

Sparse Bayesian Learning-Based Hierarchical Construction for 3D Radio Environment Maps Incorporating Channel Shadowing

Jie Wang, *Graduate Student Member, IEEE*, Qiuming Zhu, *Senior Member, IEEE*, Zhipeng Lin, *Member, IEEE*, Junting Chen, *Member, IEEE*, Guoru Ding, Qihui Wu, *Fellow, IEEE*, Guochen Gu, Qianhao Gao

Abstract—The radio environment map (REM) visually displays the spectrum information over the geographical map and plays a significant role in monitoring, management, and security of spectrum resources. In this paper, we present an efficient 3D REM construction scheme based on the sparse Bayesian learning (SBL), which aims to recover the accurate REM with limited and optimized sampling data. In order to reduce the number of sampling sensors, an efficient sparse sampling method for unknown scenarios is proposed. For the given construction accuracy and the priority of each location, the quantity and sampling locations can be jointly optimized. With the sparse sampled data, by mining the sparsity of the spectrum situation and channel propagation characteristics, a SBL-based spectrum data hierarchical recovery algorithm is developed to estimate the missing data of unsampled locations. Finally, the simulated three-dimensional (3D) REM data in the campus scenario are used to verify the proposed methods as well as to compare with the state-of-the-art. We also analyze the recovery performance and the impact of different parameters on the constructed REMs. Numerical results demonstrate that the proposed scheme can ensure the construction accuracy and improve the computational efficiency under the low sampling rate.

Index Terms—3D radio environment map, sparse Bayesian learning, sampling optimization, propagation channel model.

I. INTRODUCTION

A. Background and Motivation

WITH the rapid development of radio frequency (RF) communication technologies, various electronic devices, i.e., radio, radar, navigation, and so on, have formed a dynamic, complex, and multidimensional electromagnetic

This work was supported in part by the Jiangsu Province Frontier Leading Technology Basic Research Project under Grant BK 20222013, in part by the National Natural Science Foundation of China under Grant No. 62271250, in part by the Key Technologies R&D Program of Jiangsu (Prospective and Key Technologies for Industry) under Grants BE2022067, BE2022067-1 and BE2022067-3, in part by the Industry-university-research cooperation Fund of Eighth Research Institute, China Aerospace Science and Technology Corporation Limited Grant SAST2023-023, and in part by Practice Innovation Program of Jiangsu Province under Grant KYCX23_0381. (*Corresponding author: Qiuming Zhu*)

J. Wang, Q. Zhu, Z. Lin, Q. Wu, G. Gu and Q. Gao are with The Key Laboratory of Dynamic Cognitive System of Electromagnetic Spectrum Space, College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China (e-mail: bx2104906wangjie, zhuqiuming, linlp@nuaa.edu.cn, wuqihui2014@sina.com, guguchen, qianhao@nuaa.edu.cn).

J. Chen is 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 (e-mail: juntingc@cuhk.edu.cn).

G. Ding is with the College of Communications Engineering, Army Engineering University, Nanjing 210007, China (e-mail: guoru_ding@yeah.net).

environment [1]–[5]. Nevertheless, a large percentage of assigned spectrum remains underutilized as reported by the federal communications commission (FCC) [6]. Consequently, the dynamic spectrum access technology, i.e., cognitive radio (CR), was introduced to improve communication efficiency based on the awareness of spectrum situation [2], [7], [8]. The defense advanced research projects agency (DARPA) RadioMap program further proposed to add the spectrum awareness results on a geographical map, namely radio environment map (REM) or spectrum cartography (SC) [9]. In a word, the REM can visualize specific spectrum information, i.e., received signal strength (RSS), channel gain, power spectrum density, and radiation location, in a region of interest (ROI) at different dimensions and scales from limited observations. It is essential in dynamic spectrum access, abnormal spectrum monitoring, spectrum sharing, radiation localization, interference coordination, etc [10], [11].

Majority of SC methods focused on the two-dimensional (2D) REM construction, which usually treated it as solving a matrix completion problem or image reconstruction task based on sufficient sampled data [12]. However, with the development of space-air-ground integrated networks, the spectrum situation has been extended from 2D space to three-dimensional (3D) space [13], [14]. For example, the aerial platforms, i.e., unmanned aerial vehicle (UAV), airship, and hot air balloon, have become an important part of RF transmitters [15]. By leveraging the location flexibility of these RF devices, it is able to improve spectrum utilization [16]. Furthermore, 3D REMs provide valuable insights into the impact of these devices on the spectrum usage, which is vital for spectrum management. An illustrative example is a user handover algorithm proposed in [17], which leverages the 3D REM to predict the most favorable network connection.

Due to the limitation of sensor number and sampling time in the wide space, it is a big challenge to achieve accurate reconstruction with sparse and noisy data [18]. To address this problem, compressed sensing (CS) technology has been extensively adopted since it requires less data than the Nyquist criterion [19], [20]. By mining the underlying sparsity of spectrum situation, the CS constructs the REM by decomposing the sampling data into the linear superposition of sensing matrix and sparse signal. Nevertheless, the sensing matrix exhibits high spatial correlation and the recovery performance could be greatly deteriorated when dealing with the noisy data. Sparse Bayesian learning (SBL) can recover the exact sparse

signal under the high correlated sensing matrix and has good performance of anti-noise [21].

Moreover, the measurements are spatially correlated, and thus the sampling locations have a great impact on the accuracy and efficiency of construction. By selecting specific sampling locations, it is possible to recover the 3D REM with less data requirements. Besides, a poor propagation model may also lead to significant construction errors [22], [23]. In the specific scenarios, i.e., dense urban areas with numerous buildings, it is difficult to build an accurate channel model including the factors of multibuilding, shadowing, and antennas pattern.

To sum up, the sampling data in 3D space is usually sparse. The existing 2D construction methods require a large amount of data and are not suitable for 3D REM. Besides, the free-space propagation model is usually adopted, which is not consistent with the realistic propagation characteristic and may reduce the construction performance seriously. According to the above observations, we are motivated to exploit the inherent sparsity of spectrum situation, the radio propagation characteristics, and sampling position optimization for constructing accurate 3D REMs in unknown environments.

B. Related Work

Existing REM construction methods can be mainly divided into data-driven and model-driven ones [14], [24]–[38]. The former typically employs the spatial interpolation and machine learning (ML) techniques to recover the missing data at the unsampled positions by the available sampling data. In [33], the authors compared four typical spatial interpolation techniques, i.e., Nearest Neighbour (NN), inverse distance weighting (IDW), triangular irregular network, and Kriging algorithm [34]. The authors also mentioned that Kriging was the most accurate spatial interpolation method while IDW was robust. In [35], a method is proposed to construct the REMs with thin plate spline interpolation when no information on the radiation source is available. The authors in [36] proposed a tensor completion method to recover the incomplete spectrum data in both the spatial and temporal domains.

Recently, ML techniques recast data complementation as a learning-based optimization problem and have also been applied in REM construction. For example, a generative adversarial network-based method was proposed in [26] to recover the missing spectrum data from the simulated or previously collected training data. The authors in [25] utilized deep neural networks (DNNs) to learn the intricate underlying structure from the sampled data. In [39], the authors further proposed a DNN-based active sensing method with UAVs. However, these methods require a large amount of sampling data, in order to achieve satisfactory performance. Besides, the construction performance are usually sensitive to the sampling position distribution and measurement error [18].

The model-driven methods combine the sampling data with active transmitter information and channel propagation models to construct REMs, which are less sensitive to the measurement error as well as the number of sampling data. For example, the authors in [27] firstly estimated the transmitter

information by the sampled data and then completed the missing data with the help of propagation model. CS is a typical model-driven construction method, which can track the dynamic network topology and construct the REM with sparse sampled data. In [19], CS multispectral cartography was proposed for spectrum sensing. The least absolute shrinkage and selection operator was used in [20] to construct the REM with random sampling data. In [14], the authors proposed a compressed REM mapping method based on the improved orthogonal matching pursuit algorithm. However, both of them are easy to be affected by the measurement noise and high-correlation sensing matrix. Besides, they only focused on the REM construction under line of sight (LOS) conditions without considering the effect of non line of sight (NLOS) propagation characteristic.

The sparse Bayesian learning (SBL) can recover the sparse signal under the high correlated sensing matrix. The authors in [37] proposed an SBL-based 2D REM construction method. It considers the free space channel model and random sampling scheme. Besides, it was assumed that the RSS was only affected by a single transmitter. In [38], a scenario-dependent SBL method was proposed to construct 3D REM. It considered the realistic propagation channel by using the ray tracing (RT) technology on the known scenario. The authors also developed a sampling location optimization algorithm. In [40], a Bayesian compressive crowdsensing framework was applied to address the crowdsourced 2D REM construction.

In a nutshell, most of existing studies only focus on the 2D REM construction and the measurement positions or matrixes are either random or fixed without considering the sampling optimization. Furthermore, they are generally based on the free-space channel model or RT-based model. The free-space channel model is only applicable to flat suburban. The RT-based model not only requires detailed information of scenarios but also is computationally. To the best of our knowledge, the channel fading characteristic, i.e., shadow fading, is not considered in these studies.

C. Contributions

To fill these gaps, this paper proposes a novel SBL-based 3D REM hierarchical construction scheme. It consists of the sampling location optimization and the spectrum data recovery with consideration of shadow fading. The main novelties and contributions of this paper are summarized as follows.

- An efficient SBL-based hierarchical construction model for 3D REM in the unknown spatial heterogeneous environment is proposed. Different with previous methods, the proposed method considers the factors of sampling location optimization and channel propagation characteristic including shadow fading, which can improve the construction performance with the limitation of sampling data and time.
- A worst case SBL variance-based measurement matrix optimization method is proposed. By introducing the minimum SBL variance indicator and solving a sparse optimization problem, the optimized sampling locations and quantity are obtained under the given threshold of

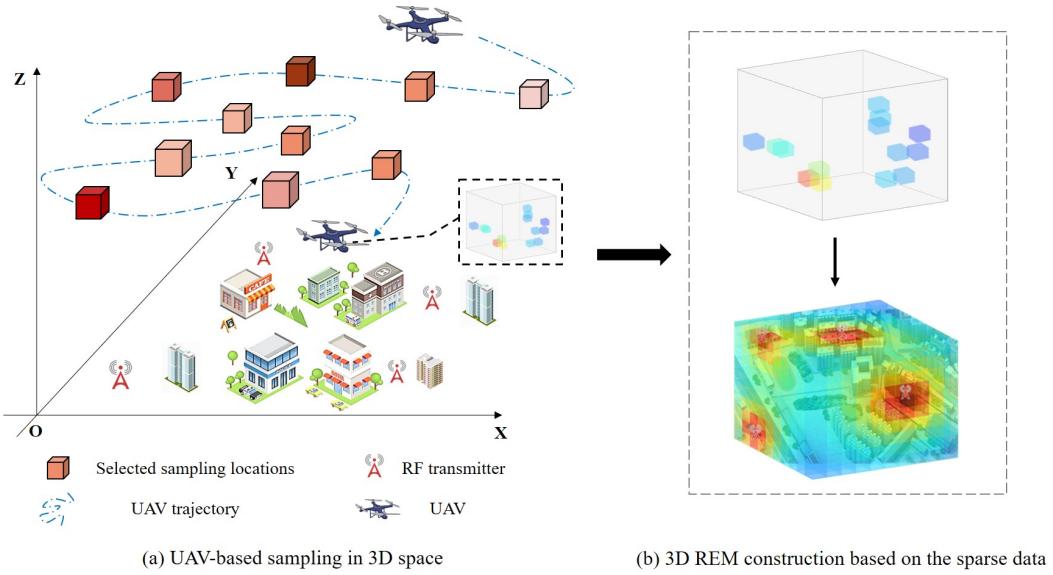


Fig. 1. Overview of UAV-based 3D REM sampling and construction.

sparse signal recovery accuracy. The proposed method outperforms other sampling methods and improves the sampling efficiency while ensuring the recovery accuracy.

- A spectrum data hierarchical recovery algorithm under unknown scenarios is developed based on the Bayesian theory. By mining the sparsity characteristics in the 3D spectrum space, the sparse signal is firstly recovered based on the SBL. In order to obtain the whole 3D REM, channel dictionary is then optimized by considering the channel characteristic and shadowing based on Gaussian process. The proposed algorithm exhibits excellent recovery performance against the traditional algorithms especially under unknown scenarios.

