Joint Sparsity Models for Wideband Array Processing

Petros T. Boufounos[a], Paris Smaragdis[b] and Bhiksha Raj[c]

[a]Mitsubishi Electric Research Laboratories, Cambridge, MA, USA;
[b]University of Illinois Urbana-Champaign, Champaign, IL, USA;
[c]Carnegie Mellon, Pittsburgh, PA, USA.

ABSTRACT

Recent work has demonstrated the power of sparse models and representations in signal processing applications and has provided the community with computational tools to use it. In this paper we explore the use of sparsity in localization and beamforming when capturing multiple broadband sources using a sensor array. Specifically, we reformulate the wideband signal acquisition as a joint/group sparsity problem in a combined frequency-space domain. Under this formulation the signal is sparse in the spatial domain but has common support in all frequencies. Using techniques from the model-based compressive sensing literature we demonstrate that it is possible to robustly capture, localize and often reconstruct multiple signals present in the scene.

Keywords: compressive sensing, joint sparsity, block sparsity, broadband array processing, localization, microphone array

1. INTRODUCTION

The recent advances in sparse representations and Compressive Sensing have had significant impact in the signal processing literature. The tools developed in this area have provided new and powerful methods to capture the structure embedded in a large number of interesting signals.

In this paper we examine how these tools impact broadband array processing. Specifically, we examine how joint and group sparsity models, can capture the scene geometry of passive broadband sensor arrays and eliminate ambiguities arising in such systems. We also provide some discussion on the limitations of such tools, and suggestions on how to overcome them.

The use of sparsity models for narrowband array processing and localization is already known in the literature.[1,2] Broadband arrays can be treated as separate narrowband problems,[2] although joint sparsity models have also been suggested and briefly explored to provide localization robustness.[1] Our work provides an extensive study of such models, incorporating recent results and algorithms from the Compressive Sensing and sparse recovery literature, in which we explore conditions both for localization and recovery of the signals of interest. While our models are formulated in the frequency domain, an alternative time-domain formulation is also possible.[3] However, a time-domain model implementation requires significant upsampling to accommodate for non-integer sample time delays, making the model very expensive computationally. The time-domain model further assumes the sources can be sparsely represented in some basis, i.e., requires additional source structure compared to our approach.

In the next section we establish the notation and provide a brief background on array processing, on sparsity and on joint sparsity models. Section 3 discusses how sparsity and joint sparsity formulations can localize and recover the source and presents an algorithm to recover the signals from the measurements. To further develop the problem intuition, the special case of linear arrays in the far-field approximation is examined in Section 4. Section 5 provides some experimental validation of our results, and Section 6 concludes.

Further author information: (Send correspondence to P.T.B.)
P.T.B.: E-mail: petrosb@merl.com, Telephone: 1-617-621-7575
P.S.: E-mail: paris@illinois.edu, Telephone: 1-217-265-6893
B.R.: E-mail: bhiksha@cs.cmu.edu, Telephone: 1-412-268-9826.

2. BACKGROUND

2.1 Microphone Arrays

Microphone arrays have been extensively studied in the literature. In this paper we adopt a commonly used discretized model of the problem. Specifically, we consider M microphones, each receiving a signal, denoted $Y_m(\omega)$, $m = 1, \ldots, M$ in the frequency domain. We discretize the area of interest to N gridpoints, indexed using $n = 1, \ldots, N$. Each of those points is a potential source location, emitting a signal $X_n(\omega)$. This signal is exactly equal to zero if there is no source in that location and non-zero if a source emits from that location. We assume that both sources and sensors are omnidirectional, although sensor directionality can be easily incorporated in the model.

