

Adaptive Sampling of RF Fingerprints for Fine-Grained Indoor Localization

Xiao-Yang Liu, Shuchin Aeron, *Senior Member, IEEE*, Vaneet Aggarwal, *Senior Member, IEEE*, Xiaodong Wang, *Fellow, IEEE*, and Min-You Wu, *Senior Member, IEEE*

Abstract—Indoor localization is a supporting technology for a broadening range of pervasive wireless applications. One promising approach is to locate users with radio frequency fingerprints. However, its wide adoption in real-world systems is challenged by the time- and manpower-consuming site survey process, which builds a fingerprint database *a priori* for localization. To address this problem, we visualize the 3-D RF fingerprint data as a function of locations (x-y) and indices of access points (fingerprint), as a *tensor* and use tensor algebraic methods for an *adaptive* tubal-sampling of this fingerprint space. In particular, using a recently proposed tensor algebraic framework in [1], we capture the complexity of the fingerprint space as a low-dimensional tensor-column space. In this formulation, the proposed scheme exploits adaptivity to identify reference points which are highly informative for learning this low-dimensional space. Further, under certain incoherency conditions, we prove that the proposed scheme achieves bounded recovery error and near-optimal sampling complexity. In contrast to several existing work that rely on random sampling, this paper shows that adaptivity in sampling can lead to significant improvements in localization accuracy. The approach is validated on both data generated by the ray-tracing indoor model which accounts for the floor plan and the impact of walls and the real world data. Simulation results show that, while maintaining the same localization accuracy of existing approaches, the amount of samples can be cut down by 71 percent for the high SNR case and 55 percent for the low SNR case.

Index Terms—Fine-grained indoor localization, RF fingerprint, adaptive sampling, tubal-sampling, tensor completion

1 INTRODUCTION

THE availability of real-time high-accuracy location-awareness in indoor environments is a key enabler for a wide range of pervasive wireless applications, such as pervasive healthcare [2] and smart-space interactive environments [3], [4]. More recently, location-based services have been used in airports, shopping malls, supermarkets, stadiums, office buildings, and homes [5], [6]. The economic impact of the indoor localization market is forecasted by ABI Research to reach \$5 billion in 2018 [7].

Indoor localization systems generally follow three approaches, viz., cell-based approach, model-based approach, and fingerprint-based approach. For cell-based approach [8], [9], a user's location is given by the access point to which it is connected and the localization error depends on the communication range and the distances between the access points. This approach, however, can only provide coarse grained localization. On the other hand model-based schemes [10], [11] exploit angle of arrival (AoA), time or time

difference of arrival (ToA, TDoA), or received signal strength (RSS) from access points. However, such model-based approach is essentially limited by the following factors: 1) low transmission power, 2) high attenuation caused by walls and furnitures, 3) complicated surface reflections, and 4) unpredictable dynamics such as disturbance by human movements.

We will focus on the fingerprint-based approach, which can be classified into three categories, namely, RF (radio frequency) fingerprint-based, non-RF fingerprint-based, and cross-technology-based. For example, RADAR [12] is considered as the first RF fingerprint-based system; Google Indoor Map [13] and Wi-FiSLAM [14] are two widely used industrial apps, while the former uses RF fingerprints, the latter incorporates user trajectories, inertial information and accelerometer; IndoorAtlas [15] utilizes the magnetic field map; SurroundSense [16] exploits ambient attributes (sound, light, etc) including RF fingerprints. The advantages of RF fingerprint-based approach include: 1) It is a passive approach that exploits Wi-Fi access points already in most buildings thus needs no extra infrastructure deployment; 2) The RSS values are provided by off-the-shelf Wi-Fi- or Zigbee-compatible devices; 3) The flourishing smartphone market indicates its upcoming wide use; 4) It assumes no radio propagation model thus is more practical than the model-based approach.

In this paper, we consider the RF fingerprint-based approach. It is a two-phase approach, a training phase (site survey) and an operating phase (location query). The fingerprints obtained are averaged over time to counteract the effects of fading, and thus usually change in a building with change of walls and/or furniture. Assuming that the fingerprint data remains statistically stable, i.e., the mean RSS

- X. Liu and X. Wang are with the Department of Electrical Engineering, Columbia University, New York, NY 10027. E-mail: yanglet@sjtu.edu.cn, wangx@ee.columbia.edu.
- S. Aeron is with the Department of Electrical and Computer Engineering, Tufts University, Medford, MA 02155. E-mail: shuchin@ece.tufts.edu.
- V. Aggarwal is with the School of Industrial Engineering, Purdue University, West Lafayette, IN 47907. E-mail: vaneet@purdue.edu.
- X. Liu and M. Wu are with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China, 200240. E-mail: mwu@sjtu.edu.cn.

Manuscript received 30 Sept. 2015; accepted 20 Nov. 2015. Date of publication 4 Dec. 2015; date of current version 29 Aug. 2016.

For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below. Digital Object Identifier no. 10.1109/TMC.2015.2505729

signal strengths from the access points don't change rapidly and that the variance in the change is small and can be captured as a small additive noise. In the training phase, an engineer uses a smartphone to record RF fingerprints within a region of interest. Thereafter, a fingerprint database is built up at the server, in which each fingerprint is associated with the corresponding reference point. In the operating phase, a user submits a location query with her current fingerprint, then the server responds by matching the query fingerprint with candidate reference points in the database.

However, wide adoption of RF fingerprint-based approach is challenged by the time- and manpower-consuming site survey process as the engineer needs to sample a large number of reference points when the designer expects a significant changes of fingerprints. For example, a region of $100\text{ m} \times 100\text{ m}$ covers 10^4 reference points with grid size $1\text{ m} \times 1\text{ m}$. If a reference point takes about 10 seconds (including moving to this point and measuring a stable fingerprint), then the site survey process takes about 27.8 hours. Setting the grid size to a finer granularity will exacerbate this problem. In the end of 2014, Google Indoor Map 6.0 provides indoor localization and navigation only at some of the largest retailers, airports and transit stations in the US and Japan [13], while its expansion is constrained by limited amount of fingerprint data of building interiors.

Recently, there are many works aiming to relieve the site survey burden, which we broadly classify into correlation-aware approach, crowdsourcing-based approach, and sparsity-aware approach. The correlation-aware approach leverages the fact that the fingerprints at nearby reference points are spatially correlated. For example, [17] utilized the kriging interpolation method while [18] adopted the kernel functions (continuous and differentiable discriminant functions) and proposed an approach based on discriminant minimization search. Such schemes try to use linear or non-linear functions to model the correlations, which are essentially model-based approaches and face similar limiting factors as discussed before.

Crowdsourcing-based approach removes the offline site survey and instead incorporates users' online cooperation. For example, LiFS [19] leverages user reports to establish a one-to-one mapping between RF fingerprints and the digital floor plan map. However, the digital map is not always available. To overcome this Jigsaw [20] exploited image processing techniques to reconstruct floor plans based on the reported pictures and Zee [21] leveraged the inertial sensors (e.g., accelerometer, compass, gyroscope) to track users as they traverse an indoor environment while simultaneously performing Wi-Fi scans. Since the crowdsourcing-based approach is fundamentally the application of machine learning techniques to large-scale data sets and will be effective only when users' reports are sufficient to cover the whole space of human activity, it requires a large number of users to cooperate with the server. Furthermore, it imposes high energy consumption on smartphones.

Recent methods for efficient RF fingerprinting [22], [23], [24], [25], [26], [27] assume that the matrix of RSS values across channels/access points and locations is low rank and nuclear norm minimization [28] under *random or deterministic non-adaptive element-wise sampling* constraints is used

for data completion. In [29] the authors assume that the signals are fast varying, model the signal variation as a first order dynamical system, and give approaches for dynamic matrix completion (MC). Below we summarize the main contributions of our work and contrast them with the approaches taken so far.

1.1 Summary of Contributions

In this paper our main technical contributions are as follows.

- We model the RF fingerprint data as a low rank *tensor* using the notion of tensor rank proposed in [1], [30]. This is in contrast to existing works [22], [23], [24], [25], [26], [27] that assume a low rank matrix model. Furthermore, our algebraic framework is different from the traditional *multilinear algebraic framework* for tensor decompositions [31] that has been considered so far in the literature for problems of completing multidimensional arrays [32], [33], [34] with different notions for tensor rank.
- We propose an *adaptive* sampling strategy that leads to a dramatic improvement in localization accuracy for the same sample complexity. Our approach adaptively samples a small subset of all reference points based on which, the fingerprint data is then reconstructed using tensor completion (TC). In this context, a major difference from existing low-rank tensor completion [30], [31], [34], [35], [36] lies in that the sampling strategy in our paper is a vector-wise sampling, while existing low-rank tensor completion deals with entry-wise sampling. This is because when measurements is done at a location, there is data for all of the access points.
- We also derive theoretical performance bounds and show that it is possible to recover data under weaker conditions than required for non-adaptive random sampling considered in [30]. This is further evidenced by our numerical simulation on (i) a software model using concepts of ray-tracing and supervised learning with real data [37], [38], which accounts for shadowing of walls, wave guiding effects in corridors due to multiple reflections, diffractions around vertical wedges, and (ii) real data. We observe orders of magnitude improvements in the completion performance, see Figs. 6 and 7, over existing methods.
- Applying the completed data for localization also results in orders of magnitude performance improvement in localization accuracy (see Figs. 8, 9 and 10) over other competing methods. Therefore our approach of exploiting adaptivity is efficient in collecting measurements while maximizing the localization accuracy.

The remainder of the paper is organized as follows. In Section 2, we present the system model and the problem statement. Section 3 describes a random sampling approach as a baseline. Section 4 provides details of our approach while the performance guarantees are given in Section 5. Simulation results are presented in Section 6. Detailed proofs are given in the Appendix, which can be found on the Computer Society Digital Library at <http://doi.ieeecomputersociety.org/10.1109/TKDE.2015.2505729>, and concluding remarks are made in Section 7.

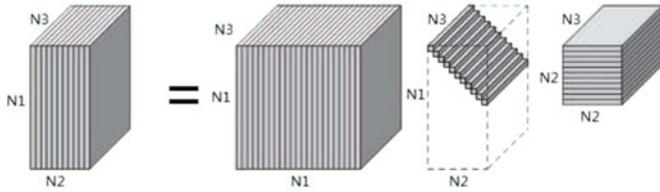


Fig. 1. The t-SVD of an $N_1 \times N_2 \times N_3$ tensor.

2 SYSTEM MODEL AND PROBLEM STATEMENT

We begin by first outlining the notations, the algebraic model and preliminary results for third-order tensors taken from [1], [30], [35].

A third-order tensor is represented by calligraphic letters, denoted as $\mathcal{T} \in \mathbb{R}^{N_1 \times N_2 \times N_3}$, and its (i, j, k) th entry is $\mathcal{T}(i, j, k)$. A *tube* (or fiber) of a tensor is a 1-D section defined by fixing all indices but one, thus a tube is a vector. In this paper, we use tube $\mathcal{T}(i, j, :)$ to denote a fingerprint at reference point (i, j) . Similarly, a *slice* of a tensor is a 2-D section defined by fixing all but two indices. *frontal*, *lateral*, *horizontal slices* are denoted as $\mathcal{T}(:, :, k)$, $\mathcal{T}(:, j, :)$, $\mathcal{T}(i, :, :)$, respectively. We use lateral slice $\mathcal{T}(:, j, :)$ to denote an $N_1 \times 1 \times N_3$ fingerprint matrix for the j th column of the grid map.

