Integrated Interpolation and Block-term Tensor Decomposition for Spectrum Map Construction

Hao Sun, *Graduate Student Member, IEEE*, Junting Chen, *Member, IEEE*

Abstract—This paper addresses the challenge of reconstructing a 3D power spectrum map from sparse, scattered, and incomplete spectrum measurements. It proposes an integrated approach combining interpolation and block-term tensor decomposition (BTD). This approach leverages an interpolation model with the BTD structure to exploit the spatial correlation of power spectrum maps. Additionally, nuclear norm regularization is incorporated to effectively capture the low-rank characteristics. To implement this approach, a novel algorithm that combines alternating regression with singular value thresholding is developed. Analytical justification for the enhancement provided by the BTD structure in interpolating power spectrum maps is provided, yielding several important theoretical insights. The analysis explores the impact of the spectrum on the error in the proposed method and compares it to conventional local polynomial interpolation. Extensive numerical results demonstrate that the proposed method outperforms state-of-the-art methods in terms of signal source separation and power spectrum map construction, and remains stable under off-grid measurements and inhomogeneous measurement topologies.

Index Terms—Integrated, interpolation, block-term tensor decomposition, alternating minimization, sparse observations, power spectrum map.

I. INTRODUCTION

Spectrum maps, or more specifically, power spectrum maps, enable various applications in wireless signal processing and communications, such as signal propagation modeling [1], [2], source localization [3], [4], wireless power transfer [5], radio resource management [6], channel characterization for terahertz communications [7], and unmanned aerial vehicle (UAV) placement and strategy optimization for low-altitude integrated aerial and terrestrial communications [8], [9]. It has been a challenging problem in constructing power spectrum maps. Firstly, measurement data is only available in a limited locations or along a few routes, but the spatial pattern of the

H. Sun and J. Chen are with the School of Science and Engineering (SSE), the Shenzhen Future Network of Intelligence Institute (FNii-Shenzhen), and the Guangdong Provincial Key Laboratory of Future Networks of Intelligence, The Chinese University of Hong Kong, Shenzhen, Guangdong 518172, China.

power spectrum map can be very complex due to possible signal reflection and attenuation from the propagation environment. Second, the location and the power spectrum of the signal source can be time-varying, and therefore, the power spectrum map should be constructed within a limited time based on limited measurements.

There has been active research on power spectrum map construction. Traditional interpolation-based approaches construct each map point as a linear combination of nearby measurements, where the weights can be found using different methods including Kriging [10], [11], local polynomial regression [12], and kernel-based methods [13], [14]. Compressive sensing inspired approaches exploit the fact that the matrix representation of a power spectrum map has a *low-rank* structure, and thus, sparse Bayesian learning [15], [16], dictionary learning [17] and matrix or tensor completion [6], [18]–[21] can be applied to interpolate or extrapolate the power spectrum information at locations without measurements. Recent work [6] proposed an orthogonal matching pursuit using a tensor model for 3D spectrum mapping. The approach in [20] reconstructed the power spectrum map through minimizing the tensor rank while also enforcing smoothness. Deep learning-based approaches [22], [23] treated the power spectrum map as one or multiple layers of 2D images, and neural networks were trained to memorize the common pattern of these images based on a huge amount of training data. Furthermore, the sampling pattern can also be optimized if some prior information on the propagation field is available. In such a case, uniform sampling can be strictly sub-optimal, whereas, leverage-score-based or energy-based sampling may perform better [24], [25].

Many of these approaches were originally developed by representing the power spectrum map as one or multiple 2D data layers, for example, each layer representing a 2D power spectrum map for a specific frequency band. Although most existing approaches can be directly extended to the case in a higher dimension, their construction performance substantially degrades due to the curse of dimensionality unless a special structure is utilized. To address this issue, recent studies proposed a block-term tensor decomposition (BTD) model to exploit the spatial correlation of the power spectrum map at different frequency bands [26], [27]. The BTD model captures the property that if a signal from a particular source in a certain frequency band is blocked at a specific location, it is likely that signals from the same source at different frequency bands will also be blocked at that location. In [26], a power spectrum map was reconstructed using a BTD model where power spectrum and spatial loss field were constructed separately. The method in [27] applies a deep neural network to the

tensor decomposition model to learn the intricate underlying structures.

However, tensor decomposition for power spectrum map construction has several limitations. First, the assumption of low-rank properties in matrix or tensor models relies on on-grid measurements, a condition rarely met in reality due to physical constraints on sensor placements. Off-grid measurements induce discretization errors that likely destroy the low-rank property, and consequently, degrade the performance of matrix or tensor completion. Second, while one can reduce the grid size to reduce the discretization error, a small grid size will translate to a large matrix dimension, and this leads to the identifiability issue as the matrix or tensor can be too sparse to complete. To address the issue of off-grid measurements, recent works proposed to joint interpolation with matrix or tensor completion [28], [29]. The work [28] applied adaptive interpolation to estimate the values at the grid points, and then, performed uncertainty-aware matrix completion according to the estimated interpolation errors. In [29], thin plate spline (TPS) interpolation was used to construct each layer of the tensor, followed by BTD to construct the power spectrum map.

However, existing approaches [28], [29] are open loop approaches, where interpolation and tensor completion are sequentially performed for different objectives under different models; but the mesh-grid data created by the interpolation may not have an ideal BTD structure. Our preliminary work [30] proposed a closed loop approach to integrate interpolation and tensor completion, where a tensor structure guided interpolation problem was formulated, but by how much the tensor structure may help the interpolation was not theoretically clear.

This paper aims to construct a 3D data array representing the spectrum information over a 2D area, based on sparse, scattered, and incomplete spectrum measurements. A tensor model is formed, where each slice of the tensor represents a power spectrum map of one frequency band. To integrate interpolation with tensor completion, we perform interpolation framed by the BTD tensor structure, and in addition, the interpolation is regularized by the low-rank property of the slices of the tensor. The BTD model captures the spatial correlation of the power spectrum maps across different frequency bands under the large-scale fading, and the small-scale frequency-selective fading is treated as noise that can be mitigated by the regression to the local polynomial interpolation model. The low-rank regularization is to enforce the interpolation to be aware of the global data structure. Analytical results are established to understand the gain for the interpolation under the BTD model. It is found that, at low signal-to-noise ratio (SNR), the power spectrum map excited by narrowband sources is easier to construct than the one excited by wideband sources, whereas, at high SNR, the one excited by a wideband source with a uniform spectrum is easier to construct. The reconstruction performance under overlapping spectrum for a two-source case is also analyzed. Our numerical experiments demonstrate that the proposed scheme works much better in separating two sources when they have overlapping spectrums as compared to state-of-the-art schemes, and the proposed algorithm is stable under off-grid measurements. An improvement of over 20% in reconstruction accuracy is observed, regardless of whether the sensors measure the full or sparse spectrum.

To summarize, the following contributions are made:

- We propose an integrated interpolation and BTD approach for constructing power spectrum maps. The method incorporates an interpolation model with the BTD structure to exploit the spatial correlation of the power spectrum maps. A nuclear norm regularization is used to exploit the low-rank property of the power spectrum maps.
- We develop an alternating regression and singular value thresholding algorithm for the integrated interpolation and BTD problem.
- We establish analytical justification on why the BTD structure may enhance the interpolation for power spectrum map construction, with several important theoretical insights obtained from the analysis.
- Extensive numerical studies are conducted and show that the proposed method surpasses existing methods in terms of signal source separation and power spectrum map construction, and it is stable under off-grid measurements and inhomogeneity of the measurement topology.

The rest of the paper is organized as follows. Section II establishes the signal model and tensor model. Section III develops an integrated interpolation and BTD approach with a matrix formulation of the proposed method is provided. An alternating regression and singular value thresholding algorithm is developed to solve the proposed method. Section IV theoretically analyzes the error of the proposed method and compares it with conventional local polynomial interpolation. It then numerically discusses low-rank regularization and the performance of signal source separation. Numerical results are presented in Section V and conclusion is given in Section VI.

Notation: Vectors are written as bold italic letters $\boldsymbol{x}$, matrices as bold capital italic letters $\boldsymbol{X}$, and tensors as bold calligraphic letters $\boldsymbol{\mathcal{X}}$. For a matrix $\boldsymbol{X}$, $[\boldsymbol{X}]_{(i,j)}$ denotes the entry in the ith row and jth column of $\boldsymbol{X}$. For a tensor $\boldsymbol{\mathcal{X}}$, $[\boldsymbol{\mathcal{X}}]_{(i,j,k)}$ denotes the entry under the index (i,j,k). The symbol '$\circ$' represents outer product, '$\otimes$' represents Kronecker product, and $\|\cdot\|_F$ represents Frobenius norm. The notation $o(x)$ means $\lim_{x\to 0} o(x)/x \to 0$, $\mathrm{diag}(\boldsymbol{x})$ represents a diagonal matrix whose diagonal elements are the entries of vector $\boldsymbol{x}$, and $\mathrm{vec}(\boldsymbol{X})$ denotes the vectorization of matrix $\boldsymbol{X}$. The symbol $\mathbb{E}\{\cdot\}$ and $\mathbb{V}\{\cdot\}$ denote expectation and variance separately.

II. SYSTEM MODEL

A. Signal Model

Consider that a bounded area $\mathcal{D} \subseteq \mathbb{R}^2$ contains R signal sources. The signals occupying K frequency bands are detected by M sensors at known locations $\boldsymbol{z}_m \in \mathcal{D}$, $m = 1, 2, \ldots, M$. Denote $\boldsymbol{s}_r \in \mathcal{D}$ as the location of the rth source. Then, the signal power from the rth signal source measured at the kth frequency band and location $\boldsymbol{z}$ is modeled as

$$\rho_k^{(r)}(\boldsymbol{z}) = \left(g_r(d(\boldsymbol{s}_r, \boldsymbol{z})) + \zeta_r(\boldsymbol{z}) + \eta_{r,k}(\boldsymbol{z})\right) \phi_k^{(r)} \quad (1)$$

where $g_r(d(\boldsymbol{s}_r, \boldsymbol{z}))$ describes the path gain of the rth source at distance $d(\boldsymbol{s}_r, \boldsymbol{z})$, the function $d(\boldsymbol{s}, \boldsymbol{z}) = \|\boldsymbol{s} - \boldsymbol{z}\|_2$ describes

the distance between a source at s and a sensor at z, $\zeta_r(z)$ captures the shadowing of the signal from the rth source, $\eta_{r,k}(z) \sim \mathcal{N}(0, \sigma_\eta^2)$ is a zero mean Gaussian random variable to model the fluctuation due to the frequency-selective fading, and $\phi_k^{(r)}$ describes the power allocation of the rth source at the kth frequency band. Note that the values of all these components are not known to the system.

The aggregated power at the kth frequency band from all the R sources measured by a sensor located at z_m is denoted as

$$\gamma_m^{(k)} = \sum_{r=1}^{R} \rho_k^{(r)}(z_m) + \epsilon, \quad \forall k \in \Omega_m \tag{2}$$

where $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$ is to model the measurement noise at each frequency band, and $\Omega_m \subseteq \{1, 2, \ldots, K\}$ contains the set of frequency bands that are measured by the mth sensor. We assume that for each source r, the total power sums to $\sum_{k=1}^{K} \phi_k^{(r)} = K$, and therefore, the SNR is normalized to $K/(K\sigma_\epsilon^2) = 1/\sigma_\epsilon^2$, where $K\sigma_\epsilon^2$ is the total noise for the entire bandwidth.

Consider to discretize the target area $\mathcal{D}$ into N_1 rows and N_2 columns that results in $N_1 \times N_2$ grid cells. Let $c_{ij} \in \mathcal{D}$ be the center location of the (i,j)th grid cell. Our goal is to reconstruct the large-scale propagation field, *i.e.*, the first two terms in (1)

$$\rho^{(r)}(z) = g_r(d(s_r, z)) + \zeta_r(z) \tag{3}$$

at grid points c_{ij} and the power spectrum $\phi_k^{(r)}$, $k = 1, 2, \ldots, K$. The reconstructed propagation field model (3) does not capture the frequency-selective fading $\eta_{r,k}$.

As a result, from the propagation model (1), the measurement model in (2) can be derived as

$$\gamma_m^{(k)} = \sum_{r=1}^{R} \rho^{(r)}(z_m)\phi_k^{(r)} + \tilde{\epsilon}_k \tag{4}$$

where $\tilde{\epsilon}_k = \sum_r \eta_{r,k}(z)\phi_k^{(r)} + \epsilon$ is a zero mean random variable that combines the randomness due to the frequency-selective small-scale fading $\eta_{r,k}(z)\phi_k^{(r)}$ and the measurement noise ϵ.

B. Tensor Model

Let $S_r \in \mathbb{R}^{N_1 \times N_2}$ be a discretized form of the propagation field for the rth source, where the (i,j)th entry is given by $[S_r]_{(i,j)} = \rho^{(r)}(c_{ij}) \in \mathbb{R}^1$. It has been widely discussed in the literature that for many common propagation scenarios, the matrix S_r tends to be low-rank [1], [26].

Let $\mathcal{H} \in \mathbb{R}^{N_1 \times N_2 \times K}$ be a tensor representation of the *target* power spectrum maps to be constructed. Based on (1) and (3), we have $[\mathcal{H}]_{(i,j,k)} = \sum_{r=1}^{R} \rho^{(r)}(c_{ij})\phi_k^{(r)}$ to represent the aggregated power of the kth frequency band from the R sources measured at location c_{ij} exempted from the small-scale fading component $\eta_{r,k}(z)\phi_k^{(r)}$. Denote $\phi^{(r)} = [\phi_1^{(r)}, \ldots, \phi_K^{(r)}]^{\mathrm{T}} \in \mathbb{R}^K$ as the power spectrum from the rth source. As a result, the tensor $\mathcal{H}$ has the following BTD structure

$$\mathcal{H} = \sum_{r=1}^{R} S_r \circ \phi^{(r)} \tag{5}$$

where '∘' represents outer product.

Conventional tensor-based power spectrum map construction approaches obtain the complete tensor $\mathcal{H}$ from the measurement $\gamma_m^{(k)}$ assuming that $\gamma_m^{(k)}$ are taken at the center of the grid cell without measurement noise [6], [20]. However, when the grid cells are too large, corresponding to small N_1 and N_2, it is hard to guarantee that the sensor at z_m is placed at the corresponding grid center c_{ij}, resulting in possibly large discretization error. When the grid cells are small, corresponding to large N_1 and N_2, there might be an identifiability issue as the dimension of the tensor is large.

Recent attempts [28], [29], [31] consider to first estimate $[\mathcal{H}]_{(i,j,k)}$ using interpolation methods based on the off-grid measurements, and then, employ matrix completion or tensor completion based on the BTD model (5) to improve the spectrum map construction. However, these methods are open-loop methods where the property that S_r are low-rank matrices is not exploited in the interpolation step; consequently, a poor open-loop interpolation may affect the performance in the tensor completion step.

