Propagation Map Reconstruction via Interpolation Assisted Matrix Completion

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

***Abstract*—Constructing a propagation map from a set of scattered measurements finds important applications in many areas, such as localization, spectrum monitoring and management. Classical interpolation-type methods have poor performance in regions with very sparse measurements. Recent advance in matrix completion has the potential to reconstruct a propagation map from sparse measurements, but the spatial resolution is limited. This paper proposes to integrate interpolation with matrix completion to exploit both the spatial correlation and the potential low rank structure of the propagation map. The proposed method first enriches matrix observations using interpolation, and develops the statistics of the interpolation error based on a local polynomial regression model. Then, two uncertainty-aware matrix completion algorithms are developed to exploit the interpolation error statistics. It is numerically demonstrated that the proposed method outperforms Kriging and other state-of-the-art schemes, and reduces the mean squared error (MSE) of propagation map reconstruction by 10%–50% for a medium to large number of measurements.**

***Index Terms*—Propagation map, interpolation, matrix completion, local polynomial regression, asymptotic analysis.**

I. Introduction

Sensing and reconstructing a propagation map find important applications in many areas. In source localization, constructing the energy propagation map may help identify the number of sources and localize the source in harsh environment [1], [2]. Constructing power spectrum maps, radio maps or coverage maps, provides assistance in wireless network planning, spectrum allocation, and interference management [3]–[5]. In antenna design, the interpolation and reconstruction of near field electromagnetic provides useful insights to understand the antenna radiation pattern which are used in testing and examining the performance and reliability of wireless devices [6], [7]. Applications of propagation map reconstruction are also found in bio-medical studies [8], [9] and environment science [10], [11].

It is challenging to reconstruct a propagation map from a set of sparse measurements due to the lack of accurate models to describe the propagation map. The problem has been studied for more than two decades, but it is still under active research. Recently developed data-driven approaches for propagation map reconstruction include Kriging [12], [13], kernel based method [14]–[17], dictionary learning [8], [18], matrix completion [19]–[22] and deep learning [23]–[26]. Kriging method predicts the signal strength at a given location by computing a weighted average of the measured signal strengths in the neighborhood of the location to be predicted. The weights are obtained through minimizing the Kriging covariance. However, the performance of the Kriging method is sensitive with the choice of the parametric form of the *semi-variogram* model in Kriging. Kernel methods [14]–[17] aim at transforming linearly inseparable data to linearly separable one using kernel functions as a proxy of the propagation map. For example, the received signal strength (RSS) at each location can be estimated using a weighted average of different radial basis function [17] where the parameters and weights are jointly optimized through an alternating minimization method. Kriging and kernel based methods rely on the assumption that the propagation map is locally smooth, and therefore, the propagation map can be interpolated and reconstructed using nearby measurements. It is not surprising that the reconstruction performance will deteriorate in regions where measurements are too sparse.

In contrast to interpolation-type methods, matrix completion approaches try to exploit the global structure of the propagation map. Specifically, the propagation map is discretized into grid cells and the signal strength measurements are arranged into a sparse matrix according to the measurements locations. The matrix is then completed using compressed sensing algorithms [19]–[22]. These methods are based on the assumption that the matrix representation of a 2D propagation map has a low rank property. Exploiting such a global structure of the propagation map, matrix completion methods improve the reconstruction performance in regions where measurements are very sparse. However, these methods have poor performance in reconstructing high resolution propagation map, because the finer the resolution, the lower percentage of the number of observed entries, and eventually, the matrix completion algorithm may fail if a complete row or column of the matrix is not observed.

Recent advance also attempts to develop deep learning techniques for propagation map reconstruction. In [23], a deep Gaussian Process is used to model the relationship between RSS measurement values and their corresponding locations. In [24], a fully convolutional deep completion autoencoder architecture is developed to learn the spatial structure of relevant

This is a preprint version. The paper has been accepted in IEEE Transactions on Signal Processing.

This work was supported in part by the National Key R&D Program of China with grant No. 2018YFB1800800, by the Basic Research Project No. HZQB-KCZYZ-2021067 of Hetao Shenzhen-HK S&T Cooperation Zone, by National Science Foundation of China No. 92067202 and No. 62171398, by the Shenzhen Science and Technology Program under Grant No. JCYJ20210324134612033 and KQTD20200909114730003, by Guangdong Research Grant No. 2019QN01X895, and by Guangdong Provincial Key Laboratory of Future Networks of Intelligence (Grant No. 2022B 1212010001).

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

propagation phenomena, and a UNet structure is employed in [25] for radio map reconstruction. However, these methods require a large amount of data to train the neural network model, and the generalization capability and analytical insights of deep learning approaches are still not very clear.

This paper proposes to *integrate* spatial interpolation with matrix completion for propagation map reconstruction. An interpolation assisted matrix completion framework is thus developed. Such an integration is highly non-trivial due to the following two challenges: (1) how to integrate the two methods such that the integrated approach works better than both of the methods applied alone; and (2) how to optimize the parameters for the integrated approach. Specifically, we discretize the area of interest into grid cells and form a sparse matrix using a windowing method, where only the grid cells that have sufficient measurements within a radius of b will be locally constructed via interpolation. Our preliminary results in [27] revealed good performance of this strategy under a handpicked parameter b. However, it was not clear how to optimize b. The key challenge is to optimize the window size b, which serves as a bridge between the interpolation and the matrix completion. To achieve such a goal, we first adopt a local polynomial regression model to estimate the selected entries of the matrix, and then, we derive analytical results to characterize the error of the interpolation. The window size b can therefore be optimized to minimize the MSE of the interpolation. Finally, we develop two new matrix completion approaches to construct the propagation map by leveraging the knowledge of uncertainty from the local interpolation step at the previous stage.

Our experiments found that by combining interpolation with matrix completion, the accuracy of propagation map reconstruction is substantially improved. Specifically, the reconstruction MSE can be reduced by 10%–50% from Kriging and other state-of-the-art schemes for a medium to large number of measurements, which translates to a saving of nearly half of the sensor measurements to achieve the same MSE. We demonstrate the application of the reconstructed propagation map to RSS-based source localization, where the root mean squared error (RMSE) of localization is reduced by more than 50% from a weighted centroid localization (WCL) [28] baseline.

To summarize, the following contributions are made:

- We develop a windowing-based integrated interpolation and matrix completion framework for propagation map reconstruction, where the window size balances the contribution between the interpolation and the matrix completion.
- Based on two local polynomial regression models, we analyze the moments and the asymptotic distribution of the interpolation error. In addition, we develop a minimum MSE approach for adjusting the window size to optimize the integrated interpolation and matrix completion framework.
- We develop two uncertainty-aware matrix completion methods to integrate the construction from the local interpolation. Our numerical results reveal that the proposed integrated approaches beat the conventional ones with substantial improvement in the accuracy of propagation map reconstruction.

The rest of the paper is organized as follows. Section II establishes the propagation reconstruction model. Section III develops a local polynomial regression method to construct a sparse matrix, analyzes the interpolation error and proposes the window size selection approach. Section IV numerically verifies the low rank property and proposes the uncertain-aware matrix completion scheme. Numerical results are presented in Section V and conclusion is given in Section VI.

Notation: Vectors are written as bold italic letters $\boldsymbol{x}$ and matrices as bold capital italic letters $\boldsymbol{X}$. For a matrix $\boldsymbol{X}$, X_{ij} denotes the entry in the ith row and jth column of $\boldsymbol{X}$. The notation $\boldsymbol{c}_{ij}$ denotes the center location of the (i,j)th grid, $o(x)$ means $\lim_{x\to 0} o(x)/x \to 0$, $O(x)$ means $|O(x)|/x \leq C$, for all $x > x_0$ with C and x_0 are positive real numbers, $\text{diag}(\boldsymbol{X})$ represents a column vector whose entries are the diagonal elements of matrix $\boldsymbol{X}$, and $\text{diag}(\boldsymbol{x})$ represents a diagonal matrix whose diagonal elements are the entries of vector $\boldsymbol{x}$. Symbol $\mathbb{E}\{\cdot\}$ and $\mathbb{V}\{\cdot\}$ denote expectation and variance separately.

II. System model

A. Reconstruction Model

Consider a propagation field that is excited by S sources located at $\boldsymbol{s}_k \in \mathcal{D}$, $k = 1, 2, \ldots, S$, in an area $\mathcal{D} \subset \mathbb{R}^2$. The signal emitted from the sources is detected by M sensors with known locations $\boldsymbol{z}_m \in \mathbb{R}^2$, $m = 1, 2, \ldots, M$, randomly deployed in $\mathcal{D}$. The propagation map to be reconstructed is modeled as

$$\rho(\boldsymbol{z}) \triangleq \sum_{k=1}^{S} g_k(d(\boldsymbol{s}_k, \boldsymbol{z})) + \zeta(\boldsymbol{z}) \qquad \boldsymbol{z} \in \mathcal{D} \tag{1}$$

where $d(\boldsymbol{s}, \boldsymbol{z}) = \|\boldsymbol{s} - \boldsymbol{z}\|_2$ describes the distance between a source at $\boldsymbol{s}$ and a sensor at $\boldsymbol{z}$, $g_k(d)$ describes the propagation function from the kth source in terms of the propagation distance d, and the term $\zeta(\boldsymbol{z})$ is a random component that captures the shadowing which is assumed to have spatial correlations.

The strength of the signal received by the mth sensor is given by

$$\gamma_m = \rho(\boldsymbol{z}_m) + \epsilon_m \tag{2}$$

where ϵ_m is a random variable with zero mean and variance σ^2 to model the measurement noise. The goal of this paper is to reconstruct $\rho(\boldsymbol{z})$ based on M RSS measurements $\{(\boldsymbol{z}_m, \gamma_m)\}$.

For a given spatial resolution $N \times N$ for the propagation map reconstruction, we consider to discretize the target area $\mathcal{D}$ into N rows and N columns that results in N^2 grid cells. Let $\boldsymbol{c}_{ij} \in \mathcal{D}$ be the center location of the (i,j)th grid cell, and $\boldsymbol{H}$ be a matrix representation of the propagation map $\rho(\boldsymbol{z})$, where the (i,j)th entry is defined as $H_{ij} = \rho(\boldsymbol{c}_{ij})$. As a result, the objective is to estimate a matrix $\bar{\boldsymbol{H}}$ based on the M RSS measurements, such that the squared error $||\bar{\boldsymbol{H}} - \boldsymbol{H}||_F^2 = \sum_{i,j}(\bar{H}_{ij} - \rho(\boldsymbol{c}_{ij}))^2$ is as low as possible.

We shall highlight that the parametric forms of the propagation models $g_k(d)$ are *unknown*, except that $g_k(d)$ are believed

to be smooth and decrease in distance d. Thus, model (1) tends to have a low rank structure. The intuition is that the path gain is usually dominated by the propagation distance $d(\boldsymbol{s}, \boldsymbol{z})$ between the source location $\boldsymbol{s}$ and the measurement location $\boldsymbol{z}$, and hence, there is a hidden homogeneity in all directions from $\boldsymbol{s}$. In addition, the statistics of the shadowing $\zeta(\boldsymbol{z})$ is also unknown. Therefore, classical parametric methods fail to apply here. On the other hand, conventional interpolation-based approaches fail to exploit the global structure of $g_k(d)$. For example, $g_k(d(\boldsymbol{s}_k, \boldsymbol{z}_1))$ may equal to $g_k(d(\boldsymbol{s}_k, \boldsymbol{z}_2))$ despite $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ being possibly far apart if the distances are equal, *i.e.*, $d(\boldsymbol{s}_k, \boldsymbol{z}_1) = d(\boldsymbol{s}_k, \boldsymbol{z}_2)$. Such a property implies that local interpolation approaches are strictly sub-optimal, and one should connect the entire set of measurements to improve propagation map reconstruction at a global scale.

B. Interpolation Assisted Matrix Completion

Denote $\boldsymbol{H}^{(k)} \in \mathbb{R}^{N\times N}$ as the propagation matrix associated with the kth source, with the (i,j)th entry given by $H^{(k)}_{i,j} = g_k(d(\boldsymbol{s}_k, \boldsymbol{c}_{ij}))$. Then, $\boldsymbol{H} = \sum_{k=1}^{S} \boldsymbol{H}^{(k)} + \boldsymbol{\zeta}$. As will be numerically evaluated in Section IV-A, $\sum_{k=1}^{S} \boldsymbol{H}^{(k)}$ is likely to be low rank, and so is $\sum_{k=1}^{S} \boldsymbol{H}^{(k)} + \boldsymbol{\zeta}$ due to the spatial correlation. As a result, propagation map reconstruction naturally leads to a sparse matrix completion problem.

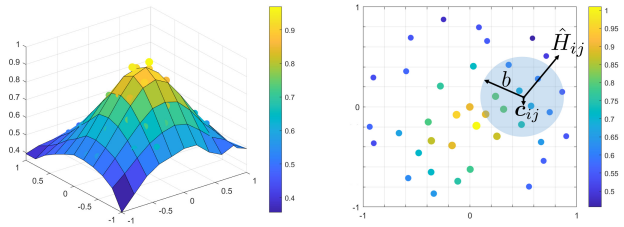
1) Constructing a Sparse Observation Matrix $\hat{\boldsymbol{H}}$: A straight-forward approach to form a sparse matrix is to assign the measurement γ_m to $\hat{H}_{ij}$ if the mth sensor $\boldsymbol{z}_m$ locates in the (i,j)th grid centered at $\boldsymbol{c}_{ij}$ and is the closest one to $\boldsymbol{c}_{ij}$. However, the number of measurements M could be substantially smaller than the number of grid cells N^2 under high resolution reconstruction, resulting in an overly sparse matrix that is difficult to complete. Moreover, there could be significant discretization error since the sensor location $\boldsymbol{z}_m$ may be away from the grid center $\boldsymbol{c}_{ij}$.

We propose to interpolate a subset of grid cells Ω to enrich the observations for matrix completion. There are two approaches. (1) Uniform sampling: we form the observation set Ω by sampling $CN\log^2(N)$ grid cells uniformly at random, where the parameter C can be empirically chosen to guarantee a sufficient number of observations for ensuring the identifiability of the matrix completion [29]. (2) Sensor-aware sampling: we form an observation set of grid cells where there are sufficient measurements nearby. Specifically, given a window size parameter b and a measurement number M_0, we define an observation set Ω as a subset of grids (i,j) such that there are at least M_0 sensors locating within a radius of b from the grid center $\boldsymbol{c}_{ij}$, *i.e.*,

$$\Omega = \left\{(i,j) : \sum_{m=1}^{M} \mathbb{I}\{\|\boldsymbol{z}_m - \boldsymbol{c}_{ij}\|_2 < b\} \geq M_0\right\} \qquad (3)$$

where $\mathbb{I}\{A\} = 1$ if condition A is satisfied, and $\mathbb{I}\{A\} = 0$ otherwise.

As a result, one only estimates $\hat{H}_{ij}$ for $(i,j) \in \Omega$ using the nearby measurements as illustrated in Fig. 1. The identifiability issue for matrix completion under the sensor-aware sampling will be discussed in Section IV-C.



(a) (b)

Figure 1. (a) Propagation map reconstruction based on local polynomial regression (b) Sensor measurements (colored dots) within a range of b (grey region) from the grid center $\boldsymbol{c}_{ij}$ are used to estimate $\hat{H}_{ij}$.

