EFFICIENT CHANNEL STATISTICS ESTIMATION FOR MILLIMETER-WAVE MIMO SYSTEMS

Yue Wang[*] *Zhi Tian*[†] *Shulan Feng*[*] *Philipp Zhang*[*]

[*]Hisilicon Research Department, Huawei Technologies Co. Ltd., Beijing, China
[†]Electrical and Computer Engineering Department, George Mason University, Fairfax, VA, USA

ABSTRACT

In millimeter-wave (mmWave) multiple-input multiple-output (MIMO) systems, channel estimation is a challenging task in terms of acquiring the instantaneous channel state information (CSI), because both the estimation complexity and the overhead required for pilot symbols and feedback grow drastically as the number of antennas increases. Alternatively, some channel statistics in the form of partial CSI are sufficient for transceiver optimization and noncoherent detection in time-varying wireless environments. To obtain such useful statistical information accurately and efficiently, this paper proposes a new channel statistics estimation method using the compressive covariance sensing technique, which directly estimates the desired second-order statistics of the channel while bypassing the intermediate recovery of the instantaneous channel matrix itself. A diagonal-search orthogonal matching pursuit (DS-OMP) algorithm is developed for fast channel estimation. The proposed algorithm has low computational complexity and reduced overhead in training and feedback, owing to its proper utilization of the joint sparsity structure of the channel covariance matrix.

Index Terms— channel statistics estimation, compressive covariance sensing, DS-OMP, millimeter-wave, MIMO.

1. INTRODUCTION

Millimeter-wave (mmWave) communications are widely considered as a viable technological option for future Gbps-datarate wireless services [1]. The deployment of mmWave systems is often coupled with massive MIMO technology, which utilizes a very large number of antenna elements to offer unprecedented spatial diversity and array gain to combat the severe channel path loss experienced at the over-28GHz mmWave bands [2]. To harness the benefits of MIMO mmWave transmissions, several prominent technical challenges need to be overcome, including the difficulty in efficient channel estimation for large-size MIMO.

For conventional MIMO, a number of successful channel estimation techniques have been developed to acquire the instantaneous CSI expressed by the channel matrix during the training period on a frame-by-frame basis, for the purpose of transmitter design and receiver demodulation. However, the very large antenna number in mmWave MIMO systems results in a much enlarged channel matrix, which causes heavy transmission needs for both pilot symbols and CSI feedback. Such heavy overheads within a given frame duration may lead to severe shortage of data transmission resources, rendering traditional channel estimation methods ineffective. To reduce the training overhead, compressive sensing (CS) has been advocated for CSI estimation in mmWave MIMO [3, 6–8]. By exploiting the sparse multipath structure of an mmWave MIMO channel [3–5], the CS approach [9] can efficiently estimate the channel matrix from a relatively small set of compressively collected pilot symbols.

While the CSI is useful and preferred for slowly-varying wireless channels, it may become costly and not quite meaningful to acquire it and feed it back to the transmitter, especially in rapidly varying wireless environments where previously acquired CSI becomes outdated quickly. Indeed, there are several situations where only some channel statistical information needs to be acquired in lieu of the CSI. First, in fast-varying MIMO channels, transmitter precoding designs and optimization have been developed based on second-order channel statistics also known as a form of partial CSI, which circumvents frequent channel updates [10, 11]. Second, channel statistics can be adequate for noncoherent detection at the receiver. For example, some channel statistic acts as a diagonal weighting matrix in noncoherent multiple-symbol differential detection [12]. All these works assume that the channel statistics are readily available either as prior knowledge or via straightforward finite sample averaging [11]. Unfortunately for mmWave MIMO, the task of channel statistics estimation can be quite computationally expensive and time consuming, as the number of antennas grows exceedingly large.

