Channel Estimation for Hybrid Architecture-Based Wideband Millimeter Wave Systems

Kiran Venugopal, *Student Member, IEEE*, Ahmed Alkhateeb, *Member, IEEE*, Nuria González Prelcic, and Robert W. Heath, Jr., *Fellow, IEEE*

***Abstract*— Hybrid analog and digital precoding allows millimeter wave (mmWave) systems to achieve both array and multiplexing gain. The design of the hybrid precoders and combiners, though, is usually based on the knowledge of the channel. Prior work on mmWave channel estimation with hybrid architectures focused on narrowband channels. Since mmWave systems will be wideband with frequency selectivity, it is vital to develop channel estimation solutions for hybrid architectures-based wideband mmWave systems. In this paper, we develop a sparse formulation and compressed sensing-based solutions for the wideband mmWave channel estimation problem for hybrid architectures. First, we leverage the sparse structure of the frequency-selective mmWave channels and formulate the channel estimation problem as a sparse recovery in both time and frequency domains. Then, we propose explicit channel estimation techniques for purely time or frequency domains and for combined time/frequency domains. Our solutions are suitable for both single carrier-frequency domain equalization and orthogonal frequency-division multiplexing systems. Simulation results show that the proposed solutions achieve good channel estimation quality, while requiring small training overhead. Leveraging the hybrid architecture at the transceivers gives further improvement in estimation error performance and achievable rates.**

***Index Terms*— Millimeter wave communications, channel estimation, frequency-selective channel, hybrid precoder-combiner, compressed sensing, sparse recovery, multi-stream MIMO, IEEE 802.11ad.**

I. Introduction

CHANNEL estimation in millimeter wave MIMO systems allows flexible design of hybrid analog/digital precoders and combiners under different optimization criteria. Unfortunately, the hybrid constraint makes it challenging to directly estimate the channels, due to the presence of the analog beamforming/combining stage to avoid the power consumption of the typical all-digital solution used at lower frequencies [2]–[4]. Further, operating at mmWave frequencies complicates the estimation of the channel because the signal-to-noise-ratio (SNR) before beamforming is low and the dimensions of the channel matrices associated with mmWave arrays [2], [5] are large. While mmWave channel estimation has extensively studied in the last few years, most prior work assumed a narrowband channel model. Since mmWave systems are attractive due to their wide bandwidth, developing efficient mmWave channel estimation for frequency-selective channels is of great importance.

A. Prior Work

To avoid the explicit estimation of the channel, analog beam training solutions were proposed [6]–[8]. In beam training, the transmitter and receiver iteratively search for the beam pair that maximizes the link SNR [6], [8], [9]. This approach is used in IEEE standards like 802.11ad [10] and 802.15.3c [11]. The directional antenna patterns can be realized using a network of phase shifters. While analog beam training works for both narrowband and wideband systems, the downside is that the solution supports mainly a single communication stream; extensions to multi-stream and multi-user communication are possible but generally incur much higher overhead.

Exploiting sparsity has been critical in formulating practical channel estimators in the hybrid MIMO architecture [4], [12]. The reason is that the analog precoding and combing stages act to perform spatial compression, reducing the dimensionality of the channel observed in the digital domain compared to the channel at the antennas. When sparsity is employed, a succession of different precoders and combiners are used to make the compressive measurements. Then, decomposing the uncompressed channel to expose the dictionary in which the channel parameters are sparse, an optimization problem is solved to determine the locations of the sparse coefficients and their values. This approach is different from beam training [6], [8], [9], which attempts to find the beams that point in the most promising directions and not estimating the channel. In compressive methods, the analog beamformers and combiners are selected to make measurements in several directions.

There are several solutions for compressive channel estimation in frequency-flat mmWave channels [12]–[15], [15]–[20]. In the frequency-flat, narrowband case, the sparsity

Manuscript received November 11, 2016; revised March 16, 2017; accepted March 30, 2017. Date of publication June 28, 2017; date of current version August 18, 2017. This work was supported in part by the Intel-Verizon 5G Research Program, in part by the National Science Foundation under Grant NSF-CCF-1319556, in part by the Agencia Estatal de Investigación, Spain, and in part by the European Regional Development Fund through projects MYRADA under Grant TEC2016- 75103-C2-2-R and COMONSENS under Grant TEC2015-69648-REDC. A portion of the results in this paper was presented at ICASSP 2017 [1]. *(Corresponding author: Robert W. Heath, Jr.)*

K. Venugopal and R. W. Heath, Jr., are with The University of Texas at Austin, Austin, TX 78701 USA (e-mail: kiranv@utexas.edu; rheath@utexas.edu).

A. Alkhateeb was with the University of Texas at Austin, Austin, TX 78701 USA. He is now with the Connectivity Lab, Facebook, Menlo Park, CA 94025 USA (e-mail: aalkhateeb@fb.com).

N. González Prelcic is with the University of Vigo, 36310 Vigo, Spain (e-mail: nuria@gts.uvigo.es).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSAC.2017.2720856

in the angular domain is exploited by making use of the extended virtual channel model [2]. Essentially, the MIMO channel is written in terms of dictionary matrices built from the transmit and receive steering vectors evaluated on a uniform grid of possible angles of arrival and departure (AoA/AoD). These dictionary matrices operate as a sparsifying basis for the channel matrix. Using that formulation, several channel estimation algorithms that use compressed sensing (CS) tools have been developed for hybrid architectures [12]–[16], where the training / measurement matrices are designed using hybrid precoders and combiners. These techniques differ in the way these measurement matrices search for the dominant angles of arrival and departure. Solutions that make use of adaptive compressed sensing [12], [17], [18], random compressed sensing [13]–[15], [19], [20], joint random and adaptive compressed sensing [15] were studied. Other non-compressed sensing techniques were also developed for mmWave channel estimation using subspace estimation [21], overlapped beams [22], and auxiliary beams [23]. A main limitation of [12]–[15] and [21]–[23] is that mmWave channel bandwidths will be large (in fact the main motivation for using millimeter wave) and the underlying channels are better modeled as frequency-selective.

B. Contributions

In this paper, we propose and evaluate a new approach for estimating wideband mmWave channels based on sparse recovery. Unlike prior work for narrowband and frequency-flat channels [12]–[15], [21]–[23], we redefine the sparsifying dictionaries to account for the sparse nature of wideband frequency-selective mmWave channels in both the angular and the delay domains. Our formulation explicitly accounts for bandlimiting filtering, which spreads the contributions of multi-paths in the channel among several discrete-time channel taps. Once the channel is written in terms of the sparsifying dictionary matrices, the hardware constraints associated with the analog precoding stage are also introduced into the formulation of the channel estimation problem. This allows application of various algorithms on sparse reconstruction from the CS literature.

In formulating our algorithm, we incorporate key system constraints associated with the analog precoders and combiners. As with prior work in the narrowband case [14], we assume that the precoders and combiners are frequency flat and are generated using a fully connected architecture with digitally controlled quantized phase shifters. Unlike other work, though, we also account for the finite switching time required to reconfigure the phase shifters from one value to another [24]. As a result, our proposed system performs training using a series of zero-padded Single Carrier - Frequency Domain Equalization (SC-FDE) frames or Orthogonal Frequency-Division Multiplexing (OFDM) symbols. Each SC-FDE frame or OFDM symbol is transmitted using a single analog precoder and received with a single analog combiner; the analog portions are reconfigured during the guard interval. By using zero padding instead of a cyclic prefix, the distortion incurred during switching can be neglected. This changes the effective received signal model and means that prior work assuming the usual cyclic prefix during training like [25]–[27] does not apply. To the best of our knowledge there are no prior work on wideband millimeter channel estimation that identify the hardware limitations and the subsequent signal processing necessary to circumvent them.

The main contributions can be summarized as follows:

- We define an appropriate sparsifying dictionary for frequency-selective mmWave channels. This dictionary depends on the transmit and receive array steering vectors evaluated on a uniform grid of possible AoAs/AoDs, and also on a band-limiting filter evaluated on a uniform grid of possible delays. This key step leads to a representation of the MIMO channel matrix that leverages the sparse structure of the mmWave channel in both the angular and delay domains. Unlike the prior work in [25] and [27], we do not limit the channel model to a virtual channel model, wherein the sparsifying dictionary is square. Note that with the virtual channel model, the additional step of shaping may be used to avoid the error due to quantization, as elaborated in [27]. Our representation is more general and gives better flexibility to define the sparsifying dictionary to get the required level of accuracy for the channel estimates.
- We present an algorithm for estimating the wideband mmWave channel in the time-domain. Important practical features critical for mmWave system modeling are incorporated in our sparse formulation. The proposed formulation simultaneously leverages the structure of the frequency-selective large antenna mmWave channel in both the delay and angular domains. This avoids the multiple measurement problem in the frequency-domain. Unlike prior work which either relies on fully-digital and/or OFDM systems for wideband channel estimation [25]–[27], our proposed approach works both for SC-FDE and OFDM based frequency-selective hybrid mmWave systems. To our knowledge there is no previous work in mmWave channel estimation that also uses a grid in the delay domain along with the angle domain.
- We present an algorithm for estimating the wideband mmWave channel in the frequency-domain, for an OFDM system. A critical step in this formulation is the use of zero padding, instead of cyclic prefixing. Zero padding helps analog circuitry reconfiguration from one OFDM symbol to the next, without corrupting the training data at the symbol edges. Most prior work, including [25] and [26], did not exploit the fact that baseband channel is band-limited and assumed perfect frequency domain equalization, or OFDM where the cyclic prefixing is not corrupted – thanks to perfect analog circuitry reconfiguration. Note also that not considering the band-limiting filter response in the effective baseband channel model artificially enhances the sparsity level in the channel as we explain later in the manuscript. In practice, these constraints necessitate additional change in the receiver chain, making the proposed frequency-domain formulation of the channel estimation problem novel.
- We propose explicit algorithms to solve the sparse recovery problems in (1) purely time-domain, (2) purely

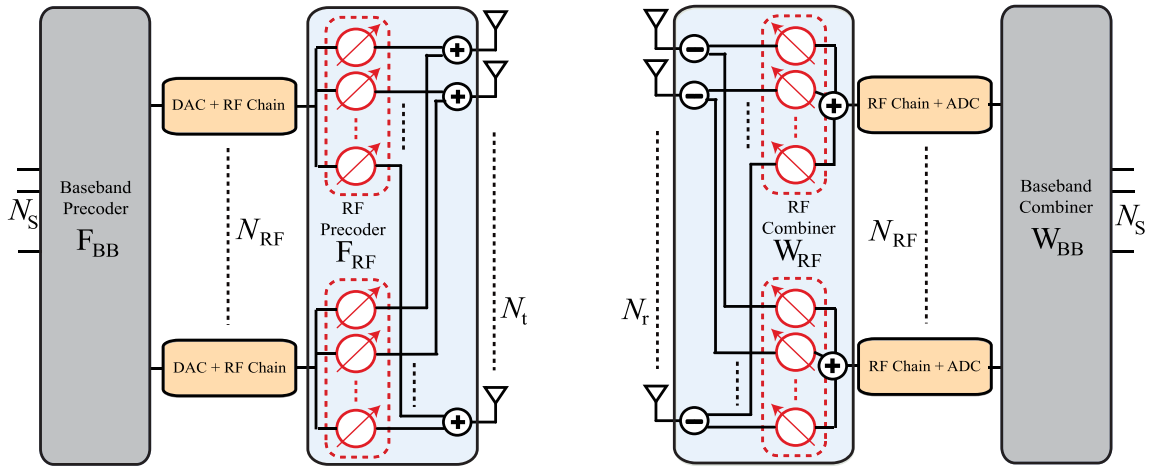


Fig. 1. Figure illustrating the transmitter and receiver structure assumed for the hybrid precoding and combining in the paper. The RF precoder and the combiner are assumed to be implemented using a network of fully connected phase shifters.

frequency-domain, and (3) combined time-frequency domains. The different approaches proposed in this paper can be suitably used for different scenarios based on system level constraints and implementation. Our proposed time-domain algorithms leverage the dictionary formulation that accounts for the sparsity in the delay domain, while the frequency-domain techniques work independent of the delay domain sparsity constraints. These proposed algorithms could serve as a baseline for future work that model a practical hardware-constrained wideband mmWave hybrid system.

