Sparsity-Adaptive Beamspace Channel Estimation for 1-Bit mmWave Massive MIMO Systems

Alexandra Gallyas-Sanhueza[1], Seyed Hadi Mirfarshbafan[1], Ramina Ghods[2], and Christoph Studer[1]

[1]*Cornell Tech, New York, NY;* [2]*Carnegie Mellon University, Pittsburgh, PA*

Abstract—We propose sparsity-adaptive beamspace channel estimation algorithms that improve accuracy for 1-bit data converters in all-digital millimeter-wave (mmWave) massive multiple-input multiple-output (MIMO) basestations. Our algorithms include a tuning stage based on Stein's unbiased risk estimate (SURE) that automatically selects optimal denoising parameters depending on the instantaneous channel conditions. Simulation results with line-of-sight (LoS) and non-LoS mmWave massive MIMO channel models show that our algorithms improve channel estimation accuracy with 1-bit measurements in a computationally-efficient manner.

I. INTRODUCTION

Millimeter-wave (mmWave) and massive multi-user (MU) multiple-input multiple-output (MIMO) will be core technologies for future wireless systems [1], [2]. The combination of these technologies enables simultaneous communication to multiple user equipments (UEs) at unprecedentedly high data rates. These advantages come at the cost of significantly increased power consumption, implementation complexity, and system costs. A viable solution to address these challenges is the use of low-resolution data converters combined with sophisticated but efficient baseband processing algorithms in all-digital basestations (BS) architectures [3]–[7].

A. Channel Estimation with Low-Resolution Data Converters

Coarse quantization of the received baseband samples, due to the use of low-resolution analog-to-digital converters (ADCs) at the BS, together with the high path loss at mmWave or terahertz (THz) frequencies [8], [9], renders the acquisition of accurate channel estimates a challenging task. Fortunately, wave propagation at mmWave or THz frequencies is directional [10] and channels typically consist only of a few dominant propagation paths [2], [11]. Both of these properties cause the channel vectors to be sparse in the beamspace domain, which can be exploited to perform denoising that improves reliability of data transmission [12]–[16].

Practical sparsity-exploiting channel denoising methods for mmWave massive MU-MIMO systems must exhibit low computational complexity due to the large number of BS antenna elements and the potentially large number of UEs that commmunicate simultaneously. A low-complexity

mmWave channel denoising algorithm called BEACHES (short for beamspace channel estimation) has been proposed recently in [16]. This method has orders-of-magnitude lower complexity than state-of-the-art denoising methods, such as atomic norm minimization (ANM) [17] and Newtonized orthogonal matching pursuit (NOMP) [18]. However, all of these existing denoising methods perform poorly when denoising channel vectors that were acquired through low-resolution data converters. Channel estimation with 1-bit ADCs has been analyzed in [4], [5], [19]–[21]. Beamspace sparsity of mmWave channels has been exploited to denoise channel vectors from 1-bit measurements in [6], [22], [23]. However, all of these denoising methods exhibit high complexity, ignore beamspace sparsity, and/or require a number of parameters that must be adapted to the instantaneous propagation conditions, such as the number of dominant propagation paths.

B. Contributions

We propose low-complexity channel estimation algorithms for mmWave massive MU-MIMO systems that operate with 1-bit data converters. By using a Bussgang-like decomposition [24] of the 1-bit measurement process, our methods adapt the optimal denoising parameters to the channel's instantaneous sparsity via Stein's unbiased risk estimate (SURE). We propose two methods that build upon BEACHES put forward in [16] and a novel method, referred to as Sparsity-Adaptive oNe-bit Denoiser (SAND), which automatically tunes two algorithm parameters to minimize the channel estimation mean-square error (MSE). To demonstrate the efficacy of our channel estimation algorithms, we perform MSE and bit error rate (BER) simulations with line-of-sight (LoS) and non-LoS mmWave channels in a massive MU-MIMO system.

C. Notation

Lowercase and uppercase boldface letters denote column vectors and matrices, respectively. The kth entry of the vector $\mathbf{a}$ is a_k; the real and imaginary parts are $[\mathbf{a}]_\mathcal{R} = \mathcal{R}\{\mathbf{a}\}$ and $[\mathbf{a}]_\mathcal{I} = \mathcal{I}\{\mathbf{a}\}$, respectively. For a matrix $\mathbf{A}$, its transpose and Hermitian transpose are $\mathbf{A}^\mathrm{T}$ and $\mathbf{A}^\mathrm{H}$, respectively. A complex Gaussian vector $\mathbf{a}$ with mean $\mathbf{m}$ and covariance $\mathbf{K}$ is written as $\mathbf{a} \sim \mathcal{CN}(\mathbf{m}, \mathbf{K})$. Expectation is denoted by $\mathbb{E}[\cdot]$.