This paper aims to develop an efficient and accurate technique for channel statistics estimation (CSE), by exploiting the special sparsity structure of mmWave MIMO propagation channels. While the CS approach is directly applicable for CSI estimation, its application to the CSE task is not straightforward, because the sparsity structure available in a CSE problem lies in channel correlations, not in the channel itself. Unfortunately, the sparse channel correlations do not have a direct linear relationship with the received (linearly compressed) samples, whereas well-known sparse signal recovery techniques hinge on linear measurement models [9]. To circumvent this obstacle, we propose a task-cognizant method that directly acquires the channel statistical information (i.e., channel covariance matrix) from compressive measurements, bypassing the intermediate step of recovering the channel matrix itself. This work is developed based on the compressive covariance sensing (CCS) framework that was initially introduced for power spectrum estimation and cyclic feature detection problems [13, 14]. Furthermore, we utilize the unique sparsity structure of mmWave MIMO channels to improve both the estimation accuracy and computational efficiency. Specifically, a diagonal-search orthogonal matching pursuit (DS-OMP) algorithm is designed that capitalizes on both the joint sparsity and the Hermitian structure of the channel covariance matrix to considerably reduce the algorithmic complexity and transmission overhead. As such, the proposed CCS-based CSE (CCS-CSE) technique is expected to offer major benefits to several useful tasks in mmWave wireless systems with large-size MIMO, including transmitter optimization with partial feedback and noncoherent detection.

Notation: a is a scalar, $\mathbf{a}$ is a vector, and $\mathbf{A}$ is a matrix. $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$ are the transpose, conjugate, and conjugate transpose of a matrix or vector, respectively. $\mathbf{A}\otimes\mathbf{B}$ is the Kronecker product of $\mathbf{A}$ and $\mathbf{B}$. $\|\mathbf{A}\|_F$ and $\mathbf{A}^\dagger$ are the Frobenius norm and the pseudoinverse of $\mathbf{A}$. The operation vec$(\cdot)$ stacks all the columns into a vector. $\mathrm{E}\{\cdot\}$ denotes expectation.

2. CHANNEL AND SIGNAL MODELS

Consider an mmWave MIMO system with N_t transmit and N_r receive antennas. We adopt a block-fading channel model in which a channel estimator observes the received training signal during a discrete time window of length T within each frame. The received signal $\mathbf{Y}\in\mathbb{C}^{N_r\times T}$ is given by

$$\mathbf{Y}=\mathbf{H}\mathbf{X}+\mathbf{W}, \tag{1}$$

where $\mathbf{H}\in\mathbb{C}^{N_r\times N_t}$ represents the mmWave MIMO channel, $\mathbf{X}\in\mathbb{C}^{N_t\times T}$ is the transmitted pilot signal, and $\mathbf{W}\in\mathbb{C}^{N_r\times T}$ is the additive Gaussian noise independent of $\mathbf{H}$ and $\mathbf{X}$.

At the mmWave frequency, a MIMO channel experiences limited scattering, resulting in a sparse multipath structure. A geometric channel model with L scatterers can be formed using ray tracing [15], where each scatterer contributes to one propagation path. Accordingly, the channel matrix $\mathbf{H}$ can be expressed as the sum of L paths in the form of [15]

$$\mathbf{H}=\sqrt{\tfrac{N_tN_r}{\rho}}\sum_{l=1}^{L}\alpha_l\mathbf{a}_r(\theta_{r,l})\mathbf{a}_t^H(\theta_{t,l}), \tag{2}$$

where ρ is the average path loss, α_l represents the gain of the l-th propagation path, $\theta_{t,l}\in[0,2\pi]$ and $\theta_{r,l}\in[0,2\pi]$ denotes the angles of departure and arrival of the l-th path at transmitter and receiver, respectively. The vectors $\mathbf{a}_t(\theta_{t,l})$ and $\mathbf{a}_r(\theta_{r,l})$ denote the array response vectors for transmitting and receiving antenna arrays, respectively.