The signal from each source n arrives to each sensor m with delay $\tau_{m,n}$. This delay is proportional to the distance of the source-sensor pair, $d_{m,n}$, and inversely proportional to the propagation speed of the transmitted wave, c, and satisfies $\tau_{m,n} = d_{m,n}/c$. The sensor receives the linear combination of all the transmitted signals, appropriately delayed. In the frequency domain, sensor m receives

$$Y_m(\omega) = \sum_{n=1}^{N} e^{-j\omega\tau_{n,m}} X_n(\omega), \tag{1}$$

compactly expressed in matrix form using

$$\mathbf{Y}(\omega) = \mathbf{A}(\omega)\mathbf{X}(\omega), \tag{2}$$

$$\mathbf{Y}(\omega) = \begin{bmatrix} Y_1(\omega) \\ \vdots \\ Y_M(\omega) \end{bmatrix}, \ \mathbf{X}(\omega) = \begin{bmatrix} X_1(\omega) \\ \vdots \\ X_N(\omega) \end{bmatrix}, \ \mathbf{A}(\omega) = \begin{bmatrix} e^{-j\omega\tau_{1,1}} & \cdots & e^{-j\omega\tau_{1,N}} \\ \vdots & \ddots & \vdots \\ e^{-j\omega\tau_{M,1}} & \cdots & e^{-j\omega\tau_{M,N}} \end{bmatrix}$$

We further assume that sources are broadband and that at most K of them exist in the scene at any time, where K is much smaller than N. The number of sources K might or might not be known in advance. Broadband sources emit signal in a wide bandwidth, i.e., $X_n(\omega)$ is non-zero for a wide interval of possible ω's.

In the array processing literature there are two problems of interest: source localization and source signal recovery. Under our formulation, the goal of source localization is to determine which of the N candidate locations in the grid contain transmitting sources (i.e., which of the acquired $X_n(\omega)$ are non-zero) but not necessarily to recover the contents of these sources (i.e., the coefficients $X_n(\omega)$). Recovering these coefficients is the goal of source signal recovery. Usually, signal recovery also requires or results to localization, but not the other way around. Localization is an easier problem and can be solved under certain conditions where recovery is not possible. We examine some example cases later in this paper.

2.2 Sparsity and Group Sparsity Models

Sparsity has recently emerged as a powerful tool in signal processing applications. A sparse model assumes that a signal, when transformed to an appropriate basis, has very few significant coefficients which can explain most of its energy. Promoting sparsity in some appropriate domain has proved to be a very efficient computational method to capture the structure of most natural and man-made signals processed by modern signal processing systems.

Sparsity is typically useful when inverting a linear system—usually underdetermined—of the form

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \tag{3}$$

where $\mathbf{y}$ is the M-dimensional measurement vector and $\mathbf{x}$ is the N dimensional sparse signal vector. Two fundamental approaches have emerged in the recent literature in to determine a sparse solution or approximate solution to (3): convex optimization and greedy algorithms. The former approach attempts to optimize a

sparsity promoting convex cost function balancing the solution sparsity with the data fidelity. For example for an underdetermined noiseless system the following optimization determines the sparsest solution

$$\widehat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_0 \text{ s.t. } \mathbf{y} = \mathbf{A}\mathbf{x}, \tag{4}$$

where the ℓ_0 norm counts the number of non-zero coefficients of $\mathbf{x}$. This is a combinatorialy complex problem but under certain conditions on $\mathbf{A}$ a convex relaxation of the ℓ_0 norm can produce the same result:[4, 5]

$$\widehat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ s.t. } \mathbf{y} = \mathbf{A}\mathbf{x}, \tag{5}$$

The geometric properties of the ℓ_1 norm, defined as $\|\mathbf{x}\|_1 = \sum_n |x_n|$, promote sparsity on $\mathbf{x}$. Under similar conditions the solution can also be computed using greedy algorithms such as orthogonal matching pursuit (OMP), Compressive Sampling Matching Pursuit (CoSaMP),[6] Subspace Pursuit (SP),[7] and iterative hard thresholding (IHT).[8]