II. 1-BIT QUANTIZED SYSTEM MODEL

We consider a mmWave massive MU-MIMO uplink system in which U single-antenna UEs transmit data to a B-antenna BS equipped with a uniform linear array (ULA). We assume

that each of the B radio-frequency (RF) chains at the BS contains a pair of 1-bit ADCs that separately quantize the in-phase and quadrature signals. A widely-used, yet simplistic channel vector model for such systems is as follows [25]:

$$\mathbf{h} = \sum_{\ell=1}^{L} \kappa_\ell \mathbf{a}(\Omega_\ell), \quad \mathbf{a}(\Omega) = [e^{j0\Omega}, e^{j1\Omega}, \ldots, e^{j(B-1)\Omega}]^{\mathrm{T}}. \quad (1)$$

Here, L stands for the number of propagation paths arriving at the BS, $\kappa_\ell \in \mathbb{C}$ is the channel gain of the ℓth path, $\mathbf{a}(\Omega_\ell) \in \mathbb{C}^B$ contains the relative phases between BS antennas, and $\Omega_\ell \in [0, 2\pi)$ is determined by the ℓth path's incident angle. We emphasize that our simulation results in Section V will use more realistic mmWave channel models.

We consider orthogonal training-based channel estimation, where only one UE transmits a pilot at a time—a generalization to other training schemes is part of ongoing work. To model 1-bit ADCs, we define $Q(z) = \mathrm{sign}(\mathcal{R}\{z\}) + j\,\mathrm{sign}(\mathcal{I}\{z\})$, which is applied element-wise to vectors. The 1-bit quantized channel vector for a given UE can be modeled as follows [5]:

$$\mathbf{r} = Q(\varrho \mathbf{h} + \mathbf{n}), \quad (2)$$

where $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, N_0 \mathbf{I}_B)$ models thermal noise. Without loss of generality, we assume $\varrho = 1$ for the rest of the paper.

All of the above vectors are in the antenna domain, where each entry is associated with one of the B BS antennas. By taking the discrete Fourier transform (DFT) across the antenna array, we can transform these vectors into the beamspace domain, where each entry corresponds to an incident angle. From (1) we see that $\mathbf{h}$ is a superposition of L complex sinusoids. Consequently, the beamspace domain representation $\hat{\mathbf{h}} = \mathbf{F}\mathbf{h}$, where $\mathbf{F}$ is the $B \times B$ unitary DFT matrix, will be sparse assuming that $L \ll B$. In what follows, all beamspace domain quantities are designated with a hat.

Fig. 1 shows examples for LoS and non-LoS channel vectors in the beamspace domain without and with 1-bit quantization. Clearly, the unquantized beamspace vectors $\hat{\mathbf{h}}$ exhibit sparsity; the 1-bit quantized beamspace vectors, which are obtained from $\hat{\mathbf{r}} = \mathbf{F}\mathbf{r}$, also exhibit sparsity but, in addition, are distorted by quantization artifacts. We also observe that the quantization artifacts differ significantly between the LoS and non-LoS channels, which exhibit different levels of sparsity. In what follows, we develop algorithms that exploit beamspace sparsity to denoise 1-bit quantized channel vectors while adapting the denoising parameters to the instantaneous channel sparsity.

III. BEACHES-BASED 1-BIT DENOISING

Before we discuss denoising methods for 1-bit measurements, we briefly review the BEACHES algorithm in [16], which was developed for systems with high-resolution data converters. Assume that we observe a noisy measurement of the channel vector $\mathbf{h}$ in the beamspace domain as

$$\hat{\mathbf{y}} = \hat{\mathbf{h}} + \hat{\mathbf{e}}, \quad (3)$$

where $\hat{\mathbf{e}} \sim \mathcal{CN}(\mathbf{0}, E_0 \mathbf{I})$ models channel estimation errors. BEACHES denoises $\hat{\mathbf{y}}$ by applying the soft-thresholding function $\hat{\mathbf{h}}' = \eta(\hat{\mathbf{y}}, \tau)$ defined as

$$[\eta(\hat{\mathbf{y}}, \tau)]_b = \frac{\hat{y}_b}{|\hat{y}_b|} \max\{|\hat{y}_b| - \tau, 0\}, \quad b = 1, \ldots, B, \quad (4)$$
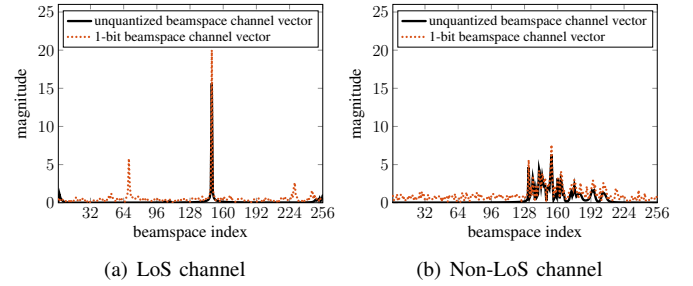

Fig. 1. Beamspace representation of an unquantized and a 1-bit quantized channel vector for a line-of-sight (LoS) and non-LoS scenario. The channel vectors are generated with the QuaDRiGa mmMAGIC UMi model [26] at 60 GHz for a 256 uniform-linear array (ULA) with $\lambda/2$ antenna spacing. The average energy has been normalized to 1 and no noise is present.

