ADAPTIVE SPARSE OPTIMIZATION FOR COHERENT AND QUASI-STATIONARY PROBLEMS USING CONTEXT-BASED CONSTRAINTS

Ananya Sen Gupta, James Preisig

Department of Applied Ocean Physics and Engineering
Woods Hole Oceanographic Institution, Woods Hole, MA 02543
Email: {asengupta,jpreisig}@whoi.edu

ABSTRACT

Stationarity of the sparse coefficients as well as the sparseness of their support, along with incoherence assumptions related to restricted isometry, are fundamental to compressive sensing and sparse optimization. However, scientific study of many sparse processes encountered in nature as well as engineering applications necessitates solving ill-conditioned optimization metrics and tracking rapidly fluctuating coefficients where such incoherence and stationarity assumptions are difficult to satisfy. We propose to close the gap between mathematical optimality of sparse reconstruction and practical constraints of real-world applications by combining contextual information as external constraints to the traditional sparse optimization problem. Specifically, we explore the unobservable dimensions in a coherent reconstruction problem by navigating the non-convex topography of a modified mixed norm metric proposed in earlier work. Investigations based on simulated and experimental field data are provided.

Index Terms — Compressive sensing, sparse optimization, adaptive signal processing.

1. INTRODUCTION

Recent years have seen unprecedented synergy from the compressive sensing community regarding the development of sparse optimization techniques [1-10] to solve ill-posed estimation problems where the number of observations (i.e., measurements) falls far short of the number of coefficients to be estimated. Most sparse optimization techniques (ref. e.g. [4-8] among others) minimize a mixture of the L_1 norm of the coefficients to be estimated and the L_2 norm of the estimation error based on an observation space significantly smaller in dimensionality compared to the total support of the coefficient space. Mathematically, this optimization problem is stated as:

$$\hat{\mathbf{x}} = \min_{\{\mathbf{x}\in C^N, 0<\lambda\le 1\}} (1-\lambda)\|\mathbf{x}\|_1 + \lambda\|\mathbf{y}-\Phi\mathbf{x}\|_2^2 \qquad (1),$$

where $\hat{\mathbf{x}}$ denotes the estimate of S-sparse components $\mathbf{x}$ defined over an N-dimensional complex field, $\mathbf{y}$ denotes the M-dimensional observation vector, Φ denotes the $M \times N$ estimation matrix and $M << N$. The coefficient distribution of $\mathbf{x}$ is assumed to have S significant components within its support.

Despite successful adaptation of compressive sensing in several applications its viability over many practical applications is challenged by the need to satisfy the restricted isometry property (RIP) [3] criteria on the coherence between the columns of Φ to guarantee precise sparse reconstruction. Recent years have seen adaptive sparse solutions for wide-sense stationary processes [5-8], strategic designs for incoherent matrices [11], mathematical bounds on restricted isometry [12-15] as well as model-based approaches [16]. While these techniques are effective for applications that allow incoherent and wide-sense stationary design, scientific study of sparse processes encountered in nature often necessitate tracking rapidly fluctuating coefficients where such incoherence and stationarity assumptions are difficult to satisfy [17-20].

The gap between mathematical feasibility of sparse reconstruction and practical constraints of real-world applications provides the motivation behind this work. We extend recent work [10] to propose a context-driven navigation of the unobservable dimensions across a non-convex cost function. This enables us to swiftly reach feasible solutions that track coherent time-varying sparse problems. Although our approach is generally applicable to any application involving sparse coefficients, we target shallow water acoustics as a case study of a physical environment where challenges discussed above are a realistic concern.

2. SHALLOW WATER ACOUSTIC CHANNELS: A CASE STUDY IN COHERENT QUASI-STATIONARY SPARSE SENSING

A classic application where the signal processing challenges outlined above manifest in real life is shallow water acoustics, particularly in the context of scientific observation of the ocean as an acoustic channel [9,10,17-20]. The rapidly fluctuating nature of multipath due to surface wave motion, particularly in rough sea conditions, as well as occurrence of oceanic phenomena such as surface wave focusing [18], makes the time-variability of the channel impulse response difficult to observe over a reasonable observation window. Therefore, realistic characterizations of the shallow water channel [9,10,17-20] have been centered around the Delay-Doppler spread function, which captures the transient shallow water channel scattering function as a two-variable function over the delay taps (i.e, impulse response) and the Doppler spread around each delay tap as shown in Figure 1.

