the estimate at the n^{th} iteration. In every iteration, it identifies a tensor element that maximally explains the residue and finds a tensor accordingly, using a least squares estimate. A detailed treatment on OMP can be found in [11].

Having obtained $\hat{\mathcal{G}}$ using Algorithm 1, we now need to split it into a matrix and a vector, corresponding to $\mathbf{C}$ and $\mathbf{p}$ respectively, as $\mathcal{G} = \mathbf{C} \odot \mathbf{p}$. This decomposition is performed using singular value decomposition (svd) of $\hat{\mathcal{G}}_{(3)}$, the mode 3 unfolding of $\hat{\mathcal{G}}$. In other words, $\hat{\mathbf{p}}$ and $\text{vec}(\hat{\tilde{\mathbf{C}}})$ are the left and right singular vectors corresponding to the maximum singular value of $\hat{\mathcal{G}}_{(3)}$. The final step of unfolding the tensor along mode 3 followed by the SVD is inspired by [12]. It may be noted that $\hat{\mathbf{p}}$ is an estimate of $(\beta_1, \beta_2, \dots, \beta_M)$, the DFT of $\mathbf{e}_\Omega$ in (9). Hence, an estimate of CFO (ω_e) can be obtained from $\hat{\mathbf{p}}$, using a kalman filter based approach, that accounts for the Wiener phase noise. Due to space constraints, we omit the discussion on CFO estimation and focus only on beam alignment in the subsequent sections.

B. Analog beamforming

With $\hat{\mathbf{C}}$, the channel estimate can be given by $\mathbf{H}_{\text{est}} = \mathbf{U}_{N_r}^* \hat{\mathbf{C}} \mathbf{U}_{N_t}^*$, upto a scale factor. For data transmission, the beamforming vectors $\mathbf{f}_{\text{est}} \in \mathbb{Q}^{N_t}$ and $\mathbf{w}_{\text{est}} \in \mathbb{Q}^{N_r}$ have to be chosen such that $|\mathbf{w}_{\text{est}}^* \mathbf{H}_{\text{est}} \mathbf{f}_{\text{est}}|$ is maximized. We find $\mathbf{f}_{\text{est}}, \mathbf{w}_{\text{est}}$ by performing the SVD of $\mathbf{H}_{\text{est}}$ followed by element-wise phase quantization of the singular vectors, corresponding to the maximum singular value.

C. Advantages of modeling using tensors

The most important advantage of modeling the MIMO channel and CFO estimation problem using tensors is that it maintains the structural information, which is otherwise lost when vectorization is used. Modeling with tensors provides scope for spectral compressed sensing [8], which deals with problem of off the grid CS. Using the techniques in [8], finer estimates of CFO and spatial frequencies of the channel matrix can be obtained. Further, it can be seen from Fig.1, that most of the mass in the tensor $\mathcal{G}$ is concentrated along a set of horizontal slabs around $\omega = \omega_e$. This information can be used to apply structured CS algorithms [13] to either reduce the number of CS measurements or provide better estimates with the same number of measurements.

D. Reducing the dimension of our problem

In typical systems $f_e \in [-f_{\max}, f_{\max}]$, where $f_{\max}$ depends on the quality of local oscillators used at TX and RX, and is typically in the order of parts per millions (ppms) of the carrier frequency. As a DFT grid of finite resolution ($\frac{2\pi}{M}$) is used to model CFO, a spectral leakage factor of $\gamma \in [1, (2Tf_{\max})^{-1}]$ is considered and the components that lie on the DFT grid within the range $[-\gamma f_{\max}, \gamma f_{\max}]$ are recovered. With $P = \lceil M\gamma f_{\max}T \rceil$, the following approximation for $\mathbf{e}_\Omega$ in (9) is used.

$$\mathbf{e}_\Omega = \sum_{k=1}^P \beta_k \mathbf{a}_M \left(\frac{2\pi k}{M} \right) + \sum_{k=M-P}^M \beta_k \mathbf{a}_M \left(\frac{2\pi k}{M} \right) \quad (15)$$

The limits of k in (10) change accordingly and we now need to solve for a tensor of dimension $N_t N_r (2P + 1)$ rather than $N_t N_r M$. It may be noted that this approximation does not hold for high values of phase noise variance (τ), as the magnitude spectrum of $\mathbf{e}_\Omega$ would not be concentrated about ω_e in such case.