The channel $\mathbf{H}$ hinges on unknown angular directions $\{\theta_{t,l},\theta_{r,l}\}_l$ that are difficult to acquire. To solve this problem, a virtual channel representation is applied to characterize the MIMO channel by fixed virtual receive and transmit directions [16], that is,

$$\mathbf{H}=\mathbf{U}_r\mathbf{H}_v\mathbf{U}_t^H, \tag{3}$$

where $\mathbf{U}_r\in\mathbb{C}^{N_r\times N_r}$ and $\mathbf{U}_t\in\mathbb{C}^{N_t\times N_t}$ are unitary discrete Fourier transform (DFT) matrices reflecting the fixed virtual receive and transmit angles that uniformly sample the unit angle space, and $\mathbf{H}_v\in\mathbb{C}^{N_r\times N_t}$ is the virtual channel matrix whose entries capture the gains of the corresponding paths. In this sense, $\mathbf{H}$ and $\mathbf{H}_v$ are unitarily equivalent, and hence $\mathbf{H}$ is linearly characterized by its virtual representation $\mathbf{H}_v$. When N_r and N_t are large, the angular resolution of the virtual receive and transmit directions can be fine enough, such that it is reasonable to assume that the angles in (2) fall along L of the N_t or N_r angular directions defined by the DFT matrices in (3). In that case, $\mathbf{H}_v$ contains only L nonzero entries corresponding to the L scattering paths. This salient property of $\mathbf{H}_v$ owes to the large antenna size in mmWave MIMO.

Given $\mathbf{Y}$, the goal of channel statistics estimation is to obtain the second-order statistics of the mmWave MIMO channel, which is captured by the channel covariance matrix

$$\mathbf{R}_h=\mathrm{E}\{\mathbf{h}\mathbf{h}^H\}, \tag{4}$$

where $\mathbf{h}$=vec$(\mathbf{H})$. From (3) and the property of Kronecker product (vec$(\mathbf{ABC})$=$(\mathbf{C}^T\otimes\mathbf{A})vec(\mathbf{B})$), $\mathbf{R}_h$ is rewritten as

$$\mathbf{R}_h=(\mathbf{U}_t^*\otimes\mathbf{U}_r)\mathbf{R}_{h_v}(\mathbf{U}_t^*\otimes\mathbf{U}_r)^H, \tag{5}$$

where $\mathbf{R}_{h_v}$=$\mathrm{E}\{\mathbf{h}_v\mathbf{h}_v^H\}$ and $\mathbf{h}_v$=vec$(\mathbf{H}_v)$. Thus, the estimation of $\mathbf{R}_h$ is equivalent to estimating $\mathbf{R}_{h_v}$.

The channel statistics matrix $\mathbf{R}_{h_v}$ possesses a special sparsity structure that is not even present in $\mathbf{H}_v$ itself. For illustration, we simulated two arbitrary realizations of $\mathbf{H}$ in (2) and its sparse representations $\mathbf{H}_v$ in (3) for N_t=10, N_r=10 and L=3, as depicted in Fig. 1. The corresponding vectorized $\mathbf{h}_v$ are shown in Fig. 2 (a-b), which result in a sparse matrix $\mathbf{R}_{h_v}$ as in Fig. 2 (c). All channel realizations of $\mathbf{h}_v$, regardless of how many, are jointly sparse with common nonzero support, which are reflected in the structure of $\mathbf{R}_{h_v}$. There are only L^2 nonzero entries in $\mathbf{R}_{h_v}$, co-located with a special geometric relationship. The L diagonal nonzero entries reveal the auto-correlations of the sparse channel and the associated off-diagonal nonzero entries indicate the cross-correlations.

3. SPARSE CHANNEL STATISTICS ESTIMATION

Sparsity is a useful feature for reducing the sampling costs and transmission overhead in data-aided channel estimation. Here, sparsity exhibits in both the channel matrix $\mathbf{H}_v$ and the channel statistics matrix $\mathbf{R}_{h_v}$, but the quantity of interest $\mathbf{R}_{h_v}$ does not have a direct linear relationship with the received samples $\mathbf{Y}$. A naive method for sparse recovery of $\mathbf{R}_{h_v}$ works in a two-step manner: first recovering $\mathbf{H}_v$ from the data collected during each frame, and then computing $\mathbf{R}_{h_v}$ via its finite sample average over F frames. This procedure follows