The rest of this paper is organized as follows. Section II gives the proposed 3D REM construction model. In Section III, the details of proposed construction scheme are given and demonstrated. Then, Section IV presents the simulation and comparison results and Section V gives some conclusions.

II. PRELIMINARIES

A. Proposed CS-Based 3D REM Construction

The 3D region of interest (ROI) is firstly discretized into several small cubes. Thus, it can be described by a 3D spectrum tensor $\chi \in \mathbb{R}^{N_x \times N_y \times N_z}$, where N_x , N_y , and N_z indicate the grid number along x , y , and z dimensions, respectively. Each element or cube is colored according to its RSS value. Technically, 3D REM construction aims to recover all RSS values, i.e., $N = N_x \times N_y \times N_z$ cubes, based on the known ones of sampled cubes. The locations of the all cubes are denoted as $\{\nu_n\}_{n=1}^N$, where $\nu_n = (x_n^v, y_n^v, z_n^v)$. Due to the low cost, mobility, and flexibility movement, an UAV equipped with a spectrum monitoring device can be used to sample RSS values in the ROI as shown in Fig. 1. Here, RSS means the received power strength. Note that this paper only focuses on optimizing sampling points. The trajectory planning is not

considered. In practical applications, optimization algorithms such as the shortest path planning can be further utilized to obtain the UAV trajectory for data acquisition.

Let us set a sparse signal vector $\omega = [\omega_1, \omega_2, \dots, \omega_n, \dots, \omega_N]^T \in \mathbb{R}^{N \times 1}$ to denote the unknown transmitting information at every cube as

$$\omega_n = \begin{cases} P_n^t, & \text{if there is a RF transmitter in the } n\text{th cube,} \\ 0, & \text{else,} \end{cases} \quad (1)$$

where P_n^t is the transmitting power of RF transmitter if there is a RF transmitter in the n th cube. It is assumed that there are K stationary RF transmitters in the 3D ROI denoted by $\{\mathbf{u}_k\}_{k=1}^K$, where $\mathbf{u}_k = (x_k^u, y_k^u, z_k^u)$ is the location of k th RF transmitter. Compared with the cube number in the 3D ROI, the number of RF transmitter is much smaller, i.e., K ($K \ll N$), which means that ω has stronger sparsity than in 2D ROI. Therefore, ω is a K -sparse signal vector with $\|\omega\|_0 = K$.

In the realistic 3D environments, the RSS in each cube is affected by RF transmitter positions and distribution, transmitting power, and channel propagation characteristic. Especially, the channel characteristic mainly considers the channel gain, which includes path loss, shadow fading, small-scale fading, and noise [41]. Path loss is a deterministic component determined by frequency and distance between the transmitter and receiver. Small-scale fading and noise can be averaged out with long-term observation, while shadow fading is related with specific scenarios [42]. Furthermore, the shadow effects are significant in the 3D scenarios due to the multiple propagation and obstruction by buildings. It is crucial to take these effects into account in the propagation model.

Based on the above analysis, the spectrum tensor χ is vectorized into $\mathbf{x} \in \mathbb{R}^{N \times 1}$, as

$$\mathbf{x} = \xi^v \circ (\varphi \omega), \quad (2)$$

where $\varphi \in \mathbb{R}^{N \times N}$ is path loss matrix and $\varphi_{i,j}$ is defined as the path loss from the i th cube to the j th cube. $\xi^v \in \mathbb{R}^{N \times 1}$ denote

the shadow fading components of all cubes, and \circ denote the element-wise (Hadamard) product. We define the term $\xi^v \circ \varphi$ as the channel propagation dictionary matrix. Note that the shadow fading is estimated based on the sampled data.

Suppose that we select M cubes as samples, denoted by $\{\mathbf{s}_m\}_{m=1}^M$, from all N cubes and $\mathbf{s}_m = (x_m^s, y_m^s, z_m^s)$ is the location of m th cube. Thus, the sampling rate is $r = M/N$. The sampled RSSs can be expressed by a vector $\tilde{\mathbf{t}} \in \mathbb{R}^{M \times 1}$. All sampled locations can be represented by a measurement matrix $\psi \in \mathbb{R}^{M \times N}$ as

$$\psi_{i,j} = \begin{cases} 1, & \text{if the } i\text{th sample is at the } j\text{th cube,} \\ 0, & \text{else,} \end{cases} \quad (3)$$

where each row of ψ has a nonzero element denoting the sampling location. Note that the strategy for 2D sampling is relatively straightforward. Uniform sampling or random sampling methods are usually adopted. As the signal varies in the 3D space and the spatial correlation is strong, the measurement matrix ψ should be carefully designed to ensure efficient sampling.

The averaged RSS vector \mathbf{t} can be expressed as

$$\begin{aligned} \mathbf{t} &= \psi(\xi^v \circ \varphi \omega) + \boldsymbol{\varepsilon} = \psi \xi^v \circ (\psi \varphi \omega) + \boldsymbol{\varepsilon} \\ &= \xi^s \circ (\Phi \omega) + \boldsymbol{\varepsilon} = \Phi \omega + \boldsymbol{\varepsilon}^*, \end{aligned} \quad (4)$$

with

$$t_m = \xi_m^s \sum_{n=1}^N \omega_n \Phi_{m,n} + \varepsilon_m, \quad (5)$$

$$\varepsilon_m^* = (\xi_m^s - 1) \sum_{n=1}^N (\Phi_{m,n} \omega_n) + \varepsilon_m, \quad m = 1, 2, \dots, M, \quad (6)$$

where $\boldsymbol{\varepsilon} \in \mathbb{R}^{M \times 1}$ is the measurement noise vector, and Φ is the sensing matrix. ξ_m^s is the shadow fading value at the sample whose logarithmic form follows normal distribution $\mathcal{N}(0, \sigma^2)$. Then, the $\xi_m^s - 1$ can be derived as

$$\xi_m^s - 1 = \exp \left\{ \frac{\ln 10}{10} \bar{\xi}_m^s \right\} - 1 \sim \frac{\ln 10}{10} \bar{\xi}_m^s, \quad (7)$$

with

$$\bar{\xi}_m^s = 10 \log_{10} \xi_m^s. \quad (8)$$

We assume that the marginal standard deviation is small enough, then the characteristic of the original multiplicative noise is very similar to that of the additive Gaussian noise [43]. Accordingly, we approximate ε_m^* as a Gaussian distribution with variance σ_0^2 since $\bar{\xi}_m^s$ follows $\mathcal{N}(0, \sigma^2)$.

In this paper, the objective of REM construction is to recover sparse signal $\hat{\omega}$ and the shadow fading components $\xi^v \in \mathbb{R}^{N \times 1}$ of all cubes to recover \mathbf{x} . Then, the 3D spectrum tensor χ is obtained by mapping \mathbf{x} to 3D space, i.e., for a $N \times 1$ dimensional vector \mathbf{x} , the matrixing operation obtains a tensor χ of $N_x \times N_y \times N_z$, where $N = N_x \times N_y \times N_z$ and χ includes all elements of \mathbf{x} .

B. SBL-Based Hierarchical REM Data Recovery

1) SBL-based sparse signal recovery

The matrix φ denotes the propagation channel characteristic, which is usually modelled by an exponential attenuation

model with respect to the distance [20]. The element of φ can be expressed by path loss function $f(\nu_i, \nu_j)$ as

$$\varphi_{i,j} = f(\nu_i, \nu_j) = \begin{cases} 1, & \text{if } d_{i,j} \leq d_0, \\ \frac{G_t G_r c_0^2}{4\pi(f_c)^2} \left(\frac{d_0}{d_{i,j}} \right)^\eta, & \text{otherwise,} \end{cases} \quad (9)$$

where G_t and G_r are the antenna gains of transceivers, respectively, c_0 is the light speed, f_c is the carrier frequency, η is the path loss exponent, $d_{i,j} = \|\nu_i - \nu_j\|_2$ is the distance between ν_i and ν_j , and d_0 is the reference distance. Similarly, $f(\mathbf{s}_m, \nu_n)$ constructs the sensing matrix Φ between samples and each cube, i.e., $\Phi_{m,n} = f(\mathbf{s}_m, \nu_n)$.

Based on the sparsity characteristic of ω , we can recover the sparse signal ω by

$$\begin{aligned} \hat{\omega} &= \arg \min \|\omega\|_1, \\ \text{s.t. } \mathbf{t} &= \psi \varphi \omega + \boldsymbol{\varepsilon}^*. \end{aligned} \quad (10)$$

The SBL approach is adopted in this paper to solve this problem, since it has good performance even with the high correlation of Φ [21].

Firstly, we introduce the probability model in SBL. The i.i.d. Gaussian noise is assumed in (4), which is denoted by $\boldsymbol{\varepsilon}^*$ with unknown variance σ_0^2 . It yields the Gaussian likelihood of \mathbf{t}

$$p(\mathbf{t} | \omega, \sigma_0^2) = (2\pi\sigma_0^2)^{-M/2} \exp \left\{ -\frac{\|\mathbf{t} - \Phi \omega\|^2}{2\sigma_0^2} \right\}. \quad (11)$$

A Gamma distribution is then posed on β ($\beta = (\sigma_0^2)^{-1}$)

$$p(\beta | c_0, d_0) = \Gamma(\beta | c_0, d_0), \quad (12)$$

with

$$\Gamma(\beta | c_0, d_0) = \Gamma(c_0)^{-1} d_0^{c_0} \beta^{c_0-1} e^{-d_0\beta}, \quad (13)$$

where $c_0 \geq 0$ and $d_0 \geq 0$ are the shape parameter and the scale parameter, respectively, $\Gamma(\cdot)$ is the Gamma function $\Gamma(c_0) = \int_0^\infty t^{c_0-1} e^{-t} dt$.

As a CS recovery algorithm based on Bayesian learning, SBL assigns a sparseness promoting prior to ω , i.e., a two-layer hierarchical sparse prior [21]. In the first layer, each element of ω is posed a zero-mean Gaussian prior as

$$p(\omega | \alpha) = \prod_{n=1}^N \mathcal{N}(\omega_n | 0, \alpha_n^{-1}), \quad (14)$$

where $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_N]^\top$. Then, in the second layer, a Gamma hyperprior over α is considered as

$$p(\alpha | a_0, b_0) = \prod_{n=1}^N \Gamma(\alpha_n | a_0, b_0). \quad (15)$$

The overall prior $p(\omega)$ can be obtained by computing the marginal integral of hyper-parameters in α as

$$p(\omega) = \int p(\omega | \alpha) p(\alpha) d\alpha, \quad (16)$$

$$\begin{aligned} p(\omega_n) &= \int p(\omega_n | \alpha_n) p(\alpha_n) d\alpha_n \\ &= \frac{b_0^{a_0} \Gamma(a_0 + 1/2)}{\sqrt{2\pi} \Gamma(a_0)} \left(b_0 + \omega_n^2 / 2 \right)^{-\left(a_0 + 1/2 \right)}. \end{aligned} \quad (17)$$

Since the integral is computable for Gamma $p(\alpha)$, the true prior $p(\omega)$ in (17) is a Student-t distribution.

Following the Bayesian inference, the estimation of $\hat{\omega}$ requires the calculation of the weight posterior of ω , which is Gaussian distribution

$$\begin{aligned} p(\omega|t, \alpha, \beta) &= \frac{p(t|\omega, \alpha, \beta)p(\omega|\alpha)}{p(t|\alpha, \beta)}, \\ &= \mathcal{N}(\omega|\mu, \Sigma_\omega), \end{aligned} \quad (18)$$

with

$$\mu = \beta \Sigma_\omega \Phi^T t, \quad (19)$$

$$\Sigma_\omega = (\beta \Phi^T \Phi + \mathcal{A})^{-1}, \quad (20)$$

where $\mathcal{A} = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_N)$. By calculating hyperparameters α and β , we can estimate the $\hat{\omega}$ with the mean μ and evaluate the recovery accuracy by the variance Σ_ω . The detailed derivations of (19) and (20) can be found in Appendix A.

2) GP-based channel dictionary optimization with shadow fading

Generally, correlated shadow fading has a very common basic statistical feature, i.e., nearby observations tend to be more similar than distant observations. Then, it is reasonable to assume that the nearby samples have high correlated shadow fading.