where we define $\hat{y}_b/|\hat{y}_b| = 0$ for $\hat{y}_b = 0$ and the parameter τ is the denoising threshold. For an optimally-chosen denoising threshold $\tau^\star$, the soft-thresholding function suppresses noise (which is typically weak) while leaving the sparse components that pertain to the channel vector mostly intact. We define $\tau^\star$ as the denoising threshold that minimizes the MSE

$$MSE = \tfrac{1}{B} \mathbb{E}[\|\hat{\mathbf{h}}' - \hat{\mathbf{h}}\|^2], \quad (5)$$

which determines the optimally-denoised channel vector $\hat{\mathbf{h}}^\star = \eta(\hat{\mathbf{y}}, \tau^\star)$ in the beamspace domain. However, the MSE expression depends on the unknown vector $\hat{\mathbf{h}}$. BEACHES circumvents this issue by using Stein's unbiased risk estimator (SURE) [27], which is an unbiased estimate of the MSE in (5) that does *not* depend on $\hat{\mathbf{h}}$. BEACHES requires (i) the channel estimation error $\hat{\mathbf{e}}$ to be i.i.d. Gaussian and (ii) knowledge of the channel estimation variance E_0 to optimally denoise the channel vector at a complexity that scales only with $B\log(B)$.

A. The 1-BEACHES Algorithm

We now present 1-BEACHES, which denoises the received 1-bit channel measurements using BEACHES. For this method, we model the 1-bit received vector $\mathbf{r}$ in (2) as

$$\mathbf{r} = Q(\mathbf{h} + \mathbf{n}) = \mathbf{h} + \mathbf{q}, \quad (6)$$

where the vector $\mathbf{q}$ depends on $\mathbf{h}$ and $\mathbf{n}$, and models quantization errors and noise. By transforming $\mathbf{r}$ into the beamspace domain, we have

$$\hat{\mathbf{r}} = \mathbf{F}\mathbf{r} = \hat{\mathbf{h}} + \hat{\mathbf{q}}. \quad (7)$$

Even though the vector $\mathbf{q}$ is not Gaussian distributed, the beamspace version $\hat{\mathbf{q}} = \mathbf{F}\mathbf{q}$ is well-approximated by a Gaussian random vector as each entry is a sum of all entries of $\mathbf{q}$ with different phases. To denoise the system in (7) with BEACHES, we need knowledge of the variance Q_0 of the entries in $\hat{\mathbf{q}}$. By assuming that $\hat{\mathbf{h}} = \mathbf{F}\mathbf{h}$ is circularly-symmetric, which is reasonable as $\mathbf{h}$ is a sum of complex sinusoids as modeled in (1), we obtain

$$Q_0 = \tfrac{1}{B} \mathbb{E}[\|\hat{\mathbf{q}}\|^2] = \tfrac{1}{B} \mathbb{E}[\|\mathbf{r}\|^2 + \|\mathbf{h}\|^2 - 2\mathcal{R}\{\mathbf{h}^{\mathrm{H}}\mathbf{r}\}]. \quad (8)$$

In order to obtain a closed-form expression of Q_0 with a minimal number of parameters, we further assume[1] that

---

[1]This assumption is accurate if the number of propagation paths L in (1) is large. As shown in Section V, this assumption is simplistic for LoS channels.

$\mathbf{h} \sim \mathcal{CN}(\mathbf{0}, E_h \mathbf{I}_B)$, which leads to $Q_0 = 2 + E_h - 4E_h/\sqrt{\pi(E_h + N_0)}$, where $E_h = \frac{1}{B}\mathbb{E}[\|\mathbf{h}\|^2]$ and $\mathbb{E}[\mathbf{h}^H\mathbf{r}]$ in (8) is computed in Appendix A. Under these assumptions, the beamspace representation (7) has the same form as (3), where $\hat{\mathbf{y}} = \hat{\mathbf{r}}$ and we model $\hat{\mathbf{e}} = \hat{\mathbf{q}} \sim \mathcal{CN}(\mathbf{0}, Q_0 \mathbf{I}_B)$, which allows us to (i) apply BEACHES to find the optimal denoising threshold $\tau^\star$ given $\hat{\mathbf{r}}$ and the variance Q_0, and (ii) compute $\hat{\mathbf{h}}^\star = \eta(\hat{\mathbf{r}}, \tau^\star)$. We call this procedure 1-BEACHES.

B. The α-BEACHES Algorithm

In the model (6), the error $\mathbf{q}$ will be large if the power of $\hat{\mathbf{h}}$ differs from the power of $\hat{\mathbf{r}}$. We now derive α-BEACHES which addresses this aspect. To this end, we use a Bussgang-like decomposition [24] that models the 1-bit ADCs as

$$\mathbf{r} = Q(\mathbf{h} + \mathbf{n}) = \alpha \mathbf{h} + \mathbf{d}, \tag{9}$$

where α is a scalar that minimizes the distortion variance $\mathbb{E}[\|\mathbf{d}\|^2]$ and also ensures $\mathbb{E}[\mathbf{d}^H\mathbf{h}] = 0$. By assuming that the vector $\mathbf{h}$ is circularly symmetric, we have

$$\alpha = \arg\min_{\alpha' \in \mathbb{C}} \mathbb{E}[\|\mathbf{r} - \alpha'\mathbf{h}\|^2] = \frac{\mathbb{E}[\mathbf{h}^H\mathbf{r}]}{\mathbb{E}[\|\mathbf{h}\|^2]}. \tag{10}$$

To obtain a closed-form expression for α, we assume $\mathbf{h} \sim \mathcal{CN}(\mathbf{0}, E_h \mathbf{I})$ as in 1-BEACHES and use the derivation of $\mathbb{E}[\mathbf{h}^H\mathbf{r}]$ in Appendix A, which yields $\alpha = 2/\sqrt{\pi(E_h + N_0)}$.