- Unlike our prior work in [1], we leverage the hybrid architecture at the both at the transmitter and the receiver of mmWave wideband systems, to show how compressive sensing, hybrid precoding and combining result in low training overhead for explicit channel estimation in frequency-selective mmWave systems in both time-domain and frequency-domain. The proposed channel estimation techniques can be used to enable MIMO and multi-user communication in 802.11ad, as a potential application area.

It is explained through simulation results that the proposed algorithms require significantly less training than when beam training (eg. IEEE 802.11ad) is used for estimating the dominant angles of arrival and departure of the channel. A strict comparison with existing beam training algorithms in terms of rate performance is not reasonable since they focus, not on estimating the explicit frequency-selective mmWave MIMO channel, but on estimating beam pairs that give good link SNR. Ensuring low estimation error rates in our proposed algorithms, however, implies that efficient hybrid precoders and combiners can be designed to support rates similar to all-digital solutions [28]. We therefore rely mainly on the average error rates to compare the efficiency of our approaches. We show that utilizing multiple RF chains at the transceivers further reduces the estimation error and the training overhead. Simulation results compare the three proposed techniques. The performance of the proposed techniques as system and channel parameters are varied are presented to identify which approach suits better for a given scenario.

Notation: We use the following notation in the rest of the paper: bold uppercase $\mathbf{A}$ is used to denote matrices, bold lower case $\mathbf{a}$ denotes a column vector, and non-bold lower case a is used to denote scalar values. We use $\mathcal{A}$ to denote a set. Further, $||\mathbf{A}||_F$ is the Frobenius norm, and $\mathbf{A}^*$, $\bar{\mathbf{A}}$ and $\mathbf{A}^T$ are the conjugate transpose, conjugate, and transpose of the matrix $\mathbf{A}$. The (i, j)th entry of matrix $\mathbf{A}$ is denoted using $[\mathbf{A}]_{i,j}$. The identity matrix is denoted as $\mathbf{I}$. Further, if $\mathbf{A}$ and $\mathbf{B}$ are two matrices, $\mathbf{A} \circ \mathbf{B}$ is the Khatri-Rao product of $\mathbf{A}$ and $\mathbf{B}$, and $\mathbf{A} \otimes \mathbf{B}$ is their Kronecker product. We use $\mathcal{N}(\mathbf{m}, \mathbf{R})$ to denote a circularly symmetric complex Gaussian random vector with mean $\mathbf{m}$ and covariance $\mathbf{R}$. We use $\mathbb{E}$ to denote expectation. Discrete time-domain signals are represented as $\mathbf{x}[n]$, with the bold lower case denoting vectors, as before. The frequency-domain signals in the kth subcarrier are represented using $\check{\boldsymbol{x}}[k]$.

II. System and Channel Models

In this section, we present the SC-FDE hybrid architecture based system model, followed by a description of the adopted wideband mmWave channel model. The time-domain channel estimation algorithm proposed in Section III operates on this kind of SCE-FDE hybrid system, while the frequency-domain approach described in Section IV can be applied to OFDM-based hybrid MIMO systems as that in [28].

A. System Model

Consider a single-user mmWave MIMO system with a transmitter having N_t antennas and a receiver with N_r antennas. Both the transmitter and the receiver are assumed to have N_{RF} RF chains as shown in Fig. 1. The hybrid precoder and combiner used in the frequency-selective mmWave system is generally of the form $\mathbf{F}^{fd}[k] = \mathbf{F}_{RF}\mathbf{F}_{BB}^{fd}[k] \in \mathbb{C}^{N_t \times N_s}$ and $\mathbf{W}^{fd}[k] = \mathbf{W}_{RF}\mathbf{W}_{BB}^{fd}[k] \in \mathbb{C}^{N_r \times N_s}$, respectively for the kth subcarrier [28]. In this paper, we focus on the channel estimation having the training precoders/combiners done in the time-domain, so we will use $\mathbf{F}$ and $\mathbf{W}$ (without k) to denote the time-domain training precoders/combiners. Accordingly, the transmitter uses a hybrid precoder $\mathbf{F} = \mathbf{F}_{RF}\mathbf{F}_{BB} \in \mathbb{C}^{N_t \times N_s}$, N_s being the number of data streams that can be transmitted.

Denoting the symbol vector at instance n as $\mathbf{s}[n] \in \mathbb{C}^{N_s \times 1}$, satisfying $\mathbb{E}[\mathbf{s}[n]\mathbf{s}[n]^*] = \frac{1}{N_s}\mathbf{I}$, the signal transmitted at discrete-time n is $\tilde{\mathbf{s}}[n] = \mathbf{F}\mathbf{s}[n]$.

The $N_r \times N_t$ channel matrix between the transmitter and the receiver is assumed to be frequency-selective, having a delay tap length N_c and is denoted as $\mathbf{H}_d,\ d = 0,\ 1, \ldots,\ N_c - 1$. With $\mathbf{v}[n] \sim \mathcal{N}(0, \sigma^2\mathbf{I})$ denoting the additive noise vector, the received signal can be written as

$$\mathbf{r}[n] = \sqrt{\rho} \sum_{d=0}^{N_c-1} \mathbf{H}_d \mathbf{F}\mathbf{s}[n-d] + \mathbf{v}[n]. \tag{1}$$

The noise sample variance $\sigma^2 = N_o B$, where B is the wideband system bandwidth, so that the received signal SNR $= \rho/\sigma^2$. The receiver applies a hybrid combiner $\mathbf{W} = \mathbf{W}_{RF}\mathbf{W}_{BB} \in \mathbb{C}^{N_r \times N_s}$, so that the post combining signal at the receiver is

$$\mathbf{y}[n] = \sqrt{\rho} \sum_{d=0}^{N_c-1} \mathbf{W}^*\mathbf{H}_d \mathbf{F}\mathbf{s}[n-d] + \mathbf{W}^*\mathbf{v}[n]. \tag{2}$$

There are several RF precoder and combiner architectures that can be implemented [13]. In this paper, we assume a fully connected phase shifting network [13]. We also consider the constraint so that only quantized angles in

$$\mathcal{A} = \left\{0,\ \frac{2\pi}{2^{N_Q}},\ \cdots,\ \frac{\left(2^{N_Q}-1\right)2\pi}{2^{N_Q}}\right\} \tag{3}$$

can be realized in the phase shifters. Here N_Q is the number of angle quantization bits. This implies $[\mathbf{F}]_{i,j} = \frac{1}{\sqrt{N_t}} e^{j\varphi_{i,j}}$ and $[\mathbf{W}]_{i,j} = \frac{1}{\sqrt{N_r}} e^{j\omega_{i,j}}$, with $\varphi_{i,j},\ \omega_{i,j} \in \mathcal{A}$.

B. Channel Model

Consider a geometric channel model [12], [27] for the frequency-selective mmWave channel consisting of N_p paths. The dth delay tap of the channel (for $d = 0,\ 1,\ \ldots,\ N_c - 1$) can be expressed as

$$\mathbf{H}_d = \sum_{\ell=1}^{N_p} \alpha_\ell\, p(dT_s - \tau_\ell)\mathbf{a}_R(\phi_\ell)\mathbf{a}_T^*(\theta_\ell), \tag{4}$$

where $p(\tau)$ denotes the band-limited pulse shaping filter response evaluated at τ, $\alpha_\ell \in \mathbb{C}$ is the complex gain of the ℓth channel path, $\tau_\ell \in \mathbb{R}$ is the delay of the ℓth path, $\phi_\ell \in [0, 2\pi)$ and $\theta_\ell \in [0, 2\pi)$ are the angles of arrival and departure, respectively of the ℓth path, and $\mathbf{a}_R(\phi_\ell) \in \mathbb{C}^{N_r \times 1}$ and $\mathbf{a}_T(\theta_\ell) \in \mathbb{C}^{N_t \times 1}$ denote the antenna array response vectors of the receiver and transmitter, respectively. Note that, the effective baseband channel is seen through the RF front end (analog processing) and hence would include the filter response used for band limiting the signal in the receiver chain, as modeled in (4).

The transmitter and the receiver are assumed to know the array response vectors. The proposed estimation algorithm applies to any arbitrary antenna array configuration. The channel model in (4) can be written compactly as

$$\mathbf{H}_d = \mathbf{A}_R\, \Delta_d \mathbf{A}_T^*, \tag{5}$$

where $\Delta_d \in \mathbb{C}^{N_p \times N_p}$ is diagonal with non-zero entries $\alpha_\ell\, p(dT_s - \tau_\ell)$, and $\mathbf{A}_R \in \mathbb{C}^{N_r \times N_p}$ and $\mathbf{A}_T \in \mathbb{C}^{N_t \times N_p}$ contain the columns $\mathbf{a}_R(\phi_\ell)$ and $\mathbf{a}_T(\theta_\ell)$, respectively. Under this notation, vectorizing the channel matrix in (5) gives

$$\mathrm{vec}(\mathbf{H}_d) = \left(\bar{\mathbf{A}}_T \circ \mathbf{A}_R\right) \begin{bmatrix} \alpha_1\, p(dT_s - \tau_1) \\ \alpha_2\, p(dT_s - \tau_2) \\ \vdots \\ \alpha_{N_p}\, p(dT_s - \tau_{N_p}) \end{bmatrix}. \tag{6}$$

Note that the ℓth column of $\bar{\mathbf{A}}_T \circ \mathbf{A}_R$ is of the form $\bar{\mathbf{a}}_T(\theta_\ell) \otimes \mathbf{a}_R(\phi_\ell)$. We define the vectorized channel

$$\mathbf{h}_c = \begin{bmatrix} \mathrm{vec}(\mathbf{H}_0) \\ \mathrm{vec}(\mathbf{H}_1) \\ \vdots \\ \mathrm{vec}(\mathbf{H}_{N_c-1}) \end{bmatrix}, \tag{7}$$

which is the unknown signal that is estimated using the channel estimation algorithms proposed in the paper. We assume that the average channel power $\mathbb{E}\left[\|\mathbf{h}_c\|_2^2\right] = N_r N_t$ to facilitate comparison of the various channel estimation approaches proposed next.

III. Time-Domain Channel Estimation via Compressed Sensing

In this section, we present our proposed time-domain explicit channel estimation algorithm that leverages sparsity in the wideband mmWave channel. The hardware constraints on the training frame structure and the precoding-combining beam patterns are also explained.