$\check{\mathcal{T}}$ is a tensor obtained by taking the Fourier transform along the third mode of \mathcal{T} , i.e., $\check{\mathcal{T}}(i, j, :) = \text{fft}(\mathcal{T}(i, j, :))$. In MATLAB notation, $\check{\mathcal{T}} = \text{fft}(\mathcal{T}, [], 3)$, and one can also compute \mathcal{T} from $\check{\mathcal{T}}$ via $\mathcal{T} = \text{ifft}(\check{\mathcal{T}}, [], 3)$.

The transpose of tensor \mathcal{T} is the $N_2 \times N_1 \times N_3$ tensor \mathcal{T}^\top obtained by transposing each of the frontal slices and then reversing the order of transposed frontal slices 2 through N_3 , i.e., for $k = 2, 3, \dots, N_3$, $\mathcal{T}^\top(:, :, k) = (\mathcal{T}(:, :, N_3 + 2 - k))^\top$ (the transpose of matrix $\mathcal{T}(:, :, N_3 + 2 - k)$). For two tubes $a, b \in \mathbb{R}^{1 \times 1 \times N_3}$, $a * b$ denotes the *circular convolution* between these two vectors.

The block diagonal matrix $\text{blkdiag}(\check{\mathcal{T}})$ is defined by placing the frontal slices in the diagonal, i.e., $\text{diag}(\check{\mathcal{T}}(:, :, 1), \check{\mathcal{T}}(:, :, 2), \dots, \check{\mathcal{T}}(:, :, N_3))$.

The algebraic development in [1] rests on defining a tensor-tensor product between two 3-D tensors, referred to as the t-product as defined below.

Definition 1 (t-product). The t-product $\mathcal{C} = \mathcal{A} * \mathcal{B}$ of $\mathcal{A} \in \mathbb{R}^{N_1 \times N_2 \times N_3}$ and $\mathcal{B} \in \mathbb{R}^{N_2 \times N_4 \times N_3}$ is a tensor of size $N_1 \times N_4 \times N_3$ whose (i, j) th tube $\mathcal{C}(i, j, :)$ is given by $\mathcal{C}(i, j, :) = \sum_{k=1}^{N_2} \mathcal{A}(i, k, :) * \mathcal{B}(k, j, :)$, for $i = 1, 2, \dots, N_1$ and $j = 1, 2, \dots, N_4$.

Owing to the relation between the circular convolution and Discrete Fourier Transform, we note the following remark that is used throughout the paper.

Remark 1. For $\mathcal{A} \in \mathbb{R}^{N_1 \times N_2 \times N_3}$ and $\mathcal{B} \in \mathbb{R}^{N_2 \times N_4 \times N_3}$, we have $\mathcal{A} * \mathcal{B} = \mathcal{C} \iff \text{blkdiag}(\check{\mathcal{A}}) \text{blkdiag}(\check{\mathcal{B}}) = \text{blkdiag}(\check{\mathcal{C}})$.

A third-order tensor of size $N_1 \times N_2 \times N_3$ can be viewed as an $N_1 \times N_2$ matrix of tubes that are in the third-dimension. So the t-product of two tensors can be regarded as multiplication of two matrices, except that the multiplication of two numbers is replaced by the circular convolution of two tubes. Further, this allows one to treat 3-D tensors as

linear operators over 2-D matrices as analyzed in [1]. Using this perspective one can define a SVD type decomposition, referred to as the tensor-SVD or t-SVD [1]. To define the t-SVD we introduce a few definitions.

Definition 2 (Identity tensor). The identity tensor $\mathcal{I} \in \mathbb{R}^{N_1 \times N_1 \times N_3}$ is a tensor whose first frontal slice $\mathcal{I}(:, :, 1)$ is the $N_1 \times N_1$ identity matrix and all other frontal slices are zero.

Definition 3 (Orthogonal tensor). A tensor $\mathcal{Q} \in \mathbb{R}^{N_1 \times N_1 \times N_3}$ is orthogonal if it satisfies $\mathcal{Q}^\top * \mathcal{Q} = \mathcal{Q} * \mathcal{Q}^\top = \mathcal{I}$.

Definition 4. The inverse of a tensor $\mathcal{U} \in \mathbb{R}^{N_1 \times N_1 \times N_3}$ is written as $\mathcal{U}^{-1} \in \mathbb{R}^{N_1 \times N_1 \times N_3}$ and satisfies $\mathcal{U}^{-1} * \mathcal{U} = \mathcal{U} * \mathcal{U}^{-1} = \mathcal{I}$.

Definition 5 (f-diagonal tensor). A tensor is called f-diagonal if each frontal slice of the tensor is a diagonal matrix, i.e., $\mathcal{T}(i, j, k) = 0$ for $i \neq j, \forall k$.

Using this definition one can obtain the t-SVD defined in the following result from [1]. Please see Fig. 1 for a graphical representation.

Theorem 1 (t-SVD). A tensor $\mathcal{T} \in \mathbb{R}^{N_1 \times N_2 \times N_3}$, can be decomposed as $\mathcal{T} = \mathcal{U} * \mathcal{\Theta} * \mathcal{V}^\top$, where \mathcal{U} and \mathcal{V} are orthogonal tensors of sizes $N_1 \times N_1 \times N_3$ and $N_2 \times N_2 \times N_3$ respectively, i.e., $\mathcal{U}^\top * \mathcal{U} = \mathcal{I}$ and $\mathcal{V}^\top * \mathcal{V} = \mathcal{I}$ and $\mathcal{\Theta}$ is a rectangular f-diagonal tensor of size $N_1 \times N_2 \times N_3$.

Definition 6 (Tensor tubal-rank). The tensor tubal-rank of a third-order tensor is the number of non-zero fibers of $\mathcal{\Theta}$ in the t-SVD.

In this framework, the principle of dimensionality reduction follows from the following result from [1].¹

Lemma 1 (Best rank- r approximation). Let the t-SVD of $\mathcal{T} \in \mathbb{R}^{N_1 \times N_2 \times N_3}$ be given by $\mathcal{T} = \mathcal{U} * \mathcal{\Theta} * \mathcal{V}^\top$ and for $r \leq \min(N_1, N_2)$ define $\mathcal{T}_r = \sum_{i=1}^r \mathcal{U}(:, i, :) * \mathcal{\Theta}(i, i, :) * \mathcal{V}^\top(:, i, :)$, then $\mathcal{T}_r = \arg \min_{\check{\mathcal{T}} \in \mathbb{T}} \|\mathcal{T} - \check{\mathcal{T}}\|_F$, where $\mathbb{T} = \{\mathcal{S} = \mathcal{X} * \mathcal{Y} \mid \mathcal{X} \in \mathbb{R}^{N_1 \times r \times N_3}, \mathcal{Y} \in \mathbb{R}^{r \times N_2 \times N_3}\}$.

We now define the notion of tensor column space. Under t-SVD in Definition 1, a tensor column subspace of \mathcal{T} is the space spanned by the lateral slices of \mathcal{U} under the t-product, i.e., the set generated by t-linear combinations like so, $\text{t-span}(\mathcal{U}) = \{\mathbf{X} = \sum_{j=1}^r \mathcal{U}(:, j, :) * \mathbf{c}_j \in \mathbb{R}^{N_1 \times 1 \times N_3}, \mathbf{c}_j \in \mathbb{R}^{N_3}\}$, where r denotes the tensor tubal-rank. We are now ready to present our main system model and assumptions.

2.1 System Model

Suppose that the region of interest is a rectangle $\mathcal{R} \in \mathbb{R}^2$. Dividing \mathcal{R} into an $N_1 \times N_2$ grid map, with each grid of the same size. The grid points are called *reference points*. Let \mathcal{G} denote the grid map and \mathcal{G} has $N_1 N_2$ reference points in total. Within \mathcal{R} , there are N_3 randomly deployed access points. We neither have access to or control over those access points, nor know their exact locations. The engineer uses a smartphone to measure the RSS values from these N_3 access points. We use a third-order tensor $\mathcal{T} \in \mathbb{R}^{N_1 \times N_2 \times N_3}$ to

1. Note that $\mathcal{\Theta}$ in t-SVD is organized in a decreasing order, i.e., $\|\mathcal{\Theta}(1, 1, :)\|_F \geq \|\mathcal{\Theta}(2, 2, :)\|_F \geq \dots$, which is implicitly defined in [30] as the algorithm for computing t-SVD is based on matrix SVD. Therefore, the best rank- r approximation of tensors is similar to PCA (principal component analysis).

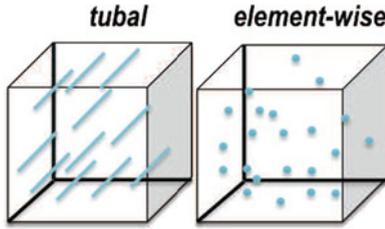


Fig. 2. Figure contrasting tubal sampling with element-wise sampling for 3D tensors.

represent the RSS map of \mathcal{G} . Each reference point $(i, j) \in \mathcal{G}$ is associated with a received signal strength vector $\mathcal{T}(i, j, \cdot)$, called a *fingerprint*, where $\mathcal{T}(i, j, k)$ is the RSS value of the k th access point. Note that the noise level is assumed to be equal to -110 dBm, $\mathcal{T}(i, j, k) = -110$ dBm [18], [19] if the signal of the k th access point cannot be detected. The fingerprint database stores the coordinates of all reference points and their corresponding RF fingerprints.

We use a third-order tensor $\mathcal{T} \in \mathbb{R}^{N_1 \times N_2 \times N_3}$ to represent the RSS map of \mathcal{G} , and the t-SVD is $\mathcal{T} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T$. As mentioned in the Introduction, the RSS value is highly correlated across space for each access point and also across these N_3 access points. We model this correlation by assuming that \mathcal{T} has low tensor tubal-rank r , in the sense defined before using the t-SVD.

We now outline the problem statement in detail. Note that as pointed out before, our sampling strategy will be to adaptively sample RF fingerprints, which correspond to the tensor fibers along the third dimension. See Fig. 2.

2.2 Problem Statement

Let $M < N_1 N_2$ denote the sampling budget, i.e., we are allowed to sample M reference points. We model the site survey process as the following *partial observation model*:

$$\mathcal{Y} = \mathcal{P}_\Omega(\mathcal{T}) + \mathcal{P}_\Omega(\mathcal{N}), \quad \Omega \subset \mathcal{G}, \quad (1)$$

where the (i, j, k) th entry of $\mathcal{P}_\Omega(\mathcal{X})$ is equal to $\mathcal{X}(i, j, k)$ if $(i, j) \in \Omega$ and zero otherwise, Ω being a subset of the grid map \mathcal{G} and of size M , and \mathcal{N} is an $N_1 \times N_2 \times N_3$ tensor with i.i.d. $N(0, \sigma^2)$ elements, representing the additive Gaussian noise. Since the engineer usually averages the recorded RSS values to get a stable fingerprint in the site survey process, while users want to get quick response from the server, the noise in the query process is much higher than that in the site survey samples \mathcal{Y} .

To cut down the grid map survey burden, we measure the RSS values of a small subset of reference points and then estimate \mathcal{T} from the samples \mathcal{Y} . There are two facts that can be exploited: the prior information that tensor \mathcal{T} is low-tubal-rank, and the estimated $\hat{\mathcal{T}}$ should equal to \mathcal{Y} on the set Ω . Therefore, we estimate \mathcal{T} by solving the following optimization problem:

$$\begin{aligned} \langle \hat{\mathcal{T}}, \hat{\Omega} \rangle = \arg \min_{\mathcal{X}, \Omega} \quad & \|\mathcal{P}_\Omega(\mathcal{Y} - \mathcal{X})\|_F^2 + \lambda \cdot \text{rank}(\mathcal{X}), \\ \text{s.t.} \quad & |\Omega| \leq M, \end{aligned} \quad (2)$$

where \mathcal{X} is the decision variable, $\text{rank}(\cdot)$ refers to the tensor tubal-rank, M is the sampling budget, and λ is a

regularization parameter. This approach aims to seek the simplest explanation fitting the samples.

