Reconstruction of Ultrasound Signals Using Randomly Acquired Samples in a Sparse Environment ⋆

Samuel Pinto * **Sean R. Sanchez** * **Liam Doran** * **Aidan Ryan** *
Sean B. Andersson *,**

* *Department of Mechanical Engineering*
** *Division of Systems Engineering*
Boston University, Boston, MA 02215 USA.
e-mail:{samcerq,srsanche,ldoran,aidanjr,sanderss}@bu.edu

Abstract: Sonar echoes can provide more than only range information, but recording the full sonar echo is challenging in resource constrained systems. This paper introduces an approach for reconstructing under-sampled sonar echo signals in environments that are not cluttered using Compressive Sensing. This technique requires sampling only around 20% of the total samples in order to achieve good reconstruction results. An experimental validation of the approach is presented.

Keywords: Compressive sensing, compression, sonar, robotics, mapping.

1. INTRODUCTION

One common goal in the problem of autonomous exploration is to take measurements of the environment being explored and to compile the information into a coherent (and ideally accurate) picture or map of the surroundings. If the pose of the robot is also unknown *a priori*, the problem is called Simultaneous Localization and Mapping (SLAM). SLAM is a fairly mature field with a deep and extensive literature; an overview of its many flavors can be found in, *e.g.*, (Cadena et al., 2016). While many of the algorithms for translating sensor data into maps are generalized in that they do not depend on the details of any given sensor, one common feature is that they rely on large amounts of data to yield accurate and reliable maps. In practice, mapping is often performed using visual sensors such as cameras and laser rangefinders which naturally produce the large volume of data needed.

Despite its proven efficacy, this familiar approach to mapping is not equally appropriate in all situations, due in part to the size and cost of the sensors, and in part to the power and storage required to gather and process data in real time. In recent years, there has been increasing interest and ongoing developments in very small, low power robots such as autonomous flying micro-robots (RoboBees) (Wood et al., 2013), micro air vehicles (Scaramuzza et al., 2014), bat-like robots (Colorado et al., 2015), and insect-scale flapping wing robots (Zou et al., 2018). Such vehicles do not have the payload capability for large sensors such as laser rangefinders nor the energy or storage budget for high resolution cameras. Additionally, given their small size, it is natural to consider deploying many such robots in an environment to collect data and then return to a central hub to generate a large scale map.

⋆ This work was supported in part by NSF through CMMI-1562031 and ECCS-1509084.

Subject to these limitations, ideally one needs a light, low-powered sensor along with algorithms which can still produce good maps from potentially heavily undersampled data. In light of this, we focus on the use of ultrasonic sensors. These are low-cost, light-weight sensors which have been used in both mapping and SLAM (Jung et al., 2009). In general, they are used to measure the range to the nearest target, but, unlike laser rangefinders, their signal expands through an approximately conic domain in space. While this can make the location of a specific target ambiguous, there is also significantly more information that can be extracted from the received signal than simply the range to the nearest object. This can be seen in the medical environment, where ultrasound can be used for safe, in-body imaging (Quistgaard, 1997). The application of ultrasound in mapping was expanded upon in recent work to infer the shape of a polyhedral room from a single source and multiple sensors (that is, to "hear the shape of a room") (Dokmanic et al., 2013) and on performing SLAM using these echoes (Krekovic et al., 2016).

In previous work by two of the authors (Sanchez and Andersson, 2018), an optimization approach was used to extract information from a single ultrasound pulse. The aim of that work was to determine (possible) locations of a sparse set of point-like objects within the cone of the ultrasound signal, in essence attempting to create a local map of the region within that cone. That work demonstrated the viability of recovering range information from multiple targets using a single ultrasonic echo signal and some of the advantages of having the full echo signal instead of only range information.

In this work, instead of focusing on translating the echo signal into a local map, we aim towards making the echo signal acquisition feasible for low power and low memory-capacity systems through the use of Compressive Sensing

(CS). The essential idea of CS is to leverage the fact that most signals, when expressed in an appropriate basis, are approximately sparse. Because of this, a small (sub-Nyquist rate) number of measurements can be acquired and reconstruction algorithms used offline that, under certain assumptions, can exactly recreate the original signal (Candès and Wakin, 2008). Because the signal can be highly undersampled, the approach can greatly reduce the amount of memory needed to store a signal without the need for computationally expensive data compression. In this work in particular, the sampling strategy is to choose at each sample time whether or not, with a given probability, to acquire a measurement. Using CS reconstruction algorithms offline, we aim to obtain a good approximation of the original signal.

