telecom

MDPI

Article

Variational Bayesian Near-Field Channel Estimation for Distributed MIMO Systems

Ling He *, Qingrui Guo, Xuerang Guo, Huiting Yang and Yanan Xin

State Grid Xinjiang Electric Power Research Institute, No. 66, East Second Lane, Changchun Middle Road, Xinshi District, Urumqi 830011, China; gur1008@163.com (Q.G.); gxrgxr820@163.com (X.G.); yanght_2016@163.com (H.Y.); 18599007359@163.com (Y.X.)

* Correspondence: 17726841506@163.com; Tel.: +86-17726841506

Abstract

Distributed multiple-input multiple-output (MIMO) is a promising architecture for future wireless systems because cooperation among geographically separated base stations (BSs) improves coverage, spectral efficiency, and link reliability. However, the large effective aperture formed by distributed BSs makes near-field effects non-negligible and complicates accurate channel state information acquisition. Existing near-field estimators often suffer from modeling errors caused by approximate angle–range decoupling or from the high storage and computational costs of dense two-dimensional sparse representations. This article proposes an off-grid variational Bayesian channel-estimation framework for the considered distributed near-field MIMO geometry, which comprises equally spaced, collinear BS reference points and aligned uniform linear arrays (ULAs) with common inter-element spacing. We establish a geometry-coupled model based on the exact geometric spherical-wave phase response and map the local direction–range parameters observed by different BSs into a common reference coordinate system, yielding a two-dimensional jointly sparse representation. An independent-vector variational Bayesian inference algorithm then decomposes the high-dimensional multiuser recovery problem into user-specific posterior subproblems. It operates directly on the received pilot matrices, avoiding pilot–matrix inversion and the resulting distortion of noise statistics. A two-dimensional skewed off-grid update is further embedded in an expectation-maximization procedure to jointly refine angle and range offsets, mitigating basis mismatch while permitting a coarser initial dictionary. Simulation results support the effectiveness of the proposed method in the evaluated scenarios.

Keywords: distributed MIMO; near-field communications; channel estimation; sparse Bayesian learning; variational Bayesian inference; off-grid refinement

Check for updates

Academic Editor: Seyeong Choi

Received: 22 July 2026
Revised: 20 August 2026
Accepted: 24 August 2026
Published: 1 September 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Multiple-input multiple-output (MIMO) is a key enabling technology for modern wireless communication systems. By equipping a base station (BS) with multiple antennas, MIMO provides array, spatial-multiplexing, and diversity gains that improve spectral efficiency and link reliability while supporting multiuser transmission and interference suppression [1,2]. Distributed MIMO further extends these benefits by allowing multiple geographically separated BSs to cooperatively serve users through coordinated signal processing. This distributed architecture provides macro-diversity, broadens spatial coverage, shortens access distances, and improves service reliability [3–5]. A representative application of distributed MIMO is a private wireless network for power-grid communications,

in which geographically separated BSs cooperatively provide reliable coverage for wide-area monitoring and control services. However, the combination of high carrier frequencies and the large effective aperture formed by the spatially separated BSs significantly enlarges the radiative near-field region. Consequently, users and dominant scatterers may lie within the Rayleigh distance of the distributed array even when they are not particularly close to any individual BS [6–8]. For example, consider the default setting with $N = 3$ BSs separated by $R = 50$ m, each equipped with $M = 15$ antennas having inter-element spacing $\Delta = \lambda/2$, where the wavelength is $\lambda = 0.01$ m. The resulting end-to-end array span is approximately $D = 100.07$ m. If D is used as a nominal aperture in the classical criterion $2D^2/\lambda$, the resulting nominal Rayleigh-distance estimate is approximately 2.003×10^6 m (2003 km), far exceeding the simulated path-range interval of 15–25 m [6,8]. This span-based value is intended only as an order-of-magnitude indicator and not as an exact near-/far-field boundary for the sparse distributed array. In this regime, the wavefront is spherical rather than planar, and the channel response depends jointly on direction and range.

This range-dependent wavefront curvature complicates accurate channel state information (CSI) acquisition. Conventional massive-MIMO channel estimators commonly adopt a plane-wave model and exploit angle-domain sparsity, making them susceptible to model mismatch under spherical-wave propagation [9]. As a result, a single physical path may spread its energy over multiple angular atoms, weakening sparsity and degrading support recovery [10]. Accurate near-field recovery must therefore resolve both angle and range, which substantially enlarges the parameter-search space. A densely sampled two-dimensional dictionary improves resolution at the expense of storage and computational complexity, whereas a coarse grid produces basis mismatch when the true parameters do not coincide with the predefined atoms. To reduce this burden, some methods decouple angle and range through Fresnel approximations [11,12]. However, neglected higher-order phase terms can introduce systematic modeling errors, and stepwise processing can propagate errors from one parameter-estimation stage to the next [13–15]. Addressing these limitations requires near-field estimators that balance modeling accuracy, parameter resolution, and computational complexity.

Recently, representative near-field channel-estimation solutions have been developed for different array geometries and propagation conditions. Han et al. divided an extremely large array into smaller subarrays and exploited local stationarity for channel estimation [16]. Qu et al. developed a two-stage beamspace multiple signal classification (MUSIC) method to reduce the complexity of a joint angle–range search [17]. Cui and Dai introduced a polar-domain sparse representation and proposed on-grid and off-grid recovery schemes that jointly account for angle and distance [10]. Zhang et al. developed a distance-parameterized angular-domain representation to reduce dictionary size and coherence [18], while Guo et al. decomposed azimuth, elevation, and range parameters for compressed channel estimation with planar arrays [19]. Huang et al. and Pan et al. developed low-complexity angle–range estimation schemes for uniform planar arrays and reconfigurable intelligent surface (RIS)-assisted systems, respectively [11,12]. In addition, Yang et al. considered hybrid-field channels and developed a practical strategy for estimating the proportions of near- and far-field paths [20].

Model-based probabilistic inference and deep learning methods have also been developed for more complex near-field channel-estimation problems. Qiao et al. combined subarray processing and Bayesian inference to jointly infer user activity, channel, and location in near-field extremely large-scale MIMO (XL-MIMO) systems [21]. Tang et al. proposed a three-layer generalized approximate message passing method for spatially non-stationary XL-MIMO channels [22], while Jin et al. used a deep generative model guided by side information to refine near-field channel estimates [23]. Closely related to the

present work, Cao et al. developed a sparse model based on the exact geometric spherical-wave phase response and a generalized approximate message passing (GAMP)-embedded sparse Bayesian learning estimator with coarse off-grid refinement for RIS-aided industrial Internet of Things (IoT) systems [15]. Their method reduces the spatial search dimension by exploiting environment-specific constraints on device locations and is tailored to planar RIS geometry.

Despite these advances, distributed near-field MIMO introduces additional structure that is not fully captured by existing methods developed for colocated arrays. First, each BS observes a different local angle and range for the same physical path, while these local parameters are coupled by the known BS geometry. Treating the BSs independently discards this geometric coupling and can lead to inconsistent support estimates. Second, a sequential estimator that first obtains local angles and then infers the range uses an intermediate hard decision; any angular error is propagated into the subsequent geometric inference. Third, a very fine two-dimensional angle–range grid can reduce basis mismatch but greatly increases the number of atoms and the computational burden. These limitations motivate the distributed near-field channel-estimation framework developed in this work.

This paper considers uplink channel estimation in a distributed MIMO system under an exact geometric spherical-wave phase model. The main contributions are as follows:

- A geometry-coupled near-field phase representation based on unapproximated geometric propagation distances is developed for distributed MIMO systems. Using the known BS locations, the local direction–range parameters observed by spatially separated BSs are mapped into a common reference coordinate system, exposing their joint angle–range sparse structure and yielding a two-dimensional sparse representation. Unlike Fresnel-based or subarray far-field models, the resulting dictionary retains the coupled direction–range dependence of spherical-wave propagation.
- An independent-vector variational Bayesian inference (IV-VBI) framework with two-dimensional skewed off-grid refinement is introduced for multiuser channel recovery. The IV-VBI procedure decomposes the original high-dimensional coupled sparse-recovery problem into user-specific posterior subproblems and operates directly on the received pilot matrices, thereby avoiding pilot–matrix inversion and the resulting distortion of the noise statistics. An adaptive off-grid update embedded in an expectation-maximization (EM) procedure jointly refines the angle and range offsets, mitigates basis mismatch, and permits a relatively coarse initial dictionary, thereby reducing the storage and computational burden of two-dimensional sparse recovery.

Numerical results under Gaussian and orthogonal pilots show that the proposed estimator consistently outperforms the practical far-field, greedy, and sequential baselines and approaches the genie-aided benchmark as the signal-to-noise ratio (SNR), the number of snapshots, the number of users, and the array size vary.

The remainder of the paper is organized as follows. Section 2 introduces the distributed near-field signal model. Section 3 develops the sparse Bayesian estimator and off-grid refinement. Section 4 presents the numerical results, and Section 5 concludes the paper.

2. System Model and Review of Conventional Estimation

As illustrated in Figure 1, we consider an uplink distributed MIMO system in which N distributed BSs jointly serve U single-antenna users. Each BS is equipped with an M-element uniform linear array (ULA). A power-grid private wireless network is a representative application of this distributed MIMO architecture; accordingly, the proposed estimator applies to distributed MIMO systems, including power-grid deployments, rather than being restricted to a power-specific scenario. At time t, user u transmits symbol $s_u(t)$, and the signal received at the nth BS is

$$\begin{aligned}\mathbf{y}_n(t) &= \sum_{u=0}^{U-1} \mathbf{h}_{u,n} s_u(t) + \mathbf{w}_n(t) \\ &= \mathbf{H}_n \mathbf{s}(t) + \mathbf{w}_n(t),\end{aligned} \tag{1}$$

where $\mathbf{y}_n(t) \in \mathbb{C}^{M\times 1}$ is the received-signal vector, $\mathbf{h}_{u,n} \in \mathbb{C}^{M\times 1}$ is the uplink channel from user u to BS n, and $\mathbf{w}_n(t) \in \mathbb{C}^{M\times 1}$ is additive white Gaussian noise. Moreover, $\mathbf{H}_n \triangleq [\mathbf{h}_{0,n}, \mathbf{h}_{1,n}, \ldots, \mathbf{h}_{U-1,n}] \in \mathbb{C}^{M\times U}$ and $\mathbf{s}(t) \triangleq [s_0(t), s_1(t), \ldots, s_{U-1}(t)]^{\mathrm{T}}$.