To model spatial phenomena, a common approach is to use Gaussian Processes (GP) in the domain of spatial statistics [44]. GP is a random process in a continuous domain, i.e., time and space. In a GP, each point in the continuous input space is associated with a normally distributed random variable. The GP effectively model the signal strength distribution. The unknown location was estimated by maximizing the joint likelihood of RSS with respect to spatial coordinates. Furthermore, Gaussian Process Regression (GPR) is utilized to model the intricate correlation between signal strength and location within various positioning systems [45]. According to [45]–[47], GP exhibits the ability to quantify uncertainty in RSS data across a continuous spatial domain, in line with Bayesian nonparametric models. Consequently, the GP serves as a valuable tool for regressing the relationship between RSS values and their corresponding spatial coordinates.

Accordingly, the spatial correlated shadow fading components of M samples obey M -dimensional Gaussian distributions and we model the uncertainty of shadow fading as a GP

$$\xi^s \sim \mathcal{GP}(\mathbf{0}, \mathcal{C}(s, s)), \quad (21)$$

where the mean vector is $\mathbf{0}$ and the covariance matrix $\mathcal{C}(s, s)$ is determined by covariance function or kernel function. To guarantee the positive-definite of the covariance matrix, we choose Matérn covariance function to capture the spatial correlations of shadow fading. The Matérn covariance function is a generalization of radial basis function (RBF), which matches the physical process by allowing the random field to have great flexibility and smoothness with manageable parameter number [48], as

$$\mathcal{C}(d) = \sigma^2 \frac{2^{1-g}}{\Gamma(g)} \left(\sqrt{2g} \frac{d}{\rho} \right)^g K_g \left(\sqrt{2g} \frac{d}{\rho} \right), \quad (22)$$

where $K_g(\cdot)$ is the Bessel function of the second kind. g and ρ are the order and the non-negative spatial decay parameter of covariance, d is the distance between samples, σ^2 is the marginal standard deviation controls the expected variation in the output. Note that the shadow fading is relatively smooth. We have conducted massive simulations and find that $g = 3/2$ is preferable. The similar results can also be found in [48]. The spatial correlations of shadow fading capture the scenario characteristics, which can be utilized to mine the information in the sensing data for the inference of unsampled cubes.

Since the sparse signal $\hat{\omega}$ is firstly estimated by (19), the RSS denoting path loss component can be computed for any location s_m as

$$t(s_m) = \sum_{n=1}^N \hat{\omega}_n \Phi_{m,n}. \quad (23)$$

The shadowing effect at $\{s_m\}_{m=1}^M$ (we denote it as s for the sake of presentation) is defined as

$$\bar{\xi}^s = 10 \log 10 \left(\frac{t}{t(s)} \right). \quad (24)$$

Note that the logarithmic form $\bar{\xi}^s$ is a GP and the shadow fading components $\bar{\xi}^{v^*}$ of unsampled cubes can be further inferred by Gaussian process regression (GPR) in Section III-C.

III. SBL BASED 3D REM HIERARCHICAL CONSTRUCTION

A. An Overview of 3D REM Construction

To address the spectrum data recovery problem, this section presents an SBL-based 3D REM hierarchical construction scheme. It introduces a worst case SBL variance-based measurement matrix optimization architecture to enhance the data acquisition efficiency. Moreover, the realistic propagation model and shadowing due to buildings are considered to improve the recovery accuracy.

The flowchart of the proposed construction scheme is shown in Fig. 2, which mainly contains two steps, i.e., spatial measurement matrix (or sampling location) optimization and 3D spectrum data recovery. Firstly, combined with the minimum SBL variance indicator and Principal Component Analysis (PCA) dimension reduction technique, we carefully select the sampling quantity and locations. Thus, the optimized measurement matrix ψ can be obtained. Secondly, according to the propagation fading characteristics in the realistic environment, we decompose the spectrum situation recovery problem into two layers, i.e., sparse signal recovery and channel dictionary optimization. The sparse signal is firstly recovered based on the SBL by mining the inherent sparsity of spectrum situation. Then, we utilize sampling data to optimize channel dictionary by considering shadow fading. Finally, the whole REM can be constructed based on the sparse signal and optimized channel dictionary.

B. Sampling Number Minimization and Location Optimization

According to the law of signal propagation, the 3D spectrum situation exhibits a significant correlation in the spatial

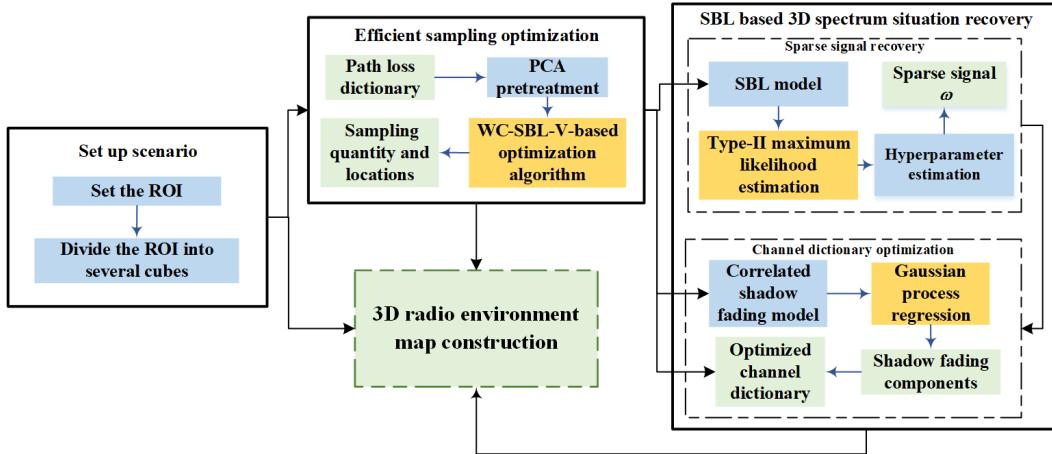


Fig. 2. The flowchart of the proposed 3D REM construction scheme.

domain, i.e., the RSS values of different cubes have a strong spatial correlation [14]. Thus the spatial sampling scheme has a significant impact on the efficiency and accuracy of REM construction [49]. Compared with random or fixed sampling scheme, we can achieve better performance if the selected sampling locations can exploit the intrinsic features of spectrum data. On the other hand, it is necessary to determine the minimum number of samples and their locations to meet the required recovery accuracy.

In the sensor network design problem, the typically challenge is to determine the optimal placement of sensor nodes. This involves selecting the minimum number of samples and the corresponding locations within a predefined spatial domain to meet specified accuracy requirements [50]. Specifically, for a given sparse dictionary φ , how to determine the measurement matrix ψ , and thus the sparse signal can be estimated within a predefined accuracy.

Let us define the index set \mathcal{S} as the candidate sampling locations. The subset $\mathcal{S}_d \subset \mathcal{S}$ is defined as the determined sampling locations. According to the recovery process of sparse signal in (19), it is a minimum variance unbiased estimation. The variance Σ_ω in (20) can be viewed as an error indicator to determine the sampling set ς ($\varsigma \subset \mathcal{S}$), and the measurement matrix is determined as $\psi = (\mathbf{I}_N)_{\varsigma \cdot}$, which consists of the rows indexed by the set ς in \mathbf{I}_N . The mean squared error (MSE) can be expressed as

$$\begin{aligned} MSE(\hat{\omega}) &= E(\|\omega - \hat{\omega}\|^2) \\ &= \text{tr}(\Sigma_\omega) \\ &= \text{tr}(\beta \Phi^T \Phi + \mathcal{A})^{-1}, \end{aligned} \quad (25)$$

where the sensing matrix $\Phi = \psi \varphi$ is consisted of the rows of φ indexed by the sampling set ς . In the initialization stage, there is no prior information of ω . We assign the same hyperparameter $\alpha_n = \alpha, n = 1, 2, \dots, N$. Then, we can convert (25)

to

$$\begin{aligned} MSE(\hat{\omega}) &= \text{tr}(\beta \Phi^T \Phi + \mathcal{A})^{-1} \\ &= \text{tr}\left[\beta \left(\Phi^T \Phi + \frac{\alpha}{\beta} \mathbf{I}_N\right)\right]^{-1} \\ &= \beta^{-1} \cdot \text{tr}\left(\Phi^T \Phi + \frac{\alpha}{\beta} \mathbf{I}_N\right)^{-1}, \end{aligned} \quad (26)$$

where α/β is the ratio of noise variance to sparse signal variance, which is sufficiently small that can be ignored [51]. Furthermore, it yields

$$\begin{aligned} MSE(\hat{\omega}) &\approx \beta^{-1} \cdot \text{tr}(\Phi^T \Phi)^{-1}, \\ &= \beta^{-1} \sum_{n=1}^N (\lambda_n^*)^{-1}, \end{aligned} \quad (27)$$

where $\lambda_1^* \geq \lambda_2^* \geq \dots \geq \lambda_N^*$ are the eigenvalues of $\mathbf{H} = \Phi^T \Phi$, here \mathbf{H} is defined as the dual sensing matrix. The parameter β is a constant and it is dependent on σ_0^2 but independent of sensing locations.

As defined in [52], the worst case error variance (WCEV), which is defined as the variance of estimation error over all directions $\ell \in \mathbb{R}^N$ with $\|\ell\| = 1$, can be computed as follows

$$\begin{aligned} WCEV_{\mathbf{H}} &= \max_{\|\ell\|=1} \ell^T \mathbf{H}^{-1} \ell \\ &= \lambda_{\max}(\mathbf{H}^{-1}), \end{aligned} \quad (28)$$

where \mathbf{H}^{-1} corresponds to the Fisher information matrix (FIM) and denotes the maximum eigenvalue of \mathbf{H}^{-1} , i.e., the minimum eigenvalue of \mathbf{H} . We denote $\lambda_{\max}(\mathbf{H}^{-1})$ as the WC-SBL-V index. This measure equals to the E-optimal experiment design (maximum eigenvalue constraint on the FIM \mathbf{H}^{-1}) [53]. It can be proved that two error indicators are equivalent since two matrix norms, i.e., 2-norm and F-norm, are equivalent [54].

In order to determine the minimum sampling number, the sparse-promoting technique is used by adding a sparsity-promoting penalty term to the cost function. The maximum acceptable eigenvalue λ_{WCEV} is defined as the WC-SBL-V

index threshold. Accordingly, the desired sample set $\varsigma \subset \mathcal{S}$ can be expressed as a sparse optimization problem,

$$\begin{aligned} \hat{\varsigma} &= \arg \min_{\varsigma \subset \mathcal{S}} |\varsigma|, \\ \text{s.t. } &\lambda_{\max}(\mathbf{H}^{-1}) \geq \lambda_{WCEV}, \\ &\mathbf{H} = \boldsymbol{\Phi}^T \boldsymbol{\Phi}, \\ &\boldsymbol{\Phi} = \boldsymbol{\psi} \boldsymbol{\varphi}, \\ &\boldsymbol{\psi} = (\mathbf{I}_N)_{\varsigma \cdot}, \end{aligned} \quad (29)$$

where $|\cdot|$ is the cardinality of a set.

The cardinality optimization problem of (29) can be solved by evaluating the maximum eigenvalue of Fisher information matrix over all potential sampling location configuration. One straightforward method is to evaluate all possible combinations of the potential sizes of the candidate sampling locations. The combination with the least number of samples that satisfies the required estimation accuracy is chosen. However, the computational cost of exhaustively searching is 2^N , which is computationally intractable for a large-scale problem.