Clearly the performance depends on the type of sampling strategy used and the algorithm for solving the optimization problem. In this paper we will consider two kinds of sampling, namely, uniform random tubal-sampling and adaptive tubal-sampling.

3 RANDOM TUBAL-SAMPLING APPROACH

Before introducing the adaptive sampling approach, we first present a non-adaptive sampling approach, i.e., tensor completion via uniform tubal-sampling. We would like to point out again that tensor completion via uniform entry-wise sampling is well-studied [30], [31], [34], however, *tubal-sampling* is required in RF fingerprint-based indoor localization. By uniform tubal-sampling, we mean that in (2) the subset Ω is simply chosen uniformly randomly (with replacement) from the grid map \mathcal{G} with $|\Omega| \leq M$. Similar to the matrix case, given a fixed Ω solving the optimization problem (2) is NP-hard. In this case one can relax the tubal-rank measure to a tensor nuclear norm (TNN) measure as proposed in [30] and solve for the resulting convex optimization problem. The tensor nuclear norm as derived from the t-SVD is defined as follows.

Definition 7. [30], [31] *The tensor-nuclear-norm (TNN) denoted by $\|\mathcal{T}\|_{\text{TNN}}$ and defined as the sum of the singular values of all the frontal slices of $\check{\mathcal{T}}$, i.e., $\|\mathcal{T}\|_{\text{TNN}} = \sum_{k=1}^{N_3} \|\check{\mathcal{T}}(:, :, k)\|_*$, where $\|\cdot\|_*$ denotes the nuclear norm.*

Extending the noiseless case considered in [30] using the tensor-nuclear-norm for tensor completion, in this paper we solve for the noisy case (2) via,

$$\hat{\mathcal{T}} = \arg \min_{\mathcal{X}} \quad \|\mathcal{P}_\Omega(\mathcal{Y} - \mathcal{X})\|_F^2 + \lambda \|\mathcal{X}\|_{\text{TNN}}. \quad (3)$$

This optimization problem can be solved using ADMM, with modifications to the algorithm in [30]. Note that in [30] the authors consider a random element-wise sampling whereas here we are considering a random tubal-sampling. It turns out that under tubal sampling the tensor completion by solving the convex optimization of Equation (3) splits into N_3 matrix completion problems in the Fourier domain. This observation leads us to derive the performance bounds as an extension of the results in matrix completion using the following notion of incoherency conditions.

Definition 8 (Tensor Incoherency Condition). *Given the t-SVD of a tensor $\mathcal{T} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T$ with tubal-rank r , \mathcal{T} is said to satisfy the tensor-incoherency conditions, if there exists $\mu_0 > 0$ such that for $k = 1, 2, \dots, N_3$.*

(Tensor-column incoherence :)

$$\mu(\mathcal{U}) \triangleq \frac{N_1}{r} \max_{i=1, \dots, N_1} \|\check{\mathcal{U}}^T(:, :, k) \mathbf{e}_i\|_F^2 \leq \mu_0, \quad (4)$$

(Tensor-row incoherence :)

$$\mu(\mathcal{V}) \triangleq \frac{N_2}{r} \max_{j=1, \dots, N_2} \|\check{\mathcal{V}}^T(:, :, k) \mathbf{e}_j\|_F^2 \leq \mu_0,$$

where $\|\cdot\|_F$ denotes the usual Frobenius norm and $\mathbf{e}_i, \mathbf{e}_j$ are standard co-ordinate basis of respective lengths.

We have the following result stated without proof.

Theorem 2. Under random tubal sampling, if $\mu(\mathcal{U}), \mu(\mathcal{V}) \leq \mu_0$ then for $M \geq CN\mu_0 r \log^2 N$ for some constant $C > 0$ and $N = \max\{N_1, N_2\}$, solving the convex optimization problem of Equation (3) recovers the tensor with high probability (for sufficiently large N).

In contrast to this result we will see that adaptive sampling as we propose below requires almost the same sampling budget but *only requires tensor column incoherence* $\mu(\mathcal{U})$ to be small (see Theorem 3), which is one of the major gains of adaptive tubal-sampling over random tubal-sampling as well as over random element-wise sampling considered in [35]. This gain is similar to the matrix case (under element-wise adaptive sampling) as shown in [36]. Adaptive sampling allows one to obtain more gains when the energy seems to be concentrated on few locations (for example, see Fig. 4), which the random sampling can miss! Further, adaptive sampling reduces the number of measurements by a constant factor for the same accuracy. This is borne out by our experimental results. Although our approach is motivated by adaptive matrix sampling strategy in [36], we would like to point out that the performance bounds for the proposed adaptive strategy do not directly follow from the results in [36] and requires a careful treatment.

4 PROPOSED ADAPTIVE SAMPLING APPROACH

We begin by revisiting the problem and providing insights into the development of the proposed adaptive strategy.

The problem (2) contains two goals: (1) For a given low-tubal-rank tensor \mathcal{X} , to select a set Ω with the smallest cardinality and the corresponding samples \mathcal{Y} , preserving most information of tensor \mathcal{X} , i.e., one can recover \mathcal{X} from Ω and \mathcal{Y} . (2) For a given set Ω and samples \mathcal{Y} , to estimate a tensor \mathcal{X} that has the least tubal-rank. However, these two goals are intertwined together and one cannot expect a computationally feasible algorithm to get the optimal solution. Therefore, we set $|\Omega| = M$ and seek to select a set Ω and the corresponding samples \mathcal{Y} that span the low-dimensional tensor-column subspace of \mathcal{T} . The focus of this section is to design an efficient sampling scheme and to provide a bound on the sampling budget M .

To achieve this, we design a two-pass sampling scheme inspired by [39]. The proposed approach exploits adaptivity to identify entries that are highly informative for learning the low-dimensional tensor-subspace of the fingerprint data. The 1st-pass sampling gathers general information about the region of interest, then the second-pass sampling concentrates on those more informative reference points. In particular the total sampling budget $M (< N_1 N_2)$ is divided into δM and $(1 - \delta)M$ for these two sampling passes and δ is called the *allocation ratio*. In the 1st-pass sampling, we randomly sample $\delta M / N_2$ out of N_1 reference points in each column of \mathcal{G} . In the second-pass sampling, the remaining $(1 - \delta)M$ samples are allocated to those highly informative columns identified by the first-pass sampling. Finally, tensor completion on those M RF fingerprints is performed to rebuild a fingerprint database.

The provable optimality of this scheme rests on these three observations.

- \mathcal{T} is embedded in an r -dimensional tensor-column subspace \mathcal{U} , $r \ll \min\{N_1, N_2\}$;

- Learning \mathcal{U} requires to know only r linearly independent lateral slices;²
- Knowing \mathcal{U} , randomly sampling a few tubes of the j th column is enough to reliably estimate the lateral slice $\mathcal{T}(:, j, :)$;

However, we do not know the value of r a priori nor the independence between any two lateral slices. Given a current estimate \mathcal{U} of the tensor-column subspace, following [39] one can *adaptively* sample each column according to the probability distribution where the probability p_j of sampling the j th lateral slice is proportional to $\|\mathcal{P}_{\mathcal{U}^\perp} \mathcal{T}(:, j, :)\|_F^2$, i.e., $p_j = \frac{\|\mathcal{P}_{\mathcal{U}^\perp} \mathcal{T}(:, j, :)\|_F^2}{\|\mathcal{P}_{\mathcal{U}^\perp} \mathcal{T}\|_F^2}$. Updating the estimate of \mathcal{U} iteratively, when cr ($c > 1$ is a small constant) columns are sampled, we can expect that with high probability, $\|\mathcal{P}_{\mathcal{U}^\perp} \mathcal{T}(:, j, :)\|_F^2 = 0, \forall j$. Note that $\mathcal{P}_{\mathcal{U}^\perp}(\cdot)$ denotes projection onto the orthogonal space of \mathcal{U} ; in t-product form, $\mathcal{P}_{\mathcal{U}} = \mathcal{U} * (\mathcal{U}^T * \mathcal{U})^{-1} * \mathcal{U}^T$, $\mathcal{P}_{\mathcal{U}^\perp} = \mathcal{I} - \mathcal{P}_{\mathcal{U}}$, and $\mathcal{U}^T * \mathcal{U}$ is invertible and can be computed according to Definition 4 and Remark 1.

The challenge is that we cannot have the exact sampling probability p_j without sampling all reference points of the grid map \mathcal{G} . Nevertheless, exploiting the spatial correlation, one can estimate the sampling probability from missing data

(sub-sampled data), as $\hat{p}_j = \frac{\|\mathcal{P}_{\mathcal{U}^\perp}(\mathcal{T}(:, \Omega_j^1, :))\|_F^2}{\|\mathcal{P}_{\mathcal{U}^\perp}(\mathcal{T}(:, \Omega^1, :))\|_F^2}$ in Algorithm 1.

Essentially Under the incoherency conditions of Equation (4) and for sufficiently large M we show in Lemma 4 that \hat{p}_j is a good estimation of p_j . Then using these estimates one can show that adaptive sampling scheme of the second pass of Algorithm 1 succeeds with high probability in estimating the correct tensor column-subspace. This is the content of Lemma 2 which analyzes a single step of second pass sampling using which the main Theorem 3 follows.

Algorithm 1. Tensor completion based on adaptive sampling

Input: parameters $\mathcal{G}, M, N_1, N_2, N_3, \delta, L$

1st-pass sampling:

Uniformly sample $\delta M / N_2$ reference points from each column of \mathcal{G} , denoted as Ω_j^1 . $\Omega^1 = \bigcup_{j=1}^{N_2} \Omega_j^1$.

2nd-pass sampling:

$\hat{\mathcal{U}} \leftarrow \emptyset, \Pi \leftarrow \{1, 2, \dots, N_2\}$.

for $l = 1 : L$ **do**

Estimate $\hat{p}_j = \frac{\|\mathcal{P}_{\hat{\mathcal{U}}^\perp}(\mathcal{T}(:, \Omega_j^1, :))\|_F^2}{\|\mathcal{P}_{\hat{\mathcal{U}}^\perp}(\mathcal{T}(:, \Omega^1, :))\|_F^2}, \forall j \in \Pi$.

Sample $s = \frac{(1-\delta)M}{N_2 L}$ columns of \mathcal{G} according to \hat{p}_j for $\forall j \in \Pi$, denoted as Π_s^l .

Calculate $\mathbb{U} \leftarrow \frac{\mathcal{P}_{\hat{\mathcal{U}}^\perp}(\mathcal{T}(:, \Pi_s^l, :))}{\|\mathcal{P}_{\hat{\mathcal{U}}^\perp}(\mathcal{T}(:, \Pi_s^l, :))\|_F}$, and update $\hat{\mathcal{U}} \leftarrow \hat{\mathcal{U}} \cup \mathbb{U}$

(concatenate \mathbb{U} to $\hat{\mathcal{U}}$).

$\Pi \leftarrow \Pi / \Pi_s^l$ (set subtraction).

end for

Approximate each lateral slice $\mathcal{T}(:, j, :)$ with $\hat{\mathcal{T}}(:, j, :) = \hat{\mathcal{U}} * (\hat{\mathcal{U}}_{\Omega_j^1}^T * \hat{\mathcal{U}}_{\Omega_j^1})^{-1} * \hat{\mathcal{U}}_{\Omega_j^1}^T * \mathcal{T}(\Omega_j^1, j, :)$.