Following the limited-scattering model [9], the channel is represented by P dominant paths as

$$\begin{aligned}\mathbf{h}_{u,n} &= \sum_{p=0}^{P-1} g_{u,p,n} \mathbf{a}(\bar{\theta}_{u,p,n}, \bar{r}_{u,p,n}) \\ &= \mathbf{A}(\bar{\boldsymbol{\theta}}_{u,n}, \bar{\mathbf{r}}_{u,n}) \mathbf{g}_{u,n},\end{aligned} \tag{2}$$

where $\bar{\theta}_{u,p,n}$ and $\bar{r}_{u,p,n}$ denote the angle and range of the pth path of user u relative to BS n, respectively, and $g_{u,p,n}$ is the corresponding complex path gain. The path-gain vector and array-manifold matrix are $\mathbf{g}_{u,n} \triangleq [g_{u,0,n}, g_{u,1,n}, \ldots, g_{u,P-1,n}]^{\mathrm{T}} \in \mathbb{C}^{P\times 1}$ and $\mathbf{A}(\bar{\boldsymbol{\theta}}_{u,n}, \bar{\mathbf{r}}_{u,n}) \triangleq [\mathbf{a}(\bar{\theta}_{u,0,n}, \bar{r}_{u,0,n}), \ldots, \mathbf{a}(\bar{\theta}_{u,P-1,n}, \bar{r}_{u,P-1,n})] \in \mathbb{C}^{M\times P}$. For the ULA, the steering vector based on the exact geometric propagation distances is

$$\mathbf{a}(\bar{\theta}_{u,p,n}, \bar{r}_{u,p,n}) \triangleq \left[1, e^{j2\pi(d_{u,p,n,1} - \bar{r}_{u,p,n})/\lambda}, \ldots, e^{j2\pi(d_{u,p,n,M-1} - \bar{r}_{u,p,n})/\lambda}\right]^{\mathrm{T}}, \tag{3}$$

where λ is the wavelength. The distance from the mth antenna element to the path endpoint is

$$d_{u,p,n,m} \triangleq \sqrt{\bar{r}_{u,p,n}^2 + (m\Delta)^2 + 2\bar{r}_{u,p,n} m\Delta \sin\bar{\theta}_{u,p,n}}, \tag{4}$$

where Δ is the inter-element spacing. Unlike a plane-wave steering vector, (3) retains the curvature-dependent phase across the complete aperture.

Collecting T snapshots gives

$$\begin{aligned}\mathbf{Y}_n &\triangleq [\mathbf{y}_n(0), \mathbf{y}_n(1), \ldots, \mathbf{y}_n(T-1)] \\ &= \mathbf{H}_n \mathbf{S} + \mathbf{W}_n \\ &= [\mathbf{A}(\bar{\boldsymbol{\theta}}_{0,n}, \bar{\mathbf{r}}_{0,n}) \mathbf{g}_{0,n}, \ldots, \mathbf{A}(\bar{\boldsymbol{\theta}}_{U-1,n}, \bar{\mathbf{r}}_{U-1,n}) \mathbf{g}_{U-1,n}] \mathbf{S} + \mathbf{W}_n,\end{aligned} \tag{5}$$

where $\mathbf{S} \triangleq [\mathbf{s}(0), \ldots, \mathbf{s}(T-1)] \in \mathbb{C}^{U\times T}$ and $\mathbf{W}_n \triangleq [\mathbf{w}_n(0), \ldots, \mathbf{w}_n(T-1)] \in \mathbb{C}^{M\times T}$.

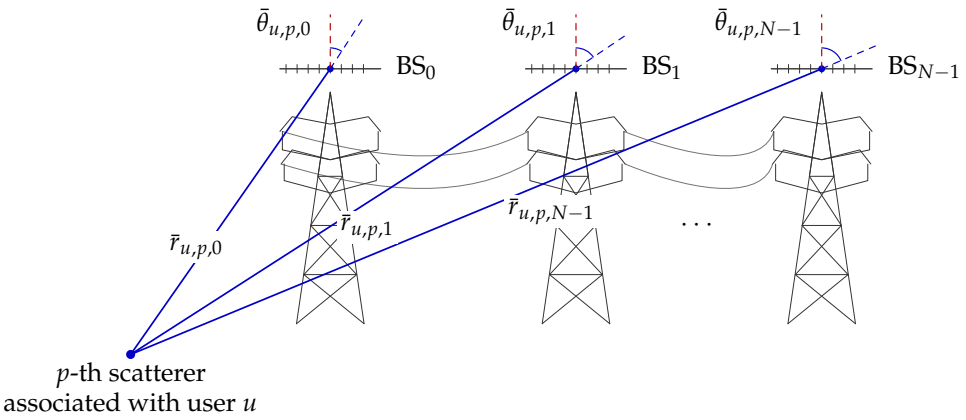


Figure 1. Distributed MIMO uplink geometry and the angle–range relationship observed by spatially separated BS arrays.

The objective is to estimate $\{\mathbf{H}_n\}_{n=0}^{N-1}$ from $\{\mathbf{Y}_n\}_{n=0}^{N-1}$. A direct least-squares (LS) estimate [24] follows from $\mathbf{Y}_n\mathbf{S}^\dagger = \mathbf{H}_n + \mathbf{W}_n\mathbf{S}^\dagger$; hence, $\widehat{\mathbf{H}}_n^{\mathrm{LS}} = \mathbf{Y}_n\mathbf{S}^\dagger$, where $(\cdot)^\dagger$ denotes the Moore–Penrose inverse. This operation neither exploits near-field sparsity nor preserves the original white-noise covariance when $\mathbf{S}$ is nonorthogonal. A subarray far-field alternative first estimates $\{\bar{\theta}_{u,n}\}$ independently at each BS and then infers range from geometry [21,25], but the local plane-wave approximation and sequential processing produce model mismatch and error propagation.

The conventional estimators used later differ primarily in how they exploit channel structure. LS ignores sparsity, far-field variational Bayesian inference (VBI) neglects range-dependent wavefront curvature, two-stage MUSIC estimates angles before inferring ranges, and orthogonal matching pursuit (OMP) makes hard support decisions on a finite grid. These concise distinctions motivate the geometry-aware Bayesian formulation developed below; the method-specific performance differences are discussed together with the numerical results in Section 4.

3. Distributed Near-Field Sparse Channel Estimation

3.1. Unified Near-Field Sparse Representation

To exploit the observations of all BSs jointly, the first antenna of a selected reference BS $\tilde{n}$ is used as the global reference. For the pth path of user u, define $\theta_{u,p} \triangleq \bar{\theta}_{u,p,\tilde{n}}$ and $r_{u,p} \triangleq \bar{r}_{u,p,\tilde{n}}$. Assuming adjacent BS reference points are separated by R, the exact distance from this path endpoint to antenna m at BS n is

$$\begin{aligned} d_{u,p,n,m} &= \widetilde{d}_{n,m}(\theta_{u,p}, r_{u,p}) \\ &\triangleq \sqrt{r_{u,p}^2 + [m\Delta + (n-\tilde{n})R]^2 + 2r_{u,p}[m\Delta + (n-\tilde{n})R]\sin\theta_{u,p}}. \end{aligned} \tag{6}$$

Equation (6) follows directly from the considered collinear geometry. Here, $m\Delta$ is the displacement of antenna m from the reference element of BS n, whereas $(n-\tilde{n})R$ is the displacement of BS n from the reference BS. Their sum gives the antenna position relative to the global reference point, and the propagation distance then follows from the law of cosines. This formulation assumes equally spaced BS reference points and aligned ULAs. Thus, the original local steering vector is related to the global-reference steering vector without approximation as

$$\mathbf{a}(\bar{\theta}_{u,p,n}, \bar{r}_{u,p,n}) = e^{-j2\pi[\widetilde{d}_{n,0}(\theta_{u,p},r_{u,p})-r_{u,p}]/\lambda}\mathbf{b}_n(\theta_{u,p}, r_{u,p}), \tag{7}$$

where $\mathbf{b}_n(\theta_{u,p}, r_{u,p}) \triangleq \left[e^{j2\pi(\widetilde{d}_{n,0}(\theta_{u,p},r_{u,p})-r_{u,p})/\lambda}, \ldots, e^{j2\pi(\widetilde{d}_{n,M-1}(\theta_{u,p},r_{u,p})-r_{u,p})/\lambda}\right]^{\mathrm{T}}$. The factor preceding $\mathbf{b}_n(\theta_{u,p}, r_{u,p})$ is a BS-dependent unit-modulus common phase and can therefore be absorbed into the complex path gain. Specifically, redefining $g_{u,p,n}$ as $g_{u,p,n}e^{-j2\pi[\widetilde{d}_{n,0}(\theta_{u,p},r_{u,p})-r_{u,p}]/\lambda}$ leaves the channel unchanged. For notational simplicity, the redefined gain is still denoted by $g_{u,p,n}$ in the sequel. Substituting (7) into (5) yields

$$\mathbf{Y}_n = [\mathbf{B}_n(\boldsymbol{\theta}_0, \mathbf{r}_0)\mathbf{g}_{0,n}, \ldots, \mathbf{B}_n(\boldsymbol{\theta}_{U-1}, \mathbf{r}_{U-1})\mathbf{g}_{U-1,n}]\mathbf{S} + \mathbf{W}_n, \tag{8}$$

where

$$\mathbf{B}_n(\boldsymbol{\theta}_u, \mathbf{r}_u) \triangleq [\mathbf{b}_n(\theta_{u,0}, r_{u,0}), \ldots, \mathbf{b}_n(\theta_{u,P-1}, r_{u,P-1})]. \tag{9}$$

Let $(\widehat{\boldsymbol{\theta}}, \widehat{\mathbf{r}}) = \{(\widehat{\theta}_j, \widehat{r}_j)\}_{j=0}^{J-1}$ be a two-dimensional angle–range grid and let $\boldsymbol{\varphi}_{u,n} \in \mathbb{C}^{J\times 1}$ be the sparse coefficient vector of user u at BS n. The corresponding dictionary is retained explicitly as

$$\mathbf{B}_n(\widehat{\boldsymbol{\theta}}, \widehat{\mathbf{r}}) \triangleq [\mathbf{b}_n(\widehat{\theta}_0, \widehat{r}_0), \mathbf{b}_n(\widehat{\theta}_1, \widehat{r}_1), \ldots, \mathbf{b}_n(\widehat{\theta}_{J-1}, \widehat{r}_{J-1})]. \tag{10}$$

The on-grid model is

$$\begin{aligned}\mathbf{Y}_n &= [\mathbf{B}_n(\widehat{\boldsymbol{\theta}},\widehat{\mathbf{r}})\boldsymbol{\varphi}_{0,n},\ldots,\mathbf{B}_n(\widehat{\boldsymbol{\theta}},\widehat{\mathbf{r}})\boldsymbol{\varphi}_{U-1,n}]\mathbf{S}+\mathbf{W}_n\\ &= \mathbf{B}_n(\widehat{\boldsymbol{\theta}},\widehat{\mathbf{r}})\boldsymbol{\Phi}_n\mathbf{S}+\mathbf{W}_n,\end{aligned} \tag{11}$$

where $\boldsymbol{\Phi}_n \triangleq [\boldsymbol{\varphi}_{0,n},\ldots,\boldsymbol{\varphi}_{U-1,n}]$. Because the true path parameters are generally off the grid [26], the model is extended as

$$\mathbf{Y}_n = \mathbf{B}_n(\widehat{\boldsymbol{\theta}}+\boldsymbol{\beta}^{\theta},\widehat{\mathbf{r}}+\boldsymbol{\beta}^{r})\boldsymbol{\Phi}_n\mathbf{S}+\mathbf{W}_n, \tag{12}$$

where $\boldsymbol{\beta}^{\theta}$ and $\boldsymbol{\beta}^{r}$ are the angle- and range-grid offsets, respectively. In this skewed off-grid representation [26,27], the two offsets are updated jointly while the exact geometric spherical-wave phase manifold is retained.

