LOCALIZATION AND BEARING ESTIMATION VIA STRUCTURED SPARSITY MODELS

Marco F. Duarte

Department of Electrical and Computer Engineering, University of Massachusetts, Amherst, MA 01003

ABSTRACT

Recent work has leveraged sparse signal models for parameter estimation purposes in applications including localization and bearing estimation. A dictionary whose elements correspond to observations for a sampling of the parameter space is used for sparse approximation of the received signals; the resulting sparse coefficient vector's support identifies the parameter estimates. While increasing the parameter space sampling resolution provides better sparse approximations for arbitrary observations, the resulting high dictionary coherence hampers the performance of standard sparse approximation, preventing accurate parameter estimation. To alleviate this shortcoming, this paper proposes the use of structured sparse approximation that rules out the presence of pairs of coherent dictionary elements in the sparse approximation of the observed data. We show through simulations that our proposed algorithms offer significantly improved performance when compared with their standard sparsity-based counterparts. We also verify their robustness to noise and applicability to both full-rate and compressive sensing data acquisition.

***Index Terms*—** localization, bearing estimation, structured sparsity, coherence, compressive sensing

1. INTRODUCTION

Sparse signal models have received significant attention in the last decade. Their most popular applications include lossy signal compression [1], denoising [2], and compressive sensing [3, 4]. Sparse signal models have also been used recently to solve parameter estimation problems such as spectral estimation, localization, and bearing estimation [5–7]. In these cases, the parameter space (e.g., the sets of possible angles of arrival, locations, or signal frequencies) is sampled in a discrete fashion, and a dictionary is formed by collecting the observations corresponding to the sampled parameter values as dictionary elements. The parameter estimation problem then reduces to finding a representation of the observed signals as the linear combination of as few of the dictionary elements as possible, a problem solved by sparse approximation algorithms [8]. The chosen elements are then linked to the corresponding values of the samples from the parameter space, which are offered as the output parameter estimates. Since sparsity is introduced as a signal model, such parameter estimation approaches are immediately compatible with compressive sensing (CS) techniques that reduce the dimensionality of the data representation [6, 7, 9–11].

Such sparsity models for parameter estimation are accurate only if the parameter values being observed are contained within the set of sampled values that generates the dictionary under use [12–14]. Since the map between parameter values and observations is often smooth, increasing the resolution of parameter sampling will enlarge both the set of events that lead to exact parameter estimation as well as the set of events that lead to accurate parameter estimation using the sparsity model, keeping the parameter estimation error low [7].

Unfortunately, it is well known that the performance of such sparse approximation algorithms greatly suffers when the dictionary involved exhibits high coherence [8]. In parameter estimation applications, the dictionary coherence increases as the resolution of the parameter space sampling grows finer. This behavior can be attributed in many cases to ambiguity or resolution issues in the underlying parameter estimation problem, and may prevent accurate estimation via sparsity — even for the simplest problems.

In this paper, we expand the approach of [7] for successful recovery of frequency-sparse signals to address the shortcomings of high-resolution sparsity-based parameter estimation. The centerpiece of our approach is the application of a structured sparsity model [15] that prevents coherent dictionary elements from appearing simultaneously in a signal approximation. The modification to baseline approaches is simple and computationally efficient, and achieves a significant improvement in the performance of the sparsity-based parameter estimation algorithm as the resolution of the parameter space sampling increases. As examples, we apply our new approach to the applications of localization and bearing estimation.

2. BACKGROUND

2.1. Sparsity-based parameter estimation

We assume that a map $\mathcal{M} : \Theta \to \mathbb{C}^L$ provides a link between a parameter value $\theta \in \Theta$ and an L-sample observation $x = \mathcal{M}(\theta)$. The goal of parameter estimation is to invert the mapping in order to obtain an estimate of the parameter $\widehat{\theta}$ from observations $y = x + n$, where n represents the observation noise. Consider the case where the signal x corresponds to a linear combination of observations corresponding to K distinct unknown parameter values. More specifically, $x = \sum_{k=1}^{K} a_k \mathcal{M}(\theta_k)$, and the goal of parameter estimation is to obtain the parameter values $\theta_1, \ldots, \theta_k$ from y. For this purpose, we can consider a dictionary $\Psi_\Omega = [\mathcal{M}(\omega_1) \ \ldots \ \mathcal{M}(\omega_N)]$ containing observations for a sampling of parameter values $\Omega = \{\omega_1, \ldots, \omega_N\} \subseteq \Theta$. If the sampling is rich enough that $\{\theta_1, \ldots, \theta_K\} \subseteq \Omega$, then we can express the observations as the matrix-vector product $x = \Psi_\Omega c$, where $c \in \mathbb{C}^L$ is a K-sparse coefficient vector (i.e., a vector containing only K nonzero entries).

