TIME-DOMAIN CHANNEL ESTIMATION FOR WIDEBAND MILLIMETER WAVE SYSTEMS WITH HYBRID ARCHITECTURE

Kiran Venugopal[†]*, Ahmed Alkhateeb*[†]*, Robert W. Heath, Jr.*[†] *and Nuria González Prelcic* [‡]

[†]The University of Texas, Austin, TX, USA, Email: {kiranv, aalkhateeb, rheath}@utexas.edu
[‡]The University of Vigo, Spain, Email: nuria@gts.uvigo.es

ABSTRACT

Millimeter wave (mmWave) systems will likely employ large antennas at both the transmitter and receiver for directional beamforming. Hybrid analog/digital MIMO architectures have been proposed previously for leveraging both array gain and multiplexing gain, while reducing the power consumption in analog-to-digital converters. Channel knowledge is needed to design the hybrid precoders/combiners, which is difficult to obtain due to the large antenna arrays and the frequency selective nature of the channel. In this paper, we propose a sparse recovery based time-domain channel estimation technique for hybrid architecture based frequency selective mmWave systems. The proposed compressed sensing channel estimation algorithm is shown to provide good estimation error performance, while requiring small training overhead. The simulation results show that using multiple RF chains at the receiver and the transmitter further reduces the training overhead.

Index Terms— Millimeter wave communication, channel estimation, hybrid architecture, compressed sensing.

1. INTRODUCTION

Millimeter wave (mmWave) based communication is a key ingredient of 5G wireless systems for realizing gigabits-per-second data rates [1–4]. To provide sufficient received power in mmWave systems, large antenna arrays need to be deployed at both the transmitter and the receiver [3,5]. Developing low-overhead mmWave channel estimation techniques is crucial for acquiring channel state information used to design precoding matrices in mmWave MIMO [6] and mmWave massive MIMO systems [7, 8]. This is, however, complicated due to hardware constraints and low signal-to-noise-ratios (SNR) without beamforming in the antenna front-end [9].

MmWave channel estimation via analog beam training [10, 11] works for both narrowband and wideband systems, but only supports a single communication stream. Channel estimation and precoding using hybrid architectures have been studied in [6, 12, 13]. While fully digital channel estimation [14] is infeasible at mmWave due to hardware constraints [9], prior work assuming hybrid architecture considered mainly narrowband mmWave systems. In practice, however, mmWave channels will be wideband and frequency selective. In [15], hybrid precoding for frequency selective mmWave systems was studied, though the channel estimation techniques were not elaborated, especially for single-carrier modulation which is a promising solution for mmWave [16]. In [7], wideband mmWave channel estimation assuming hybrid architecture and an OFDM system with ideal settings was considered.

In this paper, we propose a time-domain wideband channel estimation technique that works for both MIMO and massive MIMO mmWave systems using a hybrid architecture. The time-domain channel estimation approach is beneficial in single-carrier mmWave systems. The proposed technique can be used to enable MIMO and multi-user communication. The primary focuses of the paper is to formulate the wideband channel estimation in large antenna systems as a sparse recovery problem, leveraging the structure in the frequency selective mmWave channel while considering the hardware constraints at mmWave. These system constraints include the frame structure, finite bandwidth of the pulse shaping filter used at the transceivers, and the hybrid architecture. It is shown through simulation results that the proposed algorithm requires significantly fewer training steps to reliably estimate the channel by utilizing multiple RF chains at the transceivers.

Notation: In this paper, $||\mathbf{A}||_F$ 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 Khatri-Rao product of $\mathbf{A}$ and $\mathbf{B}$ is denoted as $\mathbf{A} \circ \mathbf{B}$ is, and $\mathbf{A} \otimes \mathbf{B}$ is their Kronecker product.

2. SYSTEM MODEL

The algorithm proposed in the next section can be used in both single-user and multi-user MIMO systems when the downlink channel is estimated independently at every user side. Due to the lack of space, we only develop the signal model for a single-user mmWave MIMO system with a transmitter having

This work was supported in part by the Intel-Verizon 5G research program and the National Science Foundation under Grant No. NSF-CCF-1319556.