3.2. Sparse Bayesian Inference

The following inference procedure adopts the sparse Bayesian learning framework used in Bayesian compressive sensing, variational Bayesian approximation, off-grid direction finding, and fast VBI-based sparse recovery [28–31]. In the present distributed near-field setting, these ideas are adapted to the dictionary constructed from the exact geometric spherical-wave phase response and to the user-wise pilot observation model.

Let ν denote the noise precision. Under complex Gaussian noise, the likelihood is

$$p(\mathbf{Y}_n \mid \boldsymbol{\Phi}_n,\nu;\boldsymbol{\beta}^{\theta},\boldsymbol{\beta}^{r}) = \prod_{t=0}^{T-1}\mathcal{CN}\Big(\mathbf{y}_n(t) \mid \mathbf{B}_n(\widehat{\boldsymbol{\theta}}+\boldsymbol{\beta}^{\theta},\widehat{\mathbf{r}}+\boldsymbol{\beta}^{r})\boldsymbol{\Phi}_n\mathbf{s}(t),\nu^{-1}\mathbf{I}\Big). \tag{13}$$

The noise precision is assigned a Gamma prior [32]

$$p(\nu) = \Gamma(\nu \mid a,b), \tag{14}$$

where a and b are small positive constants [33]. To induce sparsity, the entries of $\boldsymbol{\Phi}_n$ are assigned independent zero-mean complex Gaussian priors controlled by $\boldsymbol{\Xi}$:

$$p(\boldsymbol{\Phi}_n;\boldsymbol{\Xi}) = \prod_{j=0}^{J-1}\prod_{u=0}^{U-1}\mathcal{CN}(\varphi_{j,u,n} \mid 0,\xi_{j,u}^{-1}). \tag{15}$$

The corresponding hyperprior is

$$p(\boldsymbol{\Xi}) = \prod_{j=0}^{J-1}\prod_{u=0}^{U-1}\Gamma(\xi_{j,u} \mid a,b). \tag{16}$$

The joint density across the distributed BSs is therefore

$$p(\{\mathbf{Y}_n\}_{n=0}^{N-1},\{\boldsymbol{\Phi}_n\}_{n=0}^{N-1},\nu,\boldsymbol{\Xi};\boldsymbol{\beta}^{\theta},\boldsymbol{\beta}^{r}) = \left(\prod_{n=0}^{N-1}p(\mathbf{Y}_n \mid \boldsymbol{\Phi}_n,\nu;\boldsymbol{\beta}^{\theta},\boldsymbol{\beta}^{r})p(\boldsymbol{\Phi}_n;\boldsymbol{\Xi})\right)p(\boldsymbol{\Xi})p(\nu). \tag{17}$$

The posterior is approximated using IV-VBI [31] as

$$p(\{\boldsymbol{\Phi}_n\}_{n=0}^{N-1},\nu,\boldsymbol{\Xi} \mid \{\mathbf{Y}_n\}_{n=0}^{N-1};\boldsymbol{\beta}^{\theta},\boldsymbol{\beta}^{r}) \approx \left\{\prod_{n=0}^{N-1}q(\boldsymbol{\Phi}_n)\right\}q(\nu)q(\boldsymbol{\Xi}). \tag{18}$$

The following two identities are retained explicitly because they are used to separate the user-specific terms:

$$\begin{aligned}\mathbf{B}_n(\widehat{\boldsymbol{\theta}}+\boldsymbol{\beta}^{\theta},\widehat{\mathbf{r}}+\boldsymbol{\beta}^{r})\boldsymbol{\Phi}_n\mathbf{S} &= \sum_{u=0}^{U-1}\mathbf{B}_n(\widehat{\boldsymbol{\theta}}+\boldsymbol{\beta}^{\theta},\widehat{\mathbf{r}}+\boldsymbol{\beta}^{r})\boldsymbol{\varphi}_{u,n}\vec{\mathbf{s}}_u,\\ \mathrm{tr}\Big\{\mathbf{B}_n\boldsymbol{\varphi}_{u,n}\vec{\mathbf{s}}_u(\mathbf{B}_n\boldsymbol{\varphi}_{u,n}\vec{\mathbf{s}}_u)^{\mathrm{H}}\Big\} &= \vec{\mathbf{s}}_u\vec{\mathbf{s}}_u^{\mathrm{H}}\,\mathrm{tr}\Big\{\mathbf{B}_n\boldsymbol{\varphi}_{u,n}\boldsymbol{\varphi}_{u,n}^{\mathrm{H}}\mathbf{B}_n^{\mathrm{H}}\Big\},\end{aligned} \tag{19}$$

where $\vec{\mathbf{s}}_u\vec{\mathbf{s}}_u^{\mathrm{H}}$ is a scalar and $\mathbf{B}_n$ abbreviates $\mathbf{B}_n(\widehat{\boldsymbol{\theta}}+\boldsymbol{\beta}^{\theta},\widehat{\mathbf{r}}+\boldsymbol{\beta}^{r})$. Since the user vectors remain coupled through $\mathbf{S}$, the factorization is further written as

$$q(\boldsymbol{\Theta}) = \underbrace{\prod_{n=0}^{N-1}\prod_{u=0}^{U-1}q(\boldsymbol{\varphi}_{u,n})}_{q(\boldsymbol{\Phi}_n)}\,q(\nu)q(\boldsymbol{\Xi}), \tag{20}$$

where $\boldsymbol{\Theta}\triangleq\Big\{\{\boldsymbol{\varphi}_{u,n}\}_{u=0,n=0}^{U-1,N-1},\nu,\boldsymbol{\Xi}\Big\}$ is the set of variables. The optimal factor is obtained by minimizing the Kullback–Leibler divergence:

$$\begin{aligned}q^{\star}(\boldsymbol{\Theta}) &= \arg\min_{q(\boldsymbol{\Theta})} D_{\mathrm{KL}}\Big(q(\boldsymbol{\Theta})\|p(\boldsymbol{\Theta}\mid\{\mathbf{Y}_n\}_{n=0}^{N-1};\boldsymbol{\beta}^{\theta},\boldsymbol{\beta}^{r})\Big)\\ &\triangleq \arg\min_{q(\boldsymbol{\Theta})}\int q(\boldsymbol{\Theta})\ln\frac{q(\boldsymbol{\Theta})}{p(\boldsymbol{\Theta}\mid\{\mathbf{Y}_n\}_{n=0}^{N-1};\boldsymbol{\beta}^{\theta},\boldsymbol{\beta}^{r})}\,d\boldsymbol{\Theta}.\end{aligned} \tag{21}$$

Accordingly, each element $\boldsymbol{\Theta}_k\in\boldsymbol{\Theta}$ satisfies

$$\ln q^{\star}(\boldsymbol{\Theta}_k)\propto\Big\langle\ln p(\mathbf{Y},\boldsymbol{\Theta};\boldsymbol{\beta}^{\theta},\boldsymbol{\beta}^{r})\Big\rangle_{\prod_{j\neq k}q^{\star}(\boldsymbol{\Theta}_j)}. \tag{22}$$

The posterior factors are obtained from this update in three steps.

(i) *Update of* $q(\boldsymbol{\varphi}_{u,n})$: According to the update rule in (22), $q^{\mathrm{new}}(\boldsymbol{\varphi}_{u,n})$ is obtained as follows:

$$\begin{aligned}&\ln q^{\mathrm{new}}(\boldsymbol{\varphi}_{u,n})\\ &\propto\Big\langle\ln p(\{\mathbf{Y}_n\}_{n=0}^{N-1},\{\boldsymbol{\Phi}_n\}_{n=0}^{N-1},\nu,\boldsymbol{\Xi};\boldsymbol{\beta}^{\theta},\boldsymbol{\beta}^{r})\Big\rangle_{\prod_{u',n'\neq u,n}q(\boldsymbol{\varphi}_{u',n'})q(\nu)q(\boldsymbol{\Xi})}\\ &\propto\langle\ln p(\mathbf{Y}_n\mid\boldsymbol{\Phi}_n,\nu)\rangle_{\prod_{u',n'\neq u,n}q(\boldsymbol{\varphi}_{u',n'})q(\nu)}+\langle\ln p(\boldsymbol{\varphi}_{u,n}\mid\boldsymbol{\xi}_u)\rangle_{q(\boldsymbol{\xi}_u)}\\ &\propto-\hat{v}\sum_{t=0}^{T-1}\Big\langle\|\mathbf{y}_n(t)-\mathbf{B}_n\boldsymbol{\Phi}_n\mathbf{s}(t)\|_2^2\Big\rangle_{\prod_{u',n'\neq u,n}q(\boldsymbol{\varphi}_{u',n'})}-\boldsymbol{\varphi}_{u,n}^{\mathrm{H}}\mathrm{diag}(\hat{\boldsymbol{\xi}}_u)\boldsymbol{\varphi}_{u,n}\\ &\propto-\hat{v}\sum_{t=0}^{T-1}\|\hat{\mathbf{y}}_{n,-u}(t)-\mathbf{B}_n\boldsymbol{\varphi}_{u,n}s_u(t)\|_2^2-\boldsymbol{\varphi}_{u,n}^{\mathrm{H}}\mathrm{diag}(\hat{\boldsymbol{\xi}}_u)\boldsymbol{\varphi}_{u,n}\\ &\propto 2\hat{v}\Re(\vec{\mathbf{s}}_u\hat{\mathbf{Y}}_{n,-u}^{\mathrm{H}}\mathbf{B}_n\boldsymbol{\varphi}_{u,n})-\boldsymbol{\varphi}_{u,n}^{\mathrm{H}}\big(\hat{v}\vec{\mathbf{s}}_u\vec{\mathbf{s}}_u^{\mathrm{H}}\mathbf{B}_n^{\mathrm{H}}\mathbf{B}_n+\mathrm{diag}(\hat{\boldsymbol{\xi}}_u)\big)\boldsymbol{\varphi}_{u,n},\end{aligned} \tag{23}$$

where $s_u(t)$ denotes the symbol transmitted by user u at snapshot t, $\vec{\mathbf{s}}_u\triangleq[s_u(0),s_u(1),\ldots,s_u(T-1)]$, $\boldsymbol{\xi}_u\triangleq[\xi_{0,u},\ldots,\xi_{J-1,u}]^{\mathrm{T}}$, and $\hat{\mathbf{Y}}_{n,-u}\triangleq\mathbf{Y}_n-\sum_{k\neq u}\mathbf{B}_n\boldsymbol{\mu}_{k,n}\vec{\mathbf{s}}_k$. The quantities $\hat{v}$ and $\hat{\boldsymbol{\xi}}_u$ are defined in (29) and (32), respectively. The residual $\hat{\mathbf{Y}}_{n,-u}$ removes the contributions of the other users using their current posterior means and can therefore be interpreted as an effective single-user observation for user u. Because both the likelihood and the conditional sparse prior are complex Gaussian, completing the quadratic form yields a complex Gaussian posterior characterized by $\boldsymbol{\mu}_{u,n}$ and $\boldsymbol{\Sigma}_{u,n}$. Consequently, $q^{\mathrm{new}}(\boldsymbol{\varphi}_{u,n})$ follows a complex Gaussian distribution:

$$q^{\mathrm{new}}(\boldsymbol{\varphi}_{u,n})=\mathcal{CN}(\boldsymbol{\varphi}_{u,n}|\boldsymbol{\mu}_{u,n},\boldsymbol{\Sigma}_{u,n}), \tag{24}$$

with mean

$$\boldsymbol{\mu}_{u,n} \triangleq \hat{\nu}\boldsymbol{\Sigma}_{u,n}\mathbf{B}_n^{\mathrm{H}}\hat{\mathbf{Y}}_{n,-u}\vec{\mathbf{s}}_u^{\mathrm{H}}, \tag{25}$$

and covariance matrix

$$\boldsymbol{\Sigma}_{u,n} \triangleq \left(\hat{\nu}\vec{\mathbf{s}}_u\vec{\mathbf{s}}_u^{\mathrm{H}}\mathbf{B}_n^{\mathrm{H}}\mathbf{B}_n + \mathrm{diag}(\hat{\boldsymbol{\xi}}_u)\right)^{-1}. \tag{26}$$