A. Sparse Formulation in the Time-Domain

For the sparse formulation of the proposed time-domain approach, consider block transmission of training frames, with a zero prefix (ZP) appended to each frame [24], [29]. The frame length is assumed to be N and the ZP length is set to $N_c - 1$, with $N > N_c$, the number of discrete time MIMO channel taps. A hybrid architecture is assumed at the transmitter and the receiver as shown in Fig. 2. The use of block transmission with $N_c - 1$ zero padding is important here, since it allows reconfiguring the RF circuits from one frame to the other and avoids loss of training data during this reconfiguration. This also avoids inter-frame interference. Also note that for symboling rate of 1760 MHZ (the chip rate used in IEEE 802.11ad preamble), it is impractical to use different precoders and combiners for different symbols. It is more feasible, however, to reconfigure the RF circuitry for different frames with $N \sim 16 - 512$ symbols.

To formulate the sparse recovery problem, we assume that N_{RF} is the number of RF chains used at the transceivers. For the mth training frame, the transmitter uses an RF precoder $\mathbf{F}_m \in \mathbb{C}^{N_t \times N_{RF}}$, that can be realized using quantized angles at the analog phase shifters. Then, the nth symbol of the mth received frame is

$$\mathbf{r}_m[n] = \sum_{d=0}^{N_c-1} \mathbf{H}_d \mathbf{F}_m \mathbf{s}_m[n-d] + \mathbf{v}_m[n], \tag{8}$$

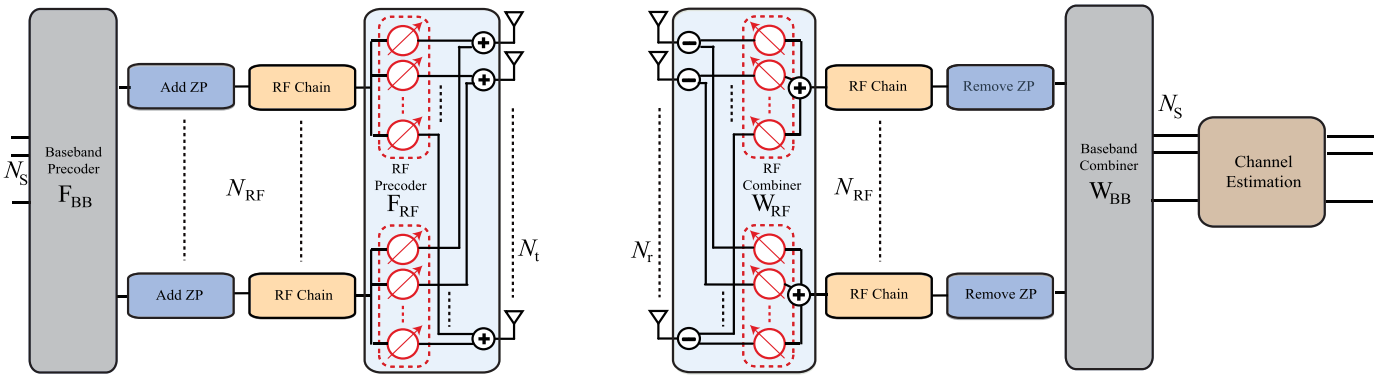


Fig. 2. Figure illustrating the transceiver chains and the frame structure assumed for the time-domain channel estimation of the frequency-selective mmWave system with N_c channel taps. Zero padding (ZP) of length at least $N_c - 1$ is prefixed to the training symbols of length N for RF chain reconfiguration across frames.

where $\mathbf{s}_m[n] \in \mathbb{C}^{N_{RF}\times 1}$ is the nth training data symbol of the mth training frame

$$\mathbf{s}_m = [\underbrace{0 \cdots 0}_{N_c-1} \ \mathbf{s}_m[1] \ \cdots \ \mathbf{s}_m[N]\ ]. \tag{9}$$

Note from (8) that due to multi-path fading the unknown parameters of the channel (contained within $\mathbf{H}_d$, $d = 0, 1, \ldots, N_c - 1$) are all entangled, and potentially also suppressed, due to adding up non-coherently. The proposed training frame structure and the sparsifying dictionary that we propose in the sequel, still expose the channel to enable sparse recovery.

At the receiver, an RF combiner $\mathbf{W}_m \in \mathbb{C}^{N_r\times N_{RF}}$ realized using quantized angles at the analog phase shifters is used during the mth training phase. The post-combining signal is

$$\begin{bmatrix} \mathbf{y}_m^T[1] \\ \mathbf{y}_m^T[2] \\ \vdots \\ \mathbf{y}_m^T[N] \end{bmatrix}^T = \mathbf{W}_m^* \begin{bmatrix} \mathbf{H}_0 & \cdots & \mathbf{H}_{N_c-1} \end{bmatrix} \left(\mathbf{I}_{N_c} \otimes \mathbf{F}_m\right) \mathbf{S}_m^T + \mathbf{e}_m^T \in \mathbb{C}^{1\times NN_{RF}}, \tag{10}$$

where

$$\mathbf{S}_m = \begin{bmatrix} \mathbf{s}_m^T[1] & 0 & \cdots & 0 \\ \mathbf{s}_m^T[2] & \mathbf{s}_m^T[1] & \cdots & . \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{s}_m^T[N] & \cdots & \cdots & \mathbf{s}_m^T[N-N_c+1] \end{bmatrix}, \tag{11}$$

is of dimension $N \times N_c N_{RF}$, and

$$\mathbb{E}\left[\mathbf{e}_m \mathbf{e}_m^*\right] = \sigma^2 \mathbf{I}_N \otimes \mathbf{W}_m^* \mathbf{W}_m. \tag{12}$$

Using the matrix equality $\text{vec}\,(\mathbf{ABC}) = \left(\mathbf{C}^T \otimes \mathbf{A}\right) \text{vec}\,(\mathbf{B})$ and the notation for the vectorized channel in (7), vectorizing (10) gives

$$\mathbf{y}_m = \begin{bmatrix} \mathbf{y}_m[1] \\ \mathbf{y}_m[2] \\ \vdots \\ \mathbf{y}_m[N] \end{bmatrix} = \underbrace{\mathbf{S}_m \left(\mathbf{I}_{N_c} \otimes \mathbf{F}_m^T\right) \otimes \mathbf{W}_m^*}_{\Phi_{td}^{(m)}} \mathbf{h}_c + \mathbf{e}_m. \tag{13}$$

Using the form in (6) and denoting $\gamma_{\ell,d} = \alpha_\ell p(dT_s - \tau_\ell)$, (13) can be expressed as

$$\mathbf{y}_m = \Phi_{td}^{(m)} \left(\mathbf{I}_{N_c} \otimes \bar{\mathbf{A}}_T \circ \mathbf{A}_R\right) \begin{bmatrix} \gamma_{1,0} \\ \vdots \\ \gamma_{N_p,0} \\ \vdots \\ \gamma_{1,(N_c-1)} \\ \vdots \\ \gamma_{N_p,(N_c-1)} \end{bmatrix} + \mathbf{e}_m. \tag{14}$$

In (14), the matrices $\mathbf{A}_T$ and $\mathbf{A}_R$, and the complex gains $\{\alpha_i\}$ and delays $\{\tau_i\}$ contained within $\gamma_{\ell,d}$ are all unknowns that need to be estimated to get the explicit multi-tap MIMO channel. Next, we recover these by estimating the columns of a suitably defined dictionary via sparse recovery.

To formulate the compressed sensing problem in the time-domain, we first exploit the sparse nature of the channel in the angular domain. Accordingly, we define the matrices $\mathbf{A}_{tx}$ and $\mathbf{A}_{rx}$ used for sparse recovery, that can be computed apriori at the receiver. The $N_t \times G_t$ matrix $\mathbf{A}_{tx}$ consists of columns $\mathbf{a}_T(\tilde{\theta}_x)$, with $\tilde{\theta}_x$ drawn from a quantized angle grid of size G_t, and the $N_r \times G_r$ matrix $\mathbf{A}_{rx}$ consists of columns $\mathbf{a}_R(\tilde{\phi}_x)$, with $\tilde{\phi}_x$ drawn from a quantized angle grid of size G_r. Neglecting the grid quantization error, we can then express (13) as

$$\mathbf{y}_m = \Phi_{td}^{(m)} \left(\mathbf{I}_{N_c} \otimes \bar{\mathbf{A}}_{tx} \otimes \mathbf{A}_{rx}\right) \hat{\mathbf{x}}_{td} + \mathbf{e}_m. \tag{15}$$

Note that the actual frequency-selective mmWave channel as seen at the baseband has angles of arrival and departure drawn from $[0, 2\pi)$. The quantization used for constructing the dictionary, when fine enough, can ensure that the dominant AoAs and AoDs are captured as columns of $\bar{\mathbf{A}}_{tx} \otimes \mathbf{A}_{rx}$. The error incurred due to the angle grid quantization is investigated in Section VI, where we assume offgrid values for the AoA/AoD in the simulations. With this, the signal $\hat{\mathbf{x}}_{td}$ consisting of the time-domain channel gains and pulse shaping filter response is more sparse than the unknown vector in (14), and is of size $N_c G_r G_t \times 1$.

Next, the band-limited nature of the sampled pulse shaping filter is used to operate with an unknown channel vector

with a lower sparsity level. For that, we look at the sampled version of the pulse-shaping filter $\mathbf{p}_d$ having entries $p_d(n) = p_{\text{rc}}\left((d - n\frac{N_\text{c}}{G_\text{c}})T_\text{s}\right)$, for $d = 1, 2, \ldots, N_\text{c}$ and $n = 1, 2, \ldots, G_\text{c}$. Then, neglecting the quantization error due to sampling in the delay domain, we can write (15) as

$$\mathbf{y}_m = \Phi_{\text{td}}^{(m)} \left(\mathbf{I}_{N_\text{c}} \otimes \bar{\mathbf{A}}_{\text{tx}} \otimes \mathbf{A}_{\text{rx}}\right) \Gamma \mathbf{x}_{\text{td}} + \mathbf{e}_m, \quad (16)$$

$$\text{where} \quad \Gamma = \begin{bmatrix} \mathbf{I}_{G_\text{r}G_\text{t}} \otimes \mathbf{p}_1^\text{T} \\ \mathbf{I}_{G_\text{r}G_\text{t}} \otimes \mathbf{p}_2^\text{T} \\ \vdots \\ \mathbf{I}_{G_\text{r}G_\text{t}} \otimes \mathbf{p}_{N_\text{c}}^\text{T} \end{bmatrix}, \quad (17)$$

and $\mathbf{x}_{\text{td}} \in \mathbb{C}^{G_\text{c}G_\text{r}G_\text{t} \times 1}$ is the N_p-sparse vector containing the time-domain complex channel gains.

