

Channel Estimation in Low-Resolution Near-Field Massive MIMO Systems

Ly V. Nguyen^{*}, Duy H. N. Nguyen[†], Italo Atzeni[‡], Antti Tölli[‡], and A. Lee Swindlehurst^{*}

^{*}Center for Pervasive Communications and Computing, University of California, Irvine, CA, USA

[†]Department of Electrical and Computer Engineering, San Diego State University, CA, USA

[‡]Centre for Wireless Communications, University of Oulu, Finland

E-mail: vanln1@uci.edu, duy.nguyen@sdsu.edu, italo.atzeni@oulu.fi, antti.tolli@oulu.fi, swindle@uci.edu

Abstract—Massive multiple-input-multiple-output (MIMO) is a core technology of current and future wireless networks. However, the very large dimension of a massive antenna array can lead to radical changes in the electromagnetic fields near the array, and the classical far-field channel model is no longer accurate. Instead, the channel should be modeled under the assumption of near-field spherical wavefronts. Furthermore, the very large dimension of the arrays can also result in high power consumption and hardware complexity. A practical solution for this problem is to use low-resolution analog-to-digital converters (ADCs). It is therefore of significance to study the near-field channel estimation problem for MIMO systems implemented with low-resolution ADCs. We propose an efficient on-grid polar-domain channel estimation method which relies on the polar-domain sparsity of the near-field channels. We first reformulate the sparse low-resolution near-field maximum-likelihood channel estimation problem by exploiting an approximation of the cumulative distribution function of a normal random variable as a logistic activation function. We then develop an on-grid polar-domain channel estimation method based on the gradient descent approach and the polar-domain sparsity of the near-field channel. Finally, we apply the deep unfolding technique to optimize the performance of the proposed method and illustrate its efficiency via several simulation studies.

I. INTRODUCTION

Communications at millimeter wave (mmWave) and sub-terahertz (sub-THz) frequencies are envisioned as a one of the enablers of future (e.g., 6G) wireless networks [1]. The abundant bandwidths in these high-frequency bands are essential for meeting the very high data rate demands of the increasing number of users and devices. However, the severe path loss of mmWave and sub-THz channels requires the deployment of truly massive antenna arrays, which fortunately is possible thanks to the very small size of the high-frequency antennas. In addition, users in mmWave and sub-THz systems are often required to be located near a base station (BS), which in conjunction with the very large size of the BS antenna array leads to radical changes in the electromagnetic field structure. In particular, the far-field planar wavefront model, where the channel depends (approximately) only on the angles of departure/arrival (AoDs/AoAs), is no longer accurate in

this case. Instead, the channel should be modeled under the assumption of near-field spherical wavefronts [2], which is a function not only of the AoDs/AoAs but also of the distances between the BS, users, and scatterers.

Research on near-field massive multiple-input-multiple-output (MIMO) systems has recently gained significant interest, e.g., [3]–[13]. The work in [4] showed that near-field channels may not be sparse in the angular domain even with relatively few signal paths and developed a polar-domain transform taking into account both the angular and range information to give a sparse polar-domain representation. The work in [5] proposed a sparse distance-parameterized angular-domain model in which the dictionary size depends only on the size of the angular grid. Exploiting the extra degrees of freedom in the distance domain, [6] proposed a location division multiple access (LDMA) approach to enhance spectrum efficiency, since LDMA can serve different users located at the same angle but at different distances. The work in [7] proposed to use circular instead of linear arrays to enable more users to benefit from near-field communications. Channel estimation for hybrid near-/far-field and mixed line-of-sight (LoS) and non-line-of-sight (NLoS) channels was studied in [8] and [9], respectively. Different beam training approaches were also proposed in [10]–[13]. Deep learning was exploited in [10], while [11] employed a conventional far-field codebook to determine the candidate angles of the users before using a customized polar-domain codebook to find their range. A near-field hierarchical beam training scheme was developed in [12] to reduce the overhead without the need for extra hardware, whereas [13] focused on wideband beam training taking into account the beam split effect over different frequencies.