III. INTEGRATED INTERPOLATION AND BTD APPROACH

In this section, we propose an integrated interpolation and BTD approach, where the BTD structure of the tensor model and the low-rank property of the tensor components are both exploited for interpolation.

A. The Integrated Interpolation and BTD Problem

Based on the BTD model in (5), we consider to fit a model $f^{(r)}(z)$ to approximate the large-scale propagation field $\rho^{(r)}(z)$ of the rth source in (3) from the multi-band measurements $\gamma_m^{(k)}$ by exploiting the structure of the tensor model $\mathcal{H}$ and the low-rank property of the tensor component S_r.

Here, we adopt a polynomial model for the propagation field $f^{(r)}(z)$ of the rth source, but note that, other conventional interpolation approaches, including Kriging and kernel regression, also work with the proposed framework. The global model $f^{(r)}(z)$ for source r can be constructed based on a number of *local* models $f_{ij}^{(r)}(z)$ on selected cells $(i,j) \in \mathcal{I}$. Without loss of generality (w.l.o.g.), a second order polynomial model for the (i,j)th grid cell centered at c_{ij} is given as follows:

$$f_{ij}^{(r)}(z) = \alpha_{ij}^{(r)} + (\beta_{ij}^{(r)})^{\mathrm{T}}(z - c_{ij}) + (z - c_{ij})^{\mathrm{T}} B_{ij}^{(r)}(z - c_{ij}). \tag{6}$$

We collect the model parameters of the local model at cell (i,j) for the rth source into a vector $\theta_{ij}^{(r)} = [\alpha_{ij}^{(r)}, (\beta_{ij}^{(r)})^{\mathrm{T}}, (\mathrm{vec}(B_{ij}^{(r)}))^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^7$, and $\Theta_{ij} = [\theta_{ij}^{(1)}; \cdots; \theta_{ij}^{(R)}] \in \mathbb{R}^{7R}$ is a collection of model parameters at cell (i,j) for all sources.

It follows that, under perfect interpolation that yields $f_{ij}^{(r)}(c_{ij}) = \rho^{(r)}(c_{ij})$, we have $[\mathcal{H}]_{(i,j,k)} = \sum_{r=1}^{R} f_{ij}^{(r)}(c_{ij})\phi_k^{(r)}$, which aligns with the BTD tensor structure in (5). Therefore, a BTD tensor structure guided least-squares local polynomial regression based on the

measurements $\gamma_m^{(k)}$ at the location z_m can be constructed through minimizing the following cost

$$
\begin{aligned}
& l_{ij}(\boldsymbol{\Theta}_{ij}, \{\phi_k^{(r)}\}) \\
& = \sum_{m=1}^{M} \sum_{k \in \Omega_m} \left(\gamma_m^{(k)} - \sum_{r=1}^{R} f_{ij}^{(r)}(z_m)\phi_k^{(r)} \right)^2 \kappa_{ij}(z_m). \quad (7)
\end{aligned}
$$

The term $\kappa_{ij}(\boldsymbol{z}) \triangleq \kappa((\boldsymbol{z} - \boldsymbol{c}_{ij})/b)$ is a kernel function with a parameter b to weight the importance of the measurements. The kernel $\kappa_{ij}(\boldsymbol{z})$ here assigns weights to the measurements $\gamma_m^{(k)}$. For those measurements far from the grid center $\boldsymbol{c}_{ij}$, the weights are small. Since our purpose is to weigh the importance of the measurements, a typical spatially smoothing kernel can be used. A possible choice of kernel function can be the Epanechnikov kernel $\kappa(\boldsymbol{u}) = \max\{0, \frac{3}{4}(1 - \|\boldsymbol{u}\|^2)\}$ [12].

The cost function for the global model f can be written as follows:

$$
l(f) = \sum_{(i,j) \in \mathcal{I}} l_{ij}(\boldsymbol{\Theta}_{ij}, \{\phi_k^{(r)}\}). \quad (8)
$$

The regression model (8) is not aware of the hidden low-rank property of the propagation field. We thus propose an integrated interpolation and BTD formulation to impose the low-rank for the global model f as follows:

$$
\begin{aligned}
\underset{\{\boldsymbol{\Theta}_{ij}, \{\phi_k^{(r)}\}\}, \{\boldsymbol{S}_r\}}{\text{minimize}} \; & l(f) + \frac{\nu}{2} \sum_{(i,j) \in \mathcal{I}} \sum_{r=1}^{R} (f_{ij}^{(r)}(\boldsymbol{c}_{ij}) - [\boldsymbol{S}_r]_{(i,j)})^2 \\
& + \mu \sum_{r=1}^{R} \|\boldsymbol{S}_r\|_* \quad (9)
\end{aligned}
$$

where $\| \cdot \|_*$ represents the nuclear norm. As a result, the regression model f not only needs to fit the measurement data $\gamma_m^{(k)}$ via minimizing the cost $l_f(\cdot)$ in (8) but is also penalized by the rank of $\boldsymbol{S}_r$ via the second and the third terms in (9).

B. Matrix Formulation of Integrated Interpolation and BTD

It turns out that the minimization problem (9) has a nice structure after reformulation with matrix representations.

Denote $\boldsymbol{x}_m(\boldsymbol{c}_{ij}) = [1, (z_m - \boldsymbol{c}_{ij})^{\mathrm{T}}, (\text{vec}((z_m - \boldsymbol{c}_{ij})(z_m - \boldsymbol{c}_{ij})^{\mathrm{T}}))^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^7$. Then, the polynomial model $f_{ij}^{(r)}(z_m)$ in (6) is rewritten as $f_{ij}^{(r)}(z_m) = \boldsymbol{x}_m^{\mathrm{T}}(\boldsymbol{c}_{ij})\boldsymbol{\theta}_{ij}^{(r)}$.

Denote $\boldsymbol{\phi}_k = [\phi_k^{(1)}, \cdots, \phi_k^{(R)}]^{\mathrm{T}} \in \mathbb{R}^R$, and recall the definition of $\boldsymbol{\Theta}_{ij}$ in (7). From (7), we have $\sum_{r=1}^{R} f_{ij}^{(r)}(z_m)\phi_k^{(r)} = \boldsymbol{\phi}_k^{\mathrm{T}} \otimes \boldsymbol{x}_m^{\mathrm{T}}(\boldsymbol{c}_{ij})\boldsymbol{\Theta}_{ij}$ where '$\otimes$' is the Kronecker product. The least-squares local polynomial regression model (7) is rewritten as follows:

$$
\begin{aligned}
& l_{ij}(\boldsymbol{\Theta}_{ij}, \{\phi_k^{(r)}\}) \\
& = \sum_{m=1}^{M} \sum_{k \in \Omega_m} \left(\gamma_m^{(k)} - \boldsymbol{\phi}_k^{\mathrm{T}} \otimes \boldsymbol{x}_m^{\mathrm{T}}(\boldsymbol{c}_{ij})\boldsymbol{\Theta}_{ij} \right)^2 \kappa_{ij}(z_m). \\
& = \sum_{m=1}^{M} \sum_{k=1}^{K} \psi_{mk} \left(\gamma_m^{(k)} - \boldsymbol{\phi}_k^{\mathrm{T}} \otimes \boldsymbol{x}_m^{\mathrm{T}}(\boldsymbol{c}_{ij})\boldsymbol{\Theta}_{ij} \right)^2 \kappa_{ij}(z_m) \quad (10)
\end{aligned}
$$

where ψ_{mk} indicates the observation status of kth frequency band at location m, if $k \in \Omega_m$, we have $\psi_{mk} = 1$, elsewise

$\psi_{mk} = 0$. For the convenience of matrix form formulation, we denote $[\boldsymbol{\psi}]_{mk} = \psi_{mk}$.

Next, we further rearrange (10) into a matrix form. Denote $\boldsymbol{X}_{ij} = [\boldsymbol{x}_1(\boldsymbol{c}_{ij}), \boldsymbol{x}_2(\boldsymbol{c}_{ij}), ..., \boldsymbol{x}_M(\boldsymbol{c}_{ij})] \in \mathbb{R}^{7 \times M}$. To concisely express the mathematical relationships, we first replace the summation term over m in (10) with its matrix form as follows:

$$
\begin{aligned}
& \sum_{m=1}^{M} \sum_{k=1}^{K} \psi_{mk} \left(\gamma_m^{(k)} - \boldsymbol{\phi}_k^{\mathrm{T}} \otimes \boldsymbol{x}_m^{\mathrm{T}}(\boldsymbol{c}_{ij})\boldsymbol{\Theta}_{ij} \right)^2 \kappa_{ij}(z_m) \\
& = \sum_{m=1}^{M} \sum_{k=1}^{K} \left(\psi_{mk}\sqrt{\kappa_{ij}(z_m)} \left(\gamma_m^{(k)} - \boldsymbol{\phi}_k^{\mathrm{T}} \otimes \boldsymbol{x}_m^{\mathrm{T}}(\boldsymbol{c}_{ij})\boldsymbol{\Theta}_{ij} \right) \right)^2 \\
& = \sum_{k=1}^{K} \left\| \mathbf{Q}_{ij}\text{diag}(\boldsymbol{\psi}(:,k)) \left(\boldsymbol{\Gamma}(:,k) - \boldsymbol{\phi}_k^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{\Theta}_{ij} \right) \right\|_2^2 \quad (11)
\end{aligned}
$$

where $\boldsymbol{\Gamma}(m,k) = \gamma_m^{(k)}$, $\boldsymbol{\Gamma}(:,k)$ and $\boldsymbol{\psi}(:,k)$ respectively represent the kth column of $\boldsymbol{\Gamma}$ and $\boldsymbol{\psi}$. The matrix $\mathbf{Q}_{ij}$ is an $M \times M$ diagonal matrix with the mth diagonal element equals to $\sqrt{\kappa_{ij}(z_m)}$.

Further, based on (11), we replace the summation term over k with its matrix form. Denote $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \cdots, \boldsymbol{\phi}_K] \in \mathbb{R}^{R \times K}$. Then, model (11) can be written as follows:

$$
\begin{aligned}
& \sum_{k \in \Omega_m} \left\| \mathbf{Q}_{ij}\text{diag}(\boldsymbol{\psi}(:,k)) \left(\boldsymbol{\Gamma}(:,k) - \boldsymbol{\phi}_k^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{\Theta}_{ij} \right) \right\|_2^2 \\
& = \left\| \boldsymbol{W}_{ij} \left(\text{vec}(\boldsymbol{\Gamma}) - \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{\Theta}_{ij} \right) \right\|_2^2
\end{aligned}
$$

where $\boldsymbol{W}_{ij}$ is an $MK \times MK$ diagonal matrix,

$$
\boldsymbol{W}_{ij} = \mathbf{I}_K \otimes \mathbf{Q}_{ij}\text{diag}(\text{vec}(\boldsymbol{\psi})) \in \mathbb{R}^{MK \times MK} \quad (12)
$$

and $\mathbf{I}_K \in \mathbb{R}^{K \times K}$ is an identity matrix.

From (6), if $z = \boldsymbol{c}_{ij}$, then we have $f_{ij}^{(r)}(\boldsymbol{c}_{ij}) = \alpha_{ij}^{(r)} = \boldsymbol{e}_r^{\mathrm{T}}\boldsymbol{\Theta}_{ij}$, where $\boldsymbol{e}_r$ is a unit vector with the $(7 \times (r-1) + 1)$th entry equals 1, and all the other elements equal 0. Thus, the integrated interpolation and BTD problem (9) can be rewritten as follows:

$$
\begin{aligned}
\underset{\boldsymbol{\Theta}_{ij}, \boldsymbol{\Phi}, \{\boldsymbol{S}_r\}}{\text{minimize}} \; & \sum_{(i,j) \in \mathcal{I}} \left\| \boldsymbol{W}_{ij} \left(\text{vec}(\boldsymbol{\Gamma}) - \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}}\boldsymbol{\Theta}_{ij} \right) \right\|_2^2 \\
& + \nu \sum_{(i,j) \in \mathcal{I}} \sum_{r=1}^{R} (\boldsymbol{e}_r^{\mathrm{T}}\boldsymbol{\Theta}_{ij} - [\boldsymbol{S}_r]_{(i,j)})^2 + \mu \sum_{r=1}^{R} \|\boldsymbol{S}_r\|_* \\
\text{subject to} \; & [\boldsymbol{\Phi}]_{ij} \geq 0 \quad\quad (13)
\end{aligned}
$$

where $\boldsymbol{\Gamma}$ is a collection of received signal strength (RSS) $\gamma_m^{(k)}$, $\boldsymbol{\Phi}$ is the power spectrum, $\boldsymbol{W}_{ij}$ is the diagonal matrix of weights defined in (12), $\boldsymbol{X}_{ij}$ is a collection of measurement locations z_m, and $\boldsymbol{S}_r$ is a discretized form of propagation field in (5). The constraint is added because the power spectrum is non-negative.

C. Alternating Regression and Singular Value Thresholding

It is observed in (13) that given the matrix components $\boldsymbol{S}_r$ and the spectrum variable $\boldsymbol{\Phi}$, the objective is a convex quadratic function in the regression parameters $\boldsymbol{\Theta}_{ij}$. Likewise, given $\boldsymbol{S}_r$ and $\boldsymbol{\Theta}_{ij}$, (13) is also a convex quadratic function in

Φ; and moreover, (13) is a convex function in S_r. Therefore, it is natural to adopt an alternating optimization approach to solve for the integrated interpolation and BTD problem (13).

