Energy-modified Leverage Sampling for Radio Map Construction via Matrix Completion

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

***Abstract*—This paper explores an energy-modified leverage sampling strategy for matrix completion in radio map construction. The main goal is to address potential identifiability issues in matrix completion with sparse observations by using a probabilistic sampling approach. Although conventional leverage sampling is commonly employed for designing sampling patterns, it often assigns high sampling probability to locations with low received signal strength (RSS) values, leading to a low sampling efficiency. Theoretical analysis demonstrates that the leverage score produces pseudo images of sources, and in the regions around the source locations, the leverage probability is asymptotically consistent with the RSS. Based on this finding, an energy-modified leverage probability-based sampling strategy is investigated for efficient sampling. Numerical demonstrations indicate that the proposed sampling strategy can decrease the normalized mean squared error (NMSE) of radio map construction by more than 10% for both matrix completion and interpolation-assisted matrix completion schemes, compared to conventional methods.**

***Index Terms*—Energy-modified leverage sampling, leverage score, sampling pattern, matrix completion, radio map.**

I. Introduction

Radio map construction is applied in various fields, such as wireless network planning [1], localization [2]–[8], UAV trajectory planning for low-altitude air-to-ground communications [9], and integrated communications and sensing for millimeter-wave and terahertz systems. Matrix and tensor completions are widely used to construct radio maps from sparse and limited samples [10]–[15]. Identifiability is crucial in these methods, as the matrix or tensor might not be completable with overly sparse measurements or improper sampling patterns [14], [16], [17]. Thus, optimizing the sampling pattern is essential for matrix or tensor completion.

This paper investigates sampling strategies for constructing radio maps through matrix completion, when some prior information is available. Prior information can often be present in various practical situations. In active sampling, a set of measurements is collected and analyzed to determine future sample locations. In power-constrained sensor networks, it may be preferable to activate only a limited number of sensors in each round and optimize the measurement locations accordingly for the next rounds. Moreover, in interpolation-assisted matrix completion [18]–[20], additional observations can be generated from a few measurements using interpolation, where optimizing the interpolation pattern is crucial.

To optimize for the sampling pattern, Bayesian method in [21] and dictionary learning method in [22] determined the sample locations in an iterative way, but these approaches did not utilize the online measurements to tune the sampling pattern on-the-fly. Apart from these methods, *leverage score* is commonly used to optimize the sampling pattern [16], [17], [23] for matrix completion. The leverage score is calculated from the singular vectors of the matrix, which can be roughly estimated assuming some prior information is available. It was shown in [16] that sampling locations are not identically informative, and therefore, uniform sampling is strictly sub-optimal. Instead, a biased sampling procedure that uses the leverage score to assess the "importance" of each observed element can recover the sparse matrix with high probability under less samples compared to uniform sampling. The work [17] proposed a two-phase sampling procedure for matrices, starting with leverage score estimation and followed by sampling for exact recovery, which requires substantially fewer samples than uniformly random sampling method to obtain a same accuracy.

However, we discover that leverage sampling, i.e., sampling based on the leverage probability formulated from the leverage score, is strictly sub-optimal for radio map construction, because it may allocate up to half of the measurements at locations where the RSS from the sources almost vanishes. Intuitively, those low RSS areas are far away from the source, and thus, less informative about the source. Specifically, leverage score [16], [17], [23] measures the "importance" of sampling a row or a column of a matrix, but *not* every entry in a row or a column has the same "importance" to be sampled. In this paper, we construct theoretical examples to show that the leverage score produces pseudo images of sources where the RSS from the sources diminishes more quickly than the leverage probability, especially when the area of the radio map scales up. This implies that the leverage score may not be a reliable metric for determining sampling probabilities in radio map construction. On the other hand, in the regions around the source locations, we show that the leverage probability is asymptotically consistent with the RSS value of the radio map.

The work was supported in part by the National Science Foundation of China (NSFC) under Grant No. 62171398, by the Basic Research Project No. HZQB-KCZYZ-2021067 of Hetao Shenzhen-HK S&T Cooperation Zone, by NSFC Grant No. 62293482, by the Shenzhen Science and Technology Program under Grant No. JCYJ20220530143804010, No. KJZD20230923115104009, and No. KQTD20200909114730003, by Guangdong Research Projects No. 2019QN01X895, No. 2017ZT07X152, and No. 2019CX01X104, by the Shenzhen Outstanding Talents Training Fund 202002, by the Guangdong Provincial Key Laboratory of Future Networks of Intelligence (Grant No. 2022B1212010001), by the National Key R&D Program of China with grant No. 2018YFB1800800, and by the Key Area R&D Program of Guangdong Province with grant No. 2018B030338001.

