

## RESEARCH ARTICLE

# Near-Field Terahertz Communications: Model-Based and Model-Free Channel Estimation

AHMET M. ELBIR<sup>1</sup>, (Senior Member, IEEE), WEI SHI<sup>2</sup>, (Member, IEEE),  
ANASTASIOS K. PAPAZAFEIROPOULOS<sup>1,3</sup>, (Senior Member, IEEE),  
PANDELIS KOURTESSIS<sup>1,3</sup>, (Member, IEEE), AND SYMEON CHATZINOTAS<sup>1</sup>, (Fellow, IEEE)

<sup>1</sup>Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg, 1855 Esch-sur-Alzette, Luxembourg

<sup>2</sup>School of Information Technology, Carleton University, Ottawa, ON K1S 5B6, Canada

<sup>3</sup>Communications and Intelligent Systems Research Group, University of Hertfordshire, AL10 9AB Hatfield, U.K.

Corresponding author: Ahmet M. Elbir (ahmetmelbir@gmail.com)

This work was supported in part by the Horizon Project Terahertz Reconfigurable Metasurfaces for Ultra-High Rate Wireless Communications (TERRAMETA), in part by the Natural Sciences and Engineering Research Council of Canada, and in part by Ericsson Canada.

**ABSTRACT** Terahertz (THz) band is expected to be one of the key enabling technologies of the sixth generation (6G) wireless networks because of its abundant available bandwidth and very narrow beamwidth. Due to high frequency operations, electrically small array apertures are employed, and the signal wavefront becomes spherical in the near-field. Therefore, near-field signal model should be considered for channel acquisition in THz systems. Unlike prior works which mostly ignore the impact of near-field beam-squint (NB) and consider either narrowband scenario or far-field models, this paper introduces both a model-based and a model-free techniques for wideband THz channel estimation in the presence of NB. The model-based approach is based on orthogonal matching pursuit (OMP) algorithm, for which we design an NB-aware dictionary. The key idea is to exploit the angular and range deviations due to the NB. We then employ the OMP algorithm, which accounts for the deviations thereby ipso facto mitigating the effect of NB. We further introduce a federated learning (FL)-based approach as a model-free solution for channel estimation in a multi-user scenario to achieve reduced complexity and training overhead. Through numerical simulations, we demonstrate the effectiveness of the proposed channel estimation techniques for wideband THz systems in comparison with the existing state-of-the-art techniques.

**INDEX TERMS** Beamsquint, channel estimation, federated learning, machine learning, near-field, orthogonal matching pursuit, sparse recovery, terahertz.

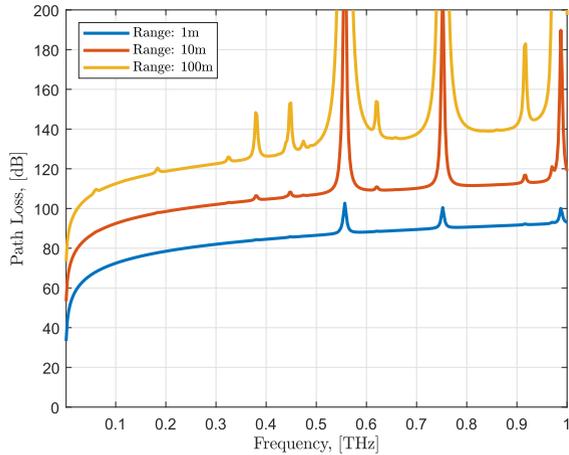
## I. INTRODUCTION

Terahertz (THz) band is expected to be a key component of the sixth generation (6G) of wireless cellular networks because of its abundant available bandwidth. In particular, THz-empowered systems are envisioned to demonstrate revolutionary enhancement under high data rate (>100Gb/s), extremely low propagation latency (<1ms) and ultra reliability (99.999%) [1], [2] and massive connectivity ( $10^3$  devices/100m<sup>2</sup>) [3].

Although demonstrating the aforementioned advantages, signal processing in THz band faces several THz-specific

challenges that should be taken into account accordingly. These challenges include, among others, severe path loss due to spreading loss and molecular absorption, extremely-sparse path model, very short transmission distance and *beam-squint* (see the full list in [4] and [5]). The path loss for THz-band transmission for various transmission ranges is illustrated in Fig. 1. In order to combat some of these challenges, e.g., path loss, analogues to massive multiple-input multiple-output (MIMO) arrays in millimeter-wave (mm-Wave) systems [6], [7], ultra-massive MIMO architectures are envisioned, wherein subcarrier-independent (SI) analog beamformers are employed. In wideband signal processing, the weights of the analog beamformers are subject to a single subcarrier frequency, i.e., carrier frequency [8], [9].

The associate editor coordinating the review of this manuscript and approving it for publication was Ronald Chang<sup>1</sup>.



**FIGURE 1.** Path loss (in dB) due to molecular absorption for various transmission ranges.

Therefore, the directions of the generated beams at different subcarriers differentiate and point to different directions causing beam-squint phenomenon [4], [10]. The impact of beam-squint in different frequency bands is presented in Fig. 2.

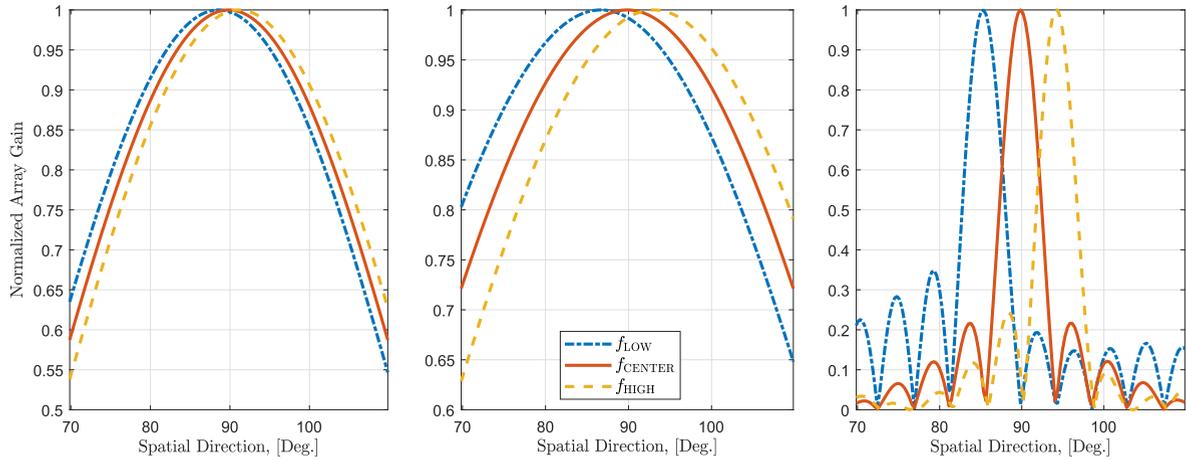
#### A. RELATED WORKS

While wideband mm-Wave channel estimation has been extensively studied in the literature [6], [11], [12], [13], [14], wideband THz channel estimation, on the other hand is relatively new [2]. Specifically, the existing solutions are categorized into two classes, i.e., hardware-based techniques [15] and algorithmic methods [16], [17]. The first category of solutions consider employing time delayer (TD) networks together with phase shifters so that true-time-delay (TTD) of each subcarrier can be obtained for beamformer design. The TD networks are used to realize virtual SD analog beamformers so that the impact of beam-squint can be mitigated for hybrid beamforming [18]. In particular, [15] devises a generalized simultaneous orthogonal matching pursuit (GSOMP) technique by exploiting the SD information collected via TD network hence achieves close to minimum mean-squared-error (MMSE) performance. However, these solutions require additional hardware, i.e., each phase shifter is connected to multiple TDs, each of which consumes approximately 100 mW, which is more than that of a phase shifter (40 mW) in THz [4]. The second category of solutions do not employ additional hardware components. Instead, advanced signal processing techniques have been proposed to compensate beam-squint. Specifically, an OMP-based beam-squint pattern detection (BSPD) approach was proposed in [16] for the recovery of the support pattern among all subcarriers in the beamspace and construct one-to-one match between the physical and spatial (i.e., deviated due to beam-squint in the beamspace) directions. Also, [19] proposed an angular-delay rotation method, which suffers from coarse beam-squint estimation and high training overhead due to

the use of complete discrete Fourier transform (DFT) matrix. However, it suffers from inaccurate support detection and low precision.

Besides the aforementioned model-based channel estimation techniques, model-free approaches, such as machine learning (ML), have also been suggested for THz channel estimation [20], [21]. For instance, ML-based learning models such as deep convolutional neural network (DCNN) [21], generative adversarial network (GAN) [22] and deep kernel learning (DKL) [20], have been proposed to lower the complexity involved during the channel inference as well as the complexity arising from the usage of the ultra-massive antenna elements. However, the works [20], [21] only consider the narrowband THz systems, which do not exploit wideband scenario, which is the main reason to climb up to the THz-band to achieve better communication performance. In [22], the wideband scenario is considered but the effect of beam-squint is ignored. Federated learning (FL)-based wideband THz channel estimation is studied in [17] and [23], wherein the labels of the model are obtained via beamspace support alignment (BSA) and sparse Bayesian learning (SBL), respectively. Specifically, the BSA approach suffers from inaccurate support detection for estimating the physical directions, and the SBL has high computational complexity as well as the requirement of the estimation of noise variance. Furthermore, both methods are based on the far-field assumption, which may not be satisfied in THz-band scenario.

The aforementioned THz works [15], [16], [17], [18], [19], [23] as well as the conventional wireless systems operating at sub-6 GHz and mm-Wave bands [11], [12], [13], [14] mostly incorporate far-field plane-wave model whereas the transmission range is shorter in THz-band such that the users are usually in the near-field region [5]. Specifically, the plane wavefront is spherical in the near-field when the transmission range is shorter than the Fraunhofer distance [24]. As a result, the channel acquisition algorithms should take into account near-field model (see, e.g., Fig. 3), which depends on both direction and range information for accurate signal processing [4]. Among the works investigating the near-field signal model, [25], [26], [27], [28], [29] consider the near-field scenario while the effect of beam-squint is ignored and only mm-Wave scenario is investigated. In particular, [28] considers the wideband mm-Wave channel estimation while the authors in [29] devise a machine learning (ML)-based approach for narrowband near-field channel estimation. In addition to near-field-only model, hybrid (near- and far-field) models are also present in the literature [25], [27], wherein only narrowband transceiver architectures are considered. On the other hand, several methods have been proposed to compensate the far-field beam-squint for both THz channel estimation [15], [16], [17], [23] and beamforming [18], [30], [31] applications. Nevertheless, THz channel estimation in the presence of *near-field beam-squint (NB)* remains relatively unexamined.