Update of Θ_{ij}: Given the values of S_r and Φ, the optimization problem (13) is equivalent to a weighted least-squares problem as follows:

$$\underset{\Theta_{ij}}{\text{minimize}} \sum_{(i,j)\in\mathcal{I}} \left\| W_{ij}\left(\text{vec}(\Gamma) - \Phi^{\mathrm{T}} \otimes X_{ij}^{\mathrm{T}}\Theta_{ij}\right)\right\|_2^2$$
$$+ \nu \sum_{(i,j)\in\mathcal{I}} \sum_{r=1}^{R}(e_r^{\mathrm{T}}\Theta_{ij} - [S_r]_{(i,j)})^2. \quad (14)$$

Note that the problem (14) is an unconstrained strictly convex problem. Hence, the solution can be obtained through setting the first order derivative of (14) to zero and we get:

$$\hat{\Theta}_{ij} = (\Phi \otimes X_{ij}W_{ij}^2\Phi^{\mathrm{T}} \otimes X_{ij}^{\mathrm{T}} + \nu\sum_{r=1}^{R}e_re_r^{\mathrm{T}})^{-1}$$
$$\times \left(\Phi \otimes X_{ij}W_{ij}^2\text{vec}(\Gamma) + \nu\sum_{r=1}^{R}e_r[S_r]_{(i,j)}\right). \quad (15)$$

Update of Φ: Similarly, given the values of S_r and Θ_{ij}, the optimization problem (13) is equivalent to a constrained weighted least-squares as follows:

$$\underset{\Phi}{\text{minimize}} \sum_{(i,j)\in\mathcal{I}} \left\| W_{ij}\left(\text{vec}(\Gamma) - \Phi^{\mathrm{T}} \otimes X_{ij}^{\mathrm{T}}\Theta_{ij}\right)\right\|_2^2$$
$$\text{subject to } [\Phi]_{ij} \geq 0. \quad (16)$$

Using the property of Kronecker product $\text{vec}(AXB) = (B^{\mathrm{T}} \otimes A)\text{vec}(X)$ and $\text{vec}(AB) = (I \otimes A)\text{vec}(B) = (B^{\mathrm{T}} \otimes I)\text{vec}(A)$ [32], problem (16) can be rewritten as:

$$\underset{\Phi}{\text{minimize}} \sum_{(i,j)\in\mathcal{I}} \left\| W_{ij}\left(\text{vec}(\Gamma)\right.\right.$$
$$\left.\left. - \left(I_K \otimes (X_{ij}^{\mathrm{T}}[\Theta_{ij}]_{7\times R})\right)\text{vec}(\Phi)\right)\right\|_2^2$$
$$\text{subject to } [\Phi]_{ij} \geq 0 \quad (17)$$

where $[\Theta_{ij}]_{7\times R}$ is the reformed matrix form of vector Θ_{ij} with dimension $7 \times R$.

Note that the mth grid in set $\mathcal{I}$ is a one-to-one mapping to the (i,j)th grid. Denote $\tilde{W}_{ij} = [W_1; \cdots; W_m; \cdots; W_{|\mathcal{I}|}]$, $Z_{ij} = W_{ij}(I_K \otimes (X_{ij}^{\mathrm{T}}[\Theta_{ij}]_{7\times R}))$ and $\tilde{Z}_{ij} = [Z_1; \cdots; Z_m; \cdots; Z_{|\mathcal{I}|}]$. We reformulate (17) to make it a general non-negative constrained least-squares problem as follows:

$$\underset{\Phi}{\text{minimize}} \left\| \tilde{W}_{ij}\left(\text{vec}(\Gamma) - \tilde{Z}_{ij}\text{vec}(\Phi)\right)\right\|_2^2$$
$$\text{subject to } [\Phi]_{ij} \geq 0. \quad (18)$$

Note that the problem (18) is a strictly convex problem. Hence, it has a unique optimal solution and a single principal pivoting algorithm [33] can be applied to solve (18).

Update of S_r: Finally, with the Θ_{ij} and Φ updated through (15) and (18), the optimization problem (13) is equivalent to a nuclear norm regularized low-rank matrix completion problem as follows:

$$\underset{\{S_r\}}{\text{minimize}} \sum_{(i,j)\in\mathcal{I}} \sum_{r=1}^{R}(e_r^{\mathrm{T}}\Theta_{ij} - [S_r]_{(i,j)})^2 + \mu\sum_{r=1}^{R}\|S_r\|_*. \quad (19)$$

It is observed that (19) can be equivalently decomposed into R parallel sub-problems each focusing on an S_r as follows:

$$\underset{\{S_r\}}{\text{minimize}} \sum_{(i,j)\in\mathcal{I}}(e_r^{\mathrm{T}}\Theta_{ij} - [S_r]_{(i,j)})^2 + \mu\|S_r\|_*. \quad (20)$$

Denote the observation matrix Ψ as $[\Psi]_{ij} = e_r^{\mathrm{T}}\Theta_{ij}$. Then, the singular value thresholding algorithm can be applied to solve (20) through the iteration as follows [34]:

$$\begin{cases} S_r^{(k)} = \mathcal{S}_\mu(Y^{(k-1)}) \\ Y^{(k)} = Y^{(k-1)} + \delta\mathcal{P}_\Omega(\Psi - S_r^{(k)}) \end{cases} \quad (21)$$

where Y is an intermediate matrix, (k) represents the index of each iteration, $[\mathcal{P}_\Omega(X)]_{ij} = [X]_{ij}$ if $(i,j) \in \Omega$ and zero otherwise, and $\mathcal{S}_\mu$ is the soft-thresholding operator which is defined as follows:

$$\mathcal{S}_\mu(Y^{(k-1)}) = U_\iota\mathcal{D}_\mu(\Sigma_\iota)V_\iota^{\mathrm{T}}$$

with $\mathcal{D}_\mu(\Sigma_\iota) = \text{diag}[(\sigma_1 - \mu)_+, \cdots(\sigma_\iota - \mu)_+]$, $\Sigma_\iota = \text{diag}[\sigma_1, \cdots, \sigma_\iota]$, $(x)_+ = \max(0, x)$, and $U_\iota\Sigma_\iota V_\iota^{\mathrm{T}}$ is the singular value decomposition (SVD) of $Y^{(k-1)}$ where ι is the rank of $Y^{(k-1)}$.

Each sub-problem of (13) is strictly convex and has a unique solution. Since the overall objective value decreases (or at least does not increase) with each iteration, and the objective (9) is lower bounded by 0, the alternating regression and singular value thresholding algorithm must converge.

Finally, the power spectrum map is constructed as $\hat{\mathcal{H}} = \sum_{r=1}^{R}\hat{S}_r \circ \hat{\phi}^{(r)}$.

IV. PERFORMANCE OF THE INTEGRATED INTERPOLATION AND BTD APPROACH

This section investigates the performance and the potential advantage of the proposed integrated interpolation and BTD approach (9) from three aspects. First, we focus on the least-squares cost term $l(f)$ by dropping the low-rank regularization in (9). By analyzing the interpolation error, we show that the proposed method indeed yields a better accuracy. Then, we numerically verify that by imposing the low-rank regularization, the reconstruction accuracy can be further improved. Finally, we discuss how the proposed structure may improve the identifiability of separating multiple sources compared to the classical tensor-based approaches.

A. Improvement from the BTD Model

We first show that exploiting the BTD model in (5) can increase the accuracy of constructing S_r.

1) Single Source: Let us start from the case of full spectrum observation for a single source, where $R = 1$ and each sensor measures all the frequency bands, i.e., ψ in (11) is an all 1's matrix. The result can be easily extended to the case where each sensor only observes a subset of frequency bands.

For the sake of notation simplicity, we omit the superscript "(r)" and adopt symbols ρ, f, and so on, to represent $\rho^{(r)}$, $f^{(r)}$, and similar variables. Recall that $f_{ij}(\boldsymbol{c}_{ij})$ is a model to approximate $\rho(\boldsymbol{c}_{ij})$, and $\alpha_{ij} = f_{ij}(\boldsymbol{c}_{ij})$ from (6). Define the interpolation error as $\xi_{ij} = \hat{\alpha}_{ij} - \rho(\boldsymbol{c}_{ij})$ where $\hat{\alpha}_{ij}$ is an estimate of α_{ij}.

Assume that the propagation field $\rho_k(\boldsymbol{z})$ in (1) is third-order differentiable. Then, if ϕ_k is available, we have the following results to characterize the construction error ξ_{ij} at a point $\boldsymbol{c}_{ij}$ of the integrated interpolation and BTD approach.

Theorem 1 (Full spectrum error analysis). *Let $\hat{\alpha}_{ij}$ be the estimate obtained from the solution to the integrated interpolation and BTD problem (14) under $\nu = 0$ based on the regression model (6). The variance $\mathcal{E}_t = \mathbb{V}\{\xi_{ij}\}$ of the interpolation error $\xi_{ij} = \hat{\alpha}_{ij} - \rho(\boldsymbol{c}_{ij})$ is given by*

$$\mathcal{E}_t = \left(\frac{\sum_{k=1}^{K} \phi_k^4}{\left(\sum_{k=1}^{K} \phi_k^2\right)^2} \sigma_\eta^2 + \frac{1}{\sum_{k=1}^{K} \phi_k^2} \sigma_\epsilon^2 \right) C\left(\{\boldsymbol{z}_m\}, b\right) \quad (22)$$

where $C(\{\boldsymbol{z}_m\}, b)$ is a constant that depends on the sensor locations $\boldsymbol{z}_m$ and the window size b.

Proof. See Appendix A. □

It is observed that the variance of the interpolation error ξ_{ij} depends on the spectrum $\boldsymbol{\Phi}$ of the source. An analysis of the coefficients is given in the following proposition.

Proposition 1 (Impact from the spectrum). *The coefficients in (22) satisfy*

$$\frac{1}{K} \leq \varpi_\eta(\boldsymbol{\Phi}) \triangleq \frac{\sum_{k=1}^{K} \phi_k^4}{\left(\sum_{k=1}^{K} \phi_k^2\right)^2} \leq 1$$

with lower bound achieved when $\boldsymbol{\Phi} = (1, 1, \ldots, 1)$ and upper bound achieved when $\boldsymbol{\Phi} = (K, 0, 0, \ldots, 0)$. In addition,

$$\frac{1}{K^2} \leq \varpi_\epsilon(\boldsymbol{\Phi}) \triangleq \frac{1}{\sum_k \phi_k^2} \leq \frac{1}{K}$$

with lower bound achieved when $\boldsymbol{\Phi} = (K, 0, 0, \ldots, 0)$ and upper bound achieved when $\boldsymbol{\Phi} = (1, 1, \ldots, 1)$.

Proof. See Appendix B. □

One can make the following observations. First, the impact from the frequency-selective fading due to the power allocation of the source is more significant than that from the measurement noise, in the sense that for $\sigma_\eta = \sigma_\epsilon$, we have $\varpi_\epsilon(\boldsymbol{\Phi}) \leq \varpi_\eta(\boldsymbol{\Phi})$ for all power spectrum $\boldsymbol{\Phi}$. The intuition is that the frequency-selective fading component $\eta_{r,k}$ is multiplied with the power allocation ϕ_k in (1) for the contribution to the measurement in (4), and hence, a power boost in the kth frequency band also enhances the frequency-selective fading component.

Second, it follows that equal-power allocation $\boldsymbol{\Phi} = (1, 1, \ldots, 1)$ minimizes the impact of the frequency-selective fading under the integrated interpolation and BTD approach. In this case, the fading component can be treated as additional measurement noise, resulting in a combined noise with variance $\sigma_\eta^2 + \sigma_\epsilon^2$. The more frequency bands that can be measured, the less impact from the frequency-selective fading.

Third, on the contrary to the impact of the frequency-selective fading, the noise term prefers a power boost in just one frequency band. This is because reducing the bandwidth also reduces the noise power, resulting in an increase in the SNR.

To summarize, in the high SNR scenario, where the measurement noise power σ_ϵ^2 is much smaller than the amount of the frequency-selective fading σ_η^2, a wide-band measurement with equal power allocation is preferred for constructing the propagation map $\boldsymbol{S}_r$. On the contrary, in the low SNR scenario, a narrow-band measurement is preferred and all the power should be allocated to a single frequency band. This is mathematically summarized in the following corollary.

Corollary 1 (Asymptotic performance). *At high SNR, the uniform spectrum $\boldsymbol{\Phi} = (1, 1, \ldots, 1)$ asymptotically minimizes $\mathcal{E}_t$ as $1/\sigma_\epsilon^2 \to \infty$, and consequently, $\mathcal{E}_t \to C\sigma_\eta^2/K$. At low SNR, the single band measurement under $\boldsymbol{\Phi} = (K, 0, 0, \ldots, 0)$ asymptotically minimizes $\mathcal{E}_t$ as $1/\sigma_\epsilon^2 \to 0$, and consequently, $\mathcal{E}_t \to C\sigma_\epsilon^2/K^2$.*

For performance bench-marking, we consider a conventional frequency-by-frequency construction for each $\rho_k(\boldsymbol{z})$, using a local polynomial interpolation technique under the same regression model (6). Specifically, for each k, we construct $\rho_k(\boldsymbol{z})$ only based on $\boldsymbol{z}_m$ and the measurement of the kth frequency band $\gamma_m^{(k)}$ using local polynomial regression under the same parameter b and the kernel function. Likewise, the spectrum $\boldsymbol{\Phi}$ is available. Denote $\hat{\alpha}_{ij}^{(k)}$ as the estimated power spectrum $\rho_k(\boldsymbol{c}_{ij})$ at location $\boldsymbol{c}_{ij}$.

Proposition 2 (Interpolation without a BTD model). *Let $\hat{\alpha}_{ij} = \sum_{k=1}^{K} \hat{\alpha}_{ij}^{(k)}/K$, and $\xi_{ij} = \sum_{k=1}^{K} \xi_{ij}^{(k)}/K$ where $\xi_{ij}^{(k)} = \hat{\alpha}_{ij}^{(k)} - \rho_k(\boldsymbol{c}_{ij})$. Under the same assumption as in Theorem 1, the variance $\mathcal{E}_p = \mathbb{V}\{\xi_{ij}\}$ of the interpolation error ξ_{ij} is given by*

$$\mathcal{E}_p = \left(\sigma_\eta^2 + \sum_{k=1}^{K} \frac{1}{\phi_k^2 K} \sigma_\epsilon^2 \right) C\left(\{\boldsymbol{z}_m\}, b\right). \quad (23)$$

Proof. See Appendix C. □

It is observed that the impact from the frequency-selective fading σ_η^2 is independent of the spectrum $\boldsymbol{\Phi}$, which is in contrast to (22) for the case with the tensor guidance. This is due to the fact that only the data $\gamma_m^{(k)}$ for the same frequency band is used. The coefficient $\sum_{k=1}^{K} 1/(\phi_k^2 K)$ of measurement noise σ_ϵ^2 in (23) is lower bounded by 1 with lower bound achieved when $\boldsymbol{\Phi} = (1, 1, \ldots, 1)$. Note that this error coefficient $\sum_{k=1}^{K} 1/(\phi_k^2 K)$ can be arbitrarily large for an arbitrarily small ϕ_k in a particular frequency band k. On the contrary, the construction performance for the integrated

interpolation and BTD approach in (22) is less affected even $\phi_k = 0$ for some k.

To make a more specific comparison, we derive the difference $\mathcal{E}_p - \mathcal{E}_t$.

Proposition 3 (Interpolation error reduction). *Under the same assumption as in Theorem 1, the proposed integrated interpolation and BTD approach reduces the variance of the interpolation error by*

$$\mathcal{E}_p - \mathcal{E}_t = \left(\varsigma_\eta(\boldsymbol{\Phi})\sigma_\eta^2 + \varsigma_\epsilon(\boldsymbol{\Phi})\sigma_\epsilon^2\right) C\left(\{\boldsymbol{z}_m\}, b\right) \quad (24)$$

where $\varsigma_\eta(\boldsymbol{\Phi}) = 2(\sum_{i \neq j} \phi_i^2 \phi_j^2)/(\sum_{k=1}^K \phi_k^2)^2$ *and* $\varsigma_\epsilon(\boldsymbol{\Phi}) = \sum_{k=1}^K ((\sum_{l \neq k}^K \phi_l^2)/(\phi_k^2 K))/\sum_{k=1}^K \phi_k^2$.