Incremental determination of sampling locations effectively reduces search space for sample configurations. Accordingly, the greedy algorithm is adopted in this paper. Each sampling location is selected after traversing all the unselected samples and find the one that maximizes the minimum nonzero eigenvalue. However, the candidate observation matrix has full row rank ($\boldsymbol{\Phi} \in \mathbb{R}^{M \times N}, M < N$). Therefore, the dimension of \mathbf{H} is $\mathbb{R}^{N \times N}$, but it is computationally expensive to calculate the eigenvalues due to the high spatial dimension. PCA is an efficient dimension reduction technology, and it is utilized to reduce the dictionary from N vectors to r vectors, which are the principle components of the N -colum vectors of $\boldsymbol{\varphi} \in \mathbb{R}^{N \times N}$, and we can obtain $\boldsymbol{\varphi}^p \in \mathbb{R}^{N \times r}$. The corresponding dual observation matrix can be converted to

$$\mathbf{H}^p = (\boldsymbol{\Phi}^p)^T \boldsymbol{\Phi}^p, \quad (30)$$

where $\boldsymbol{\Phi}^p$ is consisted of the rows indexed by set ς in $\boldsymbol{\varphi}^p$, and $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r$ are the eigenvalues of \mathbf{H}^p . Then, let $\boldsymbol{\Phi}_t^p$ denote the sensing matrix after selecting t samples as

$$\boldsymbol{\Phi}_t^p = \left[(\boldsymbol{\varphi}_{i_1}^p)^T, (\boldsymbol{\varphi}_{i_2}^p)^T, \dots, (\boldsymbol{\varphi}_{i_{t-1}}^p)^T, (\boldsymbol{\varphi}_{i_t}^p)^T \right]^T, \quad (31)$$

where i_t ($i_t \in \varsigma, i_t \in \mathcal{S}$) is the index of t th selected sample. The observation vector $\boldsymbol{\varphi}_{i_t}^p$ is the i_t th row vector of $\boldsymbol{\varphi}^p$. Accordingly, the sparse optimization problem of (29) can be converted to

$$\begin{aligned} \hat{\varsigma} &= \arg \min_{\varsigma \subset \mathcal{S}} |\varsigma|, \\ \text{s.t. } &\lambda_r \geq \lambda_{WCEV}, \\ &\mathbf{H}^p = (\boldsymbol{\Phi}^p)^T \boldsymbol{\Phi}^p, \\ &\boldsymbol{\Phi}^p = \boldsymbol{\varphi}_{\varsigma \cdot}^p. \end{aligned} \quad (32)$$

Let $\lambda_r(\mathbf{H}_{t-1}^p)$ denote the minimum eigenvalue of \mathbf{H}_{t-1}^p as

$$\lambda_r(\mathbf{H}_{t-1}^p) = \min_{\|\boldsymbol{\ell}\|=1} \boldsymbol{\ell}^T \mathbf{H}_{t-1}^p \boldsymbol{\ell}. \quad (33)$$

After the t th sample is selected, the minimum eigenvalue of \mathbf{H}_t^p can be expressed as

$$\lambda_r(\mathbf{H}_t^p) = \min_{\|\boldsymbol{\ell}\|=1, i_t \in \mathcal{S} \setminus \varsigma} \boldsymbol{\ell}^T \mathbf{H}_{t-1}^p \boldsymbol{\ell} + \left((\boldsymbol{\varphi}_{i_t}^p)^T \boldsymbol{\ell} \right)^2, \quad (34)$$

with

$$\boldsymbol{\ell} = \boldsymbol{\pi}_r^{(t-1)} \mathbf{H}_{t-1}^p, \quad (35)$$

where $\boldsymbol{\pi}_r^{(t-1)}$ is the normalized minimum eigenvector of \mathbf{H}_{t-1}^p . Moreover, the minimum eigenspace is also introduced here via optimizing the new criterion efficiently [50]. The t th sample can be determined by maximizing the projection of row vector $\boldsymbol{\varphi}_{i_t}^p$ on the minimum eigenspace of \mathbf{H}_{t-1}^p .

In the process of determining ς , the number of samples is different due to the different threshold value. Then, t can appear in two cases, i.e., $t \leq r$ and $t > r$. The minimum eigenspace of \mathbf{H}_{t-1}^p is defined for the two cases as follows.

i) For $t \leq r$, the minimum eigenspace of \mathbf{H}_{t-1}^p is the eigenspace associated with all the minimum eigenvalues of \mathbf{H}_{t-1}^p , as

$$\Pi_{t:r}^{t-1}(\mathbf{H}_{t-1}^p) = \text{span} \left(\boldsymbol{\pi}_t^{(t-1)}, \boldsymbol{\pi}_{t+1}^{(t-1)}, \dots, \boldsymbol{\pi}_r^{(t-1)} \right), \quad (36)$$

where $\boldsymbol{\pi}_j^{(t-1)}$ is the normalized eigenvector associated with the j th eigenvalue $\lambda_j(\mathbf{H}_{t-1}^p)$ of \mathbf{H}_{t-1}^p . The minimum eigenspace of \mathbf{H}_{t-1}^p of (36) is equal to the null space of \mathbf{H}_{t-1}^p as

$$\Pi_{t:r}^{t-1}(\mathbf{H}_{t-1}^p) = \text{null}(\boldsymbol{\Phi}_{t-1}^p). \quad (37)$$

ii) For $t > r$, the minimum eigenspace of \mathbf{H}_{t-1}^p is exactly the subspace spanned by the minimum eigenvector $\boldsymbol{\pi}_r^{(t-1)}$ as

$$\Pi_r^{t-1}(\mathbf{H}_{t-1}^p) = \text{span} \left(\boldsymbol{\pi}_r^{(t-1)} \right) = \Pi_r^{t-1}. \quad (38)$$

When $t \leq r$, the t th sample can be determined by

$$\hat{i}_t = \arg \max_{i_t \in \mathcal{S} \setminus \varsigma} \left\| \mathcal{O}_{t-1} \boldsymbol{\varphi}_{i_t}^p \right\|_2^2, \quad (39)$$

with

$$\mathcal{O}_{t-1} = \mathbf{I}_r - \Theta_{t-1} \Theta_{t-1}^T, \quad (40)$$

where $\Theta_{t-1} = \text{orth}((\boldsymbol{\Phi}_{t-1}^p)^T)$. \mathcal{O}_{t-1} is a projection matrix which can project $\boldsymbol{\varphi}_{i_t}^p$ onto the minimum eigenspace $\text{null}(\boldsymbol{\Phi}_{t-1}^p)$. When $t > r$, the t th sample can be obtained by

$$\hat{i}_t = \arg \max_{i_t \in \mathcal{S} \setminus \varsigma} \left\| \mathcal{R}_{t-1} \boldsymbol{\varphi}_{i_t}^p \right\|_2^2, \quad (41)$$

with

$$\mathcal{R}_{t-1} = \boldsymbol{\pi}_r^{t-1} (\boldsymbol{\pi}_r^{t-1})^T, \quad (42)$$

where \mathcal{R}_{t-1} is a projection matrix which can project $\boldsymbol{\varphi}_{i_t}^p$ onto the minimum eigenspace Π_r^{t-1} .

We evaluate all unselected observation vectors and choose the one that maximizes the eigenspace. Meanwhile, if $t \geq r$, we check the constraint in (32) after each sampling location is determined. The process of (37)-(42) is iterated until the cut-off condition $\lambda_r \geq \lambda_{WCEV}$ is satisfied. Then, the minimum and optimized sampling set $\varsigma = [i_1, i_2, \dots, i_M]$ can be obtained. It should be mentioned that the minimum number of required samples is determined by judging whether $\lambda_r^{(t)} \geq \lambda_{WCEV}$ is satisfied after the t th sample is determined. Accordingly, the constraint is only used to judge whether the number of required samples is enough. Algorithm 1 summarizes the process of optimization.

The time complexity of sparse dictionary PCA preprocessing is $O(N^3)$. The rest time complexity mainly focuses

on line 3 and line 15 in Algorithm 1. To determine the t th sample, the main computational cost is attributed to the optimization problem $\hat{i}_t = \arg \max_{i_t \in \mathcal{S} \setminus \mathcal{S}_d} \|\mathcal{O}_{t-1} \varphi_{i_t}^p\|_2^2$ and $\hat{i}_t = \arg \max_{i_t \in \mathcal{S} \setminus \mathcal{S}_d} \|\mathcal{R}_{t-1} \varphi_{i_t}^p\|_2^2$, which cost $O((N - t + 1) n^2)$ per selection. The selection of required M samples under the given threshold total costs $O(N M r^2)$. Then, the total time complexity C_1 of Algorithm 1 is

$$\begin{aligned} C_1 &= O(N^3) + O\left(\sum_{t=1}^M (N - t + 1) r^2\right) \\ &= O(N^3 + N M r^2). \end{aligned} \quad (43)$$

Algorithm 1 WC-SBL-V based measurement matrix optimization.

Input:

3D REM sparse dictionary $\varphi \in \mathbb{R}^{N \times N}$;

Initial selected sample set $\varsigma = \emptyset$;

The maximum acceptable variance λ_{WCEV} ;

Output:

Selected sample set ς ; The measurement matrix ψ ; The number of selected samples M ;

- 1: **Initialize** $\mathcal{S} = \{1, 2, \dots, N\}$, $\mathcal{S}_d = \emptyset$, $t = 0$, $\lambda_r = \infty$, $\mathcal{O}_0 = \mathbf{I}_n$;
 - 2: Obtain $\varphi^p \in \mathbb{R}^{N \times r}$ by PCA;
 - 3: **while** $t \leq r$ & $\lambda_r^{(t)} \geq \lambda_{WCEV}$ **do**;
 - 4: $t = t + 1$;
 - 5: Solve \hat{i}_t according to (39);
 - 6: Update $\varsigma = \varsigma \cup \{\hat{i}_t\}$, $\Phi_t^p = \left[(\Phi_{t-1}^p)^T \varphi_{\hat{i}_t}^p \right]^T$, $\Theta_t = \text{orth}\left((\Phi_t^p)^T\right)$, $\mathcal{O}_t = \mathbf{I}_r - \Theta_t \Theta_t^T$;
 - 7: Establish the projection matrix \mathcal{O}_t on the minimum eigenspace according to (40);
 - 8: $\mathcal{S}_d = \varsigma$;
 - 9: **end while**;
 - 10: **while** $t > r$ & $\lambda_r^{(t)} \geq \lambda_{WCEV}$ **do**;
 - 11: $t = t + 1$;
 - 12: Solve \hat{i}_t according to (41);
 - 13: Update $\varsigma = \varsigma \cup \{\hat{i}_t\}$, $\Phi_t^p = \left[(\Phi_{t-1}^p)^T \varphi_{\hat{i}_t}^p \right]^T$, $(\Phi_t^p)^T \Phi_t^p = \Pi^t \lambda^{(t)} (\Pi^t)^T$, $\lambda^{(t)} = \text{diag}[\lambda_1^{(t)}, \lambda_2^{(t)}, \dots, \lambda_r^{(t)}]$, $\mathcal{R}_t = \pi_r^t (\pi_r^t)^T$;
 - 14: Obtain the minimum eigenspace according to (38);
 - 15: Establish the projection matrix \mathcal{R}_t on the minimum eigenspace according to (42);
 - 16: $\mathcal{S}_d = \varsigma$;
 - 17: **end while**;
 - 18: **return** ς , $M = t$, $\psi = (\mathbf{I}_N)_\varsigma$;
-

C. SBL-Based Spectrum Data Recovery Incorporating Channel Shadowing

As discussed in Section II-B, the REM data is recovered hierarchically by SBL and GPR. Given the sampling data, we estimate the sparse RF transmitter signal $\hat{\omega}$ by SBL. We further derive the shadow fading components at the sampling locations

and construct a GP to estimate the channel shadowing $\bar{\xi}^{v^*}$ at the unsampled locations. Finally, we recover the spectrum data according to (2).

Firstly, we can recover ω with the mean μ and evaluate the recovery accuracy by the variance Σ_ω once the hyperparameters α and β have been estimated. They are estimated by a maximum a posterior (MAP) probability as [21]

$$\begin{aligned} (\alpha, \beta) &= \arg \max_{\alpha, \beta} p(\alpha, \beta | \mathbf{t}), \\ &= \arg \max_{\alpha, \beta} p(\mathbf{t} | \alpha, \beta) p(\alpha) p(\beta), \\ &= \arg \max_{\alpha, \beta} \ln p(\mathbf{t} | \alpha, \beta) p(\alpha) p(\beta). \end{aligned} \quad (44)$$

The hyperparameters are determined by approximating the posterior (18) from sampling data. This is equal to maximizing the evidence for the measurements $p(\mathbf{t} | \alpha, \beta)$ in (44), which is referred to type-II maximum likelihood. It intuitively selects the hyperparameters that are best supported by the observation data [55].