(ii) *Update of $q(\nu)$:* Similarly, $q^{\mathrm{new}}(\nu)$ is obtained as follows:

$$\begin{aligned}
&\ln q^{\mathrm{new}}(\nu) \\
&\quad\propto \left\langle \ln p(\{\mathbf{Y}_n\}_{n=0}^{N-1}, \{\boldsymbol{\Phi}_n\}_{n=0}^{N-1}, \nu, \boldsymbol{\Xi}; \boldsymbol{\beta}^{\theta}, \boldsymbol{\beta}^{r}) \right\rangle_{\prod_{u,n} q(\boldsymbol{\varphi}_{u,n}) q(\boldsymbol{\Xi})} \\
&\quad\propto \sum_{n=0}^{N-1} \langle \ln p(\mathbf{Y}_n \mid \boldsymbol{\Phi}_n, \nu) \rangle_{\prod_{u,n} q(\boldsymbol{\varphi}_{u,n})} + \ln p(\nu) \\
&\quad\propto \underbrace{(TMN + a - 1)}_{\triangleq a_\nu - 1} \ln \nu \\
&\qquad - \nu \underbrace{\left\{ b + \sum_{n=0}^{N-1} \left\| \mathbf{Y}_n - \sum_u \mathbf{B}_n \boldsymbol{\mu}_{u,n} \vec{\mathbf{s}}_u \right\|_{\mathrm{F}}^2 + \sum_{n=0}^{N-1} \sum_u \vec{\mathbf{s}}_u^{*} \vec{\mathbf{s}}_u^{\mathrm{T}} \, \mathrm{tr}\Big(\mathbf{B}_n \boldsymbol{\Sigma}_{u,n} \mathbf{B}_n^{\mathrm{H}} \Big) \right\}}_{\triangleq b_\nu}.
\end{aligned} \tag{27}$$

Therefore, $q^{\mathrm{new}}(\nu)$ follows a Gamma distribution:

$$q^{\mathrm{new}}(\nu) = \Gamma(\nu \mid a_\nu, b_\nu), \tag{28}$$

and its posterior mean is

$$\hat{\nu} = \langle \nu \rangle_{q^{\mathrm{new}}(\nu)} = \frac{a_\nu}{b_\nu}. \tag{29}$$

(iii) *Update of $q(\boldsymbol{\Xi})$:* Finally, $q^{\mathrm{new}}(\boldsymbol{\Xi}) = \prod_{j,u} q^{\mathrm{new}}(\xi_{j,u})$ is obtained elementwise as follows:

$$\begin{aligned}
&\ln q^{\mathrm{new}}(\xi_{j,u}) \\
&\quad\propto \left\langle \ln p(\{\mathbf{Y}_n\}_{n=0}^{N-1}, \{\boldsymbol{\Phi}_n\}_{n=0}^{N-1}, \nu, \boldsymbol{\Xi}; \boldsymbol{\beta}^{\theta}, \boldsymbol{\beta}^{r}) \right\rangle_{\prod_{u,n} q(\boldsymbol{\varphi}_{u,n}) q(\nu)} \\
&\quad\propto \sum_{n=0}^{N-1} \left\langle \ln p(\varphi_{j,u,n} \mid \xi_{j,u}) \right\rangle_{q(\boldsymbol{\varphi}_{u,n})} + \ln p(\xi_{j,u}) \\
&\quad\propto \underbrace{(N + a - 1)}_{\triangleq a_{\xi_{j,u}} - 1} \ln \xi_{j,u} \\
&\qquad - \xi_{j,u} \underbrace{\left\{ b + \sum_{n=0}^{N-1} \left(\left| [\boldsymbol{\mu}_{u,n}]_j \right|^2 + [\boldsymbol{\Sigma}_{u,n}]_{j,j} \right) \right\}}_{\triangleq b_{\xi_{j,u}}}.
\end{aligned} \tag{30}$$

Therefore,

$$q^{\mathrm{new}}(\boldsymbol{\Xi}) = \prod_{j=0}^{J-1} \prod_{u=0}^{U-1} q^{\mathrm{new}}(\xi_{j,u}) = \prod_{j=0}^{J-1} \prod_{u=0}^{U-1} \Gamma(\xi_{j,u} \mid a_{\xi_{j,u}}, b_{\xi_{j,u}}), \tag{31}$$

and the corresponding posterior mean is

$$\hat{\xi}_{j,u} = \langle \xi_{j,u} \rangle_{q^{\text{new}}(\xi_{j,u})} = \frac{a_{\xi_{j,u}}}{b_{\xi_{j,u}}}. \tag{32}$$

3.3. Two-Dimensional Off-Grid Parameter Update

The off-grid vectors are treated as deterministic parameters and updated using EM [34]:

$$\{\boldsymbol{\beta}^{\theta}, \boldsymbol{\beta}^{r}\} = \arg\max_{\boldsymbol{\beta}^{\theta}, \boldsymbol{\beta}^{r}} \sum_{n=0}^{N-1} \Big\langle \ln p(\mathbf{Y}_n \mid \boldsymbol{\Phi}_n, \nu; \boldsymbol{\beta}^{\theta}, \boldsymbol{\beta}^{r}) \Big\rangle_{q(\nu)\prod_{u=0}^{U-1} q(\boldsymbol{\varphi}_{u,n})}. \tag{33}$$

Using the first-order expansion

$$\mathbf{B}_n(\widehat{\boldsymbol{\theta}} + \boldsymbol{\beta}^{\theta}, \widehat{\mathbf{r}} + \boldsymbol{\beta}^{r}) \approx \mathbf{B}_n(\widehat{\boldsymbol{\theta}}, \widehat{\mathbf{r}}) + \mathbf{C}_n^{\theta} \operatorname{diag}(\boldsymbol{\beta}^{\theta}) + \mathbf{C}_n^{r} \operatorname{diag}(\boldsymbol{\beta}^{r}), \tag{34}$$

where

$$\mathbf{C}_n^{\theta} \triangleq \frac{\partial \mathbf{B}_n(\widehat{\boldsymbol{\theta}}, \widehat{\mathbf{r}})}{\partial \widehat{\boldsymbol{\theta}}}, \tag{35}$$

$$\mathbf{C}_n^{r} \triangleq \frac{\partial \mathbf{B}_n(\widehat{\boldsymbol{\theta}}, \widehat{\mathbf{r}})}{\partial \widehat{\mathbf{r}}}, \tag{36}$$

the expected objective has the quadratic form [30]

$$\begin{aligned}
&\sum_{n=0}^{N-1} \Big\langle \ln p(\mathbf{Y}_n \mid \boldsymbol{\Phi}_n, \nu; \boldsymbol{\beta}^{\theta}, \boldsymbol{\beta}^{r}) \Big\rangle_{q(\nu)\prod_{u=0}^{U-1} q(\boldsymbol{\varphi}_{u,n})} \\
&\propto -\sum_{n=0}^{N-1} \left\langle \left\| \mathbf{Y}_n - \Big\{ \mathbf{B}_n(\widehat{\boldsymbol{\theta}}, \widehat{\mathbf{r}}) + \mathbf{C}_n^{\theta} \operatorname{diag}(\boldsymbol{\beta}^{\theta}) + \mathbf{C}_n^{r} \operatorname{diag}(\boldsymbol{\beta}^{r}) \Big\} \sum_{u=0}^{U-1} \boldsymbol{\mu}_{u,n} \vec{\mathbf{s}}_u \right\|_{\mathrm{F}}^2 \right\rangle_{q(\nu)\prod_{u=0}^{U-1} q(\boldsymbol{\varphi}_{u,n})} \\
&\propto -\{(\boldsymbol{\beta}^{\theta})^{\mathrm{T}} \mathbf{P}^{\theta} \boldsymbol{\beta}^{\theta} + (\boldsymbol{\beta}^{r})^{\mathrm{T}} \mathbf{P}^{r} \boldsymbol{\beta}^{r} - 2(\boldsymbol{\eta}^{\theta})^{\mathrm{T}} \boldsymbol{\beta}^{\theta} - 2(\boldsymbol{\eta}^{r})^{\mathrm{T}} \boldsymbol{\beta}^{r} + 2(\boldsymbol{\beta}^{\theta})^{\mathrm{T}} \boldsymbol{\Psi} \boldsymbol{\beta}^{r}\}.
\end{aligned} \tag{37}$$