A number of properties on $\mathbf{A}$ have been shown to guarantee efficient sparse recovery of the signal $\mathbf{x}$. For example, the coherence of $\mathbf{A}$ is defined as

$$\mu = \max_{i \neq j} \frac{|\langle \mathbf{a}_i, \mathbf{a}_j \rangle|}{\|\mathbf{a}_i\|_2 \|\mathbf{a}_j\|_2}, \tag{6}$$

where $\mathbf{a}_i$ and $\mathbf{a}_j$ are columns of $\mathbf{A}$.[9] The coherence essentially computes the smallest acute angle between any pair of columns of $\mathbf{A}$. If this angle is large, i.e., the coherence is small, sparse recovery is guaranteed. Similar guarantees, including robustness to noise and modeling can be provided using the restricted isometry property (RIP)[5] or the nullspace property (NSP).[10]

Joint and group sparsity models, and their variations, provide further structure to the signal of interest.[11–22] The former can be considered a special case of the latter, thus we only describe the latter. Under this model, the signal coefficients are partitioned into groups $\mathcal{G}_i$, which partition the coefficient index set $\{1, \ldots, N\}$. The group sparsity model assumes the only a few of these groups contain non-zero coefficients, and most groups contain all-zero coefficients.

To enforce group sparsity using an optimization problem a mixed ℓ_1/ℓ_2 norm, defined as

$$\|\mathbf{x}\|_{\ell_1/\ell_2(\mathcal{G})} = \sum_i \sqrt{\sum_{j \in \mathcal{G}_i} |x_i|^2}, \tag{7}$$

is typically used as a penalty in an optimization similar to (5):[12, 14, 18–21, 23]

$$\widehat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_{\ell_1/\ell_2(\mathcal{G})} \text{ s.t. } \mathbf{y} = \mathbf{A}\mathbf{x}. \tag{8}$$

The mixed norm computes the ℓ_1 norm of the total energy of the coefficients in each group. Greedy algorithms have also been extended to handle such problems.[11, 15, 16, 22]

In joint sparsity models, several sparse signals are measured together using the same measurement matrix $\mathbf{A}$. The model assumption is that all the signal share the same sparsity pattern. In other words the significant signal coefficients are located at the same positions for all signals. By considering the whole acquisition as a linear system it is straightforward to show that these models are special case of group sparsity models and a similar ℓ_1/ℓ_2 norm promotes the appropriate sparsity.

The notions of coherence, RIP and NSP can be extended to group and joint sparsity models.[17, 20, 21] Alas, the worst-case guarantees—typical in the sparsity and compressive sensing literature—derived from those properties are quite pessimistic and do not significantly exploit the additional structure and diversity commonly expected in such models. More optimistic results can be obtained using a probabilistic analysis,[21, 23] usually more appropriate in practice.

3. JOINT SPARSITY FOR BROADBAND ARRAYS

3.1 Sparse and Jointly Sparse Reconstruction

Given the formulation of (2), the signal $\mathbf{X}(\omega)$ only contains at most K non-zero components, i.e., it is K-sparse. Thus, sparsity-related methods are immediately applicable.[2] Specifically, under this formulation, localization becomes a support recovery problem while the signal recovery problem becomes a sparse recovery problem similar to the compressive sensing reconstruction problem.[4, 5]

Specifically, for narrowband sources, the source signal can be approximated as a singe-frequency tone at its center frequency ω_0 with particular amplitude and phase. In other words, the propagation equation is a single instance of (2), evaluated at the center frequency:

$$\mathbf{Y}(\omega_0) = \mathbf{A}(\omega_0)\mathbf{X}(\omega_0). \tag{9}$$

Since the scene is sparse, the solution to

$$\widehat{\mathbf{X}}(\omega_0) = \arg\min_{\mathbf{X}(\omega_0)} \|\mathbf{X}(\omega_0)\|_0 \text{ s.t. } \mathbf{Y}(\omega_0) = \mathbf{A}(\omega_0)\mathbf{X}(\omega_0), \tag{10}$$

will, assuming a noiseless system, result to both localization and reconstruction of the source, as long as the system is identifiable. Of course, in practice, (10) is computed using the ℓ_1 norm or a greedy algorithm to handle computational complexity and by relaxing the equality of the data fidelity term to accommodate noise.