In the next section, we state our problem more formally and describe our approach to addressing it. There, we give a brief overview of CS and explain how sensing using an ultrasonic transducer fits into this framework and how reconstruction is achieved. In Sec. 3 we describe the experimental setup and our scheme for creating bases for a sparse description of the ultrasound echo signal. We develop different bases, one by measuring the signal from one ultrasonic sensor pointed directly at another (to represent a point source), one from the return echoes off a small, flat target and additional overcomplete bases described in Sec. 3.2. The experimental results are described in Sec. 4, including a comparison of the reconstruction quality between the different sparsity bases and using two different standard CS reconstruction algorithms, basis pursuit (an exact, optimal solution) and matching pursuit (a fast, greedy, sub-optimal algorithm). We conclude in Sec. 5 with a few final remarks.

2. PROBLEM STATEMENT AND APPROACH

In this paper, we address the problem of acquiring subsampled ultrasound echo data in a simple environment as a means to minimize storage requirements without significant loss in signal information. Reconstructions are performed offline with simulated subsampling to ensure the signal can be recovered. In the following subsection we present a brief background on CS. A more complete description can be found in a variety of sources in the literature, including, e.g. (Candès and Wakin, 2008) and (Nam et al., 2013). In Sec. 2.2 we describe how to fit the ultrasound signal problem into this framework.

2.1 Background on Compressive Sensing

Compressive sensing is a joint measurement and signal processing technique which can produce good (or even exact) reconstructions of signals from significantly fewer measurements than the Nyquist-Shannon sampling theorem requires. At the heart of CS is the assumption of compressibility (or true sparsity) of the signal of interest, that is, when described in an appropriate basis, most of the coefficients are negligible (compressibility) or exactly zero (sparsity).

Some of the most significant results in the CS literature rely on the matrix describing the measurements satisfying a requirement known as the Restricted Isomery Property (RIP) (Candès and Wakin, 2008). However, certifying whether a particular matrix has this property is an NP-hard problem (Bandeira et al., 2013). An alternative approach to determine the theoretical performance of CS has been developed using the concept of Mutual Coherence. Given a basis Ψ of $\mathbb{R}^n$, with normalized columns, its mutual coherence $\mu(\Psi)$ is defined as (Donoho et al., 2006)

$$\mu(\Psi) = \sqrt{n} \max_{1 \le i \ne j \le n} |\Psi_i^T \Psi_j| \tag{1}$$

where Ψ_k is the k-th column of Ψ. If a signal $\tilde{y}$ is s-sparse in the Ψ basis (i.e. it has only s non zero coefficients in the representation $\tilde{y} = \Psi x$), then, assuming

$$m \ge C\mu^2(\Psi) s \log n \tag{2}$$

where C is a constant, then the linear program

$$\min \|x\|_1 \quad \text{subject to} \quad y = \Phi\Psi x \tag{3}$$

where Φ is a matrix that extracts m random lines of Ψ and $y = \Phi\tilde{y}$, then the probability of $\tilde{y} \ne \Psi x^*$ decreases exponentially with the value of m (Bruckstein et al., 2009). The optimization problem in (3) is known as *Basis Pursuit* (BP). In other words, if you have a signal $\tilde{y} = \Psi x$, where x is s-sparse and you take m random measurements of $\tilde{y}$ where m satisfies (2) then, with an overwhelming high probability, the solution obtained by BP will be a perfect reconstruction of the original signal x. In the case where perfect reconstruction is not possible due, for example, to noise in the measurements, then the mutual coherence can be used to bound the reconstruction error.

Notice that the expression (2) depends on the value of a constant C and the sparseness assumption on x. In practice, these are both generally unknown. In fact, in real world scenarios, signals are not truly sparse and they are corrupted by noise. Thus, perfect reconstruction is not usually achievable. However, most real-world signals are approximately sparse and CS is still a powerful tool for reconstructing signal from highly under sampled data with good quality.