Then, the re-estimation rule of α_n is [56]

$$\alpha_n^* = \frac{1 + 2a}{\mu_n^2 + (\Sigma_\omega)_{nn} + 2b}. \quad (45)$$

However, by defining quantities $\Upsilon_n \equiv 1 - \alpha_n \Sigma_{nn}$, it yields

$$\alpha_n^* = \frac{\Upsilon_n + 2a}{\mu_n^2 + 2b}, \quad (46)$$

which can lead to much faster convergence than (45). The update rule of β can be expressed as

$$\beta^* = \frac{M - \sum_i \Upsilon_i + 2c}{\|\mathbf{t} - \Phi \mu\|_2^2 + 2d}. \quad (47)$$

The derivations of (45) and (47) are given in Appendix B.

If $\alpha_n^{-1} = 0$ ($\alpha_n \rightarrow \infty$), we obtain $\omega_n = 0$. It means that there is no transmitter in the n th cube [21]. Therefore, we can remove these locations to accelerate the update process. Since the matrix inversion of SBL variance Σ_ω in (20) is also computationally complex, we apply the matrix inversion lemma and write it as

$$\Sigma_\omega = (\mathcal{A})^{-1} - (\mathcal{A})^{-1} \Phi^T (\Omega)^{-1} \Phi (\mathcal{A})^{-1}, \quad (48)$$

with

$$\Omega = \Phi (\mathcal{A})^{-1} \Phi^T + \beta^{-1} \mathbf{I}. \quad (49)$$

Note that its computational complexity is lowered than $O(M^2 N)$ while the one of (20) is $O(N^3)$. By performing the iteration between (46), (47), (48), (49) and (19) until the convergence condition is satisfied, the MAP estimation of ω can be obtained from the mean of posterior, i.e., $\hat{\omega} = \mu$.

Secondly, according to correlated shadow fading model in Section II-B, we introduce the shadow fading estimation based on the GPR. In general, noise is considered in the GP model. With training the input set $\bar{\xi}^s$ calculated by (24), the corresponding model output is defined as

$$\mathbf{y} = \bar{\xi}^s + \boldsymbol{\delta}, \quad (50)$$

where $\delta_m \sim \mathcal{N}(0, \sigma_{GP}^2)$, $m = 1, 2, \dots, M$. The joint distribution of the prediction values $\bar{\xi}^{v^*}$ at the unsampled

cubes ν^* ($\nu^* = \{i | i \in \nu, i \notin s\}$) can be represented as a multi-normal distribution

$$\begin{bmatrix} \mathbf{y} \\ \bar{\xi}^{v^*} \end{bmatrix} | s, \nu^* \sim \mathcal{N} \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathcal{C}(s, s) + \sigma_{GP}^2 \mathbf{I} & \mathcal{C}(s, \nu^*) \\ \mathcal{C}(\nu^*, s) & \mathcal{C}(\nu^*, \nu^*) \end{bmatrix} \right). \quad (51)$$

According to GPR, the predictive distribution of $\bar{\xi}^{v^*}$ satisfies the multivariate Gaussian distribution

$$p(\bar{\xi}^{v^*} | \mathbf{y}, s, \nu^*) \sim \mathcal{N}(\mu_{GP}^*, \Sigma_{GP}^*), \quad (52)$$

with

$$\mu_{GP}^* = \mathcal{C}(\nu^*, s) (\mathcal{C}(s, s) + \sigma_{GP}^2 \mathbf{I})^{-1} \mathbf{y}, \quad (53)$$

$$\begin{aligned} \Sigma_{GP}^* &= \mathcal{C}(\nu^*, \nu^*) + \sigma_{GP}^2 \mathbf{I} \\ &\quad - \mathcal{C}(\nu^*, s) (\mathcal{C}(s, s) + \sigma_{GP}^2 \mathbf{I})^{-1} \mathcal{C}(s, \nu^*). \end{aligned} \quad (54)$$

We apply Bayesian inference idea to solve GPR. The marginal likelihood is given by

$$p(\mathbf{y} | s, \eta_{GP}) \sim \mathcal{N}(\mathbf{y} | \mathbf{0}, \Sigma_{\eta_{GP}}), \quad (55)$$

$$\Sigma_{\eta_{GP}} = \mathcal{C}(s, s) + \sigma_{GP}^2 \mathbf{I}, \quad (56)$$

where $\eta_{GP} = [\rho, \sigma^2, \sigma_{GP}^2]$ are the parameters of the Matérn covariance function and GP noise variance. The parameters η_{GP} can be estimated by minimizing the negative log marginal likelihood (NLML) with respect to η_{GP} , as

$$\begin{aligned} \hat{\eta}_{GP} &= \arg \min_{\eta_{GP}} \mathcal{L}(\eta_{GP}), \\ &= \arg \min_{\eta_{GP}} (-\log p(\mathbf{y} | s, \eta_{GP})), \\ &= \arg \min_{\eta_{GP}} \left(\frac{1}{2} \mathbf{y}^T \Sigma_{\eta_{GP}}^{-1} \mathbf{y} + \frac{1}{2} \log(|\Sigma_{\eta_{GP}}|) + \frac{M}{2} \log 2\pi \right). \end{aligned} \quad (57)$$

which is a non-convex problem. The optimum of problem (57) can be solved by the gradient-based optimization algorithm

$$\begin{aligned} \partial \mathcal{L}(\eta_{GP}) / \partial \eta_{GP} &= -\frac{1}{2} \mathbf{y}^T \Sigma_{\eta_{GP}}^{-1} \frac{\partial \Sigma_{\eta_{GP}}}{\partial \eta_{GP}} \Sigma_{\eta_{GP}}^{-1} \mathbf{y} \\ &\quad + \frac{1}{2} \text{tr} \left(\Sigma_{\eta_{GP}}^{-1} \frac{\partial \Sigma_{\eta_{GP}}}{\partial \eta_{GP}} \right). \end{aligned} \quad (58)$$

After obtaining the prediction model $p(\bar{\xi}^{v^*} | \mathbf{y}, s, \nu^*)$, the shadow fading components $\bar{\xi}^{v^*}$ of unsampled cubes can be estimated by μ_{GP}^* . We further obtain shadow fading $\xi^v = \bar{\xi}^v / 10$ of all cubes. Finally, we can reconstruct the 3D REM (or REM tensor) by using (2). The recovery process is summarized in Algorithm 2.

The complexity of SBL-based REM construction is mainly concentrated in line 3, 4 and 7 in Algorithm 2. Since the iteration number is a constant, this algorithm's complexity is decided by the most time-consuming step. As discussed above, the computation of (48) costs $O(M^2 N)$. The computation of (19) and (47) both cost $O(MN^2)$, which occupy the largest proportion of computational cost. Therefore, the total time complexity of SBL for sparse signal recovery is $O(MN^2)$. The time complexity of GPR for shadow fading estimation mainly lies in the process of Bayesian inference with M

samples, which is $O(M^3)$. Accordingly, the total complexity is

$$\begin{aligned} \mathcal{C}_2 &= O(MN^2) + O(M^3) \\ &= O(M^3 + MN^2). \end{aligned} \quad (59)$$

Algorithm 2 SBL-based spectrum data hierarchical recovery incorporating channel shadowing.

Input:

3D REM sensing matrix $\Phi \in \mathbb{R}^{M \times N}$;
3D REM sparse dictionary $\varphi \in \mathbb{R}^{N \times N}$;
RSS t ; The locations of samples s ; $\text{thre_}\alpha$; a_0 ; b_0 ; c_0 ;
 d_0 ;
Max iteration iter_{\max} ;
The optimized spatial measurement matrix ψ ;

Output:

Constructed 3D REM vector \mathbf{x} ; Recovered sparse signal $\hat{\omega}$; The estimated shadow fading ξ^v ;
1: **Initialize** $\alpha, \beta, \text{iter} = 1, \mu = \mathbf{0}_{N \times 1}, \Sigma_\omega$;
2: **while** $\text{iter} < \text{iter}_{\max}$ or
 $\|\mathcal{L}(\alpha^{\text{iter}}, \beta^{\text{iter}}) - \mathcal{L}(\alpha^{\text{iter}-1}, \beta^{\text{iter}-1})\| < 10^{-4}$ do;
3: Update α with (46);
4: Update β with (47);
5: Calculate μ and Σ_ω by (19), (48) and (49);
6: $\text{iter} = \text{iter} + 1$;
7: **end while**;
8: Obtain $\hat{\omega} = \mu$;
9: Calculate the shadow fading at s by (24);
10: Calculate the NLML $\mathcal{L}(\eta_{GP})$ by (55) - (57);
11: Solve $\hat{\eta}_{GP} = \arg \min_{\eta_{GP}} \mathcal{L}(\eta_{GP})$ based on (58);
12: Get prediction model $p(\bar{\xi}^{v^*} | \mathbf{y}, s, \nu^*)$ by (52) - (54);
13: Obtain $\bar{\xi}^{v^*} = \mu_{GP}^*, \mathbf{x} = \xi^v \circ (\varphi \hat{\omega})$;

TABLE I
THE MAIN SIMULATION PARAMETERS

Parameter	Value			
	Index	Height(m)	Power (dBmW)	Antenna Type
RF emitter setting	1	1.5	20	Isotropic
	2	1.5	20	Isotropic
	3	1.5	20	Isotropic
	4	1	20	Isotropic
	5	1	20	Isotropic
	6	20	20	Directional
	7	30	20	Isotropic
	8	2	20	Isotropic
Center frequency	2.45GHz			
Cartography area	1250m × 1250m × 50m.			
Grid resolution	5m × 5m × 10m.			
REM tensor size	$N_x \times N_y \times N_z = 250 \times 250 \times 6$			

IV. SIMULATION RESULTS AND DISCUSSIONS

A. Experiment Setup

In this section, the proposed 3D REM construction method is validated and evaluated by simulations. The 3D ROI is

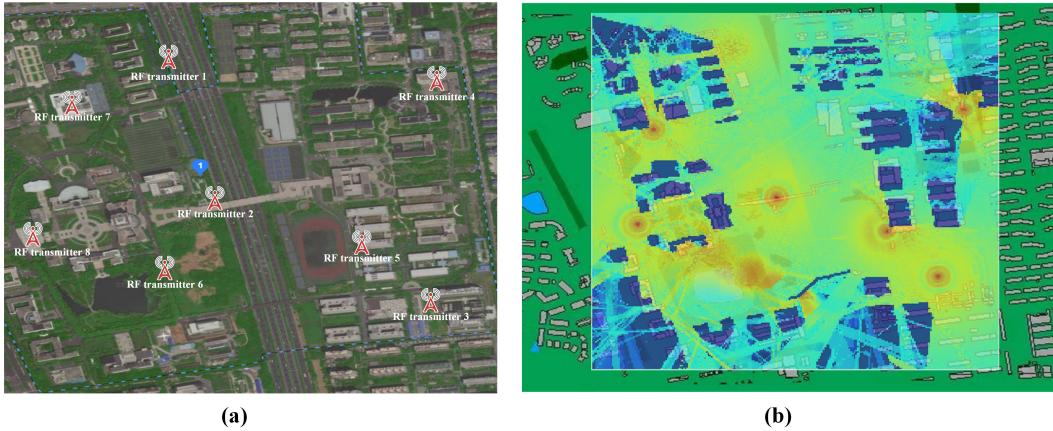


Fig. 3. The satellite map and ideal REM ($h = 2\text{m}$) of campus scenario.

a campus scenario and the size is $1250\text{m} \times 1250\text{m} \times 50\text{m}$. The satellite view is shown in Fig. 3 (a). There are many buildings densely distributed with heights from 19 to 55m and the average height is about 30 m. The terrain consists of three types, i.e., dry soil, wet soil, and vegetation. We randomly set eight RF transmitters, e.g., pedestrians, vehicles, micro base stations, and so on, and each one has a directional or isotropic antenna. The main simulation parameters are shown in the Table I.

In the simulation, we firstly discretize the ROI into $N = 250 \times 250 \times 6 = 375000$ cubes and each cube is $5\text{m} \times 5\text{m} \times 10\text{m}$. Since it is quite difficult to measure the spectrum data due to high-cost hardware system and stability of spectrum situation, we leverage the RT technique to calculate the 3D REM for the performance evaluation purpose. The RT technique has been widely used for radio propagation modeling, and has good performance on the RSS prediction under the specific area. Given detailed information of scenario and RF transmitters, the ideal REM is obtained based on the RT method, which is denoted by a spectrum tensor $\chi \in \Re^{250 \times 250 \times 6}$ as shown in Fig. 3 (b). Note that the REM is divided into cubes and we use the averaged signal strength as RSS value to remove the effect of small-scale fading [9].