The authors are with the School of Science and Engineering (SSE), the Shenzhen Future Netowrk of Intelligence Institute (FNii-Shenzhen), and the Guangdong Provincial Key Laboratory of Future Networks of Intelligence, The Chinese University of Hong Kong, Shenzhen.

Based on these analysis, we propose a probabilistic sampling strategy based on the predicted RSS values and the leverage score. The method first constructs a rough estimate of the radio map via interpolation or low resolution matrix completion. Then, it computes the leverage score of each entry of the matrix to be completed. Finally, the sampling probability is formulated based on both the RSS and the leverage score. As a result, since leverage sampling completes the matrix with high probability, the proposed probabilistic sampling based on energy-modified leverage probability may also complete the matrix with high probability and less samples. We numerically show that the proposed energy-modified leverage sampling strategy substantially increases the accuracy of radio map construction by over 10% compared to baseline methods. Integrating this strategy with interpolation-assisted matrix completion [20] reduces the construction normalized mean squared error (NMSE) by more than 10% compared to baseline methods.

II. System Model

A. Propagation Model

Consider a propagation field that is excited by K sources located at $\boldsymbol{s}_k \in \mathcal{D}$, $k = 1, \cdots, K$, in a bounded area $\mathcal{D} \subset \mathbb{R}^2$. The signal emitted from the source is captured by M sensors with known locations $\boldsymbol{z}_m \in \mathbb{R}^2$, $m = 1, 2, \ldots, M$, in $\mathcal{D}$. The radio map to be constructed is modeled as

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

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 spatially correlated shadowing.

The strength of the signal measured by the mth sensor is given by $\gamma_m = \rho(\boldsymbol{z}_m) + \epsilon_m$, where ϵ_m is a random variable with zero mean and variance σ^2 to model the measurement noise.

We consider to discretize the target area $\mathcal{D}$ into $N \times N$ 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 radio map $\rho(\boldsymbol{z})$, where the (i, j)th entry is defined as $H_{ij} = \rho(\boldsymbol{c}_{ij})$.

Our goal is to analyze and develop probabilistic sampling strategies that obtain M measurements $\rho(\boldsymbol{z}_m)$ for the completion of the matrix $\boldsymbol{H}$.

B. Leverage Sampling

Given a rank-r matrix $\boldsymbol{H} \in \mathbb{R}^{N \times N}$, the singular value decomposition (SVD) of $\boldsymbol{H}$ is defined as $\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}$. The leverage scores μ_i for the ith row and ν_j for the jth column are respectively defined as

$$\mu_i = N\|\boldsymbol{U}^{\mathrm{T}}\boldsymbol{e}_i\|_2^2/r \tag{2}$$

$$\nu_j = N\|\boldsymbol{V}^{\mathrm{T}}\boldsymbol{e}_j\|_2^2/r \tag{3}$$

where $\boldsymbol{e}_i$ is unit vector with ith element equals to 1.

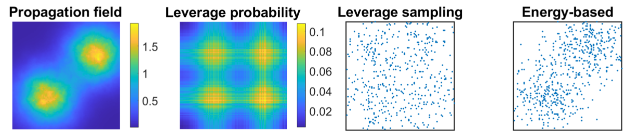

Figure 1. Visible plot of propagation field, leverage probability p_{ij}, leverage sampling, and energy-based sampling.

In the leverage sampling, one independently samples the grids based on leverage probability p_{ij} [17], [23] which is calculated as follows:

$$p_{ij} = \min\{C(\mu_i + \nu_j)r\log^2(2N)/N, 1\} \tag{4}$$

where C is a constant. It is shown in [17] that under such a probabilistic sampling strategy, an arbitrary rank-r matrix can be exactly recovered from $O(Nr\log^2(N))$ observed elements with high probability using nuclear norm minimization.

III. Energy-modified Leverage Probability-based Probabilistic Sampling

In this section, we study a specific example to illustrate the possible pseudo images in the conventional leverage sampling based on (4). Then, for the informative region of the propagation field, we show the consistency of leverage probability and the entry of $\boldsymbol{H}$.