In addition, $0 \leq \varsigma_\eta(\boldsymbol{\Phi}) \leq \frac{K-1}{K} \leq \varsigma_\epsilon(\boldsymbol{\Phi})$ where the second and third inequalities are achieved when $\boldsymbol{\Phi} = (1, 1, \ldots, 1)$ and the inequality in first inequality is achieved when $\boldsymbol{\Phi} = (K, 0, 0, \ldots, 0)$.

Proof. See Appendix D. $\qquad \square$

It is observed that the integrated interpolation and BTD approach is always of smaller variance than the conventional frequency-by-frequency local polynomial interpolation. In addition, the error from the measurement noise is more significant than that from the frequency-selective fading, in the sense that for $\sigma_\eta = \sigma_\epsilon$, we have $\varsigma_\eta(\boldsymbol{\Phi})\sigma_\eta \leq \varsigma_\epsilon(\boldsymbol{\Phi})\sigma_\epsilon$, because $\varsigma_\eta(\boldsymbol{\Phi}) \leq \frac{K-1}{K} \leq \varsigma_\epsilon(\boldsymbol{\Phi})$.

Finally, we extend our discussion to the case of sparse spectrum observation. We assume each sensor m randomly measures $|\Omega_m| = K'$ frequency bands. To simplify the discussion, we assume the source has equal power allocation $\boldsymbol{\Phi} = (1, 1, \ldots, 1)$ over all frequency bands.

Proposition 4 (Sparse spectrum error analysis). *Under the same condition of Theorem 1, the variance $\mathcal{E}_t$ of the interpolation error ξ_{ij} for the integrated interpolation and BTD approach is given by*

$$\mathcal{E}_t = \frac{\sigma_\eta^2 + \sigma_\epsilon^2}{K'} C(\{\boldsymbol{z}_m\}, b).$$

Proof. See Appendix E. $\qquad \square$

Proposition 4 verifies that the interpolation performance for the integrated interpolation and BTD approach depends on the number of observed frequency bands K' in each sensor, rather than the measurement pattern in the frequency domain, i.e., different sensors can observe different subsets of frequency bands. For instance, we may have more measurements for some frequency bands, while other frequency bands have rare measurements, and such a heterogeneous situation does not affect the performance for the integrated interpolation and BTD approach.

2) Multiple Sources: We demonstrate the result for a two-source case, where the derivation for $R > 2$ is straightforward, and similar insights can be obtained.

Consider that the two sources occupy K frequency bands. There are ηK frequency bands that are occupied by both sources where $0 \leq \eta \leq 1$ is the overlapping ratio, and $\frac{1-\eta}{2} K$ distinct frequency bands are occupied by each source

separately. Assume equal power allocation across frequency bands for each source.

Since the configuration is symmetric, it suffices to analyze the reconstruction for the rth source, $r = 1, 2$, and we have $\xi_{ij} = \xi_{ij}^{(r)}$, $\xi_{ij}^{(r)} = \hat{\alpha}_{ij}^{(r)} - \rho^{(r)}(\boldsymbol{c}_{ij})$. Assume that the propagation field $\rho_k^{(r)}(\boldsymbol{z})$ in (1) is third-order differentiable. Then, if $\phi_k^{(r)}$ is available, we have the following results to characterize the construction error ξ_{ij} at a point $\boldsymbol{c}_{ij}$.

Theorem 2 (Spectrum overlap for multiple sources). *Let $\hat{\alpha}_{ij}^{(r)}$ be the estimate obtained from the solution to the integrated interpolation and BTD problem (14) under $\nu = 0$ based on the regression model (6). The variance $\mathcal{E}_m = \mathbb{V}\{\xi_{ij}\}$ of the interpolation error ξ_{ij} for the integrated interpolation and BTD approach is given by*

$$\mathcal{E}_m = \left(\varpi_\eta(\eta)\sigma_\eta^2 + \varpi_\epsilon(\eta)\sigma_\epsilon^2\right) C(\{\boldsymbol{z}_m, b\}) \quad (25)$$

where $\varpi_\eta(\eta) = (2 - 10\eta^2 + 10\eta - 2\eta^3)/(K(1 - 3\eta^2 + 2\eta)^2)$ *and* $\varpi_\epsilon(\eta) = 2(-\eta - 1)/(K(3\eta^2 - 2\eta - 1))$.

In addition, $\varpi_\eta(\eta), \varpi_\epsilon(\eta) \geq \frac{2}{K}$ with lower bound achieved when $\eta = 0$.

Proof. See Appendix F. $\qquad \square$

It is found that when $\eta = 0$, i.e., there is no overlap of frequency bands between each source, we have the coefficients $\varpi_\eta(\eta)$ and $\varpi_\epsilon(\eta)$ to be $2/K$, which is consistent with the results in Theorem 1 for a signal case, where each source occupies $K/2$ frequency bands. Increasing the overlapping ratio η also increases the coefficients $\varpi_\eta(\eta)$ and $\varpi_\epsilon(\eta)$ in (25), leading to a worse performance. When η reaches 1, we will mathematically have the matrix in (43) to be singular, and the estimation error can be arbitrarily large. In this case, one cannot separate the two sources.

B. Low-rank Regularization

The nuclear norm regularization in the third term of (9) provides two benefits: First, it allows sparse interpolation in the first term $l(f)$, where $\mathcal{I}$ in (8) may only contain a small set of grids to interpolate, and second, it imposes low-rank of the propagation field represented by the matrix $\boldsymbol{S}_r$. Note that while the interpolation $l(f)$ in (9) exploits the locally spatial correlation of the propagation field, the low-rank regularization tries to exploit the *global* structure, where the signal strength may decrease in all directions following roughly the *same* law as distance increases.

To numerically study the performance gain due to low-rank regularization, we select two common forms of propagation model. One is under exponential scale (EXP-model) $g(d) = \alpha \exp(-d^\beta)$, with $\alpha = 1$, $\beta = 1.5$, $h = 0.1$, another is under log-scale (LOG-model) $g(d) = \alpha - \beta \times \log_{10} d$, with $\alpha = 18$, $\beta = 5$, $h = 0.1$, where $d = \sqrt{x^2 + y^2 + h^2}$ represents the distance from the source at the origin, (x, y) is the coordinate of the grid cell, and h is the height. We choose the number of sources $R = 1$, dimension $N = N_1 = N_2 = 31$, the number of frequency bands $K = 20$, and the number of
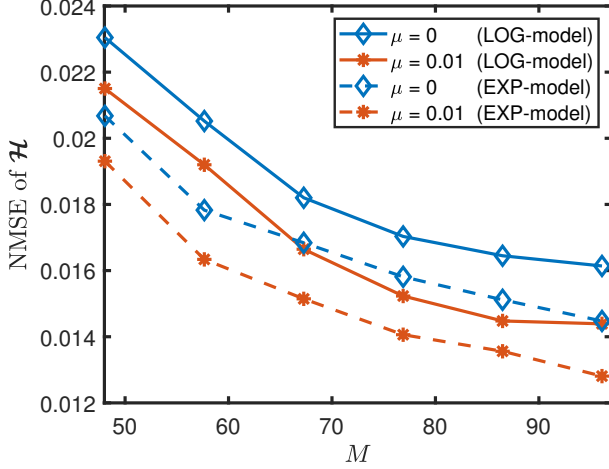
Figure 1. Comparison of the integrated interpolation and BTD approach with and without the nuclear nrom low-rank regularization, for parameters $\mu = 0.01$ and $\mu = 0$ respectively.

Figure 2. NMSE of the reconstructed source spectrum $\mathbf{\Phi}$ and the power spectrum map $\mathcal{H}$ versus the spectrum overlapping ratio η. The proposed scheme works much better in separating the two sources.

Figure 3. Reconstruction of the propagation field $\mathbf{S}_r$ of each source with the NMSE. The sources constructed by TPS-BTD tends to be over dispersed, whereas, the reconstruction of the proposed scheme appears to be more accurate.

measurements $M = \rho N^2$ with sampling ratio $\rho = 5\% - 10\%$. The performance criterion normalized mean squared error (NMSE) is the same as in Section V.

Fig. 1 compares two schemes, one with the low-rank regularization where $\mu = 0.01$ in (19), and the other without the low-rank regularization where $\mu = 0$. We choose $\nu = 10^{-4}$ in (14) for both the two schemes. Fig. 1 shows that the low-rank regularization can enhance the accuracy in the reconstruction by more than 10% for both propagation models compared to the case without regularization under the sparse observations. It is observed that the proposed scheme with low-rank regularization outperforms an interpolation scheme without the low-rank regularization.

C. Identifiability of Multiple Sources with Overlapping Spectrum

One important feature of the proposed method is the capability of identifying multiple sources possibly with overlapping spectrum. This feature was also exploited by the related works based on tensor decomposition in the literature [26], [29]. However, the proposed scheme has the advantage that it does not require a priori knowledge of the rank of $\mathbf{S}_r$. By contrast, existing schemes based on the BTD model require the rank information as $\mathbf{S}_r$ needs to be decomposed into $\mathbf{S}_r = \mathbf{A}_r \mathbf{B}_r^T$ [26], [29]. Estimating the rank parameter L for the existing schemes can be challenging, because a small L may lose some accuracy, whereas, a large L may lead to identifiability issue.

To numerically analyze the identifiability for multiple sources with overlapping spectrum, we choose the overlapping ratio $\eta = 0 - 0.6$, then, the number of overlapped frequency bands is ηK with a total $K(1+\eta)/2$ frequency bands observed of each source. We set $\phi_k^{(r)} = 1$, and then normalize $\phi^{(r)}$ for comparison. We choose $R = 2$, $K = 20$, $\rho = 5\%$, other settings are the same as in Section V. The proposed method is stable with respect to initialization. For $\phi^{(r)}$, we randomly generate its values and then normalize its l_2-norm to 1. For $\mathbf{S}_r$, we directly set all values to 0, which is sufficient. Since the
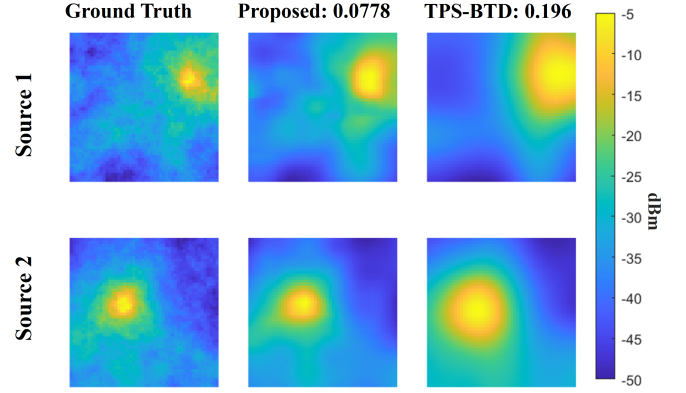
TPS method reconstructs the whole tensor without separating the source, we only compare the proposed method with the TPS-BTD baseline in Section V.

Fig. 2(a) demonstrates that the proposed method achieves more than double the improvement in spectrum $\mathbf{\Phi}$ reconstruction compared to the TPS-BTD baseline. Furthermore, as demonstrated in Fig. 2(b), it also provides an enhancement in power spectrum map $\mathcal{H}$ reconstruction exceeding 10%. This highlights its ability to identify multiple sources, even with overlapping spectra, compared to the baseline method. The TPS-BTD method underperforms primarily because selecting an appropriate rank L is challenging. Fig. 3 displays a clear visualization of the propagation field $\mathbf{S}_r$ under $\eta = 0.2$ where there is also an accuracy improvement of more than 10% in reconstructing $\mathbf{S}_r$ compared to the baseline method. Fig. 4 shows a visual plot of the reconstructed spectrum $\mathbf{\Phi}$ and Fig. 5 shows the reconstructed power spectrum map $\mathcal{H}$ with representative frequency bands. Furthermore, recoverability can also be studied following the techniques derived in [27, Theorem 2], [35]. Here, we provide numerical experiments on the recoverability in Section V.

V. NUMERICAL RESULTS

We adopt model (4) to simulate the power spectrum map in an $L \times L$ area for $L = 51$ meters, where $g_r(d) = P_r(C_0/d)^2$

Figure 4. Reconstruction of the spectrum $\mathbf{\Phi}$.

Figure 6. Reconstruction NMSE of $\mathcal{H}$ when observing the full spectrum or when randomly and sparsely observing part of the spectrum.