E. Analogy with lifting techniques

Although we started with the intuition to model spatial frequency and CFO along different dimensions, we have essentially moved to a higher dimensional space. Our framework of modeling CFO, channel using tensors is analogous to lifting techniques that exist in the literature [14] [12]. The mode 3 unfolding of $\mathcal{G}$ can also be derived using lifting [12], a method that convexifies a nonconvex optimization problem by moving to a higher dimensional space, and is illustrated below.

Let the phase error free measurements be given by $\mathbf{z}[n] = \mathbf{w}_n^* \mathbf{H} \mathbf{f}_n + \mathbf{v}[n]$. From (7), it follows that $\mathbf{H} = \mathbf{U}_{N_r}^* \mathbf{C} \mathbf{U}_{N_t}^*$ in the spatial DFT basis. Using properties of kroenecker products, it can be shown that $\mathbf{z}[n] = (\mathbf{f}_n^T \mathbf{U}_{N_t}^*) \otimes (\mathbf{w}_n^* \mathbf{U}_{N_r}^*) \text{vec}(\mathbf{C}) + \mathbf{v}[n]$. On stacking all the phase error free measurements we get $\mathbf{z} = \mathbf{A} \mathbf{x} + \mathbf{v}$, where $\mathbf{A}^{(n)} = (\mathbf{f}_n^T \mathbf{U}_{N_t}^*) \otimes (\mathbf{w}_n^* \mathbf{U}_{N_r}^*)$, and $\mathbf{x} = \text{vec}(\mathbf{C})$ is a K sparse vector. The observed measurements in (4) can now be given by,

$$\begin{aligned} \mathbf{y} &= \text{diag}(e^{j\Omega_1}, e^{j\Omega_2}, \dots, e^{j\Omega_n}) \mathbf{A} \mathbf{x} + \mathbf{v} \\ &= \text{diag}(\mathbf{U}_M^* \mathbf{p}) \mathbf{A} \mathbf{x} + \mathbf{v} \end{aligned} \quad (16)$$

Hence, we have

$$\mathbf{y}[n] = \mathbf{U}_M^{*(n)} \mathbf{p} \mathbf{A}^{(n)} \mathbf{x} + \mathbf{v}[n] = \mathbf{U}_M^{*(n)} \mathbf{p} \mathbf{x}^T \left(\mathbf{A}^{(n)} \right)^T + \mathbf{v}[n] \quad (17)$$

By defining $\mathbf{X} = \mathbf{p} \mathbf{x}^T$, we have $\mathbf{y} = \mathcal{P}_1(\mathbf{X}) + \mathbf{v}$, where $\mathcal{P}_1$ is a linear operator defined in accordance with (17). It is proposed in [12] to recover $\mathbf{X}$, by minimizing it's ℓ_1 norm and then perform SVD to obtain estimates of $\mathbf{p}, \mathbf{x}$, upto a scale factor. From a signal processing perspective, the tensor based approach can model and recover a sparse tensor with arbitrary locations of sparsity (eg:- finding three dimensional frequency of a third order tensor). Lifting can be considered as a specific instance of the tensor approach, where the locations of sparsity are constrained around a plane.

VI. SIMULATION RESULTS

In this section, we compare the performance of our tensor based algorithm with Agile Link [5] and OMP (CFO ignored). We consider the system model in Section III, with ULAs of size $N_t = 32$ and $N_r = 16$, antenna spacing of $d = \frac{\lambda}{2}$ for each of the ULAs and a narrowband mmWave channel in (6) with $N_{cl} = 2$ clusters, each comprising of $N_{ray} = 10$ paths and 3 degrees of angular spread. We consider the complex path gains $\gamma_{n,m} \stackrel{\text{iid}}{\sim} \mathcal{CN}(0, 1)$, $\forall m, n$. To increase the compressibility of the corresponding channel matrix, we choose a $2 \times$ oversampled DFT grid along the AoA and AoD dimensions, corresponding to resolutions of $\frac{\pi}{N_r}$ and $\frac{\pi}{N_t}$ respectively. We consider a carrier frequency of $f_c = 28\text{GHz}$, the maximum CFO limit to be 10ppm of f_c i.e., $f_{\max} = 280\text{KHz}$ and the spectral leakage factor(γ) to be 2. The symbol duration T is chosen to be $0.5\mu s$, which is much larger than the maximum delay spread given in [15] at 28GHz and hence our narrowband assumption is justified.

