Parametric Near-Field MMSE Channel Estimation for sub-THz XL-MIMO Systems

Wen-Xuan Long[1], Marco Moretti[1,2], Michele Morelli[1], Luca Sanguinetti[1,2], Rui Chen[3]
[1]Dipartimento di Ingegneria dell'Informazione, University of Pisa, Italy
[2]National Inter-University Consortium for Telecommunications (CNIT), Parma, Italy
[3]State Key Laboratory of Integrated Service Networks, Xidian University, China
Email: wenxuan.long@ing.unipi.it, marco.moretti@unipi.it, michele.morelli@unipi.it, luca.sanguinetti@unipi.it.

***Abstract*—Accurate channel estimation is essential for reliable communication in sub-THz extremely large (XL) MIMO systems. Deploying XL-MIMO in high-frequency bands not only increases the number of antennas, but also fundamentally alters channel propagation characteristics, placing the user equipments (UE) in the radiative near-field of the base station. This paper proposes a parametric estimation method using the multiple signal classification algorithm to extract UE location data from uplink pilot signals. These parameters are then used to reconstruct the spatial correlation matrix, followed by an approximation of the minimum mean square error channel estimator. Numerical results show that the proposed method outperforms the least-squares estimator in terms of the normalized mean-square error, even without prior UE location knowledge.**

***Index Terms*—Terahertz wireless communications, short-range communications, near-field channel estimation, spatially correlated channels, multiple signal classification, minimum mean square error.**

I. Introduction

Terahertz (THz) wireless communication is gaining increased interest in various applications, including space communications, biomedical sensing, and industrial automation, due to its potential advantages in terms of high bandwidth and low latency [1]. Several studies have investigated the transmission characteristics of wireless channels in the THz and sub-THz frequency bands [2]. In [3] and [4], the authors measured and modeled the path loss in sub-THz bands, revealing that as the frequency increases, sub-THz channels experience greater propagation loss, leading to rapid signal power attenuation. Additionally, [5] measured the power delay profile in the sub-THz band, showing that most line-of-sight (LoS) measurements lack significant multipath components, demonstrating the high directionality and narrow beamwidth of THz channels. Based on these findings, the THz wireless transmission is more likely to enable the short-range communication [6] by leveraging its strong LoS characteristics.

Considering the further expansion of array apertures in 6G base stations (BS), the user equipments (UEs) are likely to fall within the near-field (NF) region of the BS in THz short-range transmissions. In this region, the electromagnetic wavefront exhibits spherical curvature, introducing spherical phase variations across the array elements. These variations are determined by both the angle and distance between the array and the point source, fundamentally altering the channel propagation characteristics. Consequently, conventional far-field (FF) channel estimation methods [7], [8] become inadequate, necessitating the development of advanced NF channel estimation algorithms that fully exploit the unique properties of NF propagation.

A potential approach for the NF channel estimation is to first recover the location parameters of the UE, and then incorporate them into the parametric channel model to derive the channel estimate. In [9] and [10], a two-stage multiple signal classification (MUSIC) algorithm is proposed for uniform linear array (ULA)-based systems. In this approach, the direction of arrival (DoA) of the UE is first estimated, followed by the estimation of the distance between the UE and the BS using the acquired angular information. Building upon this, [11] extends the two-stage MUSIC algorithm to the NF channel estimation based on uniform planar arrays (UPAs). The estimated location parameters are then integrated into the NF channel model to derive the channel estimates. However, the aforementioned studies primarily focus on the parameter estimation for NF channels in low-frequency bands, while the applicability of parametric channel estimation approaches in THz-band short-range communications is still unexplored.

In this paper, we propose a parametric channel estimation method for the sub-THz NF channel based on the MUSIC algorithm. We assume that multiple UEs periodically transmit orthogonal training sequences so as to avoid multi-user interference at the BS. Furthermore, the signal transmitted by each UE arrives at the BS through a line-of-sight (LoS) path with a specified spatial spread. During channel estimation, the MUSIC algorithm is first applied to obtain key parameters specifying the location of the UE relative to the BS. These parameters are then used to reconstruct the spatial correlation matrix of the NF channel. In doing so, we adopt mismatched values of the angular and distance spreads, which are designed large enough to capture (almost surely) all the received UE signal energy. The reconstructed channel correlation matrix is eventually exploited to implement an approximation of the minimum mean-square error (MMSE) estimator of the UE uplink channel. Numerical simulations in the sub-THz band are conducted to evaluate the impact of transmission power, selected spatial spreads, and other parameters on the proposed solution. The results demonstrate that the proposed method significantly outperforms the conventional least-squares (LS) estimator and other existing methods in the sub-THz band.

II. System Model

We consider a communication system operating over a bandwith B at sub-THz frequencies (e.g., in the range of 0.1 THz) in which K single-antenna UEs communicate with a BS equipped with an UPA, e.g., [12, Fig. 1]. The array is placed in the $yoz-$plane of a three-dimensional space, where a spherical coordinate system is defined, with φ being the

azimuth angle, θ the elevation angle and r the distance. The array has N_H elements in each row and N_V elements in each column, resulting in a total of $N = N_H \times N_V$ elements. The horizontal and vertical inter-element spacing is δ, while $\mathbf{v}_n$ is the location of the n-th array element.