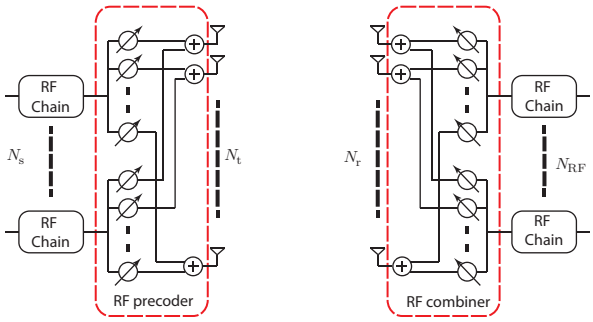

Fig. 1. Figure illustrating the transmitter and receiver structure assumed in the paper. The RF precoder and the combiner are implemented using a network of phase shifters.

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. In the time domain, the transmitter uses a hybrid precoder [15, 17] $\mathbf{F} = \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}} \in \mathbb{C}^{N_{\mathrm{t}} \times N_{\mathrm{s}}}$, N_{s} being the number of data streams that can be transmitted. Since we focus on time-domain channel estimation in this paper, note that we dropped the subcarrier index usually used with frequency selective hybrid precoders and combiners, and assume frequency flat beamforming for simplicity. Denoting the symbol vector at instance n as $\mathbf{s}[n] \in \mathbb{C}^{N_{\mathrm{s}} \times 1}$, satisfying $\mathbb{E}[\mathbf{s}[n]\mathbf{s}^*[n]] = \frac{1}{N_{\mathrm{s}}}\mathbf{I}$, the signal transmitted at discrete-time n is $\tilde{\mathbf{s}}[n] = \mathbf{F}\mathbf{s}[n]$.

The $N_{\mathrm{r}} \times N_{\mathrm{t}}$ 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 $\mathbf{H}_d,\ d = 0,\ 1,\ ...,\ N_{\mathrm{c}} - 1$. With ρ denoting the average received power and $\mathbf{v}[n] \sim \mathcal{N}\left(0, \sigma^2\mathbf{I}\right)$ denoting the circularly symmetric complex Gaussian distributed additive noise vector, the received signal is

$$\mathbf{r}[n] \quad = \quad \sqrt{\rho} \sum_{d=0}^{N_{\mathrm{c}}-1} \mathbf{H}_d\mathbf{F}\mathbf{s}[n-d] + \mathbf{v}[n]. \qquad (1)$$

In the time-domain, the receiver applies a hybrid combiner $\mathbf{W} = \mathbf{W}_{\mathrm{RF}}\mathbf{W}_{\mathrm{BB}} \in \mathbb{C}^{N_{\mathrm{r}} \times N_{\mathrm{RF}}}$ so that the post combining signal at the receiver is

$$\mathbf{y}[n] = \sqrt{\rho} \sum_{d=0}^{N_{\mathrm{c}}-1} \mathbf{W}^*\mathbf{H}_d\mathbf{F}\mathbf{s}[n-d] + \mathbf{W}^*\mathbf{v}[n]. \qquad (2)$$

A similar expression for the received signal at every user can be obtained if a mmWave massive MIMO system is considered and pilot contamination is neglected.

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

$$\mathcal{A} = \left\{0,\ \frac{2\pi}{2^{N_{\mathrm{Q}}}},\ \cdots,\ \frac{\left(2^{N_{\mathrm{Q}}} - 1\right)2\pi}{2^{N_{\mathrm{Q}}}}\right\} \qquad (3)$$

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

3. CHANNEL ESTIMATION VIA COMPRESSED SENSING

In this section, we present our channel estimation algorithm that leverages the sparse channel structure at mmWave.