All auxiliary quantities in (37) are retained below:

$$\mathbf{P}^{\theta} = \sum_{n=0}^{N-1} \Re\left\{ \left\{ (\mathbf{C}_n^{\theta})^{\mathrm{H}} \mathbf{C}_n^{\theta} \right\}^{*} \odot \sum_{u=0}^{U-1} \left\{ \vec{\mathbf{s}}_u^{*} \vec{\mathbf{s}}_u^{\mathrm{T}} (\boldsymbol{\mu}_{u,n} \boldsymbol{\mu}_{u,n}^{\mathrm{H}} + \boldsymbol{\Sigma}_{u,n}) \right\} \right\}, \tag{38}$$

$$\mathbf{P}^{r} = \sum_{n=0}^{N-1} \Re\left\{ \left\{ (\mathbf{C}_n^{r})^{\mathrm{H}} \mathbf{C}_n^{r} \right\}^{*} \odot \sum_{u=0}^{U-1} \left\{ \vec{\mathbf{s}}_u^{*} \vec{\mathbf{s}}_u^{\mathrm{T}} (\boldsymbol{\mu}_{u,n} \boldsymbol{\mu}_{u,n}^{\mathrm{H}} + \boldsymbol{\Sigma}_{u,n}) \right\} \right\}, \tag{39}$$

$$\begin{aligned}
\boldsymbol{\eta}^{\theta} = \sum_{n=0}^{N-1} \sum_{u=0}^{U-1} \Bigg\{ &- \vec{\mathbf{s}}_u^{*} \vec{\mathbf{s}}_u^{\mathrm{T}} \Re\Big\{ \operatorname{diag}\Big((\mathbf{C}_n^{\theta})^{\mathrm{H}} \mathbf{B}_n(\widehat{\boldsymbol{\theta}}, \widehat{\mathbf{r}}) \boldsymbol{\Sigma}_{u,n} \Big) \Big\} \\
&+ \Re\Big\{ (\operatorname{diag}(\boldsymbol{\mu}_{u,n}))^{*} (\mathbf{C}_n^{\theta})^{\mathrm{H}} \Big(\hat{\mathbf{Y}}_{n,-u} - \mathbf{B}_n(\widehat{\boldsymbol{\theta}}, \widehat{\mathbf{r}}) \boldsymbol{\mu}_{u,n} \vec{\mathbf{s}}_u \Big) \vec{\mathbf{s}}_u^{\mathrm{H}} \Big\} \Bigg\},
\end{aligned} \tag{40}$$

$$\begin{aligned}
\boldsymbol{\eta}^{r} = \sum_{n=0}^{N-1} \sum_{u=0}^{U-1} \Bigg\{ &- \vec{\mathbf{s}}_u^{*} \vec{\mathbf{s}}_u^{\mathrm{T}} \Re\Big\{ \operatorname{diag}\Big((\mathbf{C}_n^{r})^{\mathrm{H}} \mathbf{B}_n(\widehat{\boldsymbol{\theta}}, \widehat{\mathbf{r}}) \boldsymbol{\Sigma}_{u,n} \Big) \Big\} \\
&+ \Re\Big\{ (\operatorname{diag}(\boldsymbol{\mu}_{u,n}))^{*} (\mathbf{C}_n^{r})^{\mathrm{H}} \Big(\hat{\mathbf{Y}}_{n,-u} - \mathbf{B}_n(\widehat{\boldsymbol{\theta}}, \widehat{\mathbf{r}}) \boldsymbol{\mu}_{u,n} \vec{\mathbf{s}}_u \Big) \vec{\mathbf{s}}_u^{\mathrm{H}} \Big\} \Bigg\},
\end{aligned} \tag{41}$$

$$\boldsymbol{\Psi} = \sum_{n=0}^{N-1} \Re\left\{ \left\{ (\mathbf{C}_n^{\theta})^{\mathrm{H}} \mathbf{C}_n^{r} \right\}^{*} \odot \sum_{u=0}^{U-1} \left\{ \vec{\mathbf{s}}_u^{*} \vec{\mathbf{s}}_u^{\mathrm{T}} (\boldsymbol{\mu}_{u,n} \boldsymbol{\mu}_{u,n}^{\mathrm{H}} + \boldsymbol{\Sigma}_{u,n}) \right\} \right\}. \tag{42}$$

In (37), $\mathbf{P}^{\theta}$ and $\mathbf{P}^{r}$ characterize the quadratic curvature of the expected objective with respect to the angle and range offsets, respectively. The vectors $\boldsymbol{\eta}^{\theta}$ and $\boldsymbol{\eta}^{r}$ provide the corresponding linear update directions, while $\boldsymbol{\Psi}$ captures the coupling between the two

types of offsets. Consequently, fixing one offset vector reduces the objective to a quadratic function of the other, leading to the alternating closed-form updates below. Here, $\Re\{\cdot\}$, $(\cdot)^*$, and $\odot$ denote the real part, complex conjugation, and Hadamard product, respectively. Setting the two gradients to zero gives the alternating closed-form updates

$$(\boldsymbol{\beta}^{\theta})^{\text{new}} = (\mathbf{P}^{\theta})^{\dagger}(\boldsymbol{\eta}^{\theta} - \boldsymbol{\Psi}\boldsymbol{\beta}^{r}), \tag{43}$$

$$(\boldsymbol{\beta}^{r})^{\text{new}} = (\mathbf{P}^{r})^{\dagger}(\boldsymbol{\eta}^{r} - \boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{\beta}^{\theta}). \tag{44}$$

Finally, the sampling grid is shifted using

$$\widehat{\boldsymbol{\theta}}^{\text{new}} = \widehat{\boldsymbol{\theta}} + (\boldsymbol{\beta}^{\theta})^{\text{new}}, \tag{45}$$

$$\widehat{\mathbf{r}}^{\text{new}} = \widehat{\mathbf{r}} + (\boldsymbol{\beta}^{r})^{\text{new}}. \tag{46}$$

After sufficient VBI-EM iterations, the active grid points approach the true angle–range parameters [35], thereby mitigating two-dimensional basis mismatch without changing the original channel-variable definitions.

3.4. Algorithm Summary and Computational Remarks

The complete estimator alternates between probabilistic sparse recovery and deterministic off-grid refinement. The VBI part updates the posterior mean and covariance of each user-specific sparse vector while the remaining users are represented by their current posterior means. The EM part then uses these posterior statistics to refine the angle and range offsets of the active grid points. Using the variables defined above, the implementation is summarized in Algorithm 1.

Algorithm 1 Off-grid independent-vector VBI for distributed near-field channel estimation

1. **Input:** received pilot matrices $\{\mathbf{Y}_n\}_{n=0}^{N-1}$, pilot matrix $\mathbf{S}$, BS spacing R, antenna spacing Δ, wavelength λ, and the initial angle–range grid $(\widehat{\boldsymbol{\theta}}, \widehat{\mathbf{r}})$.
2. Construct $\mathbf{B}_n(\widehat{\boldsymbol{\theta}}, \widehat{\mathbf{r}})$ for all BSs using (6) and (7). Initialize $\boldsymbol{\beta}^{\theta} = \mathbf{0}$, $\boldsymbol{\beta}^{r} = \mathbf{0}$, $\widehat{\nu}$, and $\widehat{\boldsymbol{\Xi}}_u$.
3. For each BS n and user u, form $\hat{\mathbf{Y}}_{n,-u}$ and then update $\boldsymbol{\Sigma}_{u,n}$ and $\boldsymbol{\mu}_{u,n}$ using the expressions above.
4. Update a_{ν}, b_{ν}, $\widehat{\nu}$, $a_{\xi_{j,u}}$, $b_{\xi_{j,u}}$, and $\widehat{\xi}_{j,u}$ using the closed-form expressions above.
5. With the current posterior statistics fixed, compute $\mathbf{P}^{\theta}$, $\mathbf{P}^{r}$, $\boldsymbol{\eta}^{\theta}$, $\boldsymbol{\eta}^{r}$, and $\boldsymbol{\Psi}$, and update $\boldsymbol{\beta}^{\theta}$ and $\boldsymbol{\beta}^{r}$ using the closed-form offset equations above. Then shift the grid using (45) and (46).
6. Rebuild the dictionaries from the exact geometric spherical-wave phase response at the shifted grid locations and repeat Steps 3–5 until the relative change in the reconstructed channels is below the convergence threshold $\epsilon = 10^{-4}$ or the maximum number of iterations $I_{\max} = 500$ is reached.
7. **Output:** after absorbing the final offsets into the shifted grid, reconstruct the channels as $\widehat{\mathbf{H}}_n = \mathbf{B}_n(\widehat{\boldsymbol{\theta}}, \widehat{\mathbf{r}})\widehat{\boldsymbol{\Phi}}_n$ for $n = 0, \ldots, N-1$.

The main computational burden comes from updating the covariance matrix $\boldsymbol{\Sigma}_{u,n}$, so the dominant per-iteration cost can be described as $\mathcal{O}(NUJ^3)$. As suggested by the performance–complexity perspective in [36], estimation accuracy should be evaluated together with computational burden. We acknowledge that this complexity can be high when the number of BSs, users, or angle–range atoms becomes large. Moreover, following the GAMP extension of real-valued sparse Bayesian learning in [35], it is possible to derive a substantially faster iterative implementation by combining the proposed geometry-aware framework with approximate message passing. Such a reformulation requires a separate algorithmic derivation and convergence analysis and is therefore beyond the scope of this

article. The iterations are initialized from the on-grid dictionary, so the method does not require prior knowledge of the true support. The current formulation also assumes that each dominant path has common support across the participating BSs. It therefore does not accommodate BS-specific path visibility caused by blockage, local scattering, or spatial non-stationarity; extending the model to partial path visibility is left for future work.

4. Numerical Results

4.1. Simulation Setup

We compare the proposed method with the following benchmark methods: All reported numerical results are obtained from Monte Carlo simulations, and no over-the-air or field-measurement data are used. (i) LS: the direct estimator applied to the pilot-based model [24]; (ii) Genie-aided LS: an optimistic lower-bound benchmark with prior knowledge of the true angle–range support; (iii) two-stage MUSIC: a spatial-spectrum method that first estimates local far-field angles and then infers ranges from the BS geometry [21,25]; (iv) far-field VBI: a Bayesian inference framework that uses a conventional plane-wave angular dictionary [37]; and (v) OMP: a classical greedy algorithm applied to a discretized angle–range dictionary [10]. The path angle and range are drawn uniformly from $[-60^\circ, 60^\circ]$ and $[15, 25]$ m, respectively. The carrier frequency is 30 GHz, corresponding to $\lambda = 0.01$ m, and the antenna spacing is $\Delta = \lambda/2$. Table 1 summarizes the default simulation settings.

All methods are evaluated using the same received pilot matrices and channel realizations in each Monte Carlo trial. Genie-aided LS is included only as an ideal support-aware reference; consequently, its gap from the proposed method reflects the combined effects of support recovery, off-grid refinement, and finite-pilot interference.

Gaussian pilots do not perfectly separate users in the pilot domain, so the estimators must handle both noise and pilot interference. With orthogonal pilots, the user sequences are better conditioned and residual multiuser coupling is reduced. Both cases are relevant because practical distributed MIMO systems may operate with limited training resources, especially when the coherence interval is short or the number of active users is large. The independent and identically distributed (i.i.d.) Gaussian pilot entries are generated from a zero-mean unit-variance complex Gaussian distribution. The orthogonal pilots are obtained by applying QR orthogonalization to an independently generated complex Gaussian matrix. In both cases, each user's pilot sequence is scaled to have unit average symbol power.

Table 1. Default simulation parameters. In each figure, the swept variable takes the values shown on the horizontal axis.