2) Formulating the Matrix Completion Problem: One representative approach is the alternating least square (ALS) algorithm [30], [31], which minimizes $||\boldsymbol{y} - \mathcal{A}(\boldsymbol{X})||_2$, subject to a product model $\boldsymbol{X} = \boldsymbol{L}\boldsymbol{R}$ for $\boldsymbol{L} \in \mathbb{R}^{N\times p}$ and $\boldsymbol{R} \in \mathbb{R}^{p\times N}$, where $\boldsymbol{y} \in \mathbb{R}^M$ is the measurement vector and $\mathcal{A} : \mathbb{R}^{N\times N} \to \mathbb{R}^M$ is a sensing operator that maps the (i,j)th element of the matrix $\boldsymbol{X}$ to the mth element of a vector to be compared with $\boldsymbol{y}$ [32]. Another possibility is to minimize the nuclear norm $\|\boldsymbol{X}\|_*$ of a matrix $\boldsymbol{X}$ subject to observation noise $|X_{ij} - \hat{H}_{ij}| \leq \epsilon$. It will be shown later that these existing approaches do not perform well, since they fail to exploit the fact that the observed entries may have different uncertainty according to the methods that form the matrix observations $\hat{H}_{ij}$.

III. Local Reconstruction via Polynomial Regression

It is crucial to determine the window size b in the proposed integrated approach. Intuitively, if the observation set Ω is constructed by sampling uniformly at random, the reconstruction performance mainly depends on the quality of the interpolation for the entries in Ω. The interpolation performance depends on the parameter b as will be shown later. If Ω is constructed by a sensor-aware approach, then for a small b, matrix completion dominates the performance, since many of entries are missing; and for a large b, interpolation dominates, since the local interpolation fully constructs the matrix.

In this section, we exploit local polynomial regression[1] to construct matrix observations $\hat{H}_{ij}$. Then, we derive analytical results to characterize the construction error in terms of the parameter b, and optimize for b.

A. Interpolation based on Local Polynomial Regression

Consider to employ a *parametric model* $\hat{\rho}(\boldsymbol{z}; \boldsymbol{c})$ to locally approximate the propagation map $\rho(\boldsymbol{z})$ in the neighborhood of $\boldsymbol{c}$. When the area of interest around $\boldsymbol{c}$ is small enough compared with the variation of the propagation map $\rho(\boldsymbol{z})$, one may employ a zeroth order model

$$\hat{\rho}(\boldsymbol{z}; \boldsymbol{c}) = \alpha(\boldsymbol{c}) \qquad (4)$$

[1]Note that, in this paper, although we derive the results based on a local polynomial regression model for the local interpolation, the same methodology can be extended to the integration with other interpolation methods, such as Kriging.

or a first order model

$$\hat{\rho}(\boldsymbol{z};\boldsymbol{c}) = \alpha(\boldsymbol{c}) + \boldsymbol{\beta}^{\mathrm{T}}(\boldsymbol{c})(\boldsymbol{z}-\boldsymbol{c}) \tag{5}$$

to approximate $\rho(\boldsymbol{z})$. The coefficients $\alpha(\boldsymbol{c})$ and $\boldsymbol{\beta}(\boldsymbol{c})$ depend on the location $\boldsymbol{c}$, and can be estimated using the measurements $\{(\boldsymbol{z}_m, \gamma_m)\}$ obtained near $\boldsymbol{c}$. The extension to a higher order model is straight-forward.

Specifically, let $\boldsymbol{\theta} = \{\alpha(\boldsymbol{c}), \boldsymbol{\beta}(\boldsymbol{c}), ...\}$ be the set of coefficients for the polynomial $\hat{\rho}(\boldsymbol{z};\boldsymbol{c},\boldsymbol{\theta})$. Then, $\boldsymbol{\theta}$ can be estimated using distance-weighted least-squares regression:

$$\underset{\boldsymbol{\theta}}{\text{minimize}} \sum_{m=1}^{M} \left(\gamma_m - \hat{\rho}(\boldsymbol{z}_m;\boldsymbol{c},\boldsymbol{\theta})\right)^2 K\left(\frac{\boldsymbol{z}_m - \boldsymbol{c}}{b}\right) \tag{6}$$

where b is the window parameter as introduced in (3) and $K(\boldsymbol{u})$ is a two dimensional kernel function that satisfies the following conditions:

(i) $K(\boldsymbol{u})$ is a non-negative, symmetric, and bounded probability density function (PDF) satisfying

$$\int K(\boldsymbol{u})d\boldsymbol{u} = \iint K(\boldsymbol{u})du_x du_y = 1 \tag{7}$$

$$\iint K(\boldsymbol{u})^{p_0} u_x^{p_1} u_y^{p_2} du_x du_y = 0 \tag{8}$$

for $p_0 \in \{1,2\}$ and $p_1,\ p_2 \in \{1,3\}$.

(ii) $K(\boldsymbol{u})$ has a compact support

$$\mathcal{C} = \left\{\boldsymbol{u} \in \mathbb{R}^2 : ||\boldsymbol{u}||_2 < 1\right\} \tag{9}$$

where $K(\boldsymbol{u}) = 0$ for $\boldsymbol{u} \notin \mathcal{C}$.

Thus, the regression problem in (6) only considers the measurements that in a radius b from $\boldsymbol{c}$ and assigns a high weight if the measurement location $\boldsymbol{z}_m$ is close to $\boldsymbol{c}$, and a low weight if $\boldsymbol{z}_m$ is far away from $\boldsymbol{c}$. For instance, the Epanechnikov kernel [33] $K(\boldsymbol{u}) = \max\{0, \frac{3}{4}(1-||\boldsymbol{u}||^2)\}$ and the truncated Gaussian kernel $K(\boldsymbol{u}) = \frac{50}{17\sqrt{2\pi}}\exp(-||\boldsymbol{u}||^2/2)\mathbb{I}\{||\boldsymbol{u}||^2 \le 1\}$ satisfy the above conditions.

As a result, for a given location $\boldsymbol{c}$, the local parameter $\hat{\alpha}(\boldsymbol{c})$ under a zeroth order model is found as the solution to

$$\underset{\alpha}{\text{minimize}} \sum_{m=1}^{M} \left(\gamma_m - \alpha(\boldsymbol{c})\right)^2 K\left(\frac{\boldsymbol{z}_m - \boldsymbol{c}}{b}\right) \tag{10}$$

whereas, the local parameter $(\hat{\alpha}(\boldsymbol{c}), \hat{\boldsymbol{\beta}}(\boldsymbol{c}))$ under a first order model is found as the solution to

$$\underset{\alpha,\boldsymbol{\beta}}{\text{minimize}} \sum_{m=1}^{M} \left(\gamma_m - \alpha(\boldsymbol{c}) - \boldsymbol{\beta}^{\mathrm{T}}(\boldsymbol{c})(\boldsymbol{z}_m - \boldsymbol{c})\right)^2 K\left(\frac{\boldsymbol{z}_m - \boldsymbol{c}}{b}\right). \tag{11}$$

There exists closed-form solutions.

Proposition 1 (Interpolation)**.** *Given a location $\boldsymbol{c}$, the solution to (10) under the zeroth order model is given by*

$$\hat{\alpha}(\boldsymbol{c}) = \left(\sum_{m=1}^{M} w_m(\boldsymbol{c})\gamma_m\right) \Big/ \sum_{m=1}^{M} w_m(\boldsymbol{c}) \tag{12}$$

where $w_m(\boldsymbol{c}) = K((\boldsymbol{z}_m - \boldsymbol{c})/b)$ denotes the kernel weight for the mth measurement.

The solution to (11) under the first order model is given by

$$\begin{bmatrix} \hat{\alpha}(\boldsymbol{c}) \\ \hat{\boldsymbol{\beta}}(\boldsymbol{c}) \end{bmatrix} = (\tilde{\boldsymbol{D}}\boldsymbol{W}\tilde{\boldsymbol{D}}^{\mathrm{T}})^{-1}\tilde{\boldsymbol{D}}\boldsymbol{W}\boldsymbol{\gamma} \tag{13}$$

where $\boldsymbol{\gamma} \triangleq (\gamma_1, \gamma_2, \cdots, \gamma_M)^{\mathrm{T}}$ *is a vector form of the measurements, $\boldsymbol{W} = diag\{(w_1(\boldsymbol{c}), w_2(\boldsymbol{c}), \cdots, w_M(\boldsymbol{c}))\}$, and*

$$\tilde{\boldsymbol{D}} = \begin{bmatrix} \mathbf{1}^T \\ \boldsymbol{D} \end{bmatrix}$$

in which $\mathbf{1}$ is a vector of all 1's and $\boldsymbol{D}$ is a $2 \times M$ matrix with the mth column given by $\boldsymbol{z}_m - \boldsymbol{c}$.

Proof. Note that the problems (10) and (11) are unconstrained quadratic optimization problems and can be easily verified to be convex. Thus, the solution can be obtained by rearranging the objective into a matrix-vector form and setting the derivative to zero [33], [34]. □

As a result, to construct the (i,j)th observation $\hat{H}_{ij}$ at the grid centered at $\boldsymbol{c}_{ij}$, one can first construct a local model $\hat{\rho}(\boldsymbol{z};\boldsymbol{c}_{ij})$ at location $\boldsymbol{c}_{ij}$. Then, $\hat{H}_{ij}$ can be estimated from $\hat{\rho}(\boldsymbol{z} = \boldsymbol{c}_{ij};\boldsymbol{c}_{ij}) = \alpha(\boldsymbol{c}_{ij})$ for both zeroth order and first order models. From the solution (10) for the zeroth order model, $\hat{H}_{ij} = \left(\sum_m w_m(\boldsymbol{c}_{ij})\gamma_m\right) / \sum_m w_m(\boldsymbol{c}_{ij})$ is simply the distance-weighted strength γ_m from the measurements around the grid center $\boldsymbol{c}_{ij}$. For the first order model, $\hat{H}_{ij} = \hat{\alpha}(\boldsymbol{c}_{ij})$ as given in (13) with setting $\boldsymbol{c} = \boldsymbol{c}_{ij}$.

B. The Error of Interpolation

An unsolved yet important issue is the choice of the window parameter b in the estimation problem (6). Intuitively, the zeroth order model is easy to estimate as it has just one parameter, but it has poor inference capability away from $\boldsymbol{c}$, and as a result, it may only work well in a small local area under a small window size b where the propagation map $\rho(\boldsymbol{z})$ varies slowly. On the other hand, a higher order model may have less bias to infer $\rho(\boldsymbol{z})$ away from $\boldsymbol{c}$, but it requires a lot more measurement data as it has more parameters, and thus, it prefers a large b to include more measurements.

Recall the interpolation error $\xi_{ij} \triangleq \hat{H}_{ij} - \rho(\boldsymbol{c}_{ij})$, and $\hat{H}_{ij} = \hat{\alpha}(\boldsymbol{c}_{ij})$ which can be computed using (12) or (13) given the center location $\boldsymbol{c}_{ij}$ of the (i,j)th grid. In the following, unless specifically pointed out, we take all the expectation and variance over the random measurement noise ϵ_m given the sensor locations $\boldsymbol{z}_m$. Assume that $\rho(\boldsymbol{z})$ is second order differentiable. For a fixed sensor deployment $\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_M$, the bias $\mathbb{E}\{\xi_{ij}\}$ and variance $\mathbb{V}\{\xi_{ij}\}$ of the interpolation error for $\hat{H}_{ij}$ can be derived and summarized in the following theorem.

Theorem 1 (Error under zeroth order model)**.** *For the zeroth order model,*

$$\mathbb{E}\{\xi_{ij}\} = \sum_{m=1}^{M} \nabla\rho(\boldsymbol{c}_{ij})^T (\boldsymbol{z}_m - \boldsymbol{c}_{ij})\, \bar{w}_m(\boldsymbol{c}_{ij}) + o(b) \tag{14}$$

$$\mathbb{V}\{\xi_{ij}\} = \sum_{m=1}^{M} \bar{w}_m^2(\boldsymbol{c}_{ij})\sigma^2 \tag{15}$$

where $\bar{w}_m(\boldsymbol{c}_{ij}) \triangleq w_m(\boldsymbol{c}_{ij})/\sum_{i=1}^{M} w_i(\boldsymbol{c}_{ij})$ is the normalized weight for the mth measurement, and $\nabla\rho(\boldsymbol{c}_{ij})$ is the derivative of the propagation map $\rho(\boldsymbol{z})$ at location $\boldsymbol{z} = \boldsymbol{c}_{ij}$. The term $o(b)$ represents a residual and it satisfies $o(b)/b \to 0$ as $b \to 0$.

Proof. See Appendix A. □

For the result under the first order model, denote $\boldsymbol{D}_{ij}$ as a $2 \times M$ matrix to capture the direction from the grid center $\boldsymbol{c}_{ij}$ to the sensor location $\boldsymbol{z}_m$, and the mth column of $\boldsymbol{D}_{ij}$ is defined as $\boldsymbol{z}_m - \boldsymbol{c}_{ij}$. In addition, define a direction matrix with offsets as $\tilde{\boldsymbol{D}}_{ij} = \begin{bmatrix} \mathbf{1} & \boldsymbol{D}_{ij}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$.

Theorem 2 (Error under first order model)**.** *For the first order model,*

$$\mathbb{E}\{\xi_{ij}\} = \frac{1}{2}\left[\tilde{\boldsymbol{W}}_{ij}^{-1}\tilde{\boldsymbol{D}}_{ij}\boldsymbol{W}_{ij}diag\{\boldsymbol{D}_{ij}^{T}\boldsymbol{\Psi}_{ij}\boldsymbol{D}_{ij}\}\right]_{(1,1)} + o(b^2) \tag{16}$$

$$\mathbb{V}\{\xi_{ij}\} = \sigma^2\left[\tilde{\boldsymbol{W}}_{ij}^{-1}\left(\tilde{\boldsymbol{D}}_{ij}\boldsymbol{W}_{ij}\boldsymbol{W}_{ij}^{T}\tilde{\boldsymbol{D}}_{ij}^{T}\right)\tilde{\boldsymbol{W}}_{ij}^{-1}\right]_{(1,1)} \tag{17}$$

where $\tilde{\boldsymbol{W}}_{ij} = \tilde{\boldsymbol{D}}_{ij}\boldsymbol{W}_{ij}\tilde{\boldsymbol{D}}_{ij}^{T}$, the operation $[\boldsymbol{A}]_{(1,1)}$ returns the $(1,1)$th entry of a matrix $\boldsymbol{A}$, and $\boldsymbol{\Psi}_{ij} = \nabla^2\rho(\boldsymbol{c}_{ij})$ is the Hessian matrix of $\rho(\boldsymbol{z})$ at point $\boldsymbol{c}_{ij}$. The term $o(b^2)$ represents a residual and it satisfies $o(b^2)/b^2 \to 0$ as $b \to 0$.

Proof. See Appendix B. □

As seen from Theorems 1 and 2, the bias $\mathbb{E}\{\xi_{ij}\}$ depends on the variation of the propagation map $\rho(\boldsymbol{z})$. Specifically, under the zeroth order model, the slope $\nabla\rho(\boldsymbol{c}_{ij})$ of the propagation map contributes to the bias. It follows that, for $\boldsymbol{c}_{ij}$ far away from any source, the bias $\mathbb{E}\{\xi_{ij}\}$ tends to be small because both $\rho(\boldsymbol{c}_{ij})$ and $\nabla\rho(\boldsymbol{c}_{ij})$ are small for typical propagation maps. On the other hand, when $\boldsymbol{c}_{ij}$ locates in the area where the slope $\nabla\rho(\boldsymbol{c}_{ij})$ is large, the bias $\mathbb{E}\{\xi_{ij}\}$ will be large, and thus, the zeroth order interpolation may perform poorly. Under the first order model, the bias $\mathbb{E}\{\xi_{ij}\}$ is not affected by the slope but by the curvature of the propagation map, *i.e.*, $\nabla^2\rho(\boldsymbol{c}_{ij})$.