Seven construction methods including different sampling positions and different recovery methods are conducted via different sampling rates. The proposed hierarchical construction algorithm is denoted as SNLO-SBLHM, where SNLO represents the proposed sampling optimization algorithm in Algorithm 1 and SBLHM denotes the hierarchical recovery method in Algorithm 2. Lasso denotes the popular CS-based least-absolute shrinkage and selection operator (Lasso) [20]. The Kriging algorithm and the low-rank tensor (matrix) completion HALRTC algorithm [57], are also included as the representation of data-driven methods. Moreover, four sampling optimization schemes, i.e., SNLO, Random sampling, the determinant-based greedy (DG) sampling algorithm [58], and FrameSense sampling algorithm [59], are also considered. The channel propagation model is defined as the free-space path loss model with the path loss exponent η of 2 [20]. The hyperparameters are generally given fixed small values

to make these priors non-information, i.e., $a_0 = b_0 = 10^{-6}$, $c_0 = d_0 = 0$.

B. 3D REM Construction Performance

To demonstrate the performance of different methods, we define the mean absolute error (MAE) as the averaged difference in RSS values between the reconstructed REM and ideal REM. It can be expressed as

$$MAE^{\text{REM}} = \frac{1}{N} \sum_{n=1}^N |P_{\text{est}}^r(n) - P_{\text{true}}^r(n)|, \quad (60)$$

where $P_{\text{est}}^r(n)$ and $P_{\text{true}}^r(n)$ are the estimated and true RSS value in dBm at the n th cube, respectively.

As we can see from Fig. 4, the MAEs of SNLO-SBLHM, DG-SBLHM, FrameSense-SBLHM, Random-SBLHM, Random-Lasso, Random-HALRTC, and Random-Kriging decrease as the sampling rate increases. The proposed SNLO-SBLHM consistently outperforms other methods even at very low sampling rates. Besides, compared with the Random-Lasso, Random-HALRTC, and Random-Kriging, the proposed hierarchical method improves the construction performance substantially by considering the shadow fading under realistic scenarios. The proposed method outperforms other data-driven methods, i.e., Kriging and HALRTC, with a performance improvement of at least 50%. Notably, the proposed method achieves an improvement of about 7dB compared to Random-Lasso. These improvements are essential for practical applications.

On the other hand, the sampling location is vital for the REM construction, particularly when the number of sampling data is restricted. Compared to Random-SBLHM, Random-Lasso, Random-HALRTC, and Random-Kriging, the proposed method significantly improves the construction performance by considering sampling optimization. Two data-driven algorithms' performances are unsatisfactory at low sampling rates since they usually need sufficient data to extract the spectrum correlations.

Fig. 5 shows that the DG-SBLHM algorithm are more computational than SNLO-SBLHM, which means the PCA preprocessing and proposed sampling scheme can reduce

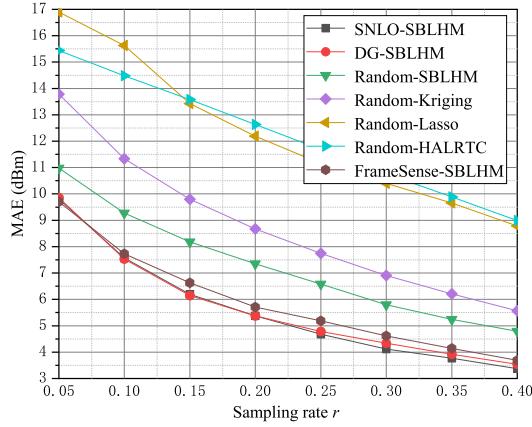


Fig. 4. 3D REM construction performance comparisons.

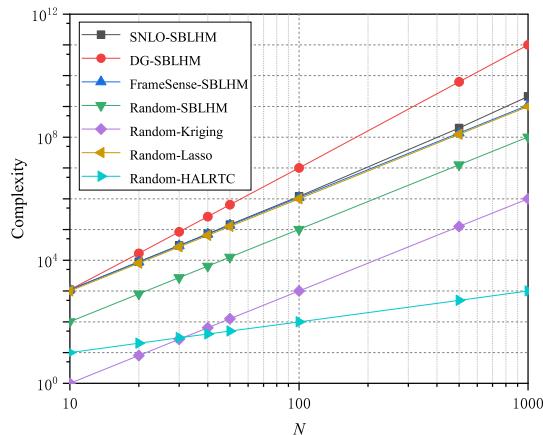


Fig. 5. Performance comparisons of complexity.

the complexity. Meanwhile, Fig. 4 shows that the construction performances of SNLO-SBLHM and DG-SBLHM are comparable as r increases, which means the PCA preprocessing can maintain the performance with low complexity. Moreover, Random-SBLHM, Random-Lasso, Random-HALRTC, and Random-Kriging present comparable complexity and are better than DG-SBLHM, FrameSense-SBLHM, and SNLO-SBLHM. Since the sampling locations of latter methods are randomly generated without optimization. Note that FrameSense-SBLHM has less complexity than SNLO-SBLHM, but the SNLO-SBLHM has better construction performance.

C. Sampling Optimization Performance

The WC-SBL-V index of different sampling schemes and the corresponding MAEs are given in Fig. 6. Four sampling schemes are compared in this paper, i.e., SNLO, DG, FrameSense, and random. It is shown that if the sampling number is the same, the SNLO provides the best performance or the

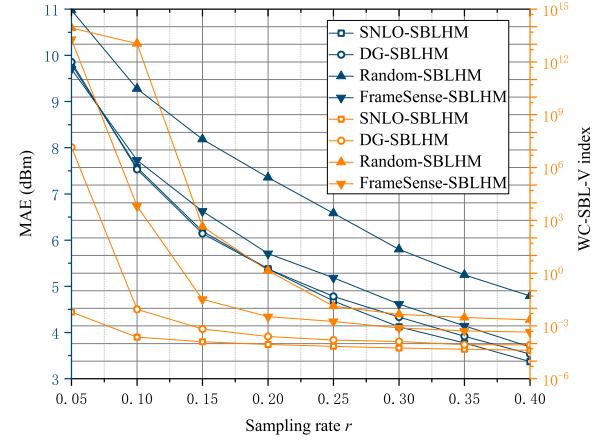


Fig. 6. Comparison of sampling optimization performance.

minimum WC-SBL-V. In other words, if we fix the WC-SBL-V index threshold, the proposed SNLO algorithm requires the least number of sampling number to meet the accuracy requirement. It is more prominent when the sampling rate is low. In Fig. 6, it is clearly shown that SNLO outperforms the FrameSense and random sampling in identifying critical sensing locations, especially when the number of samples is limited. Besides, the SNLO and DG have comparable performances. DG considers the influence of all eigenvalues' product, i.e., determinants, whereas SNLO only considers the influence of the largest eigenvalue. FrameSense considers the sum of all eigenvalues. While the maximum eigenvalue contributes the main part of determinant, especially when the sampling rate is low [50]. This also provides an explanation for why DG and SNLO outperform FrameSense. Therefore, it is sufficient to focus on the influence of maximum eigenvalue, i.e., WC-SBL-V index, to achieve satisfactory performance. Moreover, SNLO can greatly reduce the computational complexity than DG. It is apparent from Fig. 6 that random sampling yields the worst result, reaffirming the crucial importance of spatial sampling location selection.

D. Impact of Different Parameters

The impact of the proportion threshold on the REM construction performance and WC-SBL-V index is shown in Figs. 7 and 8, respectively. It can be seen that the performance of proposed algorithm generally improves with the decrease of Per_th . The WC-SBL-V index also decreases as the decrease of proportion threshold, which aligns with the objective function when optimizing sampling. In Fig. 8, it can be observed that the MAE shows minimal fluctuations with different Per_th . This suggests that the algorithm's performance remains relatively stable even with slight errors. However, it is notable that the threshold significantly influences the complexity of the algorithm. Accordingly, by retaining more principal components of sparse dictionary, we can reduce the sparse signal recovery errors, and thus obtain an improved

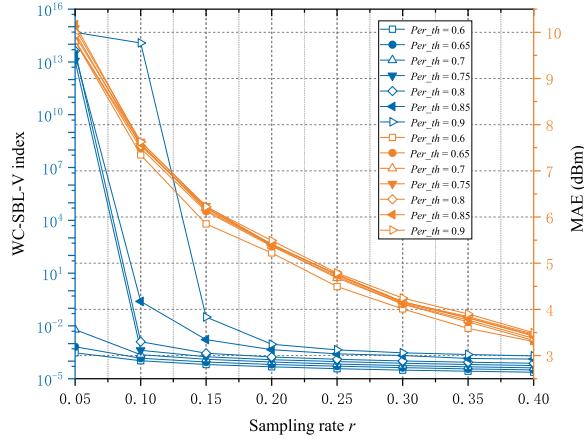


Fig. 7. Impact of Per_th on the construction performances and the WC-SBL-V index.

spectrum data recovery performance. However, with the increase of the proportion of retained principal components, the corresponding matrix dimension also increases, which results in an increase in computational complexity. Then, we need to reasonably balance the computational complexity and the performance, especially for the case of high-dimensional data. Note that when the matrix is high-dimensional, the selected sampling locations are also affected, which can result in the final selection being too close with each other and bring unsatisfactory recovery performance.

The impact of SNR on the construction performance is shown in Fig. 9. The SBL-based algorithms have better performance than others which demonstrates that the SBL-based method has an excellent noise suppression capability. Based on this feature, the proposed SNLO-SBLHM further improves the recovery accuracy by optimizing the sampling positions and considering the shadow fading. In addition, it can be observed that the Kriging algorithm has the worst performance in the presence of noise. The reason is that it relies on the geographical location and measured data to estimate the spectrum data at the unsampled positions, which cannot identify the noise influence.

E. REM Construction Performance Visualization

In order to intuitively present the construction performance of different methods under realistic scenarios, Fig. 10 compares the constructed 3D REMs at different altitudes under the campus scenario. It can be seen that the SNLO-SBLHM can achieve the most effective spectrum situation recovery. This is especially noticeable for the spectrum data near the RF transmitters or around the buildings. Fig. 10 (b)-(e) demonstrate that the recovery algorithms based on SBL exhibit superior performances in radiation source recovery. By comparing (g) and (h), it reveals that Kriging has better performance than HALRTC. This is because that HALRTC requires the low-rank characteristic of spectrum tensor, which is relatively poor when there are a substantial number of RF transmitters or buildings.

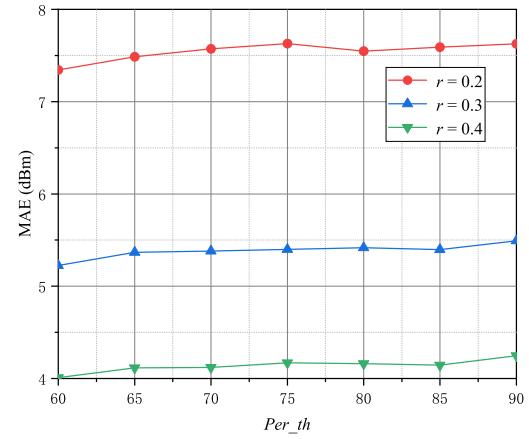


Fig. 8. Impact of Per_th on the construction performance.

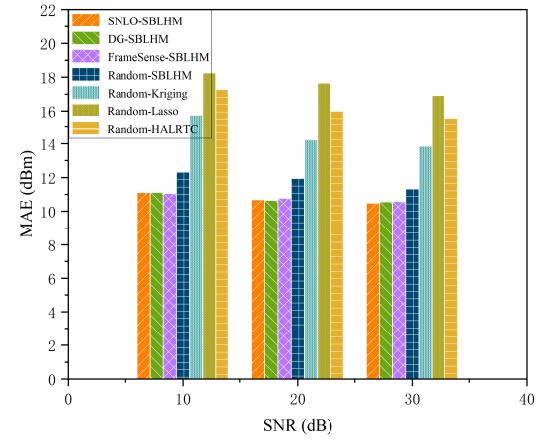


Fig. 9. Impact of SNR on the construction performance ($r = 0.05$).