All of the aforementioned works assume perfect (infinite-resolution) quantization at the BS, which requires many high-power and high-resolution analog-to-digital converters (ADCs) due to the large dimension of the massive antenna array. Although a natural approach to address this problem is to use low-resolution ADCs, there is little work in the literature on near-field channel estimation with such hardware. Recently, [14] solved the grant-free joint activity detection and channel estimation problem over near-field channels assuming mixed-resolution ADCs, where only some of the receive antennas employ low-resolution ADCs and the rest are equipped with conventional ADCs. To the best of our knowledge, this

This work was supported by the National Science Foundation under grants CCF-2225575, ECCS-2146436, CCF-2225576, and CCF-2322190, by the Research Council of Finland under grants 336449 Profi6, 346208 6G Flagship, 343586 CAMAIDE, 348396 HIGH-6G, and 357504 EETCAMD, and by the European Commission under grant 101095759 Hexa-X-II.

paper is the first attempt to address the near-field channel estimation problem in massive MIMO systems with low-resolution ADCs. In this regard, we first formulate the sparse low-resolution near-field maximum-likelihood channel estimation problem by exploiting an approximation of the cumulative distribution function (cdf) of a normal random variable as a logistic activation function. We then develop an on-grid polar-domain channel estimation method based on the gradient descent approach and the polar-domain sparsity of the near-field channel. Finally, we apply the deep unfolding technique to optimize the performance of the proposed method and illustrate its efficiency via several simulation studies.

II. SYSTEM MODEL

We consider a massive MIMO system with an N -antenna BS serving K single-antenna users. Assuming a uniform linear array (ULA) structure, the near-field channel model for user k is given by

$$\mathbf{h}_k = \sqrt{\frac{N}{L}} \sum_{\ell=1}^L \alpha_{k,\ell} \mathbf{b}(\omega_{k,\ell}, r_{k,\ell}) \in \mathbb{C}^N, \quad (1)$$

where L is the number of channel paths, $\alpha_{k,\ell}$ is the channel path gain, and $\mathbf{b}(\omega_{k,\ell}, r_{k,\ell}) \in \mathbb{C}^N$ is the steering vector given by

$$\mathbf{b}(\omega_{k,\ell}, r_{k,\ell}) = \frac{1}{\sqrt{N}} [e^{j\frac{2\pi}{\lambda}(d_{k,\ell}^{(1)} - r_{k,\ell})}, \dots, e^{j\frac{2\pi}{\lambda}(d_{k,\ell}^{(N)} - r_{k,\ell})}]^H. \quad (2)$$

Let the coordinates of scatterer ℓ for user k and the n -th receive antenna be $(r_{k,\ell} \cos \theta_{k,\ell}, r_{k,\ell} \sin \theta_{k,\ell})$ and $(0, \delta_n)$, respectively, with $\delta_n = \frac{2n-N-1}{2} d_A$, $n = 1, \dots, N$, and where d_A is the antenna spacing. The distance between scatterer ℓ for user k and the n -th receive antenna is then given by $d_{k,\ell}^{(n)} = \sqrt{r_{k,\ell}^2 + \delta_n^2 - 2\delta_n r_{k,\ell} \sin \theta_{k,\ell}}$, with $\omega_{k,\ell} = \sin \theta_{k,\ell}$.

Let $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_K] \in \mathbb{C}^{N \times K}$ denote the channel matrix and $\mathbf{X} \in \mathbb{C}^{K \times T_p}$ be the pilot sequence, where T_p is the pilot length. The unquantized received signal is $\mathbf{Z} = \mathbf{H}\mathbf{X} + \mathbf{N} \in \mathbb{C}^{N \times T_p}$, where the elements of $\mathbf{N} = [\mathbf{n}_1, \dots, \mathbf{n}_{T_p}] \in \mathbb{C}^{N \times T_p}$ are assumed to be independent and identically distributed (i.i.d.) zero-mean Gaussian noise with distribution $\mathcal{CN}(0, N_0)$. Then, the quantized received signal is given by

$$\mathbf{Y} = \mathcal{Q}_b(\Re\{\mathbf{Z}\}) + j\mathcal{Q}_b(\Im\{\mathbf{Z}\}) \in \mathbb{C}^{N \times T_p}, \quad (3)$$

where $\mathcal{Q}_b(\cdot)$ represents the b -bit ADC operation, which is applied separately to each element of its matrix or vector argument. We assume that $\mathcal{Q}_b(\cdot)$ employs uniform scalar quantization, which is characterized by a set of $2^b - 1$ thresholds denoted by $\{\tau_1, \dots, \tau_{2^b-1}\}$. Without loss of generality, we set $-\infty = \tau_0 < \tau_1 < \dots < \tau_{2^b-1} < \tau_{2^b} = \infty$. For a quantization step Δ , the quantization thresholds are given by

$$\tau_\ell = (-2^{b-1} + \ell)\Delta, \quad \text{for } \ell \in \mathcal{B} = \{1, \dots, 2^b - 1\}. \quad (4)$$

