

# COMPRESSED SENSING BASED MULTI-USER MILLIMETER WAVE SYSTEMS: HOW MANY MEASUREMENTS ARE NEEDED?

Ahmed Alkhateeb<sup>†</sup>, Geert Leus<sup>‡</sup>, and Robert W. Heath Jr.<sup>†</sup>

<sup>†</sup> The University of Texas at Austin, TX, USA, Email: {aalkhateeb, rheath},@utexas.edu

<sup>‡</sup> Delft University of Technology, The Netherlands, Email: g.j.t.leus@tudelft.nl

## ABSTRACT

Millimeter wave (mmWave) systems will likely employ directional beamforming with large antenna arrays at both the transmitters and receivers. Acquiring channel knowledge to design these beamformers, however, is challenging due to the large antenna arrays and small signal-to-noise ratio before beamforming. In this paper, we propose and evaluate a downlink system operation for multi-user mmWave systems based on compressed sensing channel estimation and conjugate analog beamforming. Adopting the achievable sum-rate as a performance metric, we show how many compressed sensing measurements are needed to approach the perfect channel knowledge performance. The results illustrate that the proposed algorithm requires an order of magnitude less training overhead compared with traditional lower-frequency solutions, while employing mmWave-suitable hardware. They also show that the number of measurements need to be optimized to handle the trade-off between the channel estimate quality and the training overhead.

**Index Terms**— Millimeter wave communication, compressed sensing, achievable rates.

## 1. INTRODUCTION

Millimeter wave (mmWave) communication is a promising technology for future cellular systems [1–4]. Directional precoding with large antenna arrays appears to be inevitable to support longer outdoor links and to provide sufficient received signal power. The design of precoding matrices, though, is usually based on complete channel state information, which is difficult to achieve in mmWave due to (i) the large number of antennas, (ii) the small signal-to-noise ratio (SNR) before beamforming, and (iii) the different hardware constraints which impacts the training signal design [5]. Therefore, developing low-complexity mmWave channel estimation techniques is crucial for the mmWave system operation.

To overcome the hardware limitations, analog beamforming solutions were proposed in [6–10], which rely on control-

ling the phase of the signal transmitted by each antenna via a network of analog phase shifters. To avoid the need for explicit channel knowledge, the algorithms in [6–10] depend on beam training which iteratively designs the analog beamforming coefficients at the transmitter and receiver. These solutions, however, have two main disadvantages: (i) they support only single-stream transmissions (2) their training overhead scales linearly with the number of users. To support multi-stream transmission, [5] proposes an efficient mmWave channel estimation algorithm by leveraging the sparse nature of the channel and *adaptive* compressed sensing tools. This solution, however, did not solve the overhead scaling issue.

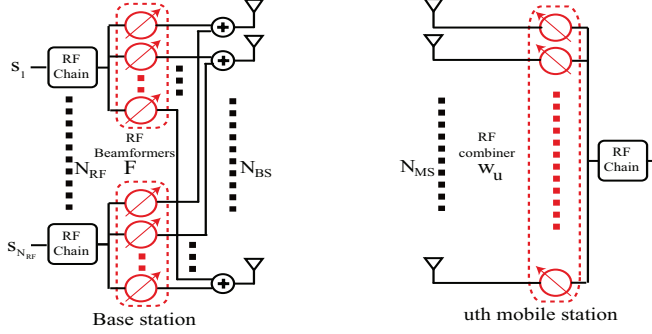
In this paper, we propose a simple downlink system operation for multi-user mmWave systems based on compressed sensing channel estimation and conjugate analog beamforming. Contrary to prior work in [5–10], the training overhead of the proposed channel estimation solution does not scale with the number of users. Hence, it is of special interest for multi-user mmWave systems. Under certain assumptions, we characterize a lower bound on the achievable rate of the proposed system operation as a function of the compressed sensing measurements. Simulation results show that the proposed algorithm requires much less training overhead compared with traditional lower-frequency solutions, while employing mmWave-suitable hardware.

