UNLABELED SENSING: RECONSTRUCTION ALGORITHM AND THEORETICAL GUARANTEES

Golnoosh Elhami, Adam Scholefield, Benjamín Béjar Haro and Martin Vetterli

School of Computer and Communication Sciences
École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland
{golnoosh.elhami, adam.scholefield, benjamin.bejarharo, martin.vetterli}@epfl.ch

ABSTRACT

It often happens that we are interested in reconstructing an unknown signal from partial measurements. Also, it is typically assumed that the location (temporal or spatial) of each sample is known and that the only distortion present in the observations is due to additive measurement noise. However, there are some applications where such location information is lost. In this paper, we consider the situation in which the order of noisy samples, taken from a linear measurement system, is missing. Previous work on this topic has only considered the noiseless case and exhaustive search combinatorial algorithms. We propose a much more efficient algorithm based on a geometrical viewpoint of the problem. We also study the uniqueness of the solution under different choices of the sampling matrix and its robustness to noise for the case of two-dimensional signals. Finally we provide simulation results to confirm the theoretical findings of the paper.

Index Terms— Unlabeled sensing, Linear sampling, Compressed sensing, SLAM
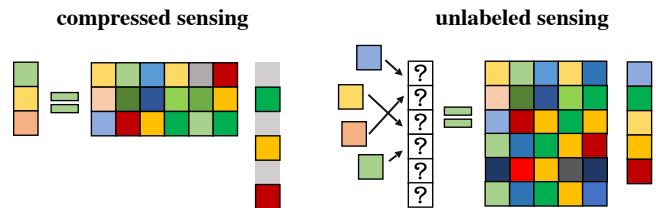
1. INTRODUCTION

Suppose you would like to measure a spatial field, such as the temperature in a room, with a moving robot that collects linear measurements of the field at some unknown spatial locations. Therefore, at the end of the measurement phase, a set of sample values with some potential locations are accessible [1–3]. The ultimate goal is to recover the underlying spatial field which, at the same time, requires knowledge about the sampling locations. In a discrete setup, this problem is called *unlabeled sensing* [4] and can be seen as a instance of the famous *simultaneous localisation and mapping* (SLAM) problem [5, 6] which has applications ranging from self-driving cars to indoor localisation.

Typically, in a linear inverse problem, we try to solve a system of the form

$$y = \Phi x, \qquad (1)$$

where $x \in \mathbb{R}^K$ represents the original unknown data vector, $y \in \mathbb{R}^N$ is a set of observed sample values, and $\Phi \in \mathbb{R}^{N \times K}$ is a tall sampling matrix. Note that in (1) the sampling matrix Φ is assumed to be perfectly known. Therefore, in the absence of any prior knowledge on x and assuming the problem is well-posed, we can find the least-squares estimate of x by using the pseudo-inverse of Φ. In unlabeled sensing however, we try to recover the original signal x given *unordered* measurements in y. Clearly, in this case, we cannot solve (1) with a simple matrix inversion. A natural question is under what conditions is the solution guaranteed to be unique, or, when this is not possible, can we find a solution and a labeling (ordering) close to the true values.

Unlabeled sensing is similar to compressed sensing in the sense that they both deal with partially revealed information in a linear measurement system. Compressed sensing [7, 8] states that if we have some prior knowledge about the sparsity of the input x, we might be able to reconstruct it uniquely when $N < K$. The sparsity assumption enables us to model real life applications using compressed sensing [9, 10]. On the other hand, in unlabeled sensing there is no specific assumption on the sparsity of data x and the missing part of the information is the correct order of sample values in y. Without any assumptions on the structure of x, we naturally need more measurements to compensate for the missing labels; thus a fat matrix for compressed sensing, and a tall matrix for the unlabeled sensing. Figure 1 illustrates the connection between the two problems.