Stacking M such measurements obtained from sending M training frames and using a different RF precoder and combiner for each frame, we have

$$\mathbf{y}_{\text{td}} = \Phi_{\text{td}} \Psi_{\text{td}} \mathbf{x}_{\text{td}} + \mathbf{e}, \quad (18)$$

$$\text{where } \mathbf{y}_{\text{td}} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_M \end{bmatrix} \in \mathbb{C}^{NMN_{\text{RF}} \times 1} \quad (19)$$

is the measured time-domain signal,

$$\mathbb{E}\left[\mathbf{e}\mathbf{e}^*\right] = \sigma^2 \text{diag}\left(\mathbf{I}_N \otimes \mathbf{W}_1^* \mathbf{W}_1, \cdots, \mathbf{I}_N \otimes \mathbf{W}_M^* \mathbf{W}_M\right), \quad (20)$$

$$\Phi_{\text{td}} = \begin{bmatrix} \mathbf{S}_1 \left(\mathbf{I}_{N_\text{c}} \otimes \mathbf{F}_1^\text{T}\right) \otimes \mathbf{W}_1^* \\ \mathbf{S}_2 \left(\mathbf{I}_{N_\text{c}} \otimes \mathbf{F}_2^\text{T}\right) \otimes \mathbf{W}_2^* \\ \vdots \\ \mathbf{S}_M \left(\mathbf{I}_{N_\text{c}} \otimes \mathbf{F}_M^\text{T}\right) \otimes \mathbf{W}_M^* \end{bmatrix} \in \mathbb{C}^{NMN_{\text{RF}} \times N_\text{c}N_\text{r}N_\text{t}} \quad (21)$$

is the time-domain measurement matrix, and

$$\Psi_{\text{td}} = \left(\mathbf{I}_{N_\text{c}} \otimes \bar{\mathbf{A}}_{\text{tx}} \otimes \mathbf{A}_{\text{rx}}\right) \Gamma = \begin{bmatrix} \left(\bar{\mathbf{A}}_{\text{tx}} \otimes \mathbf{A}_{\text{rx}}\right) \otimes \mathbf{p}_1^\text{T} \\ \left(\bar{\mathbf{A}}_{\text{tx}} \otimes \mathbf{A}_{\text{rx}}\right) \otimes \mathbf{p}_2^\text{T} \\ \vdots \\ \left(\bar{\mathbf{A}}_{\text{tx}} \otimes \mathbf{A}_{\text{rx}}\right) \otimes \mathbf{p}_{N_\text{c}}^\text{T} \end{bmatrix} \in \mathbb{C}^{N_\text{c}N_\text{r}N_\text{t} \times G_\text{c}G_\text{r}G_\text{t}} \quad (22)$$

is the dictionary in the time-domain. The beamforming and combining vectors $\mathbf{F}_m$, $\mathbf{W}_m$, $m = 1, 2, \cdots, M$ used for training have the phase angles chosen uniformly at random from the set $\mathcal{A}$ in (3).

B. AoA/AoD and Channel Gain Estimation in the Time-Domain

With the sparse formulation of the mmWave channel estimation problem in (18), compressed sensing tools can be first used to estimate the AoA and AoD. The support of $\mathbf{x}_{\text{td}}$ corresponds to a particular AoA, AoD and path delay, and hence estimating the support of $\mathbf{x}_{\text{td}}$ amounts to estimating a channel path, and the corresponding non-zero value corresponds to the path gain. Note that we can increase or decrease the angle quantization grid sizes G_r and G_t, and the delay domain quantization grid size G_c, used for constructing the time-domain dictionary to minimize the quantization error. As the sensing matrix is known at the receiver, sparse recovery algorithms can be used to estimate the AoA and AoD.

To estimate the support of the sparse vector $\mathbf{x}_{\text{td}}$, we solve the optimization problem

$$\min_{\mathbf{x}_{\text{td}}} \|\mathbf{x}_{\text{td}}\|_1 \quad \text{such that} \quad \|\mathbf{y}_{\text{td}} - \Phi_{\text{td}}\Psi_{\text{td}}\mathbf{x}_{\text{td}}\|_2 \leq \epsilon. \quad (23)$$

We consider Orthogonal Matching Pursuit (OMP) for solving (23), as used previously in [14] and [30]. There are several stopping criteria for OMP that can be used to solve (23). When the sparsity level N_p is known apriori, reaching that level could be used to stop the algorithm. When such information cannot be guessed before hand (which itself is an estimation problem), the residual error falling below a certain threshold is often used to terminate the recursive OMP algorithm. Accordingly, in the presence of noise, a suitable choice for the threshold ϵ is the noise variance. Hence, we assume the noise power as the stopping threshold, i.e., $\epsilon = \mathbb{E}[\mathbf{e}^*\mathbf{e}]$.

It is important to note that the performance of OMP algorithm depends on the properties of sensing matrix, such as mutual coherence and restricted isometry property. In this context, IID Guassian random sensing matrices are beneficial [25]. In a hybrid mmWave system, however, generating IID random sensing matrices is infeasible due to the hardware limitations of the RF precoders and combiners. Considering various hybrid architectures, [13] analyzed the mutual coherence of the sensing matrix for a narrowband sparse formulation. Due to the complicated structure of the sensing matrix, however, only a few special cases are considered in [13]. The wideband sparse formulation does not simplify things any further. Therefore, to understand the performance of the OMP algorithm, we resort to simulations later on.

Following the support estimation via sparse recovery, the channel gains can be estimated. While there are many ways to estimate the gains, even directly from OMP, we only give the details for one approach next – using least squares. The various methods are based on plugging in the columns of the sparsifying time-domain dictionary matrix corresponding to the estimated support. That is, let $\mathcal{S}^{\text{td}}$ be the estimated support set using sparse recovery in the proposed time-domain formulation. Then, using (18), we have

$$\mathbf{y}_{\text{td}} = \underbrace{[\Phi_{\text{td}}\Psi_{\text{td}}]_{:,\mathcal{S}^{\text{td}}}}_{\Omega_{\text{td}}} \hat{\mathbf{x}}_{\text{td}} + \mathbf{e}, \quad (24)$$

so that the channel coefficients via least squares is

$$\hat{\mathbf{x}}_{\text{td}}^{\text{LS}} = \left(\Omega_{\text{td}}^* \Omega_{\text{td}}\right)^{-1} \Omega_{\text{td}}^* \mathbf{y}_{\text{td}}. \quad (25)$$

IV. Frequency-Domain Channel Estimation via Compressed Sensing

In this section, we explain how the compressed sensing problem can be formulated in the frequency-domain. The additional modifications needed in the system model, and the corresponding advantages and disadvantages are also explained in this section.

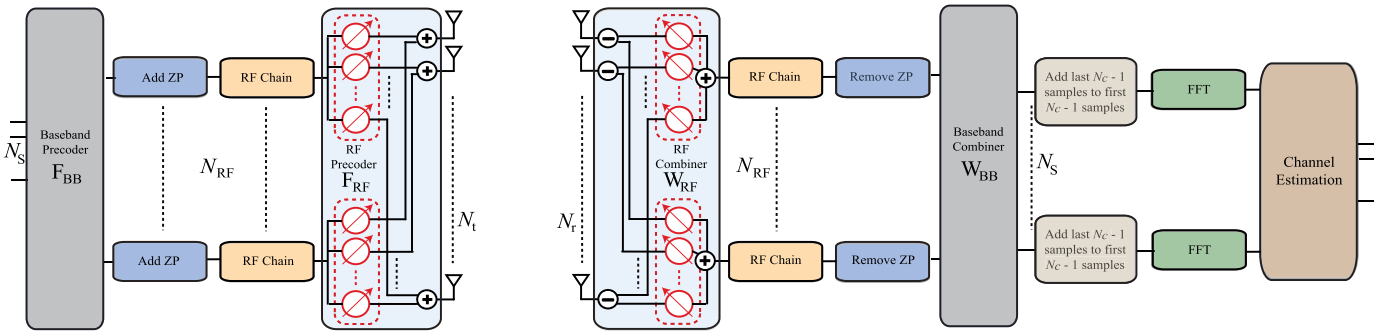


Fig. 3. Figure illustrating the transceiver chains and the frame structure assumed for the frequency-domain channel estimation of the frequency-selective mmWave system with N_c channel taps. Zero padding (ZP) of length $N_c - 1$ is prefixed to the training symbols of length N for RF chain reconfiguration across frames.

Using the geometric channel model in (4), the complex channel matrix in the frequency-domain can be written as

$$\begin{aligned}\boldsymbol{H}[k] &= \sum_{d=0}^{N_c-1} \mathbf{H}_d e^{-\mathrm{j}\frac{2\pi kd}{K}} \\ &= \sum_{\ell=1}^{N_p} \alpha_\ell \mathbf{a}_R(\phi_\ell)\mathbf{a}_T^*(\theta_\ell) \sum_{d=0}^{N_c-1} p(dT_s - \tau_\ell)e^{-\mathrm{j}\frac{2\pi kd}{K}}. \end{aligned} \quad (26)$$

Defining $\beta_{k,\ell} = \sum_{d=0}^{N_c-1} p(dT_s - \tau_\ell)e^{-\mathrm{j}\frac{2\pi kd}{K}}$ a compact expression can be derived

$$\boldsymbol{H}[k] = \sum_{\ell=1}^{N_p} \alpha_\ell \beta_{k,\ell} \mathbf{a}_R(\phi_\ell)\mathbf{a}_T^*(\theta_\ell). \quad (27)$$

Vectorizing (27) gives the unknown signal that is estimated using the frequency-domain estimation algorithm,

$$\mathrm{vec}\left(\boldsymbol{H}[k]\right) = \left(\bar{\mathbf{A}}_T \circ \mathbf{A}_R\right) \begin{bmatrix} \alpha_1\beta_{k,1} \\ \alpha_2\beta_{k,2} \\ \vdots \\ \alpha_{N_p}\beta_{k,N_p} \end{bmatrix}. \quad (28)$$

Note that the vector channel representation of the kth subcarrier in (28), is similar to the time-domain vector representation in (6). The key difference, however, is that, unlike the time-domain approach, each of the unknown vectors corresponding to the K subcarriers can be estimated separately, in parallel as explained next.

A. Sparse Formulation in the Frequency-Domain

For the sparse formulation in the proposed frequency-domain approach, appropriate signal processing is performed to convert the linear convolution occurring during the frame transmission in the system to a circular convolution in the time-domain. We would like to remind the readers that we propose ZP instead of cyclic prefix for the training frames to facilitate reconfiguration of the RF precoders and combiners from frame to frame. With ZP assumed in the frame structure of the training preamble, the overlapping and sum [31] method is used, followed by the K-point FFT to formulate the frequency-domain sparse channel estimation problem per subcarrier $k = 1, 2, \cdots, K$. The overlap and add method essentially involves adding the last $N_c - 1$ samples to first $N_c - 1$ samples as shown in Fig. 3. Fig. 3 also illustrates the proposed system model in the time-domain with the hybrid architecture and signal processing components required for the frequency-domain channel estimation in an SC-FDE system with ZP. Alternatively, for an OFDM based system, the digital processing may be implemented in the frequency-domain after the FFT operation. The advantage of the proposed frequency-domain approach is that different baseband precoders and combiners can be used for different subcarriers [28] in the frequency-domain, while the RF processing is frequency-flat. The proposed frequency-domain approach, therefore, works for both SC-FDE and OFDM systems, where the received signal is processed per-subcarrier. We next look into the received signal in the kth subcarrier.