We use the following notation:  $\mathbf{A}$  is a matrix,  $\mathbf{a}$  is a vector, and  $a$  is a scalar.  $\|\mathbf{A}\|_F$  is the Frobenius norm of  $\mathbf{A}$ , whereas  $\mathbf{A}^T$ ,  $\mathbf{A}^*$ ,  $\mathbf{A}^{-1}$ , are its transpose, Hermitian, and inverse, respectively.  $\mathbb{E}[\cdot]$  denotes expectation and  $\mathbb{1}_{\{\cdot\}}$  is an indicator.

## 2. SYSTEM MODEL

Consider a mmWave system with a base station (BS) having  $N_{\text{BS}}$  antennas and  $N_{\text{RF}}$  RF chains as shown in Fig. 1. The BS is assumed to communicate with  $U$  mobile stations (MS's), and each MS is equipped with  $N_{\text{MS}}$  antennas. We focus on the multi-user beamforming case in which the BS communicates with every MS via *only one stream*. Further, we assume that the maximum number of users that can be simultaneously served by the BS equals the number of BS RF chains, i.e.,  $U \leq N_{\text{RF}}$ . This is motivated by the spatial mul-

This work is supported in part by the National Science Foundation under Grant No. 1218338 and 1319556, and by a gift from Huawei Technologies, Inc.



**Fig. 1.** A BS with RF beamformers and  $N_{\text{RF}}$  RF chains communicating with the  $u$ th MS that employs RF combining.

timeplexing gain of the described multi-user precoding system, which is limited by  $\min(N_{\text{RF}}, U)$  for  $N_{\text{BS}} > N_{\text{RF}}$ . For simplicity, we will also assume that the BS will use  $U$  out of the  $N_{\text{RF}}$  available RF chains to serve the  $U$  users.

On the downlink, the BS applies an  $N_{\text{BS}} \times U$  RF precoder,  $\mathbf{F} = [\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_U]$ . The sampled transmitted signal is therefore  $\mathbf{x} = \mathbf{F}\mathbf{s}$ , where  $\mathbf{s} = [s_1, s_2, \dots, s_U]^T$  is the  $U \times 1$  vector of transmitted symbols, such that  $\mathbb{E}[\mathbf{s}\mathbf{s}^*] = \frac{P_T}{U} \mathbf{I}_U$ , and  $P_T$  is the average total transmitted power. Since  $\mathbf{F}$  is implemented using quantized analog phase shifters,  $[\mathbf{F}]_{m,n} = \frac{1}{\sqrt{N_{\text{BS}}}} e^{j\phi_{m,n}}$ , where  $\phi_{m,n}$  is a quantized angle, and the factor of  $\frac{1}{\sqrt{N_{\text{BS}}}}$  is for power normalization.

For simplicity, we adopt a narrowband block-fading channel model [5, 11, 12], by which the  $u$ th MS receives the signal

$$\mathbf{r}_u = \mathbf{H}_u \sum_{r=1}^U \mathbf{f}_r s_r + \mathbf{n}_u, \quad (1)$$

where  $\mathbf{H}_u$  is the  $N_{\text{MS}} \times N_{\text{BS}}$  matrix that represents the mmWave channel between the BS and the  $u$ th MS, and  $\mathbf{n}_u \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$  is a Gaussian noise vector.

At the  $u$ th MS, the RF combiner  $\mathbf{w}_u$  is used to process the received signal  $\mathbf{r}_u$  to produce the scalar

$$y_u = \mathbf{w}_u^* \mathbf{H}_u \sum_{r=1}^U \mathbf{f}_r s_r + \mathbf{w}_u^* \mathbf{n}_u, \quad (2)$$

MmWave channels are expected to have limited scattering [2]. Therefore, and to simplify the analysis, we will assume a single-path geometric channel model [9, 13]. Under this model, the channel  $\mathbf{H}_u$  can be expressed as

$$\mathbf{H}_u = \sqrt{N_{\text{BS}} N_{\text{MS}} \alpha_u} \mathbf{a}_{\text{MS}}(\theta_u) \mathbf{a}_{\text{BS}}^*(\phi_u), \quad (3)$$

where  $\alpha_u$  is the complex path gain, including the path-loss, with  $\mathbb{E}[|\alpha_u|^2] = \bar{\alpha}$ . The variables  $\theta_u$ , and  $\phi_u \in [0, 2\pi]$  are the angles of arrival and departure (AoA/AoD) respectively. Finally,  $\mathbf{a}_{\text{BS}}(\phi_u)$  and  $\mathbf{a}_{\text{MS}}(\theta_u)$  are the antenna array response vectors of the BS and  $u$ th MS respectively. The BS

and each MS are assumed to know the geometry of their antenna arrays. While the results and insights developed in the paper can be generalized to arbitrary antenna arrays, we will assume uniform arrays in the simulations of Section 5.