3.1. Frequency Selective Channel Model

Consider a geometric channel model [12, 18] for the frequency selective mmWave channel consisting of L scattering clusters. The dth delay tap of the channel can be expressed as

$$\mathbf{H}_d \quad = \quad \sum_{\ell=1}^{L} \alpha_\ell p_{\mathrm{rc}}(dT_s - \tau_\ell)\mathbf{a}_{\mathrm{R}}(\phi_\ell)\mathbf{a}_{\mathrm{T}}^*(\theta_\ell), \qquad (4)$$

where $p_{\mathrm{rc}}(\tau)$ denotes the raised cosine pulse signal evaluated at τ, $\alpha_\ell \in \mathbb{C}$ is the complex gain of the ℓth cluster, $\tau_\ell \in \mathbb{R}$ is the delay of the ℓth cluster, ϕ_ℓ and θ_ℓ are the angles of arrival and departure (AoA/AoD), respectively of the ℓth cluster, and $\mathbf{a}_{\mathrm{R}}(\phi_\ell) \in \mathbb{C}^{N_{\mathrm{r}} \times 1}$ and $\mathbf{a}_{\mathrm{T}}(\theta_\ell) \in \mathbb{C}^{N_{\mathrm{t}} \times 1}$ denote the antenna array response vectors of the receiver and transmitter, respectively.

The channel model in (4) can be written compactly as

$$\mathbf{H}_d \quad = \quad \mathbf{A}_{\mathrm{R}}\mathbf{\Delta}_d\mathbf{A}_{\mathrm{T}}^*, \qquad (5)$$

where $\mathbf{\Delta}_d \in \mathbb{C}^{L \times L}$ is diagonal with non-zero complex entries, and $\mathbf{A}_{\mathrm{R}} \in \mathbb{C}^{N_{\mathrm{r}} \times L}$ and $\mathbf{A}_{\mathrm{T}} \in \mathbb{C}^{N_{\mathrm{t}} \times L}$ contain the columns $\mathbf{a}_{\mathrm{R}}(\phi_\ell)$ and $\mathbf{a}_{\mathrm{T}}(\theta_\ell)$, respectively. Under this notation, vectorizing the channel matrix in (5) gives

$$\mathrm{vec}(\mathbf{H}_d) = \left(\bar{\mathbf{A}}_{\mathrm{T}} \circ \mathbf{A}_{\mathrm{R}}\right) \begin{bmatrix} \alpha_1 p_{\mathrm{rc}}(dT_{\mathrm{s}} - \tau_1) \\ \alpha_2 p_{\mathrm{rc}}(dT_{\mathrm{s}} - \tau_2) \\ \vdots \\ \alpha_L p_{\mathrm{rc}}(dT_{\mathrm{s}} - \tau_L) \end{bmatrix}. \qquad (6)$$

The vector form of the channel in (6) is useful for the sparse formulation that is presented next. Note that the ℓth column of $\bar{\mathbf{A}}_{\mathrm{T}} \circ \mathbf{A}_{\mathrm{R}}$ is of the form $\bar{\mathbf{a}}_{\mathrm{T}}(\theta_\ell) \otimes \mathbf{a}_{\mathrm{R}}(\phi_\ell)$.

3.2. Sparse Formulation

We assume block transmission with zero padding (ZP) appended to each transmitted frame. To formulate the sparse recovery problem, we assume single RF chains are used both the transmitter and the receiver for the ease of exposition. The formulation extends directly for multiple RF chains at the transmitter and the receiver. We assume that for the training stage the digital precoder and combiner are identity matrices. Accordingly, for the mth training frame the transmitter uses

an RF precoder $\mathbf{f}_{\mathrm{RF}}^{(m)}$ that can be realized using quantized angles at the analog phase shifters. The nth symbol of the mth received frame is

$$\mathbf{r}_m[n] \quad = \quad \sum_{d=0}^{N_{\mathrm{c}}-1} \mathbf{H}_d \mathbf{f}_{\mathrm{RF}}^{(m)} s_m[n-d] + \mathbf{v}_m[n], \qquad (7)$$

where $s_m[n]$ is the nth non-zero symbol of the mth training frame

$$\mathbf{s}_m = [\, \underbrace{0 \; \cdots \; 0}_{N_{\mathrm{c}}-1} \; s_m[1] \; \cdots \; s_m[N] \,]. \qquad (8)$$