The vector c can be accurately estimated from the observations y when the noise power $\|n\|$ is small enough and the dictionary is sufficiently *incoherent*. The worst-case coherence $\mu(\Psi_\Omega)$ corresponds to the largest absolute value of the inner product between two distinct dictionary elements:

$$\mu(\Psi_\Omega) = \max_{p \neq q} \frac{|\langle \mathcal{M}(\omega_p), \mathcal{M}(\omega_q)\rangle|}{\|\mathcal{M}(\omega_p)\|\|\mathcal{M}(\omega_q)\|}. \quad (1)$$

The majority of existing sparse approximation algorithms require $\mu(\Psi_\Omega) \leq 1/(2K-1)$ to be successful [8]; this bound limits the maximum value of the sparsity K for which the coefficient vector c can be accurately estimated from the noisy observations y. We focus in the sequel on the iterative hard thresholding (IHT) algorithm [16], which sets an initial estimate $\widehat{c}_0 = 0$ and then refines the estimate iteratively as

$$\widehat{c}_t = \mathcal{T}_K(\widehat{c}_{t-1} + \Psi_\Omega^H(y - \Psi_\Omega \widehat{c}_{t-1})) \quad (2)$$

until a convergence criterion is met. Here, $\mathcal{T}_K(c)$ describes a K-thresholding operation that preserves the K entries of the vector c with largest absolute values, while setting all other entries to zero. Such a thresholding operation provides the best K-term sparse approximation to the input vector c [1, 2]. Other algorithms such as CoSaMP, FPC, and GPSR obtain improved performance with slightly higher computational complexity for signal recovery while still relying on basic matrix product and thresholding operations [17].

For the case where the dictionary is incoherent and K is small, it is also possible to perform sparsity-based parameter estimation from compressive sensing (CS) measurements via a random projection matrix $\Phi \in \mathbb{C}^{M \times L}$ [3, 4], where our observations are now given by $y = \Phi x + n$. We simply need for the number of measurements to obey $M = \mathcal{O}(K \log L)$ and for the recovery algorithm (2) to replace the matrices Ψ_Ω and Ψ_Ω^H by $A = \Phi\Psi_\Omega$ and A^H, respectively [18, 19].

2.2. Issues with sparsity-based parameter estimation

Unfortunately, the performance of sparsity-based parameter estimation suffers significantly in the case where one of the observed parameter values $\theta_k \notin \Omega$ [7, 12–14]. It may still be possible that the closest parameter value in Ω to θ_k provides a dictionary vector $\mathcal{M}(\omega_i)$ that is a good approximation to the observation $\mathcal{M}(\theta_k)$ (i.e., $\|\mathcal{M}(\omega_i) - \mathcal{M}(\theta_k)\|$ is small). In such a case, the sparse approximation algorithm can be modified to handle such inaccuracies. Thus, we may expect sufficiently accurate parameter estimation as long as the parameter space sampling Ω is sufficiently dense.

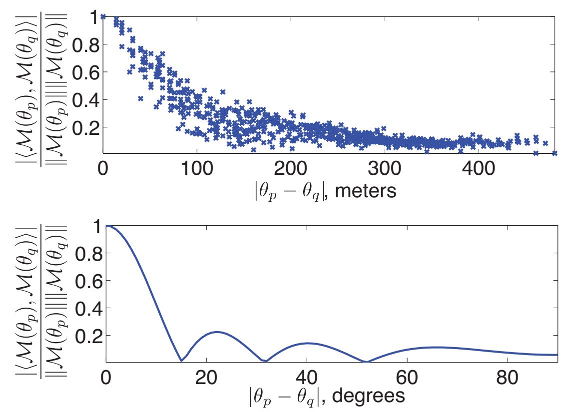

Fig. 1. *Coherence* $|\langle \mathcal{M}(\theta_p), \mathcal{M}(\theta_q)\rangle| / \|\mathcal{M}(\theta_p)\|\|\mathcal{M}(\theta_q)\rangle\|$ *as a function of the parameter distance* $|\theta_p - \theta_q|$ *for the localization (top) and bearing estimation (bottom) examples.*

It turns out that the coherence of the dictionary $\mu(\Psi_\Omega)$ often increases as the resolution of the parameter space sampling grows finer. Such increase prevents sparsity-based parameter estimation approaches from achieving good performance due to the coherence-induced shortcomings of sparse approximation. As examples, Figure 1 shows the absolute value of the inner product $|\langle \mathcal{M}(\theta_1), \mathcal{M}(\theta_2)\rangle|$ as a function of the difference $|\theta_1 - \theta_2|$ for the localization and bearing estimation examples described in Section 4. The figures show that as the sampling resolution $\min_{i,j} |\theta_i - \theta_j|$ decreases, the value of the inner product (and by extension the worst-case coherence) approaches the maximum value of 1, severely limiting the performance of sparse approximation algorithms for sparse recovery purposes.