### 3. PROPOSED DOWNLINK SYSTEM OPERATION

The proposed downlink operation for multi-user mmWave systems consists of two phases: (i) compressed sensing based downlink channel estimation and (ii) conjugate analog beamforming/combining. For the downlink channel training, random beamforming and projections are used to efficiently estimate the mmWave channel with relatively low training overhead thanks to the sparse nature of the channel. One main advantage of this technique is that all the MS's can simultaneously estimate their channels. Therefore, the training overhead does not scale with the number of users. This is contrary to the adaptive channel estimation and beamforming design techniques in [5, 6, 9], which are user-specific. The estimated channels are then used to build the analog beamformers and combiners. Extensions to hybrid analog/digital precoders are also possible [14], but our focus in this paper is on the evaluation of compressed sensing channel estimation.

#### 3.1. Compressed Sensing Based Channel Estimation

Given the geometric mmWave channel model in (3), estimating the channel is equivalent to estimating the different parameters of the channel path; namely its AoA, AoD, and the complex gain. In this section, we exploit this poor scattering nature of the mmWave channel, and formulate the channel estimation problem as a sparse problem. We then briefly show how compressed sensing can be used to estimate the channel.

**A sparse formulation:** Consider the system and mmWave channel models described in Section 2. If the BS uses a training beamforming vector  $\mathbf{p}_m$ , and the  $u$ th MS employs a training combining vector  $\mathbf{q}_n$  to combine the received signal, the resulting signal can be written as

$$y_{n,m} = \mathbf{q}_n^H \mathbf{H}_u \mathbf{p}_m s_m + \mathbf{q}_n^H \mathbf{n}_{n,m}, \quad (4)$$

where  $s_m$  is the training symbol on the beamforming vector  $\mathbf{p}_m$ , and we use  $s_m = \sqrt{P}$ , with  $P$  the average power used per transmission in the training phase. If the BS employs  $M_{\text{BS}}$  such beamforming vectors  $\mathbf{p}_m, m = 1, \dots, M_{\text{BS}}$ , at  $M_{\text{BS}}$  successive time slots, and the MS uses  $M_{\text{MS}}$  measurement vectors  $\mathbf{q}_n, n = 1, 2, \dots, M_{\text{MS}}$  at  $M_{\text{MS}}$  successive instants to detect the signal transmitted over *each* of the beamforming vectors, the resulting received matrix will be [5]

$$\mathbf{Y}_{\text{MS}} = \sqrt{P} \mathbf{Q}^H \mathbf{H}_u \mathbf{P} + \mathbf{N}, \quad (5)$$

where  $\mathbf{Q} = [\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_{M_{\text{MS}}}]$  is the  $N_{\text{MS}} \times M_{\text{MS}}$  measurement matrix,  $\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_{M_{\text{BS}}}]$  is the BS  $N_{\text{BS}} \times M_{\text{BS}}$  beamforming matrix, and  $\mathbf{N}$  is an  $M_{\text{MS}} \times M_{\text{BS}}$  noise matrix.

To exploit the sparse nature of the channel, we first vectorize the resultant matrix  $\mathbf{Y}_{\text{MS}}$  as in [5] to get

$$\mathbf{y}_{\text{MS}} = \sqrt{P} (\mathbf{P}^T \otimes \mathbf{Q}^H) (\mathbf{a}_{\text{BS}}^* (\phi_u) \otimes \mathbf{a}_{\text{MS}} (\theta_u)) \alpha_u + \mathbf{v}, \quad (6)$$