**FIGURE 2.** Normalized array gain with respect to spatial direction at low, center and high end subcarriers for (left) 3.5 GHz,  $B = 0.1$  GHz; (middle) 28 GHz,  $B = 2$  GHz; and (right) 300 GHz, 30 GHz, respectively.

## B. CONTRIBUTIONS

In this work, we introduce both a model-based and a model-free techniques for near-field THz channel estimation in the presence of NB. In the model-based approach, we propose an NB-aware (NBA) THz channel estimation technique based on OMP, henceforth called NBA-OMP. For the model-free approach, we devise a federated learning (FL) approach. Compared to conventional centralized learning (CL)-based approaches, FL is more communication-efficient since FL does not involve the transmission of datasets between the users and the server. Specifically, FL is particularly helpful for very-large antenna arrays, e.g., ultra-massive MIMO arrays which are envisioned to be employed in THz systems [4], [5].

We design a novel NBA dictionary, whose columns are composed of near-field subcarrier-dependent (SD) steering vectors spanning the whole angular spectrum and the transmission ranges up to the Fraunhofer distance. Then, we introduce our proposed NBA-OMP approach for wideband THz channel estimation.

The key idea of the proposed approach is that the degree of beam-squint is proportionally known prior to the direction-of-arrival (DoA)/range estimation while it depends on the unknown user location. For example, consider  $f_m$  and  $f_c$  to be the frequencies for the  $m$ -th and center subcarriers, respectively. When  $\theta$  is the physical user direction, the spatial direction corresponding to the  $m$ -th-subcarrier is shifted by  $\frac{f_c}{f_m}\theta$ . Thus, we employ the OMP algorithm, which accounts for this deviation thereby *ipso facto* compensating the effect of NB. We have conducted several numerical experiments to demonstrate the effectiveness of the proposed approach in comparison with the existing THz channel-estimation-based methods [16], [25].

## C. OUTLINE AND NOTATION

### 1) OUTLINE

In the remainder of the paper, we first introduce the signal model for multi-user wideband THz ultra-massive

MIMO system in Sec. II. Then, we present the proposed NBA OMP technique for a model-based and a model-free channel estimation in Sec. IV and Sec. V, respectively. The complexity and overhead analysis of the proposed approaches are discussed in Sec. VI. Sec. VII presents the numerical simulations and we finalize the paper in Sec. VIII with concluding remarks.

### 2) NOTATION

Throughout the paper, we denote the vector and matrices via bold lowercase and uppercase letters, respectively. The transpose and conjugate transpose operations are denoted by  $(\cdot)^T$  and  $(\cdot)^H$ , respectively. We denote the  $n$ -th column of a matrix  $\mathbf{A}$  as  $\mathbf{A}_n$  and  $\mathbf{A}^\dagger$  represents the Moore-Penrose pseudo-inverse. For a vector  $\mathbf{a}$ , the  $n$ -th element of  $\mathbf{a}$  is represented by  $[\mathbf{a}]_n$ .  $\lceil \cdot \rceil$  is the ceiling operator,  $\nabla$  represents the gradient operation.  $\Sigma(a) = \frac{\sin \pi a b}{b \sin \pi a}$  denotes the Dirichlet sinc function, and  $\mathbb{E}\{\cdot\}$  stands for the expectation operation.  $\|\cdot\|_2$  and  $\|\cdot\|_{\mathcal{F}}$  denote the  $l_2$  and Frobenius norms, respectively.

## II. SYSTEM MODEL

Consider a wideband THz MIMO architecture with hybrid analog/digital beamforming over  $M$  subcarriers. We assume that the base station (BS) has  $N$  antennas and  $N_{\text{RF}}$  radio-frequency (RF) chains to serve  $K$  single-antenna users. Let  $\mathbf{s}[m] = [s_1[m], \dots, s_K[m]]^T$  denote the data symbols, where  $m \in \mathcal{M} = \{1, \dots, M\}$ , which are processed via a  $K \times K$  SD baseband beamformer  $\mathbf{F}_{\text{BB}}[m] = [\mathbf{f}_{\text{BB}_1}[m], \dots, \mathbf{f}_{\text{BB}_K}[m]]$ . In order to steer the generated beams toward users in downlink, an  $N \times N_{\text{RF}}$  SI analog beamformer  $\mathbf{F}_{\text{RF}}$  ( $N_{\text{RF}} = K < N$ ) is employed. Since the analog beamformers are realized with phase-shifters, they have constant-modulus constraint, i.e.,  $[\mathbf{F}_{\text{RF}}]_{i,j} = \frac{1}{\sqrt{N}}$ . Then, the  $N \times 1$  transmitted signal becomes

$$\tilde{\mathbf{x}}[m] = \mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}[m]\mathbf{s}[m]. \quad (1)$$

The transmitted signal is passed through the wireless channel, and it is received by the  $k$ -th user at the  $m$ -th

subcarrier as

$$y_k[m] = \mathbf{h}_k^T[m] \tilde{\mathbf{x}}[m] + w_k[m] \\ = \mathbf{h}_k^T[m] \sum_{i=1}^K \mathbf{F}_{\text{RF}} \mathbf{f}_{\text{BB}_i}[m] s_i[m] + w_k[m], \quad (2)$$

where  $w_k[m] \in \mathbb{C}$  represents the complex additive white Gaussian noise with variance of  $\sigma_n^2$ , i.e.,  $w_k[m] \sim \mathcal{CN}(0, \sigma_n^2)$ .

### A. THz CHANNEL MODEL

Due to limited reflected path components and negligible scattering, the THz channel is usually constructed as the superposition of a single line-of-sight (LoS) path with a few assisting non-LoS (NLoS) paths [5], [10], [30]. In addition, multipath channel models are also widely used, especially for indoor applications [5], [32]. Hence, we consider a general scenario, wherein the  $N \times 1$  channel matrix for the  $k$ -th user at the  $m$ -th subcarrier is represented by the combination of  $L$  paths as [5]

$$\mathbf{h}_k[m] = \sqrt{\frac{N}{L}} \sum_{l=1}^L \alpha_{k,m,l} \mathbf{a}(\phi_{k,l}, r_{k,l}) e^{-j2\pi \tau_{k,l} f_m}, \quad (3)$$

where  $\tau_{k,l}$  represents the time delay of the  $l$ -th path corresponding to the array origin.  $\mathbf{a}(\phi_{k,l}, r_{k,l}) \in \mathbb{C}^N$  denotes the steering vector corresponding to the physical direction  $\phi_{k,l}$  and range  $r_{k,l}$ .  $\alpha_{k,m,l} \in \mathbb{C}$  denotes the complex path gain and the expected value of its magnitude for the indoor THz multipath model is given by

$$\mathbb{E}\{|\alpha_{k,m,l}|^2\} = \left(\frac{c_0}{4\pi f_m r_{k,l}}\right)^2 e^{-k_{\text{abs}}(f_m) r_{k,l}}, \quad (4)$$

where  $f_m$  is the  $m$ -th subcarrier frequency,  $c_0$  is speed of light,  $r_{k,l}$  represents the distance from the  $k$ -th user to the array origin and  $k_{\text{abs}}(f_m)$  is the SD medium absorption coefficient [5], [33], [34]. Furthermore,  $f_m = f_c + \frac{B}{M}(m-1 - \frac{M-1}{2})$ , where  $f_c$  and  $B$  are carrier frequency and bandwidth, respectively.

### B. NEAR-FIELD ARRAY MODEL

Due to high frequency operations as well as employing extremely small wavelength, THz-band transmission is likely to encounter near-field phenomenon for close-proximity users. Specifically, the far-field model involves the reception of the transmitted signal at the users as plane-wave. However, the plane wavefront is spherical in the near-field if the transmission range is shorter than the Fraunhofer distance  $F = \frac{2D^2}{\lambda}$ , where  $D$  is the array aperture and  $\lambda = \frac{c_0}{f_c}$  is the wavelength [23], [24]. The illustration of beam-squint for both far-field and near-field models is given in Fig. 3. For a uniform linear array (ULA), the array aperture is  $D = (N-1)d$ , where  $d = \frac{\lambda}{2}$  is the element spacing. Thus, for THz-band transmission, this distance becomes small such that the near-field signal model should be employed since  $r_{k,l} < F$ . For instance, when  $f_c = 300$  GHz and  $N = 256$ , the Fraunhofer distance is  $F = 32.76$  m.

We define the near-field steering vector  $\mathbf{a}(\phi_{k,l}, r_{k,l}) \in \mathbb{C}^N$  corresponding to the physical DoA  $\phi_{k,l}$  and range  $r_{k,l}$  as

$$\mathbf{a}(\phi_{k,l}, r_{k,l}) = \frac{1}{\sqrt{N}} [e^{-j2\pi \frac{d}{\lambda} r_{k,l}^{(1)}}, \dots, e^{-j2\pi \frac{d}{\lambda} r_{k,l}^{(N)}}]^T, \quad (5)$$

where  $\phi_{k,l} = \sin \tilde{\phi}_{k,l}$  with  $\tilde{\phi}_{k,l} \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ , and  $r_{k,l}^{(n)}$  is the distance between the  $k$ -th user and the  $n$ -th antenna as

$$r_{k,l}^{(n)} = \left(r_{k,l}^2 + 2(n-1)^2 d^2 - 2 r_{k,l} (n-1) d \phi_{k,l}\right)^{\frac{1}{2}}. \quad (6)$$