3. STRUCTURED SPARSE APPROXIMATION FOR PARAMETER ESTIMATION

In recent work, we have addressed the coherence shortcoming in parametric dictionaries for CS recovery of spectrally sparse signals [7], allowing a significant reduction in the number of measurements necessary for accurate recovery. We propose to apply the same framework to parameter estimation problems.

The goal of our modified approach is to bypass the shortcoming introduced through dictionary coherence by avoiding signal representations that employ coherent pairs of dictionary elements. More specifically, we exchange the thresholding operation in (2) with a structured sparse approximation $\widehat{c} = \mathcal{S}_K(c, \mu_0)$ that prevents coherent pairs of dictionary elements (i.e., $\frac{|\langle \mathcal{M}(\omega_p), \mathcal{M}(\omega_q)\rangle|}{\|\mathcal{M}(\omega_p)\|\|\mathcal{M}(\omega_q)\|} \geq \mu_0$) from appearing in the approximation $\widehat{x} = \Psi_\Omega \widehat{c}$. This optimal structured sparse approximation can be obtained via linear programming [7, 20]; for computational simplicity, we focus on a computationally efficient heuristic presented as Algorithm 1. The combination of Algorithm 1 and (2) provides us with a structured sparsity-based parameter estimation algorithm that prevents the shortcomings introduced by coherence. We name the resulting approaches localization via structured sparsity (LoStS) and

Algorithm 1 Structured sparse approximation $\widehat{c} = \mathcal{S}_K(c, \mu_0)$

Inputs: Coefficient vector $c \in \mathbb{C}^L$, target coherence μ_0
Outputs: Approximation $\widehat{c} \in \mathbb{C}^L$ to coefficient vector
Initialize $\widehat{c} = 0$, $i = 1$.
while $i < K$ and $c \neq 0$ **do**
  $l^* = \arg\max_{1 \leq l \leq L} |c(l)|$
  $\widehat{c}(l^*) = c(l^*)$
  $\Lambda = \{p : \frac{|\langle \mathcal{M}(\omega_p), \mathcal{M}(\omega_{l^*}) \rangle|}{\|\mathcal{M}(\omega_p)\| \|\mathcal{M}(\omega_{l^*})\|} \geq \mu_0\}$
  $c|_\Lambda = 0$
  $i = i + 1$
end while
return $\widehat{c}$

bearing estimation via structured sparsity (BEStS). Similar extensions can also be formulated for more sophisticated algorithms but fall outside the scope of this paper.

4. SIMULATIONS

Our simulations test the performance of sparsity-based parameter estimation in localization and bearing estimation leveraging the models of standard sparsity (IHT) and structured sparsity (LoStS and BEStS). For both cases, we set the number of sources to $K = 5$, which is equal to the number of parameters to be estimated. We do not present results for the approach of [7, 20] due to its computational complexity.

In the bearing estimation simulations, a linear array of 10 microphone sensors at a spacing of 25 cm records a 31.25 ms observation at a sampling rate of 256 KHz, resulting in an observation of length 8192 samples for each sensor. Each source, assumed to be far enough from the array to be observed as a point target, is located at a bearing angle randomly selected to machine precision, with sufficient spacing among sources so that the coherence limit μ_0 is not reached. Each source transmits a beacon at a frequency of 500 kHz; the linear array observations are concatenated in a single vector and a dictionary is built by collecting observations of a single target for an N-point sampling of the bearing range $[0^o, 180^o]$.

In the localization simulations, a set of 20 microphone sensors randomly deployed in a field of 340×340 m records a 2.5 s observation at a sampling rate of 200 Hz, resulting in an observation of length 512 samples for each sensor. The source locations are randomly selected within the field to machine precision, with sufficient spacing among sources so that the coherence limit μ_0 is not reached. Each source transmits a known MSK-modulated binary sequence; the sensor observations are concatenated in a single vector and a dictionary is built by collecting observations of a single target for an N^2-point two-dimensional sampling of the spatial field $[0, 340]^2$.