At the receiver, an RF combiner $\mathbf{w}_{\mathrm{RF}}^{(m)}$ is used during the mth training phase, so that the post combining signal is

$$\begin{bmatrix} y_m[1] \\ y_m[2] \\ \vdots \\ y_m[N] \end{bmatrix}^T = \mathbf{w}_{\mathrm{RF}}^{(m)*} \begin{bmatrix} \mathbf{H}_0 & \cdots & \mathbf{H}_{N_{\mathrm{c}}-1} \end{bmatrix} \left(\mathbf{S}_m^T \otimes \mathbf{f}_{\mathrm{RF}}^{(m)} \right) + \mathbf{e}^{(m)T}, \qquad (9)$$

where

$$\mathbf{S}_m = \begin{bmatrix} s_m[1] & 0 & \cdots & 0 \\ s_m[2] & s_m[1] & \cdots & . \\ \vdots & \vdots & \ddots & \vdots \\ s_m[N] & \cdots & \cdots & s_m[N-N_{\mathrm{c}}+1] \end{bmatrix}.$$

The use of block transmission with $N_{\mathrm{c}}-1$ zero padding is important here, since it allows for RF circuits reconfiguration from one frame to the other. It also avoids loss of training data during this reconfiguration and inter frame interference. Also note that for the high symbol rates at mmWave (for example, the chip rate used in IEEE 802.11ad preamble is 1760 MHz), it is impractical to use different precoders and combiners for different symbols. Vectorizing (9) gives

$$\mathbf{y}_m = \underbrace{\left(\mathbf{S}_m \otimes \mathbf{f}_{\mathrm{RF}}^{(m)T} \otimes \mathbf{w}_{\mathrm{RF}}^{(m)*} \right)}_{\mathbf{\Phi}_m} \begin{bmatrix} \mathrm{vec}(\mathbf{H}_0) \\ \mathrm{vec}(\mathbf{H}_1) \\ \vdots \\ \mathrm{vec}(\mathbf{H}_{N_{\mathrm{c}}-1}) \end{bmatrix} + \mathbf{e}_m. \qquad (10)$$

To formulate the compressed sensing problem we first exploit the sparse nature of the channel in the angular domain. Assuming the AoAs and AoDs are drawn from an angle grid on G_{r} and G_{t}, respectively and neglecting the grid quantization error, we can then express (10) as

$$\mathbf{y}_m = \mathbf{\Phi}_m \left(\mathbf{I}_{N_{\mathrm{c}}} \otimes \bar{\mathbf{A}}_{\mathrm{tx}} \otimes \mathbf{A}_{\mathrm{rx}} \right) \hat{\mathbf{x}} + \mathbf{e}_m, \qquad (11)$$

where $\mathbf{A}_{\mathrm{tx}}$ and $\mathbf{A}_{\mathrm{rx}}$ are the dictionary matrices used for sparse recovery. The $N_{\mathrm{t}} \times G_{\mathrm{t}}$ matrix $\mathbf{A}_{\mathrm{tx}}$ consists of columns $\mathbf{a}_{\mathrm{T}}(\tilde{\theta}_x)$, with $\tilde{\theta}_x$ drawn from a quantized angle grid of size G_{t}, and the $N_{\mathrm{r}} \times G_{\mathrm{r}}$ matrix $\mathbf{A}_{\mathrm{rx}}$ consists of columns $\mathbf{a}_{\mathrm{R}}(\tilde{\phi}_x)$, with $\tilde{\phi}_x$ drawn from a quantized angle grid of size G_{r}. The signal $\hat{\mathbf{x}}$ consists of the channel gains and pulse shaping filter response, and is of size $N_{\mathrm{c}} G_{\mathrm{r}} G_{\mathrm{t}} \times 1$.

Next, the band-limited nature of the sampled pulse shaping filter is used to effectively operate with an unknown channel vector of lower sparsity level. Accordingly, we define the sampled version of the pulse-shaping filter $\mathbf{p}_d$ having entries $p_d(n) = p_{\mathrm{rc}}((d-n)T_{\mathrm{s}})$, for $d = 0, 1, \cdots, N_{\mathrm{c}}-1$ and $n = 1, 2, \cdots, G_{\mathrm{c}}$. Neglecting the quantization error due to sampling in the delay domain, we can then write (11) as

$$\mathbf{y}_m = \mathbf{\Phi}_m \left(\mathbf{I}_{N_{\mathrm{c}}} \otimes \bar{\mathbf{A}}_{\mathrm{tx}} \otimes \mathbf{A}_{\mathrm{rx}} \right) \mathbf{\Gamma} \mathbf{x} + \mathbf{e}_m, \qquad (12)$$