Additionally, it has been shown that the k–sparse approximation $\hat{x}$, composed with only the k largest coefficients of x being non-zero, minimizes the error

$$\|x - \hat{x}\|_p$$

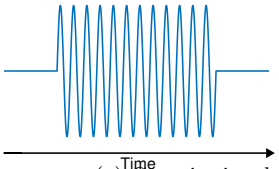
for any ℓ_p norm with an error that is bounded by an expression involving the mutual coherence (Bruckstein et al., 2009). Some CS algorithms take advantage of this and search only for the k largest coefficients of x instead of an exact optimal solution. These so called *greedy* methods may provide lower-quality reconstructions but typically require far less computational effort and can be solved much more quickly than BP. In the remainder of this paper we will use a common greedy algorithm, known as *Matching Pursuit* (MP) (Tropp and Gilbert, 2007) in addition to BP.

It is also worth mentioning that although it was established earlier that Ψ is a basis of $\mathbb{R}^n$, this notion can be extrapolated for cases in which Ψ is overcomplete or redundant, *i.e.*, it is a full rank matrix with more columns than rows. This follows from the intuitive notion that adding extra columns to Ψ cannot increase the optimal cost in (3) since the original columns remain. A more in depth discussion of the usage of overcomplete dictionaries

can be found in, *e.g.*, (Candès et al., 2011) and (Donoho et al., 2006).

2.2 Sonar Signals in Sparse Environments

In order to use sparse sampling and reconstruction in the ultrasound setting, we need to frame ultrasonic sensing in terms of a measurement model that fits the CS framework. Most ultrasonic sensors have a transducer that acts as both transmitter and receiver. During transmission, the transducer is vibrated by a pulse modulated wave, resulting in a signal such as the one shown in Fig. 1a propagating through the air. When an echo is received, the membrane fluctuates due to the air motion, yielding a corresponding voltage fluctuation that can be read out.



(a) Generic signal (b) Matrix Ψ

Fig. 1. (a) Generic shape of the transmitted sonar signal. (b) Graphical representation of the matrix ψ. Non-zero elements are shown in grayscale with darker representing larger values. The j-th column of this matrix is the signal s_j, i.e. samples of the signal in (a) delayed by j samples.

Consider a simple idealized case in which a single point-like object is located in the environment at a distance $d/2$ from the emitter. Sound waves travel through the environment, are reflected by the object, and propagate back to the emitter. If the sensor emits a pulse, $s(t)$ at time $t = 0$, then the output at the receiver is

$$y(t) = \frac{\kappa\Gamma}{d^2} s\left(t - \frac{d}{v_s}\right) \tag{4}$$

where v_s is the speed of sound, κ is a constant representing emitter and receiver gain coefficients which do not depend on the object or its relative direction from the transmitter, and Γ represents other sources of attenuation which may depend on relative orientation, such as reflectivity of the target and the receiver characteristics of the sensor.

Now suppose there are multiple point-like objects, all positioned such that the transmitted signal can reach each object (i.e. no object shadows another one). In this case, under assumptions of a linear response of the sensor, the received signal in (4) becomes

$$y(t) = \sum_{j=1}^{m} \frac{\kappa\Gamma_j}{d_j^2} s\left(t - \frac{d_j}{v_s}\right) \tag{5}$$

where m is the total number of individual signals received at the sensor. Note that this simple model can also describe more complicated scenarios such as in the case of multi-path (*e.g.* reflections of the sound wave off the floor and back (Sanchez and Andersson, 2018)) and objects with physical extent where there is a continuum of distances to points on a given object. Thus, in this framework, $s(t-\alpha)$, $\alpha \in \mathbb{R}^+$, can be viewed a basis of all possible signals $y(t)$ one might expect to receive. If the number of objects (or propagation paths) m is small, then $y(t)$ is a linear combination of a sparse subset of the basis $s(t-\alpha)$, $\alpha \in \mathbb{R}^+$, motivating the use of subsampling combined with CS to reconstruct the full signal $x(t)$ from that under sampled data.

In practice, the received signal is discretely sampled at a given rate T_s. The sampled version of $y(t)$ is

$$\tilde{y}[k] = \sum_{j=1}^{m} \frac{\kappa\Gamma_j}{d_j^2} s_j[k] \tag{6}$$

where $s_j[k] = s\left(kT_s - \frac{d_j}{v_s}\right)$. Now introduce the discrete signal x whose elements are given by

$$x[j] = \begin{cases} \frac{\kappa\Gamma_j}{\delta_j^2} & \text{if exists prop. path with dist. } \delta_j, \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

where $\delta_j = jT_sv_s$. Now, recall that s_j is a vector giving the time response at the receiver due to a pulse emitted at $t = 0$ and a target at distance δ_j. With this we can express the received (discrete) signal over N time samples as

$$\tilde{y} = \underbrace{[s_1 \; s_2 \; \cdots \; s_N]}_{\Psi} x, \tag{8}$$

Note that the discrete representation introduces a (small) error as it assumes objects only exist at the discrete distances δ_j, that is to integer multiples of v_sT_s.