Unfortunately, very often the system is not identifiable. For example, the solution to (10) always has sparsity that does not exceed the number of sensors M. Thus, if the number of sources is greater than M, the solution of (10) does not identify the system properly. Furthermore, fundamental ambiguities measured by the coherence of properties of $\mathbf{A}(\omega_0)$, can make the system non-identifiable, even if the scene contains fewer sources than microphones. We discuss some of these ambiguities in the next section.

If the sensors are broadband, then $\mathbf{X}(\omega)$ is has non-zero coefficients at the corresponding source locations for all frequencies ω emitted by the source. This means that all the vectors $\mathbf{X}(\omega)$ will have the same sparsity pattern for all ω, i.e., will be jointly sparse*. In other words, the solution should minimize a mixed ℓ_0/ℓ_2 norm, where the ℓ_2 norm is taken across frequencies—i.e., measures the total energy in all frequencies for each location—and the ℓ_0 norm counts the source positions which emit non-zero energy:

$$\|\mathbf{X}(\omega)\|_{\ell_0/\ell_2(\omega)} = \|\mathbf{X}_E\|_0, \ (\mathbf{X}_E)_n = \left\| \sum_\omega \sqrt{X_n(\omega)^2} \right\|_0, \tag{11}$$

For notational simplicity henceforth we use $\|\cdot\|_{0,2}$ to denote the mixed $\ell_0/\ell_2(\omega)$ norm. Similarly, we use $\|\cdot\|_{1,2}$ to denote the mixed $\ell_1/\ell_2(\omega)$ norm, which just substitutes the ℓ_1 for the ℓ_0 norm in (11).

Using this notation, the joint sparsity problem to localize and reconstruct the sources becomes

$$\widehat{\mathbf{X}}(\omega) = \arg\min_{\mathbf{X}(\omega)} \|\mathbf{X}(\omega)\|_{0,2} \text{ s.t. } \mathbf{Y}(\omega) = \mathbf{A}(\omega)\mathbf{X}(\omega) \text{ for all } \omega. \tag{12}$$

As with the narrowband model, in practice the solution is computed by relaxing the ℓ_0/ℓ_2 norm to an ℓ_1/ℓ_2 norm and convexifying the problem, or using one of the large number of greedy algorithms that have been developed for joint- and group-sparsity models.[11, 15, 16, 22]

Under the broadband assumption it is possible to identify the location of more than M sources using M sensors, although not necessarily reconstruct all M of them. The significant ambiguities arising in the single

---

*Technically joint sparsity in the context of compressive sensing refers to models in which the measurement matrix $\mathbf{A}(\omega)$ is the same for all the vectors $\mathbf{X}(\omega)$, which is not the same in our case. Joint sparsity models are special cases of the more general group sparsity models, and the solution approaches, using mixed-norms, are the same. Our model is also a special case of group sparsity model, with very similar structure to joint sparsity models, and the same approaches work. Because of this similarity we also refer to our model as a joint sparsity model.

Algorithm 1 The modified joint-sparsity CoSaMP algorithm used in the experiments in this paper.