A. Existence of Pseudo Images

For illustration purpose, we analyze the case of two sources, $K = 2$, where each source generates a rank-1 propagation field, denoted as $\boldsymbol{H}^{(k)}$. Specifically, assume $g_k(d)$ in (1) takes the form of $g_k(d) = \alpha e^{-\beta d^2}$, and there is no shadowing, i.e., $\zeta(\boldsymbol{z}) = 0$. To see that $\boldsymbol{H}^{(k)}$ is rank-1, we note from (1) that

$$\begin{aligned} H_{ij}^{(k)} &= \alpha e^{-\beta((s_x^{(k)} - x_i)^2 + (s_y^{(k)} - y_j)^2)} \\ &= \alpha e^{-\beta(s_x^{(k)} - x_i)^2} e^{-\beta(s_y^{(k)} - y_j)^2} \end{aligned} \tag{5}$$

where x_i and y_j are the coordinates of the rows and the columns, respectively. As a result, the matrix $\boldsymbol{H}^{(k)}$ can be written as the outer product of two rank-1 vectors $\boldsymbol{u}_k = e^{-\beta((s_x^{(k)} - \boldsymbol{x})^2)}$ and $\boldsymbol{v}_k = e^{-\beta(s_y^{(k)} - \boldsymbol{y})^2}$, scaled by α, where $\boldsymbol{x}$ and $\boldsymbol{y}$ are the vectors corresponding to the coordinates of the rows and the columns, respectively.

Without loss of generality, assume the first source locates at the origin, and the second source locates at $\boldsymbol{s}_2 = (L_1, L_1)$. We investigate the leverage sampling probability p_{ij} defined in (4) over all grid points (i, j). As shown in Fig. 1, four "sources" appears according to the values of p_{ij}, where two of them are merely pseudo images. In the pseudo images, although p_{ij} is non-zero, H_{ij} is essentially zero, indicating that sampling at (i, j) has negligible value for radio map construction. Pseudo images exist because, from (4), the leverage probability at (i, j) equals to the leverage score of the ith row plus the leverage score of the jth column of the matrix. Therefore, in an extreme case, K sources may create $K^2 - K$ pseudo images of the sources.

To analytically investigate the pseudo images, we define a region $\mathcal{I}(L_1,\delta)=\{(i,j), i\in[L_1+1-\delta, L_1+1+\delta], j\in[1,1+\delta]\}$. The following theorem implies that $\mathcal{I}(L_1,\delta)$ is one of the regions of a pseudo image of the source, where the RSS asymptotically vanishes in $\mathcal{I}(L_1,\delta)$.

Theorem 1 (Pseudo Image). *Consider that $\delta<L_1/2$. For all $(i,j)\in\mathcal{I}(L_1,\delta)$, $H_{ij}/p_{ij}\to 0$ as $\beta\to\infty$.*

Proof. Denote $\boldsymbol{H}^{(k)}$ as $\boldsymbol{H}^{(k)}=\sigma_k\boldsymbol{u}_k\boldsymbol{v}_k^{\mathrm{T}}$ where $\boldsymbol{u}_k$, $\boldsymbol{v}_k$ are singular vectors for the propagation field $\boldsymbol{H}^{(k)}$ contributed from kth source. As a result, $\boldsymbol{H}=\boldsymbol{H}^{(1)}+\boldsymbol{H}^{(2)}$. The SVD of $\boldsymbol{H}$ is given by $\boldsymbol{H}=\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}$, where $\boldsymbol{U}=[(\boldsymbol{u}_1+\boldsymbol{u}_2)/\|\boldsymbol{u}_1+\boldsymbol{u}_2\|_2,(\boldsymbol{u}_1-\boldsymbol{u}_2)/\|\boldsymbol{u}_1-\boldsymbol{u}_2\|_2]$, $\boldsymbol{V}=[(\boldsymbol{v}_1+\boldsymbol{v}_2)/\|\boldsymbol{v}_1+\boldsymbol{v}_2\|_2,(\boldsymbol{v}_1-\boldsymbol{v}_2)/\|\boldsymbol{v}_1-\boldsymbol{v}_2\|_2]$. From (5), let $\sigma_1=\alpha$, there are $\boldsymbol{u}_1=\boldsymbol{v}_1=e^{-\beta\boldsymbol{x}^2}$ and $\boldsymbol{u}_2=\boldsymbol{v}_2=e^{-\beta(\boldsymbol{x}-L_1)^2}$, for $\boldsymbol{x}=[0,1,\cdots,N-1]$.