The quantized output y is then defined as

$$y = \mathcal{Q}_b(z) = \begin{cases} \tau_\ell - \frac{\Delta}{2}, & \text{if } z \in (\tau_{\ell-1}, \tau_\ell] \text{ with } \ell \in \mathcal{B} \\ (2^b - 1)\frac{\Delta}{2}, & \text{if } z \in (\tau_{2^b-1}, \tau_{2^b}] \end{cases} \quad (5)$$

III. PROPOSED ON-GRID POLAR CHANNEL ESTIMATION

First, we rewrite the unquantized received signal as

$$\mathbf{Z} = \mathbf{W}\bar{\mathbf{H}}\mathbf{X} + \mathbf{N}, \quad (6)$$

where $\mathbf{W} \in \mathbb{C}^{N \times S_{\text{pol}}}$ is a polar-domain transformation matrix [4] that makes $\bar{\mathbf{H}}$ a sparse polar-domain channel matrix. Here, S_{pol} denotes the number of sampled near-field steering vectors in the polar domain and $\bar{\mathbf{H}} = [\bar{\mathbf{h}}_1, \dots, \bar{\mathbf{h}}_K] \in \mathbb{C}^{S_{\text{pol}} \times K}$ is a KL -sparse matrix, where $\bar{\mathbf{h}}_k$ is the L -sparse channel vector of user k , i.e., each vector $\bar{\mathbf{h}}_k$ has L non-zero elements (with $L \ll S_{\text{pol}}$). Since $\mathbf{H} = \mathbf{W}\bar{\mathbf{H}}$, we can obtain an estimate of $\mathbf{H}$ through an estimate of the sparse channel matrix $\bar{\mathbf{H}}$.

Now, we vectorize the unquantized received signal $\mathbf{Z}$ in (6) as

$$\mathbf{z} = \text{vec}(\mathbf{Z}) = (\mathbf{X}^T \otimes \mathbf{W})\bar{\mathbf{h}} + \mathbf{n} \quad (7)$$

$$= \mathbf{A}\bar{\mathbf{h}} + \mathbf{n} \in \mathbb{C}^{NT_p}, \quad (8)$$

with $\mathbf{A} = \mathbf{X}^T \otimes \mathbf{W} \in \mathbb{C}^{T_p N \times K S_{\text{pol}}}$, $\bar{\mathbf{h}} = \text{vec}(\bar{\mathbf{H}}) \in \mathbb{C}^{S_{\text{pol}} K}$, and $\mathbf{n} = \text{vec}(\mathbf{N}) \in \mathbb{C}^{NT_p}$. The problem of interest is to estimate the sparse channel vector $\bar{\mathbf{h}}$ using $\mathbf{W}$, $\mathbf{X}$, and the quantized received signal vector $\mathbf{y} = \mathcal{Q}_b(\Re\{\mathbf{z}\}) + j\mathcal{Q}_b(\Im\{\mathbf{z}\})$. Note that the sparse channel vector $\bar{\mathbf{h}}$ has KL non-zero elements. To facilitate the later derivations, we represent the term $\mathbf{A}\bar{\mathbf{h}}$ in (8) in the real domain as

$$\begin{bmatrix} \mathbf{z}^{\Re} \\ \mathbf{z}^{\Im} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{A}^{\Re} & -\mathbf{A}^{\Im} \\ \mathbf{A}^{\Im} & \mathbf{A}^{\Re} \end{bmatrix}}_{\mathbf{A}^{\Re}} \underbrace{\begin{bmatrix} \bar{\mathbf{h}}^{\Re} \\ \bar{\mathbf{h}}^{\Im} \end{bmatrix}}_{\mathbf{h}^{\Re}}, \quad (9)$$

with

$$\mathbf{A}^{\Re} = \Re\{\mathbf{X}^T\} \otimes \Re\{\mathbf{W}\} - \Im\{\mathbf{X}^T\} \otimes \Im\{\mathbf{W}\}, \quad (10)$$

$$\mathbf{A}^{\Im} = \Re\{\mathbf{X}^T\} \otimes \Im\{\mathbf{W}\} + \Im\{\mathbf{X}^T\} \otimes \Re\{\mathbf{W}\}. \quad (11)$$

Thus, we have

$$\begin{bmatrix} \mathbf{y}^{\Re} \\ \mathbf{y}^{\Im} \end{bmatrix} = \mathcal{Q}_b \left(\begin{bmatrix} \mathbf{z}^{\Re} \\ \mathbf{z}^{\Im} \end{bmatrix} + \begin{bmatrix} \mathbf{n}^{\Re} \\ \mathbf{n}^{\Im} \end{bmatrix} \right). \quad (12)$$