2. Note: A collection of r lateral slices $\mathcal{U}(:, j, :), j = 1, \dots, r$ are said to be linearly independent (in the proposed setting) if $\sum_{j=1}^r \mathcal{U}(:, j, :)*c_j = 0 \Rightarrow c_j = 0, \forall j$.

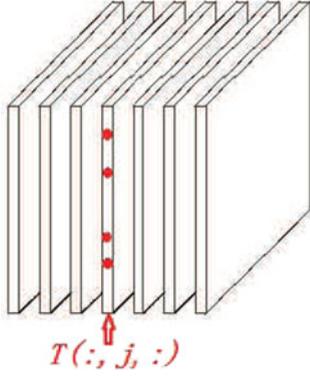


Fig. 3. After adaptively building an estimate of \mathcal{U} using two pass sampling approach of Algorithm 1, in the last step of the algorithm, the engineer performs site survey at a few random selected reference points (with red color) of the j th column of \mathcal{G} . Knowing the r -dimensional tensor-subspace \mathcal{U} , one can recover the j th lateral slice $\mathcal{T}(:, j, :)$.

The pseudo-code of our adaptive tubal-sampling approach is shown in Algorithm 1. The inputs include the grid map \mathcal{G} , the sampling budget M , the size of the tensor, N_1, N_2, N_3 , the allocation ratio δ , and the number of iterations L . The algorithm consists of three steps. The first-pass sampling is a uniform tubal-sampling, while the second-pass sampling outputs an estimate $\hat{\mathcal{U}}$ of the tensor-column subspace \mathcal{U} in L rounds, as explained below. The approach is also summarized in Fig. 3.

4.1 First-Pass Sampling

First, we gather general information of the whole region of interest, applying a uniform random sampling to avoid spatial bias. Denote these sampled δM reference points as Ω^1 , the sampled reference points in the j th column as Ω_j^1 with $m = |\Omega_j^1| = \delta M / N_2$, and the corresponding fingerprints as $\mathcal{T}(\Omega_j^1, j, :)$.

4.2 Second-Pass Sampling

Initialize with $\hat{\mathcal{U}} \leftarrow \emptyset, \Pi \leftarrow \emptyset$. In each round, we estimate the sampling probability $\hat{p}_j = \frac{\|\mathcal{P}_{\hat{\mathcal{U}}^\perp}(\mathcal{T}(:, \Omega_j^1, :))\|_F^2}{\|\mathcal{P}_{\hat{\mathcal{U}}^\perp}(\mathcal{T}(:, \Omega^1, :))\|_F^2}$ and choose s columns of \mathcal{G} according to the probability \hat{p}_j . Then, we calculate an intermediate subspace \mathbb{U} that is the space spanned by these s columns and lies outside of $\hat{\mathcal{U}}$, and then update the subspace $\hat{\mathcal{U}}$. Π_s^l denotes these s lateral slices, $\mathcal{T}(:, \Pi_s^l, :)$ denotes these RF fingerprints in the l th round, corresponding to these s columns of \mathcal{G} .

4.3 Estimation

Let Ω denote all sampled reference points (including Ω^1), Ω_j denotes the sampled reference points in the j th column of \mathcal{G} , and \mathcal{U}_{Ω_j} denotes the tensor organized by the horizontal slices of \mathcal{U} indicated by Ω_j . Define the projection operator $\mathcal{P}_{\mathcal{U}_{\Omega_j}} = \mathcal{U}_{\Omega_j} * (\mathcal{U}_{\Omega_j}^T * \mathcal{U}_{\Omega_j})^{-1} * \mathcal{U}_{\Omega_j}^T$. After L rounds, we can obtain a fairly accurate estimation of \mathcal{U} . To understand the estimator, considering the noiseless case. Since \mathcal{T} lies in \mathcal{U} , we have $\|\mathcal{T}(\Omega_j, j, :) - \mathcal{P}_{\mathcal{U}_{\Omega_j}} * \mathcal{T}(\Omega_j, j, :)\| = 0$, i.e., $\mathcal{T}(\Omega_j, j, :) = \mathcal{U}_{\Omega_j} * (\mathcal{U}_{\Omega_j}^T * \mathcal{U}_{\Omega_j})^{-1} * \mathcal{U}_{\Omega_j}^T * \mathcal{T}(\Omega_j, j, :)$. And according to Definition 1, we know that $\mathcal{T}(\Omega_j, j, :) = \mathcal{U}_{\Omega_j} * \Theta * \mathcal{V}^T(:, j, :)$.

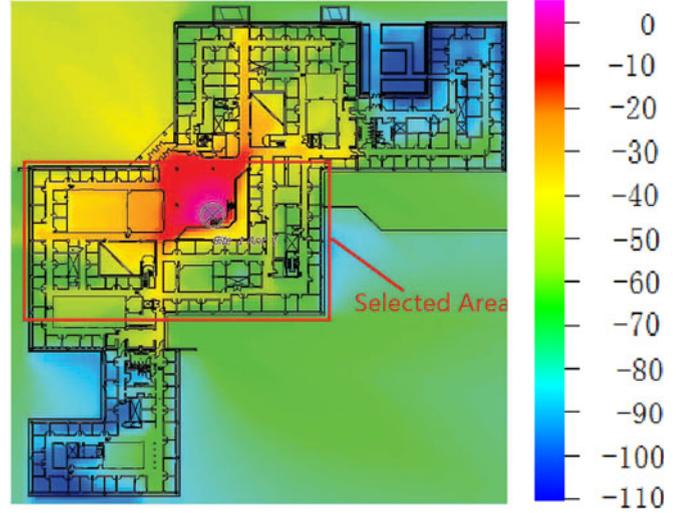


Fig. 4. We select a rectangle area within an office building for simulations. RSS is measured in dBm.

Therefore, using the estimate $\hat{\mathcal{U}}$, we approximate each lateral slice $\mathcal{T}(:, j, :)$ with $\hat{\mathcal{T}}(:, j, :) = \hat{\mathcal{U}} * (\hat{\mathcal{U}}_{\Omega_j}^T * \hat{\mathcal{U}}_{\Omega_j})^{-1} * \hat{\mathcal{U}}_{\Omega_j}^T * \mathcal{T}(\Omega_j, j, :)$ and concatenate these estimates to form $\hat{\mathcal{T}}$.

5 PERFORMANCE BOUNDS

For performance guarantee, we are interested in the recovery error and required sampling budget. We prove that Algorithm 1 has bounded recovery error and achieves near-optimal sampling budget. Since we use the estimated sampling probability in the second-pass sampling, we also prove that our estimates \hat{p}_j are relatively close to p_j .

STEP 1 - First, we analyze a single round of the second-pass sampling. Lemma 2 states that if the probability estimates are to within a constant tolerance of the true estimates (see Equation (5)), then sampling s columns of \mathcal{G} according to \hat{p}_j (estimated based on samples obtained in the first-pass sampling) will minimize the residual error within $\hat{\mathcal{U}}$ at rate $\frac{1}{s}$ ($s \geq 1$). The second term in the right-hand side of (6) denotes the residual error outside of $\hat{\mathcal{U}}$, which remains unreduced. Note that without any *prior* information (i.e., $\hat{\mathcal{U}}$ and \hat{p}_j), sampling additional s columns of \mathcal{G} will reduce the residual error at rate $\frac{s}{N_2}$ which is $\ll \frac{1}{s}$ as $s \ll \sqrt{N_2}$ in Algorithm 1. Therefore, Lemma 2 is the key to efficiently reduce recovery error. Note that $\mathcal{E} = \mathcal{P}_{\hat{\mathcal{U}}^\perp}(\mathcal{T})$ denotes the projection onto the orthogonal complement of $\hat{\mathcal{U}}$. See Appendix A, available online, for its proof.

Lemma 2. Let $\mathcal{T} = \mathcal{U} * \Theta * \mathcal{V}^T \in \mathbb{R}^{N_1 \times N_2 \times N_3}$ with tensor tubal-rank r , and $\hat{\mathcal{U}} \in \mathbb{R}^{N_1 \times r \times N_3}$ represent the estimated tensor-column subspace in a round. Let $\mathcal{E} = \mathcal{P}_{\hat{\mathcal{U}}^\perp}(\mathcal{T}) = \mathcal{T} - \mathcal{P}_{\hat{\mathcal{U}}}(\mathcal{T})$, and S be s randomly selected lateral slices of \mathcal{T} (as indicated by Π_s^l), sampled according to the distribution $\hat{p}_j = \frac{\|\mathcal{P}_{\hat{\mathcal{U}}^\perp}(\mathcal{T}(:, \Omega_j^1, :))\|_F^2}{\|\mathcal{P}_{\hat{\mathcal{U}}^\perp}(\mathcal{T}(:, \Omega^1, :))\|_F^2}$. If there exist constants $\alpha_1, \alpha_2 \in \mathbb{R}$, such that:

$$\frac{1 - \alpha_1}{1 + \alpha_2} \cdot \frac{\|\mathcal{E}(:, j, :)\|_F^2}{\|\mathcal{E}\|_F^2} \leq \hat{p}_j \leq \frac{1 + \alpha_2}{1 - \alpha_1} \cdot \frac{\|\mathcal{E}(:, j, :)\|_F^2}{\|\mathcal{E}\|_F^2}, \quad (5)$$

then, with probability $\geq 1 - \rho$ we have:

$$\begin{aligned} & \|\mathcal{T} - P_{\hat{\mathcal{U}}_{\text{t-span}(\mathcal{S}),r}}(\mathcal{T})\|_F^2 \\ & \leq \frac{r}{s\rho} \cdot \xi_0 \frac{1 + \alpha_2}{1 - \alpha_1} \cdot \|\mathcal{E}\|_F^2 + \|\mathcal{T} - \mathcal{T}_r\|_F^2, \end{aligned} \quad (6)$$

where ξ_0 is a constant such that

$$\sum_{j=1}^{N_2} |\check{\mathcal{V}}^T(i, j, k)|^2 \leq \xi_0, \quad (7)$$

$\text{t-span}(\mathcal{S})$ denotes the space spanned by the slices of \mathcal{S} , and $P_{\mathcal{H},r}(\cdot)$ denotes a projection on to the best r -dimensional subspace of \mathcal{H} .

STEP 2 - Now note that we perform the adaptive sampling scheme for L rounds. For this Lemma 3 below gives an induction argument to chain Lemma 2 across all rounds of the second-pass sampling.

Setting $\epsilon < 1$ and L sufficiently large, Lemma 3 states that our two-pass sampling scheme can approximate the low tubal-rank tensor \mathcal{T} with error comparable to that of the best rank- r approximation \mathcal{T}_r . This indicates that the second-pass sampling estimates \mathcal{U} with high accuracy.

Lemma 3. Suppose that (5) holds with $\frac{1+\alpha_2}{1-\alpha_1} \leq c$ for some constant c in each round. Let $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_L$ denote the sets of slices selected at each round and set $s = \frac{Lcr\xi_0}{\rho\epsilon}$. Then with probability $\geq 1 - \rho$ we have:

$$\begin{aligned} \|\mathcal{T} - \mathcal{P}_{\hat{\mathcal{U}}}(\mathcal{T})\|_F^2 &= \|\mathcal{T} - P_{\cup_{i=1}^L \mathcal{S}_i, r}(\mathcal{A})\|_F^2 \\ &\leq \frac{1}{1-\epsilon} \|\mathcal{T} - \mathcal{T}_r\|_F^2 + \epsilon^L \|\mathcal{T}\|_F^2. \end{aligned} \quad (8)$$

Proof. The proof follows along the same lines as the proof for the adaptive matrix completion case in [36], by applying the matrix result to N_3 frontal slices of the tensor in the Fourier domain $\check{\mathcal{T}}$. \square

Now note that Lemma 2 requires, that the estimates of \hat{p}_j are close to the actual values by a factor at most $c = \frac{1+\alpha_2}{1-\alpha_1}$. For this Lemma 4 states that this condition holds with high probability for $\alpha_1 = -1, \alpha_2 = 4$, and $c = \frac{5}{2}$.