To complete the problem formulation, we assume that the AoDs, and AoAs are taken from a grid of  $G_{\text{BS}}$  and  $G_{\text{MS}}$  points, respectively. By neglecting the grid quantization error, we can approximate  $\mathbf{y}_{\text{MS}}$  in (6) as

$$\mathbf{y}_{\text{MS}} = \sqrt{P} (\mathbf{P}^T \otimes \mathbf{Q}^H) (\bar{\mathbf{A}}_{\text{BS}}^* \otimes \bar{\mathbf{A}}_{\text{MS}}) \mathbf{z}_u + \mathbf{v}, \quad (7)$$

where the  $N_{\text{BS}} \times G_{\text{BS}}$  matrix  $\bar{\mathbf{A}}_{\text{BS}}$  and  $N_{\text{MS}} \times G_{\text{MS}}$  matrix  $\bar{\mathbf{A}}_{\text{MS}}$  are the dictionary matrices that consist of the column vectors  $\mathbf{a}_{\text{BS}} (\bar{\phi}_k)$ , and  $\mathbf{a}_{\text{MS}} (\bar{\theta}_\ell)$ , respectively, with  $\bar{\phi}_k$ , and  $\bar{\theta}_\ell$  the  $k$ th, and  $\ell$ th points of the angle grids.  $\mathbf{z}_u$  is a  $G_{\text{BS}} G_{\text{MS}} \times 1$  vector which carries the path gains of the corresponding quantized directions. Hence,  $\mathbf{z}_u$  is a sparse vector with only 1 non-zero element. Note that detecting the column of  $\bar{\mathbf{A}}_{\text{BS}}$  and  $\bar{\mathbf{A}}_{\text{MS}}$  that corresponds to this non-zero element directly implies the detection of the AoA and AoD of the channel path.

**Compressed sensing measurements:** Thanks to the sparse formulation of the mmWave channel estimation problem in (7), compressed sensing tools can be leveraged to design efficient training beamforming/combining matrices [5, 15]. Considering the measurement matrix  $\Phi = \mathbf{P}^T \otimes \mathbf{Q}^H$ , and the dictionary  $\Psi = \bar{\mathbf{A}}_{\text{BS}}^* \otimes \bar{\mathbf{A}}_{\text{MS}}$ , one interesting research direction is to study the conditions on  $\Phi, \Psi$  under which the support of the sparse vector  $\mathbf{z}_u$  can be recovered with high probability and with low training overhead. Leaving this objective for future work, we will try in this paper to get some insights into a sufficient (not necessarily the minimum) number of measurements, and the relation between this number and the achievable rate of mmWave systems. For that, we consider the following measurement matrix.

The BS will design its training beamforming matrix  $\mathbf{P}$ , such that  $[\mathbf{P}]_{m,n} = e^{j\phi_{m,n}}$  where  $\phi_{m,n}$  is randomly and uniformly selected from the set of quantized angles  $\{0, \frac{2\pi}{N_{\text{BS}}}, \dots, \frac{(N_{\text{BS}}-1)2\pi}{N_{\text{BS}}}\}$ . Each MS similarly designs its training combining matrix  $\mathbf{Q}$ , with  $N_{\text{MS}}^{\text{MS}}$  angle quantization bits. Note that for this design, each entry of the measurement matrix  $\Phi$  will also be equal to  $e^{j\zeta}$ , with the angle  $\zeta$  selected randomly from a certain quantized angle set.

**AoA/AoD estimation:** To estimate its channel AoA/AoD, each MS  $u$  needs to recover the support of its sparse vector  $\mathbf{z}_u$ . As the measurement and dictionary matrices  $\Phi, \Psi$  are known for all MS's, each MS can use sparse recovery algorithms (such as LASSO [16], Orthogonal Matching Pursuit (OMP) [17], etc.) to estimate its channel AoA/AoD. In the simulations of Section 5, we adopt OMP for low-complexity. In this case, the support of  $\mathbf{z}_u$ ,  $\text{supp}(\mathbf{z}_u)$ , is determined by solving

$$\text{supp}(\mathbf{z}_u) = \arg \max \Psi^H \Phi^H \mathbf{y}_{\text{MS}}, \quad (8)$$

which directly determines the estimated AoA/AoD,  $\hat{\theta}_u, \hat{\phi}_u$ .