In (6),  $r_{k,l}^{(n)}$  involves complicated terms as a function of the antenna index  $n$ , which makes it difficult to analyze spherical-wave model. In order to simplify (6), Fresnel approximation can be employed [24], [35], [36]. Specifically, the Fresnel approximation provides modeling the received spherical-wave as an approximated quadric surface [24], [36]. Thus, (6) becomes

$$r_{k,l}^{(n)} \approx r_{k,l} - (n-1)d\phi_{k,l} + (n-1)^2 d^2 \zeta_{k,l}, \quad (7)$$

where  $\zeta_{k,l} = \frac{1-\phi_{k,l}^2}{2r_{k,l}}$ . Rewrite (5) as

$$\mathbf{a}(\phi_{k,l}, r_{k,l}) \approx e^{-j2\pi \frac{f_c}{c_0} r_{k,l}} \tilde{\mathbf{a}}(\phi_{k,l}, r_{k,l}), \quad (8)$$

where the  $n$ -th element of  $\tilde{\mathbf{a}}(\phi_{k,l}, r_{k,l}) \in \mathbb{C}^N$  is

$$[\tilde{\mathbf{a}}(\phi_{k,l}, r_{k,l})]_n = \frac{1}{\sqrt{N}} e^{j2\pi \frac{f_c}{c_0} ((n-1)d\phi_{k,l} - (n-1)^2 d^2 \zeta_{k,l})}. \quad (9)$$

The steering vector in (8) corresponds to the physical location  $(\phi_{k,l}, r_{k,l})$ , which deviates to the spatial location  $(\tilde{\phi}_{k,m,l}, \tilde{r}_{k,m,l})$  in the beamspace arising from the absent of SD analog beamformers. Then, the  $n$ -th entry of the deviated steering vector in (9) for the spatial location is

$$[\tilde{\mathbf{a}}(\tilde{\phi}_{k,m,l}, \tilde{r}_{k,m,l})]_n = e^{j2\pi \frac{f_m}{c_0} ((n-1)d\tilde{\phi}_{k,m,l} - (n-1)^2 d^2 \tilde{\zeta}_{k,m,l})}. \quad (10)$$

### C. PROBLEM FORMULATION

The aim of this work is to estimate the wideband THz channel  $\mathbf{h}_k[m]$  in the presence of NB in downlink scenario.

1: For the proposed model-based technique, we assume that the users are synchronized and they use the received pilot signals transmitted by the BS. We also assume that all users are in the near-field region of the BS, i.e.,  $r_{k,l} \leq F$ .

2: For the proposed model-free technique, it is assumed that the datasets of all users are collected prior to the model training stage, and the labels of these datasets are determined by the model-based technique introduced in Sec. IV.

In what follows, we first introduce the NB model, then present the proposed approaches for model-based and model-free channel estimation in the presence of NB.

### III. NB MODEL

Compared to mm-Wave frequencies, in THz-band, the bandwidth is so wide that a single-wavelength assumption for beamforming cannot hold and it leads to the squint of physical DoA/ranges  $\{\phi_{k,l}, r_{k,l}\}$  in the spatial domain. Hence,

we define the relationship between the physical and spatial DoA/ranges in the following theorem:

*Theorem 1:* Denote  $\mathbf{u} \in \mathbb{C}^N$  and  $\mathbf{v}_m \in \mathbb{C}^N$  as the arbitrary near-field steering vectors corresponding to the physical (i.e.,  $\{\phi_{k,l}, r_{k,l}\}$ ) and spatial (i.e.,  $\{\bar{\phi}_{k,m,l}, \bar{r}_{k,m,l}\}$ ) locations given in (9) and (10), respectively. Then, in spatial domain at subcarrier frequency  $f_m$ , the array gain achieved by  $\mathbf{u}^H \mathbf{v}_m$  is maximized and the generated beam is focused at the location  $\{\bar{\phi}_{k,m,l}, \bar{r}_{k,m,l}\}$  such that

$$\bar{\phi}_{k,m,l} = \eta_m \phi_{k,l}, \quad \bar{r}_{k,m,l} = \frac{1 - \eta_m^2 \phi_{k,l}^2}{\eta_m (1 - \phi_{k,l}^2)} r_{k,l}, \quad (11)$$

where  $\eta_m = \frac{f_c}{f_m}$  represents the proportional deviation of DoA/ranges.

*Proof:* Define the array gain achieved by  $\mathbf{v}_m$  on an arbitrary user location  $\{\phi_{k,l}, r_{k,l}\}$  with steering vector  $\mathbf{u}$  as

$$\begin{aligned} G(\phi_{k,l}, r_{k,l}, m) &= \frac{|\mathbf{u}^H \mathbf{v}_m|^2}{N^2} \\ &= \frac{1}{N^2} \left| \sum_{n=0}^{N-1} e^{j \frac{2\pi}{c_0} [nd(f_m \bar{\phi}_{k,m,l} - f_c \phi_{k,l}) - n^2 d^2 (f_m \bar{\zeta}_{k,m,l} - f_c \zeta_{k,l})]} \right|^2 \\ &= \frac{1}{N^2} \left| \sum_{n=0}^{N-1} e^{j \frac{2\pi n}{c_0} (f_m \kappa_m - f_c \kappa)} \right|^2 \\ &= \frac{1}{N^2} \left| \frac{1 - e^{-j2\pi N (f_m \kappa_m - f_c \kappa)}}{1 - e^{-j2\pi (f_m \kappa_m - f_c \kappa)}} \right|^2 \\ &= \frac{1}{N^2} \left| \frac{\sin(\pi N (f_m \kappa_m - f_c \kappa))}{\sin(\pi (f_m \kappa_m - f_c \kappa))} \right|^2 \\ &= |\Sigma(f_m \kappa_m - f_c \kappa)|^2. \end{aligned} \quad (12)$$

Furthermore, we define

$$\bar{\zeta}_{k,m,l} = \eta_m \zeta_{k,l} = \frac{1 - \bar{\phi}_{k,m,l}^2}{2\bar{r}_{k,m,l}}, \quad (13)$$

$$\kappa_m = d(\bar{\phi}_{k,m,l} - nd\bar{\zeta}_{k,m,l}), \quad (14)$$

$$\kappa = d(\phi_{k,l} - nd\zeta_{k,l}). \quad (15)$$

The array gain in (12) implies that most of the power is focused only on a small portion of the beamspace due to the power-focusing capability of  $\Sigma(a)$ , which substantially reduces across the subcarriers as  $|f_m - f_c|$  increases. Furthermore,  $|\Sigma(a)|^2$  gives peak when  $a = 0$ , i.e.,  $f_m \bar{\phi}_{k,m,l} = f_c \phi_{k,l}$  and  $f_m \bar{\zeta}_{k,m,l} = f_c \zeta_{k,l}$ . Therefore, we have

$$\bar{\phi}_{k,m,l} = \eta_m \phi_{k,l}. \quad (16)$$

Then, by using  $\bar{\zeta}_{k,m,l} = \eta_m \zeta_{k,l}$ , we get

$$\bar{r}_{k,m,l} = \frac{1 - \eta_m^2 \phi_{k,l}^2}{\eta_m (1 - \phi_{k,l}^2)} r_{k,l}. \quad (17)$$

□

Finally, by combining (7), (11) and (17), we define the NB in terms of DoAs and the ranges of the users as

$$\begin{aligned} \Delta(\phi_{k,l}, m) &= \bar{\phi}_{k,m,l} - \phi_{k,l} = (\eta_m - 1)\phi_{k,l}, \quad (18) \\ \Delta(r_{k,l}, m) &= \bar{r}_{k,m,l} - r_{k,l} = \left( \frac{1 - \eta_m^2 \phi_{k,l}^2}{\eta_m (1 - \phi_{k,l}^2)} - 1 \right) r_{k,l}. \end{aligned} \quad (19)$$

In Fig. 3, the array gain is computed for both far- and near-field cases, wherein the coordinates with maximum array gain are achieved at different locations for different subcarriers because of beam-squint. Specifically, the physical direction/range  $(\phi_{k,l}, r_{k,l})$  is observed as  $(\bar{\phi}_{k,m,l}, \bar{r}_{k,m,l})$  at the  $m$ -th subcarrier. While the relationship between the physical and spatial directions are linear, i.e.,  $\bar{\phi}_{k,m,l} = \eta_m \phi_{k,l}$ , the spatial range information deviates nonlinearly, i.e.,  $\bar{r}_{k,m,l} = \frac{1 - \eta_m^2 \phi_{k,l}^2}{\eta_m (1 - \phi_{k,l}^2)} r_{k,l}$ .

#### IV. MODEL-BASED SOLUTION: NBA-OMP

We first introduce how the NBA dictionary is designed, then present the implementation steps of the proposed NBA-OMP technique.

##### A. NBA DICTIONARY DESIGN

The key idea of the proposed NBA dictionary design is to utilize the prior knowledge of  $\eta_m$  to obtain beam-squint-corrected steering vectors. In other words, for an arbitrary physical DoA and range, we can readily find the spatial DoA and ranges as  $\bar{\phi} = \eta_m \phi$  and  $\bar{r} = \frac{1 - \eta_m^2 \phi^2}{\eta_m (1 - \phi^2)} r$ . Using this observation, we design the NBA dictionary  $\mathcal{C}_m$  composed of steering vectors  $\mathbf{c}(\phi_m, r_m) \in \mathbb{C}^N$  as

$$\mathcal{C}_m = \{\mathbf{c}(\phi_m, r_m) | \phi_m \in [-\eta_m, \eta_m], r_m \in \mathbb{R}^+\}, \quad (20)$$

where the  $n$ -th element of  $\mathbf{c}(\phi_m, r_m)$  is

$$[\mathbf{c}(\phi_m, r_m)]_n = e^{j2\pi \frac{f_m}{c_0} ((n-1)d\phi_m - (n-1)^2 d^2 \zeta_m)}, \quad (21)$$