The matrix Ψ is a basis for the discrete signals generated by the ultrasound pulse $s(t)$. It is shown graphically in Fig. 1b where each column j represents the response at the receiver due to an object at distance δ_j. Note that the non-zero entries in the upper-right corner of the matrix are included to ensure Ψ is a circulant matrix. While these entries are nonphysical, they ensure that Ψ is a basis of $\mathbb{R}^N$ and that it has a low mutual coherence, otherwise the mutual coherence would be $\sqrt{n}$, i.e., its maximum value possible (Candès and Wakin, 2008).

If the number of objects (and the number of multi-path returns) in the environment is small, then x will be a sparse vector and, given the discussion in Subsec. 2.1, a good approximation of the signal can be made using a small number of randomly acquired samples of $\tilde{y}[k]$ using CS.

Note that the basis represented by Ψ could be generated using more complicated responses than that of a point source if such responses are more representative of the environment. The matrix could also be extended to include multiple types of responses, leading to an *overcomplete dictionary.*

3. EXPERIMENTAL SETUP

3.1 Description of the Equipment and Experimental Design

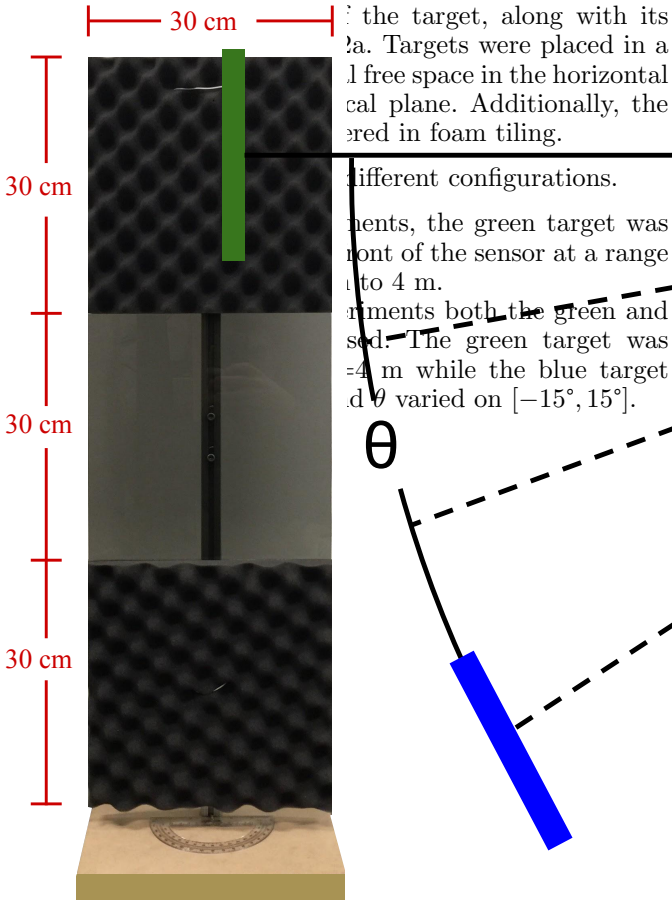
Experiments were performed using an X1 Ranging Module Pro Developers Kit (SensComp). This unit emits a 50 kHz ultrasonic pulse, either at a fixed rate or upon activation of a trigger signal. While not directly provided on the board outputs, the raw return signal of the sonar can be acquired from the kit. It is important to note that this signal is subject to a time-dependent gain (independent of the signal), presumably to enhance the signal magnitude at longer distances. As a result, the $1/d^2$ decay rate is not

observable in the signal. Since our approach relies on the *shape* of the echo signal rather than its absolute amplitude, this detail did not affect our results. If needed, the effect is easily removed in post-processing. Data was acquired using a National Instruments compact Reconfigurable Input Output system (cRIO 9076) equipped with a high speed voltage input module (NI 9233), programmed in Labview 2017 to sample at 250 kHz.