Lemma 4. Let $\mathcal{E} = \mathcal{P}_{\hat{\mathcal{U}}^\perp}(\mathcal{T})$ and \mathcal{T} satisfies the tensor-column incoherence condition (4) with $\mu(\mathcal{U}) \leq \mu_0$, with probability $\geq 1 - 6\rho$ we have

$$\frac{2}{5} \cdot \frac{\|\mathcal{E}(:, j, :)\|_F^2}{\|\mathcal{E}\|_F^2} \leq \hat{p}_j \leq \frac{5}{2} \cdot \frac{\|\mathcal{E}(:, j, :)\|_F^2}{\|\mathcal{E}\|_F^2}. \quad (9)$$

as long as the expected sampling budget M satisfies:

$$\begin{aligned} M &\geq CrN_2(\mu_0 \log(N_2/\rho) \log(rN_2/\rho) \\ &\quad + (\lceil \log(N_1N_2N_3) \rceil)^2 / (\rho\epsilon)), \end{aligned} \quad (10)$$

for some $C > 0$, and thus is of order $O(Nr \log^2 N)$ with $N = \max(N_1, N_2)$.

Proof. This result relies on three conditions: the incoherence of each lateral slice $\mathcal{T}(:, j, :)$, the tensor-column

incoherence of \mathcal{U} - Equation (4), and Lemma 12; those three conditions' failure probabilities are less than ρ, ρ , and 4ρ , respectively; therefore, Lemma 4 holds with probability $\geq 1 - 6\rho$. See Appendix C, available online, for a complete proof. \square

Remark 2. For $\mathcal{T} \in \mathbb{R}^{N_1 \times N_2 \times N_3}$ with tensor tubal-rank r , its t-SVD in Definition 1 indicates that the degree of freedom (in terms of non-zero vectors) is $N_1r + N_2r + r$ which is of order $O(Nr)$ with $N = \max(N_1, N_2)$. Therefore, our sampling budget is near-optimal within a factor of $\log^2 N$.

STEP 3 - Finally, we analyze the estimation process (Last step of the Algorithm) in Lemma 5. It states that our estimator outputs each lateral slice with bounded error, which is comparable to the energy outside of $\hat{\mathcal{U}}$, i.e., $\|\mathcal{P}_{\hat{\mathcal{U}}^\perp} \mathcal{T}(:, j, :)\|_F$.

Lemma 5. For $\mathcal{T} \in \mathbb{R}^{N_1 \times N_2 \times N_3}$ with tensor tubal-rank r , let $\mathcal{T}(:, j, :)$ denote the j th lateral slice and $\hat{\mathcal{U}}$ denote the tensor-column subspace at the L th round of the 2nd-pass sampling. Algorithm 1 estimates $\mathcal{T}(:, j, :)$ as $\hat{\mathcal{T}}(:, j, :) = \hat{\mathcal{U}} * (\hat{\mathcal{U}}_{\Omega_j}^T * \hat{\mathcal{U}}_{\Omega_j}^{-1} * \hat{\mathcal{U}}_{\Omega_j}^T * \mathcal{T}(\Omega_j, j, :))$. Then with probability $\geq 1 - 2\rho$,

$$\begin{aligned} & \|\mathcal{T}(:, j, :) - \hat{\mathcal{T}}(:, j, :)\|_F^2 \\ & \leq \left(1 + \frac{r\mu(\hat{\mathcal{U}})\beta}{m(1-\gamma)^2}\right) \|\mathcal{P}_{\hat{\mathcal{U}}^\perp} \mathcal{T}(:, j, :)\|_F^2. \end{aligned} \quad (11)$$

where $\beta = (1 + 2\sqrt{\log(1/\rho)})^2$, $\gamma = \sqrt{\frac{8r\mu(\hat{\mathcal{U}})}{3m} \log(2r/\rho)}$, and $\mu(\hat{\mathcal{U}})$ is defined in (4).

With our choice of m (and correspondingly M related via $\delta M = mN_2$), $(1 + \frac{r\mu(\hat{\mathcal{U}})\beta}{m(1-\gamma)^2})$ is smaller than $5/4$ as shown in Appendix C, available online. Therefore, $\|\hat{\mathcal{T}} - \mathcal{T}'\|_F^2 = \sum_{j=1}^{N_2} \|\mathcal{T}(:, j, :) - \hat{\mathcal{T}}(:, j, :)\|_F^2 \leq \frac{5}{4} \|\mathcal{T} - \mathcal{P}_{\hat{\mathcal{U}}}(\mathcal{T})\|_F^2$, combining Lemma 3 and Lemma 4, (see also [36]), we have Lemma 6. Note that Lemma 3, Lemma 4, and Lemma 5 each has failure probability less than $\rho, 2\rho$ and 6ρ , therefore, Lemma 6 has success probability $\geq 1 - 9\rho$.

Lemma 6. Assume that $\mathcal{T}' = \mathcal{T} + \mathcal{N}$, \mathcal{T} has tensor tubal-rank r , tensor-column incoherence $\mu(\mathcal{U}) \leq \mu_0$, $\mathcal{N}(i, j, k) \sim N(0, \sigma^2)$, and for all i , $\sum_{j=1}^{N_2} |\check{\mathcal{V}}^T(i, j, k)|^2 \leq \xi_0$. Then for $\rho, \epsilon \in (0, 1)$, sample $s = \frac{5Lr\xi_0}{2\rho\epsilon}$ columns of \mathcal{G} each round, after L rounds, compute $\hat{\mathcal{T}}$ as described. Then with probability $\geq 1 - 9\rho$,

$$\|\hat{\mathcal{T}} - \mathcal{T}'\|_F^2 \leq 5/4 \left(\frac{1}{1-\epsilon} \|\mathcal{T}' - \mathcal{T}'_r\|_F^2 + \epsilon^L \|\mathcal{T}'\|_F^2 \right). \quad (12)$$

While Lemma 6 bounds the difference between $\hat{\mathcal{T}}$ and \mathcal{T}' , we then bound the difference between $\hat{\mathcal{T}}$ and \mathcal{T} , which measures the recovery error of Algorithm 1. This error consists of two parts: the first term measures the performance of our estimation process, and the second term is essentially $\|\mathcal{N}_\Omega\|_F^2$ which measures the effect of noise in the samples \mathcal{Y} . Combining these results we have the following main Theorem.

Theorem 3. Under the partial observation model $\mathcal{P}_\Omega(\mathcal{T} + \mathcal{N})$, where $\mathcal{T} \in \mathbb{R}^{N_1 \times N_2 \times N_3}$, $\mathcal{N}(i, j, k) \sim N(0, \sigma^2)$. Assume that \mathcal{T} has tubal-rank r , $\sum_{j=1}^{N_2} |\dot{\mathcal{Y}}^T(i, j, k)|^2 \leq \xi_0, \forall i$, and tensor-column incoherence $\mu(\mathcal{U}) \leq \mu_0$. Then for $\rho, \epsilon \in (0, 1)$, $L = \lceil \log_{1/\epsilon}(N_1 N_2 N_3) \rceil$ and

$$M = \max \left\{ \frac{5N_2 L^2 r \xi_0}{2(1-\delta)\rho\epsilon}, \right. \\ \left. CrN_2(\mu_0 \log(N_2/\rho) \log(rN_2/\rho) + L^2/(\rho\epsilon)) \right\}, \quad (13)$$

with probability $\geq 1 - 9\rho$, there exist two constants c_1, c_2 such that the estimation error of Algorithm 1 obeys,

$$\|\mathcal{T} - \hat{\mathcal{T}}\|_F^2 \leq \frac{c_1}{N_1 N_2 N_3} \|\mathcal{T}\|_F^2 + c_2(MN_3 + \sqrt{8MN_3})\sigma^2. \quad (14)$$

As a consequence of this theorem, for example, assuming that the ℓ_2 -norm of each fingerprint is approximately the same, say C , then our algorithm guarantees that the recovery error of each fingerprint in ℓ_2 -norm will be bounded by

$\sqrt{\frac{c_1 C^2}{N_1 N_2 N_3} + \frac{c_2(MN_3 + \sqrt{8MN_3})\sigma^2}{N_1 N_2}}$ and the relative error is bounded

by $\sqrt{\frac{c_1}{N_1 N_2 N_3} + \frac{c_2(MN_3 + \sqrt{8MN_3})\sigma^2}{N_1 N_2 C^2}}$. Since N_1, N_2 are relatively large, $N_3 \ll \min(N_1, N_2)$ and M is provided in Lemma 4 to be in order of $O(Nr \log^2 N)$ with $N = \max(N_1, N_2)$, therefore, the relative error is small.

6 PERFORMANCE EVALUATION

We are interested in two kinds of performance: recovery error and localization error. Varying the sampling rate as 10 ~ 90 percent, we quantify the recovery error in terms of *normalized square of error (NSE)* for entries that are not sampled, i.e., recovery error for set Ω^c . The NSE is defined as:

$$NSE = \frac{\sum_{(i,j) \in \Omega^c} \|\hat{\mathcal{T}}(i, j, :) - \mathcal{T}(i, j, :)\|_F^2}{\sum_{(i,j) \in \Omega^c} \|\mathcal{T}(i, j, :)\|_F^2}, \quad (15)$$

where $\hat{\mathcal{T}}$ is the estimated tensor, Ω^c is the complement of set Ω .

In the simulations, we uniformly select 500 testing points within the selected region and then using the classic localization schemes to perform localization estimation. We measure the localization error as the Euclidean distance between the estimated location and the actual location of the testing point, i.e., $d_e = \sqrt{(\hat{x}_1 - x_1)^2 + (\hat{x}_2 - x_2)^2}$.

For tensor recovery, we consider three algorithms, tensor completion under uniformly random tubal-sampling and using the algorithm proposed in [30], [35], using the *face-wise* matrix completion algorithm in [40], and tensor completion via matricization or flattening (MC-flat) [34] under *uniform element-wise sampling* of the 3D tensor, using the AltMin algorithm for matrix completion [41]. We subsequently use the completed RSS map for localization and compare the error in location estimates.

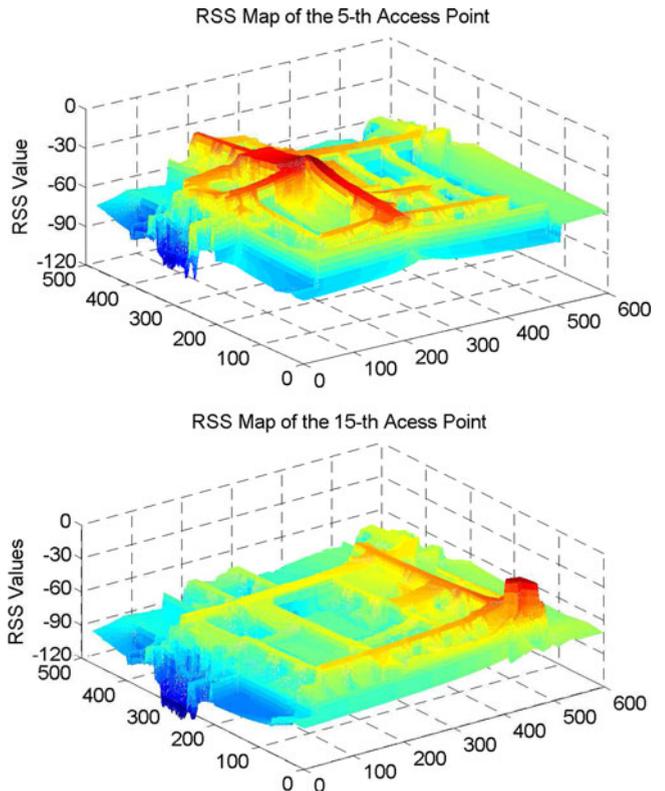


Fig. 5. The RSS radio map of the 5th and 15th access points.