The unlabeled sensing problem with restriction to band-limited signals has been studied in [11]. In this work, we continue on the introduction of Unnikrishnan et al. to unlabeled sensing without any specific assumptions on the data model [4]. The authors in [4] provide necessary and sufficient conditions on matrix Φ for unique reconstruction of the original signal with probability one in the absence of noise. There, the focus of the work is on random sampling matrices Φ and the reconstruction algorithm is assumed to be a simple combinatorial exhaustive search. In this paper, we provide an efficient algorithm for the reconstruction of the original signal, together with the complexity of the algorithm. We further extend our reconstruction algorithm to the case where the measurements are contaminated with bounded noise and focus on finding sampling matrices with robust behavior in the presence of noise. In doing so, we restrict our attention to the recovery of signals in 2D. We study the uniqueness of the reconstruction for different choices of the sampling matrix Φ. More precisely, we show that when the rows of the sampling matrix Φ correspond to equally-spaced vectors in the plane there are always multiple solutions, even in the noiseless case. We then

Fig. 1: Comparison between compressed sensing and unlabeled sensing [4].

propose a specific sampling scheme, showing that the reconstruction is unique in the noiseless case, and provide an upper bound on the reconstruction error for the noisy case. We also provide numerical simulations to illustrate and validate our findings.

2. PROBLEM STATEMENT

Consider a linear sampling system,

$$\boldsymbol{y} = \boldsymbol{P}\boldsymbol{\Phi}\boldsymbol{x} + \boldsymbol{\epsilon}, \tag{2}$$

where $\boldsymbol{x} \in \mathbb{R}^K$ represents the unknown original data vector, $\boldsymbol{y} \in \mathbb{R}^M$ is a set of noisy measured sample values, $\boldsymbol{\Phi} \in \mathbb{R}^{N \times K}$ is the known sampling matrix, $\boldsymbol{\epsilon}$ is the measurement noise and $\boldsymbol{P}$ is an unknown $M \times N$ ($M \le N$) selection matrix whose rows correspond to distinct rows of the $N \times N$ identity matrix. In this contribution, we will restrict our analysis to the case where $M = N$. Therefore, $\boldsymbol{P}$ corresponds to a permutation matrix which permutes the rows in $\boldsymbol{\Phi}$ to an arbitrary unknown order.

The authors in [4] show that if $\boldsymbol{\epsilon} = \boldsymbol{0}$ and $\boldsymbol{\Phi}$ has independent and identically distributed (i.i.d.) random entries drawn from an arbitrary continuous distribution, provided that $N \ge 2K$, every vector $\boldsymbol{x}$ can be uniquely recovered from $\boldsymbol{y}$ with probability one. It is assumed that the recovery is performed through a simple combinatorial exhaustive search.

In the following, we go further than random sampling matrices, consider the noisy case where $\boldsymbol{\epsilon} \ne \boldsymbol{0}$ and provide a novel and efficient algorithm for recovering $\boldsymbol{x}$.

3. UNLABELED SENSING RECONSTRUCTION

The focus of earlier papers on unlabeled sensing has been mostly on the uniqueness studies and not on the reconstruction algorithms [4]. Normally an exhaustive search has been considered for reconstruction. In this section, we introduce an efficient reconstruction algorithm with polynomial complexity in contrast to previous combinatorial attempts. We call it the *geometrical reconstruction* (GR) algorithm, since it is based on a geometrical view of the problem.

For simplicity, we start by explaining a simple example in 2D. Figure 2 shows the specific case of $N = 3$ and $K = 2$ ($\boldsymbol{\Phi} \in \mathbb{R}^{3 \times 2}$).

The algorithm in 2D starts by choosing two sample values, say y_1 and y_2, and assigns them to two sampling vectors, say n and m. Given these two measurements and the selected labelling, we can invert the system to find a candidate solution. In the 2D case, as depicted in Figure 2b, this simply corresponds to intersecting two lines, L_1^n and L_2^m, with distances y_1 and y_2 from the origin and perpendicular to vectors n and m.

