Channel Estimation in Millimeter Wave MIMO Systems with One-Bit Quantization

Jianhua Mo*, Philip Schniter†, Nuria González Prelcic‡ and Robert W. Heath, Jr.*

*Department of ECE, The University of Texas at Austin, Austin, TX 78712

†Department of ECE, The Ohio State University, Columbus, OH 43210

‡Department of Signal Theory and Communications, Universidade de Vigo

{jhmo, rheath}@utexas.edu, schniter@ece.osu.edu, nuria@gts.uvigo.es

Abstract—We develop channel estimation agorithms for millimeter wave (mmWave) multiple input multiple output (MIMO) systems with one-bit analog-to-digital converters (ADCs). Since the mmWave MIMO channel is sparse due to the propagation characteristics, the estimation problem is formulated as a one-bit compressed sensing problem. We propose a modified EM algorithm that exploits sparsity and has better performance than the conventional EM algorithm. We also present a second solution using the generalized approximate message passing (GAMP) algorithm to solve this optimization problem. The simulation results show that GAMP can reduce mean squared error in the important low and medium SNR regions.

I. INTRODUCTION

Millimeter wave (mmWave) communications is a promising technology for the future outdoor cellular systems due to its huge bandwidth. As larger bandwidths are used, the corresponding sampling rate of the analog-to-digital converters (ADCs) scales up. High speed (e.g., ≥ 1 GSamples/s), high precision (e.g., ≥ 6 bits) ADCs are either unavailable, or may be too costly and power-hungry for portable devices [1], [2]. A possible solution is to simply employ one-bit ADCs which have low power consumption and cost.

In this paper, we consider a multiple-input multiple-output (MIMO) systems with one-bit ADCs, which only output the signs of the inphase and quadrature baseband signals. The main advantage of this architecture is the ADCs can be implemented with very low power consumption [3]. It also simplifies the overall complexity of the circuit, for example automatic gain control may not be required. The use of one-bit ADCs results in performance loss of the communication systems compared to perfect ADCs (the ADCs with infinite precision). The capacity of one-bit quantized single-input single-output (SISO) channel was studied in [4] where it was shown that at low SNR, there is only -1.96 dB power loss due to one bit quantization. The capacity in the high SNR regime was studied in our previous work [5]. We found that the high SNR capacity of SIMO channel grows logarithmically with the number of the receiver antennas. For the mmWave MIMO channel, the high SNR capacity increases with the number of paths in the channel.

In our previous work [5], we proposed to design the input constellation to maximize the channel capacity. To perform such optimization, channel state information (CSI) is required

at the transmitter. An effective way to estimate the SISO channel is to use dithering for combating the severe non-linearity of low-precision quantization [6]–[8]. For the MIMO channel, a pilot-based channel estimation method using expectation-maximization (EM) algorithm was proposed and analyzed in [9]–[11]. These papers [7]–[11] considered the MIMO channel with small antenna arrays in lower frequency UHF (ultra high frequency) band and thus did not take into account specific features of mmWave channels.

In this paper, we propose channel estimation methods for mmWave MIMO channels when one-bit ADCs are used at the receiver. In the mmWave channel, there are fewer paths in the channel and large antenna arrays are deployed at the transmitter and receiver [12]. Therefore the channel is 'sparse' in the angular domain and assuming perfect sparsity we propose to design low complexity channel estimation algorithms. In [13], an adaptive algorithm was proposed to estimate the mmWave channel parameters with hybrid analog/digital beamforming and perfect ADCs. Here, we consider the channel estimation problem with one-bit ADCs and all digital combining at the receiver. A modified EM algorithm exploiting the sparsity is proposed and achieves better performance than the conventional EM algorithm. In addition, we apply an efficient algorithm called generalized approximate message passing (GAMP) [14] to our problem. The GAMP algorithm has the best performance in the low and medium SNR regions.

Notation : a is a scalar, $\mathbf{a}$ is a vector and $\mathbf{A}$ is a matrix. $\mathrm{tr}(\mathbf{A})$, $\mathbf{A}^T$, $\mathbf{A}^*$ and $||\mathbf{A}||_F$ represent the trace, transpose, conjugate transpose and Frobenius norm of a matrix $\mathbf{A}$, respectively. $\mathbf{A} \otimes \mathbf{B}$ denotes the Kronecker product of $\mathbf{A}$ and $\mathbf{B}$. $\mathrm{vec}(\mathbf{A})$ is a vector stacking all the columns of $\mathbf{A}$.