The ultrasonic transducer was centered on a laboratory work bench which was covered in acoustic foam panels to prevent the echo signal from bouncing off the surface of the benchtop. Two targets were built, each consisting of a 30 cm by 30 cm acrylic panel as the reflective component, mounted on a support bar to lift the center of the target, and a base. Acoustic foam panels were mounted above and below the panel to prevent the support bar from interfering with the return signal data. The center of the acrylic panels were matched to the height of the sonar sensor at 45 *cm* the target, along with its a. Targets were placed in a l free space in the horizontal cal plane. Additionally, the ered in foam tiling.

different configurations.

nents, the green target was ont of the sensor at a range to 4 m.
riments both the green and sed. The green target was =4 m while the blue target d θ varied on $[-15°, 15°]$.



(a) Flat target (b) Experimental setup

Fig. 2. Experimental setup. (a) Example of a target. (b) Position of the targets (green and blue) relative to the sonar sensor (red).

3.2 Construction of Sparsity Bases

For this work, we obtained several different Ψ matrices so as to compare their performance. These matrices were designed using combinations of data obtained from two sets of return data. The first set, shown in Fig.3(a), was gathered by facing one sensor (source) at another (receiver) and averaging the received signal over 20 pulses.

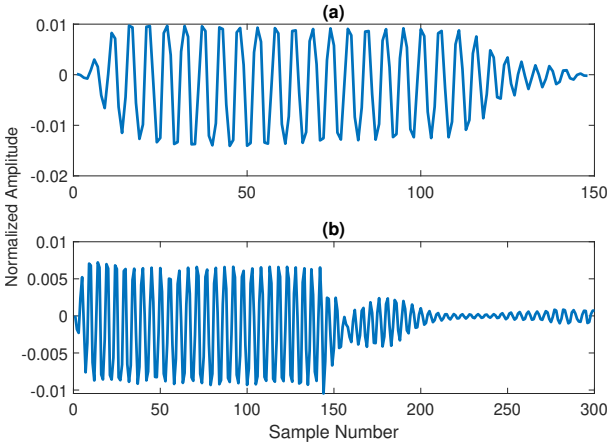


Fig. 3. Experimentally-obtained pulses using either (a) two sonar sensors pointed at each other or (b) a single sensor and a single target. Note that the time scales in the two plots are different and chosen to best show the details of the signals.

The second set, shown in Fig. 3(b), was obtained by facing the sensor at one of the flat targets at a range of 10 different distances (from 20 cm to 3.5m), averaged 20 return pulses at each distance, and then averaging those. Notice that the signal shows the expected sinusoidal shape but includes both a ramp up and final decay, likely due to dynamics of the acoustic actuator. Additionally, note that the time scale in the two graphs are different; the rectangular target return is approximately twice as long as the point source and exhibits a more complex shape.

The signal represented in Fig. 3(a) and its delayed versions were used to build a matrix Ψ referred to as the "2-sensor" basis. Similarly, the signal in Fig. 3(b) was used to build the "flat-avg" basis. We then constructed two overcomplete bases. The first, termed the "combo-2x" basis, was a composition of the 2-sensor and flat-avg bases (making Ψ $n \times 2n$). The second was a composition of the 2-sensor basis with each of the 10 (averaged) returns at each distance (which were used to make the flat-avg basis), yielding a $n \times 11n$ matrix for Ψ. This is referred to as the "combo-11x" basis. Because this matrix is very large, only MP was used to generated reconstructions from it. The computed mutual coherence values were $0.9866\sqrt{9500}$ for the 2-sensor basis, and $0.9910\sqrt{9500}$ for both the flat-avg and "combo-2x" bases. (Due to its size, we did not compute the mutual coherence of the combo-11x.) These values are all somewhat large and thus, from the underlying theory, it is not expected that they will lead to highly accurate reconstructions. Nevertheless, the experimental results below show reasonable performance in terms of reconstruction and good performance in terms of the location of those coefficients in the signal.

4. EXPERIMENTS

In each case, reconstructions were performed over a time window that started at approximately 0.001 s (the end of the emitted pulse) and ended at approximately 0.04 s (long enough to allow for multipath echoes from the far wall). With this choice of times, the length of $x(t)$ during reconstruction is $n = 9500$ samples. The length

Table 1. Reconstruction error comparisons between the combinations of the two algorithms with the 3-4 sparsity bases used for each of the 2 experiments.

Error for Experiment 1: one target, varying distance.