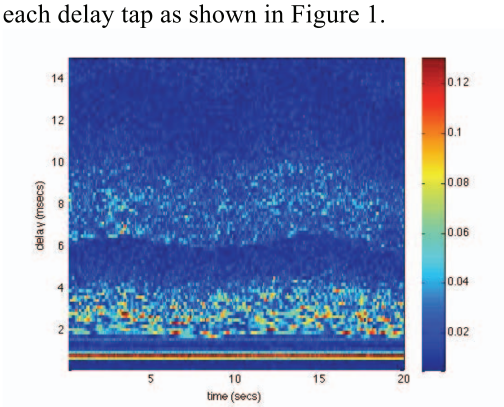

Figure 1(a): Delay spread (milliseconds) calculated using a least squared estimator along the y-axis for each instant in absolute time (x-axis) spanning 20 seconds. (Depth: 15 meters, Range between transmitter and receiver: 200 meters, rough sea conditions)

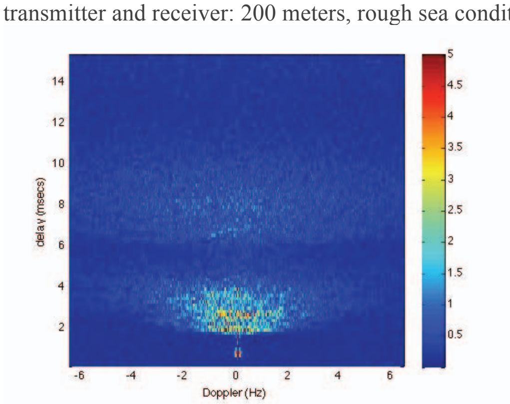

Figure 1(b): Delay-Doppler spread function calculated (long-term averaged estimate) from Figure 1(a) using direct Fourier transform along the first dimension (absolute time).

The direct arrival from the transmitter to the receiving hydrophone manifests as the straight line in Figure 1(a) and as the bright point at 0 Hz in Figure 1(b), while the bands of multipath arrivals, exhibit significant temporal fluctuations with occasional focusing (shown as bright dots) due to surface wave reflection. The sparse estimation problem of tracking the Delay-Doppler spread function coefficients, written as the KL-dimensional vector $\mathbf{u}$, may be stated in terms of the following model:

$$\mathbf{y} = \mathbf{C}\mathbf{u} + \mathbf{n}, \qquad (2)$$

where $\mathbf{y}$ denotes the M-dimensional observation vector ($M << KL$), $\mathbf{n} \sim N(\mathbf{0},\sigma^2\mathbf{I})$, and the i^{th} row of $\mathbf{C}$ is given as: $\mathbf{C}(i) = \mathbf{e}(i) \otimes \mathbf{x}(i)$, where $\otimes$ denotes the Kronecker product, $\mathbf{x}(i) = [x(i)x(i-1)\cdots x(i-K+1)]$ and $\mathbf{e}(i) = [e^{j2\pi\gamma_1 i\Delta t} e^{j2\pi\gamma_2 i\Delta t} \cdots e^{j2\pi\gamma_L i\Delta t}]$. Typically, to capture the non-stationary Doppler spread at sufficient resolution the observation window needs to be small enough to assume quasi-stationarity. This necessitates $\mathbf{C}$ to have many columns of closely spaced frequencies without enough rows to ensure sufficient inter-column incoherence to satisfy RIP constraints.

Therefore, the shallow water acoustic communications problem presents a realistic case study for developing and validating adaptive sparse optimization techniques where conditions for stationarity and incoherence are violated. In the sequel, we extend recent work in sparse optimization [10] to propose a novel method to track the unobservable dimensions that fall outside the span of columns of a coherent reconstruction matrix. We adopt the notation in Equation (1) and refer to the unobservable dimensions as the "invisible dimensions" for the rest of the paper.