Figure 5. Ground truth and reconstructed power spectrum map $\mathcal{H}$ with representative frequency bands.

follows Friis transmission equation, d represents the distance from the source. We choose the parameter $P_r = 1\text{W}$, $C_0 = 2$ for illustrative purpose. Other values are broadly similar. The number of sources is $R = 2$. The power spectrum $\phi_k^{(r)}$ is generated by $\phi_k^{(r)} = \sum_{i=1}^{2} a_i^{(r)} \text{sinc}^2(k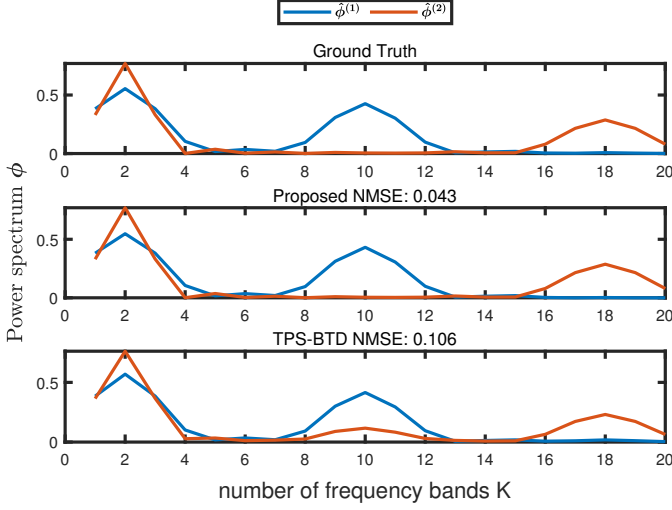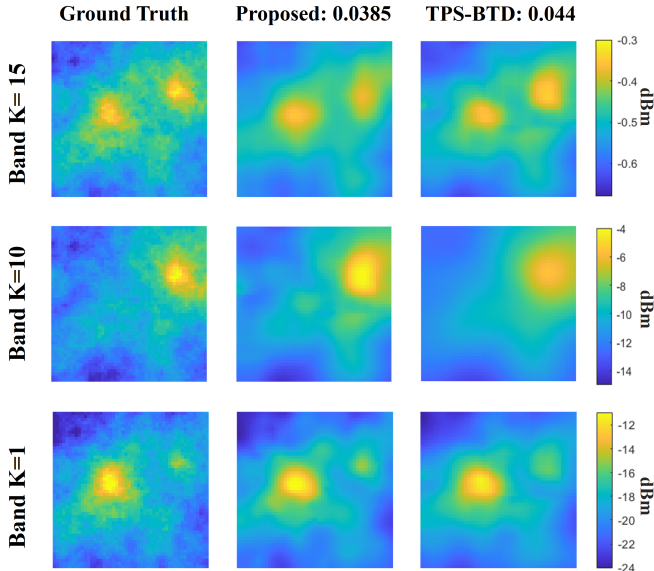 - f_i^{(r)} / b_i^{(r)})$, where $a_i^{(r)} \sim \mathcal{U}(0.5, 2)$, $f_i^{(r)} \in \{1, \cdots, K\}$ is the center of the i-th square sinc function, $b_i^{(r)} \sim \mathcal{U}(2, 4)$. The sensors are distributed uniformly at random in the $L \times L$ continuous space to collect the signal power $\gamma_m^{(k)}$, thus the measurements are likely off-grid. The shadowing component in log-scale $\log_{10}\zeta$ is modeled using a Gaussian process with zero mean and auto-correlation function $\mathbb{E}\{\log_{10}\zeta(\boldsymbol{z}_i)\log_{10}\zeta(\boldsymbol{z}_j)\} = \sigma_s^2 \exp(-||\boldsymbol{z}_i - \boldsymbol{z}_j||_2 / d_c)$, in which correlation distance $d_c = 30$ meters, shadowing variance $\sigma_s = 4$. The SNR is defines as $\text{SNR} = \gamma_m^{(k)} / \tilde{\epsilon}_k$ where the $\tilde{\epsilon}_k$ follows Gaussian distribution $\mathcal{N}(0, \sigma^2)$ and σ^2 is chosen to make the SNR 20dB. We select the parameter

b in the kernel function to ensure that it contains at least $M_0 = 14$ sensors. This is because the regression model in (6) contains 12 unknown parameters, and a number of $12 + 2 = 14$ measurements will be enough to estimate the regression model. The index set $\mathcal{I}$ is constructed through randomly and uniformly selecting among 80% of the grids, at the same time avoiding scenarios where an entire column or row of $\boldsymbol{S}_r$ is missing

We employ the NMSE of the reconstructed power spectrum map for performance evaluation. Let the NMSE of tensor $\mathcal{H}$ be $||\hat{\mathcal{H}} - \mathcal{H}||_F^2 / ||\mathcal{H}||_F^2$, the NMSE of $\mathbf{\Phi}$ and $\boldsymbol{S}_r$ is of the same form. The performance is compared with the following baselines that are recently developed or adopted in related literature. Baseline 1: Thin plate spline (TPS) [36]. Baseline 2: Nonnegative matrix factorization-TPS (NMF-TPS), we solve $||\mathbf{\Gamma} - \boldsymbol{S}_\Omega \mathbf{\Phi}^T||$, where $\mathbf{\Gamma} \in \mathbb{R}^{M \times K}$ is the collection of observations and $\boldsymbol{S}_\Omega \in \mathbb{R}^{M \times R}$, through NMF. The NMF is realized based on the successive projection algorithm (SPA) algorithm, as described in Appendix G of the supplementary materials of [27]. Subsequently, we use TPS to reconstruct R propagation fields $\boldsymbol{S}_r \in \mathbb{R}^{N_1 \times N_2}$ from $\boldsymbol{S}_\Omega$. Baseline 3: TPS-BTD [29], we first perform TPS, then, the uncertainty is derived and imposed as the restriction on the BTD method. The performance of low-rank tensor completion (LRTC) [37] is poor for power spectrum map reconstruction [26], [27]; therefore, it is not considered as a baseline.

A. The Influence of Number of Measurements M

We quantify the power spectrum map reconstruction performance of the proposed method under different number of measurements $M = 130 - 260$ which is of sampling ratio $\rho = 5\% - 10\%$ under fixed resolution $N = N_1 = N_2 = 51$. We choose the number of frequency bands being $|\Omega_m| = K = 20$ for the full observation case and $|\Omega_m| = K/2 = 10$ for the sparse observation case, separately. For the sparse observation case, we randomly select the observed frequency bands for the sensors.

Figure 7. Reconstruction NMSE of $\mathcal{H}$ under non-homogeneous observation of spectrum.

Figure 8. Reconstruction NMSE of $\mathcal{H}$ under various shadowing variance σ_s and correlation distance d_c.

Fig. 6 shows the performance of power spectrum reconstruction under both full and sparse observation scenarios. It demonstrates that the proposed method is robust when each sensor only observes random portion of the spectrum. The propos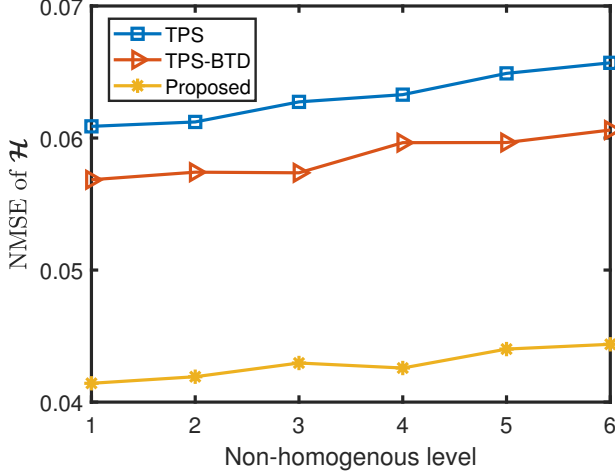ed scheme achieves over a 20% improvement in reconstruction accuracy in both cases, which translates to roughly 50% reduction of the measurements to achieve the same NMSE performance. A larger M will contribute to a performance gain of the proposed method. TPS underperforms because it fails to leverage the correlation property in the frequency domain. TPS-BTD, which relies on TPS interpolation, also falls short as it does not utilize the correlation property effectively. The NMF-TPS method executes NMF and TPS independently without an optimal integration, and the global TPS method lacks the capability to utilize local information effectively, resulting in suboptimal performance.

When the number of measurements M increase, the performance of the proposed method under the sparse observation case approaches the full observation case, revealing that exploit a tensor structure can save the number of measurements to attain a similar accuracy.

B. The Influence of Non-homogeneous Spectrum Observation

We investigate the influence of the non-homogeneous observation of spectrum on the performance under the sparse spectrum observation. In the homogeneous observation, for each sensor location z_m, we randomly select the $|\Omega_m| = K/2$ frequency bands from the whole bands. As a result, we will have a similar number of sensor measurements for each frequency band. In the non-homogeneous case, we still make that each frequency band k have a similar number of sensor measurements. But the difference lies in that for a set of sensor locations, they are with higher probability to collect the first $K/2$ bands, and for another set of sensor locations, they are with higher probability to collect the last $K/2$ bands.

To realize this, we separate the sensors locations $\{z_m\}$ to two sets $\mathcal{Z}_1$, $\mathcal{Z}_2$. Then, we choose the weighted sampling without replacement method to choose $|\Omega_m| = 10$ bands for each sensor from $K = 20$ bands. For the locations in $\mathcal{Z}_1$, we choose the weight to be $w_k^{(1)} = 1$ for $1 \leq k \leq 10$, and $w_k^{(1)} = C$ for $11 \leq k \leq 20$. For the locations in $\mathcal{Z}_2$, we choose the weight to be $w_k^{(2)} = C$ for $1 \leq k \leq 10$, and $w_k^{(2)} = 1$ for $11 \leq k \leq 20$. Therefore, the weight C serves as a measure of the level of non-homogeneity. For $C = 1$, it is homogeneous with identical weights. As C increases, the observations of the spectrum become increasingly non-homogeneous.
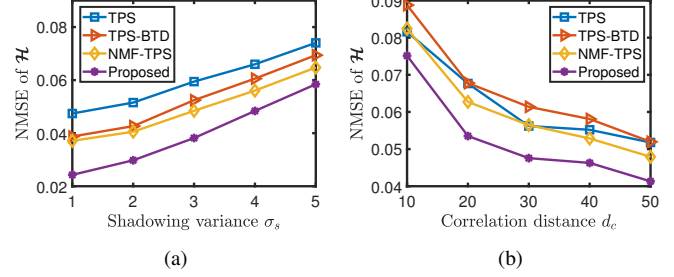
Fig. 7 illustrates that the proposed method surpasses the baseline methods and is stable under inhomogeneity of the measurement topology, achieving an improvement of around 20% in reconstruction accuracy. The NMF-TPS method is not suitable for such sparse and non-homogeneous case since a sparse Γ may hinder the separation of S_Ω and Φ^{T}. The non-homogeneous spectrum observation is different from the non-homogeneous sampling positions. Because, with a non-homogeneous spectrum, i.e., observing different spectrums at different locations, the resulting sparse Γ will hinder separation. This is because the separability of NMF is only guaranteed under a dense observation matrix. The proposed method effectively handles extremely non-homogeneous spectrum cases and still maintains a significant improvement over the baselines.

C. The Influence of Different Shadowing Parameters

We quantify the reconstruction performance of the proposed method under different shadowing variance $\sigma_s = 1 - 6$ and different correlation distance $d_c = 10 - 50$. We choose the full observation case $K = 20$, and $M = 130$ which corresponding to $\rho = 5\%$.

We first choose correlation distance $d_c = 30$ and varying σ_s from 1 to 6. Fig. 8(a) demonstrates the performance of the proposed method outperforms the baseline methods with an $18\% - 31\%$ improvement under the shadowing variances $\sigma_s = 1 - 6$. The increasing of the shadowing variance will cause a degradation in the performance of all the methods.

Then, we choose $\sigma_s = 4$ and varying d_c from 10 to 50. Fig. 8(b) demonstrates that the proposed method surpasses the baseline methods, showing an improvement of $14\% - 22\%$ across the correlation distances $d_c = 10 - 50$. Additionally, a larger correlation distance contributes to the performance gain.
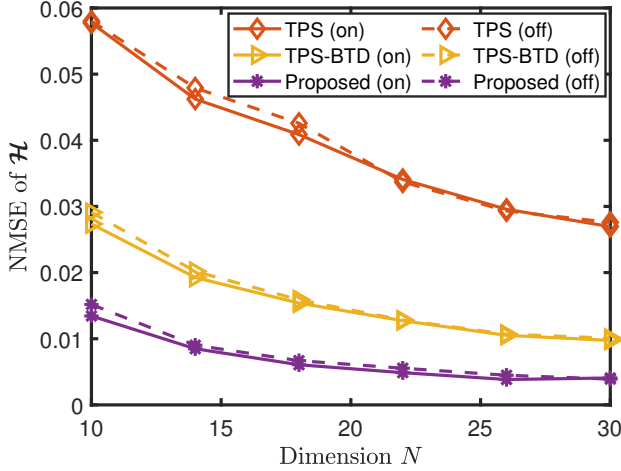
Figure 9. Reconstruction NMSE of $\mathcal{H}$ under off-grid and on-grid measurements.

Figure 10. Reconstruction NMSE of $\mathcal{H}$ under different R.

D. The Influence of Off-grid Measurements

For tensor completion, an off-grid measurements can cause large error and one generally prefers a measurement collecting from the grid center. The proposed method can naturally deal with the off-grid measurements case. Here, we numerically study the influence of off-grid measurements issue through comparing it with on-grid measurements to showcase the proposed method.

The simulation is performed under $N = 10 - 30$, $K = 20$, $M = C_1 N \log^2(N)$ with $C_1 = 2$. Fig. 9 shows that the proposed method outperforms baselines in over 20% improvement in reconstruction accuracy. In addition, the proposed method is stable under off-grid measurements which shows a similar performance between the on-grid and off-grid measurements. This is because the proposed method has an interpolation as a intermediate step, it can deal with the off-grid measurements. The performance between the off-grid and on-grid case approaches at large N.

E. The Influence of Number of Sources

We quantify the reconstruction performance of the proposed method under different number of sources $R = 2 - 3$ and different number of measurements $M = 52 - 260$ which is of sampling ratio $\rho = 2\% - 10\%$ under fixed resolution $N = N_1 = N_2 = 51$. The result as shown in Fig. 10 demonstrates that the proposed method outperforms baseline methods under different R, and an improvement of over 25% in reconstruction accuracy can be achieved. The reason that the NMSE value for $R = 3$ is smaller than $R = 2$ is that, extra source will bring into more energy in the field, thus increasing the values of $\mathcal{H}$ in $\|\hat{\mathcal{H}} - \mathcal{H}\|_F^2 / \|\mathcal{H}\|_F^2$.

F. Experiment under Real Data

We have tested our solution under the real data [38] collected in the hallways of the A5, A6 buildings on the campus of the University of Mannheim. We select the measurements from 5 frequency bands: 2.412GHz, 2.432GHz, 2.447GHz,
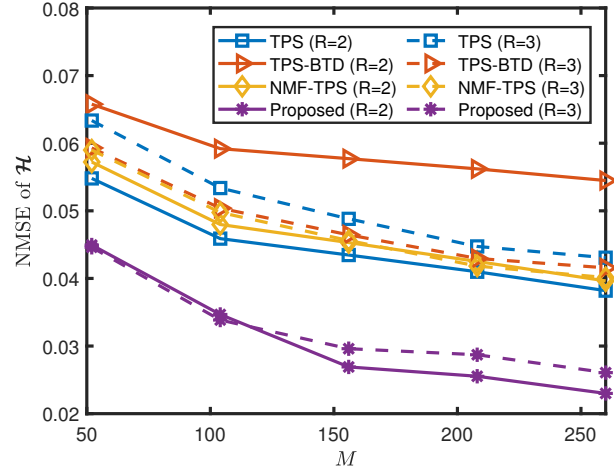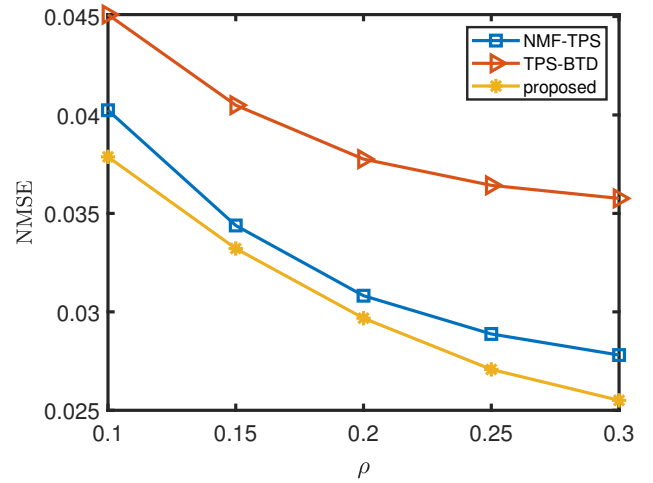
Figure 11. Reconstruction NMSE of real dataset under different sampling ratio $\rho = M/M_0$.