### 3.2. Conjugate Analog Beamforming

Given the estimated AoA, each MS  $u$  will design its analog combining vector such that  $\mathbf{w}_u = \mathbf{a}_{\text{MS}} (\hat{\theta}_u)$ . Each MS  $u$  will also feed the index of its estimated AoD back to the BS which needs  $\log_2 G_{\text{BS}}$  bits. Finally, the BS designs its analog beamforming matrix  $\mathbf{F}$  to match the effective channels (including the effect of the combining vectors), i.e., the BS sets  $\mathbf{F} = [\mathbf{a}_{\text{BS}} (\hat{\phi}_1), \mathbf{a}_{\text{BS}} (\hat{\phi}_2), \dots, \mathbf{a}_{\text{BS}} (\hat{\phi}_U)]$ .

## 4. ACHIEVABLE AND EFFECTIVE RATES

In this section, we evaluate the achievable rate of the proposed downlink mmWave system operation in a special case and try to get some insights into the relation between the performance of mmWave systems and compressed sensing measurements in more general cases.

Consider the system model in Section 2 and the proposed conjugate analog beamforming/combining in Section 3.2. For tractability, we make the following assumptions

**Assumption 1** All path gains are constants. This is relevant to mmWave LOS paths, which are dominant in dense mmWave networks [18].

**Assumption 2** The BS and MS's employ uniform arrays.

**Assumption 3** The sizes of the grids in (7) are  $G_{\text{BS}} = N_{\text{BS}}$ ,  $G_{\text{MS}} = N_{\text{MS}}$ , and the angle grid points  $\phi_k, \theta_\ell$ , of the dictionary  $\Psi$  are the virtual directions satisfying  $\frac{2\pi d}{\lambda} \sin(\phi_k) = \frac{2\pi k}{N_{\text{BS}}}$  and  $\frac{2\pi d}{\lambda} \sin(\theta_\ell) = \frac{2\pi \ell}{N_{\text{MS}}}$  [19].

Under the assumptions 2-3, the matrices  $\bar{\mathbf{A}}_{\text{BS}}$  and  $\bar{\mathbf{A}}_{\text{MS}}$  become DFT matrices [19]. Further, using the virtual channel model transformation in [19], we note that the steering vectors of both the actual and estimated AoAs/AoDs (of the dictionary  $\Psi$ ) are columns of these matrices. Therefore, the rate of user  $u$ ,  $R_u$ , can be written as

$$\log_2 \left( 1 + \frac{\mathbb{1}_{\{\hat{\theta}_u = \theta_u, \hat{\phi}_u = \phi_u\}}}{\mathbb{1}_{\{\hat{\theta}_u = \theta_u\}} \sum_{r=1}^U \mathbb{1}_{\{\hat{\phi}_u = \phi_r\}} + \frac{1}{\frac{\text{SNR}}{U} N_{\text{BS}} N_{\text{MS}} |\alpha_u|^2}} \right). \quad (9)$$

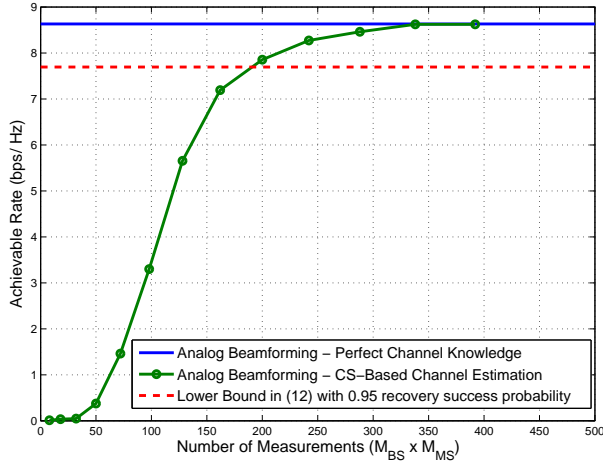
Denoting  $\hat{R}_u = \log_2 \left( 1 + \frac{\text{SNR}}{U} N_{\text{BS}} N_{\text{MS}} |\alpha_u|^2 \right)$  as the single-user rate (without interference), the average achievable rate of user  $u$ ,  $\bar{R}_u$ , can be then lower bounded as

$$\bar{R}_u \geq \mathbb{E} \left[ \hat{R}_u \mathbb{1}_{\{\cap_{r \neq u} (\phi_u \neq \hat{\phi}_r)\}} \mathbb{1}_{\{\hat{\theta}_u = \theta_u, \hat{\phi}_u = \phi_u\}} \right], \quad (10)$$