		20 cm	40 cm	60 cm	80 cm	100 cm	150 cm	200 cm	250 cm	300 cm	350 cm	400 cm	Row avg.
BP	2-sensor	0.496	0.452	0.618	0.455	0.491	0.659	0.556	0.536	0.737	0.443	0.281	0.520
	flat-avg	0.525	0.450	0.619	0.458	0.460	0.433	0.362	0.275	0.370	0.297	0.232	0.407
	combo-2x	0.461	0.399	0.595	0.419	0.423	0.465	0.370	0.269	0.367	0.278	0.196	0.386
MP	2-sensor	0.548	0.509	0.636	0.521	0.524	0.732	0.642	0.547	0.729	0.521	0.397	0.573
	flat-avg	0.578	0.502	0.750	0.503	0.554	0.501	0.447	0.345	0.453	0.365	0.329	0.484
	combo-2x	0.584	0.470	0.729	0.490	0.507	0.492	0.451	0.329	0.441	0.379	0.320	0.472
	combo-11x	0.528	0.476	0.516	0.412	0.433	0.341	0.315	0.207	0.377	0.397	0.351	0.396

Error for Experiment 2: two targets, near one at 3 m and varying angle, further one at 4 m and fixed 0° angle.

		−15°	−10°	−5°	-2.5°	0°	2.5°	5°	10°	15°	Row avg.
BP	2-sensor	0.432	0.284	0.371	0.562	0.491	0.465	0.424	0.335	0.421	0.421
	flat-avg	0.466	0.289	0.269	0.354	0.396	0.292	0.276	0.201	0.411	0.328
	combo-2x	0.417	0.246	0.253	0.350	0.338	0.271	0.275	0.187	0.378	0.302
MP	2-sensor	0.589	0.397	0.469	0.579	0.626	0.573	0.488	0.418	0.613	0.528
	flat-avg	0.540	0.349	0.297	0.401	0.493	0.329	0.331	0.284	0.511	0.393
	combo-2x	0.514	0.341	0.314	0.401	0.472	0.335	0.343	0.284	0.548	0.394
	combo-11x	0.535	0.345	0.389	0.451	0.275	0.349	0.398	0.344	0.556	0.405

of $y(t)$ is chosen to be approximately 20% of that at $m = 1900$ samples. The Φ matrix is constructed as a random permutation of the rows of the $n \times n$ identity matrix, and the Ψ matrix is constructed as described above in Sec. 3.2. The same m samples are used across all reconstructions so comparisons using different methods or bases are meaningful. When MP is performed, estimates are restricted to 50 coefficients. When BP is performed, the algorithm terminates after $10m = 19000$ iterations of the iterative (gradient descent) solver or earlier if an optimal solution is found.

Fig. 4 shows a typical reconstruction result for one of the returns from the second set of experiments. Here, one target is located 4 m away and directly in front of the sensor while the second is 3 m away and offset 2.5° from the sensor bearing. Further away, at approximately 6 m, there was a wall. Fig. 4a shows the full signal return along with the MP reconstruction using the flat-avg basis. Fig. 4b shows a closeup view of the true signal and reconstruction corresponding to the near object return signal, and Fig. 4d shows a closeup view of the same signal corresponding to the far wall return. Fig. 4c shows a stem plot of the coefficients found during the reconstruction, with the top x-axis scaled to reflect the one-way distance to a theoretical object that each coefficient represents. Notice that the reconstructions do well at capturing the timing and overall character of the signal but not the details. This is expected from the large mutual coherence of the bases. At the same time, it is important to note that the reconstruction from the echo of the far wall is qualitatively good, despite the fact that the wall looks neither like the point source represented in the 2-sensor nor the flat target represented in the flat-avg basis.

Table 1 shows the reconstruction error, scaled by the true signal size for readability, averaged over 20 trials for each combination of basis, algorithm and experimental condition. The error is calculated by

$$e = \frac{\|\hat{x} - x\|_2}{\|x\|_2}. \tag{9}$$

These tables show that, on average, BP performed better than MP, and the flat-avg basis performed better than the 2-sensor basis. In Experiment 1, the overcomplete bases performed the best, and reconstruction error with MP and the combo-11x basis was on par with the best BP reconstructions. In Experiment 2, the overcomplete basis combo-2x did help when combined with BP but the additional flexibility of the combo-11x basis was not beneficial in the MP reconstructions. It is no surprise that BP performs better than MP when the same basis is used. On the other hand, by providing MP with a basis that has more columns, it is possible to achieve good reconstructions in a fraction of the time. Indeed, average computation time for MP (including overcomplete bases) was 2 orders of magnitude lower than for BP.