II. SYSTEM MODEL

Consider the $N_\mathrm{t} \times N_\mathrm{r}$ MIMO system with one-bit quantization shown in Fig. 1. The narrowband baseband received signal is

$$\mathbf{y} = \mathbf{Hx} + \mathbf{n}, \tag{1}$$

where $\mathbf{H} \in \mathbb{C}^{N_\mathrm{r} \times N_\mathrm{t}}$ is the channel matrix, $\mathbf{x} \in \mathbb{C}^{N_\mathrm{t} \times 1}$ is the transmitted symbol with power $\mathbb{E}[\mathbf{x}^*\mathbf{x}] = P_\mathrm{T}$, $\mathbf{y} \in \mathbb{C}^{N_\mathrm{r} \times 1}$ is the received symbol and $\mathbf{n} \sim \mathcal{CN}(0, \mathbf{I})$ is the noise. With the one-bit ADCs, the receiver will obtain
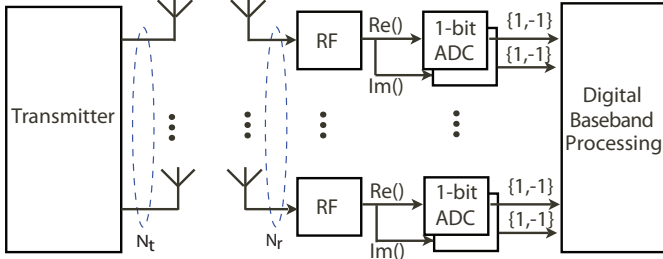
$$\mathbf{r} = \mathrm{sgn}(\mathbf{y}), \tag{2}$$

Fig. 1. A $N_\mathrm{t} \times N_\mathrm{r}$ MIMO system with one-bit quantization at the receiver. For each receiver antenna, there are two one-bit ADCs. Note that there is no limitation on the structure of the transmitter.

where $\mathrm{sgn}(\cdot)$ is the signum function applied component-wise and separately to the real and imaginary parts. The quantization output at the ith antenna is $r_i \in \{1+\mathrm{j}, 1-\mathrm{j}, -1+\mathrm{j}, -1-\mathrm{j}\}$ where $\mathrm{j} = \sqrt{-1}$. Our aim is to estimate the channel $\mathbf{H}$ based on our observations $\mathbf{r}$ and the training signal $\mathbf{x}$.

The mmWave channel can be modelled using a ray-based model with L paths. Denote α_ℓ, $\varphi_{\mathrm{r}\ell}(\mathrm{or}\,\theta_{\mathrm{r}\ell})$, $\varphi_{\mathrm{t}\ell}(\mathrm{or}\,\theta_{\mathrm{t}\ell})$ as the strengths, the azimuth (or elevation) angles of arrival and the angle of departure of the ℓth path, respectively. In the case of a uniform planar array (UPA) in the yz-plane consisting of $Y \times Z$ elements on the y and z axes respectively, the array response vector at the transmitter $\mathbf{a}_\mathrm{t}$ or receiver $\mathbf{a}_\mathrm{r}$ is given by [15]

$$\mathbf{a}(\varphi, \theta) = \frac{1}{\sqrt{N}} \Big[1, \cdots, e^{\mathrm{j}\frac{2\pi}{\lambda}d(m\sin(\varphi)\sin(\theta)+n\cos(\theta))}, \cdots,$$
$$e^{\mathrm{j}\frac{2\pi}{\lambda}d((Y-1)\sin(\varphi)\sin(\theta)+(Z-1)\cos(\theta))} \Big]^T, \quad (3)$$

where λ is the wavelength, d is the inter-element spacing, $0 \le m < Y$ and $0 \le n < Z$ are the y and z indices of an antenna element respectively. Hence, the channel matrix is,

$$\mathbf{H} = \sum_{\ell=1}^{L} \alpha_\ell \mathbf{a}_\mathrm{r}(\varphi_{\mathrm{r}\ell}, \theta_{\mathrm{r}\ell}) \mathbf{a}_\mathrm{t}^*(\varphi_{\mathrm{t}\ell}, \theta_{\mathrm{t}\ell}) \quad (4)$$