In (9), the distortion $\mathbf{d}$ is not Gaussian. By transforming into beamspace domain and dividing the result by α, we get

$$\frac{1}{\alpha}\hat{\mathbf{r}} = \frac{1}{\alpha}\mathbf{F}\mathbf{r} = \hat{\mathbf{h}} + \frac{1}{\alpha}\hat{\mathbf{d}}, \tag{11}$$

in which the distortion $\hat{\mathbf{d}}/\alpha$ is well-approximated by a Gaussian, as each entry is a scaled and phase-shifted sum of all of the entries of $\mathbf{d}$. The distortion variance D_0/α^2 is

$$\frac{1}{B}\frac{1}{\alpha^2}\mathbb{E}[\|\hat{\mathbf{d}}\|^2] = \frac{1}{B}\frac{1}{\alpha^2}\mathbb{E}[\|\mathbf{r}\|^2 - \alpha^2\|\mathbf{h}\|^2] = \frac{2}{\alpha^2} - E_h. \tag{12}$$

The model (11), enables us to apply BEACHES to $\hat{\mathbf{r}}/\alpha$ in order to determine the denoising threshold $\tau^\star$ given $\hat{\mathbf{r}}/\alpha$ and D_0/α^2. Finally, α-BEACHES computes $\hat{\mathbf{h}}^\star = \eta(\frac{\hat{\mathbf{r}}}{\alpha}, \tau^\star)$.

IV. SAND: SPARSITY-ADAPTIVE ONE-BIT DENOISER

As a generalized variant of α-BEACHES, we next develop a sparsity-adaptive method that *simultaneously* learns a prefactor γ and a denoising threshold τ in order to minimize the MSE. By defining our two-parameter estimator[2] as $\hat{\mathbf{h}}' = \gamma\eta(\hat{\mathbf{r}}, \tau)$, we aim to find the parameters $\gamma^\star$ and $\tau^\star$ that minimize the MSE in (5). Since the MSE depends on the unknown vector $\hat{\mathbf{h}}$, we select the optimal parameters $\gamma^\star$ and $\tau^\star$ that minimize SURE, which (i) is an unbiased estimator of the MSE so that $\mathbb{E}[SURE] = MSE$ and $\lim_{B \to \infty} SURE = MSE$ (see [16] for the details) and (ii) does *not* depend on $\hat{\mathbf{h}}$. For any weakly differentiable estimator $\mu(\hat{\mathbf{r}})$, using the decomposition (11) and assuming that $\hat{\mathbf{d}}$ is i.i.d. Gaussian, SURE is given by

$$SURE = \frac{1}{B}\|\mu(\hat{\mathbf{r}})\|^2 + \frac{2-D_0}{\alpha^2} - \frac{1}{B}\frac{2}{\alpha}\mathcal{R}\{\hat{\mathbf{r}}^H\mu(\hat{\mathbf{r}})\}$$
$$+ \frac{1}{B}\sum_{b=1}^{B}\frac{D_0}{\alpha}\left(\frac{\partial[\mu(\hat{r}_b)]_{\mathcal{R}}}{\partial[\hat{r}_b]_{\mathcal{R}}} + \frac{\partial[\mu(\hat{r}_b)]_{\mathcal{I}}}{\partial[\hat{r}_b]_{\mathcal{I}}}\right). \tag{13}$$

[2]This estimator is equivalent to $\hat{\mathbf{h}}' = \eta(\gamma'\hat{\mathbf{r}}, \tau')$ for $\gamma = \gamma'$ and $\tau = \frac{\tau'}{\gamma}$.