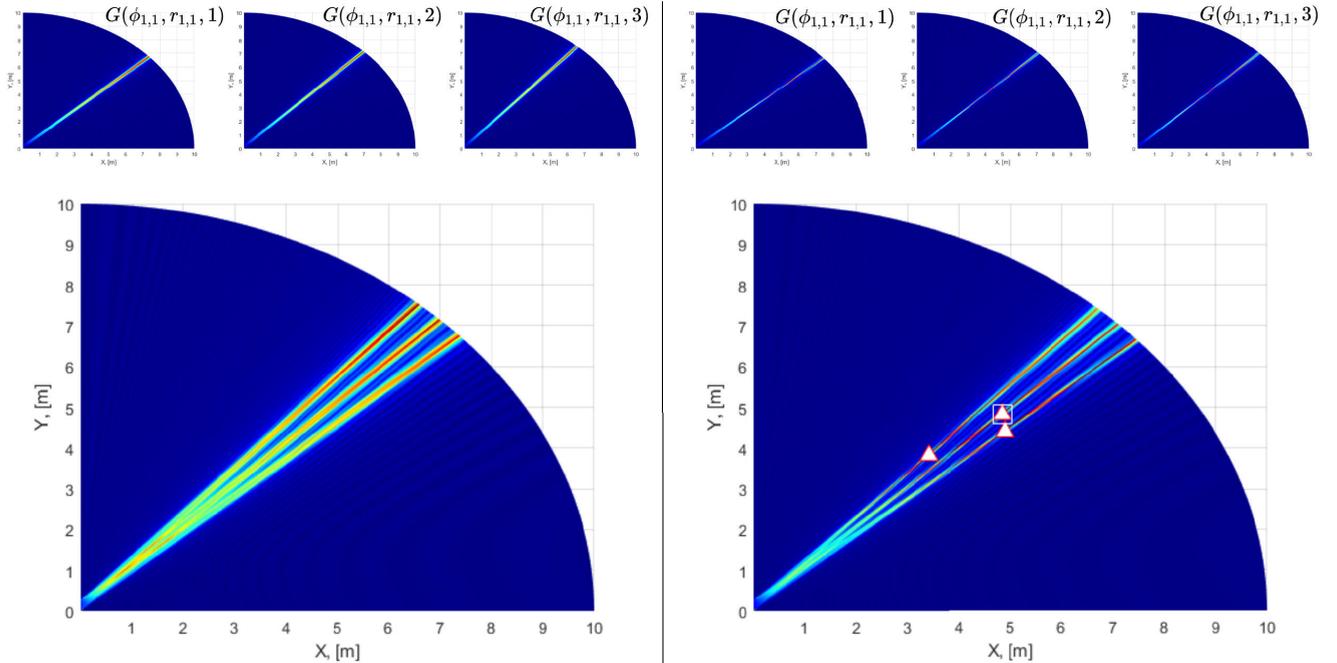
where  $\zeta_m = \frac{1 - \phi_m^2}{2r_m}$ . It is worth noting that the NBA dictionary design requires the approximated steering vector models in order to easily define and compensate the beam-squint.

Using the NBA dictionary  $\mathcal{C}_m$ , instead of SI steering vectors  $\mathbf{a}(\phi, r)$ , the SD virtual steering vectors  $\mathbf{c}(\phi_m, r_m)$  can be constructed for the OMP algorithm. Once the beamspace spectra is computed via OMP, the sparse channel support corresponding to the SD spatial DoA and ranges are obtained. Then, one can readily find the physical DoAs and ranges as  $\phi = \phi_m / \eta_m$  and  $r = \frac{\eta_m (1 - \phi_m^2)}{1 - \eta_m^2 \phi_m^2} r_m, \forall m \in \mathcal{M}$ .

The proposed NBA dictionary  $\mathcal{C}_m$  also holds spatial orthogonality as

$$\lim_{N \rightarrow +\infty} |\mathbf{c}^H(\phi_{m,i}, r_{m,i}) \mathbf{c}(\phi_{m,j}, r_{m,j})| = 0, \forall i \neq j. \quad (22)$$

In the next section, we present the channel estimation procedure with the proposed NBA dictionary.



**FIGURE 3.** Array gains  $G(\phi_{1,1}, r_{1,1}, m)$  in Cartesian coordinates for a single user ( $K = 1, L = 1$ ) located in the far-field ( $45^\circ, 800\text{m}$ ) (left) and near-field ( $45^\circ, 6\text{m}$ ) (right), respectively. Here,  $M = 3, f_c = 300\text{ GHz}$ , and  $B = 30\text{ GHz}$ . The top panel shows the gain for different subcarriers which are summed up to produce a composite array gain at the bottom for both far- and near-field cases clearly showing the beam-squint. The square represents the user location while the triangles correspond to the spatial locations (where the maximum array gain is achieved) at different subcarriers. Whereas the far-field beam-squint is only angular, the near-field beam-squint is across both range and angular domains.

**B. NEAR-FIELD CHANNEL ESTIMATION**

In downlink, the channel estimation stage is performed simultaneously during channel training by all the users, i.e.,  $k \in \mathcal{K} = \{1, \dots, K\}$ . Since the BS employs hybrid beamforming architecture, it activates only a single RF chain in each channel use to transmit the pilot signals during channel acquisition [6]. Denote  $\tilde{\mathbf{S}}[m] = \text{diag}\{\tilde{s}_1[m], \dots, \tilde{s}_P[m]\}$  as the  $P \times P$  orthogonal pilot signal matrix, then the BS employs  $P$  beamformer vectors as  $\tilde{\mathbf{F}} = [\tilde{\mathbf{f}}_1, \dots, \tilde{\mathbf{f}}_P] \in \mathbb{C}^{N \times P}$  ( $|\tilde{\mathbf{f}}_p| = 1/\sqrt{N}$ ) to send  $P$  orthogonal pilots, which are collected by the  $k$ -th user as

$$\mathbf{y}_k[m] = \tilde{\mathbf{S}}[m]\tilde{\mathbf{F}}[m]\mathbf{h}_k[m] + \mathbf{w}_k[m], \quad (23)$$

where  $\tilde{\mathbf{F}}[m] = \tilde{\mathbf{F}}^T[m] \in \mathbb{C}^{P \times N}$ . Assume that  $\tilde{\mathbf{F}}[m] = \mathbf{F} \in \mathbb{C}^{P \times N}$  and  $\tilde{\mathbf{S}}[m] = \mathbf{I}_P, \forall m \in \mathcal{M}$  [11], [15], [16], we get

$$\mathbf{y}_k[m] = \mathbf{F}\mathbf{h}_k[m] + \mathbf{w}_k[m]. \quad (24)$$

Note that the solution via traditional techniques, e.g., the least squares (LS) and MMSE estimator can be readily given respectively as

$$\mathbf{h}_k^{\text{LS}}[m] = (\mathbf{F}^H\mathbf{F})^{-1}\mathbf{F}^H\mathbf{y}_k[m], \quad (25)$$

$$\mathbf{h}_k^{\text{MMSE}}[m] = (\mathbf{R}_k^{-1}[m] + \mathbf{F}^H\mathbf{R}_k^{-1}[m]\mathbf{F})^{-1}\mathbf{F}^H\mathbf{y}_k[m], \quad (26)$$

where  $\mathbf{R}_k[m] = \mathbb{E}\{\mathbf{h}_k[m]\mathbf{h}_k^H[m]\}$  is the channel covariance matrix. Nevertheless, these methods require either prior information (e.g., MMSE) or have poor estimation performance (e.g., LS) [14], [37]. Furthermore, these methods require at

least  $P \geq N$  pilot signals, which can be heavy for channel training while our proposed approach exhibits a much lower channel training overhead (e.g, see Sec. VI).

By exploiting the THz channel sparsity, the THz channel can be represented in a sparse domain via support vector  $\mathbf{x}_k[m] \in \mathbb{C}^Q$ , which is an  $L$ -sparse vector, whose non-zero elements corresponds to the set

$$\{x_{k,l}[m]|x_{k,l}[m] \triangleq \sqrt{\frac{N}{L}}\alpha_{k,m,l}e^{-j2\pi\tau_{k,l}f_m}\}. \quad (27)$$

Then, the received signal in (24) is rewritten in sparse domain as

$$\mathbf{y}_k[m] = \mathbf{F}\mathbf{C}_m\mathbf{x}_k[m] + \mathbf{w}_k[m], \quad (28)$$

where  $\mathbf{F}$  is an  $P \times N$  fixed matrix corresponding to the hybrid beamformer weights during data transmission, and the sparse channel is represented as  $\mathbf{h}_k[m] = \mathbf{C}_m\mathbf{x}_k[m]$ , where  $\mathbf{C}_m$  is the  $N \times Q$  NBA dictionary matrix covering the spatial domain with  $\phi_{m,q} \in [-\eta_m, \eta_m]$  ( $\phi \in [-1, 1]$ ) and  $r_{m,q} \in [0, F]$  for  $q = 1, \dots, Q$  as

$$\mathbf{C}_m = [\mathbf{c}(\phi_{m,1}, r_{m,1}), \dots, \mathbf{c}(\phi_{m,Q}, r_{m,Q})]. \quad (29)$$

By utilizing the NBA dictionary matrix  $\mathbf{C}_m$ , we employ the OMP algorithm to effectively recover the sparse channel support. The proposed NBA-OMP technique is presented in Algorithm 1, which accepts  $\mathbf{y}_k[m], \mathbf{C}_m$  and  $\eta_m$  as inputs and it yields the output as the estimated channel  $\hat{\mathbf{h}}_k[m]$ , and the near-field beam-squint in terms of DoA and ranges, i.e.,  $\hat{\Delta}(\phi_{k,l}, m)$

**Algorithm 1** NBA-OMP

---

**Input:** Dictionary  $\mathbf{C}_m$ , observation  $\mathbf{y}_k[m]$ ,  $\mathbf{F}$  and  $\eta_m$   
**Output:**  $\hat{\mathbf{h}}_k[m]$ ,  $\hat{\Delta}(\phi_{k,l}, m)$  and  $\hat{\Delta}(r_{k,l}, m)$ .