5. CONCLUSIONS

In this paper, we showed that a simple random subsampling scheme of ultrasonic echoes in simple environments can be used with CS to approximately reconstruct the full unknown signal. Using experimentally acquired data and experimentally designed reconstruction bases, the results show efficacy while also indicating the importance of choosing a good representation basis for the signal. Signal reconstruction from echoes off the far wall further supports the method and the idea that an echo acquired in an unknown environment can be approximated by a collection of point-like objects and/or flat targets.

REFERENCES

Bandeira, A.S., Dobriban, E., Mixon, D.G., and Sawin, W.F. (2013). Certifying the restricted isometry property is hard. *IEEE Transactions on Information Theory*, 59(6), 3448–3450.

Bruckstein, A.M., Donoho, D.L., and Elad, M. (2009). From Sparse Solutions of Systems of Equations to Sparse Modeling of Signals and Images. *SIAM Review*, 51(1), 34–81.

Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I., and Leonard, J.J. (2016).

Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age. *IEEE Transactions on Robotics*, 32(6), 1309–1332.

Candès, E.J. and Wakin, M.B. (2008). An Introduction To Compressive Sampling. *IEEE Signal Processing Magazine*, 25(2), 21–30.

Candès, E.J., Eldar, Y.C., Needell, D., and Randall, P. (2011). Compressed sensing with coherent and redundant dictionaries. *Applied and Computational Harmonic Analysis*, 31(1), 59 – 73.

Colorado, J., Rossi, C., Zhang, C., and Barrientos, A. (2015). Towards efficient flight: insights on proper morphing-wing modulation in a bat-like robot. *Advanced Robotics*, 29(24), 1599–1610.

Dokmanic, I., Parhizkar, R., Walther, A., Lu, Y.M., and Vetterli, M. (2013). Acoustic echoes reveal room shape. *Proceedings of the National Academy of Sciences*, 110(30), 12186–12191.

Donoho, D.L., Elad, M., and Temlyakov, V.N. (2006). Stable recovery of sparse overcomplete representations in the presence of noise. *IEEE Transactions on information theory*, 52(1), 6–18.

Jung, S., Kim, J., and Kim, S. (2009). Simultaneous localization and mapping of a wheel-based autonomous vehicle with ultrasonic sensors. *Artificial Life and Robotics*, 14(2), 186.

Krekovic, M., Dokmanic, I., and Vetterli, M. (2016). EchoSLAM: Simultaneous localization and mapping with acoustic echoes. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 11–15.

Nam, S., Davies, M.E., Elad, M., and Gribonval, R. (2013). The cosparse analysis model and algorithms. *Applied and Computational Harmonic Analysis*, 34(1), 30–56.

Quistgaard, J.U. (1997). Signal acquisition and processing in medical diagnostic ultrasound. *IEEE Signal Processing Magazine*, 14(1), 67–74.