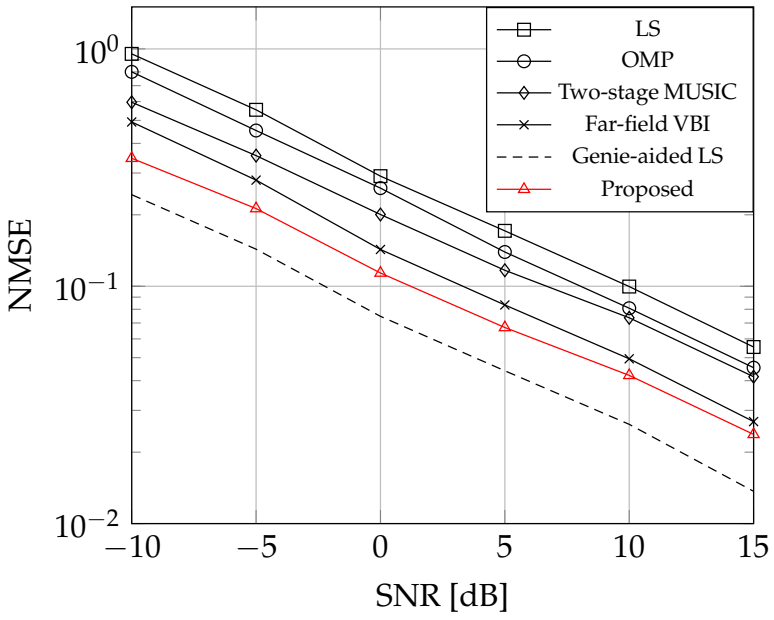
Parameter	Value	Meaning
N	3	Number of BSs
M	15	Number of antennas per BS
U	5	Number of users
T	20	Number of pilot snapshots
P	1	Number of paths per user
R	50 m	BS spacing
I	200	Number of Monte Carlo trials
f	30 GHz	Carrier frequency
λ	0.01 m	Wavelength
Δ	$\lambda/2$	Antenna spacing
θ	$\mathcal{U}[-60^\circ, 60^\circ]$	Path-angle distribution
r	$\mathcal{U}[15, 25]$ m	Path-range distribution
$\mathbf{S}$	i.i.d. complex Gaussian or orthogonal	Pilot–matrix design

Channel-estimation performance is measured using the normalized mean-squared error (NMSE), defined as

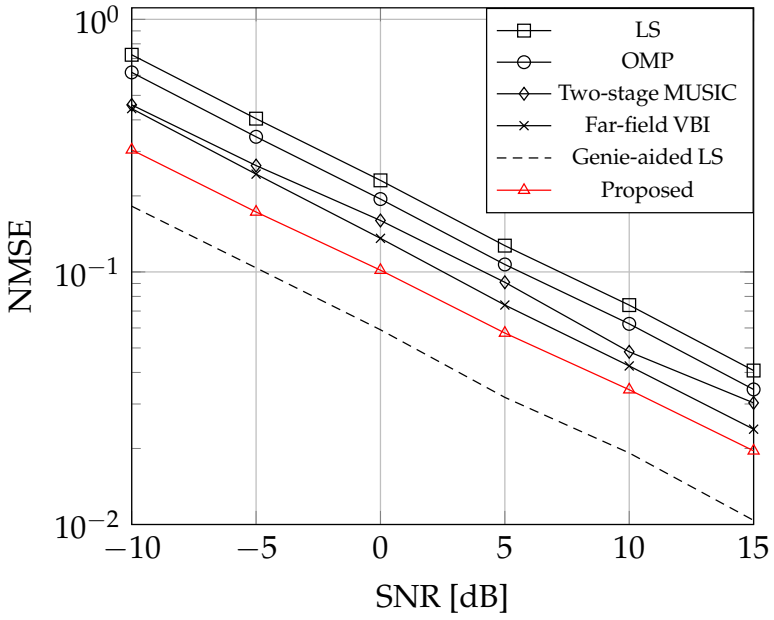
$$\frac{1}{I}\sum_{i=1}^{I}\frac{\sum_{n=0}^{N-1}\|\widehat{\mathbf{H}}_n(i)-\mathbf{H}_n(i)\|_{\mathrm{F}}^2}{\sum_{n=0}^{N-1}\|\mathbf{H}_n(i)\|_{\mathrm{F}}^2}, \tag{47}$$

where $I = 200$ is the number of Monte Carlo trials, $\widehat{\mathbf{H}}_n(i)$ is the estimated channel matrix for BS n in the ith trial, and $\mathbf{H}_n(i)$ is the corresponding true channel matrix.

For a direct side-by-side comparison, Table 2 reports the NMSE at SNR $= 0$ dB under the default i.i.d. Gaussian-pilot setting used in Figure 2a, together with the dominant computational operation and the principal robustness characteristic or limitation of each method. The NMSE values are therefore obtained under the same simulation conditions rather than being taken from studies with different system configurations.



(**a**)



(**b**)

Figure 2. NMSE versus SNR for (**a**) i.i.d. Gaussian pilots and (**b**) orthogonal pilots, with $N = 3$, $M = 15$, $U = 5$, $T = 20$, $P = 1$, and $R = 50$ m.

Table 2. Comparison of NMSE, computational characteristics, and robustness under the default i.i.d. Gaussian-pilot setting at SNR = 0 dB.

Method	Reference	NMSE	Relative Complexity and Dominant Operation	Robustness or Principal Limitation
LS	[24]	0.2910	Low; pilot–matrix pseudoinverse	Does not exploit sparsity and is sensitive to pilot conditioning
OMP	[10]	0.2591	Moderate; iterative search over a two-dimensional dictionary	Sensitive to grid mismatch and incorrect hard-support selection
Two-stage MUSIC	[21,25]	0.2006	Moderate–high; covariance eigendecomposition and spectral search	Sequential angle–range processing can propagate angular errors
Far-field VBI	[37]	0.1429	High; iterative Bayesian posterior-covariance updates	Learns sparse support but retains near-field model mismatch
Proposed	This work	0.1137	High; dominant cost $\mathcal{O}(NUJ^3)$	Mitigates basis mismatch but assumes known geometry and common support
Genie-aided LS	Ideal reference	0.0747	Support-aware LS computation	Requires the true angle–range support and is not realizable

4.2. Impact of SNR

Figure 2 compares NMSE versus SNR for independent Gaussian and orthogonal pilot matrices. In both subfigures, the NMSE of every method decreases almost monotonically as the SNR increases, confirming that noise is the dominant impairment in the low-SNR region. The trend of the proposed method is nearly parallel to that of Genie-aided LS, and the proposed method remains the best practical estimator over the entire range from −10 to 15 dB.

The gap between the proposed method and Far-field VBI demonstrates the importance of retaining the exact distance $d_{u,p,n,m}$ in (6). With a plane-wave dictionary, a single near-field path spreads over several angular atoms. Increasing the SNR cannot eliminate this deterministic model mismatch, which explains the persistent separation between the two curves at medium and high SNRs. The two-stage MUSIC estimator is further affected by error propagation: an angular error produced in the first stage changes the range inferred geometrically in the second stage. OMP makes a hard atom selection at each iteration and is therefore more sensitive than VBI to the coupled angle–range grid.

A comparison of Figure 2a,b shows that every algorithm achieves a lower NMSE with orthogonal pilots than with Gaussian pilots. Orthogonal pilots eliminate interuser pilot interference and thereby improve channel-estimation accuracy. Genie-aided LS achieves the lowest NMSE because it is supplied with the true angle and range; it should therefore be interpreted as an ideal reference.

4.3. Impact of Snapshots

Figure 3 shows the NMSE as the number of snapshots T increases from 10 to 60. The remaining simulation parameters are $N = 3$, $M = 15$, $U = 5$, SNR = 0 dB, $P = 1$, and $R = 50$ m. All curves decrease rapidly at first and then flatten, indicating diminishing returns once sufficient observations are available to identify the dominant support. Increasing T improves the conditioning of the pilot correlations $\mathbf{s}_u^* \mathbf{s}_u^{\mathrm{T}}$ in the covariance update above and provides more observations for inference.

Even when T is small, the proposed method maintains high estimation accuracy. Except for the Genie-aided LS benchmark, which relies on perfect prior knowledge of the angles and ranges, the NMSE curve of the proposed method remains below those of

all other baseline schemes over the entire range of snapshot counts. Its performance is therefore second only to the ideal lower bound.

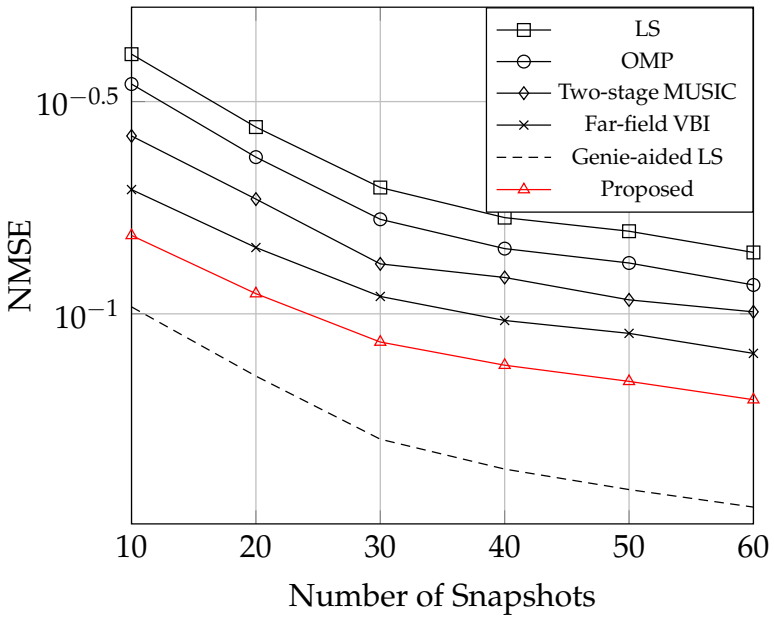


Figure 3. NMSE versus the number of snapshots, with $N = 3$, $M = 15$, $U = 5$, $P = 1$, $R = 50$ m, and SNR $= 0$ dB.

4.4. Impact of the Number of Users

Figure 4 compares the NMSE of the considered algorithms as the number of users U varies from 4 to 8. This experiment uses an i.i.d. Gaussian pilot matrix $\mathbf{S}$. The remaining simulation parameters are $N = 3$, $M = 15$, $T = 20$, SNR $= 0$ dB, $P = 1$, and $R = 50$ m.

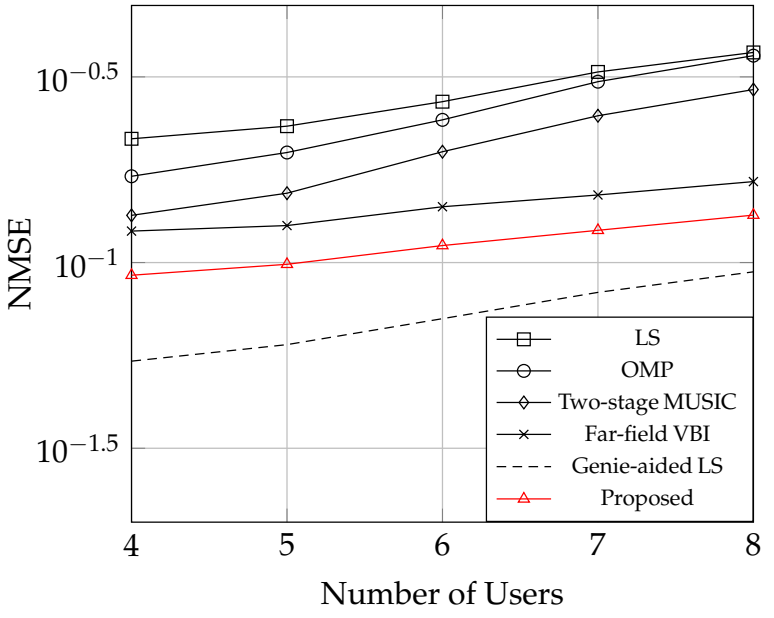


Figure 4. NMSE versus the number of users, with $N = 3$, $M = 15$, $T = 20$, $P = 1$, $R = 50$ m, and SNR $= 0$ dB.