We repeat this for all possible labellings of y_1 and y_2; i.e., for all $m, n \in \{1, \cdots, N\}$. This results in $N(N-1)$ candidate solutions, which are put in the feasible set $\mathcal{F}$. This set is guaranteed to contain the original point. For our simple example, as depicted in Figure 2b, at this point, the set $\mathcal{F}$ contains six points.

The rest of this algorithm, uses the other measurements (y_3, $\cdots$, y_N) to reduce this feasible set as much as possible. Ideally, there will be a unique solution; however, as we will see in later sections, often multiple solutions are consistent with the measurements. In this case, the algorithm will return all feasible solutions.

To reduce the feasible set, we, iteratively, consider each sample value y_n for $3 \le n \le N$. For each sample value, we check each point in $\mathcal{F}$ to see if its projected value, on any of the sampling vectors, is equal to y_n. As depicted in Figure 2c, this can be done very easily in 2D. For each sample value, we draw two tangent lines from each point in $\mathcal{F}$ to the circle $(0, y_n)$. If the tangent points on the circle lie on any of the sampling vectors, we keep the corresponding

point; otherwise, remove it from $\mathcal{F}$. This step plays a major role in reducing the complexity of the algorithm. Figure 2d depicts the single point left in the feasible set, after considering the third measurement.

Algorithm 1 explains this procedure for the general K dimensional case.

3.1. Complexity Analysis

The worst case complexity of the proposed GR algorithm is $O(KN^{K+1})$ compared to $O(N!(K^3 + NK))$ for a naive exhaustive search. To see this note that in step 4 of Algorithm 1, we have to find the inverse of $N(N-1)\cdots(N-K+1)$ different $K \times K$ matrices. The complexity for solving each equation system is $O(K^3 + K^2)$. Therefore, the total complexity of this step is $O([K^3 + K^2][N(N-1)\dots(N-K+1)])$, or simply $O((K^3 + K^2)N^K)$.

Additionally, in step 9, we check if the $N(N-1)\cdots(N-K+1)$ candidate points have projections equal to one of the sample values over one of the remaining vectors. This takes $(N-K)K$ multiplications for each candidate.

In total we have $O([K(N-K)][N(N-1)\dots(N-D+1)])$, or $O(KN^{K+1})$. A more formal analysis is given in [12].

4. RECONSTRUCTION ROBUSTNESS

The authors in [4] show that if we use random measurement matrices with an oversampling of 2 or more, the solution is unique with probability one. This statement is true when we have random measurement matrices and noiseless samples. In the following we see that, when the measurement matrix is not random or in the presence of noise, we can have non-unique solutions with non-zero measure. Later, we focus on a sampling matrix which offers unique solutions in the noiseless case and has robust behavior in the presence of noise.

4.1. Uniqueness

We will consider the uniqueness in the 2D case ($K = 2$). We believe that the results can be extended to higher dimensions as well.

In 2D, one can represent the rows of the sampling matrix with vectors on the plane, each having an angle to the x axis, φ_i:

$$\boldsymbol{\Phi} = \begin{bmatrix} \cos(\varphi_1) & \cos(\varphi_2) & \cdots & \cos(\varphi_N) \\ \sin(\varphi_1) & \sin(\varphi_2) & \cdots & \sin(\varphi_N) \end{bmatrix}^T .$$