$$= \mathbf{A}_\mathrm{r} \mathbf{\Sigma} \mathbf{A}_\mathrm{t}^*, \quad (5)$$

where $\mathbf{A}_\mathrm{r} = [\mathbf{a}_\mathrm{r}(\varphi_{\mathrm{r}1}, \theta_{\mathrm{r}1}), \mathbf{a}_\mathrm{r}(\varphi_{\mathrm{r}2}, \theta_{\mathrm{r}2}), \cdots \mathbf{a}_\mathrm{r}(\varphi_{\mathrm{r}L}, \theta_{\mathrm{r}L})]$, $\mathbf{\Sigma} = \mathrm{diag}(\alpha_1, \alpha_2, \cdots, \alpha_L)$ and $\mathbf{A}_\mathrm{t} = [\mathbf{a}_\mathrm{t}(\varphi_{\mathrm{t}1}, \theta_{\mathrm{t}1}), \mathbf{a}_\mathrm{t}(\varphi_{\mathrm{t}2}, \theta_{\mathrm{t}2}), \cdots \mathbf{a}_\mathrm{t}(\varphi_{\mathrm{t}L}, \theta_{\mathrm{t}L})]$.

The mmWave channel will mostly be line-of-sight (LOS), near LOS, or consists of a single reflected path [16]. Large antenna arrays are usually deployed to obtain beamforming gain for combating the higher path loss. It seems plausible to have 256 antennas at the base station and 32 antennas at the mobile station in future mmWave cellular networks [17]. Hence, we usually have $L \ll \min\{N_\mathrm{r}, N_\mathrm{t}\}$.

The virtual channel representation [18] of $\mathbf{H}$ is,

$$\mathbf{H} = \mathbf{U}_\mathrm{r} \mathbf{H}_\mathrm{v} \mathbf{U}_\mathrm{t}^*, \quad (6)$$

where $\mathbf{U}_\mathrm{r} \in \mathbb{C}^{N_\mathrm{r} \times N_\mathrm{r}}$ and $\mathbf{U}_\mathrm{t} \in \mathbb{C}^{N_\mathrm{t} \times N_\mathrm{t}}$ are unitary Discrete Fourier Transform (DFT) matrices, $\mathbf{H}_\mathrm{v} \in \mathbb{C}^{N_\mathrm{r} \times N_\mathrm{t}}$ is the virtual channel matrix. The Fourier transformation in (6) can be seen as a mapping from the antenna domain onto an
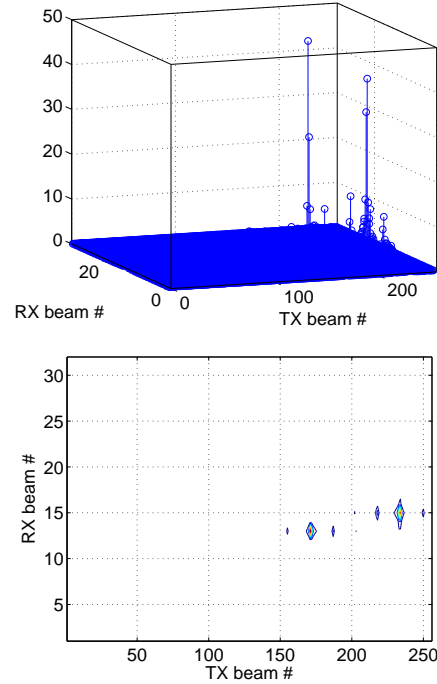

Fig. 2. This figure shows the amplitude of each entry in the virtual channel $\mathbf{H}_\mathrm{v}$ in a 256×32 mmWave channel with 2 paths. The transmitter is equipped with a 16×16 UPA and the receiver is equipped with a 16×2 UPA.

angular domain, and the entries of the matrix $\mathbf{H}_\mathrm{v}$ can be interpreted as the channel gains between the N_t transmit and the N_r receive beams. Fig. 2 shows an example of $\mathbf{H}_\mathrm{v}$ in a 256×32 mmWave channel with 2 paths. The virtual channel $\mathbf{H}_\mathrm{v}$ contains 2 clusters of entries with large amplitude, each of which corresponds to one path [18].

In this paper, we make a strong assumption that the array response vectors are along the directions defined in the DFT matrices. With this assumption, $\mathbf{H}_\mathrm{v}$ is perfectly sparse, i.e., has exactly L nonzero entries. Each element of $\mathbf{H}_\mathrm{v}$, denoted as h_{ij}, is assumed to follow the Bernoulli-Gaussian distribution. The prior information about $\mathbf{H}_\mathrm{v}$ is,

$$h_{ij} = (1-\eta)\delta(h_{ij}) + \frac{\eta}{\pi\sigma_L^2} e^{-\frac{|h_{ij}|^2}{\sigma_L^2}} \quad (7)$$

where $\delta(\cdot)$ is the delta function, $\eta = \frac{L}{N_\mathrm{t} N_\mathrm{r}}$ represents the probability that $h_{ij} \neq 0$ and σ_L^2 is determined by the average channel gain.