With $\mathbf{F}_m$ denoting the RF precoder used at the transmitter for the transmission of the mth training frame/OFDM symbol, and $\mathbf{W}_m$, the corresponding RF combiner, the post-combining signal in the kth subcarrier can be written as

$$\check{\boldsymbol{y}}_m[k] = \mathbf{W}_m^* \boldsymbol{H}[k]\mathbf{F}_m \check{\boldsymbol{s}}_m[k] + \check{\boldsymbol{e}}_m[k], \quad (29)$$

$$\text{where} \quad \check{\boldsymbol{s}}_m[k] = \sum_{n=1}^{N} \mathbf{s}_m[n] e^{-\mathrm{j}\frac{2\pi kn}{K}} \quad (30)$$

is the kth coefficient of the K-point FFT of mth time-domain transmit frame. The covariance of the frequency-domain noise vector in (29) is $\mathbb{E}\left[\check{\boldsymbol{e}}_m[k]\check{\boldsymbol{e}}_m^*[k]\right] = \sigma^2 \mathbf{W}_m^* \mathbf{W}_m$, and $\sigma^2 = N_o B$. The frequency flat RF combiners and precoders are assumed to be realized with a network of phase shifters with phase angles drawn from a finite set, as before. Vectorizing (29), and substituting (28) gives

$$\mathrm{vec}\left(\check{\boldsymbol{y}}_m[k]\right) = \underbrace{\left(\check{\boldsymbol{s}}_m^T[k]\mathbf{F}_m^T \otimes \mathbf{W}_m^*\right)}_{\Phi_{\mathrm{fd}}^{(m)}[k]} \mathrm{vec}\left(\boldsymbol{H}[k]\right) + \check{\boldsymbol{e}}_m[k]. \quad (31)$$

Assuming the AoAs and AoDs are drawn from a grid of size G_r and G_t, respectively, and neglecting the quantization error, we can write (31) in terms of the dictionary matrices defined in Section III as follows:

$$\mathrm{vec}\left(\check{\boldsymbol{y}}_m[k]\right) = \Phi_{\mathrm{fd}}^{(m)}[k]\left(\bar{\mathbf{A}}_{tx} \otimes \mathbf{A}_{rx}\right)\check{\boldsymbol{x}}[k] + \check{\boldsymbol{e}}_m[k], \quad (32)$$

with the signal $\check{\boldsymbol{x}}[k] \in \mathbb{C}^{G_r G_t \times 1}$ being N_p-sparse. Stacking M such measurements obtained over the course of M training

frame transmission, each with a different pair of RF precoder and combiner, we have the following sparse formulation for the kth subcarrier:

$$\check{\mathbf{y}}[k] = \Phi_{\text{fd}}[k]\Psi_{\text{fd}}\check{\boldsymbol{x}}[k] + \check{\boldsymbol{e}}[k], \tag{33}$$

in terms of the frequency-domain dictionary $\Psi_{\text{fd}} = \left(\bar{\mathbf{A}}_{\text{tx}} \otimes \mathbf{A}_{\text{rx}}\right) \in \mathbb{C}^{N_{\text{r}}N_{\text{t}} \times G_{\text{r}}G_{\text{t}}}$ and the measurement matrix in the frequency-domain

$$\Phi_{\text{fd}}[k] = \begin{bmatrix} \check{\boldsymbol{s}}_1^{\text{T}}[k]\mathbf{F}_1^{\text{T}} \otimes \mathbf{W}_1^* \\ \check{\boldsymbol{s}}_2^{\text{T}}[k]\mathbf{F}_2^{\text{T}} \otimes \mathbf{W}_2^* \\ \vdots \\ \check{\boldsymbol{s}}_M^{\text{T}}[k]\mathbf{F}_M^{\text{T}} \otimes \mathbf{W}_M^* \end{bmatrix} \in \mathbb{C}^{MN_{\text{RF}} \times N_{\text{r}}N_{\text{t}}}. \tag{34}$$

The covariance of the noise in (33) is

$$\mathbb{E}\left[\check{\boldsymbol{e}}[k]\check{\boldsymbol{e}}^*[k]\right] = \sigma^2 \text{diag}\left(\mathbf{W}_1^*\mathbf{W}_1, \mathbf{W}_2^*\mathbf{W}_2, \cdots, \mathbf{W}_M^*\mathbf{W}_M\right). \tag{35}$$

B. AoA/AoD and Channel Gain Estimation per Subcarrier

As discussed previously in Section III-B, we first estimate the support of $\check{\boldsymbol{x}}[k]$, that corresponds to a particular AoA and AoD, and then proceed to estimate the MIMO channel coefficients of the kth subcarrier, which correspond to the non-zero values of $\check{\boldsymbol{x}}[k]$. As with the time-domain approach, we solve the following optimization problem

$$\min_{\check{\boldsymbol{x}}[k]} \|\check{\boldsymbol{x}}[k]\|_1 \quad \text{such that} \quad \|\check{\mathbf{y}}[k] - \Phi_{\text{fd}}[k]\Psi_{\text{fd}}\check{\boldsymbol{x}}[k]\|_2 \leq \epsilon, \tag{36}$$

via OMP with the stopping threshold $\epsilon = \mathbb{E}[\check{\boldsymbol{e}}[k]^*\check{\boldsymbol{e}}[k]]$, to estimate the support of the sparse vector $\check{\mathbf{y}}[k]$, and hence the dominant angles of arrival and departure. The estimated support set is denoted as $\mathcal{S}^{\text{fd}}$, which corresponds to specific columns of the frequency-domain dictionary Ψ_{fd}. Using $\mathcal{S}^{\text{fd}}$, the channel coefficients, that correspond to the non-zero values of the sparse vector $\check{\boldsymbol{x}}[k]$, can be derived as follows. From (33), after the sparse support recovery

$$\check{\mathbf{y}}[k] = \underbrace{[\Phi_{\text{fd}}[k]\Psi_{\text{fd}}]_{:,\mathcal{S}^{\text{fd}}}}_{\Omega_{\text{fd}}}\check{\boldsymbol{x}}[k] + \check{\boldsymbol{e}}[k], \tag{37}$$

so that, using least square estimation,

$$\check{\boldsymbol{x}}^{\text{LS}}[k] = \left(\Omega_{\text{fd}}^*\Omega_{\text{fd}}\right)^{-1}\Omega_{\text{fd}}^*\check{\mathbf{y}}[\mathbf{k}]. \tag{38}$$

Note that using the sparse formulation in (33), the AoAs/AoDs and the channel coefficients of the the kth subcarrier can be estimated. Repeating the same for all the K subcarriers fully characterizes the frequency-selective mmWave channel. While the dimensions of the matrices involved in the frequency-domain compressed sensing problem is smaller in comparison to the time-domain formulation in Section III, the channel estimation should be invoked K times to fully recover the channel coefficients. Alternatively, frequency-domain correlation could be utilized to estimate the channel in a few subcarriers, and then interpolated to obtain the MIMO channel estimates of all the K subcarriers. In either case, additional pre-processing and FFT operation are required to perform the frequency-domain channel estimation.

V. Combined Time-Frequency Compressive Channel Estimation

In this section, we formulate a technique via compressed sensing for explicit channel estimation, jointly in time and frequency. The key idea is to estimate the angles of arrival and departure via compressed sensing in the frequency-domain, and then use the estimates to evaluate the channel gains and path delays in the time-domain to obtain the entire channel.

The transmitter chain for the proposed combined time-frequency compressive channel estimation approach is the same as in Fig. 2 and Fig. 3. The system model for the receiver chain in Fig. 3 can be employed to perform sparse support recovery of the angles in the frequency-domain for the proposed estimation approach in this section. Following the compressive support estimation, the pre-computed dictionary matrices in the time-domain, and the measurement matrices can be used to estimate the channel coefficients of the frequency-selective mmWave MIMO channel, as explained momentarily.

From (28) and (31), we can express the frequency-domain received signal in the kth subcarrier, in terms of the actual AoAs and AoDs in the vector form as follows

$$\text{vec}\left(\check{\mathbf{y}}_m[k]\right) = \Phi_{\text{fd}}^{(m)}[k]\left(\bar{\mathbf{A}}_{\text{T}} \circ \mathbf{A}_{\text{R}}\right)\begin{bmatrix} \alpha_1\beta_{k,1} \\ \alpha_2\beta_{k,2} \\ \vdots \\ \alpha_{N_{\text{p}}}\beta_{k,N_{\text{p}}} \end{bmatrix} + \check{\boldsymbol{e}}_m[k], \tag{39}$$

with the noise covariance $\mathbb{E}\left[\check{\boldsymbol{e}}_m[k]\check{\boldsymbol{e}}_m^*[k]\right] = \sigma^2\mathbf{W}_m^*\mathbf{W}_m$.

AoA/AoD Estimation in Frequency-Domain and Channel Gain Estimation in Time-Domain: Note from (39), the AoA and AoD information in each subcarrier k is the same, and contained in $\bar{\mathbf{A}}_{\text{T}} \circ \mathbf{A}_{\text{R}}$, whose ℓth column is of the form $\bar{\mathbf{a}}_{\text{T}}(\theta_\ell) \otimes \mathbf{a}_{\text{R}}(\phi_\ell)$. Therefore, using sparse recovery in $1 \leq P \leq K$ number of subcarriers parallely, and using the union of all the estimated angles, we can get a support set of the AoAs, denoted as $\mathcal{S}_{\mathsf{A}}$ and a set of AoD estimates denoted as $\mathcal{S}_{\mathsf{D}}$. One option to estimate the support is to use OMP as explained in Section IV-B, for P subcarriers in parallel. Since several angles might be estimated incorrectly due to sparse recovery under the influence of noise, using P parallel OMPs may not necessarily enhance performance. This is because we retain all the (potentially faulty) angle estimates on the P subcarrier. Prior work in [25], [32], and [33] have studied methods to estimate a common support set from multiple parallel measurements. Such techniques may also be employed to recover the support set containing the AoA/AoD information in the frequency-domain, both here as well as in the proposed approach in Section IV.

Following the support recovery in the frequency-domain, to recover the entire channel, we switch to the time-domain formulation in (15), (18) – (22), but restrict to the set $\mathcal{S} = \{\mathcal{S}_{\mathsf{A}}, \mathcal{S}_{\mathsf{D}}\}$. Accordingly, we can write the effective time-domain equation, conditioned on the estimated support set $\mathcal{S}$ as

$$\mathbf{y}_{\text{td}} = \underbrace{\Phi_{\text{td}}\left[\Psi_{\text{td}}\right]_{\mathcal{S}}}_{\Omega}\mathbf{x}_{\text{td}} + \mathbf{e}, \tag{40}$$

where Φ_{td} is that in (21), the noise covariance of $\mathbf{e}$ is that in (20), and

$$[\Psi_{\text{td}}]_{\mathcal{S}} = \begin{bmatrix} [\bar{\mathbf{A}}_{\text{tx}}]_{:,\mathcal{S}_{\text{D}}} \otimes [\mathbf{A}_{\text{rx}}]_{:,\mathcal{S}_{\text{A}}} \otimes \mathbf{p}_1^T \\ [\bar{\mathbf{A}}_{\text{tx}}]_{:,\mathcal{S}_{\text{D}}} \otimes [\mathbf{A}_{\text{rx}}]_{:,\mathcal{S}_{\text{A}}} \otimes \mathbf{p}_2^T \\ \vdots \\ [\bar{\mathbf{A}}_{\text{tx}}]_{:,\mathcal{S}_{\text{D}}} \otimes [\mathbf{A}_{\text{rx}}]_{:,\mathcal{S}_{\text{A}}} \otimes \mathbf{p}_{N_{\text{c}}}^T \end{bmatrix} \in \mathbb{C}^{N_{\text{c}}N_{\text{r}}N_{\text{t}} \times G_{\text{c}}|\mathcal{S}|} \tag{41}$$

is the dictionary matrix conditioned on the knowledge of the support set. The unknown $\mathbf{x}_{\text{td}}$ in (40), contains channel coefficients in the time-domain, which can now be obtained via least squares or MMSE to recover the entire MIMO channel matrices corresponding to all the delay taps. That is, from (40)

$$\mathbf{x}_{\text{td}}^{\text{LS}} = \left(\Omega^* \Omega\right)^{-1} \Omega^* \mathbf{y}_{\text{td}}. \tag{42}$$

The advantage of using the combined time-frequency approach for the wideband channel estimation is twofold. First, since the sparse recovery is done in the frequency-domain, the sizes of the measurement matrix and the dictionary are $MN_{\text{RF}} \times N_{\text{r}}N_{\text{t}}$ and $N_{\text{r}}N_{\text{t}} \times G_{\text{r}}G_{\text{t}}$, respectively, that are much smaller than the corresponding time-domain matrices Φ_{td} and Ψ_{td}. Secondly, unlike the frequency-domain approach, the channel estimates need not be separately evaluated per subcarrier, but only once in the time-domain, thus further reducing the computation complexity.