3. SPARSE OPTIMIZATION THROUGH CONTEXTUAL NAVIGATION

Recently non-convex mixed norm solver (NCMNS) was introduced [9,10] to track rapidly fluctuating sparse coefficients $\mathbf{x}$ adaptively over time for a fixed λ with precise reconstruction guaranteed for well-conditioned, and hence, incoherent Φ, using the modified cost function:

$$\hat{\mathbf{z}} = \min_{\{\mathbf{z}\in C^N\}} (1-\lambda)\|\mathbf{z}\|_2^2 + \lambda\|\mathbf{y} - \Phi\mathbf{Z}^2\mathbf{1}\|_2^2, \quad 0 < \lambda \le 1, \qquad (3)$$

where $(\mathbf{z}_i)^2 = \mathbf{x}_i, \forall_{i=1}^N, \quad \mathbf{Z} \equiv diag(\mathbf{z})$, and $\mathbf{1}$ denotes the all-ones vector. The modified mixed norm cost in Equation (3) is a non-convex function over $\mathbf{z}$, denoted as $F(\lambda,\mathbf{F},\mathbf{y},\mathbf{z})$, and mathematically equivalent to the cost function in Equation (1). The gradient of $F(\lambda,\mathbf{F},\mathbf{y},\mathbf{z})$ with respect to $\mathbf{z}^*$, which is the direction of steepest descent, has 3^N possible stationary points since for each dimension we may either have 0, or the two possible squared roots of $\hat{\mathbf{x}}_i, \forall_{i=1}^N$ to get a zero gradient at that dimension. Of these 3^N possible stationary points, 2^S (two possible squared roots of $\hat{\mathbf{x}}_i, \forall_{i=1}^N$ for each of the S dimensions) correspond the unique minimum $\hat{\mathbf{x}}$ in the original convex optimization cost in Equation (1). Sufficiency criteria for guaranteed convergence to one of the 2^S stationary points are derived in [10] and a detailed discussion is outside the scope of this work.

In this work, we improve upon the tracking precision in Equation (3) as well as include the unobservable "invisible" dimensions by initializing the tracker at each stage that contains contextual information. It is noteworthy that operating in the $\mathbf{z}$-domain allows us to turn any of the N dimensions "on" or "off" by initializing the corresponding element of $\mathbf{z}$ to zero. This switching property is a feature of the non-convex cost function in Equation (3) as its gradient with respect to $\mathbf{z}^*$ has z_i as a root for its i^{th} element. This implies we may initialize **z** appropriately to incorporate external constraints and develop a context-constrained gradient descent approach to solve Equation (3) turning "invisible dimensions" on or off as the constraints dictate.

3.1 Constraint-based navigation

We propose the following constraint-based approach to navigate the non-convex manifold of the cost function in Equation (3). If the coherence criteria for Φ dictate an S^*-sparse support for **x**, when we expect the true sparseness to be $S>S^*$, we may track the significant components using the following rule:

Table 1: Tracking initialization for gradient-based optimization that includes context-based constraints on the sparse support
Step 1: For a given time step i, locate the dominant S^* dimensions of $\hat{\mathbf{x}}[i]$ (derived from $\hat{\mathbf{z}}$) in Equation (4)
Step 2: Turn off the other dimensions of $\hat{\mathbf{x}}[i]$. Add (S-S^*) "invisible dimensions" of $\hat{\mathbf{x}}[i]$ from the physical constraints of the application (e.g. in shallow water acoustics, proximity constraints due to multipath occurring in "bands" is a reasonable physical constraint to impose based on the channel model)
Step 3: Initialize the next step of the tracking problem based on 1 and 2 and repeat.

Though relatively free of model assumptions besides well-known and measurable physical constraints of an application, the above approach may not guarantee efficient selection of the "invisible dimensions". Therefore, we adopt the normalized prediction error in estimation problems as an external constraint, as it is well known [Section III, 21] that the normalized prediction error statistically follows the actual estimation error in channel estimation applications. Step 2 in Table 1 may now be modified to include the normalized prediction error as a context-based constraint and the tracking initialization is now given in Table 2.

Table 2: Tracking initialization example for gradient-based optimization that includes context-based constraints (prediction error) on the sparse support
Step 1: For a given time step i, locate the dominant S^* dimensions of $\hat{\mathbf{x}}[i]$ (derived from $\hat{\mathbf{z}}$) in Equation (4)
Step 2: Turn off the other dimensions of $\hat{\mathbf{z}}$. Choose (S-S^*) other dimensions from $\hat{\mathbf{x}}[i]$ that minimize the normalized prediction error, as the "invisible dimensions".
Step 3: Initialize the next step of the tracking problem based on 1 and 2 and repeat.

The second step in Table 2 limits choice of the (S-S^*) dimensions to a subspace based on physical knowledge of the application to avoid combinatorial challenges. A generalized and rigorous characterization of reasonable subspaces to choose from is work in progress and out of the scope of this work.