$$\stackrel{(a)}{=} \hat{R}_u \mathbb{P}_{\{\cap_{r \neq u} (\phi_u \neq \hat{\phi}_r)\}} \mathbb{P}_{\{\hat{\theta}_u = \theta_u, \hat{\phi}_u = \phi_u\}}, \quad (11)$$

$$\geq \hat{R}_u \left( 1 - \frac{U}{N_{\text{BS}}} \right) \mathbb{P}_{\{\hat{\theta}_u = \theta_u, \hat{\phi}_u = \phi_u\}}, \quad (12)$$

where (a) follows from the independence between the estimation success event  $\{\hat{\theta}_u = \theta_u, \hat{\phi}_u = \phi_u\}$  and the single-user



**Fig. 2.** Achievable rates using the proposed system operation for different numbers of compressed sensing measurements

rate given assumption 1. Now, we note that  $P_{\{\hat{\theta}_u=\theta_u, \hat{\phi}_u=\phi_u\}}$  is the probability of correct support recovery of the sparse vector  $\mathbf{z}_u$ . This directly relates the achievable rate to the compressed sensing literature. For a general relation, assume the channel fading coherence equals  $L_C$  symbols, the effective achievable rate of user  $u$  (considering the training overhead) can be written as

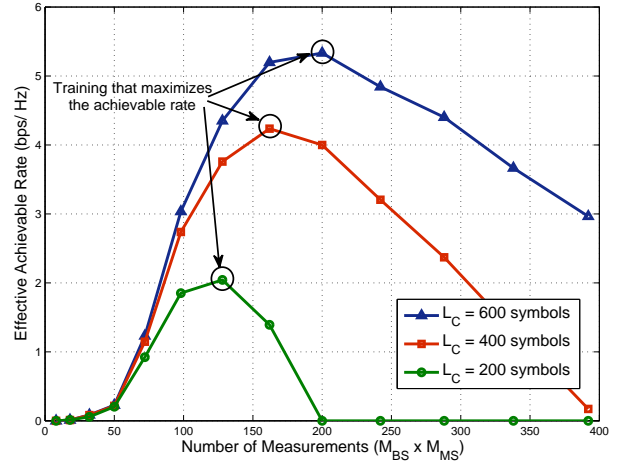
$$\bar{R}_{u,\text{eff}} \geq \hat{R}_u \left(1 - \frac{U}{N_{\text{BS}}}\right) \left(1 - \frac{M_\epsilon}{L_C}\right) (1 - \epsilon), \quad (13)$$

where  $M_\epsilon$  equals the number of measurements needed to guarantee the support recovery of the sparse vector  $\mathbf{z}_u$  with probability at least  $1 - \epsilon$ . Given this relation in (13), it is interesting to design the measurement matrix  $\Phi$  to maximize this lower bound on the achievable rate. It is also of interest to optimize the number of measurements to handle the tradeoff shown between the accurate channel estimate (small  $\epsilon$ ) and large training overhead.

## 5. SIMULATION RESULTS

In this section, we evaluate the performance of the proposed downlink mmWave system operation. We consider the system model in Section 2 with the BS employing a ULA of 64 antennas and 4 MS's, each with 32 antennas. The system operates at 28 GHz with a bandwidth of 50 MHz. The BS-MS distance is 500 m, and all channels are LOS and single-path. In the channel estimation phase, the BS and MS's apply the random beamforming and measurement procedure described in Section 3.1 with  $N_Q^{\text{BS}} = N_Q^{\text{MS}} = 4$ , and with average transmit power equal to 37 dBm. The beamformers/combiners are then built as described in Section 3.2.

In Fig. 2, the per-user achievable rate of the proposed system operation is shown versus the number of compressed



**Fig. 3.** Achievable rates using the proposed system operation for different values of channel fading coherence.

sensing measurements. This figure illustrates that  $\sim 280 - 330$  measurements are needed to approach the achievable rate with perfect channel knowledge. While this number may look large, it is actually an order of magnitude less than what is required by traditional lower frequency solutions, which is  $\sim N_{\text{BS}}N_{\text{MS}} = 2048$  symbols. Note also that this training overhead does not depend on the number of users, which makes compressed sensing of special interest to multi-user mmWave systems. Compared with adaptive compressed sensing techniques [5], they may require less training to estimate each user channel, e.g.,  $\sim 50 - 100$  in [5]. This overhead, however, scales with the number of users, which means  $200 - 400$  for 4 users, and larger if more users are served. Finally, note that these results just illustrate a sufficient number of measurements given the proposed design of the measurement matrix  $\Phi$  in Section 3.1. Therefore, optimizing this design may lead to even smaller training overhead.