A bias-variance trade-off due to the window size parameter b can be observed. First, it is not surprising to find that the absolute bias $|\mathbb{E}\{\xi_{ij}\}|$ increases for large window size b, if one numerically evaluates the expressions in (14) and (16). This is because the estimator $\hat{H}_{ij}$ tends to fit to the structure of a larger area, but the local parametric model $\hat{\rho}(\boldsymbol{z};\boldsymbol{c})$ is accurate only for a small area around $\boldsymbol{c}$. Second, by contrast, the variance decreases as b increases, which can be verified by numerically evaluating the variance expressions in (15) and (17). The reason is that a large window size b can include more sensor measurements, and thus, the measurement noise in (2) can be suppressed.

C. Asymptotic Analysis of the Interpolation Error

To explicitly analyze the bias and variance trade off, we derive the asymptotic distribution of the interpolation error ξ_{ij} under the regime of large number of sensors.

Define $f(\boldsymbol{z})$ as the density function of the sensors deployed at location $\boldsymbol{z}$. Assume that $f(\boldsymbol{z})$ is second order differentiable and the sensors are independent and identically distributed (i.i.d.) according to $f(\boldsymbol{z})$. Moreover, assume that the propagation function $\rho(\boldsymbol{z})$ is third order differentiable. We have the following results on the asymptotic distribution of the interpolation error ξ_{ij}.

Theorem 3 (Asymptotic interpolation error I)**.** *For a small enough b, the interpolation error ξ_{ij} under the zeroth order interpolation (12) converges as*

$$\xi_{ij} \xrightarrow{p} b^2 C_0\Big(\frac{\vartheta_1(\boldsymbol{c}_{ij})}{f(\boldsymbol{c}_{ij})} + \frac{1}{2}\vartheta_2(\boldsymbol{c}_{ij})\Big) + o(b^2) \tag{18}$$

as $M \to \infty$, where $\xrightarrow{p}$ means convergence in probability, and the centered error $\bar{\xi}_{ij} = \xi_{ij} - \mathbb{E}\{\xi_{ij}\}$ converges as

$$\sqrt{Mb^2}\bar{\xi}_{ij} \xrightarrow{d} \mathcal{N}\Big(0, \frac{C_1\sigma^2}{f(\boldsymbol{c}_{ij})} + o(b)\Big) \tag{19}$$

as $M \to \infty$, where $\xrightarrow{d}$ means convergence in distribution,

$$\vartheta_1(\boldsymbol{c}_{ij}) = \frac{\partial f(\boldsymbol{c}_{ij})}{\partial u_x}\frac{\partial \rho(\boldsymbol{c}_{ij})}{\partial u_x} + \frac{\partial f(\boldsymbol{c}_{ij})}{\partial u_y}\frac{\partial \rho(\boldsymbol{c}_{ij})}{\partial u_y}$$

$$\vartheta_2(\boldsymbol{c}_{ij}) = \frac{\partial^2 \rho(\boldsymbol{c}_{ij})}{\partial u_x^2} + \frac{\partial^2 \rho(\boldsymbol{c}_{ij})}{\partial u_y^2}$$

$C_0 = \int u_x^2 K(\boldsymbol{u})du_x$, and $C_1 = \int K(\boldsymbol{u})^2 d\boldsymbol{u}$.

Proof. See Appendix C. □

Following a similar technique as illustrated in Appendix C, we obtain the following result on the interpolation error ξ_{ij} under the first order interpolation (13).

Theorem 4 (Asymptotic interpolation error II)**.** *For a small enough b, the interpolation error ξ_{ij} under the first order interpolation (13) converges as*

$$\xi_{ij} \xrightarrow{p} \frac{1}{2}b^2 C_0 \vartheta_2(\boldsymbol{c}_{ij}) + o(b^2)$$

as $M \to \infty$, and the centered error $\bar{\xi}_{ij} = \xi_{ij} - \mathbb{E}\{\xi_{ij}\}$ converges as

$$\sqrt{Mb^2}\bar{\xi}_{ij} \xrightarrow{d} \mathcal{N}\Big(0, \frac{C_1\sigma^2}{f(\boldsymbol{c}_{ij})} + o(b)\Big)$$

as $M \to \infty$.

Proof. Omitted. The result is obtained following the same derivation as [35, Theorem 2.1] in a straight-forward way. □

As a result from Theorems 3 and 4, if one chooses the window size b according to M as $b = M^{-p}$, then the bias asymptotically converges to 0 at a rate $C_2 M^{-2p}$, where the coefficient C_2 depends on the choice of interpolation methods (e.g., (12) and (13)), and the variance converges to 0 at a rate $\frac{C_1\sigma^2}{f(\boldsymbol{c}_{ij})}M^{-(1-2p)}$, which is identical for both interpolation methods. In particular, under $p = 1/6$, the MSE, which equals to the squared-bias plus the variance, converges to 0 at a fast rate, where both the squared-bias and the variance terms scale as $O(M^{-2/3})$.

In addition, we discover some useful analytical insights from Theorems 3 and 4 as follows.

Advantage provided by the first order method: The advantage of the first order method over the zeroth order counterpart

is observed from the bias of the estimator. Specifically, the first term in the bias coefficient $\vartheta_1(\boldsymbol{c}_{ij})$ under the zeroth order method disappears in the bias under the first order interpolation method. As a result, the first order method is less affected by the sensor distribution $f(\boldsymbol{z})$. For example, under non-uniform sensor distribution, the term in $\vartheta_1(\boldsymbol{c}_{ij})$ yields a non-zero bias, but such a bias does not appear in Theorem 4 under the first order method.

Impact from the sensor distribution and the propagation map: The bias of the estimator depends on the sensor distribution $f(\boldsymbol{c}_{ij})$ under the zeroth order method but not for the first order method. The bias of both zeroth and first order method tend to be larger in the area where the propagation function is with larger curvature $\partial^2\rho(\boldsymbol{c}_{ij})/\partial u_x^2$, $\partial^2\rho(\boldsymbol{c}_{ij})/\partial u_y^2$. In addition, the bias under the zeroth order method also depends on the gradient of the sensor density $f(\boldsymbol{z})$. The analytical result matches with the intuition that more sensors are generally required in the region where the propagation map has larger variation.

Intuition on the variance: The variance of the interpolation error is not sensitive to model complexity as it remains the same for both zeroth order method and first order method. The variance is inversely proportional to $Mb^2f(\boldsymbol{c}_{ij})$ which represents the number of sensors within the window centered at point $\boldsymbol{c}_{ij}$ with radius b. This implies that for the area with dense measurements, the interpolation error of that area tends to have small variance.

The bias and variance trade off: It is observed from the results of Theorem 3 and Theorem 4 that the bias is proportional to the square of the window size b but the variance is inversely proportional to the square of b. This implies that there might exist an optimal window size b, thus the MSE of the estimator $\hat{\rho}$ attains smallest.

D. Window Size Optimization

Here, we explore two methods to optimize b.

1) Leave One Out cross validation: A natural way to choose b is to employ the Leave One Out cross validation [36]. The method minimizes the MSE of the interpolation by solving

$$\underset{b}{\text{minimize}} \quad \frac{1}{M}\sum_{m=1}^{M}\left(\gamma_m - \hat{\rho}_{-m}(\boldsymbol{z}_m)\right)^2 \tag{20}$$

where $\hat{\rho}_{-m}(\boldsymbol{z}_m)$ is obtained in the same way as $\hat{\rho}(\boldsymbol{z}_m)$ in (6) based on the $M-1$ samples from $\{(\boldsymbol{z}_j,\gamma_j)\}_{j=1}^{M}$ that exclude the mth one.

However, Leave One Out cross validation is computationally expensive when the data set is big with large M as it requires fitting the model $\hat{\rho}_{-m}$ M times. Moreover, the estimated minimizer $\hat{b}$ obtained from (20) is mismatched with the true minimizer on window size as shown in Fig. 2. The possible reason is that, the mth sensor plays the most significant role to predict γ_m at location $\boldsymbol{z}_m$, but the mth sensor is excluded from estimating $\hat{\rho}_{-m}(\boldsymbol{z}_m)$ in the Leave One Out cross validation.

2) Analytical minimum MSE method: We propose to optimize the window size b by minimizing the MSE of the interpolation using the analytical results from Theorems 1 and 2. The problem is formulated as follows:

$$\underset{b}{\text{minimize}} \quad \frac{1}{|\Omega|}\sum_{(i,j)\in\Omega}\mu_{ij}^2+\nu_{ij}^2 \tag{21}$$

where Ω is defined in (3) which depends on b, and μ_{ij} and ν_{ij} are respectively the bias and variance derived from Theorems 1 and 2. Specifically, μ_{ij} and ν_{ij} are given as follows:

- Under the zeroth order interpolation (12)

$$\mu_{ij} = \sum_{m=1}^{M}\hat{\boldsymbol{\beta}}^{\mathrm{T}}(\boldsymbol{c}_{ij})(\boldsymbol{z}_m-\boldsymbol{c}_{ij})\bar{w}_m \tag{22}$$

$$\nu_{ij}^2 = \sum_{m=1}^{M}\bar{w}_m^2(\boldsymbol{c}_{ij})\sigma^2 \tag{23}$$

  where $\hat{\boldsymbol{\beta}}(\boldsymbol{c}_{ij})$ is the gradient estimated from (11).
- Under the first order interpolation (13)

$$\mu_{ij} = \frac{1}{2}\left[\tilde{\boldsymbol{W}}_{ij}^{-1}\tilde{\boldsymbol{D}}_{ij}\boldsymbol{W}_{ij}\mathrm{diag}\{\boldsymbol{D}_{ij}^{\mathrm{T}}\hat{\boldsymbol{\Psi}}_{ij}\boldsymbol{D}_{ij}\}\right]_{(1,1)} \tag{24}$$

$$\nu_{ij}^2 = \sigma^2[\tilde{\boldsymbol{W}}_{ij}^{-1}\left(\tilde{\boldsymbol{D}}_{ij}\boldsymbol{W}_{ij}\boldsymbol{W}_{ij}^{\mathrm{T}}\tilde{\boldsymbol{D}}_{ij}^{\mathrm{T}}\right)\tilde{\boldsymbol{W}}_{ij}^{-1}]_{(1,1)} \tag{25}$$

  where $\hat{\boldsymbol{\Psi}}$ is the estimated Hessian obtained from solving problem (6) under a second order model.

In addition, the parameter M_0 in (3) can be specified as $M_0 = 4$ for the zeroth order interpolation method, since there are 3 parameters to estimate including $\hat{\alpha}$ and the gradient $\hat{\boldsymbol{\beta}}(\boldsymbol{c}_{ij})$ for evaluating (22) and (23), and we need one additional measurement since the support of the kernel in (9) has a radius of b and the measurement on the edge of the window will receive a zero weight. Similarly, we may configure $M_0 = 7$ for the first order interpolation, since there are 6 parameters to estimate for (24) and (25).

Fig. 2 shows a numerical example on the MSE of propagation map reconstruction under various choices of window size b. In the experiment, we use the same propagation model as in Section V. The dimension is $N = 30$ and the number of sensors is $M = 400$. The MSE of the propagation map reconstruction is calculated through $1/N^2||\bar{\boldsymbol{H}}-\boldsymbol{H}||_F^2$, where $||\cdot||_F$ represents Frobenius norm, and $\bar{\boldsymbol{H}}$ is the reconstructed propagation map based on solving the matrix completion

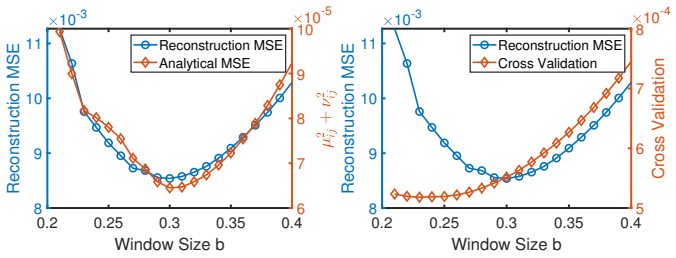


Figure 2. (a) The MSE of the propagation map reconstruction and the analytical MSE of the interpolation based on (21); the minimizers $\hat{b}$ are matched; (b) The MSE of the propagation map reconstruction and the cost from cross validation based on (20); the minimizers $\hat{b}$ are mismatched.

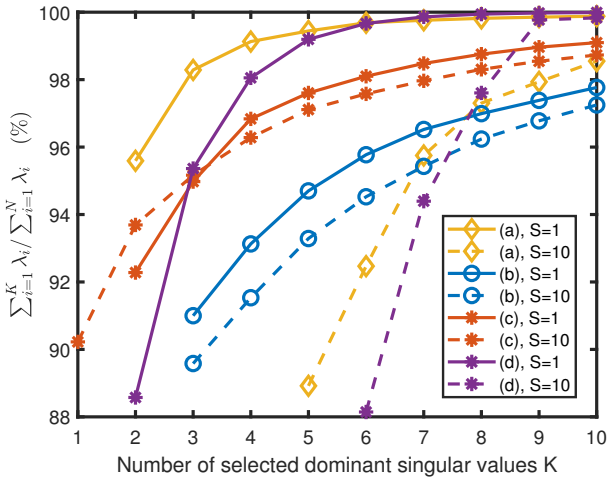


Figure 3. Percentage of the sum of the first K dominant singular values over the sum of all singular values for a 100 by 100 propagation matrix.

problem (26) in which the $\hat{H}_{ij}$ is constructed from the local interpolation and $\mathcal{I}_{ij}$ is the uncertainty interval based on the local interpolation error. More details to be discussed in Section IV. The μ_{ij} and ν_{ij} are calculated according to (24) and (25). Fig. 2 (a) shows the reconstruction MSE of matrix completion and the analytical MSE derived in (21) in terms of b. It is observed that the minimizers $\hat{b}$ of both curves are close. Fig. 2 (b) shows the reconstruction MSE of matrix completion and the MSE in (20) from cross validation. It is observed that the minimizers are significantly mismatched, confirming that the proposed analytical minimum MSE method works better in optimizing the window size b. A slight adjustment of window size b in a ± 0.1 range may lead to 20% performance difference. It is thus critical to optimize b.

IV. Global Reconstruction via Matrix Completion

In this section, we study matrix completion methods for propagation map reconstruction that exploits the global structure of the propagation map. We first numerically verify the low rank property of a propagation map, and then, we develop uncertainty-aware matrix completion methods that exploit the uncertainty from the local reconstruction.

A. The Low Rank Property

It is believed that the matrix representation of the propagation map as constructed in Section II-B is likely low rank. The intuition is as follows: First, a propagation map tends to have a unimodal structure where the longer the distance away from the source, the smaller the RSS; second, there is usually significant spatial correlation and hence $\rho(\boldsymbol{z})$ is smooth, leading to the low rankness.