Algorithm 1 GR Algorithm in K dimensions without noise

1: **Initialize:** $\mathcal{P} = \emptyset$, feasible set: $\mathcal{F} = \emptyset$.
2: Take every $N(N-1)\cdots(N-K+1)$ selection of K rows of $\boldsymbol{\Phi}$, call them $\boldsymbol{\Phi}_i$ and put them in set $\mathcal{P}$.
3: **for** each $\boldsymbol{\Phi}_i \in \mathcal{P}$, $i = 1, \cdots, N(N-1)\cdots(N-K+1)$ **do**
4:      Solve $K \times K$ equations $[y_1, \cdots, y_K]^T = \boldsymbol{\Phi}_i \boldsymbol{x}^i$.
5:      $\mathcal{F} \leftarrow \mathcal{F} \cup \{\boldsymbol{x}^i\}$.
6: **end for**
7: **for** each sample value y_n, $n = K+1, \cdots, N$ **do**
8:      **for** every $\boldsymbol{x}^i \in \mathcal{F}$ **do**
9:          Check if $\boldsymbol{x}^i$ has the projection value equal to y_n over any of the remaining $N - (n-1)$ rows of $\boldsymbol{\Phi}$.
10:          **if** such a row is not found **then** $\mathcal{F} \leftarrow \mathcal{F} \setminus \{\boldsymbol{x}^i\}$.
11:          **end if**
12:      **end for**
13: **end for**
14: **return** $\boldsymbol{x} \in \mathcal{F}$.
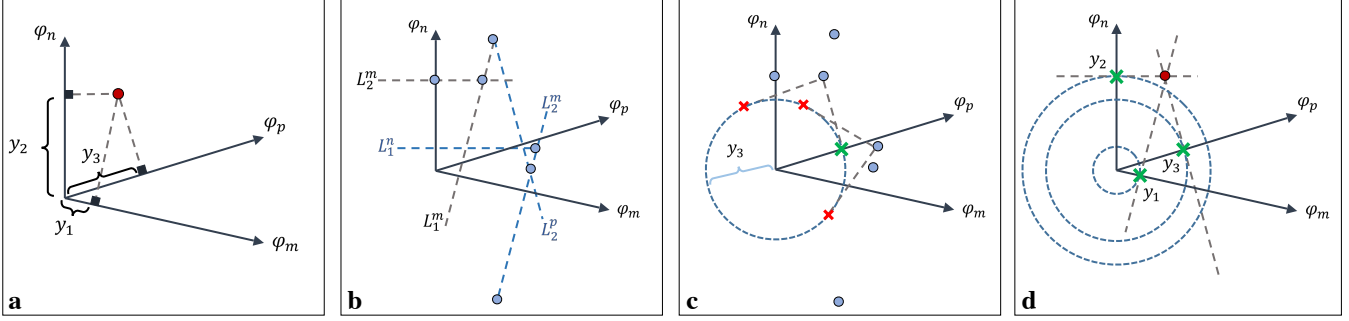
Fig. 2: Geometrical view of the reconstruction. **a)** original data point and its projections. **b)** construction of the feasible set $\mathcal{F}$. **c)** tangent lines from candidate points to the circle $(0, y_3)$. **d)** only one point has a tangent lying on a sampling vector.
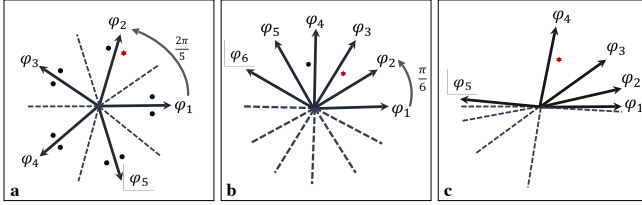
Fig. 3: Solutions of a) 5 equally-spaced, b) 6 equally spaced and c) 5 exponentially-spaced settings. The original point is shown by a red star while black points lead to the same sample values.

Suppose that, as depicted in Figures 3a and 3b, the N sampling vectors are equally spaced. In other words, $\varphi_i = \pi(i-1)/N$, for even N, and $\varphi_i = 2\pi(i-1)/N$, for odd N. Here, $i \in \{1, \cdots, N\}$. These sampling vectors have an appealing symmetry which could be favorable in noisy scenarios. However, it is easy to see that these sampling matrices do not offer unique solutions in the noiseless case. Figure 3 shows these two settings and their non-unique solutions. When N is odd, we have multiple solutions because of the symmetry in rotation and reflection. This results in $2N$ solutions. When N is even, we have a symmetry with respect to the reflection line of the vectors, which results in 2 solutions.

In contrast to the random sampling setting where the non-unique solutions had a measure zero, in these settings for every point in space, we have more than one solution to (2). With these examples in mind we study the following sampling setting which offers a unique solution to (2) and has a robust behavior in the presence of noise.