V. CONCLUSION

This paper has addressed the issue of 3D REM construction with limited sample positions under unknown environments. We have formulated the 3D REM construction as the sparse sampling optimization and data recovery problems, by exploiting the sparsity of spectrum data and the characteristics of channel propagation. Then, a worst case SBL variance-based measurement matrix optimization mechanism has been developed to determine the required sampling quantity and sampling locations under the given recovery threshold. Meanwhile, PCA preprocessing has been performed on the sparse dictionary to further improve the sampling efficiency by reducing the matrix dimension. We have also proposed a hierarchical SBL model for the spectrum data recovery. The sparse signal has been recovered by SBL in the first layer and the shadow fading components have been estimated by GPR in the second layer. The construction performance of typical algorithms, i.e., SNLO-SBLHM, DG-SBLHM, FrameSense-SBLHM, Random-SBLHM, Random-Lasso, Random-HALRTC, and Random-Kriging have been

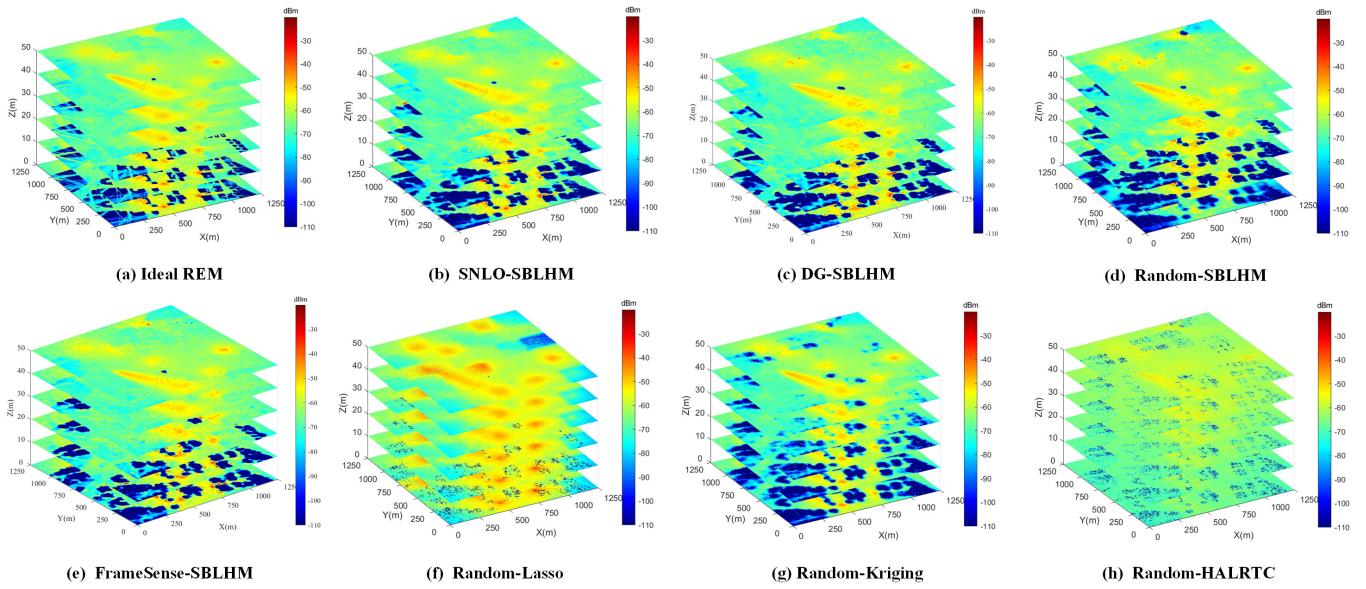


Fig. 10. Visualization of constructed 3D REMs with different algorithms ($r = 0.1$). (a) Ideal REM; and (b)-(h) SNLO-SBLHM, DG-SBLHM, Random-SBLHM, FrameSense-SBLHM, Random-Lasso, Random-Kriging, and Random-HALRTC, respectively.

fully compared and analyzed. The impacts of sparsity and proportion threshold on the recovery accuracy have also been studied. It has been shown that the proposed sampling and recovery methods are effective under unknown scenarios and superior in 3D REM construction against other methods. In the future work, we will apply the hierarchical recovery method on the UAV-assisted 3D REM mapping system [30], [60] and optimize the UAV trajectory based on the proposed sampling scheme.

APPENDIX A

DERIVATION OF EQUATIONS (19) AND (20)

To obtain (19) and (20), we need to compute marginal likelihood $p(\mathbf{t}|\boldsymbol{\alpha}, \beta)$ term in (18). We first derive the term Λ by exploiting the determinant identity

$$|\mathcal{A}| |\beta^{-1}\mathbf{I} + \Phi\mathcal{A}^{-1}\Phi^T| = |\beta^{-1}\mathbf{I}| |\mathcal{A} + \beta\Phi^T\Phi|. \quad (61)$$

We have

$$\begin{aligned} |\Lambda|^{-1/2} &= |\beta^{-1}\mathbf{I} + \Phi\mathcal{A}^{-1}\Phi^T|^{-1/2}, \\ &= \prod_M \beta^{1/2} |\mathcal{A}|^{1/2} |\mathcal{A} + \beta\Phi^T\Phi|^{-1/2}, \\ &= \beta^{M/2} \prod_{n=1}^N \alpha_n^{1/2} |\mathcal{A} + \beta\Phi^T\Phi|^{-1/2}. \end{aligned} \quad (62)$$

Then, the Woodbury inversion identity is employed to the term $(\Lambda)^{-1}$ as

$$\begin{aligned} (\Lambda)^{-1} &= (\beta^{-1}\mathbf{I} + \Phi\mathcal{A}^{-1}\Phi^T)^{-1}, \\ &= \beta\mathbf{I} - \beta\Phi(\beta\Phi^T\Phi + \mathcal{A})^{-1}\Phi^T\beta, \end{aligned} \quad (63)$$

we have

$$\begin{aligned} \mathbf{t}^T(\Lambda)^{-1}\mathbf{t} &= \mathbf{t}^T \left(\beta\mathbf{I} - \beta\Phi(\beta\Phi^T\Phi + \mathcal{A})^{-1}\Phi^T\beta \right) \mathbf{t}, \\ &= \beta\mathbf{t}^T\mathbf{t} - \beta\mathbf{t}^T\Phi(\beta\Phi^T\Phi + \mathcal{A})^{-1}\Phi^T\beta\mathbf{t}, \\ &= \beta\mathbf{t}^T\mathbf{t} - \beta\mathbf{t}^T\Phi\Sigma_\omega\Phi^T\beta\mathbf{t}, \\ &= \beta\mathbf{t}^T(\mathbf{t} - \Phi\mu), \end{aligned} \quad (64)$$

where we define

$$\begin{aligned} \mu &= \beta\Sigma_\omega\Phi^T\mathbf{t}, \\ \Sigma_\omega &= (\beta\Phi^T\Phi + \mathcal{A})^{-1}. \end{aligned} \quad (65)$$

With (14), (62), and (65), the marginal likelihood $p(\mathbf{t}|\boldsymbol{\alpha}, \beta)$ is

$$\begin{aligned} p(\mathbf{t}|\boldsymbol{\alpha}, \beta) &= \int p(\mathbf{t}|\boldsymbol{\omega}, \beta) p(\boldsymbol{\omega}|\boldsymbol{\alpha}) d\boldsymbol{\omega}, \\ &= (2\pi)^{-M/2} \beta^{M/2} \prod_{n=1}^N \alpha_n^{1/2} |\Sigma_\omega|^{1/2} \exp \left\{ -\frac{1}{2} \beta\mathbf{t}^T(\mathbf{t} - \Phi\mu) \right\}, \end{aligned} \quad (66)$$

Then, the weight posterior of $\boldsymbol{\omega}$ in (18) is

$$\begin{aligned} p(\boldsymbol{\omega}|\mathbf{t}, \boldsymbol{\alpha}, \beta) &= \frac{p(\mathbf{t}|\boldsymbol{\omega}, \boldsymbol{\alpha}, \beta) p(\boldsymbol{\omega}|\boldsymbol{\alpha})}{p(\mathbf{t}|\boldsymbol{\alpha}, \beta)}, \\ &= (2\pi)^{-N/2} |\Sigma_\omega|^{-1/2} \exp \left\{ -\frac{1}{2} (\boldsymbol{\omega} - \mu)^T (\Sigma_\omega)^{-1} (\boldsymbol{\omega} - \mu) \right\}, \\ &= \mathcal{N}(\boldsymbol{\omega}|\mu, \Sigma_\omega), \end{aligned} \quad (67)$$

where the mean and variance is equivalent to (19) and (20).

APPENDIX B

DERIVATION OF EQUATIONS (45) AND (47)

The objective (44) is equivalent to maximize the product of hyper-priors and marginal likelihood $p(\mathbf{t}|\boldsymbol{\alpha}, \beta)$, namely the

"evidence for the hyper-parameters". It is also known as the type-II maximum likelihood method, as

$$\begin{aligned} p(\mathbf{t}|\boldsymbol{\alpha}, \beta) &= \int p(\mathbf{t}|\boldsymbol{\omega}, \beta) p(\boldsymbol{\omega}|\boldsymbol{\alpha}) d\boldsymbol{\omega}, \\ &= (2\pi)^{-M/2} |\Lambda|^{-1/2} \exp \left\{ -\frac{1}{2} \mathbf{t}^T (\Lambda)^{-1} \mathbf{t} \right\}, \end{aligned} \quad (68)$$

where $\Lambda = \beta^{-1} \mathbf{I} + \Phi \mathcal{A}^{-1} \Phi^T$. By ignoring the irrelevant terms, we can obtain the objective function of the priors over hyper-parameters in the logarithmic case as

$$\begin{aligned} \mathcal{L}(\boldsymbol{\alpha}, \beta) &= -\frac{1}{2} \left\{ \log |\Lambda| + \mathbf{t}^T (\Lambda)^{-1} \mathbf{t} \right\} \\ &\quad + \sum_{n=1}^N (a_0 \log \alpha_n - b_0 \alpha_n) + c_0 \log \beta - d_0 \beta. \end{aligned} \quad (69)$$

Then, we calculate the derivative with respect to $\log \alpha_i$ and let $\partial \mathcal{L}(\boldsymbol{\alpha}, \beta) / \partial \log \alpha_i = 0$ to find the re-estimation rule of α_i to obtain (45). Similarly, calculating the derivative of (69) with respect to $\log \beta$ and let $\partial \mathcal{L}(\boldsymbol{\alpha}, \beta) / \partial \log \beta = 0$ to obtain the update rule of β in (47).