The likelihood function of $\bar{\mathbf{h}}$ is given by

$$p(\mathbf{y}|\bar{\mathbf{h}}) = \sum_{m=1}^{NT_p} \left[\log(\Phi(s_m^{\Re, \text{up}}) - \Phi(s_m^{\Re, \text{low}})) + \log(\Phi(s_m^{\Im, \text{up}}) - \Phi(s_m^{\Im, \text{low}})) \right], \quad (13)$$

where $\Phi(\cdot)$ is the cdf of a normal random variable and

$$s_m^{\Re, \text{up}} = \sqrt{2\rho}(y_m^{\Re, \text{up}} - z_m^{\Re}), \quad s_m^{\Re, \text{low}} = \sqrt{2\rho}(y_m^{\Re, \text{low}} - z_m^{\Re}), \quad (14)$$

$$s_m^{\Im, \text{up}} = \sqrt{2\rho}(y_m^{\Im, \text{up}} - z_m^{\Im}), \quad s_m^{\Im, \text{low}} = \sqrt{2\rho}(y_m^{\Im, \text{low}} - z_m^{\Im}), \quad (15)$$

with $\rho = 1/N_0$. Here, $y_m^{\Re, \text{up}}$ and $y_m^{\Re, \text{low}}$ (or $y_m^{\Im, \text{up}}$ and $y_m^{\Im, \text{low}}$) are the lower and upper thresholds of the quantization bin to which $\Re\{y_m\}$ (or $\Im\{y_m\}$) belongs, respectively.

We are interested in solving the following sparsity-constrained ML problem:

$$\begin{aligned} & \underset{\{\bar{\mathbf{h}}\}}{\text{maximize}} && p(\mathbf{y}|\bar{\mathbf{h}}) \\ & \text{subject to} && \|\bar{\mathbf{h}}_{\mathcal{I}(k)}\|_0 = L, \quad \forall k = 1, \dots, K, \end{aligned} \quad (16)$$

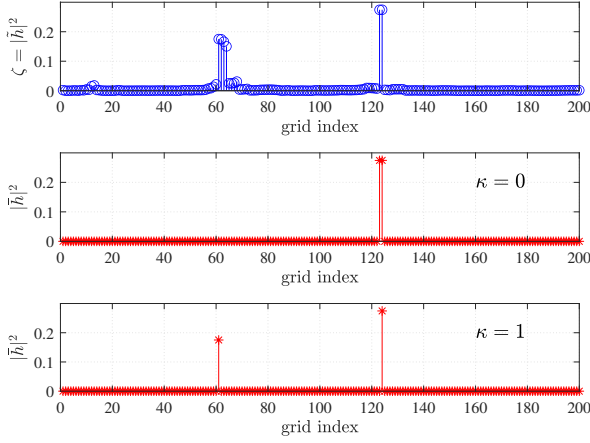


Fig. 1: Illustration of the effect of the parameter κ with $L = 2$.

with $\mathcal{I}(k) = \{(k-1)S_{\text{pol}} + 1, \dots, kM_{\text{pol}}\}$. However, the gradient of $p(\mathbf{y}|\bar{\mathbf{h}})$, given by

$$\nabla p(\mathbf{y}|\bar{\mathbf{h}}) = - \sum_{m=1}^{NT_p} \left[\frac{\sqrt{2\rho} (\mathbf{A}_{m,:}^{\mathbb{R}})^T (\phi(s_m^{\mathbb{R},\text{up}}) - \phi(s_m^{\mathbb{R},\text{low}}))}{\Phi(s_m^{\mathbb{R},\text{up}}) - \Phi(s_m^{\mathbb{R},\text{low}})} + \frac{\sqrt{2\rho} (\mathbf{A}_{2m,:}^{\mathbb{R}})^T (\phi(s_m^{\mathbb{S},\text{up}}) - \phi(s_m^{\mathbb{S},\text{low}}))}{\Phi(s_m^{\mathbb{S},\text{up}}) - \Phi(s_m^{\mathbb{S},\text{low}})} \right], \quad (17)$$

may be undefined (i.e., the denominators in (17) may be zero) since the function $\Phi(\cdot)$ rapidly approaches zero or one [15]. To address this issue, we employ the approximation $\Phi(s) \approx \sigma(cs) = 1/(1 + e^{-cs})$, where $c = 1.702$ (see [16], [17]), and approximate the likelihood function $p(\mathbf{y}|\bar{\mathbf{h}})$ as

$$p(\mathbf{y}|\bar{\mathbf{h}}) \approx \tilde{p}(\mathbf{y}|\bar{\mathbf{h}}) = \sum_{m=1}^{NT_p} [\log(\sigma(s_m^{\mathbb{R},\text{up}}) - \sigma(s_m^{\mathbb{R},\text{low}})) + \log(\sigma(s_m^{\mathbb{S},\text{up}}) - \sigma(s_m^{\mathbb{S},\text{low}}))] . \quad (18)$$