III. MAIN RESULTS

Noting that the channel $\mathbf{H}$ has the structure in (6), we propose a heuristic method to design the training sequence, denoted as $\mathbf{X} \in \mathbb{C}^{N_\mathrm{t} \times K}$ where K is the length of the training sequence. Let $\mathbf{X} = \mathbf{U}_\mathrm{t} \mathbf{Z}$. Then, the received signal at the receiver is

$$\mathbf{R} = \mathrm{sgn}(\mathbf{H}\mathbf{X} + \mathbf{N}) \quad (8)$$

$$= \mathrm{sgn}(\mathbf{U}_\mathrm{r} \mathbf{H}_\mathrm{v} \mathbf{U}_\mathrm{t}^* \mathbf{U}_\mathrm{t} \mathbf{Z} + \mathbf{N}) \quad (9)$$

$$= \mathrm{sgn}(\mathbf{U}_\mathrm{r} \mathbf{H}_\mathrm{v} \mathbf{Z} + \mathbf{N}), \quad (10)$$

where $\mathbf{N} \in \mathbb{C}^{N_t \times K}$ is the i.i.d. Gaussian noise.

Through vectorization,

$$
\begin{aligned}
\mathrm{vec}(\mathbf{R}) &= \mathrm{sgn}\left(\mathrm{vec}\left(\mathbf{U}_r \mathbf{H}_v \mathbf{Z} + \mathbf{N}\right)\right) \\
&\overset{(a)}{=} \mathrm{sgn}\left(\left(\mathbf{Z}^T \otimes \mathbf{U}_r\right)\mathrm{vec}(\mathbf{H}_v) + \mathrm{vec}(\mathbf{N})\right), \quad (11)
\end{aligned}
$$

where (a) follows from the equality $\mathrm{vec}(\mathbf{ABC}) = \left(\mathbf{C}^T \otimes \mathbf{A}\right)\mathrm{vec}(\mathbf{B})$.

We rewrite the complex-valued (11) in real-valued form as

$$
\bar{\mathbf{r}} = \mathrm{sgn}\left(\mathbf{Wh} + \bar{\mathbf{n}}\right), \quad (12)
$$

where

$$
\bar{\mathbf{r}} = \begin{bmatrix} \mathbf{Re}(\mathrm{vec}(\mathbf{R})) \\ \mathbf{Im}(\mathrm{vec}(\mathbf{R})) \end{bmatrix}, \mathbf{h} = \begin{bmatrix} \mathbf{Re}(\mathrm{vec}(\mathbf{H}_v)) \\ \mathbf{Im}(\mathrm{vec}(\mathbf{H}_v)) \end{bmatrix},
$$

$$
\bar{\mathbf{n}} = \begin{bmatrix} \mathbf{Re}(\mathrm{vec}(\mathbf{N})) \\ \mathbf{Im}(\mathrm{vec}(\mathbf{N})) \end{bmatrix},
$$

and
$$
\mathbf{W} = \begin{bmatrix} \mathbf{Re}(\mathbf{Z}^T \otimes \mathbf{U}_r) & -\mathbf{Im}(\mathbf{Z}^T \otimes \mathbf{U}_r) \\ \mathbf{Im}(\mathbf{Z}^T \otimes \mathbf{U}_r) & \mathbf{Re}(\mathbf{Z}^T \otimes \mathbf{U}_r) \end{bmatrix}.
$$

The channel estimation problem is: estimate $\mathbf{h}$ given the equivalent training matrix $\mathbf{Z}$, DFT matrix $\mathbf{U}_r$ and the quantized noisy received signal $\bar{\mathbf{r}}$. The problem is similar to the *one-bit compressed sensing problem* [19]–[21].