2.457GHz, and 2.462GHz. In this dataset, the indoor region is a $14 \times 34\text{m}^2$ area which is divided into $1 \times 1\text{m}^2$ grids and there are $M_0 = 166$ RSS measurements for each emitter. The sampling ratio was varied from 10% to 30%. We choose a value of $\mu = 3.2$ to ensure the threshold of SVT algorithm preserves at least 99% information of the matrix. The simulation result is shown in Fig. 11. To see the visual plot, we choose a sampling ratio of $\rho = 10\%$, utilizing measurements from 17 locations, as shown in Fig. 12.

The proposed method outperforms the TPS-BTD and NMF-TPS baselines with an improvement over 3% under different sampling ratio. The visual plot is based on a sampling ratio of 10%, utilizing measurements from only 17 locations to reconstruct the entire tensor. The visual performance of the proposed method is promising, and more details can be reconstructed.

VI. CONCLUSION

This paper developed a novel integrated interpolation and BTD approach for reconstructing power spectrum maps. This
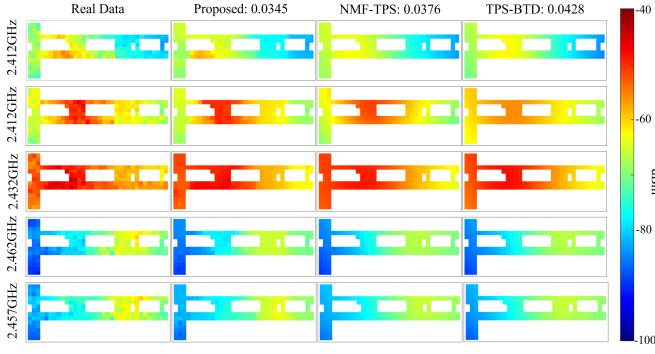
Figure 12. Visual plot of reconstruction performance under real dataset.

integrated approach addressed the spatial correlation of power spectrum maps, captured their low-rank characteristics, and reduced errors from off-grid measurements. An algorithm based on alternating least squares and singular value thresholding was developed. Theoretical analysis showed that the incorporation of the BTD structure improved interpolation. Extensive simulations were conducted, demonstrating that the proposed method achieved over 10% improvement in reconstruction accuracy across various scenarios, such as different numbers of sensors, measurement topology homogeneity, shadowing parameters, and off-grid scenarios.

APPENDIX A
PROOF OF THEOREM 1

Under $\nu = 0$, the solution for the weighted least-squares problem (14) is

$$\hat{\boldsymbol{\Theta}}_{ij} = (\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}})^{-1} \boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \mathrm{vec}(\boldsymbol{\Gamma}). \quad (26)$$

Denote $\boldsymbol{\gamma}_k = [\gamma_1^{(k)}, \gamma_2^{(k)}, \cdots, \gamma_M^{(k)}]^{\mathrm{T}} \in \mathbb{R}^M$ and recall that $\boldsymbol{\Gamma}(m,k) = \gamma_m^{(k)}$. From (4), we have,

$$\begin{aligned}
&\mathrm{vec}(\boldsymbol{\Gamma}) \\
&= [\boldsymbol{\gamma}_1^{\mathrm{T}}, \boldsymbol{\gamma}_2^{\mathrm{T}}, \cdots, \boldsymbol{\gamma}_K^{\mathrm{T}}]^{\mathrm{T}} \\
&= [\rho(\boldsymbol{z}_1)\phi_1 \cdots \rho(\boldsymbol{z}_M)\phi_1, \cdots, \rho(\boldsymbol{z}_1)\phi_K \cdots \rho(\boldsymbol{z}_M)\phi_K]^{\mathrm{T}} + \tilde{\boldsymbol{\epsilon}} \\
&= \tilde{\boldsymbol{\rho}} + \tilde{\boldsymbol{\epsilon}} \quad (27)
\end{aligned}$$

where $\tilde{\boldsymbol{\epsilon}} = [\tilde{\epsilon}_1, \cdots, \tilde{\epsilon}_1, \cdots, \tilde{\epsilon}_K, \cdots, \tilde{\epsilon}_K]^{\mathrm{T}} \in \mathbb{R}^{MK \times 1}$.

Applying third order Taylor's expansion to $\rho(\boldsymbol{z}_m)$ at the neighborhood of $\boldsymbol{c}_{ij}$, we have

$$\begin{aligned}
\rho(\boldsymbol{z}_m) = {} & \rho(\boldsymbol{c}_{ij}) + (\nabla\rho(\boldsymbol{c}_{ij}))^{\mathrm{T}}(\boldsymbol{c}_{ij})(\boldsymbol{z}_m - \boldsymbol{c}_{ij}) \\
& + \frac{1}{2}(\boldsymbol{z}_m - \boldsymbol{c}_{ij})^{\mathrm{T}}\nabla^2\rho(\boldsymbol{c}_{ij})(\boldsymbol{z}_m - \boldsymbol{c}_{ij}) \\
& + P_3(\boldsymbol{z}_m - \boldsymbol{c}_{ij}) + o(\|\boldsymbol{z}_m - \boldsymbol{c}_{ij}\|^3) \quad (28)
\end{aligned}$$

where $\nabla\rho(\boldsymbol{c}_{ij})$ is the first order derivative of $\rho(\boldsymbol{z})$ evaluated at point $\boldsymbol{c}_{ij}$, $\nabla^2\rho(\boldsymbol{c}_{ij})$ is the Hessian matrix of $\rho(\boldsymbol{z})$ evaluated at point $\boldsymbol{c}_{ij}$, $o(x)$ means $\lim_{x \to 0} o(x)/x \to 0$, $P(\boldsymbol{z}_m - \boldsymbol{c}_{ij}) = \sum_{k_1+k_2=3} \frac{\nabla^{(k_1+k_2)}\rho(\boldsymbol{c}_{ij})}{k_1!k_2!}(\boldsymbol{z}_m(1) - \boldsymbol{c}_{ij}(1))^{k_1}(\boldsymbol{z}_m(2) - \boldsymbol{c}_{ij}(2))^{k_2}$ represents third order component, $\boldsymbol{z}_m(n)$ represent the nth element in $\boldsymbol{z}_m$.

Denote $\boldsymbol{P}_{ij} = [P(\boldsymbol{z}_1 - \boldsymbol{c}_{ij}), \cdots, P(\boldsymbol{z}_M - \boldsymbol{c}_{ij})]^{\mathrm{T}}$ and recall $\boldsymbol{\Phi} = [\phi_1, \phi_2, \cdots, \phi_K]$. Then, the expression of $\tilde{\boldsymbol{\rho}}$ in (27) can be rearranged into the following matrix form

$$\tilde{\boldsymbol{\rho}} = \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}} \begin{bmatrix} \rho(\boldsymbol{c}_{ij}) \\ \nabla\rho(\boldsymbol{c}_{ij}) \\ \nabla^2\rho(\boldsymbol{c}_{ij}) \end{bmatrix} + \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{P}_{ij} + \boldsymbol{r}_{ij} \quad (29)$$

where $\boldsymbol{r}_{ij}$ is a vector of the residual terms $o(\|\boldsymbol{z}_m - \boldsymbol{c}_{ij}\|^3)$.

Since $\mathbb{E}\{\tilde{\epsilon}\} = 0$, the expectation of $\hat{\boldsymbol{\Theta}}_{ij}$ in (26) can be written as

$$\begin{aligned}
& \mathbb{E}\{\hat{\boldsymbol{\Theta}}_{ij}\} \\
&= \mathbb{E}\{(\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}})^{-1} \boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \mathrm{vec}(\boldsymbol{\Gamma})\} \\
&= (\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}})^{-1} \boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \tilde{\boldsymbol{\rho}}.
\end{aligned}$$

The variance of $\hat{\boldsymbol{\Theta}}_{ij}$ can be derived as follows:

$$\begin{aligned}
& \mathbb{V}\{\hat{\boldsymbol{\Theta}}_{ij}\} \\
&= \mathbb{E}\{(\hat{\boldsymbol{\Theta}}_{ij} - \mathbb{E}\{\hat{\boldsymbol{\Theta}}_{ij}\})^2\} \\
&= \mathbb{E}\{((\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}})^{-1} \boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \mathrm{vec}(\boldsymbol{\Gamma}) \\
&\quad\quad - (\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}})^{-1} \boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \tilde{\boldsymbol{\rho}})\} \\
&= \mathbb{E}\{((\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}})^{-1} \boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \tilde{\boldsymbol{\epsilon}})^2\} \\
&= \mathbb{E}\{(\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}})^{-1}(\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \tilde{\boldsymbol{\epsilon}} \\
&\quad \times \tilde{\boldsymbol{\epsilon}}^{\mathrm{T}} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}})(\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}})^{-1}\}.
\end{aligned}$$
$$(30)$$

Since each sensor measures all the frequency bands, $\boldsymbol{W}_{ij}$, defined in (12), is identity matrix. Through simple matrix multiplication, the matrix $\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}}$ can be derived as follows:

$$(\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}})^{-1} = \left(\sum_{k=1}^K \phi_k^2 \boldsymbol{A}_1 \right)^{-1} \quad (31)$$

where

$$\boldsymbol{A}_1 = \begin{bmatrix} a_1 & \boldsymbol{b}_1 & \boldsymbol{c}_1 \\ \boldsymbol{b}_1^{\mathrm{T}} & \boldsymbol{D}_1 & \boldsymbol{E}_1 \\ \boldsymbol{c}_1^{\mathrm{T}} & \boldsymbol{E}_1^{\mathrm{T}} & \boldsymbol{F}_1 \end{bmatrix},$$

with $a_1 = \sum_m \kappa_{ij}(\boldsymbol{z}_m)$, $\boldsymbol{b}_1 = \sum_m \kappa_{ij}(\boldsymbol{z}_m)(\boldsymbol{z}_m - \boldsymbol{c}_{ij})^{\mathrm{T}}$, $\boldsymbol{c}_1 = \sum_m \kappa_{ij}(\boldsymbol{z}_m)\mathrm{vec}((\boldsymbol{z}_m - \boldsymbol{c}_{ij})(\boldsymbol{z}_m - \boldsymbol{c}_{ij})^{\mathrm{T}})^{\mathrm{T}}$, $\boldsymbol{D}_1 = \sum_m (\boldsymbol{z}_m - \boldsymbol{c}_{ij})\kappa_{ij}(\boldsymbol{z}_m)(\boldsymbol{z}_m - \boldsymbol{c}_{ij})^{\mathrm{T}}$, $\boldsymbol{E}_1 = \sum_m (\boldsymbol{z}_m - \boldsymbol{c}_{ij})\kappa_{ij}(\boldsymbol{z}_m)\mathrm{vec}((\boldsymbol{z}_m - \boldsymbol{c}_{ij})(\boldsymbol{z}_m - \boldsymbol{c}_{ij})^{\mathrm{T}})^{\mathrm{T}}$, $\boldsymbol{F}_1 = \sum_m \mathrm{vec}((\boldsymbol{z}_m - \boldsymbol{c}_{ij})(\boldsymbol{z}_m - \boldsymbol{c}_{ij})^{\mathrm{T}})\kappa_{ij}(\boldsymbol{z}_m)\mathrm{vec}((\boldsymbol{z}_m - \boldsymbol{c}_{ij})(\boldsymbol{z}_m - \boldsymbol{c}_{ij})^{\mathrm{T}})^{\mathrm{T}}$. Here, $\boldsymbol{A}_1$ is a topology matrix that only depends on the sample locations $\boldsymbol{z}_m$, the grid position $\boldsymbol{c}_{ij}$, and the kernel function κ_{ij}. In addition,

$$\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \tilde{\boldsymbol{\epsilon}} \tilde{\boldsymbol{\epsilon}}^{\mathrm{T}} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}} = \sum_{k=1}^K \phi_k^2(\phi_k^2 \eta_{r,k}^2 + \epsilon^2) \boldsymbol{A}_2 \quad (32)$$

where

$$\boldsymbol{A}_2 = \begin{bmatrix} a_2 & \boldsymbol{b}_2 & \boldsymbol{c}_2 \\ \boldsymbol{b}_2^{\mathrm{T}} & \boldsymbol{D}_2 & \boldsymbol{E}_2 \\ \boldsymbol{c}_2^{\mathrm{T}} & \boldsymbol{E}_2^{\mathrm{T}} & \boldsymbol{F}_2 \end{bmatrix},$$

with $a_2 = \sum_m \kappa_{ij}^2(\boldsymbol{z}_m)$, $\boldsymbol{b}_2 = \sum_m \kappa_{ij}^2(\boldsymbol{z}_m)(\boldsymbol{z}_m - \boldsymbol{c}_{ij})^{\mathrm{T}}$, $\boldsymbol{c}_2 = \sum_m \kappa_{ij}^2(\boldsymbol{z}_m)\mathrm{vec}((\boldsymbol{z}_m - \boldsymbol{c}_{ij})(\boldsymbol{z}_m - \boldsymbol{c}_{ij})^{\mathrm{T}})^{\mathrm{T}}$, $\boldsymbol{D}_2 = \sum_m (\boldsymbol{z}_m - \boldsymbol{c}_{ij})\kappa_{ij}^2(\boldsymbol{z}_m)(\boldsymbol{z}_m - \boldsymbol{c}_{ij})^{\mathrm{T}}$, $\boldsymbol{E}_2 = \sum_m (\boldsymbol{z}_m - \boldsymbol{c}_{ij})\kappa_{ij}^2(\boldsymbol{z}_m)\mathrm{vec}((\boldsymbol{z}_m - \boldsymbol{c}_{ij})(\boldsymbol{z}_m - \boldsymbol{c}_{ij})^{\mathrm{T}})^{\mathrm{T}}$, $\boldsymbol{F}_2 = $

$\sum_m \text{vec}((\boldsymbol{z}_m - \boldsymbol{c}_{ij})(\boldsymbol{z}_m - \boldsymbol{c}_{ij})^{\mathsf{T}})\kappa_{ij}^2(\boldsymbol{z}_m)\text{vec}((\boldsymbol{z}_m - \boldsymbol{c}_{ij})(\boldsymbol{z}_m - \boldsymbol{c}_{ij})^{\mathsf{T}})^{\mathsf{T}}$. Likewise, $\boldsymbol{A}_2$ is also a topology matrix that captures the impact from the sample locations.