The gradient of (18), given by

$$\nabla_{\bar{\mathbf{h}}^{\mathbb{R}}} \tilde{p}(\mathbf{y}|\bar{\mathbf{h}}) = \gamma (\mathbf{A}^{\mathbb{R}})^T \left[\mathbf{1} - \sigma(\gamma (\mathbf{A}^{\mathbb{R}} \bar{\mathbf{h}}^{\mathbb{R}} - \mathbf{y}^{\mathbb{R},\text{up}})) - \sigma(\gamma (\mathbf{A}^{\mathbb{R}} \bar{\mathbf{h}}^{\mathbb{R}} - \mathbf{y}^{\mathbb{R},\text{low}})) \right], \quad (19)$$

with $\gamma = c\sqrt{2\rho}$, does not suffer from the above divide-by-zero issue. Hence, we obtain the reformulated problem

$$\begin{aligned} & \underset{\{\bar{\mathbf{h}}\}}{\text{maximize}} && \tilde{p}(\mathbf{y}|\bar{\mathbf{h}}) \\ & \text{subject to} && \|\bar{\mathbf{h}}_{\mathcal{I}(k)}\|_0 = L, \quad \forall k = 1, \dots, K. \end{aligned} \quad (20)$$

At this stage, we employ the projected gradient descent approach to solve problem (20) as

$$\tilde{\mathbf{h}}^{\mathbb{R},(i)} = \bar{\mathbf{h}}^{\mathbb{R},(i-1)} + \eta^{(i)} \nabla_{\bar{\mathbf{h}}^{\mathbb{R}}} \tilde{p}(\mathbf{y}|\bar{\mathbf{h}}^{\mathbb{R},(i-1)}), \quad (21)$$

$$\bar{\mathbf{h}}^{\mathbb{R},(i)} = f_{\kappa}(\tilde{\mathbf{h}}^{\mathbb{R},(i)}), \quad (22)$$

Algorithm 1 Function $f_{\kappa}(\cdot)$

Input: $\tilde{\mathbf{h}}^{\mathbb{R}} = [(\tilde{\mathbf{h}}^{\mathbb{R}})^T, (\tilde{\mathbf{h}}^{\mathbb{S}})^T]^T$ and κ

- 1: **Initialization:** $\bar{\mathbf{h}}^{\mathbb{R}} = \mathbf{0}$ and $\bar{\mathbf{h}}^{\mathbb{S}} = \mathbf{0}$
- 2: **for** $k = 1, \dots, K$ **do**
- 3: $\zeta_k = \text{diag}(\tilde{\mathbf{h}}_{\mathcal{I}(k)}^{\mathbb{R}}) \tilde{\mathbf{h}}_{\mathcal{I}(k)}^{\mathbb{R}} + \text{diag}(\tilde{\mathbf{h}}_{\mathcal{I}(k)}^{\mathbb{S}}) \tilde{\mathbf{h}}_{\mathcal{I}(k)}^{\mathbb{S}}$
- 4: $\nu_k = \zeta_k$
- 5: **for** $\ell = 1, \dots, L$ **do**
- 6: $\varrho = f_{\text{one-hot}}(\nu_k)$
- 7: $\bar{\mathbf{h}}_{\mathcal{I}(k)}^{\mathbb{R}} = \bar{\mathbf{h}}_{\mathcal{I}(k)}^{\mathbb{R}} + \text{diag}(\varrho) \tilde{\mathbf{h}}_{\mathcal{I}(k)}^{\mathbb{R}}$
- 8: $\bar{\mathbf{h}}_{\mathcal{I}(k)}^{\mathbb{S}} = \bar{\mathbf{h}}_{\mathcal{I}(k)}^{\mathbb{S}} + \text{diag}(\varrho) \tilde{\mathbf{h}}_{\mathcal{I}(k)}^{\mathbb{S}}$
- 9: $\nu_k = \nu_k - \text{diag}(\varrho_{\kappa}) \zeta_k$
- 10: **end for**
- 11: **end for**
- 12: **return** $\bar{\mathbf{h}}^{\mathbb{R}} = [(\bar{\mathbf{h}}^{\mathbb{R}})^T, (\bar{\mathbf{h}}^{\mathbb{S}})^T]^T$