The Fraunhofer distance of the array is computed as $d_F = 2D^2/\lambda$ [13], where $D = \sqrt{(N_V^2 + N_H^2)}\delta$ is the array aperture length and λ is the wavelength. This distance is used to distinguish between the FF and radiative NF regions of the array [13]. When the array aperture becomes large, reaching tens or even hundreds of times the wavelength, the expansion of the Fraunhofer distance makes it highly probable that the UEs would be located within the relative NF region of the array. This is typically the case of sub-THz MIMO systems for short-range communications [6]. As an example, consider a half-wavelength-spaced array with size 0.15×0.075 m^2, and operating at 0.1 THz (i.e., $\lambda = 0.003$ m). This array contains 5000 antennas in the configuration 100×50, so that we have $d_F = 18.75$ m. In short-range communications (e.g., on the order of tens of meters), the UEs will likely be located below d_F. This implies that the FF approximation cannot be used, and the exact propagation model for the channel must be considered instead. This model is introduced next.

A. Channel Model

Based on [3]–[5], a reasonable assumption in sub-THz short-range communications is that the signal transmitted by each UE arrives at the BS array within a small solid angle centered around the LoS path. We concentrate on a given UE and denote by (φ, θ) the LoS angles from the UE to the array. Assuming a conventional correlated Rayleigh fading channel, the channel vector $\mathbf{h} \in \mathbb{C}^N$ can be modeled as [14]

$$\mathbf{h} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}_N, \mathbf{R}), \tag{1}$$

which is fully characterized by the spatial correlation matrix $\mathbf{R}$. The latter is given by $\mathbf{R} = \beta\mathbf{A}$, where $\beta = \frac{1}{N}\text{tr}\{\mathbf{R}\}$ is the average channel power (capturing pathloss and shadowing), and $\mathbf{A}$ is expressed as

$$\mathbf{A} = \int_{r-\triangle_r}^{r+\triangle_r}\int_{\varphi-\triangle_\varphi}^{\varphi+\triangle_\varphi}\int_{\theta-\triangle_\theta}^{\theta+\triangle_\theta} f(\tilde{r},\tilde{\varphi},\tilde{\theta}) \mathbf{a}(\tilde{r},\tilde{\varphi},\tilde{\theta})\mathbf{a}^{\mathrm{H}}(\tilde{r},\tilde{\varphi},\tilde{\theta}) d\tilde{\theta}d\tilde{\varphi}d\tilde{r}, \tag{2}$$

where (r, φ, θ) are the distance, the azimuth and elevation angles of the considered UE, while the triplet $(\triangle_r, \triangle_\varphi, \triangle_\theta)$ accounts for the corresponding distance and angular spreads. Also, $\mathbf{a}(\tilde{r}, \tilde{\varphi}, \tilde{\theta})$ is the array response vector [15]:

$$\mathbf{a}(\tilde{r},\tilde{\varphi},\tilde{\theta}) = \left[1, e^{j\frac{2\pi}{\lambda}(\tilde{r}_1-\tilde{r})}, \ldots, e^{j\frac{2\pi}{\lambda}(\tilde{r}_N-\tilde{r})}\right]^{\mathrm{T}}, \tag{3}$$

while $\tilde{r}_n$ and $\tilde{r}$ are the distances from the n-th array element and the reference element to a point within the spatially spread region. The distance $\tilde{r}_n$ can be computed as [15]

$$\tilde{r}_n = \tilde{r}\sqrt{1 - \frac{2\mathbf{k}^{\mathrm{T}}(\tilde{\varphi},\tilde{\theta})\mathbf{v}_n}{\tilde{r}} + \frac{\|\mathbf{v}_n\|^2}{\tilde{r}^2}}, \tag{4}$$

where $\mathbf{k}(\tilde{\varphi}, \tilde{\theta}) = [\cos\tilde{\theta}\cos\tilde{\varphi}, \cos\tilde{\theta}\sin\tilde{\varphi}, \sin\tilde{\theta}]^{\mathrm{T}}$ is the radiation direction from the point to the array. Finally, $f(\cdot)$ is the normalized spatial scattering function [14]. Unlike in FF conditions, the array response vector in (3) is influenced not only by $\tilde{\varphi}$ and $\tilde{\theta}$, but also by the distance $\tilde{r}$.