3.1 Tracking sparseness of the underlying distribution

We may measure the match between the estimated and expected sparseness of the solution through the prediction error, which statistically follows the true estimation error [21]. An investigation over experimental field data [19] collected at 200 meters range and 15 meters range for rough sea conditions demonstrated high time-variability over the value of the optimal weighting ratio λ. Therefore, direct measurement of the optimal sparseness against prediction error necessitates multiple solutions of the coefficients using different λ in Equation (3), which conflicts with our goal of efficient tracking. Therefore, we propose the following data-driven measure to update λ based on the tracked estimates:

$$\lambda = \sqrt{\|\hat{\mathbf{x}}\|_1^2 \Big/ \|\hat{\mathbf{x}}\|_2^2} \qquad (4),$$

i.e. the observed sparseness itself is used to update λ over the field data.

3. RESULTS

We provide representative numerical and data-driven evidence for adopting the NCNMS technique [10] among other existing methods for employing the proposed context-based sparse optimization approach. Figure 2 provides numerical evidence of the relative ability of the NCMNS technique [10] to track rapidly time-varying coefficients of shifting support against popular sparse sensing techniques [5-7]. Figure 3 illustrates the Delay-Doppler spread function computed using NCNMS [10] measured across 12 milliseconds delay spread and approximately double the Doppler resolution of Figure 1(b) across a much smaller observation window of 680 milliseconds, i.e. a more appropriate window length to assume quasi-stationarity in shallow water acoustics. The consistent success of the NCMNS technique [10] in high-precision tracking at comparable or less computation time over other methods combined with its ability to correctly reconstruct the bands of multipath arrivals in the shallow water acoustic channel, provide the basis for imposing the contextual constraints in Tables 1 and 2 using non-convex navigation.

3. CONCLUDING REMARKS

We explore the signal processing challenges posed by real environments to reliably employ established sparse optimization techniques that assume incoherence and

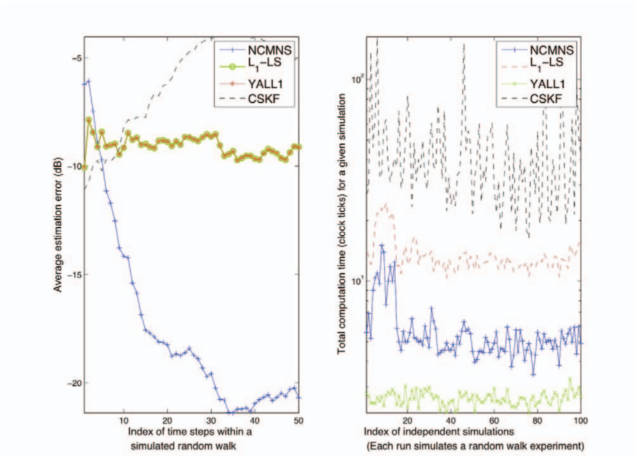

Figure 2: Comparison between NCNMS and other established techniques in tracking rapidly shifting support of the tracked coefficients.

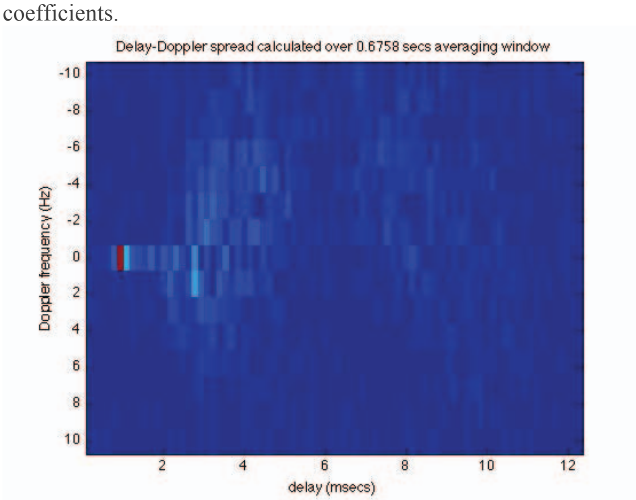

Figure 3: Delay-Doppler spread function calculated over 12 milliseconds delay spread and ± 10 Hz and a quasi-stationary window length of 0.6758 milliseconds. (Depth: 15 meters, Range between transmitter and receiver: 200 meters, rough sea conditions)