We assume a digital phase control of 3 bits ($q = 8$) at the BS and MS. For each measurement $n \in [M]$, the vectors $\mathbf{w}_n, \mathbf{f}_n$, have unit ℓ_2 norm and are chosen independently and uniformly at random from a quantized set. Element wise phase quantization of 3 bits is performed on the unquantized beamforming weights, also of unit norm. The achievable rate $R = \log_2 \left(1 + \frac{\rho |\mathbf{w}_{\text{est}}^* \mathbf{H} \mathbf{f}_{\text{est}}|^2}{\sigma^2} \right)$ was averaged over 1000 realizations of $\mathbf{H}$ for all the three algorithms.

Agile Link was evaluated with the same system model and beamtraining vectors as proposed in [5]. In addition, we perform 3 bit quantization on these vectors and normalize them. The number of CS measurements required for Agile Link[5] is given by $B_r B_t N_{\text{hash}}$, with $B_r, B_t = O(K)$ and $N_{\text{hash}} = O(\log(N_t N_r))$. Under the same settings, we also evaluate the standard OMP [11] when phase noise and CFO are ignored. We evaluate our algorithm in the worst case scenario i.e, when f_e is close to $f_{\max} = 280\text{KHz}$ and is maximally off-grid (considering a DFT grid of resolution $\frac{1}{MT} = 31.25\text{KHz}$, for $M = 64$). Hence $f_e = 265.625\text{KHz}$ is chosen. For Agile Link $B_r, B_t, N_{\text{hash}}$ were optimized and set to (4, 4, 4) and (4, 4, 8) for 64 and 128 measurements respectively. The proposed tensor based approach, however, demands $2P + 1$ times higher complexity in memory and

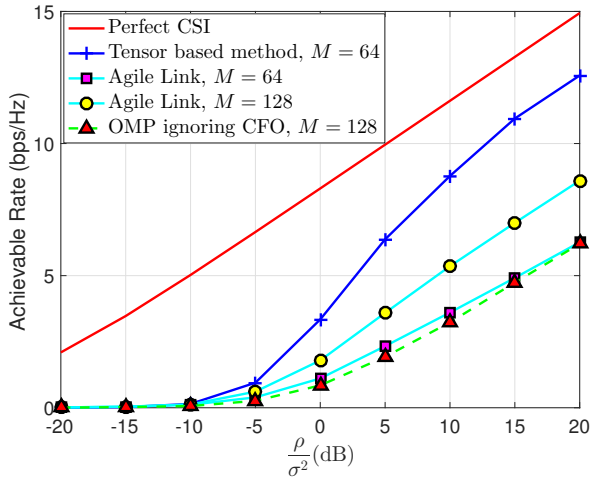


Fig. 2. The figure shows the achievable rate versus SNR for our tensor based method, OMP (CFO ignored) and Agile Link for $f_e = 265.625\text{KHz}$, $\tau = 0.27\text{rad}$

time than [5] or standard OMP, which is well justified by the significant performance gains relative to the other two as seen in Fig.2. Agile Link considers the magnitude of the noisy measurements which deteriorates its performance at low SNRs. It may be noted that although Agile Link is robust to CFO errors at high SNRs, it needs larger number of measurements to identify the optimal beamforming weights compared to our tensor based approach, because the former ignores the phase of the measurements. Furthermore, our proposed method estimates the channel matrix unlike Agile Link which finds just the beam steering vectors. From Fig.3,