Algorithm 1 SAND: Sparsity-Adaptive oNe-bit Denoiser

1: **input** $\mathbf{r}$, α and D_0
2: $\hat{\mathbf{r}} = \text{FFT}(\mathbf{r})$, $SURE_{\min} = \infty$, $\tau = 0$
3: $\hat{\mathbf{r}}^s = \text{sort}\{|\hat{\mathbf{r}}|, \text{'ascend'}\}$, $\hat{r}_{B+1}^s = \hat{r}_{B+2}^s = \infty$
4: $a = \sum_{k=1}^{B}(\hat{r}_k^s)^2$, $b = \sum_{k=1}^{B}\hat{r}_k^s$, $c = \sum_{k=1}^{B}(\hat{r}_k^s)^{-1}$
5: **for** $k = 0, \ldots, B+1$ **do**
6: $\quad \gamma = \frac{2(a-\tau b) - D_0(2(B-k)-\tau c)}{2\alpha(a - 2\tau b + \tau^2(B-k))}$
7: $\quad SURE = \frac{1}{B}\gamma^2\left(a - 2\tau b + \tau^2(B-k)\right) + \frac{2-D_0}{\alpha^2} - \frac{1}{B}\frac{\gamma}{\alpha}\left(2(a-\tau b) - D_0(2(B-k) - \tau c)\right)$
8: $\quad$ **if** $SURE < SURE_{\min}$ **and** $\tau < \sqrt{2D_0\log(B)}$ **then**
9: $\quad\quad SURE_{\min} = SURE$, $\tau^\star = \tau$, $\gamma^\star = \gamma$
10: $\quad$ **end if**
11: $\quad \tau = \hat{r}_{k+1}^s$, $a = a - (\hat{r}_{k+1}^s)^2$, $b = b - \hat{r}_{k+1}^s$, $c = c - (\hat{r}_{k+1}^s)^{-1}$
12: **end for**
13: $\hat{h}_k^\star = \gamma^\star \frac{\hat{r}_k}{|\hat{r}_k|}\max\{|\hat{r}_k| - \tau^\star, 0\}$, $k = 1, \ldots, B$
14: **return** $\hat{\mathbf{h}}^\star = \text{IFFT}(\hat{\mathbf{h}}^\star)$

Refer to Appendix B for the proof. Since SURE is an unbiased estimator of the MSE, we use SURE in (13) with $\mu(\hat{\mathbf{r}}) = \gamma\eta(\hat{\mathbf{r}}, \tau)$, in order to find the optimal parameters $\gamma^\star$ and $\tau^\star$. While a naïve approach could perform a two-dimensional grid search over the tuple (γ, τ), we next show that we can efficiently find $\gamma^\star$ and $\tau^\star$ with $\mathcal{O}(B\log(B))$ complexity.

Let $\hat{\mathbf{r}}^s$ be a vector containing the absolute values of $\hat{\mathbf{r}}$ sorted in ascending order. For a given τ, let k be the number of entries in $\hat{\mathbf{r}}^s$ that are smaller than τ. For the denoiser $\mu(\hat{\mathbf{r}}) = \gamma\eta(\hat{\mathbf{r}}, \tau)$, following the derivations in [16, App. B], SURE in (13) is

$$SURE = \frac{1}{B}\gamma^2\sum_{b=k+1}^{B}(\hat{r}_b^s - \tau)^2 + \frac{2-D_0}{\alpha^2} \tag{14}$$
$$- \frac{1}{B}\frac{\gamma}{\alpha}\sum_{b=k+1}^{B}\left(2\hat{r}_b^s(\hat{r}_b^s - \tau) - D_0\left(2 - \frac{\tau}{\hat{r}_b^s}\right)\right).$$

By defining the quantities $a = \sum_{b=k+1}^{B}(\hat{r}_b^s)^2$, $b = \sum_{b=k+1}^{B}\hat{r}_b^s$ and $c = \sum_{b=k+1}^{B}(\hat{r}_b^s)^{-1}$, we can rewrite (14) as

$$SURE = \frac{1}{B}\gamma^2\left(a - 2\tau b + \tau^2(B-k)\right) + \frac{2-D_0}{\alpha^2}$$
$$- \frac{1}{B}\frac{\gamma}{\alpha}\left(2(a-\tau b) - D_0(2(B-k) - \tau c)\right). \tag{15}$$

For a fixed τ, the optimal $\gamma^\star$ that minimizes (15) is given by

$$\gamma^\star = \frac{2(a-\tau b) - D_0(2(B-k) - \tau c)}{2\alpha(a - 2\tau b + \tau^2(B-k))}. \tag{16}$$

The optimal threshold $\tau^\star$ could take any value between 0 and $\hat{r}_B^s$. However, as in the derivation of BEACHES [28], we restrict the search to values in $\hat{\mathbf{r}}^s$, as it significantly reduces the complexity, without sacrificing performance. We also set an upper limit for τ of $\sqrt{2D_0\log(B)}$, which ensures (with high probability) that the threshold is lower than the largest noise realization [27]. For each $\tau = \hat{r}_k^s$, $k = 0, \ldots, B$ (with $\hat{r}_0^s = 0$), and for its associated $\gamma^\star$ given by (16), we evaluate SURE in (15), and then pick $\gamma^\star$ and $\tau^\star$ that result in the minimum value of SURE. We call the resulting algorithm Sparsity-Adaptive oNe-bit Denoiser (SAND), which is summarized in Algorithm 1. Since the complexity of a fast Fourier transform (FFT) and sorting scale with $\mathcal{O}(B\log(B))$, and the operations in each iteration (lines 6 to 11) have complexity $\mathcal{O}(1)$, the overall complexity of SAND scales with $\mathcal{O}(B\log(B))$.
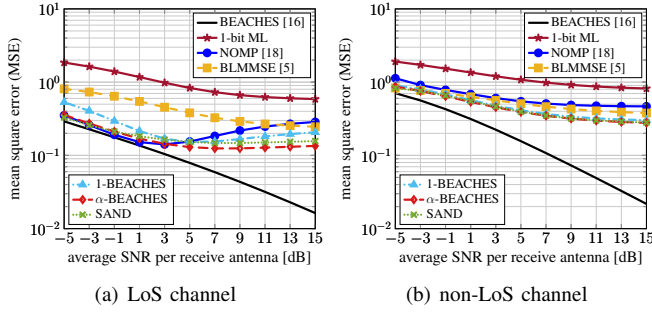
Fig. 2. Mean square error (MSE) of the considered channel denoising methods for mmWave LoS and non-LoS channels. The proposed sparsity-adaptive denoising methods significantly outperform naïve 1-bit ML channel estimation.