The simulation results demonstrate that as the number of users U increases, the NMSE of all algorithms exhibits an upward trend. This is because a larger number of users exacerbates multiuser interference and increases the dimension of the channel matrix, which increases the difficulty of parameter estimation. Nevertheless, when the number of users is

small, all methods achieve satisfactory estimation performance. As U increases, the NMSE of the proposed method increases at the slowest rate and remains significantly lower than those of the LS, two-stage MUSIC, far-field VBI, and OMP algorithms, demonstrating its robustness against multiuser interference. Furthermore, the Genie-aided LS benchmark, which assumes that the true angles and ranges are perfectly known, yields the lowest NMSE across all user counts, with the performance of the proposed method closely following it.

4.5. Impact of the Number of Antennas

Figure 5 compares the NMSE of the considered algorithms as the number of antennas M varies. This experiment uses an i.i.d. complex Gaussian pilot matrix $\mathbf{S}$. The remaining simulation parameters are $N = 3$, $U = 5$, $T = 20$, SNR $= 0$ dB, $P = 1$, and $R = 50$ m.

The numerical results show that increasing the antenna array size M reduces the NMSE of all schemes except the conventional LS estimator. This improvement occurs chiefly because a larger physical aperture provides sharper spatial resolution and greater array gain, thereby facilitating more accurate channel-parameter estimation. The LS method does not achieve the same improvement because it cannot exploit the inherent angular sparsity. The proposed estimator consistently outperforms the LS, two-stage MUSIC, far-field VBI, and OMP benchmarks and yields the lowest NMSE among the practical methods. The Genie-aided LS scheme, which benefits from perfect prior knowledge of the propagation angles and distances, establishes the performance floor across all antenna configurations.

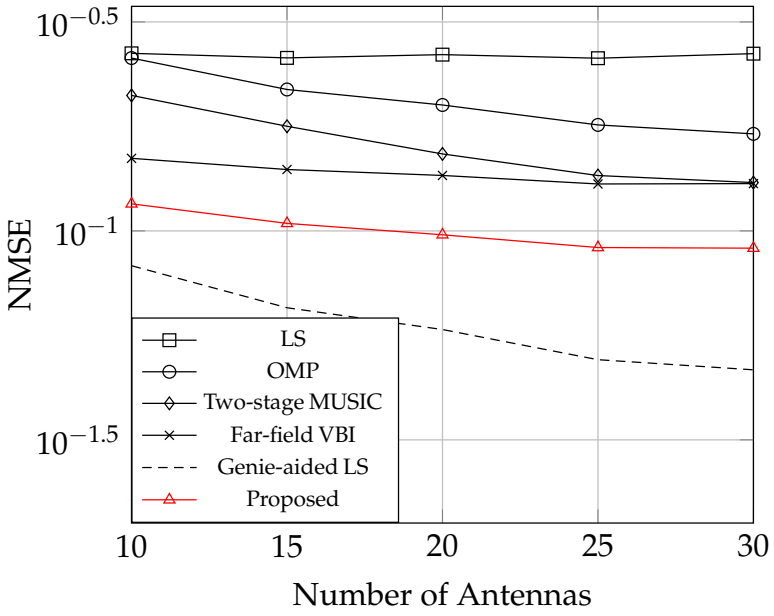


Figure 5. NMSE versus the number of antennas, with $N = 3$, $T = 20$, $U = 5$, $P = 1$, $R = 50$ m, and SNR $= 0$ dB.

4.6. Impact of the Number of BSs

Figure 6 evaluates the NMSE as the number of cooperating BSs N increases from 2 to 6. The experiment uses an i.i.d. Gaussian pilot matrix and 200 independent Monte Carlo trials. The remaining parameters are $M = 15$, $U = 5$, $T = 20$, $P = 1$, $R = 50$ m, and SNR $= -5$ dB.

The NMSE of the proposed method decreases monotonically from approximately 0.209 at $N = 2$ to 0.159 at $N = 6$ and remains lower than those of all practical baselines. Increasing the number of BSs provides additional spatially distributed observations of the common angle–range support, enabling the proposed method to exploit both observation diversity

and the known inter-BS geometry. Two-stage MUSIC also benefits from the additional BSs, but its sequential angle–range processing remains sensitive to angular errors. In contrast, the improvements of LS, OMP, and far-field VBI are comparatively limited because they do not fully exploit the common geometry-coupled near-field structure. Genie-aided LS remains the lower-bound benchmark because it is supplied with the true support.

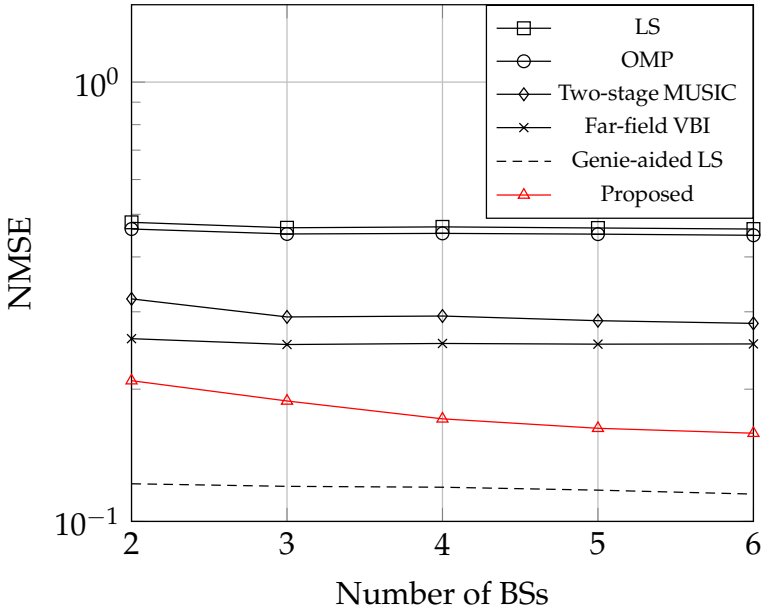


Figure 6. NMSE versus the number of BSs, with $M = 15$, $T = 20$, $U = 5$, $P = 1$, $R = 50$ m, and SNR $= -5$ dB.

4.7. Robustness to Carrier-Frequency Offset

Figure 7 evaluates robustness to a common carrier-frequency offset (CFO) ranging from 0 to 10 kHz. The CFO is introduced across the pilot snapshots through the phase rotation $\exp(j2\pi\Delta f t T_s)$, where T_s corresponds to a 20 MHz bandwidth. The experiment uses an i.i.d. Gaussian pilot matrix and 200 independent Monte Carlo trials. The remaining parameters are $N = 3$, $M = 15$, $U = 5$, $T = 20$, $P = 1$, $R = 50$ m, and SNR $= 5$ dB.

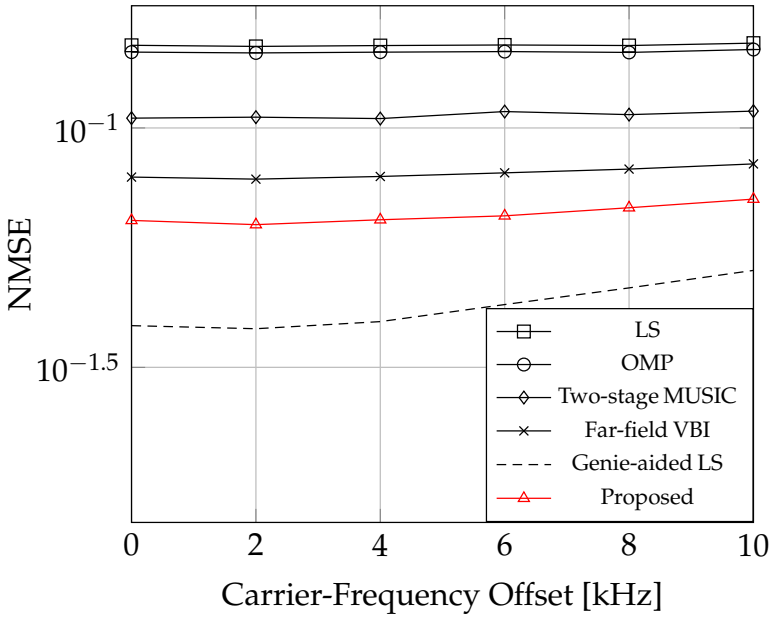


Figure 7. NMSE versus the common carrier-frequency offset, with $N = 3$, $M = 15$, $T = 20$, $U = 5$, $P = 1$, $R = 50$ m, and SNR $= 5$ dB.

The proposed method exhibits only moderate degradation over the tested CFO range: its NMSE changes from approximately 0.0641 without CFO to 0.0710 at 10 kHz and remains below those of all practical baselines. Far-field VBI shows a similar but slightly larger relative increase, whereas LS, OMP, and two-stage MUSIC remain nearly flat at higher NMSE levels. The genie-aided benchmark also degrades because the unmodeled CFO perturbs the pilot-domain observations even when the true angle–range support is known. These results demonstrate robustness to the considered common CFO values, but they do not cover BS-location, array-orientation, or phase/calibration errors, which require separate mismatch models and are left for future work.

5. Conclusions

This paper investigated near-field channel estimation for the considered distributed MIMO geometry with equally spaced, collinear BS reference points and aligned ULAs with common inter-element spacing. By exploiting the known inter-BS geometry, the local direction–range parameters observed by spatially separated BSs were mapped into a common reference coordinate system. This mapping, together with the exact geometric spherical-wave phase response, enabled the distributed channels to be represented by a two-dimensional jointly sparse angle–range model. Based on this representation, an independent-vector VBI algorithm was developed to perform user-specific posterior updates directly from the received pilot matrices, thereby avoiding pilot–matrix inversion and the associated alteration of the noise statistics. A two-dimensional skewed off-grid update was further incorporated into an EM procedure to jointly refine the angle and range offsets and mitigate basis mismatch. All reported numerical results were obtained from Monte Carlo simulations, and no over-the-air or field-measurement data were used. Under the default i.i.d. Gaussian-pilot setting at SNR $= 0$ dB, the proposed estimator achieved an NMSE of 0.1137, compared with 0.1429 for far-field VBI, 0.2006 for two-stage MUSIC, 0.2591 for OMP, and 0.2910 for LS; the ideal Genie-aided LS value was 0.0747. When the number of BSs increased from $N = 2$ to $N = 6$ at SNR $= -5$ dB, the proposed method's NMSE decreased from approximately 0.209 to 0.159. Under a common CFO increasing from 0 to 10 kHz at SNR $= 5$ dB, its NMSE changed from approximately 0.0641 to 0.0710. These results support the effectiveness of combining geometry-aware near-field modeling, Bayesian sparse recovery, and joint off-grid refinement in the evaluated aligned distributed-MIMO scenarios. Future work will consider more general BS geometries, multipath and spatially non-stationary propagation, and lower-complexity implementations for large angle–range dictionaries.