6.1 Experiment Setup - Model-Based Data

We select a region of 47.5 m \times 59.7 m in a real office building, as shown in Fig. 4. It is divided into a 476 \times 598 grid map. There are 15 access points randomly deployed within this region. The indoor radio channel is characterized by multi-path propagation with dominant propagation phenomena: the shadowing of walls, wave guiding effects in corridors due to multiple reflections, and diffractions around vertical wedges. The ray-tracing based model [37], [38] is adopted, which considers all these effects leading to highly accurate prediction results. Further, the model parameters were found by supervised learning with the real collected measurements using a professional software [37]. We generated a 476 \times 598 \times 15 RSS tensor as the *ground truth* for our simulations. Note that the RSS values are measured in dBm. For example, the RSS radio map for the 5th and 15th access points are shown in Fig. 5.

Radio map recovery performance. Fig. 6 shows the RSS tensor recovery performance for varying sampling rate. Compared schemes are matrix completion and tensor completion via uniform sampling, and adaptive sampling with allocation ratio $\delta = 1/4$ and $\delta = 1/2$. We find that all tensor approaches are better than matrix completion, this is because tensor exploits the cross correlations among access points while matrix completion only takes advantage of correlation within each access point. Both AS schemes outperform tensor completion via uniform sampling since adaptivity can guide the sampling process to concentrate on more informative entries. Allocating equal sampling budget for the first-pass and the second-pass gives better performance than uneven allocation. This shows that the first-pass and the second-pass have equal importance. The

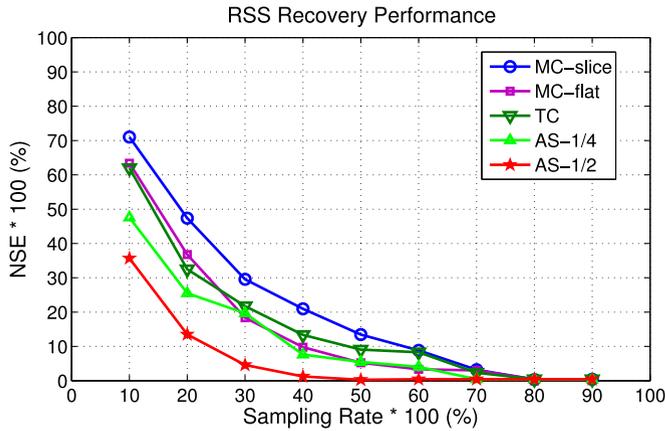


Fig. 6. Tensor recovery for varying sampling rate. MC denotes matrix completion via uniform sampling, TC denotes tensor completion via uniform sampling, while AS-1/4 and AS-1/2 denote our adaptive sampling scheme with allocation ratios $\delta = 1/4$ and $\delta = 1/2$, respectively. MC-flat denotes completion by *flattening* the 3-D data into a matrix of locations \times APs and using matrix completion, an approach followed in [25].

proposed scheme (AS with $\delta = 1/2$) rebuilds a fingerprint data with 5 percent error using less than 30 percent samples.

6.2 Experiment Setup - Real Data

We collected a Wi-Fi RSS data set in the same office. The data set contains 89 selected locations and 31 access points. Since the locations are not exactly on a grid, we set the grid size to be $3 \text{ m} \times 5 \text{ m}$, and apply the KNN method to extract a full third-order tensor as the ground truth. To be specific, for each grid point, we set its RSS vector by averaging the RSS vectors from the nearest three ($k = 3$) locations. The ground truth tensor has dimension $10 \times 10 \times 31$. This tensor serves as a complement to our model-generated data while in the next section, we want to test the localization performance at a finer granularity and covering the whole region of interest.

Radio map recovery performance. Fig. 7 shows the RSS tensor recovery performance for the real-world data set. First, compared with Fig. 6, we see that the recovery performance on real-world data is consistent with that of simulated data. Second, for real-world data set, tensor model is superior to matrix model. In our case, a major ingredient for the recovery improvement may be the large number of access points (i.e., 31), compared with the dimension of the grid (i.e., 10×10). Third, as expected, the propose adaptive scheme achieves better recovery performance.

6.3 Localization Performance

An important factor that influences the localization error is the measurement noise. Note that for site survey, the engineer stays for a while to obtain a stable fingerprint. Therefore, we only consider measurement noise in the query fingerprint. The noise may come from the measuring process or the dynamics in the environment. High SNR and low SNR cases are considered. For the high SNR case we add 3 dBm Gaussian noise, while 10 dBm Gaussian noise is added for the low SNR case.

We choose three representative localization techniques for comparison, namely, weighted KNN, the kernel approach, and support vector machine (SVM). KNN

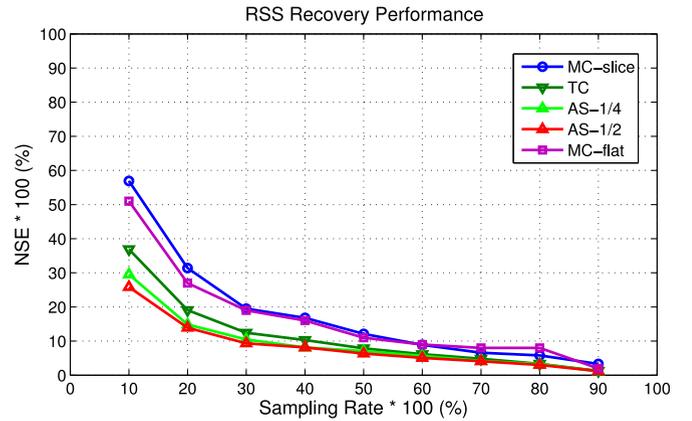


Fig. 7. Tensor recovery in the real-world data set. The same legend as in Fig. 6.

(weighted KNN) is the most widely used technique since it is simple and is reported to have good performance in indoor localization systems [12], [19]. The kernel approach is an improved scheme over weighted KNN, which can be regarded as the basic principle of all machine learning-based localization approaches. The kernel function encapsulates the complicated relationship between RF fingerprints and physical positions. SVM is an efficient machine learning method widely used in fingerprint-based indoor localization. SVM uses kernel functions and tries to learn the complicated relationship between RF fingerprints and physical positions by regression.

In addition to localization based on the estimated tensor \hat{T} , we also test the above three localization techniques on the samples \mathcal{Y} (without doing any reconstruction or estimation). Let DL denote *direct localization* (DL) on uniformly sampled fingerprints, while DL-1/4 and DL-1/2 denote direct localization on adaptively sampled fingerprints with allocation ratios $\delta = 1/4$ and $\delta = 1/2$, respectively.

6.3.1 Weighted KNN

Let k be a fixed positive integer which are usually set to be 1, 3, 5, 7, etc., consider a sampled fingerprint r_p at reference point p . Find within the fingerprint database \mathcal{F} the reference point locations p_1, p_2, \dots, p_k whose fingerprints are nearest to r_p . Then, estimate the location p by weighted averaging p_1, p_2, \dots, p_k as follows:

$$\hat{p} = \sum_{i=1}^k p_i \frac{\frac{1}{d(r_{p_i}, r_p) + d_0}}{\sum_{i=1}^k \frac{1}{d(r_{p_i}, r_p) + d_0}}, \quad (16)$$

where $d(r_{p_i}, r_p)$ is the Euclidean distance between the two fingerprints, and d_0 is a small real constant used to avoid division by zero. In the simulations, we find that $k = 5$ and $d_0 = 0.01$ dBm are the best since the reference points in our experiments are located on a grid map with each reference point having four neighbors (plus itself leads to $k = 5$). Note that we estimate the x-coordinate and y-coordinate separately.

Fig. 8 shows the empirical CDFs of localization error of weighted KNN with 20 percent samples for 3 dBm noise and 10 dBm noise. Besides the recovery schemes TC and MC, we also consider direct localization with samples

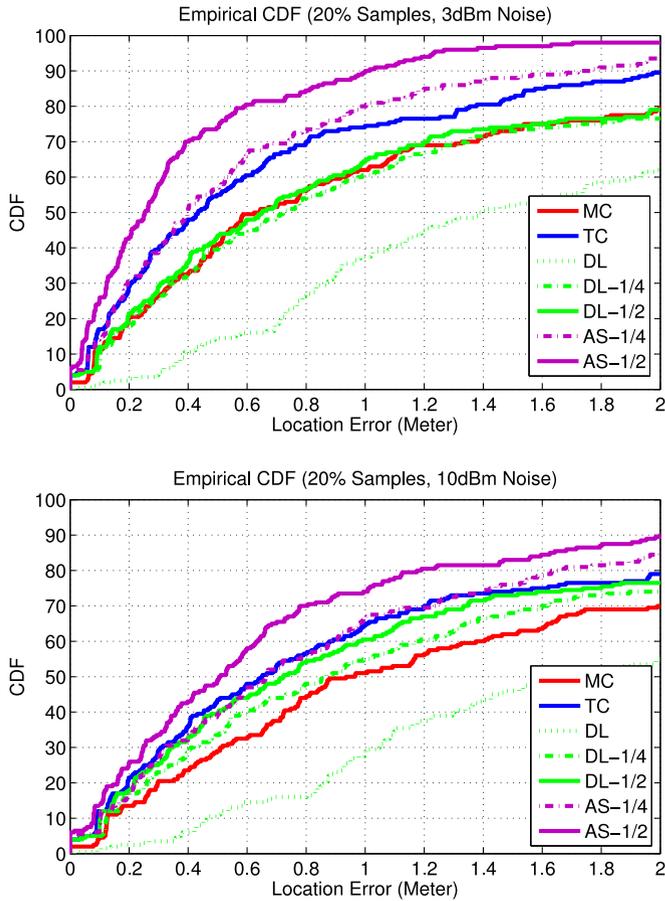


Fig. 8. The empirical CDFs of localization error of KNN with 20 percent samples for 3 dBm noise case and 10 dBm noise case. MC denotes matrix completion via uniform sampling, TC denotes tensor completion via uniform sampling, DL denotes direct localization on uniformly sampled fingerprints, while AS denotes our adaptive sampling scheme. For adaptive DL and AS, there are two allocation ratios $\delta = 1/4$ and $\delta = 1/2$.

drawn uniformly from the grid map and those output by the adaptive sampling approach. We find that the proposed adaptive sampling approach (both $\delta = 1/4$ and $\delta = 1/2$) dramatically outperforms other schemes. In the high SNR case, 97 percent (90 percent respectively) of the reference points can be localized with error less than 2 m (1 m), while 90 percent (95 percent respectively) in the low SNR case. Such performance improvements may be explained by the performance of direct localization (uniform, adaptive with $\delta = 1/4$ and $\delta = 1/2$). Since in both high and low SNR cases, adaptive sampled RF fingerprints are more useful for localization than non-adaptive sampled RF fingerprints. To some extent, it shows that our adaptive sampling approach has identified entries that are highly informative.