V. RESULTS

We now demonstrate the efficacy of 1-BEACHES, α-BEACHES, and SAND. As reference methods, we consider perfect channel state information (CSI), where $\mathbf{h}^\star = \mathbf{h}$, BEACHES [16], which denoises the infinite-resolution (unquantized) measurements $\mathbf{y} = \mathbf{h} + \mathbf{n}$, and 1-bit maximum-likelihood (ML) channel estimation, where $\mathbf{h}^\star = \mathbf{r}$ is the 1-bit observation in (2). In addition, we compare the performance to state-of-the-art denoising methods, including (i) Newtonized orthogonal matching pursuit (NOMP) [18] with an equivalent noise variance Q_0 and a false alarm rate $P_{\mathrm{fa}} = 0.5$ (using E_0 results in poor performance; P_{fa} has been tuned to achieve low MSE at low and high SNR) and (ii) the 1-bit Bussgang linear MMSE estimator (BLMMSE) [5], [19], which corresponds to $\mathbf{h}^\star = \frac{E_h}{\sqrt{\pi(E_h + N_0)}}\mathbf{r}$ for the used orthogonal pilots.

A. Simulation Setup

We simulate a mmWave massive MIMO system with $B = 256$ BS antennas and $U = 16$ single-antenna UEs. We generate LoS and non-LoS channel matrices using the QuaDRiGa mmMAGIC UMi model [26] at a carrier frequency of $60\,\mathrm{GHz}$ for a BS with $\lambda/2$-spaced antennas arranged in a ULA. The UEs are placed randomly in a $120°$ circular sector around the BS between a distance of $10\,\mathrm{m}$ and $110\,\mathrm{m}$, and the UEs are separated by at least $4°$. We model UE-side power control to ensure that the highest receive power is at most $6\,\mathrm{dB}$ higher than that of the weakest UE.

B. Mean-Square Error (MSE) Performance

Fig. 2 shows the channel estimation MSE of the proposed 1-bit denoising algorithms and the considered baseline methods. We observe that the three proposed methods, 1-BEACHES, α-BEACHES, and SAND significantly outperform 1-bit ML channel estimation. Furthermore, we see that α-BEACHES and SAND have a slight advantage over 1-BEACHES in LoS scenarios. Surprisingly, SAND has a slightly higher MSE than α-BEACHES, which we attribute to the fact that SAND has to learn two parameters, whereas α-BEACHES only learns the optimal denoising threshold. For that reason, SAND is more sensitive to the assumptions made in footnote 1. NOMP and BLMMSE also outperform 1-bit ML, but their MSE is higher than that of our algorithms, especially at high SNR.

C. Bit Error Rate (BER) Performance

To assess the impact of the proposed 1-bit denoising algorithms on the uncoded BER performance during the data detection phase, we use the 1-bit Bussgang linear MMSE equalizer proposed in [29], which operates on the 1-bit quantized received data using the channel estimates provided by our denoising methods and the considered baseline algorithms. We consider QPSK and 16-QAM transmission.

Fig. 3 shows that the proposed sparsity-adaptive denoising algorithms significantly outperform naïve 1-bit ML channel estimation. We furthermore see that for QPSK, all three methods, 1-BEACHES, α-BEACHES, and SAND, perform equally well under both LoS and non-LoS scenarios. For 16-QAM, where it is important to get an accurate estimate of the channel gain, α-BEACHES and SAND outperform 1-BEACHES and NOMP, which directly operate with the received 1-bit measurements. Hence, correcting the scale of the received data is critical for higher-order constellation sets. While BLMMSE adjusts for the scale, it is unable to exploit sparsity which results in rather poor BER performance. For non-LoS channels, NOMP performs inferior to the proposed methods. In addition, NOMP requires high complexity [28]. Since the propagation conditions (such as the number of propagation paths L) are typically unknown in practice, SAND and α-BEACHES are the preferred denoising methods.

VI. CONCLUSIONS

We have presented three sparsity-adaptive channel vector denoising algorithms for 1-bit mmWave massive MIMO systems. Two of our algorithms denoise 1-bit measurements of the channel estimates using BEACHES [16] in order to automatically adapt the denoising parameter to the instantaneous channel realization. While 1-BEACHES applies BEACHES to the 1-bit measurements using the effective noise variance (which also includes the quantization noise variance), α-BEACHES uses a Bussgang-like scaling factor [24], which results in superior performance. We have also introduced SAND (short for Sparsity-Adaptive oNe-bit Denoiser), a novel denoising algorithm with $\mathcal{O}(B \log(B))$ complexity, which jointly optimizes the thresholding parameter and the scaling factor in a nonparametric fashion. Our simulations have shown that α-BEACHES and SAND perform equally well under the considered LoS and non-LoS mmWave channels and outperform 1-BEACHES as well as other considered baseline methods in the case of 16-QAM transmission.