Author Contributions: Conceptualization, L.H.; methodology, L.H.; software, L.H.; validation, L.H.; formal analysis, L.H.; resources, L.H.; data curation, L.H.; writing—original draft preparation, L.H.; writing—review and editing, L.H., Q.G., X.G., H.Y. and Y.X.; supervision, Y.X.; project administration, L.H.; funding acquisition, H.Y. and Y.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science and Technology Project of State Grid Xinjiang Electric Power Co., Ltd., with the grant number DQ30DK26Z005.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data and code supporting the findings of this study are available from the corresponding author upon reasonable request. The data and code are not publicly available due to institutional and intellectual property restrictions associated with the funded project.

Conflicts of Interest: All authors were employed by the State Grid Xinjiang Electric Power Research Institute. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Sanguinetti, L.; Björnson, E.; Hoydis, J. Toward Massive MIMO 2.0: Understanding Spatial Correlation, Interference Suppression, and Pilot Contamination. *IEEE Trans. Commun.* **2020**, *68*, 232–257. [CrossRef]
2. Din, I.U.; Khan, D.; Ullah, A.; Ouameur, M.A.; Razampoosh, B. A Compact Four-Element Multiple-Input Multiple-Output Array with an Integrated Frequency Selective Surface for Millimeter-Wave Applications. *Telecom* **2025**, *6*, 73. [CrossRef]
3. Gesbert, D.; Hanly, S.V.; Huang, H.; Shamai, S.; Simeone, O.; Yu, W. Multi-Cell MIMO Cooperative Networks: A New Look at Interference. *IEEE J. Sel. Areas Commun.* **2010**, *28*, 1380–1408. [CrossRef]
4. Onireti, O.; Héliot, F.; Imran, M.A. On the Energy Efficiency–Spectral Efficiency Trade-Off of Distributed MIMO Systems. *IEEE Trans. Commun.* **2013**, *61*, 3741–3753. [CrossRef]
5. Yazar, A. Machine-Learning-Based Adaptive Wireless Network Selection for Terrestrial and Non-Terrestrial Networks in 5G and Beyond. *Telecom* **2025**, *6*, 71. [CrossRef]
6. Cui, M.; Wu, Z.; Lu, Y.; Wei, X.; Dai, L. Near-Field MIMO Communications for 6G: Fundamentals, Challenges, Potentials, and Future Directions. *IEEE Commun. Mag.* **2023**, *61*, 40–46. [CrossRef]
7. Ouyang, C.; Wang, Z.; Chen, Y.; Mu, X.; Zhu, P. A Primer on Near-Field Communications for Next-Generation Multiple Access. *Proc. IEEE* **2024**, *112*, 1527–1565. [CrossRef]
8. Lu, H.; Zeng, Y.; You, C.; Han, Y.; Zhang, J.; Wang, Z.; Dong, Z.; Jin, S.; Wang, C.X.; Jiang, T.; et al. A Tutorial on Near-Field XL-MIMO Communications Toward 6G. *IEEE Commun. Surv. Tutor.* **2024**, *26*, 2213–2257. [CrossRef]
9. Dai, J.; Liu, A.; Lau, V.K.N. FDD Massive MIMO Channel Estimation with Arbitrary 2D-Array Geometry. *IEEE Trans. Signal Process.* **2018**, *66*, 2584–2599. [CrossRef]
10. Cui, M.; Dai, L. Channel Estimation for Extremely Large-Scale MIMO: Far-Field or Near-Field? *IEEE Trans. Commun.* **2022**, *70*, 2663–2677. [CrossRef]
11. Huang, C.; Xu, J.; Xu, W.; You, X.; Yuen, C.; Chen, Y. Low-Complexity Channel Estimation for Extremely Large-Scale MIMO in Near Field. *IEEE Wirel. Commun. Lett.* **2024**, *13*, 671–675. [CrossRef]
12. Pan, Y.; Pan, C.; Jin, S.; Wang, J. RIS-Aided Near-Field Localization and Channel Estimation for the Terahertz System. *IEEE J. Sel. Top. Signal Process.* **2023**, *17*, 878–892. [CrossRef]
13. Hsu, Y.S.; Wong, K.T.; Yeh, L. Mismatch of Near-Field Bearing–Range Spatial Geometry in Source Localization by a Uniform Linear Array. *IEEE Trans. Antennas Propag.* **2011**, *59*, 3658–3667. [CrossRef]
14. Li, Y.; Zhou, L.; Fu, H.; Dai, J. Decoupled Message Passing Approach for Exact Mixed-Field Sources Localization. *IEEE Sens. J.* **2024**, *24*, 25986–25997. [CrossRef]
15. Cao, H.; Dai, J.; Jiang, X.Q.; Xu, W.; Zhou, B. Near-Field Channel Estimation for RIS-Aided Industrial IoT Systems. *IEEE Internet Things J.* **2026**, *13*, 21440–21453. [CrossRef]
16. Han, Y.; Jin, S.; Wen, C.K.; Ma, X. Channel Estimation for Extremely Large-Scale Massive MIMO Systems. *IEEE Wirel. Commun. Lett.* **2020**, *9*, 633–637. [CrossRef]
17. Qu, K.; Guo, S.; Ye, J.; Zhao, H. Two-Stage Beamspace MUSIC-Based Near-Field Channel Estimation for Hybrid XL-MIMO. *IEEE Commun. Lett.* **2024**, *28*, 1949–1953. [CrossRef]
18. Zhang, X.; Zhang, H.; Eldar, Y.C. Near-Field Sparse Channel Representation and Estimation in 6G Wireless Communications. *IEEE Trans. Commun.* **2024**, *72*, 450–464. [CrossRef]
19. Guo, X.; Chen, Y.; Wang, Y. Compressed Channel Estimation for Near-Field XL-MIMO Using Triple Parametric Decomposition. *IEEE Trans. Veh. Technol.* **2023**, *72*, 15040–15045. [CrossRef]
20. Yang, W.; Li, M.; Liu, Q. A Practical Channel Estimation Strategy for XL-MIMO Communication Systems. *IEEE Commun. Lett.* **2023**, *27*, 1580–1583. [CrossRef]
21. Qiao, L.; Liao, A.; Li, Z.; Wang, H.; Gao, Z.; Gao, X.; Su, Y.; Xiao, P.; You, L.; Ng, D.W.K. Sensing User's Activity, Channel, and Location with Near-Field Extra-Large-Scale MIMO. *IEEE Trans. Commun.* **2024**, *72*, 890–906. [CrossRef]
22. Tang, A.; Wang, J.B.; Pan, Y.; Zhang, W.; Chen, Y.; Yu, H.; de Lamare, R.C. Spatially Non-Stationary XL-MIMO Channel Estimation: A Three-Layer Generalized Approximate Message Passing Method. *IEEE Trans. Signal Process.* **2025**, *73*, 356–371. [CrossRef]
23. Jin, Z.; You, L.; Ng, D.W.K.; Xia, X.G.; Gao, X. Near-Field Channel Estimation for XL-MIMO: A Deep Generative Model Guided by Side Information. *IEEE Trans. Cogn. Commun. Netw.* **2026**, *12*, 628–643. [CrossRef]

24. Biguesh, M.; Gershman, A.B. Training-Based MIMO Channel Estimation: A Study of Estimator Tradeoffs and Optimal Training Signals. *IEEE Trans. Signal Process.* **2006**, *54*, 884–893. [CrossRef]
25. Xia, L.; Wang, R.; Yuan, X.; Teng, B.; Rodríguez-Piñeiro, J. Near-Field Localization for Reconfigurable Intelligent Surface Aided XL-MIMO Systems Harnessing the NLoS Components. *IEEE Trans. Wirel. Commun.* **2026**, *25*, 1857–1874. [CrossRef]
26. Liu, D.; Zhao, Y. Real-Valued Sparse Bayesian Learning Algorithm for Off-Grid DOA Estimation in the Beamspace. *Digit. Signal Process.* **2022**, *121*, 103322. [CrossRef]
27. Cao, Z.; Zhou, L.; Dai, J. Sparse Bayesian Approach for DOD and DOA Estimation with Bistatic MIMO Radar. *IEEE Access* **2019**, *7*, 155335–155346. [CrossRef]
28. Ji, S.; Xue, Y.; Carin, L. Bayesian Compressive Sensing. *IEEE Trans. Signal Process.* **2008**, *56*, 2346–2356. [CrossRef]
29. Tzikas, D.G.; Likas, A.C.; Galatsanos, N.P. The Variational Approximation for Bayesian Inference. *IEEE Signal Process. Mag.* **2008**, *25*, 131–146. [CrossRef]
30. Yang, Z.; Xie, L.; Zhang, C. Off-Grid Direction of Arrival Estimation Using Sparse Bayesian Inference. *IEEE Trans. Signal Process.* **2013**, *61*, 38–43. [CrossRef]
31. Cao, Z.; Dai, J.; Xu, W.; Chang, C. Fast Variational Bayesian Inference for Temporally Correlated Sparse Signal Recovery. *IEEE Signal Process. Lett.* **2021**, *28*, 214–218. [CrossRef]
32. Choo, Y.; Yang, H. Broadband Off-Grid DOA Estimation Using Block Sparse Bayesian Learning for Nonuniform Noise Variance. *IEEE J. Ocean. Eng.* **2022**, *47*, 1024–1040. [CrossRef]
33. Karimi, M.; Zare, M.; Derakhtian, M. Sparse Bayesian Learning with Jeffreys' Non-Informative Prior for Off-Grid DOA Estimation. *Signal Process.* **2025**, *230*, 109809. [CrossRef]
34. Zou, Q.; Zhang, H.; Yang, H. Expectation-Maximization-Aided Hybrid Generalized Expectation Consistent for Sparse Signal Reconstruction. *IEEE Signal Process. Lett.* **2021**, *28*, 648–652. [CrossRef]
35. Dai, J.; So, H.C. Real-Valued Sparse Bayesian Learning for DOA Estimation with Arbitrary Linear Arrays. *IEEE Trans. Signal Process.* **2021**, *69*, 4977–4990. [CrossRef]
36. Miuccio, L.; Panno, D.; Riolo, S. A Flexible Encoding/Decoding Procedure for 6G SCMA Wireless Networks via Adversarial Machine Learning Techniques. *IEEE Trans. Veh. Technol.* **2023**, *72*, 3288–3303. [CrossRef]
37. Ji, C.; Dai, J.; Fu, H.; Xu, W. Joint Device Activity Detection and Channel Estimation for mMTC. *IEEE Internet Things J.* **2024**, *11*, 27232–27244. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.