- 1: **for**  $k \in \mathcal{K}$
- 2:    $l = 1, \mathcal{I}_{l-1} = \emptyset, \mathbf{r}_{l-1}[m] = \mathbf{y}_k[m], \forall m \in \mathcal{M}$ .
- 3:   **while**  $l \leq L$  **do**
- 4:      $q^* = \operatorname{argmax}_q \sum_{m=1}^M |\mathbf{c}^H(\phi_{m,q}, r_{m,q}) \mathbf{F}^H \mathbf{r}_{l-1}[m]|$ .
- 5:      $\mathcal{I}_l = \mathcal{I}_{l-1} \cup \{q^*\}$ .
- 6:      $\hat{\phi}_{k,l} = \frac{\phi_{m,q^*}}{\eta_m}$ .
- 7:      $\hat{r}_{k,l} = \frac{\eta_m(1-\phi_{m,q^*}^2)}{1-\eta_m^2\phi_{m,q^*}} r_{m,q^*}$ .
- 8:      $\hat{\Delta}(\phi_{k,l}, m) = (\eta_m - 1)\hat{\phi}_{k,l}, \forall m \in \mathcal{M}$ .
- 9:      $\hat{\Delta}(r_{k,l}, m) = (\eta_m - 1) \frac{1-\eta_m^2\hat{\phi}_{k,l}^2}{\eta_m(1-\hat{\phi}_{k,l}^2)} \hat{r}_{k,l}, \forall m \in \mathcal{M}$ .
- 10:      $\Psi_m(\mathcal{I}_l) = \mathbf{F}\mathbf{C}_m(\mathcal{I}_l)$ .
- 11:      $\mathbf{r}_l[m] = (\mathbf{I}_P - \Psi_m(\mathcal{I}_l)\Psi_m^\dagger(\mathcal{I}_l)) \mathbf{y}_k[m]$ .
- 12:      $l = l + 1$ .
- 13:   **end while**
- 14:    $\Xi_k = [\mathbf{a}(\hat{\phi}_{k,1}, \hat{r}_{k,1}), \dots, \mathbf{a}(\hat{\phi}_{k,L}, \hat{r}_{k,L})]$ .
- 15:   **for**  $m \in \mathcal{M}$
- 16:      $\hat{\mathbf{u}}_k[m] = \Psi_m^\dagger(\mathcal{I}_{l-1}) \mathbf{y}_k[m]$ .
- 17:      $\hat{\mathbf{h}}_k[m] = \Xi_k \hat{\mathbf{u}}_k[m]$ .
- 18: **end for**

---

and  $\hat{\Delta}(r_{k,l}, m)$ , respectively. In steps 4–6 of Algorithm 1, the orthogonality between the residual observation vector  $\mathbf{r}_l[m]$  and the columns of the dictionary matrix is checked as

$$q^* = \operatorname{argmax}_q \sum_{m=1}^M |\mathbf{c}^H(\phi_{m,q}, r_{m,q}) \mathbf{F}^H \mathbf{r}_{l-1}[m]|, \quad (30)$$

where  $q^*$  denotes the support index for the  $l$ th path iteration. Then, the DoA and range estimates can be found as  $\hat{\phi}_{k,l} = \frac{\phi_{m,q^*}}{\eta_m}$  and  $\hat{r}_{k,l} = \frac{\eta_m(1-\phi_{m,q^*}^2)}{1-\eta_m^2\phi_{m,q^*}} r_{m,q^*}$ , respectively. In a similar way, the corresponding beam-squint terms are  $\hat{\Delta}(\phi_{k,l}, m) = (\eta_m - 1)\hat{\phi}_{k,l}$  and  $\hat{\Delta}(r_{k,l}, m) = (\eta_m - 1) \frac{1-\eta_m^2\hat{\phi}_{k,l}^2}{\eta_m(1-\hat{\phi}_{k,l}^2)} \hat{r}_{k,l}$ . Once the DoA and ranges are obtained, the THz channel is constructed from the estimated support as  $\hat{\mathbf{h}}_k[m] = \Xi_k \hat{\mathbf{u}}_k[m]$  where  $\Xi_k = [\mathbf{a}(\hat{\phi}_{k,1}, \hat{r}_{k,1}), \dots, \mathbf{a}(\hat{\phi}_{k,L}, \hat{r}_{k,L})]$  and  $\hat{\mathbf{u}}_k[m] = \Psi_m^\dagger(\mathcal{I}_L) \mathbf{y}_k[m]$ , wherein  $\Psi_m(\mathcal{I}_L) = \mathbf{F}\mathbf{C}_m(\mathcal{I}_L)$  and  $\mathcal{I}_L$  includes the estimated support indices.

**V. MODEL-FREE SOLUTION: NBA-OMP-FL**

Due to operating at small wavelengths and severe path loss, ultra-massive arrays are envisioned to be employed at THz communications. This results in very large number of array data to be processed. For instance, when employing learning-based techniques, the size of the datasets is proportional to the number of antennas. Thus, huge communication overhead occurs due to the transmission of these datasets in CL-based techniques [17], [38], [39]. To alleviate this overhead,

we present an FL-based near-field THz channel estimation scheme in this section.

Denote  $\boldsymbol{\theta} \in \mathbb{R}^Z$  and  $\mathcal{D}_k$  as the vector of learnable parameters and the local dataset for the  $k$ -th user, respectively. Then, we have the relationship  $f(\boldsymbol{\theta})$  between the input of the learning model represented by  $\boldsymbol{\theta}$  and its prediction as  $\mathcal{Y}_k^{(i)} = f(\boldsymbol{\theta}) \mathcal{X}_k^{(i)}$ . Here, the sample index for the local dataset is  $i = 1, \dots, \mathbf{D}_k$ , where  $\mathbf{D}_k = |\mathcal{D}_k|$  denotes the number of samples in the  $k$ -th local dataset. Furthermore, we have  $\mathcal{X}_k^{(i)}$  and  $\mathcal{Y}_k^{(i)}$  representing the input and output, respectively, such that the input-output tuple of the local dataset is  $\mathcal{D}_k^{(i)} = (\mathcal{X}_k^{(i)}, \mathcal{Y}_k^{(i)})$ .

The input data is the received pilots, i.e.,  $\mathbf{y}_k[m]$ . For the output (label) data, we consider the channel estimates obtained via the proposed NBA-OMP approach in Algorithm 1. Then, by using  $\hat{\mathbf{h}}_k[m]$ , the output data is constructed as

$$\mathcal{Y}_k = \left[ \operatorname{Re}\{\hat{\mathbf{h}}_k[m]\}^\top, \operatorname{Im}\{\hat{\mathbf{h}}_k[m]\}^\top \right]^\top \in \mathbb{R}^{2N}. \quad (31)$$

When designing the input data, the real, imaginary and angle information of  $\mathbf{y}_k[m]$  are incorporated in order to improve the feature extraction performance of the learning model. Therefore, each input sample has a ‘‘three-channel’’ structure with the size of  $N_{\text{RF}} \times 3$  [38]. Thus, we have  $[\mathcal{X}_k]_1 = \operatorname{Re}\{\mathbf{y}_k[m]\}$ ,  $[\mathcal{X}_k]_2 = \operatorname{Im}\{\mathbf{y}_k[m]\}$  and  $[\mathcal{X}_k]_3 = \angle\{\mathbf{y}_k[m]\}$ .

In order to provide all the users to participate in the training in a distributed manner, each-user computes their model parameters and aims to minimize an average MSE cost. Thus, the FL-based model training problem can be cast as

$$\begin{aligned} & \underset{\boldsymbol{\theta}}{\text{minimize}} \quad \frac{1}{K} \sum_{k=1}^K \mathcal{L}_k(\boldsymbol{\theta}) \\ & \text{subject to: } f(\mathcal{X}_k^{(i)}|\boldsymbol{\theta}) = \mathcal{Y}_k^{(i)}, \end{aligned} \quad (32)$$

where  $\mathcal{L}_k(\boldsymbol{\theta})$  is the loss function for the learning model computed at the  $k$ -th user, and it is defined as

$$\mathcal{L}_k(\boldsymbol{\theta}) = \frac{1}{\mathbf{D}_k} \sum_{i=1}^{\mathbf{D}_k} \|f(\mathcal{X}_k^{(i)}|\boldsymbol{\theta}) - \mathcal{Y}_k^{(i)}\|_{\mathcal{F}}^2, \quad (33)$$

where  $f(\mathcal{X}_k^{(i)}|\boldsymbol{\theta})$  corresponds to the output of the learning model once it is fed with  $\mathcal{X}_k^{(i)}$ . In order to efficiently solve the optimization problem presented in (32), iterative techniques can be conducted such that the learning model parameters are computed at each iteration based on local datasets. Specifically, first the learning model parameters are computed at the users, and then they are transmitted to the server, at which they are aggregated.

Prior to the transmission of the learning models to the BS,  $E$  local updates are performed [39]. In particular, define  $\boldsymbol{\theta}_t^{(e)} \in \mathbb{R}^Z$  and  $\boldsymbol{\beta}_k(\boldsymbol{\theta}_t^{(e)}) = \nabla \mathcal{L}_k(\boldsymbol{\theta}_t^{(e)})$  as the model parameter and gradient vectors computed at the  $e$ -th local update, respectively. Then, after performing  $E$  local updates, we denote the  $E$ -th computed gradient vector by  $\boldsymbol{\beta}_k(\boldsymbol{\theta}_t) = \boldsymbol{\beta}_k(\boldsymbol{\theta}_t^{(E)})$  for notational convenience. Next, the

TABLE 1. Computational complexity of Algorithm 1.

Step	Complexity Order
4	$O(QMNP(NP + P))$
10	$O(PNL(L + 1)/2)$
11	$O(P^2L(L + 1)/2 + P^2)$
16	$O(L(P + N))$

model parameter updates  $\beta_k(\theta_t)$  are transmitted to the BS, at which they are received in a synchronous fashion. While there are different user scheduling mechanisms, e.g., random, round robin and proportional fair [40], [41], we assume that the model updates  $\{\beta_k(\theta_t)\}_{k \in \mathcal{K}}$  are aligned through a synchronization procedure and they are received in the same state. Thus, the model update rule for the  $t$ -th iteration is described by

$$\theta_{t+1} = \theta_t - \varepsilon \frac{1}{K} \sum_{k=1}^K \beta_k(\theta_t), \quad (34)$$

for  $t = 1, \dots, T$  and  $\varepsilon$  is the learning rate.

During model training, each user simply transmits their model update vector  $\beta_k(\theta_t)$  to the BS for  $k \in \mathcal{K}$ . Assume that the model  $\beta_k(\theta_t)$  is divided into  $U = \frac{Z}{M}$  equal-length blocks for each subcarrier as

$$\beta_k(\theta_t) = [\mathbf{x}_K^{(t)\top} [1], \dots, \mathbf{x}_K^{(t)\top} [M]]^\top, \quad (35)$$

where  $\mathbf{x}_k^{(t)} [m] \in \mathbb{C}^U$  denotes data symbols (model updates) of the  $k$ -th user, that are transmitted to the server to be aggregated. The BS, then, receives the following  $N \times U$  uplink signal as

$$\mathbf{Y}^{(t)} [m] = \sum_{k=1}^K \mathbf{h}_k^{(t)} [m] \mathbf{x}_k^{(t)\top} [m] + \mathbf{N} [m], \quad (36)$$

where  $\mathbf{h}_k^{(t)} [m] \in \mathbb{C}^N$  represents the uplink channel during model transmission and  $\mathbf{N} [m] \in \mathbb{C}^{N \times U}$  denotes the noise term added onto the transmitted data.

## VI. COMPLEXITY AND OVERHEAD

In this section, we analyze the computational complexity of the proposed NBA-OMP approach, channel training overhead as well as the communication-overhead of the FL-based model training.