We perform parameter estimation from (i) full-length observations compressed and decompressed via transform coding to an $K = 100$-coefficient representation, and (ii) $M = 100$ CS random projections using (i) standard sparsity (IHT) and (ii) structured sparsity (LoStS and BEStS) with the corresponding dictionaries for localization and bearing estimation. In both applications, we set the coherence target level $\mu_0 = 0.2$, which is also provided as input to the LoStS and BEStS algorithms.

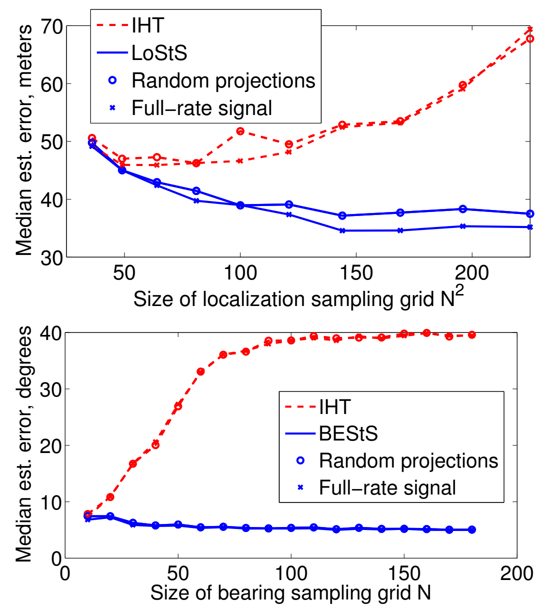

Fig. 2. *Performance of sparse and structured sparse approximation for localization (top) and bearing estimation (bottom) as a function of the size of the parameter space sampling grid.*

Figure 2 shows the performance of parameter estimation (in terms of median estimation error among 100 trials) as a function of the size of the parameter space sampling grid. While the performance of standard sparsity methods suffers greatly due to the increasing coherence that is induced by the finer parameter space sampling, the use of structured sparsity enables the desired improvement in performance as N increases. The performance behavior is similar for the full-length signal and CS measurement cases, which highlights the fact that the results are due to the parameter sampling rather than the particular signal representation chosen. In addition, Figure 3 shows the performance of parameter estimation as a function of the measurement-to-noise ratio for a fixed resolution of the parameter space sampling. The figure shows that while the performance degrades gracefully in all cases as the noise power increases, the performance of structured sparsity-based algorithms is significantly improved over their standard sparsity-based counterparts. Additionally, the full-length signal setup exhibits better robustness to noise than the CS setup (i.e., CS performs equally or better than full-rate sampling only for sufficiently high SNRs), a behavior that has received increasing attention in the CS community [21, 22] and translates to the sparsity-based parameter estimation setup.

5. CONCLUSIONS

The integration of sparse signal models in parameter estimation introduces a tradeoff between the fidelity of the sparse

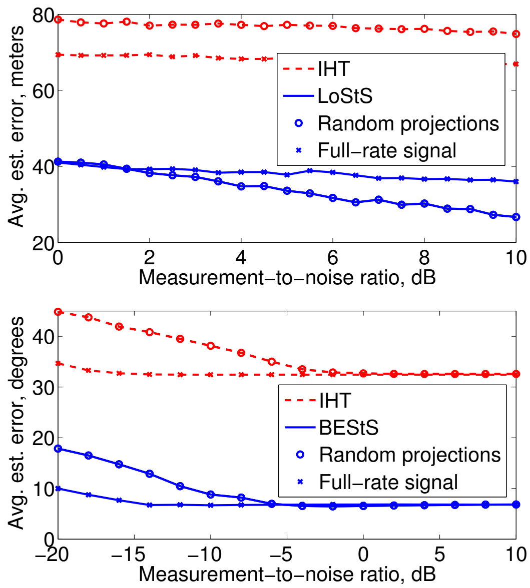

Fig. 3. *Performance of sparse and structured sparse approximation for localization ($N = 256^2$, top) and bearing estimation ($N = 256$, bottom) as a function of the measurement-to-noise ratio.*

approximation of the observations (which requires fine-rate parameter space sampling) and the decreasing performance of standard sparse approximation algorithms (which restrict the parameter space sampling rate that will still allow accurate estimation). We have proposed a structured sparse approximation algorithm that can achieve the promised parameter estimation performance improvements with higher-resolution parametric dictionaries by preventing pairs of coherent dictionary elements from appearing together in the output of the sparse approximation algorithm. While our algorithm provides significant improvements for observations that do not manifest coherence (i.e., with pairwise sufficiently-spaced parameters), it shares a resolution limitation with many legacy parameter estimation algorithms that is manifested in the coherence of the parametric dictionary. Additional refinement of the structure applied by the sparse approximation algorithm on the parameter estimation output to address such resolution ambiguities remains an interesting topic for future work.