stationarity constraints on the underlying sparse estimation problem. Specifically, we propose a fundamentally non-traditional approach to adaptive sparse estimation of quasi-stationary variables that exploits context-driven proximity relationships over non-convex surfaces to navigate unobservable dimensions in coherent detection problems. We also provide a data-driven measure to track the sparseness of the underlying distribution. We further provide numerical and data-driven basis for adopting the NCNMS technique [10] over existing methods to employ context-based constraints for sparse optimization. Ongoing directions include exploiting proposed context-based external constraints to measure the true sparseness of the underlying distribution as well as generalize the concept of context-driven navigation beyond proximity and prediction error constraints.

REFERENCES

[1] E. Candes, T. Tao, Decoding by linear programming, IEEE Trans. Inform. Theory 51 (2005) 4203–4215.
[2] E. Candès, J. Romberg, T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. IEEE Trans. on Info.Th., 52(2)(2006) 489 – 509.
[3] E. J. Candes, The restricted isometry property and its implications for compressed sensing, Acad. des Sc. (2008).
[4] J. Tropp, A. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit., IEEE Trans. Info.Th. 53 (2007) 4655– 4666.
[5] S.-J.Kim, K.Koh, M.Lustig, S.Boyd, D.Gorinevsky, An interior- point method for large-scale l1-regularized least squares, IEEE J. Select. Topics Sig. Proc. 1 (2008) 606–617.
[6] J. Yang, Y. Zhang, Alternating direction algorithms for L1-problems in compressive sensing, Technical Report, Dept. of Math., Nanjing Univ., Nanjing, China, 2009.
[7] N. Vaswani, Kalman filtered compressed sensing, ICIP 2008.
[8] N. Vaswani, W. Lu, Modified-CS: Modifying compressive sensing for problems with partially known support, in: ISIT 2009.
[9]A. Sen Gupta, J. Preisig, 'A Geometric Mixed Norm Approach to Shallow Water Acoustic Channel Estimation', Proc. Eur. Conf. on Underwater Acoustics, Istanbul, Turkey, 2010.
[10] A. Sen Gupta, J. Preisig, "A Geometric Mixed Norm Approach to Shallow Water Acoustic Channel Estimation and Tracking", Elsevier Physical Communication Journal, Special Issue on Compressive Sensing, December 2011.
[11] M. A. Iwen, Simple deterministically constructible RIP matrices with sublinear Fourier sampling requirements, CISS 2009, 870-875.
[12] Jeffrey Blanchard, Coralia Cartis, and Jared Tanner, Decay properties of restricted isometry constants. (IEEE Signal Processing Letters, 16(7) (2009): 572-575.
[13] Jeffrey D. Blanchard, Andrew Thompson, On Support Sizes of Restricted Isometry Constants, Applied and Computational Harmonic Analysis, 29(3) (2010): 382-390.
[14] T. Tony Cai, Lie Wang, Guangwu Xu: New bounds for restricted isometry constants. IEEE Transactions on Information Theory 56(9) (2010): 4388-4394.
[15] Waheed U. Bajwa, Robert Calderbank, and Sina Jafarpour, Model selection: Two fundamental measures of coherence and their algorithmic significance. Proc. IEEE Int. Symp. Info. Th., 2010.
[16] R. Baraniuk, V. Cevher, M. F. Duarte, C. Hegde: Model-based compressive sensing. IEEE Transactions on Information Theory 56(4): 1982-2001 (2010).
[17] W. Li, J. C. Preisig, Estimation of rapidly time-varying sparse channels, IEEE J. Ocean. Eng. 32 (2007) 927 – 939.
[18] J. Preisig, G. Deane, Surface wave focusing and acoustic communications in the surf zone, J. Acoust. Soc. Am. 116 (2004) 2067–2080.
[19] A. Sen Gupta, J. Preisig, 'Tracking the Time-varying Sparsity of Channel Coefficients in Shallow Water Acoustic Communications', Proc. Asil. Conf. Sig., Sys. Comp., CA, 2010.
[20] P. Bello, Characterization of randomly time-variant linear channels, IEEE Trans. Commun. Sys. CS-11 (1963) 360–393.
[21] R. Nadakuditi, J. Preisig, A channel subspace post-filtering approach to adaptive least-squares estimation, IEEE Trans. Sig. Proc. 52 (2004).