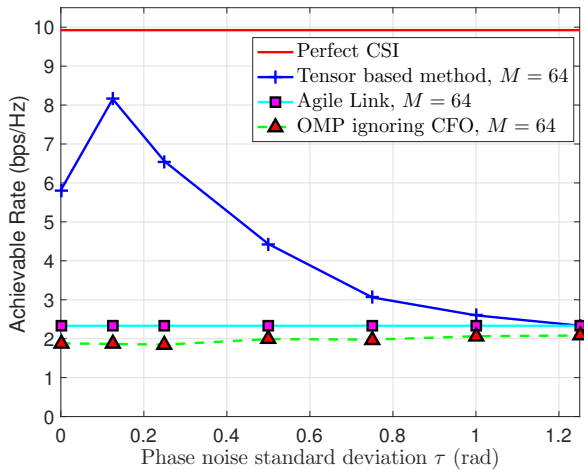


Fig. 3. The figure shows the achievable rate versus the standard deviation of phase noise process (τ), for all the three methods, at an SNR of 5dB

it can be seen that the tensor based algorithm is better than Agile Link in a wide range of τ . The practical value of τ is given by $2\pi f_c \sqrt{cT_s}$ [7], where $c = 4.7 \times 10^{-18} (\text{rad.Hz})^{-1}$, and it can be verified that $\tau = 0.27\text{rad}$ for our settings.

VII. CONCLUSION AND FUTURE WORK

We have proposed a compressive channel estimation technique for narrowband mmWave systems using analog beamforming, that is robust to synchronization impairments. The essence of our paper is to model CFO and channel in a tensor, and to recover the tensor, while still exploiting the sparsity of the mmWave channel. With few measurements, our method is able to do beam alignment better than the existing ones, in addition to estimating the carrier frequency offset. Considering timing mismatch, investigating performance bounds for the proposed method as a function of synchronization impairments, reducing the computational complexity using structured CS algorithms and developing robust CS algorithms for wideband mmWave systems are interesting directions for future work.

REFERENCES

- [1] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," in *Proceedings of the IEEE*, vol. 102, no. 3, pp. 366–385, 2014.
- [2] 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.
- [3] D. Ramasamy, S. Venkateswaran, and U. Madhow, "Compressive parameter estimation in AWGN," *IEEE Trans. Signal Process.*, vol. 62, no. 8, pp. 2012–2027, Aug. 2014.
- [4] A. Alkhateeb, G. Leus, and R. W. Heath, "Compressed sensing based multi-user millimeter wave systems: How many measurements are needed?" in *Proc. of the IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 2015, pp. 2909–2913.
- [5] O. Abari, H. Hassanieh, M. Rodreguez, and D. Katabi, "Millimeter wave communications: From point-to-point links to agile network connections," in *Proc. of the ACM Workshop on Hot Topics in Networks*, 2016, pp. 169–175.
- [6] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *arXiv preprint arXiv:1607.01668*, Jul. 2016.
- [7] A. Pitarokoilis, S. K. Mohammed, and E. G. Larsson, "Effect of oscillator phase noise on uplink performance of large MU-MIMO systems," in *Proc. of the 50th Allerton Conf. on Commun., Control, and Computing (Allerton)*, 2012, pp. 1190–1197.
- [8] M. F. Duarte and Y. C. Eldar, "Structured compressed sensing: From theory to applications," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4053–4085, 2011.
- [9] S. Gandy, B. Recht, and I. Yamada, "Tensor completion and low-rank tensor recovery via convex optimization," *Inverse Problems*, vol. 27, no. 2, p. 025010, 2011.
- [10] S. Friedland, Q. Li, and D. Schonfeld, "Compressive sensing of sparse tensors," *IEEE Trans. on Image Process.*, vol. 23, no. 10, pp. 4438–4447, 2014.
- [11] T. T. Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery with noise," *IEEE Trans. on Inf. Theory*, vol. 57, no. 7, pp. 4680–4688, 2011.
- [12] S. Ling and T. Strohmer, "Self-calibration and biconvex compressive sensing," *Inverse Problems*, vol. 31, no. 11, p. 115002, 2015.
- [13] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Trans. on Inf. Theory*, vol. 56, no. 4, pp. 1982–2001, 2010.
- [14] E. J. Candes, T. Strohmer, and V. Voroninski, "Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming," *Commun. on Pure and Applied Mathematics*, vol. 66, no. 8, pp. 1241–1274, 2013.
- [15] S. Deng, M. K. Samimi, and T. S. Rappaport, "28 GHz and 73 GHz millimeter-wave indoor propagation measurements and path loss models," in *Proc. of the IEEE Intl. Conf. on Commun. Workshop*, 2015, pp. 1244–1250.