Then, from (2) and (3), there are

$$\mu_i\geq N\frac{2(\|\boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{e}_i\|_2^2+\|\boldsymbol{u}_2^{\mathrm{T}}\boldsymbol{e}_i\|_2^2)}{\|\boldsymbol{u}_1+\boldsymbol{u}_2\|_2^2}/r \tag{6}$$

and

$$\nu_j\geq N\frac{2(\|\boldsymbol{v}_1^{\mathrm{T}}\boldsymbol{e}_j\|_2^2+\|\boldsymbol{v}_2^{\mathrm{T}}\boldsymbol{e}_j\|_2^2)}{\|\boldsymbol{v}_1+\boldsymbol{v}_2\|_2^2}/r. \tag{7}$$

Thus, for $(i,j)\in\mathcal{I}(L_1,\delta)$, there are $H_{ij}=\sum_{k=1}^{2}H_{ij}^{(k)}\leq 2\alpha e^{-\beta(L_1-\delta)^2}\triangleq\bar{H}_{ij}$ and

$$\begin{aligned}p_{ij}&=\min\{C(\mu_i+\nu_j)r\log^2(2N)/N,1\}\\&\geq 2C\left(\frac{e^{-\beta(L_1-\delta)^2}}{\|\boldsymbol{u}_1+\boldsymbol{u}_2\|_2^2}+\frac{e^{-\beta\delta^2}}{\|\boldsymbol{v}_1+\boldsymbol{v}_2\|_2^2}\right)\log^2(2N)\triangleq\bar{p}_{ij}.\end{aligned}$$

As a consequence, there are

$$\begin{aligned}\frac{H_{ij}}{p_{ij}}&\leq\frac{\bar{H}_{ij}}{\bar{p}_{ij}}=\frac{2\alpha e^{-\beta(L_1-\delta)^2}}{2C\left(\frac{e^{-\beta(L_1-\delta)^2}}{\|\boldsymbol{u}_1+\boldsymbol{u}_2\|_2^2}+\frac{e^{-\beta\delta^2}}{\|\boldsymbol{v}_1+\boldsymbol{v}_2\|_2^2}\right)\log^2(2N)}\\&=\frac{\alpha}{C\log^2(2N)}\left(\frac{1}{\frac{1}{\|\boldsymbol{u}_1+\boldsymbol{u}_2\|_2^2}+\frac{e^{-\beta\delta^2}}{\|\boldsymbol{v}_1+\boldsymbol{v}_2\|_2^2e^{-\beta(L_1-\delta)^2}}}\right)\to 0\end{aligned}$$

as $\beta\to\infty$. $\square$

According to Theorem 1, when $\beta\to\infty$, corresponding to a sharp propagation field, both p_{ij} and H_{ij} tend to 0, but H_{ij} tends to 0 more rapidly than p_{ij}. This implies that if one uses the leverage probability p_{ij} in (4) as the sampling probability, the probability p_{ij} in $\mathcal{I}(L_1,\delta)$ may still be non-negligible but the measurement H_{ij}, which is essentially 0, contains almost no information of the propagation field. As a result, sampling at the pseudo image is highly inefficient.

While Theorem 1 examines the case of increasing β to a large value resulting in a sharp propagation field, this situation is analogous to increasing the distance L_1 while keeping β constant. Likewise, pseudo images appear where the RSS of the sources vanishes, and sampling at these pseudo images is inefficient.

B. Consistency of the Leverage Probability p_{ij} and H_{ij}

For the regions around the sources, the leverage probability p_{ij} defined in (4) is essentially consistent with H_{ij}.

Define $\mathcal{J}(L_1,\delta)=\{(i,j), i,j\in[L_1+1-\delta,L_1+1+\delta]\}$. The following theorem implies that $\mathcal{J}(L_1,\delta)$ is one of the regions of high importance, where the leverage probability is strongly correlated with $\boldsymbol{H}$ in $\mathcal{J}(L_1,\delta)$.