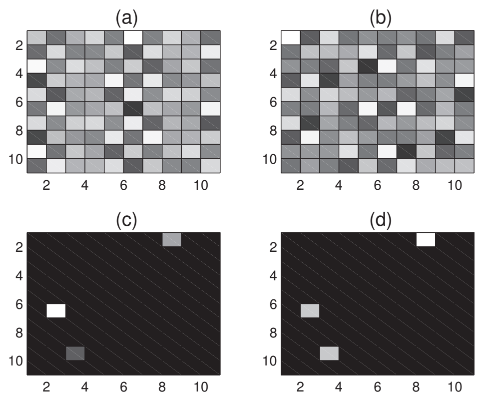

Fig. 1. Illustration of $\mathbf{H}$ and $\mathbf{H}_v$ for two frames: (a-b) display $\mathbf{H}$, and (c-d) show the resulting $\mathbf{H}_v$, for $L = 3$ channel paths.

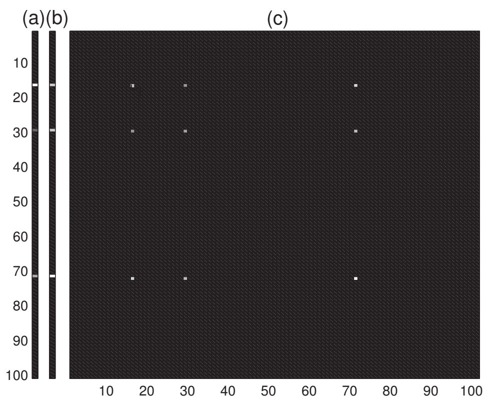

Fig. 2. Illustration of the structure of $\mathbf{R}_{h_v}$: (a-b) depict $\mathbf{h}_v$ of the two frames; (c) is the square $\mathbf{R}_{h_v}$ that implies the joint sparsity among $\mathbf{h}_v$'s and the Hermitian property of $\mathbf{R}_{h_v}$.

the basic CS approach for CSE (CS-CSE). Herein, expensive sparsity-enforcing inverse problems need to be performed F times for a large number of F frames in order to acquire reliable statistics, which is highly wasteful of the training and computing resources. Further, the estimation can have inferior quality, because it only utilizes the sparsity of $\mathbf{H}_v$, but ignores some unique properties of the statistics $\mathbf{R}_{h_v}$. To overcome these drawbacks and estimate $\mathbf{H}_v$ more accurately and efficiently, next we develop a new channel statistics estimation approach using the CCS technique [13, 14]. This new technique, termed CCS-CSE, directly estimates $\mathbf{R}_{h_v}$ while bypassing the recovery of $\mathbf{H}_v$ itself.

3.1. Linear Relationships

According to the CCS framework [13, 14], a key step is to identify a proper linear relationship between the observations and the desired channel statistics. To this end, we first vectorize $\mathbf{Y}$ in (1) as

$$\mathbf{y} = \text{vec}(\mathbf{Y}) = \text{vec}(\mathbf{HX} + \mathbf{W}) = \text{vec}(\mathbf{HX}) + \mathbf{w}, \quad (6)$$

where $\mathbf{w}$=vec($\mathbf{W}$). From (3), $\mathbf{y}$ can be expressed as

$$\mathbf{y} = \text{vec}(\mathbf{U}_r\mathbf{H}_v\mathbf{U}_t^H\mathbf{X}) + \mathbf{w}. \quad (7)$$

Using the property of Kronecker product, we have

$$\mathbf{y}=((\mathbf{U}_t^H\mathbf{X})^T\otimes\mathbf{U}_r)\,\text{vec}(\mathbf{H}_v)+\mathbf{w}=\mathbf{\Psi}\mathbf{h}_v+\mathbf{w}, \quad (8)$$

where $\mathbf{\Psi} = (\mathbf{U}_t^H\mathbf{X})^T \otimes \mathbf{U}_r$ and $\mathbf{h}_v = \text{vec}(\mathbf{H}_v)$.