The conditional probability of $\bar{\mathbf{r}}$ given $\mathbf{h}$ is

$$
p\left(\bar{\mathbf{r}}|\mathbf{h}\right) = \prod_{i=1}^{2N_r} \Phi\left(\frac{\bar{r}_i \mathbf{w}_i^T \mathbf{h}}{\sigma}\right), \quad (13)
$$

where $\mathbf{w}_i^T$ is the ith row of $\mathbf{W}$, $\sigma^2 = \frac{1}{2}$ is the variance of the real-valued noise $\bar{\mathbf{n}}$ and $\Phi(\cdot)$ is the normal cumulative distribution function. By setting the derivative of the $\log p\left(\bar{\mathbf{r}}|\mathbf{h}\right)$ to zero, the maximal-likelihood (ML) estimator $\widehat{\mathbf{h}}_{\mathrm{ML}}$ should satisfy the following equation,

$$
\sum_{i=1}^{2N_r} \mathbf{w}_i \bar{r}_i \frac{e^{-\frac{(\mathbf{w}_i^T \widehat{\mathbf{h}})^2}{2\sigma^2}}}{\Phi\left(\frac{\bar{r}_i \mathbf{w}_i^T \widehat{\mathbf{h}}}{\sigma}\right)} = \mathbf{0}. \quad (14)
$$

It is hard to find a closed form expression of $\widehat{\mathbf{h}}_{\mathrm{ML}}$ except for simple SISO and 2×2 MIMO channel [10], [11].

In this paper, we implement three iterative algorithms to estimate $\mathbf{h}$. The first one is the EM algorithm proposed in [9]–[11]. The steps of the EM algorithm are shown in Algorithm 1. The EM algorithm is sensitive to the initial value and may converge to a local optimum. Note that the prior information that $\mathbf{h}$ is sparse is not used in the EM algorithm.

The second algorithm is the modified EM algorithm. The proposed algorithm is same as the EM algorithm except that in the maximization step, matching pursuit method is used to find the major nonzero components of $\mathbf{h}$. The sparsity of $\mathbf{h}$ is exploited to reduce the mean estimation error. In the simulations, we also find this modified EM algorithm is insensitive to the initialization.

The last one is the *generalized approximate message passing* (GAMP) algorithm [14], which decomposes the vector-valued estimation problem into a sequence of scalar problems. GAMP is applicable to estimation problems with linear transform and component-wise nonlinearities. In addition, GAMP converges very fast (less than 25 steps in most cases). In the GAMP algorithm, the prior information about the distribution of $\mathbf{h}$ in (7) is exploited. The parameters η and σ_L^2 can be learned by EM-GAMP algorithm [22] if they are unknown.

Algorithm 1 Expectation-Maximization algorithm

1: **Initialize** $\widehat{\mathbf{h}}$ to a random vector.
2: **Repeat**
  (a) Expectation step:
    Compute the minimum mean square estimator of $\mathbf{y}$:

$$
\begin{aligned}
\widehat{y}_i &= \mathbb{E}\left[y_i | \bar{\mathbf{r}}, \mathbf{W}, \mathbf{h} = \widehat{\mathbf{h}}\right] \quad (15) \\
&= \mathbf{w}_i^T \widehat{\mathbf{h}} + \bar{r}_i \frac{\sigma}{\sqrt{2\pi}} \frac{e^{-\frac{(\mathbf{w}_i^T \widehat{\mathbf{h}})^2}{2\sigma^2}}}{\Phi\left(\frac{\bar{r}_i \mathbf{w}_i^T \widehat{\mathbf{h}}}{\sigma}\right)}. \quad (16)
\end{aligned}
$$

  (b) Maximization step:
    Compute the maximum likelihood estimator of $\mathbf{h}$:

$$
\widehat{\mathbf{h}} = \left(\mathbf{W}^T \mathbf{W}\right)^{-1} \mathbf{W}^T \mathbf{y}. \quad (17)
$$

3: **Until** $\widehat{\mathbf{h}}$ converges.

Algorithm 2 Modified Expectation-Maximization algorithm

1: **Initialize** $\widehat{\mathbf{h}}$ to a random vector.
2: **Repeat**
  (a) Expectation step:
    Compute the minimum mean square estimator of $\mathbf{y}$ as in the EM algorithm.
  (b) Use matching pursuit method to find $\widehat{\mathbf{h}}$.
3: **Until** $\widehat{\mathbf{h}}$ converges.

IV. SIMULATION RESULTS

We consider the channel estimation problem in a 32×8 mmWave MIMO system. The transmitter is equipped with a 8×4 UPA and the receiver is equipped with a 4×2 UPA. Throughout the simulations, $\mathbb{E}[||\mathbf{H}||_F^2] = N_t N_r$.

The training matrix $\mathbf{Z} \in \mathbb{C}^{N_t \times K}$ is composed of rows of Walsh-Hadamard matrix. Therefore, $\mathbf{W}^T \mathbf{W} = \frac{K P_T}{N_t} \mathbf{I}_{2N_t N_r}$ is a diagonal matrix and the computation of (17) can be simplified.