where i is the iteration index, $\eta^{(i)}$ is the step size, and $f_{\kappa}(\cdot)$ is a projector function to ensure the sparsity constraint for $\bar{\mathbf{h}}$ in (20). Thus, based on the coarse estimate $\tilde{\mathbf{h}}^{\mathbb{R}}$, the function $f_{\kappa}(\cdot)$ will produce a real vector of $2KL$ non-zero elements corresponding to the KL channel paths. A detailed description of the function $f_{\kappa}(\cdot)$ is given in Algorithm 1. The channel paths are found separately for different users, as indicated by the for loop over the user index in line 2. For each user k , we rely on the magnitude of the elements in the coarse estimate $\tilde{\mathbf{h}}_{\mathcal{I}(k)}$ (line 3) to find the L channel paths of user k . In particular, we sequentially find the most dominant path by means of the one-hot vector function $f_{\text{one-hot}}(\cdot)$ (line 6) and then remove it from the magnitude vector ν_k (line 9). The function $f_{\text{one-hot}}(\nu)$ produces a one-hot vector ϱ whose non-zero element is 1 at the index of ν associated with the largest value. The vector ϱ_{κ} in line 9 is constructed from ϱ by setting κ samples to the left and right of the non-zero element in ν to be 1. This is to address the coherence issue illustrated in the top plot in Fig. 1, where several consecutive grids have a non-zero amplitude but the channel path is in only one of these grids. It can be seen from Fig. 1 that, when $\kappa = 0$, the L channel paths are determined by the L largest samples in the magnitude vector ν_k . However, this can lead to incorrect path estimates due to the coherence issue, as can be observed in the middle plot of Fig. 1. This problem can be efficiently addressed by setting $\kappa = 1$, since not only the largest peak but also its adjacent peak are removed from the magnitude vector ν_k before searching for the second channel path.

Furthermore, we propose to use the deep unfolding technique [18] to unfold each iteration of the method in (21) and (22) as a layer of a deep neural network. We treat the scaling parameter γ inside the logistic activation function and the step size η as the trainable parameters of the proposed deep neural network. Let $\hat{\mathbf{h}}$ denote the sparse channel estimate, which is set to be the output of the last layer of the deep neural network, i.e., $\hat{\mathbf{h}} = \bar{\mathbf{h}}^{(I)}$, where I is the number of network layers. The cost function to be minimized is $\|\hat{\mathbf{h}} - \mathbf{h}\|^2$, with $\hat{\mathbf{h}} = (\mathbf{I}_K \otimes \mathbf{W}) \hat{\mathbf{h}}$. We can obtain a training sample consisting

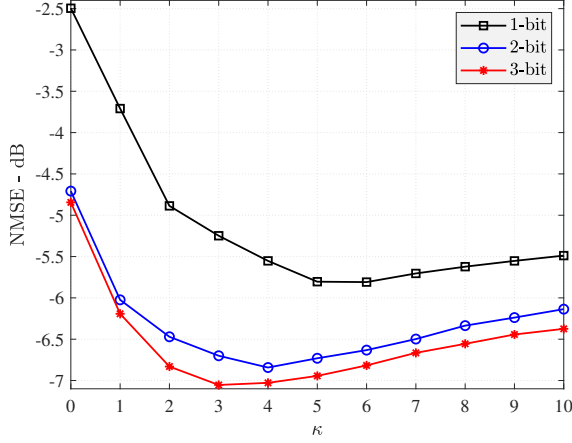


Fig. 2: NMSE performance versus parameter κ with $\text{SNR} = 20$ dB.

of the pilot matrix $\mathbf{X}$, a channel vector realization $\mathbf{h}$, and a noise vector $\mathbf{n}$, which can be randomly generated. To address the gradient backpropagation of the one-hot vector function $f_{\text{one-hot}}(\nu)$ in the training process, we replace it by a softmax function $f_{\text{softmax}}(c_2\nu)$, where c_2 is a parameter that should be set to a large value to match the one-hot vector function.