Theorem 2 (Consistency). *For $(i,j)\in\mathcal{J}(L_1,\delta)$, $H_{ij}/p_{ij}\to C'/C$ with $C'=\alpha/(4\log^2(2N))$, as $\delta\to 0$.*

Proof. From (5), there are

$$H_{ij}=\alpha e^{-\beta(i-1)^2}e^{-\beta(j-1)^2}+\alpha e^{-\beta(i-L_1-1)^2}e^{-\beta(j-L_1-1)^2}.$$

For the leverage probability p_{ij} defined in (4), from (6) and (7), there are

$$\begin{aligned}p_{ij}=2C\log^2(2N)(&e^{-\beta 2(i-1)^2}+e^{-\beta 2(i-L_1-1)^2}\\&+e^{-\beta 2(j-1)^2}+e^{-\beta 2(j-L_1-1)^2}).\end{aligned} \tag{8}$$

Then, for $(i,j)\in\mathcal{J}(L_1,\delta)$, and $\delta\to 0$, there are $e^{-\beta 2(i-1)^2}/e^{-\beta 2(j-1)^2}\to 1$ and $(e^{-\beta 2(i-1)^2}+e^{-\beta 2(j-1)^2})/2e^{-\beta(i-1)^2}e^{-\beta(j-1)^2}\to 1$. Thus, there are $H_{ij}/p_{ij}\to C'/C$, where $C'=\alpha/(4\log^2(2N))$. $\square$

Thus, for the region of high importance, the H_{ij} itself is essentially consistent as the leverage probability p_{ij}.

The results in Theorems 1 and 2 motivate the proposed probabilistic sampling based on the energy-modified leverage probability $\tilde{p}_{ij}=C_1\sqrt{H_{ij}p_{ij}}$, where p_{ij} is from (4) and C_1 is a constant that depends on the matrix dimension and the rank structure.

It follows that at the region of high importance $\mathcal{J}$, there is $\tilde{p}_{ij}\propto p_{ij}$, according to Theorem 2. Therefore, $\tilde{p}_{ij}$ is essentially the leverage probability in $\mathcal{J}$. In the pseudo images $\mathcal{I}$ where the source signal almost vanishes, there is $\tilde{p}_{ij}\ll p_{ij}$, thus, the sampling probability is significantly reduced, and therefore, frequently sample at locations where the source signal vanishes can be avoided.

C. Implementation of Energy-modified Leverage Sampling

Suppose that the expected number of measurements is M. First, we interpolate a matrix $\hat{\boldsymbol{H}}$, using ιM measurements uniformly random taken in the area of interest, for a small $\iota<1$, as the prior information. Algorithms such as k-nearest neighbor (KNN), Kriging, or regression can be used for the construction of $\hat{\boldsymbol{H}}$. Then, we calculate the SVD of $\hat{\boldsymbol{H}}=\hat{\boldsymbol{U}}\hat{\boldsymbol{\Sigma}}\hat{\boldsymbol{V}}^{\mathrm{T}}$ to obtain the leverage scores, $\hat{\mu}_i$ and $\hat{\nu}_j$ as in (2)–(3), for the ith row and jth column, and we further obtain $\hat{p}_{ij}=\hat{\mu}_i+\hat{\nu}_j$.

Next, we establish the energy-modified leverage probability $\tilde{p}_{ij}=C_1\sqrt{\hat{H}_{ij}\hat{p}_{ij}}$, where $C_1=(1-\iota)M(\sum_{i,j}\sqrt{\hat{H}_{ij}\hat{p}_{ij}})$. Then, in the second round of sampling, we independently sample each grid according to the probability $\tilde{p}_{ij}$. One can verify that the expected number of samples equals to $(1-\iota)M$. Finally, matrix completion is performed via nuclear norm minimization using the total M measurements obtained from the two rounds of sampling, to obtain the reconstructed ratio map $\tilde{\boldsymbol{H}}$.