B. Pilot Signal Model

We assume a block fading channel model where the channel vectors remain static within a coherence block of τ_c channel uses. Within each block, $\tau_p \geq K$ uses are allocated for uplink channel estimation. We assume that the UEs transmit with power p and are separated by means of *orthogonal* pilot sequences of length τ_p. Focusing on a generic UE k and omitting the index for notational simplicity, the interference-free observation vector $\mathbf{y} \in \mathbb{C}^N$ is given by [16]:

$$\mathbf{y} = \mathbf{h} + \mathbf{w}, \tag{5}$$

where $\mathbf{w} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}_N, \sigma_w^2\mathbf{I}_N)$ with $\sigma_w^2 = N_0B/(p\tau_p)$. The MMSE estimate of $\mathbf{h}$ is [16]

$$\widehat{\mathbf{h}}^{\mathrm{mmse}} = \mathbf{R}\left(\mathbf{R} + \sigma_w^2\mathbf{I}_N\right)^{-1}\mathbf{y}. \tag{6}$$

Denoting by μ the rank of $\mathbf{A}$, the eigenvalue decomposition (EVD) of $\mathbf{A}$ yields

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathrm{H}}, \tag{7}$$

where $\mathbf{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_\mu\}$ collects the μ non-zero eigenvalues of $\mathbf{A}$, while $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_\mu]^T$ collects the corresponding *unit-norm* eigenvectors. Plugging (7) into (6) yields the following equivalent expression:

$$\widehat{\mathbf{h}}^{\mathrm{mmse}} = \mathbf{U}\mathbf{\Lambda}\left(\mathbf{\Lambda} + \frac{\sigma_w^2}{\beta}\mathbf{I}_\mu\right)^{-1}\mathbf{U}^{\mathrm{H}}\mathbf{y}, \tag{8}$$

from which we see that the MMSE estimator requires knowledge of both $\mathbf{A}$ (or its EVD decomposition) and $\{\beta, \sigma_w^2\}$. This problem is addressed next.

III. Parametric-based Channel Estimation

From (2), we see that, for a given array manifold and spatial scattering model, the spatial correlation matrix $\mathbf{R}$ is fully determined by the average channel power β, the UE location (r, φ, θ) and the triplet of intervals $(\triangle_r, \triangle_\varphi, \triangle_\theta)$. In practice, these parameters are not known. A possible solution consists of estimating the location of the UE and fixing sufficiently large values of the aforementioned intervals, say $(\bar{\triangle}_r, \bar{\triangle}_\varphi, \bar{\triangle}_\theta)$, which allows us to design a *refined* correlation matrix $\bar{\mathbf{R}}$. Assuming $\bar{\triangle}_r > \triangle_r$, $\bar{\triangle}_\varphi > \triangle_\varphi$, $\bar{\triangle}_\theta > \triangle_\theta$ and perfect estimation of (r, φ, θ), from [14, Lemma 1] the subspace spanned by the columns of $\bar{\mathbf{R}}$ contains the subspace spanned by the columns of $\mathbf{R}$. Hence, the refined matrix $\bar{\mathbf{R}}$ contains all *plausible* channel dimensions and can replace $\mathbf{R}$ in the MMSE estimator (8). As for the UE location (r, φ, θ), it can be estimated by using the sample covariance matrix and the MUSIC algorithm, as shown later.

A. Parametric-based Reconstruction of $\mathbf{R}$

We begin by observing that $\mathbf{R}$ captures macroscopic effects such as spatial channel correlation and average path loss. Hence, it changes slowly compared to $\mathbf{h}$ and maintains constant over τ_s coherence blocks, where τ_s can be at the order of hundreds or more [17], [18]. For this reason, we suppose that the BS has received the pilot signal in (5) over $M \leq \tau_s$ coherence blocks and denote by

$$\mathbf{y}(m) = \mathbf{h}(m) + \mathbf{w}(m), \ m = 1, 2, \ldots, M, \tag{9}$$

TABLE I: System parameters.

Parameter	Value
Carrier frequency	$f_0 = 0.1\,\text{THz}$
Wavelength	$\lambda = 3\,\text{mm}$
Number of antennas	$N_H \times N_V = 64 \times 32$
Antenna spacing	$\delta = \lambda/2$
Transmit power	$p = -4$ dBm
Pilot length	$\tau_p = 10$
Number of observation vectors	$M = 10$
Bandwidth	$B = 100$ MHz
Noise power	$N_0 B = -84$ dBm
UE position	$(r,\varphi,\theta) = (4\text{ m}, -20^\circ, -30^\circ)$
Average channel power	$\beta = -90$ dB
Angular spread of elevation angle	$\Delta_\theta = 1.5^\circ$
Assumed angular spread	$\bar{\Delta}_\theta = 5^\circ$

the M observations, with the channel and noise vectors being statistically independent for different values of *m*. Notice that such observations can be obtained from pilots already used for channel estimation in previous coherence blocks, so that no extra pilots are needed. We use vectors $\{\mathbf{y}(m)\}$ to obtain the sample correlation matrix as

$$\widehat{\mathbf{R}}_y^{\text{sample}} = \frac{1}{M}\sum_{m=1}^{M} \mathbf{y}(m)\mathbf{y}^{\text{H}}(m), \tag{10}$$

and find a feasible method to get an estimate of parameters (r,φ,θ) from $\widehat{\mathbf{R}}_y^{\text{sample}}$. Among several existing schemes, the MUSIC is a powerful approach for estimating the parameters of complex sinusoidal signals embedded in additive white Gaussian noise [19]. The EVD of (10) can be expressed as