### A. COMPLEXITY OF NBA-OMP

The computational complexity of NBA-OMP is the same as traditional OMP techniques [6], and it is mainly due to the matrix multiplications in step 4 ( $O(QMNP(NP + P))$ ), step 10 ( $O(PNL\bar{L})$ ), step 11 ( $O(P^2L\bar{L} + P^2)$ ) and step 16 ( $O(L(P + N))$ ), where  $\bar{L} = (L + 1)/2$ . Hence, the overall complexity is  $O(QMNP(NP + P) + (PL\bar{L} + 1)(N + P))$ . Table 1 summarizes the complexity orders of Algorithm 1.

### B. LEARNING MODEL COMPLEXITY

We examine the computational complexity of the proposed CNN model comprised of convolutional and fully connected layers given in Table 2. The time complexity of the convolutional layers (i.e.,  $l = 2, 4, 6$ ) is

$$\mathbf{C}_{\text{CL}} = O \left( \sum_{l \in \{2,4,6\}} D_x^{(l)} D_y^{(l)} W_x^{(l)} W_y^{(l)} N_F^{(l-1)} N_F^{(l)} \right), \quad (37)$$

where  $D_x^{(l)}, D_y^{(l)}$  are the column and row sizes of each output feature map of the convolutional layer,  $W_x^{(l)}, W_y^{(l)}$  are the 2-D filter size of the  $l$ -th layer.  $N_F^{(l-1)}$  and  $N_F^{(l)}$  represent the number of input and output feature maps for the  $l$ -th layer, respectively. According to the parameters in Table 2, we have  $D_x^{(l)} = 1, D_y^{(l)} = N$  and  $W_x^{(l)} = W_y^{(l)} = 3$  for  $l = 2, 4, 6$ . Furthermore, the tuple  $(N_F^{(l-1)}, N_F^{(l)})$  equals to  $(3, 128), (128, 128), (128, 128)$  for  $l = 2, 4, 6$ , respectively. As a result, the computational complexity order of the convolutional layers in (37) is given approximately as

$$\begin{aligned} \mathbf{C}_{\text{CL}} &= O \left( N \cdot 3^2 \cdot (3 \cdot 128 + 128^2 + 128^2) \right) \\ &\approx O(N \cdot 3^2 \cdot 2^{15}). \end{aligned} \quad (38)$$

Next, the complexity order for the fully connected layers (i.e.,  $l = 8, 10$ ) is given by

$$\mathbf{C}_{\text{FCL}} = O \left( \sum_{l \in \{8,10\}} D_x^{(l)} D_y^{(l)} \rho N_{\text{FCL}}^{(l)} \right), \quad (39)$$

where  $(D_x^{(l)}, D_y^{(l)})$  equals to  $(128N, 1)$  and  $(1024, 1)$  for  $l = 8, 10$ , respectively. Furthermore, we have  $N_{\text{FCL}}^{(l)} = 1024$  for  $l = 8, 10$  and  $\rho = 0.5$  is the dropout ratio. Then, (39) becomes

$$\begin{aligned} \mathbf{C}_{\text{FCL}} &= O(128 \cdot N \cdot 1024 \cdot 0.5 + 1024 \cdot 1024 \cdot 0.5) \\ &= O(2^{16}(N + 2^3)) \\ &= O(2^{15}(2N + 2^4)). \end{aligned} \quad (40)$$

Finally, combining (38) and (40) yields the total time complexity of the CNN model approximately as

$$\begin{aligned} \mathbf{C} &= \mathbf{C}_{\text{CL}} + \mathbf{C}_{\text{FCL}} \\ &\approx O(2^{15}(3^2N + 2N + 2^4)). \end{aligned} \quad (41)$$

### C. CHANNEL TRAINING OVERHEAD

During channel training, the BS with  $K$  RF chains transmits  $P$  pilot signals, which are received simultaneously by the signal antenna users with a single RF chain. Therefore, the channel training stage requires  $P$  channel usage. Since the proposed approach relies on the sparsity of the THz channel to reconstruct the channel vector  $\mathbf{h}_k [m]$ , the NBA-OMP approach does not require  $P \geq N$ . Instead, much fewer channel pilots can be used to reconstruct the channel accurately. Specifically, the channel training overhead of the NBA-OMP requires  $P \ll N$  (8 times lower, see Sec. VII)

**TABLE 2.** Learning model parameters.

Layer	Type	Parameters
1	Input	Size: $N_{\text{RF}} \times 3$
2	Convolutional	Size: $3 \times 3 @ 128$
3	Normalization	-
4	Convolutional	Size: $3 \times 3 @ 128$
5	Normalization	-
6	Convolutional	Size: $3 \times 3 @ 128$
7	Normalization	-
8	Fully Connected	Size: $1024 \times 1$
9	Dropout	Ratio: 0.5
10	Fully Connected	Size: $1024 \times 1$
11	Dropout	Ratio: 0.5
12	Output (Regression)	Size: $2N \times 1$

channel usage for pilot signaling while the traditional approaches, e.g., LS and MMSE estimation, need at least  $N$  channel usage.

#### D. MODEL TRAINING OVERHEAD

The communication overhead of a learning-based approach can be measured in terms of amount of data symbols transmitted during/for model training [17], [37], [39], [42], [43]. Thus, the communication overhead (i.e.,  $\mathcal{T}_{\text{CL}}$ ) for the CL involves the amount of dataset transmitted from the users to the server whereas it can be computed for FL (i.e.,  $\mathcal{T}_{\text{FL}}$ ) as the amount of model parameters exchanged between the users and the server during training. Then, we have

$$\mathcal{T}_{\text{CL}} = \sum_{k=1}^K \mathbf{D}_k (3N_{\text{RF}} + \xi 2N), \quad (42)$$

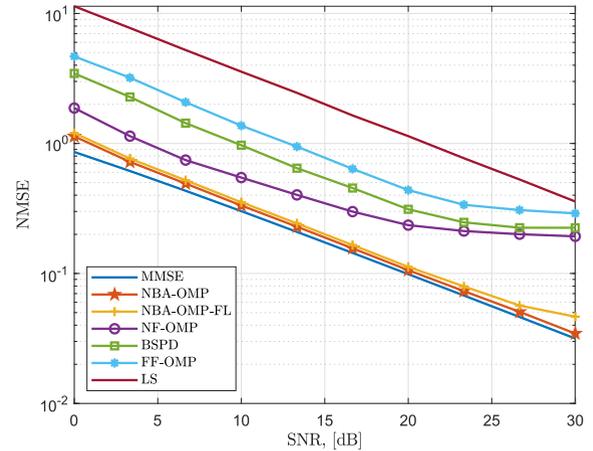
where  $\mathbf{D}_k$  is the number of samples of the  $k$ -th dataset and  $3N_{\text{RF}} + 2N$  is the number of symbols in each input-output (i.e.,  $\mathcal{X}_k^{(i)} - \mathcal{Y}_k^{(i)}$ ) tuple. In (42),  $\xi$  denotes whether the labels are also transmitted from the users to the server. Thus, if the labels are computed at the users, they are transmitted to the server with additional overhead, and we have  $\xi = 1$ . On the other hand,  $\xi = 0$  (labels are not computed at the user) if the received inputs are transmitted to the server, at which the labeling (channel estimation) is handled, thereby less communication overhead is achieved. Next, we defined the communication overhead for FL as

$$\mathcal{T}_{\text{FL}} = 2ZTK, \quad (43)$$

which involves a two-way (user  $\rightleftarrows$  server) transmission of the learnable parameters from  $K$  users to the server for  $T$  consecutive iterations.

#### VII. NUMERICAL SIMULATIONS

We evaluate the performance of our proposed channel estimation approaches, in comparison with the state-of-the-art channel estimation techniques, e.g., far-field OMP (FF-OMP) [44], near-field OMP (NF-OMP) [25], beam-squint pattern detection (BSPD) [16] as well as LS and MMSE. Note that, during the simulations, the MMSE estimator is

**FIGURE 4.** Near-field THz wideband channel estimation NMSE versus SNR.  $N = 256$ ,  $f_c = 300$  GHz,  $M = 128$  and  $B = 30$  GHz.

computed subcarrier-wise such that it has a beam-squint-free benchmark performance.

Throughout the simulations, unless stated otherwise, the signal model is constructed with  $f_c = 300$  GHz,  $B = 30$  GHz,  $M = 128$ ,  $K = N_{\text{RF}} = 8$ ,  $L = 3$ ,  $P = 8$  and  $N = 256$ , for which the Fraunhofer distance is  $F = 32.76$  m. The NBA dictionary matrix is constructed with  $Q = 10N$ , and the user directions and ranges are selected as  $\phi_{k,l} \in \text{unif}[-\frac{\pi}{2}, \frac{\pi}{2}]$ ,  $r_{k,l} \in \text{unif}[5, 30]$  m, respectively.

A convolutional neural network (CNN) is designed with parameters enlisted in Table 2 through a hyperparameter optimization [39], [43]. The learning model is, then, trained for  $E = 10$  local updates and  $T = 100$  iterations, which involves the exchange of model parameter updates between the users and the server. The learning rate is selected as  $\varepsilon = 0.001$ . For dataset generation, we first generated  $V = 1000$  independent channel realizations. Then, synthetic noise with AWGN is added onto the generated channel data in order to resemble the characteristics related to the imperfections/distortions in the wireless channels. This synthetic noise level is obtained over pre-determined noise level for three signal-to-noise ratio (SNR) levels, i.e.,  $\text{SNR}_{\text{TRAIN}} = \{15, 20, 25\}$  dB such that  $G = 1000$  realizations of each channel data are obtained in order to improve robustness [37]. During channel generation, we considered the local datasets with non-identical distributions such that the  $k$ -th user's dataset is selected with the DoA information as  $\vartheta_{k,l} \in [-\frac{\pi}{2} + \frac{\pi}{K}(k-1), -\frac{\pi}{2} + \frac{\pi}{K}k]$ .