In the simulations, we compare several different algorithms:
i) GAMP with one-bit ADCs,
ii) EM algorithm with one-bit ADCs,
iii) Modified EM algorithm with one-bit ADCs,
iv) Lasso algorithm [23] with perfect ADCs,
v) ML estimation with perfect ADCs.

In the last two schemes, we assume that there are perfect ADCs. In other words, the receiver knows $\mathbf{y}$ instead of $\mathbf{r}$.

The performance metric is the normalized mean square error (MSE), given by $\mathbb{E}\left[\frac{||\widehat{\mathbf{h}} - \mathbf{h}||^2}{||\mathbf{h}||^2}\right]$ where $\widehat{\mathbf{h}}$ is the estimation of $\mathbf{h}$.

In Figs. 3, 4 and 5, we compare these algorithms when the length of the training sequence is 64, 128 and 256,
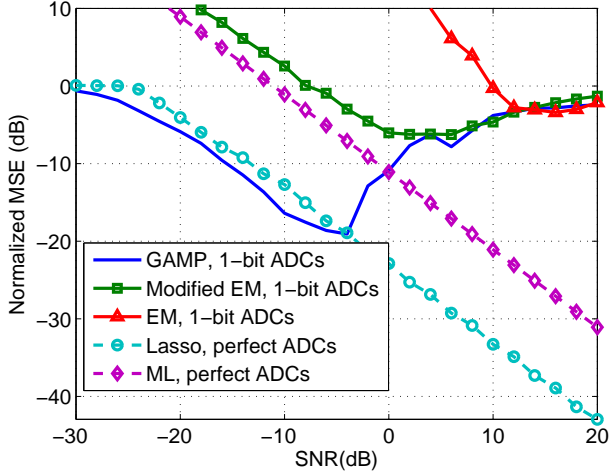
Fig. 3. Performance comparison of different algorithms when there are $L = 2$ paths in the channel and the training signal has length $K = 64$.
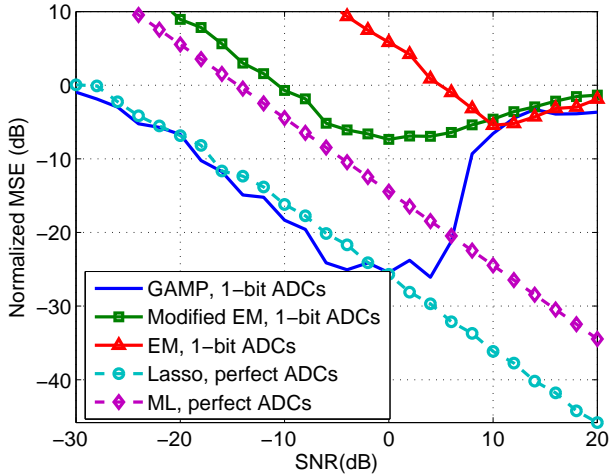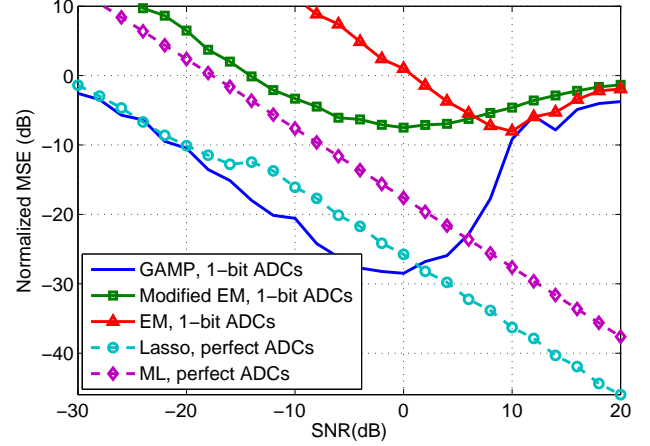

Fig. 5. Performance comparison of different algorithms when there are $L = 2$ paths in the channel and the training signal has length $K = 256$.

Fig. 4. Performance comparison of different algorithms when there are $L = 2$ paths in the channel and the training signal has length $K = 128$.