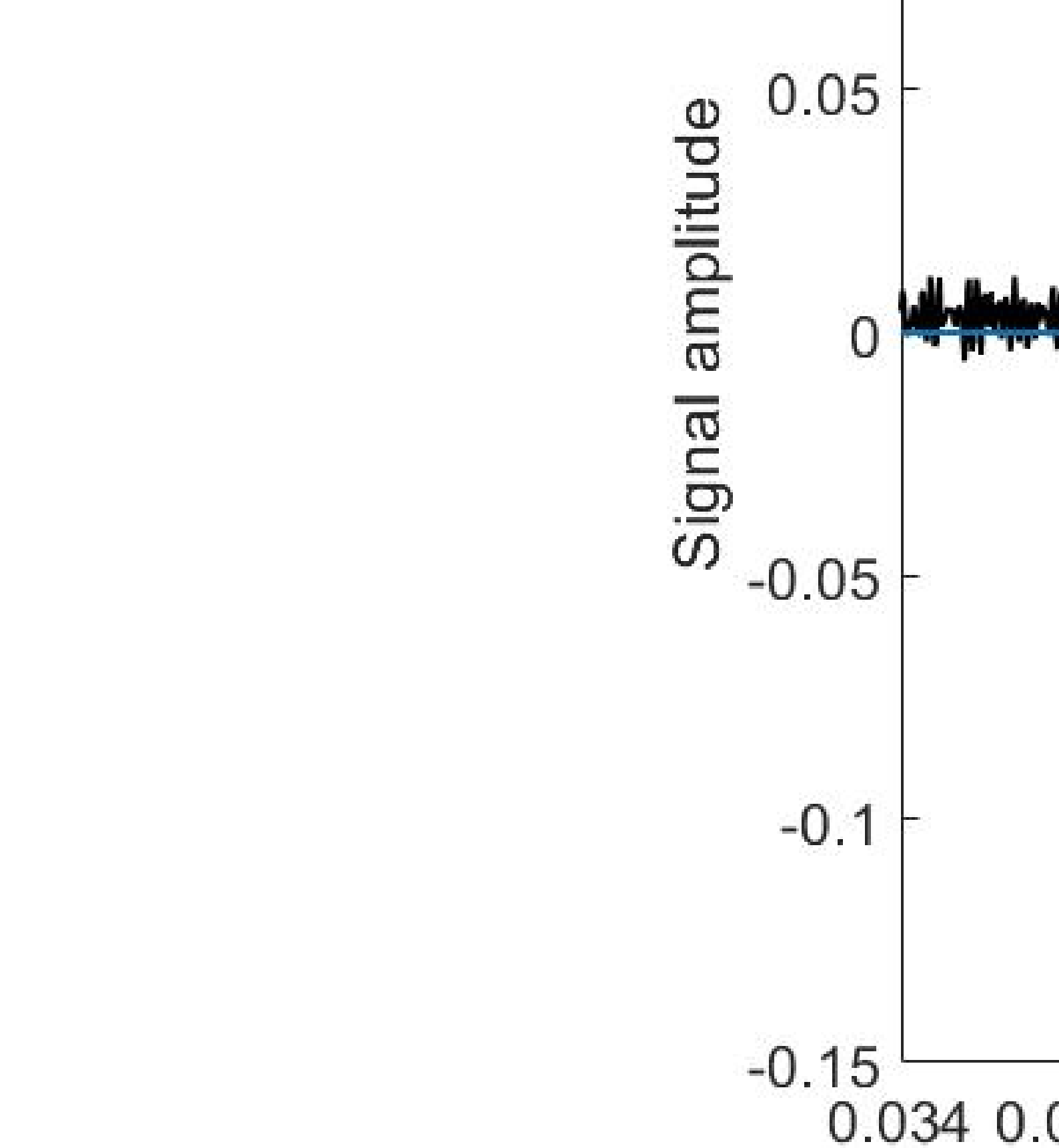
Sanchez, S.R. and Andersson, S.B. (2018). Using Compressive Sensing With In-Air Ultrasonic Measurements for Robotic Mapping. *ASME 2018 Dynamic Systems and Control Conference*.

Scaramuzza, D., Achtelik, M.C., Doitsidis, L., Friedrich, F., Kosmatopoulos, E., Martinelli, A., Achtelik, M.W., Chli, M., Chatzichristofis, S., Kneip, L., Gurdan, D., Heng, L., Lee, G.H., Lynen, S., Pollefeys, M., Renzaglia, A., Siegwart, R., Stumpf, J.C., Tanskanen, P., Troiani, C., Weiss, S., and Meier, L. (2014). Vision-Controlled Micro Flying Robots: From System Design to Autonomous Navigation and Mapping in GPS-Denied Environments. *IEEE Robotics and Automation Magazine*, 21(3), 26–40.

Tropp, J.A. and Gilbert, A.C. (2007). Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit. *IEEE Transactions on Information Theory*, 53(12), 4655–4666.

Wood, R., Nagpal, R., and Wei, G.Y. (2013). Flight of the Robobees. *Scientific American*, 308(3), 60–65.

Zou, Y., Zhang, W., Zhou, S., Ke, X., Cui, F., and Liu, W. (2018). Monolithic fabrication of an insect-scale self-lifting flapping-wing robot. *Micro & Nano Letters*, 13(2), 267–269.



(c) Stem plot of reconstruction coefficients

(d) Zoom view of wall echo

Fig. 4. Example signal reconstruction using MP with the $n \times n$ flat-avg basis. One target is at $(d, \theta) = (4\,\mathrm{m}, 0^\circ)$, another at $(3\,\mathrm{m}, 2.5^\circ)$. (a) Ground truth, reconstruction and error signals. Error is calculated to be 0.305 (b) Zoom of the near target return and (d) zoom of the wall return. (c) Reconstruction coefficient stem plot; zero-coefficients are not shown for clarity. The top axis is included to correlate coefficients to distance from the sensor.