VI. Simulation Results

In this section, the performance of the three proposed channel estimation algorithm are provided. For the compressed sensing estimation of the angles of arrival and departure, orthogonal matching pursuit is used. The channel gains are then estimated using least squares.

We assume uniform linear array (ULA) with half wavelength antenna element separation for the simulations. For such a ULA,

$$\mathbf{a}_{\text{R}}(\phi_\ell) = \frac{1}{\sqrt{N_{\text{r}}}}\left[1 \quad e^{\text{j}\pi\cos(\phi_\ell)} \quad \cdots \quad e^{\text{j}(N_{\text{r}}-1)\pi\cos(\phi_\ell)}\right]^T,$$

and

$$\mathbf{a}_{\text{T}}(\theta_\ell) = \frac{1}{\sqrt{N_{\text{t}}}}\left[1 \quad e^{\text{j}\pi\cos(\theta_\ell)} \quad \cdots \quad e^{\text{j}(N_{\text{t}}-1)\pi\cos(\theta_\ell)}\right]^T.$$

The AoA and AoD quantization used for constructing the transmitter and receiver dictionary matrices are taken from an angle grid of size G_{r} and G_{t}, respectively. This implies that the ℓth column of $\mathbf{A}_{\text{tx}}$ is $\mathbf{a}_{\text{T}}(\tilde{\theta}_\ell)$, where $\tilde{\theta}_\ell = \frac{(\ell-1)\pi}{G_{\text{t}}}$ and the kth column of $\mathbf{A}_{\text{rx}}$ is $\mathbf{a}_{\text{R}}(\tilde{\phi}_k)$, where $\tilde{\phi}_k = \frac{(k-1)\pi}{G_{\text{r}}}$. The angle quantization used in the phase shifters is assumed to have N_{Q} quantization bits, so that the entries of $\mathbf{F}_m$, $\mathbf{W}_m$, $m = 1, 2, \cdots, M$ are drawn from $\mathcal{A}$, as defined in (3), with equal probability. The N_{p} paths of the wideband mmWave channel are assumed to be independently and identically distributed, with delay τ_ℓ chosen uniformly at random from $[0, (N_{\text{c}}-1)T_{\text{s}}]$, where T_{s} is the sampling interval and N_{c} is the number of delay taps of the channel. The angles of arrival and departure for each of the channel paths are assumed to be distributed independently and uniformly in $[0, \pi]$. The raised cosine pulse shaping signal is assumed to have a roll-off factor of 0.8.

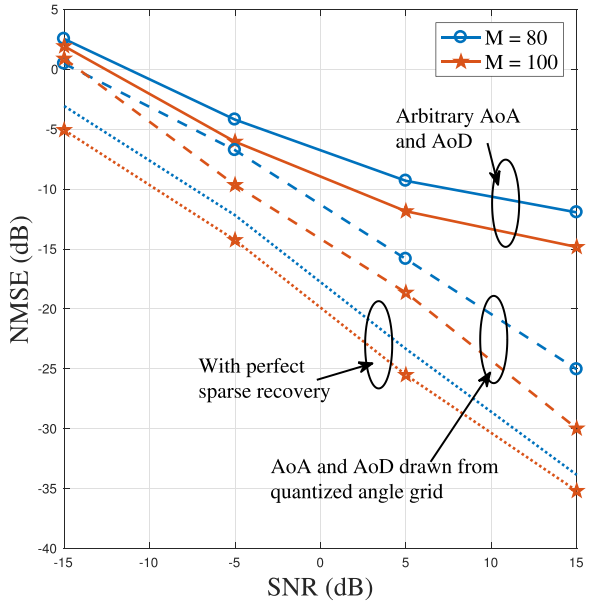


Fig. 4. Average NMSE as a function of SNR for different training length M when $N_s = 1$ and $N_{\text{RF}} = 1$ using the proposed time-domain channel estimation technique. We assume $N = 16$ symbols per frame for a frequency-selective channel of 4 taps. Using the proposed approach, training length of $80 - 100$ is sufficient to ensure very low estimation error, processing completely in the time-domain.

Let $\hat{\mathbf{h}}_{\text{c}} \in \mathbb{C}^{N_{\text{c}}N_{\text{r}}N_{\text{t}} \times 1}$ denote the estimated channel vector. We use the following metrics to compare the performance of our proposed channel estimation algorithms:

1) the normalized mean squared error (NMSE) of the channel estimates defined as

$$\text{NMSE} = \frac{\|\mathbf{h}_{\text{c}} - \hat{\mathbf{h}}_{\text{c}}\|_2^2}{\|\mathbf{h}_{\text{c}}\|_2^2} = \frac{\sum_{d=0}^{N_{\text{c}}-1} ||\mathbf{H}_d - \hat{\mathbf{H}}_d||_F^2}{\sum_{d=0}^{N_{\text{c}}-1} ||\mathbf{H}_d||_F^2}. \tag{43}$$

2) the ergodic spectral efficiency as defined in [27].

Fig. 4 shows the NMSE as a function of the post-combining received signal SNR using the proposed time-domain channel estimation approach. Here we assume $N_{\text{r}} = 16$, $N_{\text{t}} = 32$, $N_{\text{c}} = 4$, $N = 16$ and $N_{\text{p}} = 2$. The time-domain dictionary is constructed with the parameters $G_{\text{r}} = 32$, $G_{\text{t}} = 64$, and $G_{\text{c}} = 8$. From Fig. 4, it can be seen that with training length of even $80 - 100$ frames, sufficiently low channel estimation error can be ensured. For comparing the recovery performance of the sensing matrix using OMP, we plot the NMSE for the case where the sparse recovery is perfect (i.e., the exact angles and delays are provided by a genie). In Fig. 4, we also provide the NMSE when the AoA and AoD of the mmWave channel are drawn from quantized grids with $G_{\text{t}} = 64$ and $G_{\text{r}} = 32$ that are used to construct the dictionaries. The gap between the genie-aided perfect sparse recovery curve and that obtained when the parameters fall exactly on the grid elements captures the performance of the recovery algorithm used. Future work could try to narrow this gap using better recovery algorithms.

Fig. 5 shows how employing multiple RF chains at the transmitter and receiver can give good improvement in the

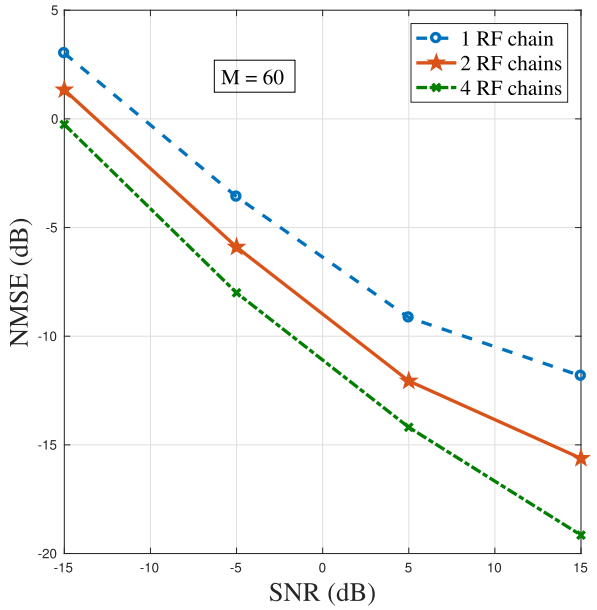


Fig. 5. Average NMSE for the proposed time-domain channel estimation approach as a function of SNR for different numbers of RF chains used at the transceivers. By employing multiple RF chains at the transceivers, the NMSE performance is improved.

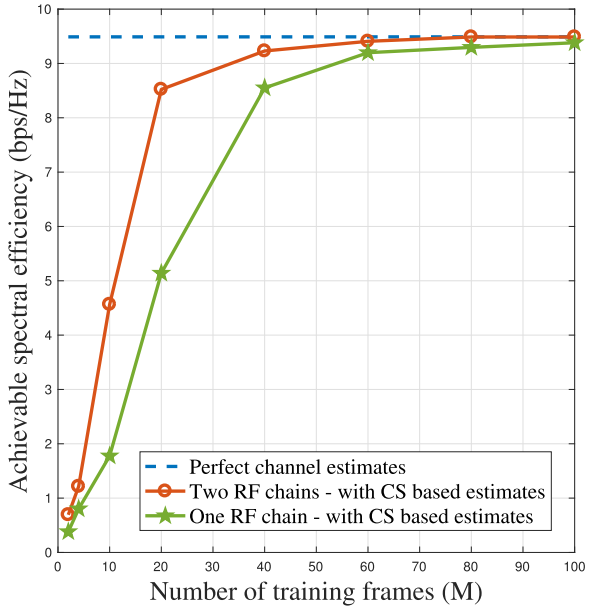


Fig. 6. Achievable spectral efficiency using the proposed time-domain channel estimation approach as a function of the number of training frames used M for different numbers of RF combiners N_{RF} used at the receiver. Employing multiple RF chains at the transceivers significantly reduces the number of training steps.

estimation performance while requiring fewer number of training frame transmissions. We assume the same set of parameters as those used for generating Fig. 4, and the proposed time-domain channel estimation approach, while altering the number of RF chains used at the transceivers. In Fig. 5, we assume $M = 60$ frames are transmitted for training. The improvement in NMSE performance occurs thanks to a larger number of measurements per training frame, resulting in smaller estimation error via compressed sensing. Further, employing multiple RF chains also increases the number of effective combining/precoding beam patterns that scale with N_{RF}. So, larger N_{RF} is preferred to decrease the estimation error and to fully leverage the hybrid architecture in wideband mmWave systems.

In both Fig. 4 and Fig. 5, we considered averaged NMSE to highlight the effectiveness of the proposed time-domain channel estimation algorithm and the performance gain when multiple RF chains are employed at the transmitter and receiver. In Fig. 6, the achievable spectral efficiency is plotted as a function of number of training steps M. We assume the same set of parameters as that in Fig. 4, and the ergodic spectral efficiency computation used in [27] for SNR = 5 dB. It is observed that having more RF combiners results in fewer number of training frame transmissions to achieve the same spectral efficiency. This also shows that the NMSE levels in Fig. 4 and Fig. 5 are sufficient to obtain achievable spectral efficiency provided by an all-digital system. Note that we compare all-digital spectral efficiency results in Fig. 6 because it is known that optimal frequency-selective hybrid precoder and combiner, designed based on the MIMO channel knowledge can achieve spectral efficiency similar to that guaranteed by an all-digital system [28].