In Fig. 3, the effective achievable rate is shown for different numbers of fading coherence values,  $L_C$ . This figure indicates that it is important to compromise between the accuracy of the channel estimate and the training overhead to maximize the achievable rate, especially with fast channels.

## 6. CONCLUSION

In this paper, we proposed and evaluated a low-complexity downlink system operation for multi-user mmWave systems based on compressed sensing channel estimation. Simulation results showed that the proposed system operation requires a relatively small training overhead (w.r.t. the channel matrix dimensions) to achieve a very good performance. Results also indicated that the number of measurements need to be wisely selected to maximize the effective system sum-rate.

## 7. REFERENCES

- [1] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 101–107, 2011.
- [2] T.S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G.N. Wong, J.K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!," *IEEE Access*, vol. 1, pp. 335–349, 2013.
- [3] T. Bai, A. Alkhateeb, and R.W. Heath, "Coverage and capacity of millimeter-wave cellular networks," *IEEE Communications Magazine*, vol. 52, no. 9, pp. 70–77, September 2014.
- [4] F. Boccardi, R.W. Heath, A. Lozano, T.L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, February 2014.
- [5] A. Alkhateeb, O. El Ayach, G. Leus, and R.W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 831–846, Oct 2014.
- [6] J. Wang, Z. Lan, C.W. Pyo, T. Baykas, C.S. Sum, M.A. Rahman, J. Gao, R. Funada, F. Kojima, H. Harada, et al., "Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 8, pp. 1390–1399, 2009.
- [7] IEEE Standard 802.15.3c, "Wireless medium access control (MAC) and physical layer (PHY) specifications for high rate wireless personal area networks (WPANs), amendment 2: Millimeter-wave-based alternative physical layer extension," October 2009.
- [8] L. Chen, Y. Yang, X. Chen, and W. Wang, "Multi-stage beamforming codebook for 60GHz WPAN," in *Proc. of 6th International ICST Conference on Communications and Networking in China*, China, 2011, pp. 361–365.
- [9] 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 Transactions on Communications*, vol. 61, no. 10, pp. 4391–4403, 2013.
- [10] Y.M. Tsang, A.S.Y. Poon, and S. Addepalli, "Coding the beams: Improving beamforming training in mmwave communication system," in *Proc. of 2011 IEEE Global Telecommunications Conference (GLOBECOM)*, Houston, TX, 2011, pp. 1–6.
- [11] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R.W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1499–1513, March 2014.
- [12] J. Brady, N. Behdad, and A.M. Sayeed, "Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis, and measurements," *IEEE Trans. on Ant. and Propag.*, vol. 61, no. 7, pp. 3814–3827, July 2013.
- [13] C. Nicolas Barati, S. Amir Hosseini, S. Rangan, P. Liu, T. Korakis, and S. S Panwar, "Directional cell search for millimeter wave cellular systems," *arXiv preprint arXiv:1404.5068*, 2014.
- [14] A. Alkhateeb, G. Leus, and R. W Heath Jr, "Limited feedback hybrid precoding for multi-user millimeter wave systems," *submitted to IEEE Transactions on Wireless Communications*, *arXiv preprint arXiv:1409.5162*, 2014.
- [15] D. Ramasamy, S. Venkateswaran, and U. Madhow, "Compressive adaptation of large steerable arrays," in *Proc. of 2012 Information Theory and Applications Workshop (ITA)*, CA, 2012, pp. 234–239.
- [16] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.
- [17] T. Tony Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery with noise," *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4680–4688, 2011.
- [18] T. Bai and R. W. Heath Jr., "Coverage analysis in dense millimeter wave cellular networks," in *Proc. of 47th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, November 2013.
- [19] A.M. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2563–2579, Oct 2002.