Suppose that the N sampling vectors are distributed with exponential spacing between each other, that is

$$\varphi_i = \frac{2^{i-1}-1}{2^N-1} 2\pi, \quad i \in \{1, \cdots, N\}. \tag{3}$$

It can be shown that, for the exponentially-spaced sampling setting, every $\boldsymbol{x}$ will not be confused with any other point in the plane. The exponentially-spaced setting achieves this by eliminating the rotational and reflective symmetries, which plagued the uniformly spaced vectors.

Note, however, that if a point lies on a bi-sector of two sampling vectors, it will still have two different labellings, even though both labellings lead to the same point.

These results are formalised in the following lemma, which is proved in [12]:

Lemma 1. *Consider the sampling matrix $\boldsymbol{\Phi}$ with $N \geq 4$ exponentially-spaced sampling vectors. Suppose that for some $\boldsymbol{x}$ and $\boldsymbol{x}'$, there exists a permutation matrix $\boldsymbol{P} \neq \boldsymbol{I}$ such that $\boldsymbol{\Phi x} = \boldsymbol{P \Phi x}'$. If*

$\|\boldsymbol{x}\| = \|\boldsymbol{x}'\|$, *then $\boldsymbol{x}$ and $\boldsymbol{x}'$ coincide and lie on the bi-sector of two of the sampling vectors.*

The next natural question to ask is how robust this sampling and reconstruction scheme is in the presence of noise.

4.2. Noise

In order to study the robustness of the sampling settings and the algorithms, we first generalize our proposed algorithm in 2D to the noisy case.

Here we assume that the noise is bounded between $[-\epsilon_{\max}, \epsilon_{\max}]$. After constructing the feasible set $\mathcal{F}$ like in the GR algorithm, we remove those points whose projections are further than a certain distance threshold from the remaining samples. Then, for each of these possible labelings, we draw bands of width $2\epsilon_{\max}$ normal to the vectors and centered on the remaining sample values. The intersection of these bands defines a convex region which we call a *consistent region*. If the intersection is empty, we remove that labeling from the candidate set. At the end, all the points in the consistent region, could produce the given samples with the bounded noise. We skip explaining the details of the generalized algorithm in the interest of space.

In Figure 4 we show examples of consistent regions for different sampling structures.

Now we are ready to show the robustness of the exponentially-spaced setting in presence of noise.

Lemma 2. *Consider the exponentially-spaced vector setting (3) with $\|\boldsymbol{\epsilon}\|_\infty \leq \epsilon_{max}$. If there exist two solutions $\boldsymbol{x}$ and $\boldsymbol{x}'$ with different labelings, then*

$$\|\boldsymbol{x} - \boldsymbol{x}'\| \leq \frac{4\epsilon_{max}}{\cos\left(\frac{2^{N-1}-1}{2^N-1}\pi\right)}. \tag{4}$$

Proof. Refer to [12] for the proof. □

Lemma 2 provides an upper bound on the distance between the consistent regions as the outputs of the generalized GR algorithm. It also offers a criteria for choosing the number of sampling vectors N for a given robustness. This result brings a trade-off between having more vectors for collecting more samples, and having less vectors for a better robustness bound. We leave finding the optimal number of sampling vectors to future work.

The above lemma combined with the fact that points on the bi-sectors of the vectors produce multiple labelings, show that the biggest deviations between the reconstruction will concentrate around the bi-sectors of the sampling vectors which are closest to each other (have smaller angles). We verify this result in the next section.