$$\text{where} \qquad \mathbf{\Gamma} = \begin{bmatrix} \mathbf{I}_{G_{\mathrm{r}}G_{\mathrm{t}}} \otimes \mathbf{p}_0^T \\ \mathbf{I}_{G_{\mathrm{r}}G_{\mathrm{t}}} \otimes \mathbf{p}_1^T \\ \vdots \\ \mathbf{I}_{G_{\mathrm{r}}G_{\mathrm{t}}} \otimes \mathbf{p}_{N_{\mathrm{c}}-1}^T \end{bmatrix}, \qquad (13)$$

and $\mathbf{x}$ is $G_{\mathrm{c}} G_{\mathrm{r}} G_{\mathrm{t}} \times 1$ sparse vector containing the complex channel gains.

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

$$\mathbf{y} = \mathbf{\Phi}\mathbf{\Psi}\mathbf{x} + \mathbf{e}, \qquad (14)$$

where $\mathbf{y} = \left[\mathbf{y}_1^T, \mathbf{y}_2^T, ..., \mathbf{y}_M^T \right]^T \in \mathbb{C}^{NM \times 1}$ is the measured signal,

$$\mathbf{\Phi} = \begin{bmatrix} \mathbf{S}_1 \otimes \mathbf{f}_{\mathrm{RF}}^{(1)T} \otimes \mathbf{w}_{\mathrm{RF}}^{(1)*} \\ \mathbf{S}_2 \otimes \mathbf{f}_{\mathrm{RF}}^{(2)T} \otimes \mathbf{w}_{\mathrm{RF}}^{(2)*} \\ \vdots \\ \mathbf{S}_M \otimes \mathbf{f}_{\mathrm{RF}}^{(M)T} \otimes \mathbf{w}_{\mathrm{RF}}^{(M)*} \end{bmatrix} \in \mathbb{C}^{NM \times N_{\mathrm{c}} N_{\mathrm{r}} N_{\mathrm{t}}} \qquad (15)$$

is the measurement matrix, and

$$\mathbf{\Psi} = \left(\mathbf{I}_{N_{\mathrm{c}}} \otimes \bar{\mathbf{A}}_{\mathrm{tx}} \otimes \mathbf{A}_{\mathrm{rx}} \right) \mathbf{\Gamma} \qquad (16)$$

$$= \begin{bmatrix} \left(\bar{\mathbf{A}}_{\mathrm{tx}} \otimes \mathbf{A}_{\mathrm{rx}} \right) \otimes \mathbf{p}_0^T \\ \left(\bar{\mathbf{A}}_{\mathrm{tx}} \otimes \mathbf{A}_{\mathrm{rx}} \right) \otimes \mathbf{p}_1^T \\ \vdots \\ \left(\bar{\mathbf{A}}_{\mathrm{tx}} \otimes \mathbf{A}_{\mathrm{rx}} \right) \otimes \mathbf{p}_{N_{\mathrm{c}}-1}^T \end{bmatrix} \in \mathbb{C}^{N_{\mathrm{c}} N_{\mathrm{r}} N_{\mathrm{t}} \times G_{\mathrm{c}} G_{\mathrm{r}} G_{\mathrm{t}}} \qquad (17)$$

is the dictionary. Note that the size of the dictionary dictates the complexity of the sparse recovery algorithm and is independent of M.

AoA/AoD estimation With the sparse formulation of the mmWave channel estimation problem in (14), compressed sensing tools can be first used to estimate the AoA and AoD. Note that we can increase or decrease G_{r}, G_{t} and G_{c} to meet the required level of sparsity. As the sensing matrix is known at the receiver, sparse recovery algorithms can be used to estimate the AoA and AoD. Following this, the channel gains can be estimated to minimize the mean squared error or via least squares by plugging in the columns of the dictionary matrices corresponding to the estimated AoA and AoD.