$$\widehat{\mathbf{R}}_y^{\text{sample}} = \sum_{n=1}^{N} \xi_n \mathbf{q}_n \mathbf{q}_n^{\text{H}} = \mathbf{Q}_y \mathbf{\Sigma}_y \mathbf{Q}_y^{\text{H}}, \tag{11}$$

where $\{\xi_n | n = 1, 2, \ldots, N\}$ are the eigenvalues of $\widehat{\mathbf{R}}_y^{\text{sample}}$ and $\mathbf{q}_n \in \mathbb{C}^N$ is the unit-norm eigenvector corresponding to ξ_n. Furthermore, we have $\mathbf{\Sigma}_y = \text{diag}\{\xi_1, \xi_2, \ldots, \xi_N\}$ and $\mathbf{Q}_y = [\mathbf{q}_1, \ldots, \mathbf{q}_N]$. Since the channel and noise vectors in (9) are independent, $\widehat{\mathbf{R}}_y^{\text{sample}}$ can be decomposed as

$$\widehat{\mathbf{R}}_y^{\text{sample}} = \mathbf{Q}_y^s \mathbf{\Sigma}_y^s (\mathbf{Q}_y^s)^{\text{H}} + \mathbf{Q}_y^n \mathbf{\Sigma}_y^n (\mathbf{Q}_y^n)^{\text{H}} \tag{12}$$

based on the rank $\widehat{\mu}_y = \text{rank}\{\widehat{\mathbf{R}}_y^{\text{sample}}\}$[1], where $\mathbf{\Sigma}_y^s \in \mathbb{C}^{\widehat{\mu}_y \times \widehat{\mu}_y}$ contains the $\widehat{\mu}_y$ largest eigenvalues of $\widehat{\mathbf{R}}_y^{\text{sample}}$ and $\mathbf{Q}_y^s \in \mathbb{C}^{N \times \widehat{\mu}_y}$ spans the signal subspace of $\widehat{\mathbf{R}}_y^{\text{sample}}$ formed by the eigenvectors corresponding to these $\widehat{\mu}_y$ largest eigenvalues. Furthermore, $\mathbf{\Sigma}_y^n \in \mathbb{C}^{(N-\widehat{\mu}_y)\times(N-\widehat{\mu}_y)}$ is the diagonal matrix composed of the remaining eigenvalues and $\mathbf{Q}_y^n \in \mathbb{C}^{N\times(N-\widehat{\mu}_y)}$ is the noise subspace, spanned by the eigenvectors corresponding to these smaller eigenvalues. Then, recalling that $\widehat{\mathbf{R}}_y^{\text{sample}}$ converges (almost surely) to the true correlation matrix $\mathbf{R}_y = \text{E}\{\mathbf{y}(m)\mathbf{y}^{\text{H}}(m)\}$ as $M \to \infty$, and that the channel vector $\mathbf{h}$ is orthogonal to the noise subspace of $\mathbf{R}_y$, we construct the MUSIC metric as [19]

$$P(\tilde{r},\tilde{\varphi},\tilde{\theta}) = \frac{1}{\mathbf{a}^{\text{H}}(\tilde{r},\tilde{\varphi},\tilde{\theta})\mathbf{Q}_y^n(\mathbf{Q}_y^n)^{\text{H}}\mathbf{a}(\tilde{r},\tilde{\varphi},\tilde{\theta})}, \tag{13}$$

where $\mathbf{a}(\tilde{r},\tilde{\varphi},\tilde{\theta})$ is the search vector with $\tilde{r} \in [0, d_F]$, $\tilde{\varphi} \in [-\pi/2, \pi/2]$ and $\tilde{\theta} \in [-\pi/2, \pi/2]$. The estimated UE location $(\hat{r},\hat{\varphi},\hat{\theta})$ is then obtained through the following three-dimensional spectral peak search

$$(\hat{r},\hat{\varphi},\hat{\theta}) = \underset{(\tilde{r},\tilde{\varphi},\tilde{\theta})}{\arg\max}\{P(\tilde{r},\tilde{\varphi},\tilde{\theta})\} \tag{14}$$

and used, together with the appropriately selected triplet $(\bar{\Delta}_r, \bar{\Delta}_\varphi, \bar{\Delta}_\theta)$, to compute the parametric-based correlation matrix as $\bar{\mathbf{R}} = \widehat{\beta}\bar{\mathbf{A}}$, where

$$\bar{\mathbf{A}} = \int_{\hat{r}-\bar{\Delta}_r}^{\hat{r}+\bar{\Delta}_r}\int_{\hat{\varphi}-\bar{\Delta}_\varphi}^{\hat{\varphi}+\bar{\Delta}_\varphi}\int_{\hat{\theta}-\bar{\Delta}_\theta}^{\hat{\theta}+\bar{\Delta}_\theta} f(\tilde{r},\tilde{\varphi},\tilde{\theta})\cdot \mathbf{a}(\tilde{r},\tilde{\varphi},\tilde{\theta})\mathbf{a}^{\text{H}}(\tilde{r},\tilde{\varphi},\tilde{\theta})d\tilde{\theta}d\tilde{\varphi}d\tilde{r}, \tag{15}$$

and $\widehat{\beta}$ is an estimate of β, which is computed next.