In the absence of a direct linear relationship between $\mathbf{y}$ and the unknown channel statistics $\mathbf{R}_{h_v}$, we turn to the observation statistics $\mathbf{R}_y$=E$\{\mathbf{y}\mathbf{y}^H\}\in\mathbb{C}^{N_rT\times N_rT}$, which can be approximated by its finite-sample average across K frames as $\hat{\mathbf{r}}_y = \text{vec}(\frac{1}{F}\sum_{f=1}^{F}\mathbf{y}(k)\mathbf{y}^H(k))$. Apart from the noise term $\mathbf{w}$ in (8), the observation covariance matrix $\mathbf{R}_y$ holds

$$\mathbf{R}_y = \mathbf{\Psi}\text{E}\{\mathbf{h}_v\mathbf{h}_v^H\}\mathbf{\Psi}^H = \mathbf{\Psi}\mathbf{R}_{h_v}\mathbf{\Psi}^H. \quad (9)$$

After vectorizing $\mathbf{R}_y$ and $\mathbf{R}_{h_v}$ in (9) and using the property of Kronecker product, we reach a linear representation for the relevant second-order statistics:

$$\mathbf{r}_y=\text{vec}(\mathbf{R}_y)=(\mathbf{\Psi}^*\otimes\mathbf{\Psi})\,\text{vec}(\mathbf{R}_{h_v})=\mathbf{\Phi}\mathbf{r}_{h_v}, \quad (10)$$

where $\mathbf{\Phi} = \mathbf{\Psi}^* \otimes \mathbf{\Psi}$ is of size $(N_rT)^2 \times (N_rN_t)^2$.

In the CCS-based CSE formulation (10), the observations $\mathbf{Y}$ are compressive with respect to the CSE task when setting $T < N_t$, which renders $\mathbf{\Phi}$ rank-deficient. Choosing a small value of T can effectively reduce the training overhead, but calls for sparsity-cognizant channel recovery. In that case, random data $\mathbf{X}$ can be used to induce a well-behaved sampling matrix $\mathbf{\Phi}$ for sparse signal recovery. When compressive measurements are involved, the computational complexity of the CCS-CSE problem in (10) has been drastically reduced from that of the CS-CSE, because only one inverse problem needs to be solved to reconstruct $\mathbf{r}_{h_v}$ from $\mathbf{r}_y$, as opposed to solving F inverse problems in CS-CSE.

3.2. DS-OMP Algorithm

The problem of CSE boils down to solving for $\mathbf{r}_{h_v}$ from $\mathbf{r}_y$ in (10). Since $\mathbf{R}_{h_v}$ is highly sparse as illustrated in Fig. 2, its vectorized form $\mathbf{r}_{h_v}$ is sparse as well, which permits the application of existing sparse signal recovery algorithms. Here we develop a new diagonal-search orthogonal matching pursuit (DS-OMP) algorithm to offer fast reconstruction of the channel statistics at low complexity. OMP is a greedy search algorithm for linear regression, which iteratively searches for a most significant regressor and then updates the residue signal through orthogonal projection before the next search [17, 18]. The key idea of DS-OMP is to effectively reduce the search space of OMP by making judicious use of the highly structured sparsity pattern of $\mathbf{R}_{h_v}$, which is unique to the channel statistics estimation problem for large-size MIMO as encountered in mmWave communications.

As illustrated in Fig. 2 (c), $\mathbf{R}_{h_v}$ has only a small number of L nonzero diagonal elements corresponding to the autocorrelation of channel paths, and an off-diagonal cross-correlation term $[\mathbf{R}_{h_v}]_{i,j}$ is nonzero only when the i-th and j-th diagonals are nonzero. This sparsity pattern is induced by the joint sparsity structure of $\mathbf{H}_v$ across all frames and the Hermitian property of $\mathbf{R}_{h_v}$. DS-OMP judiciously utilizes this special sparsity pattern to reconstruct $\mathbf{R}_{h_v}$, without invoking the solutions of multiple-measurement vectors for the joint sparsity issue [19]. Specifically, DS-OMP searches for the strongest L components along the set of N_rN_t diagonal entries Λ of $\mathbf{R}_{h_v}$ only, rather than searching the full dimension $N_rN_t \times N_rN_t$. At the i-th iteration, a new nonzero diagonal is identified via correlation matching with the residue signal $\mathbf{u}$, and its index $t^{(i)}$, as appeared in the vectorized $\mathbf{r}_{h_v}$, is added to the diagonal index set Γ. Meanwhile, DS-OMP updates the off-diagonal set Ω by inferring from current and previously identified nonzero diagonals. The computational complexity of DS-OMP is reduced from that of standard OMP by a factor of N_tN_rL, because the search space is reduced from the full dimension $N_rN_t \times N_rN_t$ to only N_rN_t diagonals of $\mathbf{R}_{h_v}$ at each iteration and the number of iterations is reduced from L^2 to L. The iteration steps are explained in Algorithm 1.