REFERENCES

- [1] G. Ding, Q. Wu, L. Zhang, Y. Lin, T. A. Tsiftsis, and Y.-D. Yao, "An amateur drone surveillance system based on the cognitive internet of things," *IEEE Communications Magazine*, vol. 56, no. 1, pp. 29–35, 2018.
- [2] A. Ahmad, S. Ahmad, M. H. Rehmani, and N. U. Hassan, "A survey on radio resource allocation in cognitive radio sensor networks," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 888–917, 2015.
- [3] B. Wang and K. R. Liu, "Advances in cognitive radio networks: A survey," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 1, pp. 5–23, 2011.
- [4] J. Liu, J. Yu, D. Niyato, R. Zhang, X. Gao, and J. An, "Covert ambient backscatter communications with multi-antenna tag," *IEEE Transactions on Wireless Communications*, vol. 22, no. 9, pp. 6199–6212, 2023.
- [5] X. Ding, K. Zhou, G. Li, K. Yang, X. Gao, J. Yuan, and J. An, "Customized joint blind frame synchronization and decoding methods for analog ldpc decoder," *IEEE Transactions on Communications*, vol. 72, no. 2, pp. 756–770, 2024.
- [6] P. Kolodzy, "Spectrum policy task force," *FCC, Washington, DC, USA, Tech. Rep. 02-135*, Nov 2002.
- [7] Y. Huang, X. Zhu, and Q. Wu, "Intelligent spectrum anti-jamming with cognitive software-defined architecture," *IEEE Systems Journal*, vol. 17, no. 2, pp. 2686–2697, 2023.
- [8] Y. Huang, H. Cui, Y. Hou, C. Hao, W. Wang, Q. Zhu, J. Li, Q. Wu, and J. Wang, "Space-based electromagnetic spectrum sensing and situation awareness," *Space: Science & Technology*, vol. 4, p. 0109, 2024.
- [9] H. B. Yilmaz, T. Tugcu, F. Alagöz, and S. Bayhan, "Radio environment map as enabler for practical cognitive radio networks," *IEEE Communications Magazine*, vol. 51, no. 12, pp. 162–169, 2013.
- [10] C. He, Y. Dong, and Z. J. Wang, "Radio map assisted multi-UAV target searching," *IEEE Transactions on Wireless Communications*, vol. 22, no. 7, pp. 4698–4711, 2023.
- [11] J. Pan, N. Ye, H. Yu, T. Hong, S. Al-Rubaye, S. Mumtaz, A. Al-Dulaimi, and I. Chih-Lin, "AI-driven blind signature classification for IoT connectivity: A deep learning approach," *IEEE Transactions on Wireless Communications*, vol. 21, no. 8, pp. 6033–6047, 2022.
- [12] D. Romero and S.-J. Kim, "Radio map estimation: A data-driven approach to spectrum cartography," *IEEE Signal Processing Magazine*, vol. 39, no. 6, pp. 53–72, 2022.
- [13] G. Ding, J. Wang, Q. Wu, Y. Yao, F. Song, and T. A. Tsiftsis, "Cellular-base-station-assisted device-to-device communications in TV white space," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 1, pp. 107–121, 2016.
- [14] F. Shen, Z. Wang, G. Ding, K. Li, and Q. Wu, "3D compressed spectrum mapping with sampling locations optimization in spectrum-heterogeneous environment," *IEEE Transactions on Wireless Communications*, vol. 21, no. 1, pp. 326–338, 2022.
- [15] M. Shi, K. Yang, D. Niyato, H. Yuan, H. Zhou, and Z. Xu, "The meta distribution of SINR in UAV-assisted cellular networks," *IEEE Transactions on Communications*, vol. 71, no. 2, pp. 1193–1206, 2023.
- [16] Z. Na, Y. Liu, J. Shi, C. Liu, and Z. Gao, "Uav-supported clustered noma for 6g-enabled internet of things: Trajectory planning and resource allocation," *IEEE Internet of Things Journal*, vol. 8, no. 20, pp. 15 041–15 048, 2021.
- [17] C. Suarez-Rodriguez, Y. He, and E. Dutkiewicz, "Theoretical analysis of rem-based handover algorithm for heterogeneous networks," *IEEE Access*, vol. 7, pp. 96 719–96 731, 2019.
- [18] S. Bi, J. Lyu, Z. Ding, and R. Zhang, "Engineering radio maps for wireless resource management," *IEEE Wireless Communications*, vol. 26, no. 2, pp. 133–141, 2019.
- [19] S. K. Sharma, E. Lagunas, S. Chatzinotas, and B. Ottersten, "Application of compressive sensing in cognitive radio communications: A survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1838–1860, 2016.
- [20] J. A. Bazerque and G. B. Giannakis, "Distributed spectrum sensing for cognitive radio networks by exploiting sparsity," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1847–1862, 2010.
- [21] M. E. Tipping, "Sparse Bayesian Learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, 2001.
- [22] X. Pang, M. Sheng, N. Zhao, J. Tang, D. Niyato, and K.-K. Wong, "When UAV meets IRS: Expanding air-ground networks via passive reflection," *IEEE Wireless Communications*, vol. 28, no. 5, pp. 164–170, 2021.
- [23] J. Chen and D. Gesbert, "Efficient local map search algorithms for the placement of flying relays," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 1305–1319, 2020.
- [24] E. Dall'Anese, S.-J. Kim, and G. B. Giannakis, "Channel gain map tracking via distributed kriging," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 3, pp. 1205–1211, 2011.
- [25] S. Shrestha, X. Fu, and M. Hong, "Deep spectrum cartography: Completing radio map tensors using learned neural models," *IEEE Transactions on Signal Processing*, vol. 70, pp. 1170–1184, 2022.
- [26] X. Han, L. Xue, Y. Xu, and Z. Liu, "A two-phase transfer learning-based power spectrum maps reconstruction algorithm for underlay cognitive radio networks," *IEEE Access*, vol. 8, pp. 81 232–81 245, 2020.
- [27] H. B. Yilmaz and T. Tugcu, "Location estimation-based radio environment map construction in fading channels," *Wireless communications & mobile computing*, vol. 15, no. 3, pp. 561–570, 1 2015.
- [28] M. Pesko, T. Javornik, L. Vidmar, A. Kosir, M. Stular, and M. Mohorcic, "The indirect self-tuning method for constructing radio environment map using omnidirectional or directional transmitter antenna," *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 1, pp. 1–12, 2015.
- [29] Y. Teganya and D. Romero, "Deep completion autoencoders for radio map estimation," *IEEE Transactions on Wireless Communications*, vol. 21, no. 3, pp. 1710–1724, 2022.
- [30] Q. Zhu, Y. Zhao, Y. Huang, Z. Lin, L. H. Wang, Y. Bai, T. Lan, F. Zhou, and Q. Wu, "Demo abstract: An UAV-based 3D spectrum real-time mapping system," in *Proceedings of IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2022, pp. 1–2.
- [31] Y. Zhao, Q. Zhu, Z. Lin, L. Guo, Q. Wu, J. Wang, and W. Zhong, "Temporal prediction for spectrum environment maps with moving radiation sources," *IET Communications*, vol. 17, no. 5, pp. 538–548, 2023.
- [32] H. Sun and J. Chen, "Propagation map reconstruction via interpolation assisted matrix completion," *IEEE Trans. Signal Process.*, vol. 70, pp. 6154–6169, 2022.
- [33] J. D. Naranjo, A. Ravanshid, I. Viering, R. Halfmann, and G. Bauch, "Interference map estimation using spatial interpolation of mdt reports in cognitive radio networks," in *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC)*, 2014, pp. 1496–1501.
- [34] K. Sato and T. Fujii, "Kriging-based interference power constraint: Integrated design of the radio environment map and transmission power," *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 1, pp. 13–25, 2017.
- [35] B. Fuchs, L. Le Coq, and M. D. Migliore, "On the interpolation of electromagnetic near field without prior knowledge of the radiating source," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 7, pp. 3568–3574, 2017.
- [36] M. Tang, G. Ding, Q. Wu, Z. Xue, and T. A. Tsiftsis, "A joint tensor completion and prediction scheme for multi-dimensional spectrum map construction," *IEEE Access*, vol. 4, pp. 8044–8052, 2016.
- [37] D.-H. Huang, S.-H. Wu, W.-R. Wu, and P.-H. Wang, "Cooperative radio source positioning and power map reconstruction: A Sparse

- Bayesian Learning Approach.” *IEEE Transactions on Vehicular Technology*, vol. 64, no. 6, pp. 2318–2332, 2015.
- [38] J. Wang, Q. Zhu, Z. Lin, Q. Wu, Y. Huang, X. Cai, W. Zhong, and Y. Zhao, “Sparse Bayesian Learning-Based 3-D Radio Environment Map Construction—Sampling Optimization, Scenario-Dependent Dictionary Construction, and Sparse Recovery,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 10, no. 1, pp. 80–93, 2024.
- [39] R. Shrestha, D. Romero, and S. P. Chepuri, “Spectrum surveying: Active radio map estimation with autonomous UAVs,” *IEEE Transactions on Wireless Communications*, vol. 22, no. 1, pp. 627–641, 2023.
- [40] S. He and K. G. Shin, “Steering crowdsourced signal map construction via Bayesian compressive sensing,” in *Proceedings of IEEE Conference on Computer Communications*, 2018, pp. 1016–1024.
- [41] B. Hua, H. Ni, Q. Zhu, C.-X. Wang, T. Zhou, K. Mao, J. Bao, and X. Zhang, “Channel modeling for UAV-to-ground communications with posture variation and fuselage scattering effect,” *IEEE Transactions on Communications*, vol. 71, no. 5, pp. 3103–3116, 2023.
- [42] A. Zanella, “Best practice in RSS measurements and ranging,” *IEEE Communications Surveys & Tutorials*, vol. 18, no. 4, pp. 2662–2686, 2016.
- [43] X. Jiang, N. Li, Y. Guo, D. Yu, and S. Yang, “Localization of multiple RF sources based on Bayesian compressive sensing using a limited number of uavs with airborne RSS sensor,” *IEEE Sensors Journal*, vol. 21, no. 5, pp. 7067–7079, 2021.
- [44] A. Ranganathan, M.-H. Yang, and J. Ho, “Online sparse gaussian process regression and its applications,” *IEEE Transactions on Image Processing*, vol. 20, no. 2, pp. 391–404, 2011.
- [45] S. He and S.-H. G. Chan, “Towards crowdsourced signal map construction via implicit interaction of iot devices,” in *2017 14th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, 2017, pp. 1–9.
- [46] X. Wang, X. Wang, S. Mao, J. Zhang, S. C. G. Periaswamy, and J. Patton, “Deepmap: Deep gaussian process for indoor radio map construction and location estimation,” in *2018 IEEE Global Communications Conference (GLOBECOM)*, 2018, pp. 1–7.
- [47] M. Dashti, S. Yiu, S. Yousefi, F. Perez-Cruz, and H. Claussen, “Rssi localization with gaussian processes and tracking,” in *2015 IEEE Globecom Workshops (GC Wkshps)*, 2015, pp. 1–6.
- [48] Y. Xu, B. Zhang, G. Ding, B. Zhao, S. Li, and D. Guo, “Radio environment map construction based on spatial statistics and Bayesian hierarchical model,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 3, pp. 767–779, 2021.
- [49] E. Clark, T. Askham, S. L. Brunton, and J. Nathan Kutz, “Greedy sensor placement with cost constraints,” *IEEE Sensors Journal*, vol. 19, no. 7, pp. 2642–2656, 2019.
- [50] C. Jiang, Y. C. Soh, and H. Li, “Sensor placement by maximal projection on minimum eigenspace for linear inverse problems,” *IEEE Transactions on Signal Processing*, vol. 64, no. 21, pp. 5595–5610, 2016.
- [51] Y. Saito, T. Nonomura, K. Yamada, K. Nakai, T. Nagata, K. Asai, Y. Sasaki, and D. Tsubakino, “Determinant-based fast greedy sensor selection algorithm,” *IEEE Access*, vol. 9, pp. 68 535–68 551, 2021.
- [52] S. Joshi and S. Boyd, “Sensor selection via convex optimization,” *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 451–462, 2009.
- [53] W. Ma, X. Dang, Q. Cheng, and H. Zhu, “Sensor selection based on sparse sensing in the presence of sensor position error,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 6, pp. 8915–8930, 2023.
- [54] R. Vandebril, M. V. Barel, and N. Mastronardi, *Matrix computations and semiseparable matrices. Vol. 2: Eigenvalue and singular value methods*. Baltimore, USA: Johns Hopkins University Press, 2008.
- [55] F. Kiaee, H. Sheikhzadeh, and S. Eftekhari Mahabadi, “Relevance vector machine for survival analysis,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 3, pp. 648–660, 2016.
- [56] Z. Zhang, T.-P. Jung, S. Makeig, Z. Pi, and B. D. Rao, “Spatiotemporal Sparse Bayesian Learning with applications to compressed sensing of multichannel physiological signals,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 6, pp. 1186–1197, 2014.
- [57] J. Liu, P. Musalski, P. Wonka, and J. Ye, “Tensor completion for estimating missing values in visual data,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, pp. 208–220, 2013.
- [58] F. Shen, G. Ding, and Q. Wu, “Efficient Remote Compressed Spectrum Mapping in 3D spectrum-heterogeneous environment with inaccessible areas,” *IEEE Wireless Communications Letters*, vol. 11, no. 7, pp. 1488–1492, 2022.
- [59] J. Ranieri, A. Chebira, and M. Vetterli, “Near-optimal sensor placement for linear inverse problems,” *IEEE Transactions on Signal Processing*, vol. 62, no. 5, pp. 1135–1146, 2014.
- [60] K. Mao, Q. Zhu, Y. Qiu, X. Liu, M. Song, W. Fan, A. B. J. Kokkeler, and Y. Miao, “A UAV-aided real-time channel sounder for highly dynamic nonstationary A2G scenarios,” *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–15, 2023.