B. Parametric-based Channel Estimator

Letting $\bar{\mu} = \text{rank}\{\bar{\mathbf{R}}\}$, the EVD of the refined correlation matrix $\bar{\mathbf{R}}$ takes the form

$$\bar{\mathbf{R}} = \widehat{\beta}\bar{\mathbf{U}}\bar{\mathbf{\Lambda}}\bar{\mathbf{U}}^{\text{H}}, \tag{16}$$

where $\bar{\mathbf{\Lambda}} = \text{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_{\bar{\mu}}\}$ contains the $\bar{\mu}$ non-zero eigenvalues of $\bar{\mathbf{A}}$, and $\bar{\mathbf{U}} = [\mathbf{u}_1, \cdots, \mathbf{u}_{\bar{\mu}}]$ contains the eigenvectors corresponding to these eigenvalues. As mentioned previously, when $\bar{\Delta}_r - \Delta_r \geq |\widehat{r} - r|$, $\bar{\Delta}_\varphi - \Delta_\varphi \geq |\widehat{\varphi} - \varphi|$ and $\bar{\Delta}_\theta - \Delta_\theta \geq |\widehat{\theta} - \theta|$, the subspace spanned by the columns of $\bar{\mathbf{U}}$ covers the subspace spanned by the columns of the true $\mathbf{R}$. Hence, the parametric-based channel estimator based on $\bar{\mathbf{R}}$ takes the form

$$\widehat{\mathbf{h}}^{\text{parametric}} = \bar{\mathbf{U}}\bar{\mathbf{\Lambda}}\left(\bar{\mathbf{\Lambda}} + \frac{\widehat{\sigma}_w^2}{\widehat{\beta}}\mathbf{I}_{\bar{\mu}}\right)^{-1}\bar{\mathbf{U}}^{\text{H}}\mathbf{y}, \tag{17}$$

where $\widehat{\sigma}_w^2$ is an estimate of σ_w^2, which is computed next.

C. Estimation of β and σ_w^2

We define $\bar{\mathbf{U}}^n \in \mathbb{C}^{N\times(N-\bar{\mu})}$ as the noise subspace of $\bar{\mathbf{R}}$, spanned by the eigenvectors corresponding to the $N - \bar{\mu}$ smaller eigenvalues. Then, we exploit the observation vectors $\mathbf{y}(m)$ in (9) to compute

$$\bar{\mathbf{x}}_n(m) = \left(\bar{\mathbf{U}}^n\right)^{\text{H}}\mathbf{y}(m),\ m = 1, 2, \ldots, M. \tag{18}$$

Observing that $\left(\bar{\mathbf{U}}^n\right)^{\text{H}}\mathbf{h}$ is ideally zero, after substituting (9) into (18) we obtain $\bar{\mathbf{x}}_n(m) = \left(\bar{\mathbf{U}}^n\right)^{\text{H}}\mathbf{w}(m) = \mathbf{n}(m)$, where $\mathbf{n}(m) \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}_{N-\bar{\mu}}, \sigma_w^2\mathbf{I}_{N-\bar{\mu}})$. An estimate of σ_w^2 is thus given by

$$\widehat{\sigma}_w^2 = \frac{1}{M(N-\bar{\mu})}\sum_{m=1}^{M}\|\bar{\mathbf{x}}_n(m)\|^2. \tag{19}$$

Finally, observing that

$$\text{E}\left\{\|\mathbf{y}(m)\|^2\right\} = N\left(\beta + \sigma_w^2\right), \tag{20}$$

an estimate of β can be obtained as

$$\widehat{\beta} = \frac{1}{N}\text{tr}\{\widehat{\mathbf{R}}_y^{\text{sample}}\} - \widehat{\sigma}_w^2, \tag{21}$$

[1]In the simulations, we utilize $\xi_{\widehat{\mu}_y} \geq 10^{-2}$ to determine the rank of $\widehat{\mathbf{R}}_y^{\text{sample}}$.

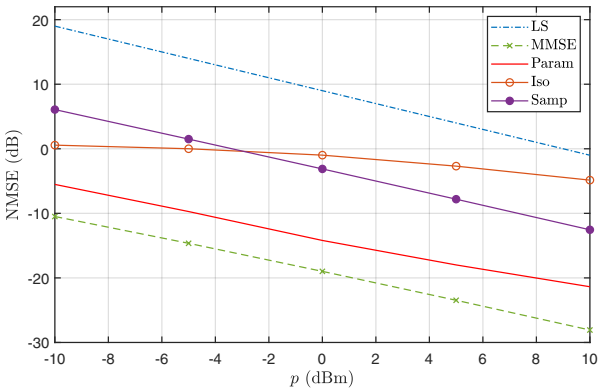

Fig. 1: The NMSE vs. p for different estimators using the simulation parameters provided in Table I.