APPENDIX A: DERIVATION OF $\frac{1}{B}\mathbb{E}\big[\mathbf{h}^H\mathbf{r}\big]$

Since $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, N_0\mathbf{I}_B)$ and $\mathbf{h}$ is assumed circularly symmetric, the imaginary part of $\mathbb{E}\big[\mathbf{h}^H\mathbf{r}\big]$ is zero, and

$$\frac{1}{B}\mathbb{E}\big[\mathbf{h}^H\mathbf{r}\big] = \frac{1}{B}\mathbb{E}[[\mathbf{h}]_\mathcal{R}[\mathbf{r}]_\mathcal{R}] + \frac{1}{B}\mathbb{E}[[\mathbf{h}]_\mathcal{I}[\mathbf{r}]_\mathcal{I}]. \quad (17)$$

By assuming $\mathbf{h} \sim \mathcal{CN}(\mathbf{0}, E_h\mathbf{I}_B)$, $\frac{1}{B}\mathbb{E}[[\mathbf{h}]_\mathcal{R}[\mathbf{r}]_\mathcal{R}]$ becomes

$$\frac{1}{B}\sum_{b=1}^{B}\mathbb{E}\left[\int_{-\infty}^{-[n_b]_\mathcal{R}} \frac{-[h_b]_\mathcal{R}}{\sqrt{\pi E_h}}e^{-\frac{([h_b]_\mathcal{R})^2}{E_h}}d[h_b]_\mathcal{R}\right] \quad (18)$$

$$+\frac{1}{B}\sum_{b=1}^{B}\mathbb{E}\left[\int_{-[n_b]_\mathcal{R}}^{-\infty} \frac{[h_b]_\mathcal{R}}{\sqrt{\pi E_h}}e^{-\frac{([h_b]_\mathcal{R})^2}{E_h}}d[h_b]_\mathcal{R}\right] = \frac{E_h}{\sqrt{\pi(E_h+N_0)}}.$$
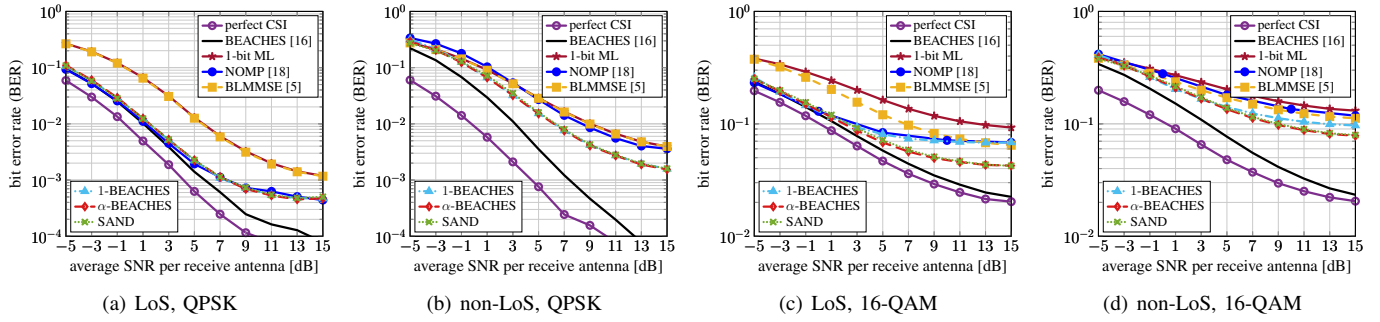
Fig. 3. Uncoded bit error rate (BER) of 1-bit channel estimation and 1-bit data detection in mmWave LoS and non-LoS channels. We see that α-BEACHES and SAND outperform 1-BEACHES and 1-bit ML for LoS and non-LoS channel conditions for 16-QAM transmission.

Following the same procedure for the imaginary part, we get

$$\frac{1}{B}\mathbb{E}\left[\mathbf{h}^{\mathrm{H}}\mathbf{r}\right] = \frac{2E_h}{\sqrt{\pi(E_h+N_0)}}. \qquad (19)$$

Appendix B: Derivation of SURE in (13)

For deriving SURE as in (13), we follow the procedure in [16, App. A], with the following modifications: Instead of $\hat{\mathbf{y}} = \hat{\mathbf{h}} + \hat{\mathbf{e}}$, we use $\hat{\mathbf{r}} = \alpha\hat{\mathbf{h}} + \hat{\mathbf{d}}$. In other words, where [16] uses $\hat{\mathbf{y}} \sim \mathcal{CN}(\hat{\mathbf{h}}, E_0\mathbf{I}_B)$, we replace it by $\hat{\mathbf{r}} \sim \mathcal{CN}(\alpha\hat{\mathbf{h}}, D_0\mathbf{I}_B)$. Instead of $g(\hat{\mathbf{y}}) = \mu(\hat{\mathbf{y}}) - \hat{\mathbf{y}}$, we use $g(\hat{\mathbf{r}}) = \mu(\hat{\mathbf{r}}) - \hat{\mathbf{r}}/\alpha$.