One special example to understand the low rank property is the exponential 2-D propagation map given by $g(d(\boldsymbol{s},\boldsymbol{z})) = e^{-d^2/\sigma}$ over $\boldsymbol{z} = (x, y)$, which can be decomposed as the product of two 1-D functions $g(d(\boldsymbol{s},\boldsymbol{z})) = u(x)v(y)$. This is due to the fact that $d(\boldsymbol{s},\boldsymbol{z})^2 = (x - s_{1,1})^2 + (y - s_{1,2})^2$ and $e^{-(x^2+y^2)/\sigma} = e^{-x^2/\sigma}e^{-y^2/\sigma} = u(x)v(y)$. As a result, the matrix representation of the propagation map can be written as the outer product of two vectors $\boldsymbol{H} = \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$, indicating that $\boldsymbol{H} \in \mathbb{R}^{N\times N}$ is always rank-1 regardless of the matrix dimension N.

To numerically study the rank property of more propagation maps, we select four common forms of propagation models and pick the parameters that *maximize* the rank of the propagation matrix $\boldsymbol{H}$ using a brute force search. The resulting models are listed as follows:

(a) $g(d) = \alpha d^{-\beta},\ \beta = 2.2,\ h = 0.09$
(b) $g(d) = \alpha \exp(-d^{\beta}),\ \beta = 1.8,\ h = 0.05$
(c) $g(d) = \alpha - \beta \times \log_{10} d,\ \beta = 1.8,\ h = 0.01$
(d) $g(d) = \alpha d^{-\gamma}\beta^{-d},\ \beta = 2.8,\ \gamma = 1.5,\ h = 0.01$

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, h is the elevation, and the parameters α are selected to normalize the total energy in the area. Note that models (a), (b) and (c) correspond to energy in linear scale, exponential scale, and log-scale, respectively, in radio propagation; in addition, model (d) corresponds to a model for underwater acoustic propagation [37]. The shadowing component is with the same setting as in Section V.

Fig. 3 shows the percentage of the sum of the first K dominant singular values over the sum of all singular values λ_i for $N = 100$. It clearly demonstrates the low rank property, where 5 dominant eigen-mode combined already reconstruct over 95% energy of the propagation map for the single source scenario, $S = 1$. For the scenario where $S = 10$ sources are randomly placed in the area and each source emits signal with the power randomly generated from an exponential distribution with the rate parameter $\kappa = 1$, it is observed that there is still a low rank property for the aggregated propagation map, and the rank increases much slower than the increasing of the number of sources.

B. Uncertainty-aware Matrix Completion

The remaining challenge is to design new matrix completion methods that exploit the knowledge that the observed entries may have different uncertainty according to the local interpolation. We explore two uncertainty-aware matrix completion formulations to demonstrate the core idea.

1) Nuclear norm minimization with trust region (NNM-t): We introduce *trust regions* in the following nuclear norm minimization problem to exploit the local uncertainty in matrix completion

$$\begin{aligned} \underset{\boldsymbol{X}\in\mathbb{R}^{N\times N}}{\text{minimize}} \quad & \|\boldsymbol{X}\|_* \\ \text{subject to} \quad & X_{ij} - \hat{H}_{ij} \in \mathcal{I}_{ij}, \qquad \forall (i,j) \in \Omega \end{aligned} \tag{26}$$

where $\mathcal{I}_{ij}$ is the interval to specify the trust region of the estimate $\hat{H}_{ij}$ obtained from the local reconstruction in Section III, and Ω is defined in (3).

The trust region $\mathcal{I}_{ij}$ for the estimation $\hat{H}_{ij}$ in (26) can be constructed using Theorems 1–4. Specifically, Theorems 3 and 4 suggest that the interpolation error $\xi_{ij} = \hat{H}_{ij} - \rho(\boldsymbol{c}_{ij})$ asymptotically follows a Gaussian distribution, where the mean μ_{ij}

Algorithm 1 Propagation map reconstruction under NNM-t

Initialize M, N, $\boldsymbol{z}_m$, γ_m, δ, $b_{\min}$, $b_{\max}$ and interpolation order r.

1) Distance-weighted least-squares regression: Solving (6) with $(r+1)$th order model.
2) Bias and variance approximation: μ_{ij} and ν_{ij}.
   a) If $r=0$, approximate (14), (15) with (22), (23).
   b) If $r=1$, approximate (16), (17) with (24), (25).
3) Optimize b: Find the minimizer $\hat{b}$ of (21) subject to $b_{\min} \le b \le b_{\max}$.
4) Estimate $\hat{H}_{ij}$. Solving (6) to get $\hat{\alpha}(\boldsymbol{c}_{ij})$ and $\hat{H}_{ij} = \hat{\alpha}(\boldsymbol{c}_{ij})$.
   a) If $r=0$,
   $$\hat{H}_{ij} = \Big(\sum_{m=1}^{M} w_m(\boldsymbol{c}_{ij})\gamma_m\Big) \Big/ \sum_{m=1}^{M} w_m(\boldsymbol{c}_{ij}).$$
   b) If $r=1$,
   $$\hat{H}_{ij} = \Big[(\tilde{\boldsymbol{D}}_{ij}\boldsymbol{W}_{ij}\tilde{\boldsymbol{D}}_{ij}^{\mathrm{T}})^{-1}\tilde{\boldsymbol{D}}_{ij}\boldsymbol{W}_{ij}\boldsymbol{\gamma}\Big]_{(1,1)}.$$
5) Repeat step 1.
6) Confidence interval $\mathcal{I}_{ij}$ construction: $\mathcal{I}_{ij} = \big(\mu_{ij} \pm \Phi^{-1}(1-\frac{\delta}{2})\nu_{ij}^{\frac{1}{2}}\big)$.
7) Nuclear norm matrix completion. Get the recovered matrix $\bar{\boldsymbol{H}}$ through solving (26).

and variance ν_{ij}^2 can be practically computed via (22)–(25). Then, using Gaussian approximation, the trust region $\mathcal{I}_{ij}$ with a $(1-\delta)$-confident level, *i.e.*, $\mathbb{P}\{\hat{H}_{ij} - \rho(\boldsymbol{c}_{ij}) \in \mathcal{I}_{ij}\} = 1-\delta$, can be approximately constructed as

$$\mathcal{I}_{ij} = \Big(\mu_{ij} \pm \Phi^{-1}\Big(1-\frac{\delta}{2}\Big)\nu_{ij}\Big) \tag{27}$$

where $\Phi(x)$ is the cumulative distribution function of the standard Gaussian distribution, and the notation $(a \pm b)$ represents an interval $(a-b, a+b)$. The overall steps can be referred to Algorithm 1.

2) Weighted ALS (wALS): We design a wALS formulation to exploit the uncertainty of the observation based on the conventional ALS algorithm developed in [31]. We propose to assign weights, which are inversely proportional to the standard variance of the interpolation error, to the matrix observation. The weights $\boldsymbol{w}$ are chosen with the intuition that the estimate with smaller variance is believed to have higher accuracy and thus plays a more significant role in the matrix completion problem.

Let $M' = |\Omega|$ be the number of observed entries via local reconstruction from interpolation. For the nth element in Ω that corresponds to the (i,j)th entry of the matrix, construct an observation $y_n = \hat{H}_{ij} - \mu_{ij}$, and construct a weight $w_n = \nu_{ij}^{-1}$. Let $\mathcal{A}: \mathbb{R}^{N\times N} \to \mathbb{R}^{M'}$ be a sensing operator that maps the (i,j)th element of a matrix $\boldsymbol{X}$ to the corresponding nth element of the vector $\boldsymbol{y}$. The wALS problem is formulated as follows

$$\begin{aligned} \underset{\boldsymbol{X}\in\mathbb{R}^{N\times N}}{\text{minimize}} \quad & ||\boldsymbol{w} \circ (\boldsymbol{y} - \mathcal{A}(\boldsymbol{X}))||_2 \\ \text{subject to} \quad & \boldsymbol{X} = \boldsymbol{L}\boldsymbol{R} \\ & \boldsymbol{L} \in \mathbb{R}^{N\times p} \\ & \boldsymbol{R} \in \mathbb{R}^{p\times N} \end{aligned} \tag{28}$$

where '$\circ$' means element-wise product.

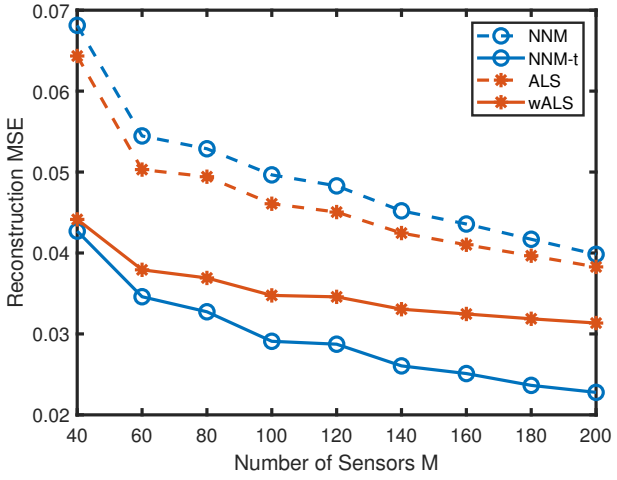


Figure 4. Reconstruction MSE for $S=1$ and $N=30$. The uncertainty-aware scheme works.

To numerically evaluate the performance of the proposed uncertainty-aware matrix completion algorithms with their conventional counterparts, consider the same experiment configuration in Section V with number of sources $S=1$, number of measurements $M = 40-200$ for reconstructing a propagation map with resolution $N=30$. For a conventional NNM scheme, it solves a similar NNM problem except that it replaces the constraint in (26) with $|X_{ij} - \hat{H}_{ij}| \le \epsilon$, where $\epsilon = 2/M$ is the best parameter we found via cross validation. For a conventional ALS, the weighting vector $\boldsymbol{w}$ in (28) is $\mathbf{1}$. For the proposed methods, the confident level parameter $\delta = 0.05$ is chosen for the NNM-t method, and the rank parameter $p=1$ is chosen for both wALS and the conventional ALS. Fig. 4 compares the MSE in propagation map reconstruction under different matrix completion algorithms. The results demonstrate that both of the proposed NNM-t and wALS methods outperform their conventional counterparts, which only exploit the local reconstruction $\hat{H}_{ij}$ from Section III, but not the bias and variance derived from our analysis. The results confirm the advantage of explicitly exploiting the observation uncertainty for enhancing the reconstruction performance.

C. Identifiability of the Matrix Completion

Recall that under the sensor-aware approach, the observation set Ω is formed according to the deterministic sensor deployment $\{\boldsymbol{z}_m\}$ and the windows size parameter b. Thus, it is important to understand whether the matrix is completable based on Ω. Although establishing a rigorous approach to determining the best observation pattern of $\hat{\boldsymbol{H}}$ with *both* the interpolation performance and the identifiability conditions

[38]–[40] taken into account still remains an open question that goes beyond the scope of this paper, we have obtained the following insights with a simple strategy to guarantee the identifiability for matrix completion with high probability.

First, it is found that, for uniformly random sensor deployment, the proposed scheme tends to naturally meet the identifiability conditions required by the low rank matrix completion [38]–[41]. For sensor-aware deployment, our observations are as follows. For a moderate to large M, recall that a subset of entries $\hat{H}_{ij}$ with M_0 measurements within a window size b are constructed via interpolation, where M_0 was chosen based on the identifiability of the local polynomial regression model, and b was optimized such that the interpolated error for $\hat{H}_{ij}$ is minimized. This step naturally forms a large number of observations $\hat{H}_{ij}$, and such a phenomenon is consistent with the identifiability condition of low rank matrix completion as observed from our experiments. For a small M, the window size optimization tends to select a large window size b to meet the identifiability condition of the local polynomial regression model (see Section III-D), and meanwhile, a large b is also needed for reducing the local interpolation error. As a result, a large number of observed entries are constructed from the interpolation under a large window size b; this, again, is consistent with the identifiability condition as observed from our experiments.

Then, following the above insight, we introduce a simple strategy in our implementation to guarantee the identifiability for matrix completion with high probability. Specifically, we introduce a constraint $b_{\min} \leq b \leq b_{\max}$ in (21) to control the window size parameter b, where $b_{\min}$ is chosen such that the number of elements in the observation set Ω formed in (3) is no fewer that ωN^2, where ω is a target sampling ratio which is typically chosen to meet the asymptotic scaling rule $\omega N^2 \geq CN\log^2(N)$ [29] with a sufficiently large parameter C to guarantee sufficient observations for small to medium N. In addition, after obtaining b from solving (21), we adopt a simple identifiability check for the observation set Ω, where we require each row and each column of the sparse observation matrix $\hat{\boldsymbol{H}}$ has at least one entry in Ω. If such a criterion is not satisfied, then b is repeatedly increased by 20% until such a criterion is met. We have observed that, following this strategy, the proposed matrix completion is robust for small to large M.

D. Complexity

The computation of the interpolation assisted matrix completion method consists of two parts. First, the interpolation step has a computational complexity $\mathcal{O}(M^2\omega N^2)$ as M is the number of measurements and N represents the resolution of the reconstructed map. Here, $\omega < 1$ is the interpolation rate, since only a subset of grid points are constructed by the interpolation, and moreover, the interpolation of each grid point requires only a *subset* of M measurements nearby. More specifically, if a uniform sampling approach is adopted to form Ω as discussed in Section II-B, then only $CN\log^2(N)$ grid cells are needed to be constructed. As such, ω can be upper bounded by $CN\log^2(N)/N^2$, and the complexity of the interpolation step is $\mathcal{O}(M^2N\log^2(N))$.

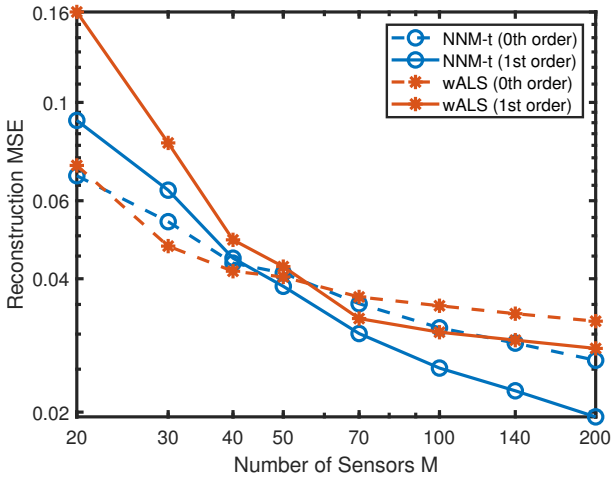


Figure 5. Reconstruction MSE for $S = 1$ and $N = 30$. Zeroth order interpolation performs better for a small number sensors, whereas, first order method prefers more sensors.

Second, for the NNM-t approach, the matrix completion step has complexity $\mathcal{O}(CN\log^3(2N))$, where C is a constant that captures the *incoherence* of the matrix and the accuracy requirement of the matrix completion algorithm [42]. Thus, the total complexity of the NNM-t method is $\mathcal{O}(M^2\omega N^2 + CN\log^3(2N))$. In particular, under uniform sampling, the total complexity can be simplified to $\mathcal{O}(M^2N\log^2(N))$ for large N and M.

To benchmark, Kriging has a computational complexity of $\mathcal{O}(M^3N^2)$ for constructing N^2 points [43], because, first, Kriging needs to learn a semi-variogram, and second, Kriging requires to use all the M measurements to reconstruct every points. A numerical comparison on the computational time is reported in Table I.