IV. NUMERICAL RESULTS

In this section, we present numerical results to show the efficiency of the proposed on-grid near-field channel estimation method. In our simulations, we consider $N = 200$ antennas at the BS serving $K = 2$ users. We set the pilot length to $T_p = 2K$, the dictionary size to $S_{\text{pol}} = 1000$, the number of channel paths to $L = 4$, and the number of iterations (network layers) to $I = 10$. We assume a carrier frequency $f_c = 60$ GHz, antenna spacing $d_A = \lambda_c/2$, angles $\theta_{k,\ell} \in [-\pi/3, \pi/3]$, and ranges $r_{k,\ell} \in [10, 50]$ m. The channel estimation performance is measured in terms of normalized mean squared error (NMSE), defined as $\text{NMSE} = \mathbb{E}[\|\hat{\mathbf{h}} - \mathbf{h}\|^2] / \mathbb{E}[\|\mathbf{h}\|^2]$, where $\hat{\mathbf{h}}$ is an estimate of $\mathbf{h}$.

First, we examine the effect of the parameter κ in Fig. 2. We observe that, when κ is too small or too large, it can degrade the estimation performance. When κ is too small, the function $f_\kappa(\cdot)$ cannot efficiently address the coherence issue and this leads to an incorrect path estimation, as previously illustrated in Fig. 1. However, if κ is too large, the peak removal step (line 9 of Algorithm 1) in the function $f_\kappa(\cdot)$ may remove another actual channel path close to the dominant path. Hence, one should choose an appropriate value of κ to obtain the best performance. For example, in this case, we choose $\kappa = 5, 4, 3$ for $b = 1, 2, 3$, respectively.

Next, we show the NMSE performance with 3-bit ADCs versus the SNR in Fig. 3. We compare the proposed deep unfolded network with the case where we heuristically set the step sizes η . It can be seen that the deep unfolded channel estimator with trained step sizes and scaling parameters for the logistic activation function gives better performance than when we heuristically set the step sizes. In Fig. 4, we show the

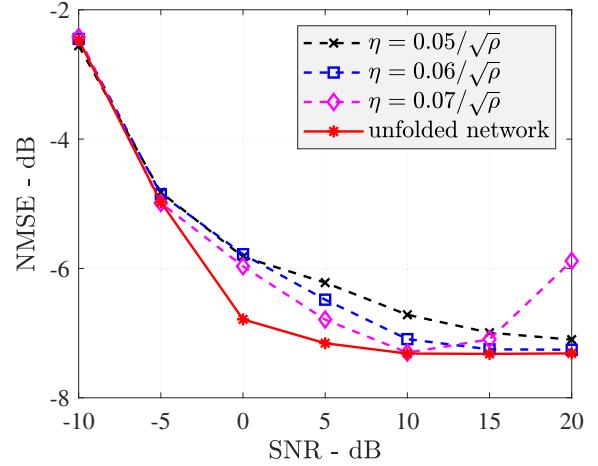


Fig. 3: NMSE performance versus SNR with $b = 3$ bits.

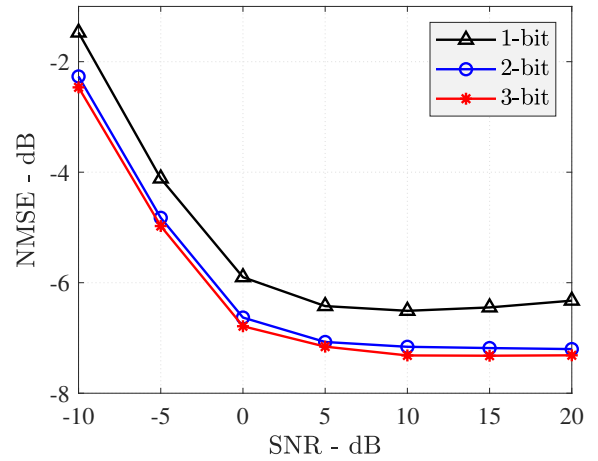


Fig. 4: NMSE performance of the deep unfolded network versus SNR for different ADC resolutions.

performance of the deep unfolded network versus the SNR for different ADC resolutions. We observe that 2- and 3-bit ADCs yield very similar performances, both significantly better than the case of 1-bit ADCs.

V. CONCLUSION

In this paper, we proposed an efficient on-grid near-field channel estimation method for massive MIMO systems with low-resolution ADCs by exploiting the polar-domain sparsity of the near-field channel. A reformulated sparse low-resolution near-field ML channel estimation problem was formulated by exploiting the approximation of the cdf of a normal random variable as a logistic activation function. In addition, the deep unfolding technique was applied to optimize the proposed algorithm and the efficiency of the proposed method was demonstrated via numerical results.