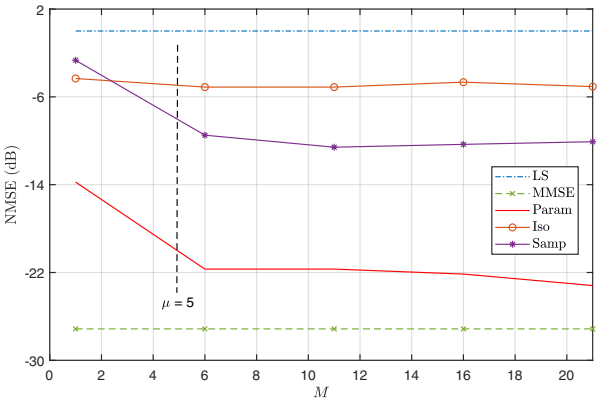

Fig. 2: The NMSE vs. M for different estimators using the simulation parameters provided in Table I.

where $\widehat{\sigma}_w^2$ is given in (19). The estimates $(\widehat{\beta}, \widehat{\sigma}_w^2)$ are eventually used in (17) to evaluate $\widehat{\mathbf{h}}^{\text{parametric}}$.

IV. Numerical Results

We now evaluate the performance of the parametric channel estimation scheme (17) in terms of the normalized mean squared error (NMSE). To better reflect sub-THz channel performance, we adopt parameters consistent with existing studies: the channel power is set to $\beta = -90$ dB [4], the transmit power to $p = -4$ dBm (except for Fig. 1) [20], the uplink channel bandwidth $B = 100$ MHz [21] and the noise power spectral density to $N_0 = -174$ dBm/Hz [6]. The BS employs a UPA consisting of 64×32 elements, with an array aperture given by $D = \sqrt{(N_V^2 + N_H^2)\delta} = 0.107$ m. This configuration corresponds to a Fraunhofer distance of $d_F = 7.680$ m. We consider a specific UE located at $(r, \varphi, \theta) = (4\text{ m}, -20°, -30°)$, which lies within the radiative NF of the array. In the MUSIC algorithm, we conduct a three-dimensional search over the space defined by $\tilde{r} \in [0.5\text{m}, d_F]$, $\tilde{\varphi} \in [-90°, 90°]$ and $\tilde{\theta} \in [-90°, 90°]$. The search is performed with the step sizes of 0.5m for $\tilde{r}$ and $0.5°$ for both $\tilde{\varphi}$ and $\tilde{\theta}$. For the refined $\bar{\mathbf{R}}$, the elevation angle spread is set to $\bar{\triangle}_\theta = 5°$, while the distance and azimuth spreads are computed using $\bar{\triangle}_r = \hat{r}(\cos(\hat{\theta} - \bar{\triangle}_\theta) - \cos(\hat{\theta} + \bar{\triangle}_\theta))/2$ and $\bar{\triangle}_\varphi = \arctan(\bar{\triangle}_r / (\hat{r}\cos\hat{\theta}))$. Unless otherwise stated, other system parameters are listed in Table I.

The parametric estimator (17), labeled as 'Param', is compared with the MMSE estimator (8), labeled as 'MMSE', as well as with three alternative methods:

- The LS estimator: This method directly sets the channel estimate as $\widehat{\mathbf{h}}^{\text{ls}} = \mathbf{y}$, labeled as 'LS';
- The sample correlation matrix-based estimator, i.e.,

$$\widehat{\mathbf{h}}^{\text{sample}} = \left(\widehat{\mathbf{R}}_y^{\text{sample}} - \widehat{\sigma}_w^2 \mathbf{I}_N\right)\left(\widehat{\mathbf{R}}_y^{\text{sample}}\right)^{-1}\mathbf{y}, \quad (22)$$

  which is labeled as 'Samp';
- The isotropic correlation matrix-based estimator:

$$\widehat{\mathbf{h}}^{\text{iso}} = \widehat{\mathbf{R}}^{\text{iso}}\left(\widehat{\mathbf{R}}^{\text{iso}} + \widehat{\sigma}_w^2 \mathbf{I}_N\right)^{-1}\mathbf{y} \quad (23)$$

  with $\widehat{\mathbf{R}}^{\text{iso}}$ being given by [14, Eq. (8)]. This estimator is labeled as 'Iso'.

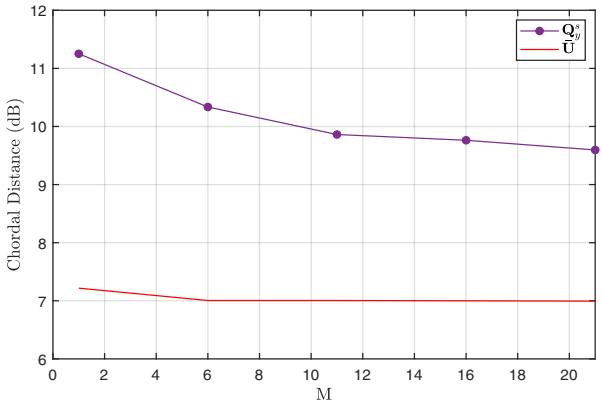

Fig. 3: The chordal distance vs. M for different correlation matrices using the simulation parameters provided in Table I.