V. Numerical Results

We adopt model (1) to simulate the propagation map in an $L \times L$ area for $L = 2$ kilometers, where $g_k(d) = P_k d^{-1.5} A(f)^{-d}$, with parameter $A(f) = 0.8$, corresponding to an empirical energy field of underwater acoustic signal at frequency $f = 5$ kHz [37], [44], $d = \sqrt{x^2 + y^2 + h^2}$ represents the distance from the source, (x, y) is the coordinate, $h = 400$ meters is the depth of interest for a multiple source scenario we test later, and $h = 1$ kilometers is for a single source scenario. The parameter P_k follows an independent exponential distribution with rate parameter $\kappa = 1$ to model the power emitted from source k. 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 $d_c = 200$ meters, $\sigma_s^2 = 1$. The sensors are randomly and independently distributed in the area following a uniform distribution. The measurement model in (2) is adopted, where the measurement noise ϵ is modeled as a zero mean Gaussian random variable with standard deviation $\sigma = 0.02$.

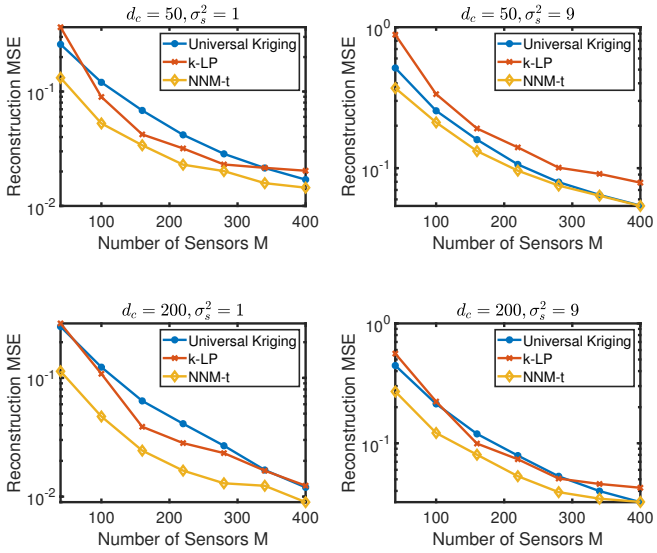


Figure 6. Performance under different shadowing parameters.

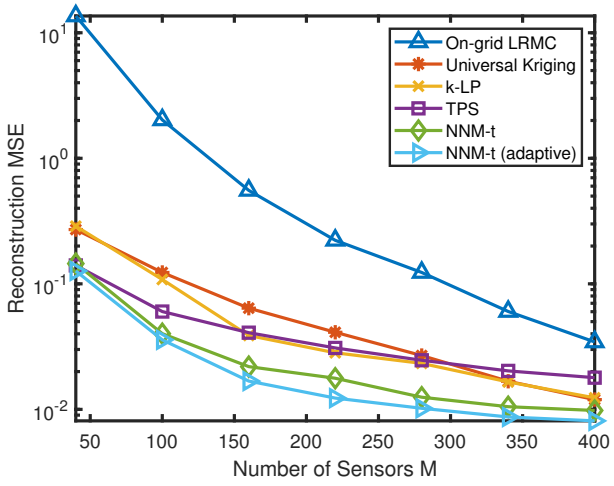


Figure 7. Reconstruction MSE for $S = 3$ and $N = 30$.

The MSE of the reconstructed propagation map is employed for performance evaluation, which is calculated through $1/N^2||\bar{\boldsymbol{H}} - \boldsymbol{H}||_F^2$ as stated in Section III-D2.

The proposed interpolation assisted matrix completion is implemented following Algorithm 1 under $M_0 = 7$ and confident level parameter $\delta = 0.05$ as explained in Sections III-D and IV-B2, respectively. We evaluate two versions of implementation of the proposed scheme. Scheme 1) NNM-t: The observation set Ω is constructed using uniformly random sampling with $C = 1.6$ as described in Section II-B, and $b_{\max} = 0.35L$ to guarantee that the interpolation only focuses on a local area. To circumvent the possible singularity issue of estimating the interpolation parameters via computing (24) and (25) for a too small b, where b serves as the kernel parameter in (6), we lower bound the kernel parameter as $b' = \max\{b, b_{\min}^{(i,j)}\}$ for each specific entry $(i,j) \in \Omega$, where $b_{\min}^{(i,j)}$ is the minimum radius from the grid center $\boldsymbol{c}_{ij}$ to include M_0 sensor measurements. Scheme 2) NNM-t (adaptive): The observation set Ω is constructed using sensor-aware sampling with an optimized b via solving (21), where the parameter of identifiability check is set as $C = 2.5$.[2] In addition, for each grid cell, we optimize an individual window size b_{ij} by solving (21) while dropping the summation. The observation $\hat{H}_{ij}$ and the confident interval $\mathcal{I}_{ij}$ are computed based on b_{ij}.

The performance is compared with the following baselines that are recently developed or adopted in related literature. Baseline 1: Kriging method [12], [45], the propagation map is reconstructed by Universal Kriging $\hat{\rho}(\boldsymbol{c}) = \hat{m}(\boldsymbol{c}) + \epsilon(\boldsymbol{c})$, where $\hat{m}(\boldsymbol{c})$ is the interpolated deterministic part, and $\epsilon(\boldsymbol{c})$ is the interpolated residual which can be estimated with an Ordinary Kriging method. Baseline 2: k-nearest neighbor local polynomial interpolation (k-LP) [46], a first order local polynomial regression method is used to estimate the RSS using k nearest measurements, where k is chosen through cross validation. Baseline 3: On-grid low rank matrix completion (On-grid LRMC) [38], based on the sensor topology, we first form a refined on-grid measurement set as follows: if there is a sensor in a grid cell, we re-sample the grid cell by sampling at the grid center, and hence, there will be no discretization error. Note that since this approach is not practical in real life, this scheme is for performance benchmarking only. Then, the matrix is completed by solving a conventional NNM problem. Baseline 4: Thin plate spline (TPS) [47], a thin plate spline method is used to construct $\hat{\rho}(\boldsymbol{c})$ from the scattered measurements $\{\boldsymbol{z}_m, \gamma_m\}$.

[2]The parameter $C = 1.6$ and 2.5 correspond to 60% and over 90% observation ratio, respectively, at matrix dimension $N = 30$. Hence, the NNM-t (adaptive) scheme requires high complexity in the interpolation step.

A. Zeroth Order or First Order Interpolation

We evaluate the performance of the proposed schemes over different interpolation models under different number of measurement samples. For both NNM-t and wALS schemes, the zeroth order interpolation (12) and first order interpolation (13) are used to construct $\hat{H}_{ij}$. In a single source propagation map, $S = 1$, and for $N = 30$, Fig. 5 shows that when there are only a few measurements, the methods based on zeroth order interpolation attain a smaller MSE than their counterparts based on first order models. This is because zeroth order models have less parameters to estimate. On the other hand, for medium to large number of measurements, methods based on first order interpolation perform better, because first order interpolation can substantially eliminate the bias in the local interpolation as revealed in our asymptotic analysis stated in Theorems 3 and 4.

B. Performance under Different Shadowing Parameters

We quantify the propagation map reconstruction performance of the proposed schemes under different shadowing parameters. The simulations are performed under $d_c = \{50, 200\}$ meters and $\sigma_s^2 = \{1, 9\}$ separately with the baselines are Universal Kriging and k-LP. The experiment results shown in Fig. 6 reveal that the increasing of shadowing variance

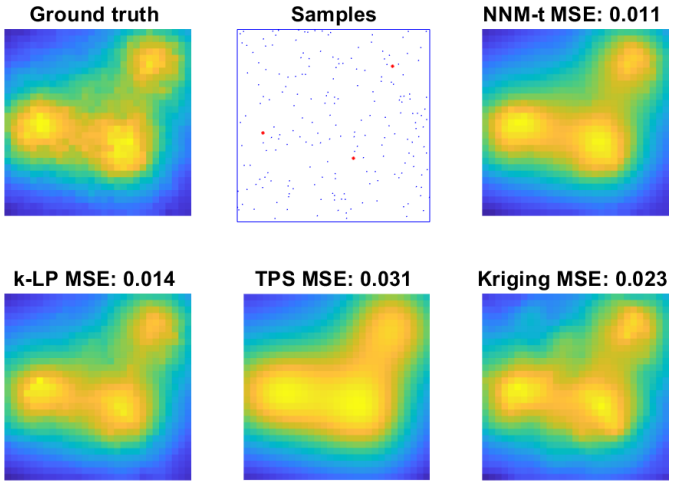


Figure 8. Ground truth, sampling pattern, and the reconstruction MSE of different methods.

decreases the performance of all baselines. The larger the correlation distance, the better the performance.

C. Performance under Different Measurements

We quantify the propagation map reconstruction performance of the proposed schemes under different number of measurements $M = 40 - 400$ with fixed resolution $N = 30$ and number of sources $S = 3$.

Fig. 7 demonstrates the reconstruction MSE versus the number of sensors M. A visual plot of a realization of the propagation field, sampling pattern, and the reconstruction results under $M = 200$ are demonstrated in Fig. 8. It is observed that the proposed NNM-t method can realize 10%–50% MSE reduction from all baseline schemes under medium to large M, which translates to a saving of roughly $1/3$ of sensor measurements at a medium M. At small M, *i.e.*, $M = 40$, its performance is similar to that of the TPS scheme. The NNM-t (adaptive) method works better than the NNM-t method and outperforms all baseline schemes because it optimizes the window size parameter b_{ij} for each grid cell in an adaptive way. The Universal Kriging and k-LP schemes perform well at a large M, but their performance degrades at small M, whereas, the TPS scheme works well at small M, but suffers from substantial performance degradation at large M. Finally, the On-grid LRMC scheme performs badly because the number of observations $M = 40$–400 is substantially smaller than the total number of entries $N^2 = 900$ to be recovered, and the matrix is likely non-identifiable in this regime.

Table I summarizes the computation time based on our implementation using Matlab. The Universal Kriging baseline is implemented by the "mGstat" function provided in the Geostatistical toolbox in Matlab. The NNM-t problem is solved by CVX using an interior point method. The complexity of k-LP and LRMC schemes is for benchmarking because they are the building blocks of the proposed NNM-t. As observed in Table I, the proposed method has a lower complexity than

Table I
COMPUTATION TIME (SECONDS) AT DIFFERENT M UNDER $N = 30$

M	40	100	160	220	280	340	400
k-LP	0.39	0.53	0.79	1.19	1.64	1.85	2.05
LRMC	6.63	6.39	6.60	6.76	6.45	7.11	6.91
U-Kriging	2.51	5.77	10.63	19.49	28.32	40.37	55.76
NNM-t	8.26	6.10	6.67	8.57	6.72	6.98	8.73

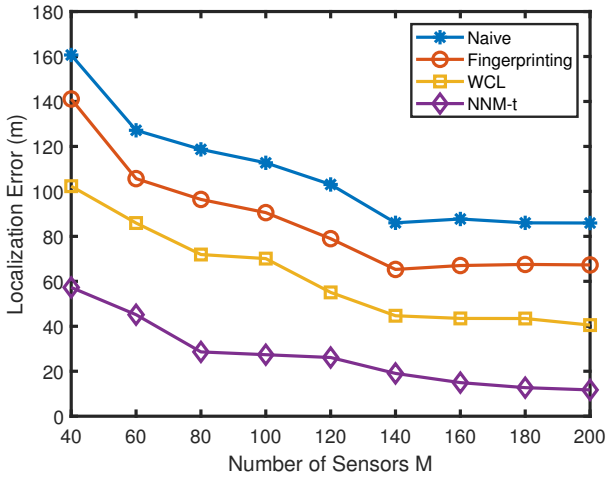


Figure 9. Source localization error for $S = 1$ and $N = 30$.

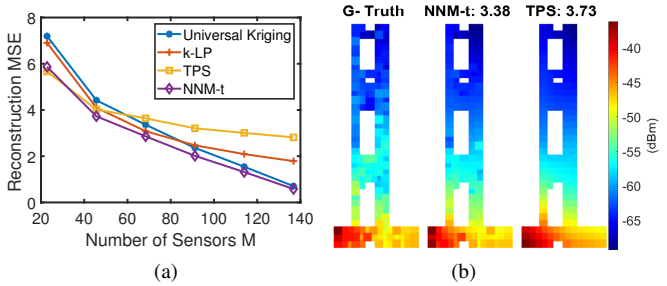
Universal Kriging at medium to large M as expected from our analysis.

The gain of the proposed method over a pure interpolation scheme can be interpreted from the following two aspects. First, although there is shadowing, the matrix of the propagation field is still relatively low rank when choosing a large N for the matrix dimension. In addition, the proposed NNM-t method can dynamically adjust the windows size b such that there will always be enough observations to guarantee the recoverability of the matrix completion. Second, a pure interpolation method suffers from performance loss in the regions where measurements are locally too sparse, typically, at the edge of the area of interest. By contrast, the proposed interpolation assisted matrix completion approach can fill in the missing values in these regions using the global information constructed from the other part of the propagation field.

D. Application in RSS-based Source Localization

Based on the reconstructed propagation map, we demonstrate the application of propagation map reconstruction under NNM-t for $S = 1$ to RSS-based source localization. The localization algorithm based on matrix factorization developed in [2] is adopted. Specifically, based on the reconstructed propagation map matrix $\bar{\boldsymbol{H}}$, the dominant singular vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ of $\bar{\boldsymbol{H}}$ are computed. Then, the coordinates of the source are estimated via peak localization of the dominant singular vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ [2].

Three baseline schemes for localization are evaluated. Baseline 1: WCL [28], we use $\hat{\boldsymbol{s}}_{\text{WCL}} = \sum_{m=1}^{M} w_m \boldsymbol{z}_m / \sum_{m=1}^{M} w_m$



(a) (b)

Figure 10. The reconstruction performance evaluated on real data set. (a) The reconstruction MSE versus the number of measurements used for propagation map construction. (b) Visualization of the propagation field in an indoor environment, showing the ground truth, the reconstruction from NNM-t, and the reconstruction from TPS.

to estimate the location of the source, where $w_m = \gamma_m$ serves as the weight. Baseline 2: naive method, the location of the source is chosen as the location of the sensor with largest measurements. Baseline 3: Fingerprinting [48], the propagation map is first constructed using the Delaunay triangulation, and then, the location with the largest RSS in the reconstructed propagation map is selected as the location estimate of the source.

Fig. 9 shows the RMSE of RSS-based source localization based on the propagation map reconstructed from the proposed NNM-t method. The RMSE is reduced by roughly a half and both baseline schemes require 2 times more sensors to achieve the same localization accuracy.

E. Experiment on a Real Dataset

We extend the experiment to a real dataset [49]. In this dataset, there are $M = 166$ RSS measurements for each emitter. The measurements are taken in a 14×34 m^2 indoor area. In the implementation of the proposed algorithm, the area is divided into 1×1 m^2 grid cells such that the 166 measurements are in the grid center. Among the 166 measurements, $M' = 20$ – 140 measurements are randomly selected for reconstructing the prorogation field. The results in reconstruction MSE are reported in Fig. 10. It is shown that the proposed NNM-t method also achieves a promising performance on a real dataset.