Fig. 4 shows the channel estimation normalized MSE (NMSE) performance under various SNR levels. We can see that the proposed NBA-OMP and NBA-OMP-FL techniques outperform the competing methods and closely follows the MMSE performance. The superior performance of NBA-OMP can be attributed to accurately compensating the effect of beam-squint over both direction and range parameters via NBA dictionary, which is composed of SD steering vectors. On the other hand, the remaining methods fail

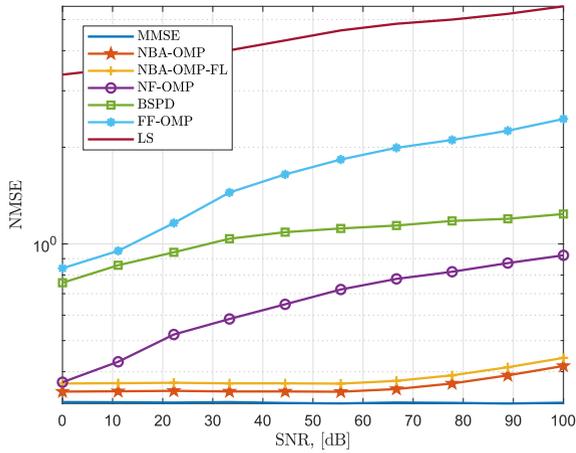
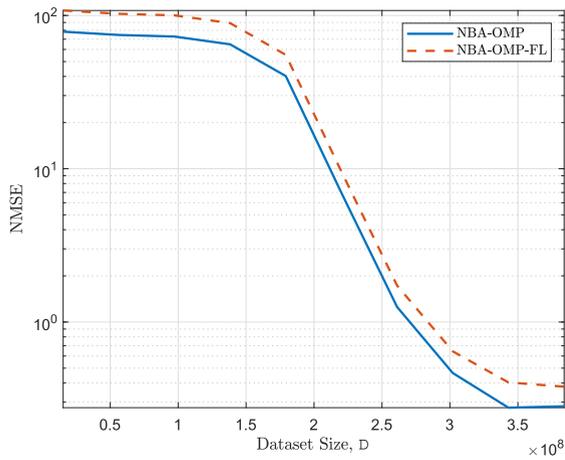
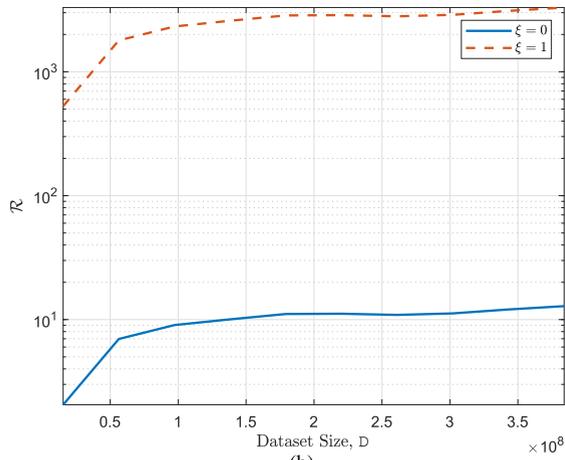


FIGURE 5. Near-field THz wideband channel estimation NMSE versus bandwidth when SNR = 10 dB.



(a)



(b)

FIGURE 6. Channel estimation NMSE (a) and communication-efficiency ratio (b) when SNR = 10 dB.

to exhibit such high precision in high SNR regime. This is because these methods either fail to take into account the NB [25] or only consider the far-field signal model [8], [16]. Furthermore, the proposed approach does not

require an additional hardware components in order to realize SD dictionary matrices steering vectors, hence it is hardware-efficient. Compared to NBA-OMP, the FL-based approach has slight performance loss due to model training with unevenly distributed datasets. As a result, the performance of NBA-OMP behaves like a yardstick for the FL approach since the learning-based methods cannot perform better than their labels [38].

Fig. 5 compares the NMSE performance with respect to the bandwidth  $B \in [0, 100]$  GHz. We can see that the proposed NBA-OMP approaches effectively compensate the impact of beam-squint for a large portion of the bandwidth up to  $B < 70$  GHz.

Finally, we analyze the communication overhead of the proposed FL approach. To this end, we first compute the overhead for CL as  $\mathcal{T}_{CL} = 8 \cdot \mathbf{D}_k \cdot (3 \cdot 8) = 24.57 \times 10^9$  when  $\xi = 0$  (note that  $\mathcal{T}_{CL} = 6.31 \times 10^{12}$  when  $\xi = 1$ ), where the number of input-output tuples in the dataset is  $\mathbf{D}_k = 3MVG = 3 \cdot 128 \cdot 1000 \cdot 1000 = 384 \times 10^6$ . On the other hand, the communication overhead for FL is  $\mathcal{T}_{FL} = 2 \cdot 1,196,928 \cdot 100 \cdot 8 = 1.91 \times 10^9$ , which exhibits approximately 12 (3300) times lower than that of CL when  $\xi = 0$  ( $\xi = 1$ ). This observation shows the effectiveness of the FL approach for channel estimation tasks, especially when the number of antennas is very high. It is also worthwhile mentioning that the gain achieved by employing FL is huge if the labeling is handled at the server, i.e.,  $\xi = 0$ . Based on this analysis, Fig. 6 demonstrates the channel estimation NMSE and communication-efficiency ratio, i.e.,  $\mathcal{R} = \frac{\mathcal{T}_{CL}}{\mathcal{T}_{FL}}$  with respect to the dataset size  $\mathbf{D} = \sum_{k=1}^K \mathbf{D}_k$  when SNR = 10 dB. We can see that smaller datasets can be used for training the learning model for faster convergence at the cost of low NMSE performance. A satisfactory channel estimation performance can be achieved (i.e.,  $NMSE \approx 10^{-1}$ ) if  $\mathbf{D} \approx 3 \times 10^8$ . As expected, larger datasets result in longer training times as we observed from the simulations that the number of iterations is approximately  $T = 100$  for such settings. In this scenario, the communication-efficiency ratio  $\mathcal{R}$  is about 12 and 3300 when  $\xi = 0$  and  $\xi = 1$ , respectively.

VIII. CONCLUSION

We considered the near-field wideband THz channel estimation problem in the presence of near-field beam-squint. We introduced both model-based and model-free approaches to effectively estimate the channel with lower complexity and overhead. The model-based approach is based on the OMP technique, for which an NBA dictionary is designed such that the SD DoA and range parameters are accurately matched, hence, the beam-squint-corrected channel estimate is obtained. The model-free approach is based on FL scheme such that the data labels are obtained from the model-based approach. The proposed approach has lower channel and model training overhead as compared to existing techniques. Specifically, the proposed model-based approach achieves close-to-MMSE performance with 8 times less channel usage than the conventional techniques whereas the proposed

model-free technique enjoys 12 times less communication overhead as compared to the centralized schemes. The proposed NBA-OMP is particularly useful for THz-band applications, wherein ultra-massive number of antennas are used leading to high computational complexity, and the near-field beam-squint may cause degradations in the system performance.