Thus, Equation (30) is further derived as follows:

$$\mathbb{V}\{\hat{\boldsymbol{\Theta}}_{ij}\}$$

$$= \mathbb{E}\left\{ \left(\sum_{k=1}^K \phi_k^2 \boldsymbol{A}_1\right)^{-1} \sum_{k=1}^K \phi_k^2(\phi_k^2 \eta_{r,k}^2 + \epsilon^2)\boldsymbol{A}_2 \right.$$

$$\left. \times \left(\sum_{k=1}^K \phi_k^2 \boldsymbol{A}_1\right)^{-1} \right\}$$

$$= \frac{1}{(\sum_{k=1}^K \phi_k^2)^2} \left[\boldsymbol{A}_1^{-1}\boldsymbol{A}_2\boldsymbol{A}_1^{-1}\right]_{(1,1)} \mathbb{E}\left\{\sum_{k=1}^K \phi_k^2(\phi_k^2 \eta_{r,k}^2 + \epsilon^2)\right\}. \tag{33}$$

Since $\eta_{r,k}(\boldsymbol{z}) \sim \mathcal{N}(0, \sigma_\eta^2)$ and $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$ in (1) and (2), we have

$$\mathbb{E}\left\{\sum_{k=1}^K \phi_k^2(\phi_k^2 \eta_{r,k}^2 + \epsilon^2)\right\}$$

$$= \sum_{k=1}^K \phi_k^2 \phi_k^2 \mathbb{E}\left\{\eta_{r,k}^2\right\} + \sum_{k=1}^K \phi_k^2 \mathbb{E}\left\{\epsilon^2\right\}$$

$$= \sum_{k=1}^K \phi_k^4 \sigma_\eta^2 + \sum_{k=1}^K \phi_k^2 \sigma_\epsilon^2. \tag{34}$$

Since $\boldsymbol{\Theta}_{ij}$ is deterministic, we have $\mathbb{V}\{\hat{\boldsymbol{\Theta}}_{ij} - \boldsymbol{\Theta}_{ij}\} = \mathbb{V}\{\hat{\boldsymbol{\Theta}}_{ij}\}$. Under single source case, we have $[\boldsymbol{\Theta}_{ij}]_{(1,1)} = \alpha_{ij}$. Thus, from (33) and (34), the variance of interpolation error ξ_{ij} becomes

$$\mathbb{V}\{\xi_{ij}\} = \left[\mathbb{V}\{\hat{\boldsymbol{\Theta}}_{ij} - \boldsymbol{\Theta}_{ij}\}\right]_{(1,1)}$$

$$= \left(\frac{\sum_{k=1}^K \phi_k^4}{\left(\sum_{k=1}^K \phi_k^2\right)^2}\sigma_\eta^2 + \frac{1}{\sum_{k=1}^K \phi_k^2}\sigma_\epsilon^2\right) C(\{\boldsymbol{z}_m\}, b)$$

where $C(\{\boldsymbol{z}_m\}, b) = \left[\boldsymbol{A}_1^{-1}\boldsymbol{A}_2\boldsymbol{A}_1^{-1}\right]_{(1,1)}$ is a function of the sensor locations $\boldsymbol{z}_m$ and window size b. This leads to the result in Theorem 1.

APPENDIX B
PROOF OF PROPOSITION 1

For the error term related to frequency-selective fading σ_η^2, we have

$$\varpi_\eta(\boldsymbol{\Phi}) \triangleq \frac{\sum_{k=1}^K \phi_k^4}{\left(\sum_{k=1}^K \phi_k^2\right)^2} = \frac{1}{1 + 2\frac{\sum_{i \neq j}\phi_i^2\phi_j^2}{\sum_{k=1}^K \phi_k^4}}.$$

Then, a larger value of $\sum_{i \neq j}\phi_i^2\phi_j^2/\sum_{k=1}^K \phi_k^4$ will contribute to a smaller value of $\varpi_\eta(\boldsymbol{\Phi})$, vice versa. It can be verified easily that when there is equal-power allocation, i.e., $\phi_i = \phi_j$, $\forall i \neq j$, $\sum_{i \neq j}\phi_i^2\phi_j^2/\sum_{k=1}^K \phi_k^4$ attains its largest value $(K-1)/2$, thus $\varpi_\eta(\boldsymbol{\Phi})$ attains its smallest value $\frac{1}{K}$. When there is a power boost in kth frequency band, i.e., ϕ_k approaches K and others approaches 0, $\sum_{i \neq j}\phi_i^2\phi_j^2/\sum_{k=1}^K \phi_k^4$ will attain its smallest value 0, thus $\varpi_\eta(\boldsymbol{\Phi})$ attains its largest value 1.

For the error term related to measurement noise σ_ϵ^2, we have

$$\varpi_\epsilon(\boldsymbol{\Phi}) \triangleq \frac{1}{\sum_{k=1}^K \phi_k^2}.$$

It can be verified easily that it attains the largest value $1/K$ when there is equal-power allocation, i.e., $\phi_i = \phi_j$, $\forall i \neq j$ and the smallest value $1/K^2$ when there is a power boost in kth frequency band, i.e., ϕ_k approaches K and others approaches 0. This leads to the result in Proposition 1.

APPENDIX C
PROOF OF PROPOSITION 2

The derivation follows the same approach for proving Theorem 1, resulting in the local polynomial interpolation error for the kth frequency band as

$$\mathbb{V}\{\xi_{ij}^{(k)}\} = \left(\sigma_\eta^2 + \frac{\sigma_\epsilon^2}{\phi_k^2}\right)C(\{\boldsymbol{z}_m\}, b)$$

where the result coincides with (22) under $K = 1$. Note that since the frequency-selective fading $\eta_{r,k}$ in (1) and the measurement noise ϵ in (2) are assumed to be independent across k, given the sensor topology $\boldsymbol{z}_m$ and $\boldsymbol{\phi}$, the interpolation error $\xi_{ij}^{(k)}$ is also independent across k. Therefore, the averaged interpolation error of all the K frequency bands is

$$\mathbb{V}\{\xi_{ij}\} = \frac{\sum_{k=1}^K \mathbb{V}\{\xi_{ij}^{(k)}\}}{K} = \left(\sigma_\eta^2 + \sum_{k=1}^K \frac{1}{\phi_k^2 K}\sigma_\epsilon^2\right)C(\{\boldsymbol{z}_m\}, b).$$

APPENDIX D
PROOF OF PROPOSITION 3

We have the difference value of the variance of interpolation error between the integrated interpolation and BTD approach and conventional frequency-by-frequency interpolation as follows:

$$\mathcal{E}_\text{p} - \mathcal{E}_\text{t}$$

$$= \sigma_\eta^2 + \sum_{k=1}^K \frac{1}{\phi_k^2 K}\sigma_\epsilon^2 - \frac{\sum_{k=1}^K \phi_k^4}{\left(\sum_{k=1}^K \phi_k^2\right)^2}\sigma_\eta^2 - \frac{1}{\sum_{k=1}^K \phi_k^2}\sigma_\epsilon^2$$

$$= \left(1 - \frac{\sum_{k=1}^K \phi_k^4}{(\sum_{k=1}^K \phi_k^2)^2}\right)\sigma_\eta^2 \tag{35}$$

$$+ \left(\sum_{k=1}^K \frac{1}{\phi_k^2 K} - \frac{1}{\sum_{k=1}^K \phi_k^2}\right)\sigma_\epsilon^2. \tag{36}$$

For (35), the coefficient of σ_η^2 can be derived as follows:

$$1 - \frac{\sum_{k=1}^K \phi_k^4}{(\sum_{k=1}^K \phi_k^2)^2} = \frac{2\sum_{i \neq j}\phi_i^2\phi_j^2}{(\sum_{k=1}^K \phi_k^2)^2} > 0.$$

For (36), the coefficient of σ_ϵ^2 can be derived as follows:

$$\left(\sum_{k=1}^K \frac{1}{\phi_k^2 K} - \frac{1}{\sum_{k=1}^K \phi_k^2}\right)$$

$$= \frac{1}{\sum_{k=1}^K \phi_k^2}\left(\sum_{k=1}^K \left(\frac{\sum_{l \neq k}\phi_l^2}{\phi_k^2 K} + \frac{1}{K}\right) - 1\right)$$

$$= \frac{1}{\sum_{k=1}^K \phi_k^2}\sum_{k=1}^K \frac{\sum_{l \neq k}\phi_l^2}{\phi_k^2 K} > 0.$$

Thus, $\mathcal{E}_{\mathrm{p}} > \mathcal{E}_{\mathrm{t}}$ and the difference is

$$\mathcal{E}_{\mathrm{p}} - \mathcal{E}_{\mathrm{t}} = \frac{2\sum_{i \neq j} \phi_i^2 \phi_j^2}{(\sum_{k=1}^{K} \phi_k^2)^2} \sigma_\eta^2 + \frac{\sigma_\epsilon^2}{\sum_{k=1}^{K} \phi_k^2} \sum_{k=1}^{K} \frac{\sum_{l \neq k}^{K} \phi_l^2}{\phi_k^2 K}.$$

APPENDIX E
PROOF OF PROPOSITION 4

The derivation is the same as proving Theorem 1 in Appendix A, except that, due to the sparse observation, (31) becomes

$$(\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}})^{-1} = (K' \boldsymbol{A}_1)^{-1}. \quad (37)$$

To verify the above result, using Equation (12) and the definition of $\boldsymbol{\Phi}$ and $\boldsymbol{X}_{ij}$, the first element in the matrix $\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}}$ is derived as follows:

$$\begin{aligned}
&\left[\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}}\right]_{(1,1)} \\
&= \sum_{m=1}^{M} \sum_{k=1}^{K} \phi_k^2 \psi(m,k) \kappa_{ij}(\boldsymbol{z}_m) \\
&= \sum_{m=1}^{M} \sum_{k=1}^{K} \psi(m,k) \kappa_{ij}(\boldsymbol{z}_m) \\
&= K' \sum_{m=1}^{M} \kappa_{ij}(\boldsymbol{z}_m)
\end{aligned} \quad (38)$$

which equals to $K' a_1$, where a_1 is defined in (31) and the last equation is because $|\Omega_m| = \sum_{k=1}^{K} \psi(m,k) = K'$.

Similarly, one can easily show that all the other elements in the matrix $\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}}$ equals to the corresponding element in $\boldsymbol{A}_1$ scaled by K'. Moreover, considering $\sum_{k=1}^{K} \psi(m,k) = K'$, Equation (32) becomes

$$\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \tilde{\boldsymbol{\epsilon}} \tilde{\boldsymbol{\epsilon}}^{\mathrm{T}} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}} = K'(\eta_{r,k}^2 + \epsilon^2) \boldsymbol{A}_2. \quad (39)$$

Then, the remaining part in Appendix A can be directly applied based on the modified quantities (37) and (39), leading to the following result $\mathbb{V}\{\xi_{ij}\} = ((\sigma_\eta^2 + \sigma_\epsilon^2)/K')C(\{\boldsymbol{z}_m\}, b)$ in Proposition 4.

APPENDIX F
PROOF OF THEOREM 2

Recall $\boldsymbol{\gamma}_k = [\gamma_1^{(k)}, \gamma_2^{(k)}, \cdots, \gamma_M^{(k)}]^{\mathrm{T}} \in \mathbb{R}^M$ and $\boldsymbol{\Gamma}(m,k) = \gamma_m^{(k)}$. From (4), under $R = 2$, we have $\gamma_m^{(k)} = \rho^{(1)}(\boldsymbol{z}_m)\phi_k^{(1)} + \tilde{\epsilon}_k + \rho^{(2)}(\boldsymbol{z}_m)\phi_k^{(2)} + \tilde{\epsilon}_k$. Thus,

$$\begin{aligned}
\mathrm{vec}(\boldsymbol{\Gamma}) &= [\boldsymbol{\gamma}_1^{\mathrm{T}}, \boldsymbol{\gamma}_2^{\mathrm{T}}, \cdots, \boldsymbol{\gamma}_K^{\mathrm{T}}]^{\mathrm{T}} \\
&= [\rho^{(1)}(\boldsymbol{z}_1)\phi_1^{(1)} + \rho^{(2)}(\boldsymbol{z}_1)\phi_1^{(2)}, \cdots \\
&\quad \rho^{(1)}(\boldsymbol{z}_M)\phi_1^{(1)} + \rho^{(2)}(\boldsymbol{z}_M)\phi_1^{(2)}, \cdots \\
&\quad \rho^{(1)}(\boldsymbol{z}_1)\phi_K^{(1)} + \rho^{(2)}(\boldsymbol{z}_1)\phi_K^{(2)}, \cdots \\
&\quad \rho^{(1)}(\boldsymbol{z}_M)\phi_K^{(1)} + \rho^{(2)}(\boldsymbol{z}_M)\phi_K^{(2)}] + \tilde{\boldsymbol{\epsilon}} \\
&= \tilde{\boldsymbol{\rho}} + \tilde{\boldsymbol{\epsilon}}
\end{aligned} \quad (40)$$

where $\tilde{\boldsymbol{\epsilon}} = [\tilde{\epsilon}_1, \cdots, \tilde{\epsilon}_1, \cdots, \tilde{\epsilon}_K, \cdots, \tilde{\epsilon}_K]^{\mathrm{T}} \in \mathbb{R}^{MK \times 1}$.