VI. Conclusion

In this paper, an interpolation assisted matrix completion method is proposed to reconstruct a propagation map from a sparse set of signal strength measurements. The observation matrix is enriched through interpolation and the bias and variance of the interpolation method are analyzed to build uncertainty-aware matrix completion algorithms which integrate interpolation with matrix completion. Furthermore, a minimum MSE method to optimize the window size for local interpolation is proposed. The numerical results demonstrate that the reconstruction MSE of the propagation map can be reduced by 10%–50% compared to the state-of-the-art schemes which corresponds to a saving of nearly half of measurements under a not-so-dense sampling. In the application of RSS-based source localization, the RMSE of localization is reduced by more than 50% from a WCL baseline.

Appendix A
Proof of Theorem 1

Applying first-order Taylor's expansion to $\rho(\boldsymbol{z})$ at the neighborhood of $\boldsymbol{c}_{ij}$, the observation model in (2) becomes

$$\begin{aligned}\gamma_m &= \rho(\boldsymbol{z}_m) + \epsilon_m \\ &= \rho(\boldsymbol{c}_{ij}) + \nabla\rho^{\mathrm{T}}(\boldsymbol{c}_{ij})(\boldsymbol{z}_m - \boldsymbol{c}_{ij}) + o(||\boldsymbol{z}_m - \boldsymbol{c}_{ij}||) + \epsilon_m\end{aligned} \tag{29}$$

due to the first order differentiability of $\rho(\boldsymbol{z})$.

Recall that $\bar{w}_m(\boldsymbol{c}_{ij}) \triangleq w_m(\boldsymbol{c}_{ij}) / \sum_{i=1}^{M} w_i(\boldsymbol{c}_{ij})$, and therefore, $\sum_{m=1}^{M} \bar{w}_m(\boldsymbol{c}_{ij}) = 1$. The expectation of the zeroth order solution $\hat{\alpha}(\boldsymbol{c}_{ij})$ in (12) can be written as

$$\begin{aligned}&\mathbb{E}\left\{\hat{\alpha}(\boldsymbol{c}_{ij})\right\} \\ &= \mathbb{E}\left\{\frac{\sum_{m=1}^{M} w_m(\boldsymbol{c}_{ij})\gamma_m}{\sum_{m=1}^{M} w_m(\boldsymbol{c}_{ij})}\right\} \\ &= \mathbb{E}\Bigg\{\sum_{m=1}^{M} \bar{w}_m(\boldsymbol{c}_{ij})(\rho(\boldsymbol{c}_{ij}) + \nabla\rho^{\mathrm{T}}(\boldsymbol{c}_{ij})(\boldsymbol{z}_m - \boldsymbol{c}_{ij}) \\ &\qquad + o(||\boldsymbol{z}_m - \boldsymbol{c}_{ij}||) + \epsilon_m)\Bigg\} \\ &= \mathbb{E}\Bigg\{\sum_{m=1}^{M} \bar{w}_m(\boldsymbol{c}_{ij})(\rho(\boldsymbol{c}_{ij}) + \nabla\rho^{\mathrm{T}}(\boldsymbol{c}_{ij})(\boldsymbol{z}_m - \boldsymbol{c}_{ij}) \\ &\qquad + o(||\boldsymbol{z}_m - \boldsymbol{c}_{ij}||))\Bigg\} + \sum_{m=1}^{M} \bar{w}_m(\boldsymbol{c}_{ij})\mathbb{E}\left\{\epsilon_m\right\} \\ &= \rho(\boldsymbol{c}_{ij}) + \sum_{m=1}^{M} \bar{w}_m(\boldsymbol{c}_{ij})(\nabla\rho^{\mathrm{T}}(\boldsymbol{c}_{ij})(\boldsymbol{z}_m - \boldsymbol{c}_{ij})) + o(b)\end{aligned} \tag{30}$$

where $o(b)$ in the last equation is due to the fact that $o(||\boldsymbol{z}_m - \boldsymbol{c}_{ij}||)/b \to 0$ as $b \to 0$ since we are only interested in the (i,j)th grid such that $||\boldsymbol{z}_m - \boldsymbol{c}_{ij}|| < b$ according to the interpolation strategy.

As a result, the bias

$$\begin{aligned}\mathbb{E}\{\xi_{ij}\} &= \mathbb{E}\left\{\hat{\alpha}(\boldsymbol{c}_{ij})\right\} - \rho(\boldsymbol{c}_{ij}) \\ &= \sum_{m=1}^{M} \bar{w}_m(\boldsymbol{c}_{ij})(\nabla\rho^{\mathrm{T}}(\boldsymbol{c}_{ij})(\boldsymbol{z}_m - \boldsymbol{c}_{ij})) + o(b).\end{aligned}$$

For the variance of $\xi_{ij} = \hat{\alpha}(\boldsymbol{c}_{ij}) - \rho(\boldsymbol{c}_{ij})$, we note that $\mathbb{V}\left\{\xi_{ij}\right\} = \mathbb{V}\left\{\hat{\alpha}(\boldsymbol{c}_{ij})\right\}$ because $\rho(\boldsymbol{c}_{ij})$ is deterministic. From (12) and (2),

$$\hat{\alpha}(\boldsymbol{c}_{ij}) = \sum_{m=1}^{M} \bar{w}_m\rho(\boldsymbol{c}_{ij}) + \sum_{m=1}^{M} \bar{w}_m\epsilon_m.$$

So, $\mathbb{V}\left\{\xi_{ij}\right\} = \mathbb{V}\left\{\sum_{m=1}^{M} \bar{w}_m\epsilon_m\right\} = \sum_{m=1}^{M} \bar{w}_m^2(\boldsymbol{c}_{ij})\sigma^2$, since $\mathbb{V}\left\{\epsilon_m\right\} = \sigma^2$ from the observation model in (2). The results in Theorem 1 are thus proven.

APPENDIX B
PROOF OF THEOREM 2

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

$$\begin{aligned}\rho\left(\boldsymbol{z}_m\right) &= \rho(\boldsymbol{c}_{ij}) + \nabla\rho^{\mathrm{T}}(\boldsymbol{c}_{ij})(\boldsymbol{z}_m - \boldsymbol{c}_{ij}) \\ &\quad + \frac{1}{2}(\boldsymbol{z}_m - \boldsymbol{c}_{ij})^{\mathrm{T}}\boldsymbol{\Psi}_{ij}(\boldsymbol{z}_m - \boldsymbol{c}_{ij}) + o(\|\boldsymbol{z}_m - \boldsymbol{c}_{ij}\|^2)\end{aligned} \tag{31}$$

where $\boldsymbol{\Psi}_{ij} = \nabla^2\rho(\boldsymbol{c}_{ij})$ is the Hessian matrix of $\rho(\boldsymbol{z})$ evaluated at point $\boldsymbol{c}_{ij}$.

Denote $\boldsymbol{\varrho} \triangleq [\rho(\boldsymbol{z}_1), \cdots, \rho(\boldsymbol{z}_M)]^{\mathrm{T}}$. Then the expression in (31) can be rearranged into the following matrix form

$$\boldsymbol{\varrho} = \tilde{\boldsymbol{D}}_{ij}^{\mathrm{T}}\begin{bmatrix}\rho(\boldsymbol{c}_{ij}) \\ \nabla\rho(\boldsymbol{c}_{ij})\end{bmatrix} + \frac{1}{2}\mathrm{diag}\{\boldsymbol{D}_{ij}^{\mathrm{T}}\boldsymbol{\Psi}_{ij}\boldsymbol{D}_{ij}\} + \boldsymbol{r}_{ij} \tag{32}$$

where $\boldsymbol{r}_{ij}$ is a vector of the residual terms $o(||\boldsymbol{z}_m - \boldsymbol{c}_{ij}||^2)$. Noticing that the residual from the Taylor's expansion scales as $||\boldsymbol{z}_m - \boldsymbol{c}_{ij}||^2$ with a bound $||\boldsymbol{z}_m - \boldsymbol{c}_{ij}|| < b$ due to the interpolation strategy, thus we have $o(||\boldsymbol{z}_m - \boldsymbol{c}_{ij}||^2)/b^2 \to 0$ as $b \to 0$. Therefore, $\boldsymbol{r}_{ij} \sim o(b^2)$.

From the measurement model (2), we have $\boldsymbol{\gamma} = \boldsymbol{\varrho} + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon} = [\epsilon_1, \cdots, \epsilon_M]^{\mathrm{T}}$. The expectation of the first order solution $\hat{\boldsymbol{\theta}} = \begin{bmatrix}\hat{\alpha}(\boldsymbol{c}_{ij}) & \hat{\boldsymbol{\beta}}(\boldsymbol{c}_{ij})\end{bmatrix}^{\mathrm{T}}$ in (13) can be written as

$$\mathbb{E}\{\hat{\boldsymbol{\theta}}\} = \mathbb{E}\left\{\tilde{\boldsymbol{W}}_{ij}^{-1}\tilde{\boldsymbol{D}}_{ij}\boldsymbol{W}_{ij}\boldsymbol{\gamma}\right\} = \tilde{\boldsymbol{W}}_{ij}^{-1}\tilde{\boldsymbol{D}}_{ij}\boldsymbol{W}_{ij}\boldsymbol{\varrho} \tag{33}$$

since the noise $\boldsymbol{\epsilon}$ has zero mean, where $\tilde{\boldsymbol{W}}_{ij} = \tilde{\boldsymbol{D}}_{ij}\boldsymbol{W}_{ij}\tilde{\boldsymbol{D}}_{ij}^{\mathrm{T}}$.

Substitute $\boldsymbol{\varrho}$ in (33) with (32), the expectation of $\hat{\boldsymbol{\theta}}$ can be rewritten as

$$\begin{aligned}&\mathbb{E}\{\hat{\boldsymbol{\theta}}\} \\ =&\tilde{\boldsymbol{W}}_{ij}^{-1}\tilde{\boldsymbol{D}}_{ij}\boldsymbol{W}_{ij} \\ &\times\left(\tilde{\boldsymbol{D}}_{ij}^{\mathrm{T}}\begin{bmatrix}\rho(\boldsymbol{c}_{ij}) \\ \nabla\rho(\boldsymbol{c}_{ij})\end{bmatrix} + \frac{1}{2}\mathrm{diag}\{\boldsymbol{D}_{ij}^{\mathrm{T}}\boldsymbol{\Psi}_{ij}\boldsymbol{D}_{ij}\} + \boldsymbol{r}_{ij}\right) \\ =&\begin{bmatrix}\rho(\boldsymbol{c}_{ij}) \\ \nabla\rho(\boldsymbol{c}_{ij})\end{bmatrix} \\ &+ \tilde{\boldsymbol{W}}_{ij}^{-1}\tilde{\boldsymbol{D}}_{ij}\boldsymbol{W}_{ij}\left(\frac{1}{2}\mathrm{diag}\{\boldsymbol{D}_{ij}^{\mathrm{T}}\boldsymbol{\Psi}_{ij}\boldsymbol{D}_{ij}\}\right) + o(b^2).\end{aligned}$$

As a result, the bias

$$\begin{aligned}\mathbb{E}\{\xi_{ij}\} &= \left[\mathbb{E}\{\hat{\boldsymbol{\theta}}\} - \begin{bmatrix}\rho(\boldsymbol{c}_{ij}) \\ \nabla\rho(\boldsymbol{c}_{ij})\end{bmatrix}\right]_{(1,1)} \\ &= \frac{1}{2}\left[\tilde{\boldsymbol{W}}_{ij}^{-1}\tilde{\boldsymbol{D}}_{ij}\boldsymbol{W}_{ij}\mathrm{diag}\{\boldsymbol{D}_{ij}^{\mathrm{T}}\boldsymbol{\Psi}_{ij}\boldsymbol{D}_{ij}\}\right]_{(1,1)} + o(b^2)\end{aligned}$$

where the operation $[\boldsymbol{A}]_{(1,1)}$ returns the $(1,1)$th entry of a matrix $\boldsymbol{A}$.

The variance of $\hat{\boldsymbol{\theta}}$ can be derived following

$$\begin{aligned}\mathbb{V}\{\hat{\boldsymbol{\theta}}\} &= \mathbb{E}\left\{\left(\hat{\boldsymbol{\theta}} - \mathbb{E}\{\hat{\boldsymbol{\theta}}\}\right)^2\right\} \\ &= \mathbb{E}\left\{\left(\tilde{\boldsymbol{W}}_{ij}^{-1}\tilde{\boldsymbol{D}}_{ij}\boldsymbol{W}_{ij}\boldsymbol{\gamma} - \tilde{\boldsymbol{W}}_{ij}^{-1}\tilde{\boldsymbol{D}}_{ij}\boldsymbol{W}_{ij}\boldsymbol{\varrho}\right)^2\right\} \\ &= \mathbb{E}\left\{\left(\tilde{\boldsymbol{W}}_{ij}^{-1}\tilde{\boldsymbol{D}}_{ij}\boldsymbol{W}_{ij}(\boldsymbol{\gamma} - \boldsymbol{\varrho})\right)^2\right\} \\ &= \mathbb{E}\left\{\left(\tilde{\boldsymbol{W}}_{ij}^{-1}\tilde{\boldsymbol{D}}_{ij}\boldsymbol{W}_{ij}\boldsymbol{\epsilon}\right)^2\right\} \\ &= \sigma^2\tilde{\boldsymbol{W}}_{ij}^{-1}\left(\tilde{\boldsymbol{D}}_{ij}\boldsymbol{W}_{ij}\boldsymbol{W}_{ij}^{\mathrm{T}}\tilde{\boldsymbol{D}}_{ij}^{\mathrm{T}}\right)\tilde{\boldsymbol{W}}_{ij}^{-1}.\end{aligned}$$

Since $\boldsymbol{\theta}$ is deterministic, $\mathbb{V}\{\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\} = \mathbb{V}\{\hat{\boldsymbol{\theta}}\}$, therefore $\mathbb{V}\{\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\} = \sigma^2\tilde{\boldsymbol{W}}_{ij}^{-1}\left(\tilde{\boldsymbol{D}}_{ij}\boldsymbol{W}_{ij}\boldsymbol{W}_{ij}^{\mathrm{T}}\tilde{\boldsymbol{D}}_{ij}^{\mathrm{T}}\right)\tilde{\boldsymbol{W}}_{ij}^{-1}$.