(a) 5 equally spaced   (b) 6 equally spaced

(c) 5 exponentially spaced   (d) 5 randomly spaced
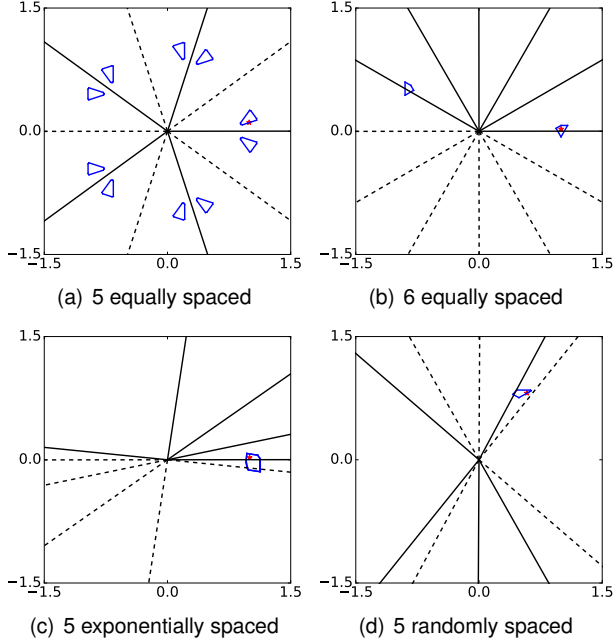
Fig. 4: Consistent regions found by the generalized GR algorithm for SNR = 25 dB. The original point is shown by a red star.

5. SIMULATION RESULTS

We start this section with an attempt to illustrate what the consistent regions look like in different sampling settings. We consider the problem in 2D and take 5 vectors (6 for evenly spaced sampling vectors) which are either equally spaced, exponentially spaced or randomly spaced. We suppose that the noise vector ϵ has i.i.d. elements uniformly distributed in $[-\epsilon_{max}, \epsilon_{max}]$. We assume that $\|x\| = 1$, and with the uniform noise define SNR $= -10\log(\epsilon_{max}^2/3)$. Figure 4 shows the output of the generalized GR algorithm for the four mentioned sampling setups with SNR = 25 dB. As expected, the number of consistent regions is $2N = 10$ for the 5 equally spaced setting and 2 for the 6 equally spaced setting. Also, in accordance with theoretical findings, the consistent regions are unique for the points that lie outside of the bi-sector regions in the exponentially and randomly spaced settings. In the proposed exponentially spaced setting, however, we are also able to bound the distance between consistent regions in (4).

In the next simulation we aim at confirming the results of Lemma 2. We consider 5 exponentially spaced sampling vectors. We change the SNR in the range of 10 dB to 80 dB. For each value of the SNR we run 20000 trials and compute the probability that all the points in the consistent regions are closer to the original x than a given bound (e.g. $\|x - x'\| \leq 0.2$). Figure 5 shows the simulation result for different bounds. As the figure suggests, when the value of the SNR grows, consistent regions shrink and get closer to the original x, and thus with probability one all the consistent regions are in the vicinity of the original point. In comparison to the result of Lemma 2, we compute the SNR that would correspond to the RHS of (4) to be equal to 0.2. We show this SNR by a dashed vertical line in Figure 5. As we expected, this SNR value is larger than the SNR at which the probability reaches one (around 30 dB). This is because the bound in (4) considers the worst case by taking only two axes and not considering the intersection of regions from different axes.

Finally, we run a simple simulation in which we change the original data x for a fixed $\epsilon_{max} = 0.02$ and highlight the points in space
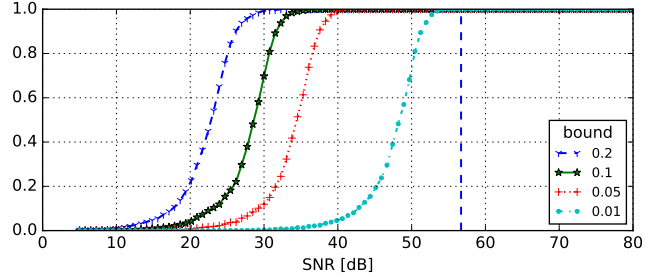
Fig. 5: Probability that all the points in the consistent regions are closer to x than a given bound (for different values of such bounds). The vertical dashed line shows the SNR which would make the RHS of (4) equal to 0.2.
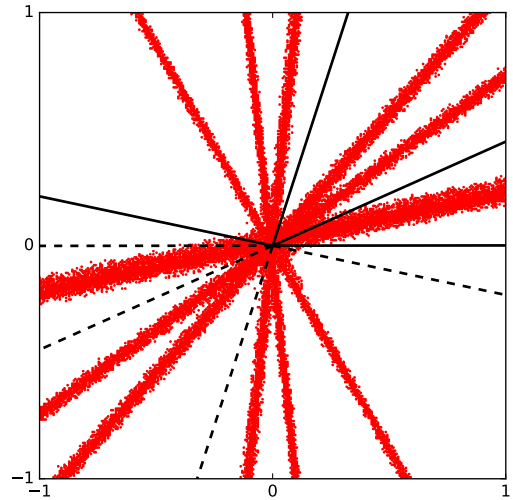
Fig. 6: When the original point x is a red dot, there are more than one labelings for $\epsilon_{max} = 0.02$ and exponentially-spaced vectors.