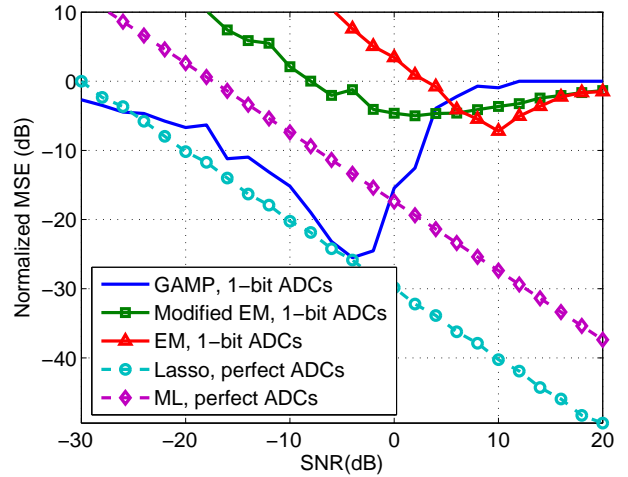

Fig. 6. Performance comparison of different algorithms when there are $L = 1$ paths in the channel and the training signal has length $K = 128$.

respectively. In these figures, $L = 2$ which means that there are 2 non-zero entries in the virtual channel realization $\mathbf{H_v}$.

First, we find that the GAMP algorithm has much better performance than EM and modified EM algorithm, especially in the low and medium SNR regions. Actually, when the SNR is less than 0 dB, the normalized MSE of GAMP algorithm is even close to that of Lasso algorithm with perfect ADCs. Thus, GAMP is a promising algorithm for channel estimation with one-bit ADCs.

Second, the modified EM algorithm has better performance than the conventional EM algorithm in the low and medium SNR regimes. The reason is that the sparsity of $\mathbf{h}$ is taken into account in the modified EM algorithm. For the same reason, the Lasso algorithm with perfect ADCs performs better than the ML estimation with perfect ADCs.

Third, as the length increases from $K = 64$ to 128 and 256,

there is about 3 dB and 6 dB power gain, respectively.

Last, we find that for the GAMP and EM and modified EM algorithms, there exists an optimal SNR achieving the minimum normalized MSE. The reason is that the one-bit ADC is a highly nonlinear system and noises may be helpful in this kind of nonlinear systems. At high SNR, $\alpha\mathbf{h}$ ($\alpha > 1$) and $\mathbf{h}$ will lead to almost the same one-bit quantization output $\mathbf{r}$. Therefore, $\alpha\mathbf{h}$ and $\mathbf{h}$ are not distinguishable and the amplitude information is lost. However, in the low and medium SNR, $\alpha\mathbf{h}$ and $\mathbf{h}$ will lead to different quantization outputs and the amplitude of $\mathbf{h}$ can be somehow recovered. This phenomenon is known as *stochastic resonance* [24]. For the ML estimator without quantization, which is a linear system, the MSE decreases as SNR increases as expected.

Figs. 6 and 7 show the performance of these three schemes when there is 1 path and 4 paths in the channel, respectively. We find that the performance of the algorithms is similar to
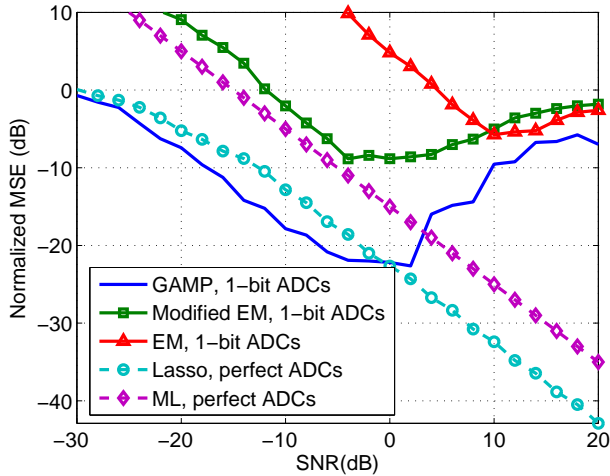
Fig. 7. Performance comparison of different algorithms when there are $L = 4$ paths in the channel and the training signal has length $K = 128$.

the case of $L = 2$ shown in Fig. 4.

V. CONCLUSION

In the paper, we proposed a solution for channel estimation in the mmWave MIMO channel with one-bit ADCs. The mmWave MIMO channel has the property of sparsity and the compressed sensing techniques can be used. We proposed a modified EM algorithm, which take into account the sparsity and has better performance than the conventional EM algorithm. We also proposed to use a computationally efficient algorithm, i.e., GAMP algorithm. In the simulations, we found that GAMP has the best performance in the low and medium SNR regions.

A key assumption in our paper is that each entry in the virtual channel $\mathbf{H}_v$ follows i.i.d. Bernoulii Gaussian distribution. If the array response vector does not exactly follow the directions defined in the two DFT matrices, there is 'leakage' and each path does not correspond to a single element in $\mathbf{H}_v$. In addition, each path has angular spread in the mmWave channel and may correspond to several adjacent entries. Our future work will consider these cases. A possible solution is to make $\mathbf{H}_v$ sparse by a windowing approach [25].