As a result,

$$\begin{aligned}\mathbb{V}\{\xi_{ij}\} &= \left[\mathbb{V}\{\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\}\right]_{(1,1)} \\ &= \sigma^2\left[\tilde{\boldsymbol{W}}_{ij}^{-1}\left(\tilde{\boldsymbol{D}}_{ij}\boldsymbol{W}_{ij}\boldsymbol{W}_{ij}^{\mathrm{T}}\tilde{\boldsymbol{D}}_{ij}^{\mathrm{T}}\right)\tilde{\boldsymbol{W}}_{ij}^{-1}\right]_{(1,1)}.\end{aligned}$$

APPENDIX C
PROOF OF THEOREM 3

The observation model in (2) can be rewritten as

$$\begin{aligned}\gamma_m &= \rho(\boldsymbol{z}_m) + \epsilon_m \\ &= \rho(\boldsymbol{c}_{ij}) + (\rho(\boldsymbol{z}_m) - \rho(\boldsymbol{c}_{ij})) + \epsilon_m.\end{aligned} \tag{34}$$

From (34), we have

$$\begin{aligned}&\frac{1}{Mb^2}\sum_{m=1}^{M} w_m(\boldsymbol{c}_{ij})\gamma_m \\ &= \frac{1}{Mb^2}\sum_{m=1}^{M} w_m(\boldsymbol{c}_{ij})\rho(\boldsymbol{c}_{ij}) \\ &\quad + \frac{1}{Mb^2}\sum_{m=1}^{M} w_m(\boldsymbol{c}_{ij})(\rho(\boldsymbol{z}_m) - \rho(\boldsymbol{c}_{ij})) \\ &\quad + \frac{1}{Mb^2}\sum_{m=1}^{M} w_m(\boldsymbol{c}_{ij})\epsilon_m\end{aligned} \tag{35}$$

where in the rest of the proof, the three terms in (35) are respectively denoted as

$$\hat{f}(\boldsymbol{c}_{ij}) \triangleq \frac{1}{Mb^2}\sum_{m=1}^{M} w_m(\boldsymbol{c}_{ij}) \tag{36}$$

which serves as an approximation of the sensor density at point $\boldsymbol{c}_{ij}$,

$$\varepsilon_1(\boldsymbol{c}_{ij}) \triangleq \frac{1}{Mb^2}\sum_{m=1}^{M} w_m(\boldsymbol{c}_{ij})(\rho(\boldsymbol{z}_m) - \rho(\boldsymbol{c}_{ij})) \tag{37}$$

which quantifies the weighted average bias of the estimate at location $\boldsymbol{c}_{ij}$, and

$$\varepsilon_2(\boldsymbol{c}_{ij}) \triangleq \frac{1}{Mb^2}\sum_{m=1}^{M} = w_m(\boldsymbol{c}_{ij})\epsilon_m \tag{38}$$

which quantifies the weighted average error due to the measurement noise.

Then, dividing $\hat{f}(\boldsymbol{c}_{ij})$ on both sides of (35) and according to the zeroth order solution (12), it yields

$$\begin{aligned}\hat{\rho}(\boldsymbol{c}_{ij}) &= \frac{\sum_{m=1}^{M} w_m(\boldsymbol{c}_{ij})\gamma_m}{\sum_{m=1}^{M} w_m(\boldsymbol{c}_{ij})} \\ &= \rho(\boldsymbol{c}_{ij}) + \frac{\varepsilon_1(\boldsymbol{c}_{ij})}{\hat{f}(\boldsymbol{c}_{ij})} + \frac{\varepsilon_2(\boldsymbol{c}_{ij})}{\hat{f}(\boldsymbol{c}_{ij})}\end{aligned}$$

which implies that the error equals to

$$\xi_{ij} = \hat{\rho}(\boldsymbol{c}_{ij}) - \rho(\boldsymbol{c}_{ij}) = \frac{\varepsilon_1(\boldsymbol{c}_{ij})}{\hat{f}(\boldsymbol{c}_{ij})} + \frac{\varepsilon_2(\boldsymbol{c}_{ij})}{\hat{f}(\boldsymbol{c}_{ij})}. \tag{39}$$

To analyze $\varepsilon_1(\boldsymbol{c}_{ij})$ and $\varepsilon_2(\boldsymbol{c}_{ij})$, we derive the following lemmas.

Lemma 1. *Suppose that $f(\boldsymbol{z})$ is second order differentiable and bounded. Then, for a sufficiently small b,*

$$\int K\Big(\frac{\boldsymbol{x}-\boldsymbol{c}_{ij}}{b}\Big)^2 f(\boldsymbol{x})d\boldsymbol{x} = C_1 b^2 f(\boldsymbol{c}_{ij}) + o\left(b^3\right).$$

In addition, for a function $g(\boldsymbol{x})$ that satisfies $g(b\boldsymbol{u})/b \to 0$ as $b \to 0$, uniformly for all $||\boldsymbol{u}|| \leq 1$, there is

$$\frac{1}{b}\int K(\boldsymbol{u})g(b\boldsymbol{u})f(\boldsymbol{z})d\boldsymbol{u} \to 0$$

as $b \to 0$, for all $\boldsymbol{z}$, where $C_1 = \int K(\boldsymbol{u})^2 d\boldsymbol{u}$.

Proof. Let $\boldsymbol{u} = (\boldsymbol{x}-\boldsymbol{c}_{ij})/b$. Note that $\int d\boldsymbol{x} = \int db\boldsymbol{u} = \iint dbu_x dbu_y = \iint b^2 du_x du_y = \int b^2 d\boldsymbol{u}$. We have

$$\int K\Big(\frac{\boldsymbol{x}-\boldsymbol{c}_{ij}}{b}\Big)^2 f(\boldsymbol{x})d\boldsymbol{x} = b^2\int K(\boldsymbol{u})^2 f(b\boldsymbol{u}+\boldsymbol{c}_{ij})d\boldsymbol{u}. \tag{40}$$

Consider the first order Taylor's expansion

$$f(b\boldsymbol{u}+\boldsymbol{c}_{ij}) = f(\boldsymbol{c}_{ij}) + \nabla f(\boldsymbol{c}_{ij})^{\mathrm{T}} b\boldsymbol{u} + R_f(b\boldsymbol{u}) \tag{41}$$

where ∇f is the first order derivative of f, $R_f(b\boldsymbol{u})$ is the remainder term that satisfies $R_f(b\boldsymbol{u})/b \to 0$ as $b \to 0$ uniformly for all $\boldsymbol{u}$ with $\|\boldsymbol{u}\| \leq 1$ due to the Taylor's theorem and the fact that $f(\boldsymbol{z})$ is second order differentiable. Thus,

$$\begin{aligned}
&\int K(\boldsymbol{u})^2 f(b\boldsymbol{u}+\boldsymbol{c}_{ij})d\boldsymbol{u} \\
&= \int K(\boldsymbol{u})^2 \left(f(\boldsymbol{c}_{ij}) + \nabla f(\boldsymbol{c}_{ij})^{\mathrm{T}} b\boldsymbol{u} + R_f(b\boldsymbol{u})\right) d\boldsymbol{u} \\
&= f(\boldsymbol{c}_{ij})\int K(\boldsymbol{u})^2 d\boldsymbol{u} + b\int K(\boldsymbol{u})^2 \nabla f(\boldsymbol{c}_{ij})^{\mathrm{T}}\boldsymbol{u} d\boldsymbol{u} \quad (42)\\
&\quad + \int K(\boldsymbol{u})^2 R_f(b\boldsymbol{u})d\boldsymbol{u} \quad (43)\\
&= C_1 f(\boldsymbol{c}_{ij}) + o(b) \quad (44)
\end{aligned}$$

where the second term in (42) equals 0, since $\iint K(\boldsymbol{u})^2 (a_1 u_x + a_2 u_y) du_x du_y = 0$ due to (8). For (43), since the support of $K(\boldsymbol{u})$ is $\mathcal{C} = \left\{\boldsymbol{u} \in \mathbb{R}^2 : ||\boldsymbol{u}||_2 < 1\right\}$, we have

$$\begin{aligned}
&\frac{1}{b}\int K(\boldsymbol{u})^2 R_f(b\boldsymbol{u})d\boldsymbol{u} \\
&\leq \frac{1}{b}\Big|\sup_{\boldsymbol{u}\in\mathcal{C}} R_f(b\boldsymbol{u})\Big| \cdot \Big|\int_{\mathcal{C}} K(\boldsymbol{u})^2 d\boldsymbol{u}\Big|
\end{aligned} \tag{45}$$

which converges to 0 because the term $R_f(b\boldsymbol{u})/b$ uniformly converges to 0 and the integral term is bounded.

Multiplying (44) with b^2 and substituting the result in (40) confirms the result in Lemma 1. □

Define a random variable $X_m = w_m(\boldsymbol{c}_{ij})\epsilon_m$, where recall that $w_m(\boldsymbol{c}_{ij}) = K\big(\frac{\boldsymbol{z}_m - \boldsymbol{c}_{ij}}{b}\big)$. Then, using Lemma 1, the random variable X_m can be shown to have the following property.

Lemma 2. *The mean and variance of X_m are given by $\mathbb{E}\{X_m\} = 0$ and $\mathbb{V}\{X_m\} = C_1\sigma^2 b^2 f(\boldsymbol{c}_{ij}) + o(b^3)$.*

Proof. Since $\mathbb{E}\{\epsilon_m\} = 0$ and ϵ_m is independent of the term $w_m(\boldsymbol{c}_{ij})$, it follows that $\mathbb{E}\{X_m\} = 0$.

Moreover,

$$\begin{aligned}
\mathbb{V}\{X_m\} &= \mathbb{E}\left\{K\Big(\frac{\boldsymbol{z}_m-\boldsymbol{c}_{ij}}{b}\Big)^2 \epsilon_m^2\right\} \\
&= \mathbb{E}\left\{K\Big(\frac{\boldsymbol{z}_m-\boldsymbol{c}_{ij}}{b}\Big)^2\right\}\mathbb{E}\{\epsilon_m^2\} \\
&= \sigma^2\mathbb{E}\left\{K\Big(\frac{\boldsymbol{z}_m-\boldsymbol{c}_{ij}}{b}\Big)^2\right\} \\
&= \sigma^2\int K\Big(\frac{\boldsymbol{z}_m-\boldsymbol{c}_{ij}}{b}\Big)^2 f(\boldsymbol{x})d\boldsymbol{x} \\
&= C_1\sigma^2 b^2 f(\boldsymbol{c}_{ij}) + o(b^3)
\end{aligned} \tag{46}$$

where the last equation is from Lemma 1. □

Define a random variable $Y_m = w_m(\boldsymbol{c}_{ij})(\rho(\boldsymbol{z}_m) - \rho(\boldsymbol{c}_{ij}))$. We show the property of Y_m as follows.

Lemma 3. *The random variable Y_m has mean $\mathbb{E}\{Y_m\} = b^4 C_0\big(\vartheta_1(\boldsymbol{c}_{ij}) + \frac{1}{2}f(\boldsymbol{c}_{ij})\vartheta_2(\boldsymbol{c}_{ij})\big) + o(b^4)$ and its variance satisfies $\mathbb{V}\{Y_m\}/b^3 \to 0$ as $b \to 0$, where $\vartheta_1(\boldsymbol{c}_{ij}) = \frac{\partial f(\boldsymbol{c}_{ij})}{\partial u_x}\frac{\partial \rho(\boldsymbol{c}_{ij})}{\partial u_x} + \frac{\partial f(\boldsymbol{c}_{ij})}{\partial u_y}\frac{\partial \rho(\boldsymbol{c}_{ij})}{\partial u_y}$ and $\vartheta_2(\boldsymbol{c}_{ij}) = \frac{\partial^2 \rho(\boldsymbol{c}_{ij})}{\partial u_x^2} + \frac{\partial^2 \rho(\boldsymbol{c}_{ij})}{\partial u_y^2}$.*

Proof. Let $\boldsymbol{u} = (\boldsymbol{x}-\boldsymbol{c}_{ij})/b$. We have

$$\begin{aligned}
\mathbb{E}\{Y_m\} &= \int K\Big(\frac{\boldsymbol{x}-\boldsymbol{c}_{ij}}{b}\Big)(\rho(\boldsymbol{x}) - \rho(\boldsymbol{c}_{ij}))f(\boldsymbol{x})d\boldsymbol{x} \\
&= b^2\int K(\boldsymbol{u})(\rho(\boldsymbol{c}_{ij}+b\boldsymbol{u}) - \rho(\boldsymbol{c}_{ij}))f(\boldsymbol{c}_{ij}+b\boldsymbol{u})d\boldsymbol{u}.
\end{aligned} \tag{47}$$

Consider the first order Taylor expansion of $f(\boldsymbol{c}_{ij}+b\boldsymbol{u})$ as (41) and the second order Taylor expansion

$$\rho(\boldsymbol{c}_{ij}+b\boldsymbol{u}) = \rho(\boldsymbol{c}_{ij}) + \nabla\rho(\boldsymbol{c}_{ij})^{\mathrm{T}} b\boldsymbol{u} + \frac{1}{2}b^2\boldsymbol{u}^{\mathrm{T}}\boldsymbol{\Psi}_{ij}\boldsymbol{u} + R_\rho(b\boldsymbol{u}).$$

We have

$$\begin{aligned}
&\int K(\boldsymbol{u})(\rho(\boldsymbol{c}_{ij}+b\boldsymbol{u}) - \rho(\boldsymbol{c}_{ij}))f(\boldsymbol{c}_{ij}+b\boldsymbol{u})d\boldsymbol{u} \\
&= \int K(\boldsymbol{u})\big(\nabla\rho(\boldsymbol{c}_{ij})^{\mathrm{T}} b\boldsymbol{u} + \frac{1}{2}b^2\boldsymbol{u}^{\mathrm{T}}\boldsymbol{\Psi}_{ij}\boldsymbol{u} + R_\rho(b\boldsymbol{u})\big) \\
&\quad \times \left(f(\boldsymbol{c}_{ij}) + \nabla f(\boldsymbol{c}_{ij})^{\mathrm{T}} b\boldsymbol{u} + R_f(b\boldsymbol{u})\right) d\boldsymbol{u}.
\end{aligned} \tag{48}$$

Recall that $K(\boldsymbol{u})$ has a bounded support and the terms $R_f(b\boldsymbol{u})/b$ and $R_\rho(b\boldsymbol{u})/b^2$ uniformly converge to 0 as $b \to 0$ due to the Taylor expansion. Then, the terms involving $R_f(b\boldsymbol{u})$ and $R_\rho(b\boldsymbol{u})$ from expanding the product in (48) are $o(b^2)$ due to Lemma 1. As a result, equation (48) simplifies to

$$\begin{aligned}
&b^2\int K(\boldsymbol{u})\nabla\rho(\boldsymbol{c}_{ij})^{\mathrm{T}}\boldsymbol{u}\nabla f(\boldsymbol{c}_{ij})^{\mathrm{T}}\boldsymbol{u}d\boldsymbol{u} \\
&\quad + \frac{b^2}{2}\int K(\boldsymbol{u})\boldsymbol{u}^{\mathrm{T}}\boldsymbol{\Psi}_{ij}\boldsymbol{u}f(\boldsymbol{c}_{ij})d\boldsymbol{u} \\
&= b^2\int K(\boldsymbol{u})\left(\frac{\partial\rho(\boldsymbol{c}_{ij})}{\partial u_x}u_x + \frac{\partial\rho(\boldsymbol{c}_{ij})}{\partial u_y}u_y\right) \\
&\qquad \times \left(\frac{\partial f(\boldsymbol{c}_{ij})}{\partial u_x}u_x + \frac{\partial f(\boldsymbol{c}_{ij})}{\partial u_y}u_y\right)du_x du_y
\end{aligned}$$

$$
\begin{aligned}
&+\frac{b^2}{2}\int K(\boldsymbol{u})\Big(\frac{\partial^2\rho(\boldsymbol{c}_{ij})}{\partial u_x^2}u_x^2+2\frac{\partial^2\rho(\boldsymbol{c}_{ij})}{\partial u_x\partial u_y}u_xu_y\\
&\qquad+\frac{\partial^2\rho(\boldsymbol{c}_{ij})}{\partial u_y^2}u_y^2\Big)f(\boldsymbol{c}_{ij})du_xdu_y+o(b^2)\\
&=b^2C_0\Bigg\{\frac{\partial f(\boldsymbol{c}_{ij})}{\partial u_x}\frac{\partial\rho(\boldsymbol{c}_{ij})}{\partial u_x}+\frac{\partial f(\boldsymbol{c}_{ij})}{\partial u_y}\frac{\partial\rho(\boldsymbol{c}_{ij})}{\partial u_y}\\
&\qquad+\frac{1}{2}\frac{\partial^2\rho(\boldsymbol{c}_{ij})}{\partial u_x^2}f(\boldsymbol{c}_{ij})+\frac{1}{2}\frac{\partial^2\rho(\boldsymbol{c}_{ij})}{\partial u_y^2}f(\boldsymbol{c}_{ij})\Bigg\}+o(b^2)\\
&=b^2C_0\big(\vartheta_1(\boldsymbol{c}_{ij})+\frac{1}{2}f(\boldsymbol{c}_{ij})\vartheta_2(\boldsymbol{c}_{ij})\big)+o(b^2).
\end{aligned}\tag{49}
$$