for which the algorithm finds more than one labeling. The results are depicted in Figure 6. Observe that the highlighted regions concentrate around the bi-sectors of pairs of vectors. Also, the closer the two sampling vectors, the bigger the region around their bi-sector. Again, we remind the reader that these regions are related to the points with several possible labelings. Therefore, if the point lies in one of the highlighted regions, given a set of noisy samples, we are able to find the points of interest with several valid labelings using the generalized GR algorithm. Otherwise, if the point is outside these regions, there is only one valid labeling and we can recover the labeling and the related consistent region by the proposed algorithm.

6. CONCLUSION

In this paper, we proposed an efficient algorithm with polynomial complexity for data reconstruction in the unlabeled sensing problem. We showed that some sampling settings result in multiple solutions. In the search for unique solutions, we designed a sampling scheme which offers unique solutions in absence of noise. We extended our algorithm to the noisy case in 2D and showed that our proposed sampling scheme is robust under noise. Extending the algorithm to noisy cases in higher dimensions remains as future work.

References

[1] R. Smith, M. Self, and P. Cheeseman, "Estimating uncertain spatial relationships in robotics," in *Autonomous robot vehicles*, pp. 167–193, Springer, 1990.

[2] J. J. Leonard and H. F. Durrant-Whyte, "Mobile robot localization by tracking geometric beacons," *IEEE Transactions on Robotics and Automation*, vol. 7, pp. 376–382, Jun 1991.

[3] J. Unnikrishnan and M. Vetterli, "Sampling and reconstruction of spatial fields using mobile sensors," *IEEE Transactions on Signal Processing*, vol. 61, pp. 2328–2340, May 2013.

[4] J. Unnikrishnan, S. Haghighatshoar, and M. Vetterli, "Unlabeled sensing with random linear measurements," *arXiv:1512.00115*, 2015.

[5] S. Thrun and J. J. Leonard, "Simultaneous localization and mapping," in *Springer handbook of robotics*, pp. 871–889, Springer, 2008.

[6] S. Thrun and M. Montemerlo, "The graph SLAM algorithm with applications to large-scale mapping of urban structures," *The International Journal of Robotics Research*, vol. 25, no. 5-6, pp. 403–429, 2006.

[7] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, pp. 1289–1306, April 2006.

[8] E. Candes, "Compressive sampling," in *Int. Congress of Mathematics*, pp. 1433–1452, 2006.

[9] J. Yoo, S. Becker, M. Monge, M. Loh, E. Candes, and A. Emami-Neyestanak, "Design and implementation of a fully integrated compressed-sensing signal acquisition system," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5325–5328, 2012.

[10] Y. Chi, "Joint sparsity recovery for spectral compressed sensing," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3938–3942, 2014.

[11] P. Marziliano and M. Vetterli, "Reconstruction of irregularly sampled discrete-time bandlimited signals with unknown sampling locations," *IEEE Transactions on Signal Processing*, vol. 48, no. 12, pp. 3462–3471, 2000.

[12] G. Elhami, "Unlabeled sensing: Algorithm and guarantees," tech. rep., EPFL, 2016. http://infoscience.epfl.ch/record/221297.