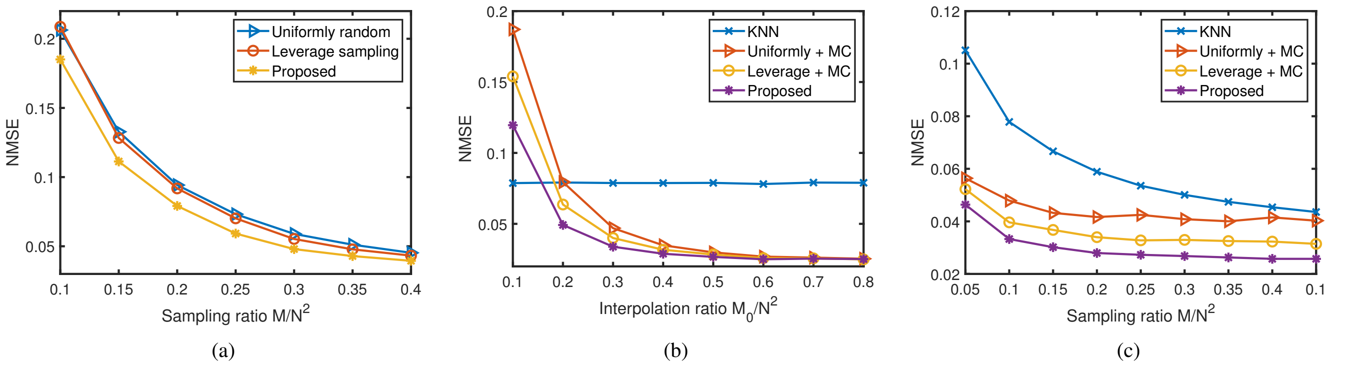

Figure 2. (a) Construction NMSE of matrix completion under different sampling strategy versus different sampling ratio M/N^2. (b) Construction NMSE of interpolation-assisted matrix completion under different interpolation radio M_0/N^2. (c) Construction NMSE of interpolation-assisted matrix completion under different sampling ratio M/N^2. The dimension N of matrix is fixed as 100, and we only vary M or M_0 in the simulation.

IV. Numerical Results

We adopt model (1) to simulate the radio map in an $L \times L$ area for $L = 2$ kilometers, $K = 3$ sources, where $g_k(d) = Pd^{-1.5}A(f)^{-d}$, with parameter $A(f) = 0.8$, corresponding to an empirical energy field of underwater acoustic signal at frequency $f = 5$ kHz, $d = \sqrt{x^2 + y^2 + h^2}$ represents the distance from the source, (x, y) is the coordinate, and $h = 400$ meters is the depth of interest [24]. The shadowing component in log-scale $\log_{10}\zeta$ is modeled using a Gaussian process with zero mean and auto-correlation function $\mathbb{E}\{\log_{10}\zeta(\boldsymbol{z}_i)\log_{10}\zeta(\boldsymbol{z}_j)\} = \sigma_s^2 \exp(-||\boldsymbol{z}_i - \boldsymbol{z}_j||_2/d_{\mathrm{c}})$, in which $d_{\mathrm{c}} = 200$ meters, $\sigma_s^2 = 1$. We choose $\epsilon \sim \mathcal{N}(0, \sigma)$ with $\sigma = 0.1$ to model the measurement noise.

The NMSE of the reconstructed radio map is employed for performance evaluation, which is calculated through $||\tilde{\boldsymbol{H}} - \boldsymbol{H}||_F^2/||\boldsymbol{H}||_F^2$. We evaluate the performance of radio map construction under matrix dimension $N = 100$.

We first test the energy-modified leverage sampling under the matrix completion scheme in [25], and choose M measurements with the sampling ratio $M/N^2 = 10\% - 40\%$. We compare three sampling schemes, the proposed energy-modified leverage sampling, the uniformly random sampling, and the leverage sampling. For the leverage sampling, we first uniformly random sample ιM grids to obtain p_{ij}, then we sample $(1-\iota)M$ grids accordingly, where we choose $\iota = 0.7$. Fig. 2 (a) shows that the proposed energy-modified leverage sampling outperforms the leverage sampling and uniformly random sampling in construction NMSE with larger than 10% improvement under small sampling ratio.

We then test the effectiveness of the proposed energy-modified leverage sampling in interpolation-assisted matrix completion [20]. In this method, we choose $\iota = 1$ to estimate $\hat{H}_{ij}$ and $\hat{p}_{ij}$, based on all the measurements M, then, we interpolate M_0 grids based on the probability $\tilde{p}_{ij}$ with $C_1 = M_0(\sum_{i,j}\sqrt{\hat{H}_{ij}\hat{p}_{ij}})$. After that, we perform matrix completion. We compare this method to the following baseline methods. Baseline 1: Uniformly random interpolation followed by matrix completion (Uniformly + MC). In this approach, we uniformly and randomly interpolate grids based on the M measurements, and then, use the singular value thresholding (SVT) algorithm to solve the matrix completion problem. Baseline 2: Leverage interpolation followed by matrix completion (Leverage + MC). This approach interpolates each grids independently, based on the leverage probability p_{ij}. Baseline 3: KNN method, with $k = 3$. The KNN has a computational complexity of $O(MN^2)$. The computational complexity for Baselines 1 and 2, as well as the proposed method, is $O(MM_0 + N^3)$. This similarity arises because all methods incorporate both interpolation and matrix completion components.