Algorithm 1 Diagonal Search Orthogonal Matching Pursuit

Input: $\mathbf{r}_y$, $\mathbf{\Phi}$, L
Initialize: i=0, $\mathbf{r}^{(0)}$=$\mathbf{0}$, $\mathbf{u}^{(0)}$=$\mathbf{r}_y$, $\Gamma^{(0)}$=$\varnothing$, Λ=$\varnothing$, Ω=$\varnothing$
Initialize the full diagonal search space:
for $k = 1, \ldots, L$ **do**
  $\Lambda = \Lambda \cup \{(k-1)N_rN_t + k\}$
end for
Iterate:
while $i < L$ **do**
  $i = i+1;\ \ t^{(i)} = \arg\max_{j\in\Lambda} |<\mathbf{u}^{(i-1)}, \mathbf{\Phi}_{\cdot,j}>|$
  if $i = 1$ **then**
    $\Gamma^{(i)} = \Gamma^{(i-1)} \cup t^{(i)}$
  else
    for $m = 1, \ldots, i-1$ **do**
      $\tau = \mod(|t^{(i)} - t^{(m)}|, N_rN_t)$
      if $t^{(i)} > t^{(m)}$ **then**
        $\Omega = \Omega \cup \{t^{(m)} + \tau, t^{(i)} - \tau\}$
      else
        $\Omega = \Omega \cup \{t^{(i)} + \tau, t^{(m)} - \tau\}$
      end if
    end for
    $\Gamma^{(i)} = \Gamma^{(i-1)} \cup \{t^{(i)}, \Omega\}$
  end if
  $\mathbf{r}^{(i)} = \mathbf{\Phi}^{\dagger}_{\cdot,\Gamma^{(i)}}\mathbf{r}_y;\ \ \mathbf{u}^{(i)} = \mathbf{r}_y - \mathbf{\Phi}_{\cdot,\Gamma^{(i)}}\mathbf{r}^{(i)}$
end while
Output: $\hat{\mathbf{r}}_{h_v} = \mathbf{r}^{(L)}$

4. SIMULATION RESULTS

An mmWave MIMO system operating at 28GHz is simulated for N_t=10 and N_r=10. The performance of channel statistics estimation is evaluated by the probability of correct estimation expressed by $\Pr(\|\mathbf{R}_h - \hat{\mathbf{R}}_h\|_F^2/(N_rN_t)^2 \le \eta)$, where η is a predefined threshold (e.g., η=0.001). The proposed CCS-CSE algorithm using DS-OMP is compared with CS-CSE.

We first test the algorithms in various wireless environments. Fig. 3(a) shows that the CCS-CSE achieves better accuracy than CS-CSE in noisy channels, especially at the low signal-to-noise ratio (SNR) regime. Fig. 3(b) indicates that CCS-CSE outperforms CS-CSE in different scattering environments indicated by L, and the performance gap enlarges as L increases. This is because CS-CSE can only make use of the channel sparsity that unfortunately fades away as L increases, whereas CCS-CSE utilizes not only the joint sparsity but also the Hermitian property of the channel statistics.

Fig. 4 (a) and (b) test the two methods at various observation costs in T (the length of the training sequence within each frame) and F (the total number of involved frames), respectively. Evidently, the CCS-CSE consumes less observation resources than the CS-CSE to achieve the same accuracy. Such benefits in reduced transmission overhead come at a much lower computational cost as well.