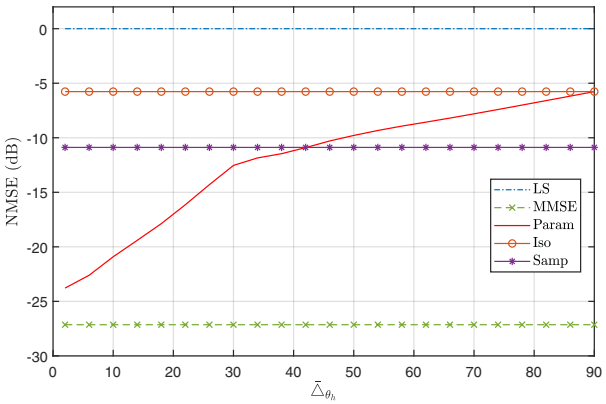

Fig. 4: The NMSE vs. $\bar{\triangle}_\theta$ for different estimators using the simulation parameters provided in Table I.

Fig. 1 plots the NMSE of the different estimators as a function of the transmit power p. As expected, the NMSE of all estimators decreases as the transmit power increases, reflecting the general trend that a higher SNR improves the channel estimation accuracy. More importantly, compared to existing schemes, the proposed estimator exhibits a significant performance advantage, ranking second only to the conventional MMSE estimator, which validates the effectiveness of the proposed method.

Fig. 2 illustrates the NMSE as a function of the number of received pilot vectors, M. The figure also indicates $\mu = \text{rank}\{\mathbf{R}\}$. As observed, the proposed estimator is applicable for any $M \geq 1$ and naturally improves as M increases. The most significant NMSE reduction occurs when $M > \mu$, i.e.,

$M = 6$, after which the performance gradually stabilizes, indicating that additional pilot vectors provide only marginal improvement.

Fig. 3 plots the chordal distance[2] [16, Eq. (7.21)] between the subspaces of the sample correlation matrix $\mathbf{Q}_y^s$ and the refined correlation matrix $\bar{\mathbf{U}}$ with respect to the actual subspace $\mathbf{U}$. From Fig.3, we observe that although the chordal distance of $\mathbf{Q}_y^s$ decreases as M increases, it remains consistently higher than that of $\bar{\mathbf{U}}$. This further explains the performance advantage of 'Param' over 'Samp' as observed in Fig. 2.

Fig. 4 evaluates the impact of the selected elevation angle spread $\bar{\triangle}_\theta$ on the proposed estimator for values ranging from 2° to 90°. However, it is important to note that for sub-THz band channels, the angular spread is smaller than 10°. Here, we merely provide an example illustrating how the proposed estimator varies with $\bar{\triangle}_\theta$. As expected, the performance of the proposed estimator gradually deteriorates as $\bar{\triangle}_\theta$ increases. Notably, when the selected angular spread approaches isotropy, i.e., $\bar{\triangle}_\theta \geq 30^\circ$, the accuracy degradation gradually decreases. When the selected $\bar{\triangle}_\theta = 90^\circ$, i.e., the refined $\bar{\mathbf{R}}$ under the fully isotropic assumption, the proposed estimator achieves the same accuracy of the one presented in [14].

V. Conclusions

We considered the parametric NF channel estimation problem in the sub-THz band without any prior information. The MUSIC algorithm was used to estimate the location of the UE relative to the BS. Using the estimated location, the spatial correlation matrix of the UE-BS channel was reconstructed and incorporated into an approximation of the MMSE estimator to derive the NF channel estimate in the sub-THz band. Numerical results showed that the proposed method significantly outperforms the conventional LS estimator and other existing methods. Given the computational complexity of 3-D MUSIC, our future work will focus on developing a low-complexity UE localization approach tailored for the THz band. This method will be integrated with a parametric channel estimation framework to not only reduce the computational overhead, but also to enhance the estimation performance.

Acknowledgment

This work has been performed in the framework of the HORIZON-JU-SNS-2022 project TIMES, cofunded by the European Union. This work is also supported in part by Italian Ministry of Education and Research (MUR) in the Framework of the FoReLab Project (Departments of Excellence) and the Project GARDEN funded by EU in NextGenerationEU Plan through Italian "Bando Prin 2022-D.D.1409 del 14-09-2022".

References

[1] Z. Chen, C. Han, Y. Wu, L. Li, C. Huang, Z. Zhang, G. Wang, and W. Tong, "Terahertz wireless communications for 2030 and beyond: A cutting-edge frontier," *IEEE Commun. Mag.*, vol. 59, no. 11, pp. 66–72, Nov. 2021.