## REFERENCES

- [1] T. S. Rappaport, Y. Xing, O. Kanhere, S. Ju, A. Madanayake, S. Mandal, A. Alkhateeb, and G. C. Trichopoulos, "Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond," *IEEE Access*, vol. 7, pp. 78729–78757, 2019.
- [2] I. F. Akyildiz, C. Han, Z. Hu, S. Nie, and J. M. Jornet, "Terahertz band communication: An old problem revisited and research directions for the next decade," *IEEE Trans. Commun.*, vol. 70, no. 6, pp. 4250–4285, Jun. 2022.
- [3] X. You, "Toward 6G TKμ extreme connectivity: Architecture, key technologies and experiments," Aug. 2022, *arXiv:2208.01190*.
- [4] A. M. Elbir, K. V. Mishra, S. Chatzinotas, and M. Bennis, "Terahertz-band integrated sensing and communications: Challenges and opportunities," 2022, *arXiv:2208.01235*.
- [5] H. Sarrieddeen, M.-S. Alouini, and T. Y. Al-Naffouri, "An overview of signal processing techniques for terahertz communications," *Proc. IEEE*, vol. 109, no. 10, pp. 1628–1665, Oct. 2021.
- [6] R. W. Heath Jr., N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Feb. 2016.
- [7] A. M. Elbir, K. V. Mishra, S. A. Vorobyov, and R. W. Heath Jr., "Twenty-five years of advances in beamforming: From convex and nonconvex optimization to learning techniques," 2022, *arXiv:2211.02165*.
- [8] A. Alkhateeb and R. W. Heath Jr., "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1801–1818, May 2016.
- [9] X. Yu, J.-C. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 485–500, May 2016.
- [10] J. Tan and L. Dai, "Wideband beam tracking in THz massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1693–1710, Jun. 2021.
- [11] A. Alkhateeb, G. Leus, and R. W. Heath Jr., "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6481–6494, Nov. 2015.
- [12] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Mar. 2013.
- [13] Z. Marzi, D. Ramasamy, and U. Madhoo, "Compressive channel estimation and tracking for large arrays in mm-wave picocells," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 514–527, Apr. 2016.
- [14] J. Wang, Z. Lan, C. W. Pyo, T. Baykas, C. S. Sum, M. A. Rahman, J. Gao, R. Funada, F. Kojima, H. Harada, and S. Kato, "Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 8, pp. 1390–1399, Oct. 2009.
- [15] K. Dovelos, M. Matthaiou, H. Q. Ngo, and B. Bellalta, "Channel estimation and hybrid combining for wideband terahertz massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1604–1620, Jun. 2021.
- [16] J. Tan and L. Dai, "Wideband channel estimation for THz massive MIMO," *China Commun.*, vol. 18, no. 5, pp. 66–80, May 2021.
- [17] A. M. Elbir, W. Shi, K. V. Mishra, and S. Chatzinotas, "Federated multi-task learning for THz wideband channel and DoA estimation," 2022, *arXiv:2207.06017*.
- [18] A. M. Elbir, "A unified approach for beam-split mitigation in terahertz wideband hybrid beamforming," *IEEE Trans. Veh. Technol.*, early access, Apr. 7, 2023, doi: [10.1109/TVT.2023.3265395](https://doi.org/10.1109/TVT.2023.3265395).
- [19] B. Wang, F. Gao, S. Jin, H. Lin, and G. Y. Li, "Spatial- and frequency-wideband effects in millimeter-wave massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 66, no. 13, pp. 3393–3406, May 2018.
- [20] S. Nie and I. F. Akyildiz, "Deep kernel learning-based channel estimation in ultra-massive MIMO communications at 0.06–10 THz," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2019, pp. 1–6.
- [21] Y. Chen and C. Han, "Deep CNN-based spherical-wave channel estimation for terahertz ultra-massive MIMO systems," in *Proc. GLOBECOM IEEE Global Commun. Conf.*, Dec. 2020, pp. 1–6.
- [22] E. Balevi and J. G. Andrews, "Wideband channel estimation with a generative adversarial network," *IEEE Trans. Wireless Commun.*, vol. 20, no. 5, pp. 3049–3060, May 2021.
- [23] A. M. Elbir, W. Shi, A. K. Papazafeiropoulos, P. Kourtessis, and S. Chatzinotas, "Terahertz-band channel and beam split estimation via array perturbation model," *IEEE Open J. Commun. Soc.*, early access, Mar. 31, 2023, doi: [10.1109/OJCOMS.2023.3263625](https://doi.org/10.1109/OJCOMS.2023.3263625).
- [24] E. Björnson, Ö. T. Demir, and L. Sanguinetti, "A primer on near-field beamforming for arrays and reconfigurable intelligent surfaces," 2021, *arXiv:2110.06661*.
- [25] X. Wei and L. Dai, "Channel estimation for extremely large-scale massive MIMO: Far-field, near-field, or hybrid-field?" *IEEE Commun. Lett.*, vol. 26, no. 1, pp. 177–181, Jan. 2022.
- [26] 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, Apr. 2022.
- [27] W. Yu, Y. Shen, H. He, X. Yu, J. Zhang, and K. B. Letaief, "Hybrid far- and near-field channel estimation for THz ultra-massive MIMO via fixed point networks," 2022, *arXiv:2205.04944*.
- [28] X. Zhang, H. Zhang, and Y. C. Eldar, "Near-field sparse channel representation and estimation in 6G wireless communications," 2022, *arXiv:2212.13527*.
- [29] X. Zhang, Z. Wang, H. Zhang, and L. Yang, "Near-field channel estimation for extremely large-scale array communications: A model-based deep learning approach," 2022, *arXiv:2211.15440*.
- [30] A. M. Elbir, K. V. Mishra, and S. Chatzinotas, "Terahertz-band joint ultra-massive MIMO radar-communications: Model-based and model-free hybrid beamforming," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 6, pp. 1468–1483, Nov. 2021.
- [31] L. Dai, J. Tan, Z. Chen, and H. V. Poor, "Delay-phase precoding for wideband THz massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7271–7286, Sep. 2022.
- [32] S. Tarboush, H. Sarrieddeen, H. Chen, M. H. Loukil, H. Jemaa, M. S. Alouini, and T. Y. Al-Naffouri, "TeraMIMO: A channel simulator for wideband ultra-massive MIMO terahertz communications," 2021, *arXiv:2104.11054*.
- [33] H. Yuan, N. Yang, K. Yang, C. Han, and J. An, "Hybrid beamforming for terahertz multi-carrier systems over frequency selective fading," *IEEE Trans. Commun.*, vol. 68, no. 10, pp. 6186–6199, Oct. 2020.
- [34] H. Yuan, N. Yang, X. Ding, C. Han, K. Yang, and J. An, "Cluster-based multi-carrier hybrid beamforming for massive device terahertz communications," *IEEE Trans. Commun.*, vol. 70, no. 5, pp. 3407–3420, May 2022.
- [35] M. Cui, L. Dai, Z. Wang, S. Zhou, and N. Ge, "Near-field rainbow: Wideband beam training for XL-MIMO," *IEEE Trans. Wireless Commun.*, early access, Nov. 21, 2022, doi: [10.1109/TWC.2022.3222198](https://doi.org/10.1109/TWC.2022.3222198).
- [36] A. M. Elbir and T. E. Tuncer, "Far-field DOA estimation and near-field localization for multipath signals," *Radio Sci.*, vol. 49, pp. 765–776, Sep. 2014.
- [37] A. M. Elbir and S. Coleri, "Federated learning for channel estimation in conventional and RIS-assisted massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 4255–4268, Jun. 2022.
- [38] A. M. Elbir, A. K. Papazafeiropoulos, and S. Chatzinotas, "Federated learning for physical layer design," *IEEE Commun. Mag.*, vol. 59, no. 11, pp. 81–87, Nov. 2021.
- [39] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Y. Arcas, "Communication-efficient learning of deep networks from decentralized data," 2016, *arXiv:1602.05629*.
- [40] H. H. Yang, Z. Liu, T. Q. S. Quek, and H. V. Poor, "Scheduling policies for federated learning in wireless networks," *IEEE Trans. Commun.*, vol. 68, no. 1, pp. 317–333, Jan. 2020.
- [41] A. M. Elbir and S. Coleri, "Federated learning for hybrid beamforming in mm-wave massive MIMO," *IEEE Commun. Lett.*, vol. 24, no. 12, pp. 2795–2799, Dec. 2020.
- [42] M. M. Amiri and D. Gündüz, "Federated learning over wireless fading channels," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3546–3557, May 2020.

- [43] M. M. Amiri and D. Gündüz, "Machine learning at the wireless edge: Distributed stochastic gradient descent over-the-air," *IEEE Trans. Signal Process.*, vol. 68, pp. 2155–2169, 2020.
- [44] J. Rodríguez-Fernández, N. González-Prelcic, K. Venugopal, and R. W. Heath Jr., "Frequency-domain compressive channel estimation for frequency-selective hybrid millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2946–2960, May 2018.



**AHMET M. ELBIR** (Senior Member, IEEE) received the B.S. degree (Hons.) in electrical engineering from Firat University, in 2009, and the Ph.D. degree in electrical engineering from Middle East Technical University (METU), Ankara, Turkey, in 2016. He was a Visiting Postdoctoral Researcher with Koc University, Istanbul, Turkey, from 2020 to 2021, and a Senior Researcher with Duzce University, Turkey, from 2016 to 2022. He is currently a Research Fellow with the University of Luxembourg, Luxembourg; Carleton University, Canada; and the University of Hertfordshire, U.K. His research interests include array signal processing, sparsity-driven convex optimization, signal processing for communications, and deep learning for array signal processing. He was a recipient of the 2016 METU Best Ph.D. Thesis Award for his Ph.D. studies and the IET Radar, Sonar and Navigation Best Paper Award, in 2022. He has been an Associate Editor of IEEE ACCESS, since 2018, and a Lead Guest Editor of the IEEE JOURNAL OF SELECTED TOPICS ON SIGNAL PROCESSING and IEEE WIRELESS COMMUNICATIONS.



**WEI SHI** (Member, IEEE) received the bachelor's degree in computer engineering from the Harbin Institute of Technology (HIT) and the master's and Ph.D. degrees in computer science from Carleton University, Ottawa, ON, Canada. She is currently an Associate Professor with the School of Information Technology, cross appointed to the Department of Systems and Computer Engineering, Faculty of Engineering and Design, Carleton University. She is specialized in the design and analysis of fault tolerance algorithms addressing security issues in distributed environments, such as data-center networks, clouds, mobile agents and actuator systems, wireless sensor networks, and critical infrastructures, such as power grids and smart cities. She has also been conducting research in data privacy and big data analytics. She is a Professional Engineer licensed in Canada.



**ANASTASIOS K. PAPAZAFEIROPOULOS** (Senior Member, IEEE) received the B.Sc. degree (Hons.) in physics, the M.Sc. degree (Hons.) in electronics and computers science, and the Ph.D. degree from the University of Patras, Greece, in 2003, 2005, and 2010, respectively. From 2011 to 2012 and from 2016 to 2017, he was a Postdoctoral Research Fellow with the Institute for Digital Communications, The University of Edinburgh, Edinburgh, U.K. From 2012 to 2014, he was a Research Fellow with Imperial College London, London, U.K. He is currently a Vice-Chancellor Fellow with the University of Hertfordshire, Hertfordshire, U.K. He is also a Visiting Research Fellow with the SnT University of Luxembourg, Luxembourg. He has been involved in several EPSRC and EU FP7 projects, including HIATUS and HARP. His research interests include machine learning for wireless communications, intelligent reflecting surfaces, massive MIMO, heterogeneous networks, 5G and beyond wireless networks, full-duplex radio, mm-wave communications, random matrix theory, hardware-constrained communications, and the performance analysis of fading channels. He was a recipient of the Marie Curie Fellowship (IEF-IAWICOM).



**PANDELIS KOURTEASSIS** (Member, IEEE) is currently the Director of the Centre for Engineering Research and a Reader in communication networks with the University of Hertfordshire, U.K., leading the activities of the Networks Engineering Research Group into Communications and Information Engineering, including next generation passive optical networks, optical and wireless MAC protocols, 5G RANs, software-defined networks and networks virtualization 5G and satellite networks and more recently machine learning for next generation networks. His funding ID includes EU COST, FP7, H2020, European Space Agency (ESA), UKRI, and industrially funded projects. He has published more than 80 papers with peer-reviewed journals, peer-reviewed conference proceedings, and international conferences. His research has received coverage at scientific journals, magazines, white papers, and international workshops. He has served as the general chair, the co-chair, and a technical program committee member with the scientific committees and expert groups for IEEE workshops and conferences, European technology platforms, and European networks of excellence. He has been the co-editor of a Springer book and a chapter editor of an IET book on softwarization for 5G.



**SYMEON CHATZINOTAS** (Fellow, IEEE) received the M.Eng. degree in telecommunications from the Aristotle University of Thessaloniki, Thessaloniki, Greece, in 2003, and the M.Sc. and Ph.D. degrees in electronic engineering from the University of Surrey, Guildford, U.K., in 2006 and 2009, respectively. He is currently a Full Professor, the Deputy Head of the SIGCOM Research Group, Interdisciplinary Centre for Security, Reliability, and Trust, University of Luxembourg, Esch-sur-Alzette, Luxembourg, and a Visiting Professor with the University of Parma, Parma, Italy. He has been involved in numerous research and development projects with the Institute of Informatics Telecommunications, National Center for Scientific Research "Demokritos," Institute of Telematics and Informatics, Center of Research and Technology Hellas, and the Mobile Communications Research Group, Center of Communication Systems Research, University of Surrey. He has coauthored more than 400 technical papers in refereed international journals, conferences, and scientific books. His research interests include multiuser information theory, cooperative/cognitive communications, and wireless network optimization. He was a co-recipient of the 2014 IEEE Distinguished Contributions to Satellite Communications Award, the CROWNCOM 2015 Best Paper Award, and the 2018 *EURASIP Journal on Wireless Communications and Networking* Best Paper Award. He is on the editorial board of the IEEE OPEN JOURNAL OF VEHICULAR TECHNOLOGY and the *International Journal of Satellite Communications and Networking*.

...