Similar to the derivation in (29), the expression of $\tilde{\boldsymbol{\rho}}$ in (40) can be rearranged into the following matrix form

$$\tilde{\boldsymbol{\rho}} = \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}} \boldsymbol{\rho}_{ij} + \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{P}_{ij} + \boldsymbol{r}_{ij} \quad (41)$$

where $\boldsymbol{P}_{ij} = [P^{(1)}(\boldsymbol{z}_1 - \boldsymbol{c}_{ij}), \cdots, P^{(1)}(\boldsymbol{z}_M - \boldsymbol{c}_{ij}), P^{(2)}(\boldsymbol{z}_1 - \boldsymbol{c}_{ij}), \cdots, P^{(2)}(\boldsymbol{z}_M - \boldsymbol{c}_{ij})]^{\mathrm{T}}$, $\boldsymbol{r}_{ij}$ is a vector of the residual terms $o(\|\boldsymbol{z}_m - \boldsymbol{c}_{ij}\|^3)$, and

$$\begin{aligned}
\boldsymbol{\rho}_{ij} = &[\rho^{(1)}(\boldsymbol{c}_{ij}), \mathrm{vec}(\nabla\rho^{(1)}(\boldsymbol{c}_{ij})), \mathrm{vec}(\nabla^2\rho^{(1)}(\boldsymbol{c}_{ij})), \\
&\rho^{(2)}(\boldsymbol{c}_{ij}), \mathrm{vec}(\nabla\rho^{(2)}(\boldsymbol{c}_{ij})), \mathrm{vec}(\nabla^2\rho^{(2)}(\boldsymbol{c}_{ij}))]^{\mathrm{T}}.
\end{aligned}$$

Similar as in Appendix A, the variance of the interpolation error ξ_{ij} is as follows:

$$\begin{aligned}
\mathbb{V}\{\xi_{ij}\} = &\Big[(\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}})^{-1} \\
&\times \mathbb{E}\{(\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \tilde{\boldsymbol{\epsilon}} \tilde{\boldsymbol{\epsilon}}^{\mathrm{T}} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}})\} \\
&\times (\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}})^{-1}\Big]_{(1,1)}. \quad (42)
\end{aligned}$$

Following a similar derivation in (37) and (39), we have

$$\begin{aligned}
&(\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}})^{-1} \\
&= \begin{bmatrix} K_1 \boldsymbol{A}_1 & K_2 \boldsymbol{A}_1 \\ K_2 \boldsymbol{A}_1 & K_1 \boldsymbol{A}_1 \end{bmatrix}^{-1}
\end{aligned} \quad (43)$$

where $K_1 = (1 + \eta)/2 \times K$, $K_2 = \eta K$, and

$$\begin{aligned}
&\mathbb{E}\{(\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \tilde{\boldsymbol{\epsilon}} \tilde{\boldsymbol{\epsilon}}^{\mathrm{T}} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}})\} \\
&= \begin{bmatrix} (M_1\sigma_\eta^2 + K_1\sigma_\epsilon^2)\boldsymbol{A}_2 & (M_2\sigma_\eta^2 + K_2\sigma_\epsilon^2)\boldsymbol{A}_2 \\ (M_2\sigma_\eta^2 + K_2\sigma_\epsilon^2)\boldsymbol{A}_2 & (M_1\sigma_\eta^2 + K_1\sigma_\epsilon^2)\boldsymbol{A}_2 \end{bmatrix}
\end{aligned}$$

where $M_1 = K_1 + K_2$, $M_2 = 2K_2$.

Assume $\boldsymbol{A}_1$ is invertible and the Schur complement of $\begin{bmatrix} K_1\boldsymbol{A}_1 & K_2\boldsymbol{A}_1 \\ K_2\boldsymbol{A}_1 & K_1\boldsymbol{A}_1 \end{bmatrix}$ is also invertible, then we have

$$\begin{aligned}
&(\boldsymbol{\Phi} \otimes \boldsymbol{X}_{ij} \boldsymbol{W}_{ij}^2 \boldsymbol{\Phi}^{\mathrm{T}} \otimes \boldsymbol{X}_{ij}^{\mathrm{T}})^{-1} \\
&= \begin{bmatrix} \frac{K_1}{K_1^2 - K_2^2}\boldsymbol{A}_1^{-1} & -\frac{K_2}{K_1^2 - K_2^2}\boldsymbol{A}_1^{-1} \\ -\frac{K_2}{K_1^2 - K_2^2}\boldsymbol{A}_1^{-1} & \frac{K_1}{K_1^2 - K_2^2}\boldsymbol{A}_1^{-1} \end{bmatrix} \triangleq \mathbf{A}.
\end{aligned}$$

Then, (42) equals

$$\left[\mathbf{A} \begin{bmatrix} (M_1\sigma_\eta^2 + K_1\sigma_\epsilon^2)\boldsymbol{A}_2 & (M_2\sigma_\eta^2 + K_2\sigma_\epsilon^2)\boldsymbol{A}_2 \\ (M_2\sigma_\eta^2 + K_2\sigma_\epsilon^2)\boldsymbol{A}_2 & (M_1\sigma_\eta^2 + K_1\sigma_\epsilon^2)\boldsymbol{A}_2 \end{bmatrix} \mathbf{A}\right]_{(1,1)}.$$

Since σ_η^2 and σ_ϵ^2 are independent, we can analyze $\mathbb{V}\{\xi_{ij}\}$ separately. Through simple matrix operations, we have

$$\begin{aligned}
&\left[\mathbf{A} \begin{bmatrix} K_1\sigma_\epsilon^2\boldsymbol{A}_2 & K_2\sigma_\epsilon^2\boldsymbol{A}_2 \\ K_2\sigma_\epsilon^2\boldsymbol{A}_2 & K_1\sigma_\epsilon^2\boldsymbol{A}_2 \end{bmatrix} \mathbf{A}\right]_{(1,1)} \\
&= \frac{-2\eta - 2}{K(3\eta^2 - 2\eta - 1)}\sigma_\epsilon^2 C(\{\boldsymbol{z}_m\}, b)
\end{aligned}$$

and

$$\begin{aligned}
&\left[\mathbf{A} \begin{bmatrix} M_1\sigma_\eta^2\boldsymbol{A}_2 & M_2\sigma_\eta^2\boldsymbol{A}_2 \\ M_2\sigma_\eta^2\boldsymbol{A}_2 & M_1\sigma_\eta^2\boldsymbol{A}_2 \end{bmatrix} \mathbf{A}\right]_{(1,1)} \\
&= \frac{2 - 10\eta^2 + 10\eta - 2\eta^3}{K(1 - 3\eta^2 + 2\eta)^2}\sigma_\eta^2 C(\{\boldsymbol{z}_m\}, b).
\end{aligned}$$

This concludes the result in Theorem 2.

REFERENCES

[1] D. Romero and S.-J. Kim, "Radio map estimation: A data-driven approach to spectrum cartography," *IEEE Signal Process. Mag.*, vol. 39, no. 6, pp. 53–72, 2022.

[2] D. Gesbert, O. Esrafilian, J. Chen, R. Gangula, and U. Mitra, "UAV-aided RF mapping for sensing and connectivity in wireless networks," *IEEE Wireless Commun.*, vol. 30, no. 4, pp. 116–122, 2023.

[3] J. Gao, D. Wu, F. Yin, Q. Kong, L. Xu, and S. Cui, "Metaloc: Learning to learn wireless localization," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 12, pp. 3831–3847, 2023.

[4] J. Chen and U. Mitra, "Unimodality-constrained matrix factorization for non-parametric source localization," *IEEE Trans. Signal Process.*, vol. 67, no. 9, pp. 2371–2386, May 2019.

[5] X. Mo, Y. Huang, and J. Xu, "Radio-map-based robust positioning optimization for UAV-enabled wireless power transfer," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 179–183, 2019.

[6] F. Shen, Z. Wang, G. Ding, K. Li, and Q. Wu, "3D compressed spectrum mapping with sampling locations optimization in spectrum-heterogeneous environment," *IEEE Trans. Wireless Commun.*, vol. 21, no. 1, pp. 326–338, 2022.

[7] C. Han and Y. Chen, "Propagation modeling for wireless communications in the terahertz band," *IEEE Commun. Mag.*, vol. 56, no. 6, pp. 96–101, 2018.

[8] B. Li and J. Chen, "Large timescale optimization for communications over aerial ad hoc networks with predetermined trajectories," *IEEE Trans. Commun.*, pp. 1–1, 2024.

[9] Y. Zeng, J. Chen, J. Xu, D. Wu, X. Xu, S. Jin, X. Gao, D. Gesbert, S. Cui, and R. Zhang, "A tutorial on environment-aware communications via channel knowledge map for 6G," *IEEE Commun. Surv. Tutor.*, pp. 1–1, 2024.

[10] K. Sato, K. Suto, K. Inage, K. Adachi, and T. Fujii, "Space-frequency-interpolated radio map," *IEEE Trans. Veh. Technol.*, vol. 70, no. 1, pp. 714–725, 2021.

[11] K. Sato and T. Fujii, "Kriging-based interference power constraint: Integrated design of the radio environment map and transmission power," *IEEE Trans. Cogn. Commun. Netw.*, vol. 3, no. 1, pp. 13–25, 2017.

[12] J. Fan, *Local polynomial modelling and its applications: Monographs on statistics and applied probability 66.* Routledge, 1996.

[13] Y. Zhang and S. Wang, "K-nearest neighbors gaussian process regression for urban radio map reconstruction," *IEEE Commun. Lett.*, vol. 26, no. 12, pp. 3049–3053, 2022.

[14] M. Hamid and B. Beferull-Lozano, "Non-parametric spectrum cartography using adaptive radial basis functions," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 2017, pp. 3599–3603.

[15] J. Wang, Q. Zhu, Z. Lin, Q. Wu, Y. Huang, X. Cai, W. Zhong, and Y. Zhao, "Sparse bayesian learning-based 3-D radio environment map construction-sampling optimization, scenario-dependent dictionary construction, and sparse recovery," *IEEE Trans. Cogn. Commun. Netw.*, vol. 10, no. 1, pp. 80–93, 2024.

[16] J. Wang, Q. Zhu, Z. Lin, J. Chen, G. Ding, Q. Wu, G. Gu, and Q. Gao, "Sparse bayesian learning-based hierarchical construction for 3D radio environment maps incorporating channel shadowing," *IEEE Trans. Wireless Commun.*, pp. 1–1, 2024.

[17] S.-J. Kim and G. B. Giannakis, "Cognitive radio spectrum prediction using dictionary learning," in *Proc. IEEE Global Telecomm. Conf.*, 2013, pp. 3206–3211.

[18] M. D. Migliore, "A compressed sensing approach for array diagnosis from a small set of near-field measurements," *IEEE Trans. Antennas Propag.*, vol. 59, no. 6, pp. 2127–2133, 2011.

[19] H. Sun and J. Chen, "Grid optimization for matrix-based source localization under inhomogeneous sensor topology," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 2021, pp. 5110–5114.

[20] D. Schaufele, R. L. G. Cavalcante, and S. Stanczak, "Tensor completion for radio map reconstruction using low rank and smoothness," in *Proc. IEEE Int. Workshop Signal Process. Adv. Wireless Commun.*, 2019, pp. 1–5.

[21] H. Xia, S. Zha, J. Huang, J. Liu, and P. Liu, "Spectrum cartography based on dynamic compressed sensing by using multiple domains information," *J. Commn. Net*, vol. 25, no. 4, pp. 507–515, 2023.

[22] Y. Teganya and D. Romero, "Deep completion autoencoders for radio map estimation," *IEEE Trans. Wireless Commun.*, vol. 21, no. 3, pp. 1710–1724, 2021.

[23] T. Hu, Y. Huang, J. Chen, Q. Wu, and Z. Gong, "3D radio map reconstruction based on generative adversarial networks under constrained aircraft trajectories," *IEEE Trans. Veh. Technol.*, vol. 72, no. 6, pp. 8250–8255, 2023.

[24] Y. Chen, S. Bhojanapalli, S. Sanghavi, and R. Ward, "Completing any low-rank matrix, provably," *J. Mach. Learn. Res.*, vol. 16, no. 1, pp. 2999–3034, 2015.

[25] H. Sun and J. Chen, "Energy-modified leverage sampling for radio map construction via matrix completion," *IEEE Signal Process. Lett.*, 2024, to appear, arXiv:2402.17138.

[26] G. Zhang, X. Fu, J. Wang, X.-L. Zhao, and M. Hong, "Spectrum cartography via coupled block-term tensor decomposition," *IEEE Trans. Signal Process.*, vol. 68, pp. 3660–3675, 2020.

[27] S. Shrestha, X. Fu, and M. Hong, "Deep spectrum cartography: Completing radio map tensors using learned neural models," *IEEE Trans. Signal Process.*, vol. 70, pp. 1170–1184, 2022.

[28] H. Sun and J. Chen, "Propagation map reconstruction via interpolation assisted matrix completion," *IEEE Trans. Signal Process.*, vol. 70, pp. 6154–6169, 2022.

[29] X. Chen, J. Wang, G. Zhang, and Q. Peng, "Tensor-based parametric spectrum cartography from irregular off-grid samplings," *IEEE Signal Process. Lett.*, vol. 30, pp. 513–517, 2023.

[30] H. Sun, J. Chen, and Y. Luo, "Tensor-guided interpolation for off-grid power spectrum map construction," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 2024, pp. 7295–7299.

[31] H. Sun and J. Chen, "Regression assisted matrix completion for reconstructing a propagation field with application to source localization," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Singapore, May 2022, pp. 3353–3357.

[32] A. Graham, *Kronecker products and matrix calculus with applications.* Courier Dover Publications, 2018.

[33] L. F. Portugal, J. J. Judice, and L. N. Vicente, "A comparison of block pivoting and interior-point algorithms for linear least squares problems with nonnegative variables," *Mathematics of Computation*, vol. 63, no. 208, pp. 625–643, 1994.

[34] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.

[35] S. Timilsina, S. Shrestha, and X. Fu, "Quantized radio map estimation using tensor and deep generative models," *IEEE Trans. Signal Process.*, vol. 72, pp. 173–189, 2024.

[36] J. A. Bazerque, G. Mateos, and G. B. Giannakis, "Group-lasso on splines for spectrum cartography," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4648–4663, 2011.

[37] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 208–220, 2013.

[38] T. King, S. Kopf, T. Haenselmann, C. Lubberger, and W. Effelsberg, "CRAWDAD dataset mannheim/compass (v. 2008-04-11)," Downloaded from https://crawdad.org/mannheim/compass/20080411/signalstrength, Apr. 2008, traceset: signalstrength.

Hao Sun (Student Member, IEEE) received the B.S. degree in Biomedical Engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2018. He received two national scholarships during his undergraduate study. He is currently working towards the Ph.D. degree with the School of Science and Engineering (SSE) and the Shenzhen Future Network of Intelligence Institute (FNii-Shenzhen), at the Chinese University of Hong Kong, Shenzhen (CUHK–Shenzhen), Guangdong, China. His research interests mainly include matrix completion and tensor decomposition methods with application to channel estimation.

Junting Chen (S'11–M'16) received the Ph.D. degree in Electronic and Computer Engineering from the Hong Kong University of Science and Technology (HKUST), Hong Kong SAR China, in 2015, and the B.Sc. degree in Electronic Engineering from Nanjing University, Nanjing, China, in 2009. From 2014–2015, he was a visiting student with the Wireless Information and Network Sciences Laboratory at MIT, Cambridge, MA, USA.

He is an Assistant Professor with the School of Science and Engineering and the Future Network of Intelligence Institute (FNii) at The Chinese University of Hong Kong, Shenzhen (CUHK–Shenzhen), Guangdong, China. Prior to joining CUHK–Shenzhen, he was a Postdoctoral Research Associate with the Ming Hsieh Department of Electrical Engineering, University of Southern California (USC), Los Angeles, CA, USA, from 2016–2018, and with the Communication Systems Department of EURECOM, Sophia–Antipolis, France, from 2015–2016. His research interests include channel estimation, MIMO beamforming, machine learning, and optimization for wireless communications and localization. His current research focuses on radio map sensing, construction, and application for wireless communications. Dr. Chen was a recipient of the HKTIIT Post-Graduate Excellence Scholarships in 2012. He was nominated as the Exemplary Reviewer of IEEE WIRELESS COMMUNICATIONS LETTERS in 2018. His paper received the Charles Kao Best Paper Award from WOCC 2022.