VI. ACKNOWLEDGEMENT

REFERENCES

[1] B. Le, T. Rondeau, J. Reed, and C. Bostian, "Analog-to-digital converters," *IEEE Signal Process. Mag.*, vol. 22, no. 6, pp. 69–77, 2005.

[2] B. Murmann, "ADC performance survey 1997-2014." [Online]. Available: http://www.stanford.edu/ murmann/adcsurvey.html

[3] T. Sundstrom, B. Murmann, and C. Svensson, "Power dissipation bounds for high-speed Nyquist analog-to-digital converters," *IEEE Trans. Circuits Syst. I*, vol. 56, no. 3, pp. 509–518, 2009.

[4] A. Mezghani and J. Nossek, "On ultra-wideband MIMO systems with 1-bit quantized outputs: Performance analysis and input optimization," in *IEEE International Symposium on Information Theory*, 2007, pp. 1286–1289.

[5] J. Mo and R. Heath, "High SNR capacity of millimeter wave MIMO systems with one-bit quantization," in *Proc. of Information Theory and Applications (ITA) Workshop*, 2014.

[6] O. Dabeer and A. Karnik, "Signal parameter estimation using 1-bit dithered quantization," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5389–5405, 2006.

[7] O. Dabeer and U. Madhow, "Channel estimation with low-precision analog-to-digital conversion," in *IEEE International Conference on Communications (ICC)*, 2010, pp. 1–6.

[8] G. Zeitler, G. Kramer, and A. Singer, "Bayesian parameter estimation using single-bit dithered quantization," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 2713–2726, 2012.

[9] T. Lok and V.-W. Wei, "Channel estimation with quantized observations," in *IEEE International Symposium on Information Theory*, Aug 1998, p. 333.

[10] M. T. Ivrlac and J. A. Nossek, "On MIMO channel estimation with single-bit signal-quantization," in *ITG Smart Antenna Workshop*, 2007.

[11] A. Mezghani, F. Antreich, and J. Nossek, "Multiple parameter estimation with quantized channel output," in *2010 International ITG Workshop on Smart Antennas (WSA)*, 2010, pp. 143–150.

[12] M. Akdeniz, Y. Liu, M. Samimi, S. Sun, S. Rangan, T. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, June 2014.

[13] A. Alkhateeb, O. El Ayach, G. Leus, and R. 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.

[14] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *2011 IEEE International Symposium on Information Theory Proceedings (ISIT)*, July 2011, pp. 2168–2172.

[15] C. A. Balanis, *Antenna Theory: analysis and design*. John Wiley & Sons, 2012.

[16] A. Ghosh, T. Thomas, M. Cudak, R. Ratasuk, P. Moorut, F. Vook, T. Rappaport, G. Maccartney, S. Sun, and S. Nie, "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, June 2014.

[17] W. Hong, K.-H. Baek, Y. Lee, Y. Kim, and S.-T. Ko, "Study and prototyping of practically large-scale mmWave antenna systems for 5G cellular devices," *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 63–69, September 2014.

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

[19] P. Boufounos and R. Baraniuk, "1-bit compressive sensing," in *42nd Annual Conference on Information Sciences and Systems*, March 2008, pp. 16–21.

[20] L. Jacques, J. Laska, P. Boufounos, and R. Baraniuk, "Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors," *IEEE Trans. Inf. Theory*, vol. 59, no. 4, pp. 2082–2102, April 2013.

[21] J. Ziniel, P. Schniter, and P. Sederberg, "Binary linear classification and feature selection via generalized approximate message passing," in *Information Sciences and Systems (CISS), 2014 48th Annual Conference on*, March 2014, pp. 1–6.

[22] J. Vila and P. Schniter, "Expectation-Maximization Gaussian-Mixture Approximate Message Passing," *IEEE Trans. Signal Process.*, vol. 61, no. 19, pp. 4658–4672, Oct 2013.

[23] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.

[24] M. D. McDonnell, N. G. Stocks, C. E. M. Pearce, and D. Abbott, *Stochastic resonance: from suprathreshold stochastic resonance to stochastic signal quantization*. Cambridge University Press, 2008.

[25] P. Schniter and A. Sayeed, "A sparseness-preserving virtual MIMO channel model," in *Proc. 38th Annu. Conf. Inf. Sci. Syst.(CISS04)*, pp. 36–41.