6.3.2 Kernel Approach

We estimate the coordinates as $\hat{p} = \sum_{i \in H(p)} \phi(r_{p_i}, r_p) p_i + \alpha_0$, where $H(p)$ denotes the set of reference points whose fingerprints have the smallest Euclidean distances to the query fingerprint, $\phi(\cdot)$ is the kernel function which can be polynomial functions, Gaussian functions, and exponential functions. In the simulations, we set $\phi(\cdot)$ to be the square distance, the cardinality of $H(p)$ be 50. Therefore, we have:

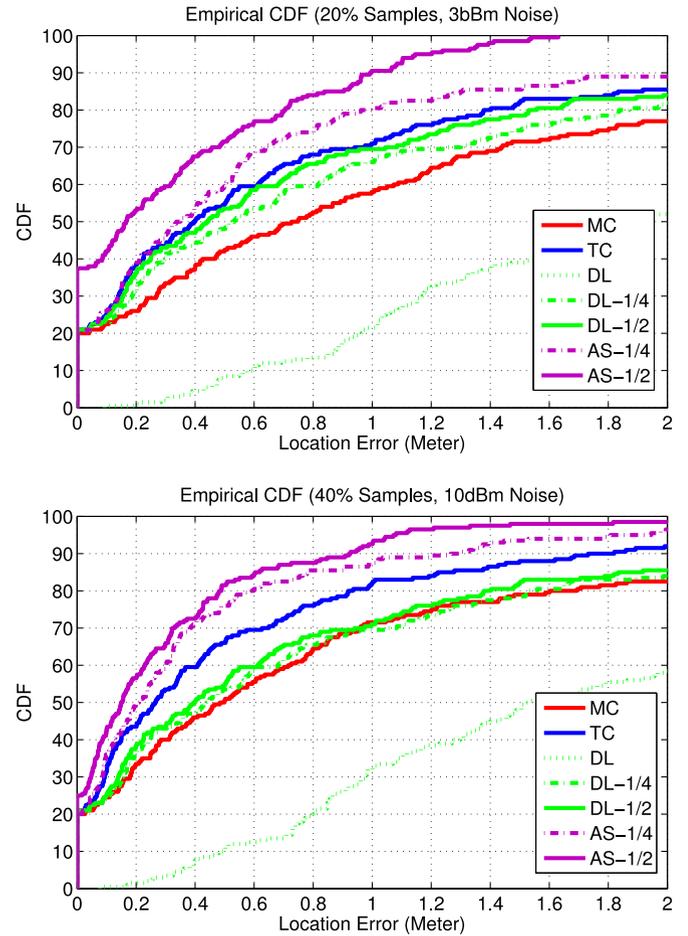


Fig. 9. The empirical CDFs for the kernel approach with the same legend as Fig. 8.

$$\hat{p} = \sum_{i \in H(p)} p_i \frac{\frac{1}{d^2(r_{p_i}, r_p) + d_0^2}}{\sum_{i=1}^k \frac{1}{d^2(r_{p_i}, r_p) + d_0^2}}. \quad (17)$$

We can see the strong similarity with the weighted KNN. A major difference is that the number of reference points used by the kernel method is much larger than the number of neighboring reference points used by weighted KNN.

Fig. 9 shows the empirical CDFs of localization error of the kernel approach with 20 percent samples for 3 and 10 dBm noise. Similar to KNN, the proposed scheme outperforms all other schemes. There are two interesting things to notice. First, for the kernel approach, 38 percent testing points has zero error for adaptive sampling with $\delta = 1/2$ and 3 dBm noise. Second, compared with Fig. 8, the kernel approach is more robust to noise. Surprisingly, the kernel approach performs better in the 10 dBm noise case than in the 3 dBm noise case. The reason is that with lower SNR, the kernel function captures the location centroid of Wi-Fi access points, i.e., the coordinates of access points. Because (17) is dominant by nearby (relative to the user's location) access points, and with 50 ($|H(p)| = 50$) reference points rather than $k = 5$ reference points, (17) essentially captures the location of the dominating access points [18].

6.3.3 SVM Approach

Support vector machines separate reference points using linear hyperplanes in the fingerprint space. The x-coordinate

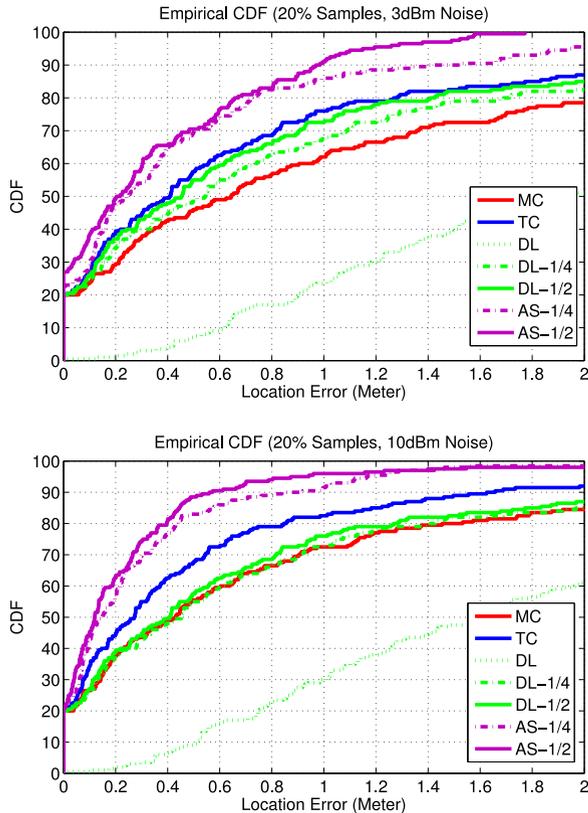


Fig. 10. The empirical CDFs for SVM with the same legend as Fig. 8.

and y -coordinate are estimated separately. Let (x_j, y_j) denote the coordinates of reference point p_j and r_{p_j} denote the fingerprint. For x_j , the training data $\{(x_j, y_j), r_{p_j}\}$ become labeled-pairs $\{\Delta_i, r_{p_i}\}$ where the input is r_{p_i} , and the output is $\Delta_i = 1$ if $x_i = x_j$ and $\Delta_i = -1$ if $x_i \neq x_j$. SVM construct a classifier of the form $\Delta_i(r_{p_i}) = \text{sign}[\sum_{i=1}^{N_1 N_2} w_i \Delta_i \phi(r_{p_i}, r_{p_j}) + b]$, where w_i are positive real constants and b is a real constant. The kernel function $\phi(\cdot, \cdot)$ typically has the following choices: $\phi(r_{p_i}, r_{p_j}) = r_{p_i}^T r_{p_j}$ (linear SVM); $\phi(r_{p_i}, r_{p_j}) = (r_{p_i}^T r_{p_j} + 1)^d$ (polynomial SVM of degree d); $\phi(r_{p_i}, r_{p_j}) = \exp\{-\|r_{p_i} - r_{p_j}\|_2^2 / \sigma^2\}$ (radius basis function kernel SVM); $\phi(r_{p_i}, r_{p_j}) = \tanh[\kappa r_{p_i}^T r_{p_j} + \theta]$ (two-layer neural SVM), where σ, κ, θ are constants. We use linear SVM in the experiments, and according to the structural risk minimization principle [42], we seek the solution of the following optimization problem: $\min_{w, b, \xi} \frac{1}{2} \|w\|_2 + C \sum_{i=1}^N \xi_i$ such that

$$\begin{cases} w^T \cdot r_{p_i} + b \geq 1 - \xi_i, & \text{if } x_j = x_i, \\ w^T \cdot r_{p_i} + b \leq -1 + \xi_i, & \text{if } x_j \neq x_i, \\ \xi_i > 0. \end{cases} \quad (18)$$

where C is the penalty parameter of the error term and ξ_i are introduced in case a separating hyperplane in this higher dimensional space does not exist. We use the SVM library [43].

Fig. 10 shows the empirical CDFs of localization error of the SVM approach with 20 percent samples for 3 and 10 dBm noise. We can see that the performance of SVM is quite similar to the kernel approach in Fig. 9. Notice that for adaptive sampling, the allocation ratio δ does not affect the

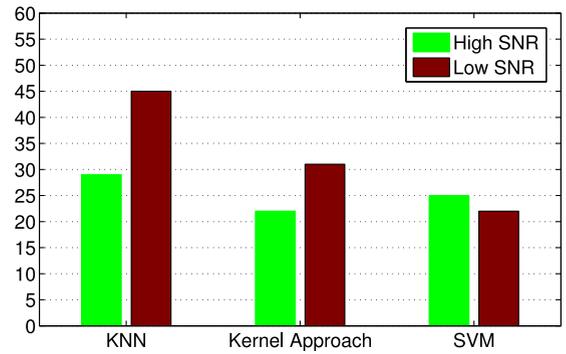


Fig. 11. The sampling budget needed for KNN, kernel approach and SVM.

localization error that much as in the KNN approach and the kernel approach. It seems that for high SNR case, we should allocate more sampling budget to the first-pass sampling, while more sampling budget to the second-pass sampling for low SNR case. Since for high SNR, the first-pass sampling is able to better locate more informative columns of \mathcal{G} (essentially those more informative columns lies near Wi-Fi access points), while for low SNR it is better to have an accurate estimation of the low-dimensional subspace which is the aim of the second-pass sampling.

6.3.4 Reduction of Sampled Reference Points

We are interested in cutting down the sampling budget. We apply the adaptive sampling scheme with $\delta = 1/2$ and then run recovery-and-localization experiment. Through 20 simulations, keeping these with [94% ~ 96%]-percentile localization error being 1m, and calculate the average sampling rate. Fig. 11 shows the results for both the high and low SNR cases. For high SNR, the kernel approach needs the least sampling budget which is 23 percent, while SVM needs 22 percent for low SNR. Maintaining similar localization accuracy of KNN, the amount of samples required by the adaptive sampling approach is cut down by 71 percent for the high SNR case and 55 percent for the low SNR case.

7 CONCLUSIONS

In this paper, an adaptive sampling approach is proposed to relieve the site survey burden of fingerprint-based indoor localization. It cuts down the sampling budget by 71 percent for the high SNR case and 55 percent for the low SNR case while maintaining a similar localization error performance of widely used localization schemes (KNN, the kernel approach and SVM). The performance gain comes from the basic observation that the RF fingerprints are highly correlated across space and across access points and it is possible to adaptively locate more informative entries. Since RF fingerprint calibration is tubal-sampling, we used low tubal-rank tensors to model RF fingerprints instead of low CP-rank, and proposed an algorithm for reconstruct the fingerprint database which adaptively sampling a subset of the reference points and then performing tensor completion. We show that the proposed scheme achieves near-optimal sampling complexity. Our results solve a major challenge of fingerprint-based indoor localization and we advocate its wide adoption in real-world systems.

ACKNOWLEDGMENTS

The authors would like to thank Rittwik Jana, Sarat Puthenpura, and N. K. Shankaranayanan at AT&T Labs-Research, for helpful discussions on propagation models and localization used in this paper. Xiao-Yang Liu gratefully acknowledges financial support from China Scholarship Council. This work was supported in part by the US National Science Foundation (NSF) CCF award 1319653, and NSFC (China) projects - 61373155, 61329101, and 91438121.