[2]The chordal distance $d_C(\mathbf{X}, \mathbf{Y})$ between two matrices $\mathbf{X}$ and $\mathbf{Y}$ is defined as $d_C = \|\mathbf{X}\mathbf{X}^{\mathrm{H}} - \mathbf{Y}\mathbf{Y}^{\mathrm{H}}\|_F^2$

[2] C. Han, Y. Wang, Y. Li, Y. Chen, N. A. Abbasi, T. Kürner, and A. F. Molisch, "Terahertz wireless channels: A holistic survey on measurement, modeling, and analysis," *IEEE Commun. Surv. Tutor.*, vol. 24, no. 3, pp. 1670–1707, June 2022.

[3] D. Bodet, V. Petrov, S. Petrushkevich, and J. M. Jornet, "Sub-terahertz near field channel measurements and analysis with beamforming and Bessel beams," *Sci. Rep.*, vol. 14, no. 1, p. 19675, Aug. 2024.

[4] D. G. Selimis, M. F. De Guzman, K. N. Manganaris, F. I. Lazarakis, K. Haneda, and K. P. Peppas, "Path loss, angular spread and channel sparsity modeling for indoor and outdoor environments at the sub-THz band," *Phys. Commun.*, vol. 66, p. 102453, Feb. 2024.

[5] N. A. Abbasi, A. Hariharan, A. M. Nair, and A. F. Molisch, "Channel measurements and path loss modeling for indoor THz communication," in *2020 14th European Conference on Antennas and Propagation (EuCAP)*, 2020, pp. 1–5.

[6] R. Okura, Y. Koda, and H. Harada, "Performance evaluation of low sub-THz 5G NR sidelink for ultra-wideband short-range communication," in *2024 IEEE 100th Vehicular Technology Conference (VTC2024-Fall)*, 2024, pp. 1–5.

[7] W.-X. Long, M. Moretti, L. Sanguinetti, and R. Chen, "Channel estimation in RIS-aided communications with interference," *IEEE Wireless Commun. Lett.*, vol. 12, no. 10, pp. 1751–1755, Oct. 2023.

[8] W.-X. Long, M. Moretti, A. Abrardo, L. Sanguinetti, and R. Chen, "MMSE design of RIS-aided communications with spatially-correlated channels and electromagnetic interference," *IEEE Trans. Wireless Commun.*, vol. 23, no. 11, pp. 16 992–17 006, Nov. 2024.

[9] J. He, M. N. S. Swamy, and M. O. Ahmad, "Efficient application of MUSIC algorithm under the coexistence of far-field and near-field sources," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 2066–2070, Apr. 2012.

[10] K. Qu, S. Guo, J. Ye, and H. Zhao, "Two-stage beamspace MUSIC-based near-field channel estimation for hybrid XL-MIMO," *IEEE Commun. Lett.*, vol. 28, no. 8, pp. 1949–1953, Aug. 2024.

[11] A. Kosasih, Ö. T. Demir, and E. Björnson, "Parametric near-field channel estimation for extremely large aperture arrays," in *2023 57th Asilomar Conference on Signals, Systems, and Computers*, 2023, pp. 162–166.

[12] E. Björnson and L. Sanguinetti, "Rayleigh fading modeling and channel hardening for reconfigurable intelligent surfaces," *IEEE Wireless Commun. Lett.*, vol. 10, no. 4, pp. 830–834, Apr. 2021.

[13] M. Cui, Z. Wu, Y. Lu, X. Wei, and L. Dai, "Near-field MIMO communications for 6G: Fundamentals, challenges, potentials, and future directions," *IEEE Commun. Mag.*, vol. 61, no. 1, pp. 40–46, Jan. 2023.

[14] Ö. T. Demir, A. Kosasih, and E. Björnson, "Spatial correlation modeling and RS-LS estimation of near-field channels with uniform planar arrays," in *2024 IEEE 25th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2024, pp. 236–240.

[15] Y. Liu, Z. Wang, J. Xu, C. Ouyang, X. Mu, and R. Schober, "Near-field communications: A tutorial review," *IEEE Open J. Commun. Soc.*, vol. 4, pp. 1999–2049, Aug. 2023.

[16] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017.

[17] K. Upadhya and S. A. Vorobyov, "Covariance matrix estimation for massive MIMO," *IEEE Signal Processing Letters*, vol. 25, no. 4, pp. 546–550, Apr. 2018.

[18] A. K. Kocharlakota, K. Upadhya, and S. A. Vorobyov, "Impact of pilot overhead and channel estimation on the performance of massive MIMO," *IEEE Trans. Commun.*, vol. 69, no. 12, pp. 8242–8255, Dec. 2021.

[19] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[20] J. Yu, J. Chen, P. Zhou, Z. Li, H. Li, P. Yan, D. Hou, and W. Hong, "A 300 GHz transmitter front end with 4.1 dBm peak output power for sub-THz communication using 130-nm SiGe BiCMOS technology," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 11, pp. 4925–4936, Aug. 2021.

[21] Y. Xing, O. Kanhere, S. Ju, and T. S. Rappaport, "Sub-terahertz wireless coverage analysis at 142 GHz in urban microcell," in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 3942–3947.