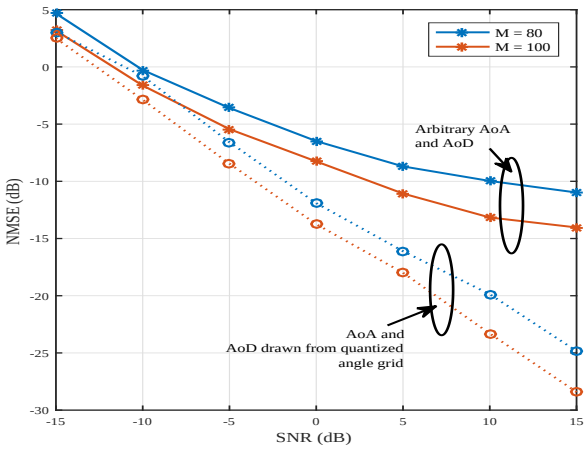

Fig. 2. Normalized mean squared error (NMSE) as a function of SNR for different training length M when $N_{\mathrm{RF}} = 1$. Using the proposed approach, training length of $80-100$ is sufficient to ensure very low estimation error

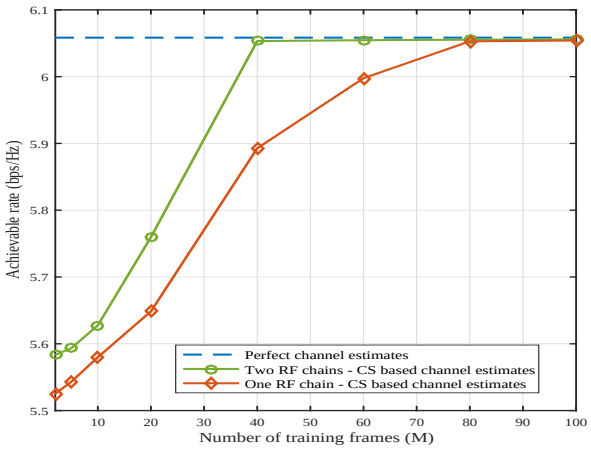

Fig. 3. Achievable spectral efficiency for SNR = 0 dB and $N_{\mathrm{s}} = 1$, as a function of the number of training steps M for different numbers of RF chains N_{RF} used at the transceivers.

4. SIMULATION RESULTS

In this section, the performance of the proposed channel estimation algorithm is provided for the single-user scenario. We consider a system with $N_{\mathrm{t}} = 32$ transmitter antennas and $N_{\mathrm{r}} = 32$ receiver antennas for illustration. Uniform linear arrays with half wavelength separation are assumed. The AoA and AoD quantization used for construction of the transmitter and receiver dictionary matrices are taken to be $G_{\mathrm{r}} = 64$ and $G_{\mathrm{t}} = 64$. The angle quantization used in the phase shifters is assumed to have $N_{\mathrm{Q}} = 2$ quantization bits so that the entries of the RF precoders and combiners are drawn from $\{1, -1, \mathrm{j}, -\mathrm{j}\}$ with equal probability. The frame length is assumed to be $N = 16$ and the tap length of the frequency selective channel is assumed to be uniform $N_{\mathrm{c}} = 4$ for illustration. The scattering cluster centers are assumed to be independently and uniformly distributed with delay $\tau_\ell \in [0, (N_{\mathrm{c}} - 1)T_{\mathrm{s}}]$, and angles ϕ_ℓ and θ_ℓ from $(0, \pi)$. The raised cosine pulse shaping signal is assumed to have a roll-off factor of 0.8.

Fig. 2 shows the normalized mean squared error (NMSE) of the channel estimates as a function of the post combining received signal SNR. Here we define NMSE as

$$\texttt{NMSE} = \frac{\sum_{d=0}^{N_{\mathrm{c}}} ||\mathbf{H}_d - \hat{\mathbf{H}}_d||_F^2}{\sum_{d=0}^{N_{\mathrm{c}}} ||\mathbf{H}_d||_F^2} \tag{18}$$

for comparing the effectiveness of our proposed channel estimation algorithm. From Fig. 2, it can be seen that with training length of even $80 - 100$ frames, sufficiently low channel estimation error can be ensured. Increasing the training length M leads to smaller NMSE, while the size of the dictionary in this setup is a constant 4096×32768, independent of M. For comparing the impact of angle quantization error, we show the NMSE for the case when the AoAs/AoDs are drawn from quantized grids and also the case when the AoAs/AoDs are unrestricted. 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. More elaborate plots showing the performance gains with different dictionary sizes and complexity are provided in the extended version of this paper [19].