REFERENCES

- [1] N. Rajatheva, I. Atzeni, E. Björnson *et al.*, “White paper on broadband connectivity in 6G,” <https://oulurepo.oulu.fi/handle/10024/36799>, 2020.
- [2] Z. Zhou, X. Gao, J. Fang, and Z. Chen, “Spherical wave channel and analysis for large linear array in LoS conditions,” in *Proc. IEEE Globecom Workshops*, San Diego, CA, USA, 2015.
- [3] N. J. Myers, J. Kaleva, A. Tölli, and R. W. Heath, “Message passing-based link configuration in short range millimeter wave systems,” *IEEE Transactions on Communications*, vol. 68, no. 6, pp. 3465–3479, 2020.
- [4] M. Cui and L. Dai, “Channel estimation for extremely large-scale MIMO: Far-field or near-field?” *IEEE Trans. Commun.*, vol. 70, no. 4, pp. 2663–2677, 2022.
- [5] X. Zhang, H. Zhang, and Y. C. Eldar, “Near-field sparse channel representation and estimation in 6G wireless communications,” *IEEE Trans. Commun.*, vol. 72, no. 1, pp. 450–464, 2024.
- [6] Z. Wu and L. Dai, “Multiple access for near-field communications: SDMA or LDMA?” *IEEE J. Select. Areas Commun.*, vol. 41, no. 6, pp. 1918–1935, 2023.
- [7] Z. Wu, M. Cui, and L. Dai, “Enabling more users to benefit from near-field communications: From linear to circular array,” *IEEE Trans. Wireless Commun. (Early Access)*, 2023.
- [8] X. Wei and L. Dai, “Channel estimation for extremely large-scale massive MIMO: Far-field, near-field, or hybrid-field?” *IEEE Commun. Letters*, vol. 26, no. 1, pp. 177–181, 2022.
- [9] Y. Lu and L. Dai, “Near-field channel estimation in mixed LoS/NLoS environments for extremely large-scale MIMO systems,” *IEEE Trans. Commun.*, vol. 71, no. 6, pp. 3694–3707, 2023.
- [10] W. Liu, H. Ren, C. Pan, and J. Wang, “Deep learning based beam training for extremely large-scale massive MIMO in near-field domain,” *IEEE Commun. Letters*, vol. 27, no. 1, pp. 170–174, 2023.
- [11] Y. Zhang, X. Wu, and C. You, “Fast near-field beam training for extremely large-scale array,” *IEEE Wireless Commun. Letters*, vol. 11, no. 12, pp. 2625–2629, 2022.
- [12] Y. Lu, Z. Zhang, and L. Dai, “Hierarchical beam training for extremely large-scale MIMO: From far-field to near-field,” *IEEE Trans. Commun. (Early Access)*, 2023.
- [13] M. Cui, L. Dai, Z. Wang, S. Zhou, and N. Ge, “Near-field rainbow: Wideband beam training for XL-MIMO,” *IEEE Trans. Wireless Commun.*, vol. 22, no. 6, pp. 3899–3912, 2023.
- [14] Y. Mei, Z. Gao, D. Mi, M. Zhou, D. Zheng, M. Matthaiou, P. Xiao, and R. Schober, “Massive access in extra large-scale MIMO with mixed-ADC over near-field channels,” *IEEE Trans. Veh. Technol.*, vol. 72, no. 9, pp. 12 373–12 378, 2023.
- [15] L. V. Nguyen, D. H. N. Nguyen, and A. L. Swindlehurst, “Deep learning for estimation and pilot signal design in few-bit massive MIMO systems,” *IEEE Trans. Wireless Commun.*, vol. 22, no. 1, pp. 379–392, 2023.
- [16] S. R. Bowling, M. T. Khasawneh, S. Kaewkuekool, and B. R. Cho, “A logistic approximation to the cumulative normal distribution,” *J. Industrial Engineering and Management*, vol. 2, no. 1, pp. 114–127, 2009.
- [17] L. V. Nguyen, A. L. Swindlehurst, and D. H. N. Nguyen, “Linear and deep neural network-based receivers for massive MIMO systems with one-bit ADCs,” *IEEE Trans. Wireless Commun.*, vol. 20, no. 11, pp. 7333–7345, 2021.
- [18] J. R. Hershey, J. L. Roux, and F. Weninger, “Deep unfolding: Model-based inspiration of novel deep architectures,” *arXiv:1409.2574*, 2014.