References

[1] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[2] T. S. Rappaport, R. W. Heath Jr., R. C. Daniels, and J. N. Murdock, *Millimeter Wave Wireless Communications*. Prentice Hall, 2015.

[3] S. Dutta, C. N. Barati, A. Dhananjay, D. A. Ramirez, J. F. Buckwalter, and S. Rangan, "A case for digital beamforming at mmWave," *arXiv preprint arXiv:1901.08693*, Jan. 2019.

[4] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, "Throughput analysis of massive MIMO uplink with low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 4038–4051, Jun. 2017.

[5] Y. Li, C. Tao, G. Seco-Granados, A. Mezghani, A. L. Swindlehurst, and L. Liu, "Channel estimation and performance analysis of one-bit massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 4075–4089, Aug. 2017.

[6] J. Mo, P. Schniter, and R. W. Heath Jr., "Channel estimation in broadband millimeter wave MIMO systems with few-bit ADCs," *IEEE Trans. Signal Process.*, vol. 66, no. 5, pp. 1141–1154, Mar. 2016.

[7] P. Skrimponis, S. Dutta, M. Mezzavilla, S. Rangan, S. H. Mirfarshbafan, C. Studer, J. Buckwalter, and M. Rodwell, "Power consumption analysis for mobile mmwave and sub-THz receivers," in *2020 IEEE 6G Wireless Summit*, Mar. 2020, pp. 1–5.

[8] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3029–3056, Sep. 2015.

[9] Z. Gao, C. Hu, L. Dai, and Z. Wang, "Channel estimation for millimeter-wave massive MIMO with hybrid precoding over frequency-selective fading channels," *IEEE Commun. Lett.*, vol. 20, no. 6, pp. 1259–1262, Jun. 2016.

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

[11] 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, May 2013.

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

[13] J. Mo, P. Schniter, N. González Prelcic, and R. W. Heath Jr., "Channel estimation in millimeter wave MIMO systems with one-bit quantization," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Nov. 2014, pp. 957–961.

[14] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7465–7490, Nov. 2013.

[15] J. Brady, N. Behdad, and A. M. Sayeed, "Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis, and measurements," *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3814–3827, Jul. 2013.

[16] R. Ghods, A. Gallyas-Sanhueza, S. H. Mirfarshbafan, and C. Studer, "BEACHES: Beamspace channel estimation for multi-antenna mmWave systems and beyond," in *Proc. IEEE Int. Workshop Signal Process. Advances Wireless Commun. (SPAWC)*, Jul. 2019, pp. 1–5.

[17] B. N. Bhaskar, G. Tang, and B. Recht, "Atomic norm denoising with applications to line spectral estimation," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 5987–5999, Dec. 2013.

[18] B. Mamandipoor, D. Ramasamy, and U. Madhow, "Newtonized orthogonal matching pursuit: Frequency estimation over the continuum," *IEEE Trans. Signal Process.*, vol. 64, no. 19, pp. 5066–5081, Oct. 2016.

[19] Y. Li, C. Tao, L. Liu, G. Seco-Granados, and A. L. Swindlehurst, "Channel estimation and uplink achievable rates in one-bit massive MIMO systems," in *IEEE Sensor Array and Multichannel Signal Process. Workshop (SAM)*, Rio de Janeiro, Brazil, Jul. 2016.

[20] C. Mollén, J. Choi, E. G. Larsson, and R. W. Heath Jr., "Uplink performance of wideband massive MIMO with one-bit ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 87–100, Jan. 2017.

[21] C. Studer and G. Durisi, "Quantized massive MU-MIMO-OFDM uplink," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2387–2399, Jun. 2016.

[22] X. Huang and B. Liao, "One-bit MUSIC," *IEEE Signal Process. Lett.*, vol. 26, no. 7, pp. 961–965, Jul. 2019.

[23] A. Kaushik, E. Vlachos, J. Thompson, and A. Perelli, "Efficient channel estimation in millimeter wave hybrid MIMO systems with low resolution ADCs," in *Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2018, pp. 1825–1829.

[24] J. J. Bussgang, "Crosscorrelation functions of amplitude-distorted Gaussian signals," Res. Lab. Elec., Cambridge, MA, USA, Tech. Rep. 216, Mar. 1952.

[25] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge Univ. Press, 2005.

[26] S. Jaeckel, L. Raschkowski, K. Börner, L. Thiele, F. Burkhardt, and E. Eberlein, "QuaDRiGa - quasi deterministic radio channel generator, user manual and documentation," Fraunhofer Heinrich Hertz Institute, Tech. Rep. v2.2.0, Jun. 2019.

[27] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *Journal of the American Statistical Association*, vol. 90, no. 432, pp. 1200–1224, Dec. 1995.

[28] S. H. Mirfarshbafan, A. Gallyas-Sanhueza, R. Ghods, and C. Studer, "Beamspace channel estimation for massive MIMO mmWave systems: Algorithm and VLSI design," Oct. 2019. [Online]. Available: https://arxiv.org/abs/1910.00756

[29] L. V. Nguyen and D. H. Nguyen, "Linear receivers for massive MIMO systems with one-bit ADCs," Jul. 2019. [Online]. Available: https://arxiv.org/abs/1907.06664