Fig. 3 shows the achievable rates using the channel estimates, for different number of RF chains at the transceivers. In Fig. 3, we compute rates as in [18], without water filling. The improvement in rate performance, even with smaller training steps occurs thanks to a larger number of effective measurements per training sent, that scales with the number of RF combiners N_{RF} at the receiver. Similarly, employing multiple RF chains at the transmitter contributes to a larger set of random precoders, resulting in smaller estimation error via compressed sensing. So, a larger N_{RF} is preferred to decrease the training overhead and to fully leverage the hybrid architecture in wideband mmWave systems.

5. CONCLUSION

In this paper, we proposed a time-domain channel estimation algorithm for frequency selective mmWave systems using hybrid architecture at the transmitter and receiver. The proposed channel training protocol can be used to support single-carrier MIMO operation in systems like IEEE 802.11ad, since the entire channel is estimated after the beam training phase. Simulation results showed that the proposed algorithm required very few training frames to ensure low estimation error, and further reduction can be obtained by using multiple RF chains at the transmitter and the receiver.

6. REFERENCES

[1] F. Boccardi *et al*, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, pp. 74–80, Feb. 2014.

[2] J. Andrews *et al*, "What will 5G be?," *IEEE J. Sel. Areas Commun.*, vol. 32, pp. 1065–1082, June 2014.

[3] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, vol. 49, pp. 101–107, June 2011.

[4] T. Rappaport *et al*, "Millimeter wave mobile communications for 5G cellular: It will work!," *IEEE Access*, vol. 1, pp. 335–349, May 2013.

[5] W. Roh *et al*, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results," *IEEE Commun. Mag.*, vol. 52, pp. 106–113, Feb 2014.

[6] A. Alkhateeb, G. Leus, and R. W. Heath, "Compressed sensing based multi-user millimeter wave systems: How many measurements are needed?," in *Proc. IEEE Int. Conf. Acoustics, Speech and Sig. Process. (ICASSP)*, pp. 2909–2913, April 2015.

[7] Z. Gao, L. Dai, and Z. Wang, "Channel estimation for mmwave massive MIMO based access and backhaul in ultra-dense network," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, pp. 1–6, May 2016.

[8] T. E. Bogale and L. B. Le, "Massive MIMO and mmwave for 5G wireless hetnet: Potential benefits and challenges," *IEEE Veh. Technol. Mag.*, vol. 11, pp. 64–75, March 2016.

[9] R. W. Heath, N. Gonzlez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 10, pp. 436–453, April 2016.

[10] J. Wang, "Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems," *IEEE J. Sel. Areas Commun.*, vol. 27, pp. 1390–1399, Oct 2009.

[11] 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, pp. 4391–4403, October 2013.

[12] A. Alkhateeb, O. E. 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, pp. 831–846, Oct 2014.

[13] R. M. Rial, C. Rusu, A. Alkhateeb, N. G. Prelcic, and R. W. Heath Jr., "Hybrid MIMO architectures for millimeter wave communications: Phase shifters or switches?," *IEEE Access*, Jan 2016.

[14] S. Kashyap, C. Mollen, E. Bjornson, and E. G. Larsson, "Frequency-domain interpolation of the zero-forcing matrix in massive MIMO-OFDM," in *IEEE Int. Workshop Signal Process. Advances Wireless Commun. (SPAWC)*, pp. 1–5, July 2016.

[15] A. Alkhateeb and R. W. Heath Jr., "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Trans. Commun.*, vol. 64, pp. 1801–1818, May 2016.

[16] 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, pp. 1152–1163, June 2014.

[17] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath Jr., "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Commun.*, vol. 13, pp. 1499–1513, Mar. 2014.

[18] P. Schniter and A. Sayeed, "Channel estimation and precoder design for millimeter-wave communications: The sparse way," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, pp. 273–277, Nov 2014.

[19] K. Venugopal, A. Alkhateeb, N. G. Prelcic, and R. W. Heath Jr., "Channel estimation for hybrid architecture based wideband millimeter wave systems," *CoRR*, vol. abs/1611.03046, 2016. `http://arxiv.org/abs/1611.03046`.