---

initialize Iteration count $l = 0$, Initial estimate $\widehat{\mathbf{X}}^l(\omega) = (0)$, for all ω.
while not converged **do**
   Increase iteration count: $l \leftarrow l + 1$
   Compute residual: $\mathbf{R}(\omega) = \mathbf{Y}(\omega) - \widehat{\mathbf{X}}^{l-1}(\omega)$, for all ω
   Compute gradient: $\mathbf{G}^l(\omega) = \mathbf{A}^H(\omega)\mathbf{R}(\omega)$
   Compute proxy: $\mathbf{p} = \mathbf{G}_E^l$ {The energy vector $\mathbf{G}_E^l$ is computed similar to $\mathbf{X}_E$ in (11)}
   Identify support candidate: $\mathcal{T}^l = \mathrm{supp}\left(\mathbf{p}|_{2K}\right) \cup \mathrm{supp}\left(\widehat{\mathbf{X}}^{l-1}(\omega)\right)$ {$\mathrm{supp}(\cdot)$ determines the support set of a vector; $\mathbf{x}|_K$ keeps the K coefficients of $\mathbf{x}$ with the largest magnitude and sets the remaining to zero.}
   Line search: Find τ to minimize $\sum_\omega \left\| \mathbf{Y}(\omega) - \mathbf{A}(\omega)\left(\widehat{\mathbf{X}}^{l-1}(\omega) - \tau\mathbf{G}^l(\omega)\right) \right\|_2^2$, s.t. $\mathbf{X}(\omega)|_{\mathcal{T}^c} = 0$ for all ω.
   Form temporary estimate: $\mathbf{B}^l(\omega) = \left(\widehat{\mathbf{X}}^{l-1}(\omega) - \tau\mathbf{G}^l(\omega)\right)\Big|_{\mathcal{T}}$
   Compute final support: $\mathcal{S} = \mathrm{supp}(\mathbf{B}_E^l|_K)$
   Truncate and update estimate: $\widehat{\mathbf{X}}^l(\omega) = \mathbf{B}^l(\omega)\big|_{\mathcal{S}}$
end while
return Signal estimate $\widehat{\mathbf{X}}^l(\omega)$

---

frequency, i.e., narrowband, case manifest themselves for different locations at different frequencies. The joint sparsity model exploits these differences to resolve these ambiguities and localize the sources.

Alas, without prior information on the sources, system inversion is not always possible unless the system satisfies the narrowband identifiability conditions for all frequency. Thus, for some frequencies there is significant ambiguity between the location of two sources, it is not possible reconstruct the sources at these frequencies, even if we can localize them accurately. Furthermore, if there are more than M sources, reconstruction is fundamentally impossible. The fundamental cause for this limitation is that reconstruction eventually requires inversion at all frequencies ω of (2), with $\mathbf{A}(\omega)$ and $\mathbf{X}(\omega)$ restricted to the elements corresponding the K locations of the sources. If more than M sources exist, this inversion is rank-deficient. A separate model of the sources, e.g. source sparsity on some basis that is incoherent with the DFT basis or on a fusion frame basis[21] can resolve these ambiguities and enable reconstruction; however, we do not discuss such an extension in this paper.

3.2 Reconstruction Algorithms and Guarantees

As mentioned above, two fundamental algorithmic approaches exist to compute (12): convex relaxation and greedy algorithms. Convex relaxation has been studied in a number of recent publications.[12, 14, 18–21, 23] Inspired by compressive sensing results, most of these publication focus on worst-case analysis based on extensions of well-known system, properties such as the restricted isometry property (RIP), the nullspace property (NSP), or the system coherence. Similar guarantees can be provided for greedy algorithms, such as model-based CoSaMP or IHT.[11, 13, 15, 16, 22]

Unfortunately, these guarantees are very pessimistic for the broadband array case. In particular, the worst-case analysis degenerates to the analysis of a narrowband source, with center frequency such that the corresponding propagation equation (2) exhibits significant ambiguities. With true broadband sources, this analysis is inappropriate. Instead, a probabilistic average-case analysis[21, 23] would be more appropriate. Such analysis provides significantly more optimistic recovery conditions for joint and group sparsity models, assuming a probabilistic model on the sparsity pattern and the signal coefficients. Adaptation of this model to the specifics of the array system depends on the specific array geometry and is beyond the scope of this paper.