To show the influence of interpolation ratio M_0/N^2 on the performance of the proposed method, we choose the number of measurements $M = 1000$ to satisfy the sampling ratio $M/N^2 = 10\%$ and vary the interpolation ratio $M_0/N^2 = 10\% - 80\%$. Simulation results in Fig. 2 (b) demonstrates that under the same sampling ratio, the proposed energy-modified leverage sampling significantly outperforms the baseline methods by more than 20% under a low interpolation ratio. The proposed method only needs a interpolation ratio of among 50% to attain the best performance.

To show the influence of the sampling ratio M/N^2 on the performance of the proposed method, we choose the number of measurements M to satisfy the sampling ratio M/N^2 ranging from $5\% - 45\%$ and set the number of grids to be interpolated $M_0 = 3000$ with $M_0/N^2 = 30\%$. Fig. 2 (c) illustrates that interpolation based on energy-modified leverage probability exhibits an over 10% improvement of construction accuracy, as compared to the baseline methods. It demonstrates superior performance compared to the conventional matrix completion approach, which lacks a designed interpolation pattern.

V. Conclusion

This paper proposed an energy-modified leverage sampling method. It was theoretically shown that the leverage scores were not efficient since the existence of pseudo images. In addition, the leverage probability p_{ij} was shown to be consistent with the RSS at the regions around source locations. Then, an energy-modified leverage probability $\tilde{p}_{ij}$ was formulated. Simulation results demonstrated the proposed method have over 10% improvement in construction NMSE.

References

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

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

[4] 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.

[5] Z. Xing, J. Chen, and Y. Tang, "Integrated segmentation and subspace clustering for RSS-based localization under blind calibration," in *Proc. IEEE Global Telecomm. Conf.*, 2022, pp. 5360–5365.

[6] Z. Xing and J. Chen, "Constructing indoor region-based radio map without location labels," *IEEE Trans. Signal Process.*, pp. 1–16, 2024.

[7] R. Liu, B. P. L. Lau, K. Ismail, A. Chathuranga, C. Yuen, S. X. Yang, Y. L. Guan, S. Mao, and U.-X. Tan, "Exploiting radio fingerprints for simultaneous localization and mapping," *IEEE Pervasive Comput.*, 2023.

[8] R. Liu, S. H. Marakkalage, M. Padmal, T. Shaganan, C. Yuen, Y. L. Guan, and U.-X. Tan, "Collaborative slam based on wifi fingerprint similarity and motion information," *IEEE Internet Things J.*, vol. 7, no. 3, pp. 1826–1840, 2020.

[9] W. Liu and J. Chen, "UAV-aided radio map construction exploiting environment semantics," *IEEE Trans. Wireless Commun.*, vol. 22, no. 9, pp. 6341–6355, 2023.

[10] L. Ma, W. Zhao, Y. Xu, and C. Li, "Radio map efficient building method using tensor completion for WLAN indoor positioning system," in *Proc. IEEE Int. Conf. Commun.*, 2018, pp. 1–6.

[11] 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.

[12] Y. Zhang and L. Ma, "Radio map crowdsourcing update method using sparse representation and low rank matrix recovery for WLAN indoor positioning system," *IEEE Wireless Commun. Lett.*, vol. 10, no. 6, pp. 1188–1191, 2021.

[13] H. Sun and J. Chen, "Integrated interpolation and block-term tensor decomposition for spectrum map construction," 2024.

[14] 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.

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

[16] X. Huang, W. Liu, B. Du, and D. Tao, "Leveraged matrix completion with noise," *IEEE Trans. Cybern.*, pp. 1–11, 2023.

[17] 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.

[18] 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.

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

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

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

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

[23] A. Eftekhari, M. B. Wakin, and R. A. Ward, "MC2: A two-phase algorithm for leveraged matrix completion," *Information and Inference: A Journal of the IMA*, vol. 7, no. 3, pp. 581–604, 2018.

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

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