In Fig. 4, we see that an error gap exists between the cases where the AoA/AoD fall within the quantized grid and when

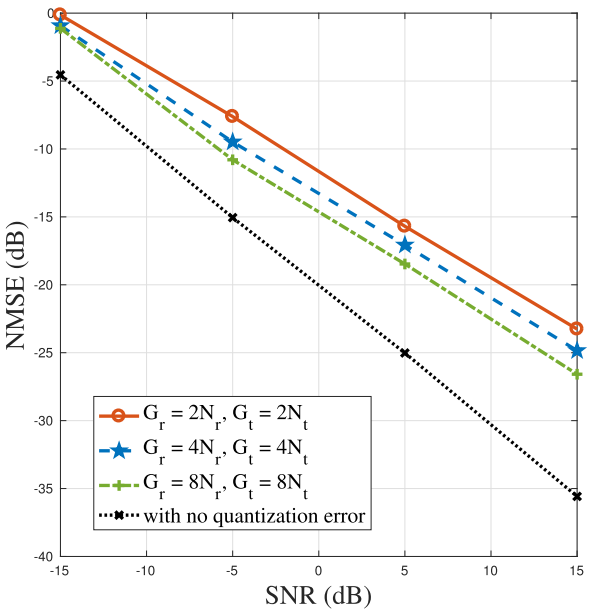


Fig. 7. Average NMSE versus SNR (assuming $N_{\mathrm{r}} = N_{\mathrm{t}} = 16$) using the proposed frequency-domain channel estimation approach. The number of angles in the quantized grid used for generating the dictionary is denoted as G_{r} (for AoA) and G_{t} (for AoD). The figure shows how increasing G_{t} and G_{r} reduces the grid quantization error for a given antenna size.

the AoA/AoD are arbitrary. This gap is especially enhanced at higher SNR levels. Intuitively, thanks to our flexible and generic sparsifying dictionary construction, choosing larger values for G_{r} (G_{t}) in comparison with N_{r} (N_{t}) can further narrow the error gap between the two cases. Similarly, increasing G_{c} in comparison to N_{c} also helps, as the dictionary will become more and more *robust*. This capability of our dictionary design is shown in Fig. 7 where we plot the NMSE

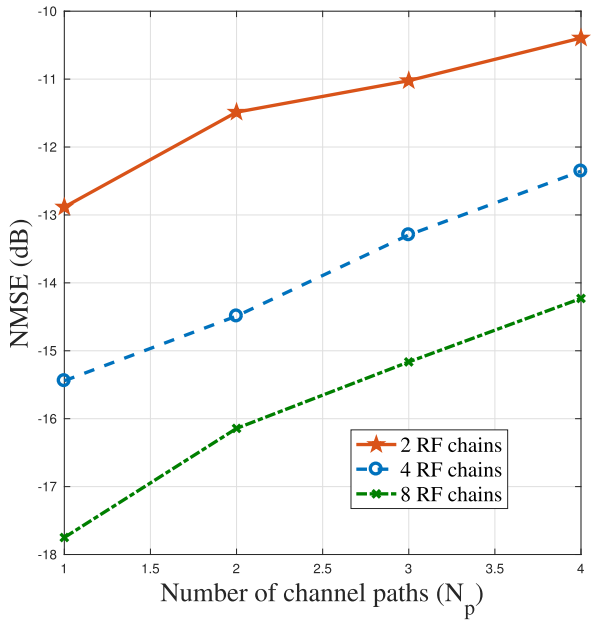


Fig. 8. Average NMSE versus the number of paths N_p, for different hybrid configurations at the transceivers using the proposed combined time-frequency domain channel estimation approach. Increasing N_p increases the number of unknown parameters of the channel, and hence higher number of compressive measurements are required to get the required target estimation error performance.

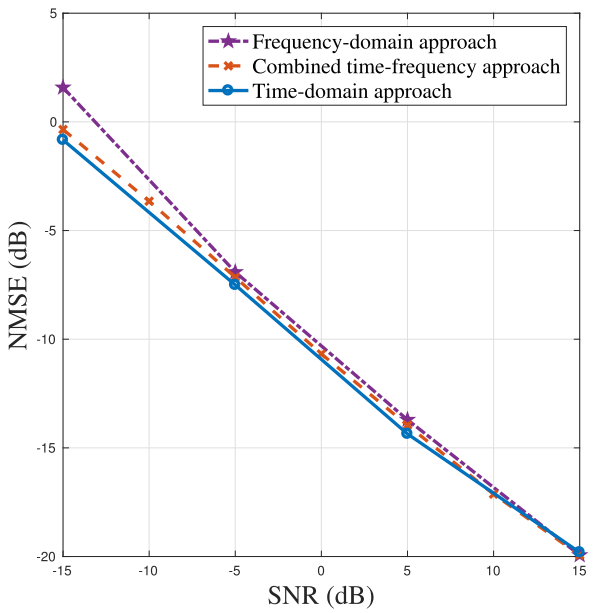


Fig. 9. Plot showing the error performance of the three compressed sensing based channel estimation approaches proposed in the paper as a function of SNR. At low SNR the time-domain approach has the least average NMSE, while at higher SNRs, all the three proposed approaches give similar performance.

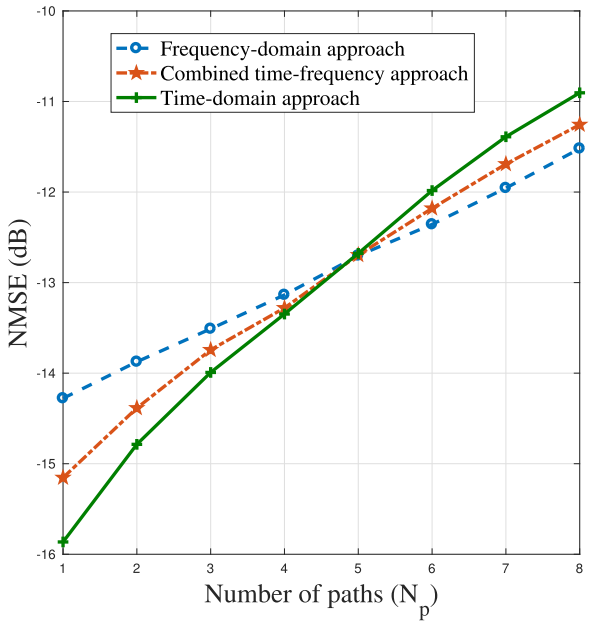


Fig. 10. Plot showing the error performance of the three proposed approaches, as a function of the number of paths N_p in the channel. Increasing N_p degrades the average NMSE performance. While the proposed time-domain approach gives the minimum average NMSE when the number of paths is small, the frequency-domain approach gives the best error performance for larger N_p.

for different values of the grid sizes along with the NMSE obtained when there is no error due to angle quantization. For plotting this baseline curve, we assume the exact angles are provided by a genie, similar to how we plot the reference curve in Fig. 4.

In Fig. 8, we plot the NMSE as a function of the number of paths in the channel for various RF combiner setups at the receiver. The proposed combined time-frequency domain channel estimation approach is used with $N_r = N_t = 32$, $G_r = G_t = 64$, $N = 32$, $N_c = 8$ and $M = 60$ compressive training steps. In Fig. 8, we use sparse recovery in $P = 1$ subcarrier in the frequency-domain to estimate the AoAs and AoDs, before switching to the time-domain to estimate all the gain coefficients of the channel paths. As N_p is increased, the number of unknown parameters in the channel increases, thus increasing the estimation error for a given number of training steps and hardware configuration. Increasing the number of RF combiners, however, helps reduce the NMSE to meet a target estimation error performance.

In Fig. 9, we look the error performance of the three proposed channel estimation approaches by plotting the NMSE as a function of SNR. We assume $N_r = N_t = 32$, $G_r = G_t = 64$, $M = 60$, $N_c = 4$ and $N_p = 2$. The number of RF chains at the transceivers N_{RF} is assumed to be 4. The combined time-frequency approach is assumed to use OMP on $P = 1$ subcarrier in the frequency-domain to recover the angles of arrival and departure. For constructing the time-domain dictionary, we assume $G_c = 2N_c$ delay quantization parameter. It can be seen that the time-domain and combined time-frequency approaches give the best error performance whereas the proposed frequency-domain approach results is large estimation error, especially at lower SNRs. This is mainly due to the accumulation of error incurred due to K parallel OMPs in the frequency, which is avoided in the combined time-frequency approach and the proposed time-domain approach, which invoke the sparse recovery algorithm only once (when $P = 1$). At higher SNRs, however, the three proposed approaches give similar estimation error performance.

In Fig. 10, we compare the three proposed approaches' error performance as the number of channel paths is increased.

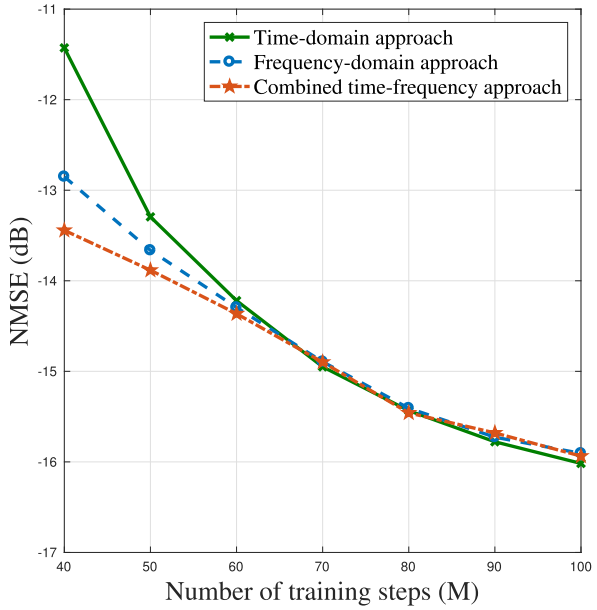


Fig. 11. Plot showing the error performance of the three proposed approaches, as a function of the number of training steps M. More number of compressive measurements lead to better estimation error performance at the expense of higher signaling overhead. The combined time-frequency approach gives the best trade-off between low training overhead and minimum average NMSE performance.

We set $G_{\mathrm{r}} = G_{\mathrm{t}} = 2N_{\mathrm{r}} = 2N_{\mathrm{t}} = 64$, and assume 4 RF chains at the transceivers. The training frame length of $N = 16$ is assumed for the wideband channel of $N_{\mathrm{c}} = 4$ delay taps. For each case, $M = 60$ training steps are assumed. As the number of channel paths is increased, all the approaches perform worse. In particular, for a given delay quantization parameter $G_{\mathrm{c}} = 2N_{\mathrm{c}}$, assumed for the time-domain approach's plot in Fig. 10, the degradation in NMSE is significant as the delay estimation results in more error. For smaller number of channel paths, however, the time-domain approach gives lower channel estimation error. The plot also shows that when the number of delay taps is smaller compared to the number of paths in the channel, frequency-domain techniques perform better.

A comparison of the performance of the three proposed approaches as a function of the number of training steps is shown in Fig. 11. We set the SNR to 5 dB here and assume $G_{\mathrm{r}} = G_{\mathrm{t}} = 2N_{\mathrm{r}} = 2N_{\mathrm{t}} = 64$. Each training frame is assumed to be of length 16 symbols, for a frequency-selective mmWave channel of tap length 4, and channel paths 2. The NMSE plots in Fig. 11 is assumed 4 RF chains at the transceivers with 2 bit quantization at the phase shifters during channel estimation. It can be seen that while, with a few training steps the combined time-frequency approach and the proposed frequency-domain approach outperform the time-domain approach, with larger number of training steps, the time-domain approach gives the least NMSE.

To compare and comment on the overhead in channel training in the proposed compressive sensing based approaches, consider the short preamble structure used in IEEE 802.11ad [10], which is of duration 1.891μs. At a chip rate of 1760 MHz, this short preamble consisting of the short training frame (STF) and the channel estimation frame (CEF) amounts to more than 3200 symbols. After the end of this short preamble transmission, IEEE 802.11ad beamforming protocol then switches to a different beam pair combination, and the process is repeated recursively to estimate the best set of beamforming directions. For the setting in Fig. 11, however, $MN = 1600$ symbols are only required for the proposed approaches to achieve low average NMSE and explicit estimation of the frequency-selective MIMO channel.