REFERENCES

- [1] M. Kilmer, K. Braman, and N. Hao, "Third order tensors as operators on matrices: A theoretical and computational framework with applications in imaging," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 1, pp. 148–172, 2013.
- [2] S. Kang, J. Lee, H. Jang, Y. Lee, and J. Song, "A scalable and energy-efficient context monitoring framework for mobile personal sensor networks," *IEEE Trans. Mobile Comput.*, vol. 9, no. 5, pp. 686–702, May 2010.
- [3] N. Roy, S. K. Das, and C. Julien, "Resource-optimized quality-assured ambiguous context mediation in pervasive environments," *IEEE Trans. Mobile Comput.*, vol. 11, no. 2, pp. 218–229, Feb. 2012.
- [4] F. Adib, Z. Kabelac, D. Katabi, and R. C. Miller, "3D tracking via body radio reflections," in *Proc. 11th USENIX Conf. Netw. Syst. Des. Implementation*, 2014, pp. 317–329.
- [5] I. A. Junglas and R. T. Watson, "Location-based services," *Commun. ACM*, vol. 51, no. 3, pp. 65–69, 2008.
- [6] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu, "E-eyes: Device-free location-oriented activity identification using Fine-grained Wi-Fi signatures," in *Proc. 20th Annu. Int. Conf. Mobile Comput. Netw.*, 2014, pp. 617–628.
- [7] (2013, Mar. 25). Retail indoor location market to reach \$5 billion in 2018, as apple acquires Wi-FiSLAM for \$20 million [Online]. Available: <https://www.abiresearch.com/press/retail-indoor-location-market-to-reach-5-billion-i/>
- [8] L. Liao, D. Fox, J. Hightower, H. Kautz, and D. Schulz, "Voronoi tracking: Location estimation using sparse and noisy sensor data," *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2003, pp. 723–728.
- [9] A.-L. Barabasi, "The origin of bursts and heavy tails in human dynamics," *Nature*, vol. 435.7039, pp. 207–211, 2005.
- [10] S. Gezici, Z. Tian, G. B. Giannakis, H. Kobayashi, A. F. Molisch, H. V. Poor, and Z. Sahinoglu, "Localization via ultra-wideband radios: A look at positioning aspects for future sensor networks," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 70–84, Jul. 2005.
- [11] M. Z. Win, A. Conti, S. Mazuelas, Y. Shen, W. M. Gifford, D. Dardari, and M. Chiani, "Network localization and navigation via cooperation," *IEEE Commun. Mag.*, vol. 49, no. 5, pp. 56–62, May 2011.
- [12] P. Bahl and V. N. Padmanabhan, "RADAR: An In-building RF-based user location and tracking system," in *Proc. IEEE INFOCOM*, 2000, pp. 775–784.
- [13] (2011, Nov. 29). A new frontier for Google Maps: Mapping the indoors [Online]. Available: <http://googleblog.blogspot.com/2011/11/new-frontier-for-google-maps-mapping.html>
- [14] B. Ferris, D. Fox, and N. Lawrence, "Wi-fi-SLAM using gaussian process latent variable models," in *Proc. Int. Joint Conf. Artif. Intell.*, 2007, pp. 2480–2485.
- [15] (2013, Jun.). IndoorAtlas [Online]. Available: <https://www.indooratlas.com/>
- [16] M. Azizyan, I. Constandache, and R. R. Choudhury, "SurroundSense: Mobile phone localization via ambient fingerprinting," in *Proc. 15th Annu. Int. Conf. Mobile Comput. Netw.*, 2009, pp. 261–272.
- [17] B. Li, Y. Wang, H. K. Lee, A. Dempster, and C. Rizos, "Method for yielding a database of location fingerprints in WLAN," *IEE Proc. Commun.*, vol. 152, no. 5, pp. 580–586, Oct. 2005.
- [18] S. P. Kuo and Y. C. Tseng, "Discriminant minimization search for large-scale RF-based localization systems," *IEEE Trans. Mobile Comput.*, vol. 10, no. 2, pp. 291–304, Feb. 2011.
- [19] Z. Yang, C. Wu, and Y. Liu, "Locating in fingerprint space: Wireless indoor localization with little human intervention," in *Proc. 18th Annu. Int. Conf. Mobile Comput. Netw.*, 2012, pp. 269–280.
- [20] R. Gao, M. Zhao, T. Ye, et al., "Jigsaw: Indoor floor plan reconstruction via mobile crowdsensing," in *Proc. ACM MobiCom*, 2014, pp. 249–260.
- [21] A. Rai A, K. Chintalapudi, V. Padmanabhan, and R. Sen, "Zee: Zero-effort crowdsourcing for indoor localization," in *Proc. ACM MobiCom*, 2012, pp. 293–304.
- [22] Y. Zhang, Y. Zhu, M. Lu, and A. Chen, "Using compressive sensing to reduce fingerprint collection for indoor localization," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2013, pp. 4540–4545.
- [23] Y. Hu, W. Zhou, Z. Wen, Y. Sun, and B. Yin, "Efficient radio map construction based on Low-rank approximation for indoor positioning," *Math. Problems Eng.*, vol. 2013, pp. 1–9, 2013.
- [24] D. Milioris, G. Tzagkarakis, and A. Papakonstantinou, "Low-dimensional signal-strength fingerprint-based positioning in wireless LANs," *Ad Hoc Netw.*, vol. 12, pp. 100–114, 2014.
- [25] S. Nikitaki, G. Tsagkatakis, and P. Tsakalides, "Efficient multi-channel signal strength based localization via matrix completion and Bayesian sparse learning," *IEEE Trans. Mobile Comput.*, vol. 14, no. 11, pp. 2244–2256, Nov. 2015.
- [26] S. Nikitaki, G. Tsagkatakis, and P. Tsakalides, "Efficient training for fingerprint based positioning using matrix completion," in *Proc. 20th Eur. Signal Process. Conf.*, Bucharest, Romania, 2012, pp. 195–199.
- [27] D. Milioris, M. Bradonjic, and P. Muhlethaler, "Building complete training maps for indoor location estimation," in *Proc. IEEE INFOCOM*, 2015, pp. 75–76.
- [28] E. J. Candès and T. Tao, "The power of convex relaxation: Near-optimal matrix completion," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2053–2080, May 2010.
- [29] G. Tsagkatakis, and P. Tsakalides, "Efficient recalibration via dynamic matrix completion," in *Proc. 23th IEEE Mach. Learn. Signal Process. Workshop*, Southampton, UK, 2013, pp. 1–6.
- [30] Z. Zhang, G. Ely, S. Aeron, N. Hao, and M. Kilmer, "Novel methods for multilinear data completion and de-noising based on tensor-SVD," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2014, pp. 3842–3849.
- [31] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.
- [32] S. Bhojanapalli and S. Sanghavi, "A new sampling technique for tensors," arXiv:1502.05023, Feb. 2015.
- [33] P. Jain and S. Oh, "Provable tensor factorization with missing data," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 1431–1439.
- [34] S. Gandy, B. Recht, and I. Yamada, "Tensor completion and low-rank tensor recovery via convex optimization," *Inverse Probl.*, vol. 27, no. 2, pp. 1–19, 2011.
- [35] Z. Zhang and S. Aeron, "Exact tensor completion using t-SVD," arXiv:1502.04689v2, Feb. 2015.
- [36] A. Krishnamurthy and A. Singh, "Low-rank matrix and tensor completion via adaptive sampling," *Neural Inf. Process. Syst.*, 2013, pp. 836–844.
- [37] (2013, Aug.). 3D Indoor Standard Ray Tracing (software) [Online]. Available: <http://awe-communications.com/Propagation/Indoor/SRT/index.htm>
- [38] A. Kaya, L. Greenstein, and W. Trappe, "Characterizing indoor wireless channels via ray tracing combined with stochastic modeling," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4165–4175, Aug. 2009.
- [39] A. Deshpande, L. Rademacher, S. Vempala, and G. Wang, "Matrix approximation and projective clustering via volume sampling," in *Proc. 17th Annu. ACM-SIAM Symp. Discrete Algorithm*, 2006, pp. 1117–1126.
- [40] L. Kong, M. Xia, X.-Y. Liu, M.-Y. Wu, and X. Liu, "Data loss and reconstruction in sensor networks," in *Proc. IEEE INFOCOM*, 2013, pp. 1654–1662.
- [41] P. Jain, P. Netrapalli, and S. Sanghavi, "Low-rank matrix completion using alternating minimization," in *Proc. 45th Annu. ACM Symp. Theory Comput.*, 2013, pp. 665–674.
- [42] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Process. Lett.*, vol. 9, no. 3, pp. 293–300, 1999.
- [43] C.-C. Chang and C.-J. Lin, "A library for support vector machines" *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, 2011, [Online]. Available: <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [44] L. Balzano, B. Recht, and R. Nowak, "High-dimensional matched subspace detection when data are missing," in *Proc. IEEE ISIT*, 13–18 Jun. 2010, pp. 1638–1642.

- [45] Y. Wang and A. Singh, "Column subset selection with missing data via active sampling," in *Proc. 18th Int. Conf. Artif. Intell. Statist.*, 2015, pp. 1033–1041.



Xiao-Yang Liu received the BEng degree in computer science from the Huazhong University of Science and Technology, China, in 2010. He is currently working toward the PhD degree in the Department of Computer Science and Engineering at Shanghai Jiao Tong University. His research interests include wireless communication, distributed systems, big data analysis, cyber security, and sparse optimization.



Shuchin Aeron (M'08-SM'15) received the PhD degree in ECE from Boston University in 2009. He is currently an assistant professor in the Department of ECE at Tufts University. From 2009 to 2011, he was a post-doctoral research fellow at Schlumberger-Doll Research (SDR), Cambridge, MA, where he worked on signal processing answer products for borehole acoustics. He has several patents and the proposed workflows are currently implemented in logging while drilling tools. He is the recipient of the best thesis award from both the College of Engineering and the Department of ECE in 2009. He received the Center of Information and Systems Engineering (CISE) award from Boston University in 2006 and a Schlumberger-Doll Research grant in 2007. His main research interest lie in statistical signal processing, information theory, and optimal sampling and recovery. He is a senior member of the IEEE.



Vaneet Aggarwal (S'08-M'11-SM'15) received the BTech degree in 2005 from the Indian Institute of Technology, Kanpur, India, and the MA and PhD degrees in 2007 and 2010, respectively from Princeton University, Princeton, NJ, all in electrical engineering. He is currently an assistant professor at Purdue University, West Lafayette, IN. Prior to this, he was a senior member of Technical Staff Research at AT&T Labs-Research, NJ, and an adjunct assistant professor at Columbia University, NY. His research interests are in

applications of information and coding theory to wireless systems, machine learning, and distributed storage systems. He was the recipient of Princeton University's Porter Ogden Jacobus Honorific Fellowship in 2009. He is a senior member of the IEEE.



Xiaodong Wang (S'98-M'98-SM'04-F'08) received the PhD degree in electrical engineering from Princeton University. He is a professor of electrical engineering at Columbia University in New York. His research interests fall in the general areas of computing, signal processing, and communications, and has published extensively in these areas. Among his publications is a book entitled *Wireless Communication Systems: Advanced Techniques for Signal Reception*, published by Prentice Hall in 2003. His current research interests include wireless communications, statistical signal processing, and genomic signal processing. He received the 1999 US National Science Foundation (NSF) CAREER Award, the 2001 IEEE Communications Society and Information Theory Society Joint Paper Award, and the 2011 IEEE Communication Society Award for Outstanding Paper on New Communication Topics. He has served as an associate editor for the *IEEE Transactions on Communications*, the *IEEE Transactions on Wireless Communications*, the *IEEE Transactions on Signal Processing*, and the *IEEE Transactions on Information Theory*. He is a fellow of the IEEE and listed as an ISI Highly-cited Author.



Min-You Wu (S'84-M'85-SM'96) received the MS degree from the Graduate School of Academia Sinica, Beijing, China, in 1981 and the PhD degree from Santa Clara University, CA in 1984. He is a professor in the Department of Computer Science and Engineering at Shanghai Jiao Tong University and a research professor with the University of New Mexico. He serves as the chief scientist at the Grid Center of Shanghai Jiao Tong University. His research interests include wireless networks, sensor networks, multimedia network-

ing, parallel and distributed systems, and compilers. He is a senior member of the IEEE.

▷ For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.