6. REFERENCES

[1] R. A. DeVore, "Nonlinear approximation," *Acta Numerica*, vol. 7, pp. 51–150, 1998.

[2] D. L. Donoho, "Denoising by soft thresholding," *IEEE Trans. on Inform. Theory*, vol. 41, no. 3, 1995.

[3] D. L. Donoho, "Compressed sensing," *IEEE Trans. Info. Theory*, vol. 52, no. 4, pp. 1289–1306, Sept. 2006.

[4] E. J. Candès, "Compressive sampling," in *Int. Congress of Mathematicians*, Madrid, Spain, 2006, vol. 3, pp. 1433–1452.

[5] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm," *IEEE Trans. Signal Proc.*, vol. 45, no. 3, pp. 600–616, 1997.

[6] L. C. Potter, E. Ertin, J. T. Parker, and M. Cetin, "Sparsity and compressed sensing in radar imaging," *Proc. IEEE*, vol. 98, no. 6, pp. 1006–1020, June 2010.

[7] M. F. Duarte and R. G. Baraniuk, "Spectral compressive sensing," 2011, preprint.

[8] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Review*, vol. 51, no. 1, pp. 34–81, 2009.

[9] V. Cevher, M. F. Duarte, and R. G. Baraniuk, "Localization via spatial sparsity," in *European Signal Proc. Conf. (EUSIPCO)*, Lausanne, Switzerland, 2008.

[10] V. Cevher, A. C. Gurbuz, J. H. McClellan, and R. Chellappa, "Compressive wireless arrays for bearing estimation," in *IEEE Int. Conf. Acoustics, Speech, and Signal Proc. (ICASSP)*, Las Vegas, NV, Apr. 2008, pp. 2497–2500.

[11] T. J. Shaw and G. C. Valley, "Angle of arrival detection using compressive sensing," in *European Signal Proc. Conf. (EUSIPCO)*, Aalborg, Denmark, Aug. 2010, pp. 1424–1428.

[12] Y. Chi, L. Scharf, A. Pezeshki, and R. Calderbank, "The sensitivity to basis mismatch of compressed sensing in spectrum analysis and beamforming," *IEEE Trans. Signal Proc.*, vol. 59, no. 5, pp. 2182–2195, May 2011.

[13] M. A. Herman and T. Strohmer, "General deviants: An analysis of perturbations in compressed sensing," *IEEE J. Selected Topics in Signal Proc.*, vol. 4, no. 2, pp. 342–349, Apr. 2010.

[14] C. D. Austin, J. N. Ash, and R. L. Moses, "Parameter estimation using sparse reconstruction with dynamic dictionaries," in *IEEE Int. Conf. Acoustics, Speech, and Signal Proc. (ICASSP)*, Prague, Czech Republic, May 2011, pp. 2852–2855.

[15] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Trans. Info. Theory*, vol. 56, no. 4, pp. 1982–2001, Apr. 2010.

[16] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Appl. Comput. Harmon. Anal.*, vol. 27, no. 3, pp. 265–274, Nov. 2009.

[17] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," *Proc. IEEE*, vol. 98, no. 6, pp. 948 –958, June 2010.

[18] H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed sensing and redundant dictionaries," *IEEE Trans. Info. Theory*, vol. 54, no. 5, pp. 2210–2219, May 2008.

[19] E. J. Candès, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Appl. Comput. Harmon. Anal.*, Oct. 2010, To appear.

[20] C. Hegde, M. F. Duarte, and V. Cevher, "Compressive sensing recovery of spike trains using a structured sparsity model," in *Workshop on Signal Proc. with Adaptive Sparse Structured Representations (SPARS)*, Saint Malo, France, Apr. 2009.

[21] M. A. Davenport, J. N. Laska, J. R. Treichler, and R. G. Baraniuk, "The pros and cons of compressive sensing for wideband signal acquisition: Noise folding vs. dynamic range," 2011, Preprint.

[22] E. Arias-Castro and Y. C. Eldar, "Noise folding in compressed sensing," 2011, Preprint.