VII. Conclusion

In this paper, we proposed wideband channel estimation algorithms for frequency-selective mmWave systems using a hybrid architecture at the transmitter and receiver. The system model adopts zero padding that allows enough time for switching the analog beams and, hence, well matches the hybrid architectures. The proposed channel estimation algorithms are based on sparse recovery and can support MIMO operation in mmWave systems since the entire channel is estimated after the beam training phase. Three different approaches - in purely time, in purely frequency and a combined time-frequency approach were proposed, that can be used in both SC-FDE and OFDM based wideband mmWave systems. Leveraging the frame structure and the hybrid architecture at the transceivers, it was shown that compressed sensing tools can be used for mmWave channel estimation. Simulation results showed that the proposed algorithms required very few training frames to ensure low estimation error. It was shown that further reduction in the training overhead and estimation error can be obtained by employing multiple RF chains at the transceivers.

References

[1] K. Venugopal, A. Alkhateeb, R. W. Heath, and N. González Prelcic, "Time-domain channel estimation for wideband millimeter wave systems with hybrid architecture," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Sep. 2016, pp. 6493–6497.

[2] R. W. Heath, N. González-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, Apr. 2016.

[3] J. Singh, S. Ponnuru, and U. Madhow, "Multi-Gigabit communication: The ADC bottleneck," in *Proc. IEEE Int. Conf. Ultra-Wideband (ICUWB)*, Vancouver, BC, Canada, Sep. 2009, pp. 22–27.

[4] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, Jr., "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[5] A. Alkhateeb, J. Mo, N. González-Prelcic, and R. W. Heath, "MIMO precoding and combining solutions for millimeter-wave systems," *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 122–131, Dec. 2014.

[6] J. Wang *et al.*, "Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 8, pp. 1390–1399, Oct. 2009.

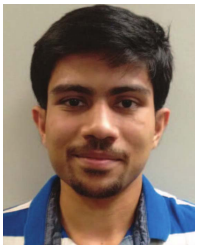
[7] L. Chen, Y. Yang, X. Chen, and W. Wang, "Multi-stage beamforming codebook for 60 GHz WPAN," in *Proc. 6th Int. ICST Conf. Commun. Netw. China (CHINACOM)*, Aug. 2011, pp. 361–365.

[8] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4391–4403, Oct. 2013.

[9] Y. M. Tsang, A. S. Y. Poon, and S. Addepalli, "Coding the beams: Improving beamforming training in mmWave communication system," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Houston, TX, USA, Dec. 2011, pp. 1–6.

[10] E. Perahia, C. Cordeiro, M. Park, and L. L. Yang, "IEEE 802.11ad: Defining the next generation multi-Gbps Wi-Fi," in *Proc. 7th IEEE Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2010, pp. 1–5.

[11] *IEEE 802.15 WPAN Millimeter Wave Alternative PHY Task Group 3C (TG3c)*. [Online]. Available: http://www.ieee802.org/15/pub/TG3c contributions.html
[12] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, Jr., "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.
[13] R. Méndez-Rial, C. Rusu, N. González-Prelcic, A. Alkhateeb, and R. W. Heath, "Hybrid MIMO architectures for millimeter wave communications: Phase shifters or switches?" *IEEE Access*, vol. 4, pp. 247–267, 2016.
[14] A. Alkhateeb, G. Leusz, and R. W. Heath, "Compressed sensing based multi-user millimeter wave systems: How many measurements are needed?" in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Brisbane, QLD, Australia, Apr. 2015, pp. 2909–2913.
[15] Y. Han and J. Lee, "Two-stage compressed sensing for millimeter wave channel estimation," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2016, pp. 860–864.
[16] J. Lee, G.-T. Gil, and Y. H. Lee, "Exploiting spatial sparsity for estimating channels of hybrid MIMO systems in millimeter wave communications," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec. 2014, pp. 3326–3331.
[17] M. A. Iwen and A. H. Tewfik, "Adaptive strategies for target detection and localization in noisy environments," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2344–2353, May 2012.
[18] M. L. Malloy and R. D. Nowak, "Near-optimal compressive binary search," *CoRR*, Mar. 2012. [Online]. Available: http://arxiv.org/abs/1203.1804
[19] D. Ramasamy, S. Venkateswaran, and U. Madhow, "Compressive adaptation of large steerable arrays," in *Proc. Inf. Theory Appl. Workshop (ITA)*, San Diego, CA, USA, Feb. 2012, pp. 234–239.
[20] D. E. Berraki, S. M. D. Armour, and A. R. Nix, "Application of compressive sensing in sparse spatial channel recovery for beamforming in mmWave outdoor systems," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2014, pp. 887–892.
[21] H. Ghauch, T. Kim, M. Bengtsson, and M. Skoglund, "Subspace estimation and decomposition for large millimeter-wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 528–542, Apr. 2015.
[22] M. Kokshoorn, P. Wang, Y. Li, and B. Vucetic, "Fast channel estimation for millimetre wave wireless systems using overlapped beam patterns," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 1304–1309.
[23] D. Zhu, J. Choi, and R. W. Heath, Jr. (2016). "Auxiliary beam pair enabled AoD and AoA estimation in closed-loop large-scale mmWave mimo systems." [Online]. Available: https://arxiv.org/abs/1610.05587
[24] A. Ghosh *et al.*, "Millimeter-wave enhanced local area systems: A high-data-rate approach for future wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1152–1163, Jun. 2014.
[25] Z. Gao, C. Hu, L. Dai, and Z. Wang, "Channel estimation for millimeter-wave massive MIMO with hybrid precoding over frequency-selective fading channels," *IEEE Commun. Lett.*, vol. 20, no. 6, pp. 1259–1262, Jun. 2016.
[26] B. Gao, Z. Xiao, L. Su, Z. Chen, D. Jin, and L. Zeng, "Multi-device multi-path beamforming training for 60-GHz millimeter-wave communications," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 1328–1333.
[27] P. Schniter and A. Sayeed, "Channel estimation and precoder design for millimeter-wave communications: The sparse way," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Nov. 2014, pp. 273–277.
[28] A. Alkhateeb and R. W. Heath, Jr., "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1801–1818, May 2016.
[29] Z. Wang, X. Ma, and G. B. Giannakis, "OFDM or single-carrier block transmissions?" *IEEE Trans. Commun.*, vol. 52, no. 3, pp. 380–394, Mar. 2004.
[30] G. Taubböck, F. Hlawatsch, D. Eiwen, and H. Rauhut, "Compressive estimation of doubly selective channels in multicarrier systems: Leakage effects and sparsity-enhancing processing," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 255–271, Apr. 2010.
[31] Y. Chen, J. Zhang, and A. D. S. Jayalath, "New training sequence structure for zero-padded SC-FDE system in presence of carrier frequency offset," in *Proc. IEEE 68th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2008, pp. 1–4.
[32] Z. Gao, L. Dai, W. Dai, B. Shim, and Z. Wang, "Structured compressive sensing-based spatio-temporal joint channel estimation for FDD massive MIMO," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 601–617, Feb. 2016.
[33] B. Wang, L. Dai, T. Mir, and Z. Wang, "Joint user activity and data detection based on structured compressive sensing for NOMA," *IEEE Commun. Lett.*, vol. 20, no. 7, pp. 1473–1476, Jul. 2016.



Kiran Venugopal (S'15) received the B.Tech. degree in electronics and communication engineering from the National Institute of Technology, Calicut, India, in 2011, and the M.E. degree in telecommunication engineering from the Indian Institute of Science, Bangalore, India, in 2013. He is currently pursuing the Ph.D. degree with the Wireless Networking and Communications Group, Department of Electrical and Computer Engineering, The University of Texas at Austin, USA. From 2013 to 2014, he was a Systems Design Staff Engineer with Broadcom Communication Technologies Pvt., Ltd., Bangalore. He is currently involved in the analysis and physical layer development of wideband mmWave systems. His research interests include areas in wireless communication and signal processing. He received the Prof. S. V. C. Aiya Medal for the best M.E. (Telecommunications) student from the Indian Institute of Science. In 2011, he received the Er. M. L. Bapna Gold Medal from the National Institute of Technology, for scoring the highest CGPA among all the under graduate programs.

Ahmed Alkhateeb (S'08–M'17) received the B.S. (Hons.) and M.S. degrees from Cairo University, Egypt, in 2008 and 2012, respectively, and the Ph.D. degree in electrical and computer engineering from The University of Texas at Austin in 2016. He is currently a Wireless Communications Researcher with Facebook, CA, USA. His research interests are in the broad area of communication theory, applied math, and signal processing. In the context of wireless communication, his interests include mmWave communication and massive MIMO systems. He was a recipient of the 2012–2013 UT Austin's MCD Fellowship and the 2016 IEEE Signal Processing Society Young Author Best Paper Award.

Nuria González Prelcic has held visiting positions at The University of New Mexico in 2011 and The University of Texas at Austin from 2014 to 2016. She has been the Founder Director of the Atlantic Research Center for Information and Communication Technologies, University of Vigo, since 2010. She is currently an Associate Professor with the Signal Theory and Communications Department, University of Vigo, Spain. She has co-authored around 20 journal and conference papers in the topic of signal processing for mmWave communications since 2014, including a tutorial published in the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING. Her main research interests include signal processing theory and signal processing for wireless communications: filter banks, compressive sampling and estimation, multicarrier modulation, channel estimation, and MIMO processing for millimeter wave communications. She is also serving as a Guest Editor for the forthcoming special issue of this journal on signal processing for mmWave wireless communications.

Robert W. Heath, Jr. (S'96–M'01–SM'06–F'11) received the B.S. and M.S. degrees from the University of Virginia, Charlottesville, VA, USA, in 1996 and 1997, respectively, and the Ph.D. degree from Stanford University, Stanford, CA, USA, in 2002, all in electrical engineering. From 1998 to 2001, he was a Senior Member of the Technical Staff and then a Senior Consultant with Iospan Wireless Inc, San Jose, CA, USA, where he was involved in the design and implementation of the physical and link layers of the first commercial MIMO-OFDM communication system. Since 2002, he has been with the Department of Electrical and Computer Engineering, The University of Texas at Austin, where he is currently a Cullen Trust for Higher Education Endowed Professor, and is a member of the Wireless Networking and Communications Group. He is also the President and the CEO of MIMO Wireless Inc. He authored *Introduction to Wireless Digital Communication* (Prentice Hall, 2017) and *Digital Wireless Communication: Physical Layer Exploration Lab Using the NI USRP* (National Technology and Science Press, 2012), and co-authored *Millimeter Wave Wireless Communications* (Prentice Hall, 2014).

Dr. Heath has been a Co-Author of 15 award winning conference and journal papers, including recently the 2010 and 2013 *EURASIP Journal on Wireless Communications and Networking* Best Paper Awards, the 2012 *Signal Processing Magazine* Best Paper Award, the 2013 Signal Processing Society Best Paper Award, the 2014 *EURASIP Journal on Advances in Signal Processing* Best Paper Award, the 2014 *Journal of Communications and Networks* Best Paper Award, the 2016 IEEE Communications Society Fred W. Ellersick Prize, and the 2016 IEEE Communications and Information Theory Societies Joint Paper Award. He received the 2017 EURASIP Technical Achievement Award. He was a Distinguished Lecturer in the IEEE Signal Processing Society. He is an ISI Highly Cited Researcher. He is also an elected member of the Board of Governors for the IEEE Signal Processing Society, a licensed Amateur Radio Operator, a Private Pilot, and a registered Professional Engineer in Texas.