Thus, multiply (49) with b^2, we obtain

$$
\mathbb{E}\left\{Y_m\right\}=b^4C_0\big(\vartheta_1(\boldsymbol{c}_{ij})+\frac{1}{2}f(\boldsymbol{c}_{ij})\vartheta_2(\boldsymbol{c}_{ij})\big)+o(b^4).\tag{50}
$$

Similarly,

$$
\begin{aligned}
&\mathbb{V}\left\{Y_m\right\}\\
&=\mathbb{E}\left\{Y_m^2\right\}-\mathbb{E}\left\{Y_m\right\}^2\\
&=\int\Big(K\Big(\frac{\boldsymbol{x}-\boldsymbol{c}_{ij}}{b}\Big)(\rho(\boldsymbol{x})-\rho(\boldsymbol{c}_{ij}))\Big)^2f(\boldsymbol{x})d\boldsymbol{x}-\mathbb{E}\left\{Y_m\right\}^2\\
&=b^2\int\big(K\left(\boldsymbol{u}\right)\left(\nabla\rho(\boldsymbol{c}_{ij})^{\mathrm{T}}b\boldsymbol{u}+\frac{1}{2}b^2\boldsymbol{u}^{\mathrm{T}}\boldsymbol{\Psi}_{ij}\boldsymbol{u}+R_\rho(b\boldsymbol{u})\right)\big)^2\\
&\qquad\times f(\boldsymbol{c}_{ij}+b\boldsymbol{u})d\boldsymbol{u}-\mathbb{E}\left\{Y_m\right\}^2\\
&=b^4\int\big(K\left(\boldsymbol{u}\right)\left(\nabla\rho(\boldsymbol{c}_{ij})^{\mathrm{T}}\boldsymbol{u}+\frac{1}{2}b\boldsymbol{u}^{\mathrm{T}}\boldsymbol{\Psi}_{ij}\boldsymbol{u}+R_\rho(b\boldsymbol{u})\right)/b\big)^2\\
&\qquad\times f(\boldsymbol{c}_{ij}+b\boldsymbol{u})d\boldsymbol{u}-\mathbb{E}\left\{Y_m\right\}^2
\end{aligned}\tag{51}
$$

and thus, $\mathbb{V}\left\{Y_m\right\}/b^3\to 0$, as $b\to 0$. □

Next, it is observed that $\mathrm{Cov}(X_m,Y_m)=0$. This is because that $\mathbb{E}\{X_mY_m\}=\mathbb{E}\{w_m^2(\boldsymbol{c}_{ij})(\rho(\boldsymbol{z}_m)-\rho(\boldsymbol{c}_{ij}))\}\mathbb{E}\{\epsilon_m\}=0$ due to the independence between the zero mean noise ϵ_m and $w_m^2(\boldsymbol{c}_{ij})(\rho(\boldsymbol{z}_m)-\rho(\boldsymbol{c}_{ij}))$. As a result, $\mathrm{Cov}(X_m,Y_m)=\mathbb{E}\{X_mY_m\}-\mathbb{E}\{X_m\}\mathbb{E}\{Y_m\}=0$ due to $\mathbb{E}\{X_m\}=0$ from Lemma 2. Therefore, $\mathbb{V}\{X_m+Y_m\}=\mathbb{V}\{X_m\}+\mathbb{V}\{Y_m\}$.

In addition, from (37) and (38), it is observed that $\varepsilon_1(\boldsymbol{c}_{ij})+\varepsilon_2(\boldsymbol{c}_{ij})=\frac{1}{b^2}\frac{1}{M}\sum_{m=1}^{M}(X_m+Y_m)$ where the variable $Z_m\triangleq X_m+Y_m$ is i.i.d., since both $\boldsymbol{z}_m$ and ϵ_m are i.i.d.. As a result, the law of large number implies $b^2(\varepsilon_1(\boldsymbol{c}_{ij})+\varepsilon_2(\boldsymbol{c}_{ij}))\xrightarrow{p}\mathbb{E}\{Z_m\}$ as $M\to\infty$.

In addition, the central limit theorem yields

$$
\sqrt{M}\Big(b^2(\varepsilon_1(\boldsymbol{c}_{ij})+\varepsilon_2(\boldsymbol{c}_{ij}))-\mathbb{E}\{Z_m\}\Big)\xrightarrow{d}\mathcal{N}(0,\mathbb{V}\{Z_m\})
$$

as $M\to\infty$, where $\mathbb{E}\{Z_m\}=\mathbb{E}\{X_m\}+\mathbb{E}\{Y_m\}$ and $\mathbb{V}\{Z_m\}=\mathbb{V}\{X_m+Y_m\}=\mathbb{V}\{X_m\}+\mathbb{V}\{Y_m\}$ as shown before.

Finally, it is known that $\hat{f}(\boldsymbol{z})\xrightarrow{p}f(\boldsymbol{z})$ for every $\boldsymbol{z}$ [50]. Then, using Slutsky's theorem [51], the mean

$$
\frac{b^2(\varepsilon_1(\boldsymbol{c}_{ij})+\varepsilon_2(\boldsymbol{c}_{ij}))}{\hat{f}(\boldsymbol{c}_{ij})}\xrightarrow{p}\frac{\mathbb{E}\{Z_m\}}{f(\boldsymbol{c}_{ij})}\tag{52}
$$

for every $\boldsymbol{c}_{ij}$. Then, substituting $(\varepsilon_1(\boldsymbol{c}_{ij})+\varepsilon_2(\boldsymbol{c}_{ij}))/\hat{f}(\boldsymbol{c}_{ij})$ with ξ_{ij} as in (39) and $\mathbb{E}\{Z_m\}$ with the mean of X_m and Y_m in Lemmas 2 and 3, the result (18) in Theorem 3 is obtained.

In addition, defining $\bar{\xi}_{ij}=\xi_{ij}-\mathbb{E}\{\xi_{ij}\}$, it follows that

$$
\sqrt{Mb^2}\bar{\xi}_{ij}\xrightarrow{d}\frac{\mathcal{N}\Big(0,\mathbb{V}\{Z_m\}\Big)}{f(\boldsymbol{c}_{ij})}\tag{53}
$$

for every $\boldsymbol{c}_{ij}$. Substituting the variance of X_m and Y_m in Lemmas 2 and 3 to $\mathbb{V}\{Z_m\}$ in (53), the result (19) in Theorem 3 is obtained.

REFERENCES

[1] D. Zhu, X. Peng, and G. Li, "A matrix completion based method for RFI source localization in microwave interferometric radiometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 9, pp. 7588–7602, 2021.

[2] 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.

[3] D. Romero, S.-J. Kim, G. B. Giannakis, and R. López-Valcarce, "Learning power spectrum maps from quantized power measurements," *IEEE Trans. Signal Process.*, vol. 65, no. 10, pp. 2547–2560, 2017.

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

[5] S. Chouvardas, S. Valentin, M. Draief, and M. Leconte, "A method to reconstruct coverage loss maps based on matrix completion and adaptive sampling," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 2016, pp. 6390–6394.

[6] B. Fuchs, L. Le Coq, and M. D. Migliore, "On the interpolation of electromagnetic near field without prior knowledge of the radiating source," *IEEE Trans. Antennas Propag.*, vol. 65, no. 7, pp. 3568–3574, 2017.

[7] A. C. M. Austin and M. J. Neve, "Efficient field reconstruction using compressive sensing," *IEEE Trans. Antennas Propag.*, vol. 66, no. 3, pp. 1624–1627, 2018.

[8] S. G. Sathyanarayana, Z. Wang, N. Sun, B. Ning, S. Hu, and J. A. Hossack, "Recovery of blood flow from undersampled photoacoustic microscopy data using sparse modeling," *IEEE Trans. Med. Imag.*, vol. 41, no. 1, pp. 103–120, 2022.

[9] Z. Ke, W. Huang, Z.-X. Cui, J. Cheng, S. Jia, H. Wang, X. Liu, H. Zheng, L. Ying, Y. Zhu, and D. Liang, "Learned low-rank priors in dynamic MR imaging," *IEEE Trans. Med. Imag.*, vol. 40, no. 12, pp. 3698–3710, 2021.

[10] J. Chen, T. Li, J. Wang, and C. W. d. Silva, "WSN sampling optimization for signal reconstruction using spatiotemporal autoencoder," *IEEE Sensors Journal*, vol. 20, no. 23, pp. 14 290–14 301, 2020.

[11] Z. Liu, P. Wu, S. Duan, W. Zhan, X. Ma, and Y. Wu, "Spatiotemporal reconstruction of land surface temperature derived from fengyun geostationary satellite data," *IEEE J.. Sel. Topics Appl. Earth Observ.*, vol. 10, no. 10, pp. 4531–4543, 2017.

[12] 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.

[13] A. Yakhot, T. Anor, and G. E. Karniadakis, "A reconstruction method for gappy and noisy arterial flow data," *IEEE Trans. Med. Imag.*, vol. 26, no. 12, pp. 1681–1697, 2007.

[14] Y. Teganya, D. Romero, L. M. L. Ramos, and B. Beferull-Lozano, "Location-free spectrum cartography," *IEEE Trans. Signal Process.*, vol. 67, no. 15, pp. 4013–4026, 2019.

[15] J. A. Bazerque and G. B. Giannakis, "Nonparametric basis pursuit via sparse kernel-based learning: A unifying view with advances in blind methods," *IEEE Signal Process. Mag.*, vol. 30, no. 4, pp. 112–125, 2013.

[16] Y.-Q. Xu, B. Zhang, X. Zhang, J. Hu, and D. Guo, "Radio environment map construction with Gaussian process and kernel transformation," in *in Proc. IEEE Int. Conf. Commun., Image and Signal Processing*, 2021, pp. 350–355.

[17] 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.

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

[19] A. Balachandrasekaran, A. L. Cohen, O. Afacan, S. K. Warfield, and A. Gholipour, "Reducing the effects of motion artifacts in fMRI: A structured matrix completion approach," *IEEE Trans. Med. Imag.*, vol. 41, no. 1, pp. 172–185, 2022.

[20] 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.

[21] 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.

[22] 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.

[23] X. Wang, X. Wang, S. Mao, J. Zhang, S. C. Periaswamy, and J. Patton, "Deepmap: Deep Gaussian process for indoor radio map construction and location estimation," in *Proc. IEEE Global Telecomm. Conf.* IEEE, 2018, pp. 1–7.

[24] Y. Teganya and D. Romero, "Deep completion autoencoders for radio map estimation," *IEEE Trans. on Wireless Commun.*, 2021.

[25] R. Levie, Ç. Yapar, G. Kutyniok, and G. Caire, "RadioUNet: Fast radio map estimation with convolutional neural networks," *IEEE Trans. on Wireless Commun.*, vol. 20, no. 6, pp. 4001–4015, 2021.

[26] 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.

[27] 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.

[28] J. Wang, P. Urriza, Y. Han, and D. Cabric, "Weighted centroid localization algorithm: theoretical analysis and distributed implementation," *IEEE Trans. Wireless Commun.*, vol. 10, no. 10, pp. 3403–3413, 2011.

[29] E. J. Candes and Y. Plan, "Matrix completion with noise," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 925–936, 2010.

[30] G. Tang and A. Nehorai, "Lower bounds on the mean-squared error of low-rank matrix reconstruction," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4559–4571, 2011.

[31] D. Zachariah, M. Sundin, M. Jansson, and S. Chatterjee, "Alternating least-squares for low-rank matrix reconstruction," *IEEE Signal Process. Lett.*, vol. 19, no. 4, pp. 231–234, Apr 2012.

[32] S. Foucart and H. Rauhut, "An invitation to compressive sensing," in *A mathematical introduction to compressive sensing*. Springer, 2013, pp. 1–39.

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

[34] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[35] D. Ruppert and M. P. Wand, "Multivariate locally weighted least squares regression," *The annals of statistics*, vol. 22, no. 3, pp. 1346–1370, 1994.

[36] A. Vehtari, A. Gelman, and J. Gabry, "Practical bayesian model evaluation using leave-one-out cross-validation and waic," *Statistics and computing*, vol. 27, no. 5, pp. 1413–1432, 2017.

[37] L. Brekhovskikh and Y. P. Lysanov, *Fundamentals of ocean acoustics*. Acoustical Society of America, 2004.

[38] E. Candes and B. Recht, "Exact matrix completion via convex optimization," *Commun. of the ACM*, vol. 55, no. 6, pp. 111–119, 2012.

[39] D. L. Pimentel-Alarcón, N. Boston, and R. D. Nowak, "A characterization of deterministic sampling patterns for low-rank matrix completion," *IEEE J. Sel. Top. Signal Process.*, vol. 10, no. 4, pp. 623–636, 2016.

[40] Y. Chen, Y. Chi, J. Fan, C. Ma, and Y. Yan, "Noisy matrix completion: Understanding statistical guarantees for convex relaxation via nonconvex optimization," *SIAM journal on optimization*, vol. 30, no. 4, pp. 3098–3121, 2020.

[41] 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.

[42] P. Jain and P. Netrapalli, "Fast exact matrix completion with finite samples," in *Conference on Learning Theory*. PMLR, 2015, pp. 1007–1034.

[43] B. V. Srinivasan, R. Duraiswami, and R. Murtugudde, "Efficient kriging for real-time spatio-temporal interpolation," in *Proceedings of the 20th Conference on Probability and Statistics in the Atmospheric Sciences*. American Meteorological Society Atlanta GA, 2010, pp. 228–235.

[45] V.-P. Chowdappa, C. Botella, J. J. Samper-Zapater, and R. J. Martinez, "Distributed radio map reconstruction for 5G automotive," *IEEE Intell. Transp. Syst. Mag.*, vol. 10, no. 2, pp. 36–49, 2018.

[44] M. Stojanovic and J. Preisig, "Underwater acoustic communication channels: Propagation models and statistical characterization," *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 84–89, 2009.

[46] A. Verdin, C. Funk, B. Rajagopalan, and W. Kleiber, "Kriging and local polynomial methods for blending satellite-derived and gauge precipitation estimates to support hydrologic early warning systems," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 2552–2562, 2016.

[47] 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.

[48] J. Talvitie, M. Renfors, and E. S. Lohan, "Distance-based interpolation and extrapolation methods for RSS-based localization with indoor wireless signals," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1340–1353, 2015.

[49] 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.

[50] E. Parzen, "On estimation of a probability density function and mode," *The annals of mathematical statistics*, vol. 33, no. 3, pp. 1065–1076, 1962.

[51] A. DasGupta, *Asymptotic theory of statistics and probability*. Springer, 2008, vol. 180.

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 toward the Ph.D. degree 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. His research interests mainly include matrix completion and tensor decomposition with application to radio map reconstruction and source localization.

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.