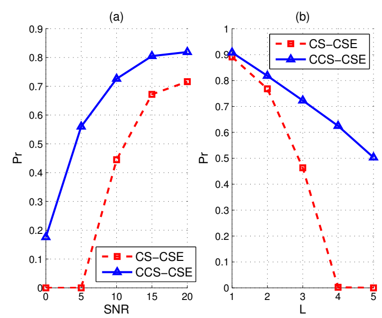

Fig. 3. Probability of correct estimation vs. (a) SNR (L=3, T=6, F=10), and (b) L (T=6, F=10, SNR=10).

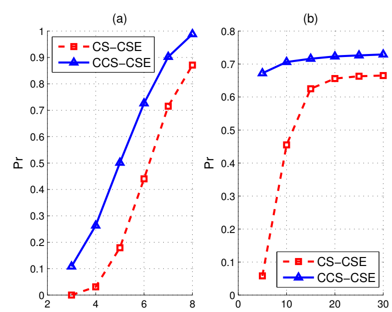

Fig. 4. Probability of correct estimation vs. (a) T (L=3, F=10, SNR=10), and (b) F (L=3, T=6, SNR=10).

5. CONCLUSION

An efficient channel statistics estimation method is developed for mmWave MIMO communications. Large-size mmWave MIMO gives rise to a highly structured sparsity pattern in the channel covariance matrix, which is utilized by the proposed DS-OMP algorithm for fast and accurate channel statistics estimation. As a result, both the computational cost and transmission overhead are considerably reduced.

6. REFERENCES

[1] T. S. Rappaport *et al.*, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, no. 1, pp. 335–349, Aug. 2013.

[2] 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, no. 2, pp. 106–113, Feb. 2014.

[3] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Procss.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

[4] H. Zhang, S. Venkateswaran, and U. Madhow, "Channel modeling and MIMO capacity for outdoor millimeter wave links," in *Proc. Wireless Commun. Netw. Conf.*, Apr. 2010, pp. 1–6.

[5] T. S. Rappaport *et al.*, "Broadband millimeter-wave propagation measurements and models using adaptive-beam antennas for outdoor urban cellular communications," *IEEE Trans. Antennas Propag.*, vol. 61, no. 4, pp. 1850–1859, Apr. 2013.

[6] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proc. IEEE*, vol. 98, no. 6, pp. 1058–1076, June 2010.

[7] S.L.H. Nguyen, and A. Ghrayeb, "Compressive sensing-based channel estimation for massive multiuser MIMO systems," in *Proc. Wireless Commun. Netw. Conf.*, Apr. 2013, pp. 2890–2895.

[8] P. Schniter and A. M. 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.

[9] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[10] S. A. Jafar, S. Vishwanath, and A. Goldsmith, "Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback," in *Proc. Int. Conf. Commun.*, June 2001, pp. 2266–2270.

[11] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel correlations," *IEEE Trans. Inform. Theory.*, vol. 49, no. 7, pp. 1673–1690, July 2003.

[12] A. Schenk and R. F. H. Fischer, "Noncoherent detection in massive MIMO systems," in *Proc. Int. ITG Workshop Smart Antennas*, Mar. 2013, pp. 1–8.

[13] Z. Tian, "Cyclic feature based wideband spectrum sensing using compressive sampling," in *Proc. Int. Conf. Commun.*, June 2011, pp. 1–5.

[14] D. D. Ariananda and G. Leus, "Compressive wideband power spectrum estimation," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4775–4789, Sep. 2012.

[15] A. M. Sayeed and V. Raghavan, "Maximizing MIMO capacity in sparse multipath with reconfigurable antenna arrays," *IEEE J. Sel. Topics Signal Procss.*, vol. 1, no. 1, pp. 156–166, June 2007.

[16] A. M. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2563–2579, Oct. 2002.

[17] Y. Pati, R. Rezaiifar, P. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with application to wavelet decomposition," in *Proc. Asilomar Conf. Signals Syst. Comput.*, Nov. 1993, pp. 40–44.

[18] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[19] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2477–2488, July 2005.