For the experiments in the remainder of this paper we use a slight variation of model-based CoSaMP,[6, 15] described in Algorithm 1. The main difference is the use of line-search[24] to reduce the cost in steps 7 and 8 instead of a full least-squares minimization using the pseudoinverse. This makes the algorithm significantly more efficient if the propagation matrix $\mathbf{A}(\omega)$ is only available in functional form and not explicitly. The modified algorithm requires the applications of $\mathbf{A}(\omega)$ and its adjoint, which is usually easily computable in array applications.
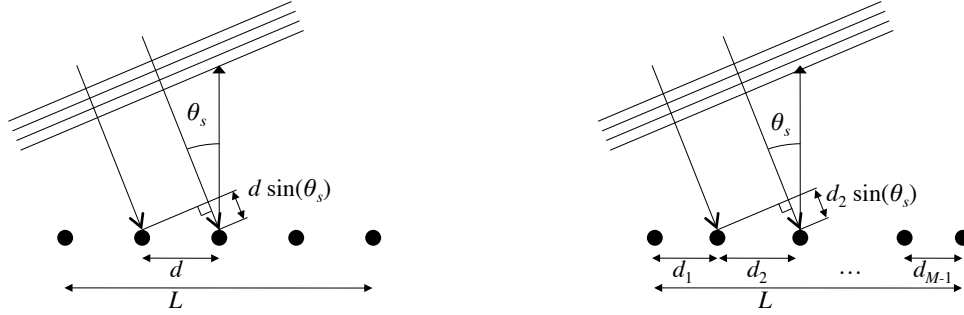
Figure 1. Uniform (left) and non-uniform (right) linear array geometry.

4. LINEAR ARRAYS

4.1 Linear Array Geometry

To provide further intuition, this section focuses on linear arrays under the far-field approximation. We examine both uniform and non-uniform arrays. Their geometry is summarized in Figure 1. As shown in the figure, all array elements (sensors) are collinear, placed at distance d apart for uniform arrays and pairwise distances $d_1, \ldots, d_{M-1}$ for non-uniform arrays. The length of the array, also known as the aperture size, is denoted by L and is equal to $(M-1)d$ for a uniform array and $\sum_{m=1}^{M-1} d_m$ for a non-uniform array.

The signal is assumed to arrive from far field, which means the signal travels in plane waves incident to the array at angle θ_S, the angle of the source with respect to the broadside of the array. This approximation is accurate for point sources located far from the array. The relative delay of the signal to the mth array element, compared to the time of arrival to the 1st element is equal to $\sum_{i=1}^{m-1} d_i \sin \theta_S / c$—which is equal to $(m-1)d \sin \theta_S / c$ for a uniform array. For $\theta_S = \pi/2$, $\sin \theta_S = 1$, we obtain the maximum delay of the signal to the mth element, denoted τ_m:

$$\tau_m = \sum_{i=1}^{m-1} \frac{d_i}{c}, \ 0 \text{ for } m = 1. \tag{13}$$

The distance of the source to the array does not figure in the far field approximation. Localization, therefore, aims to determine the angles θ in which sources are located. Given that this angle does not figure directly in the manifold, but always through its sine, we define $\psi = \sin \theta$, which is a more natural coordinate for the system. Thus, to form the array propagation matrix $\mathbf{A}(\omega)$ and solve the sparsity problem described in the previous section we discretize uniformly in the u-space[†]. Thus, we assume ψ takes values in an N-point grid in $[-1, 1]$, i.e.,

$$\psi_n = -1 + (n-1)\psi_0, \ \psi_0 = \frac{2}{N-1}, \ n = 1, \ldots, N, \tag{14}$$

where ψ_0 is the grid size.

Using this discretization, the array manifold matrix becomes

$$\mathbf{A}(\omega) = \begin{bmatrix} 1 & \cdots & 1 & \cdots & 1 \\ e^{-j\omega\tau_2} & \cdots & e^{-j\omega\tau_2\psi_n} & \cdots & e^{+j\omega\tau_2} \\ \vdots & \ddots & \vdots & & \\ e^{-j\omega\tau_m} & \cdots & e^{-j\omega\tau_m\psi_n} & \cdots & e^{+j\omega\tau_m} \\ \vdots & \ddots & \vdots & & \\ e^{-j\omega\tau_M} & \cdots & e^{-j\omega\tau_M\psi_n} & \cdots & e^{+j\omega\tau_M} \end{bmatrix}, \tag{15}$$

[†]It is often more convenient to define the k-space, in which the wavelength of the transmitted signal factors in the transformation. However, in our case the array is broadband and the wavelength is not fixed. Furthermore, the array is not necessarily uniform, therefore the distance of the array elements is not a single constant.
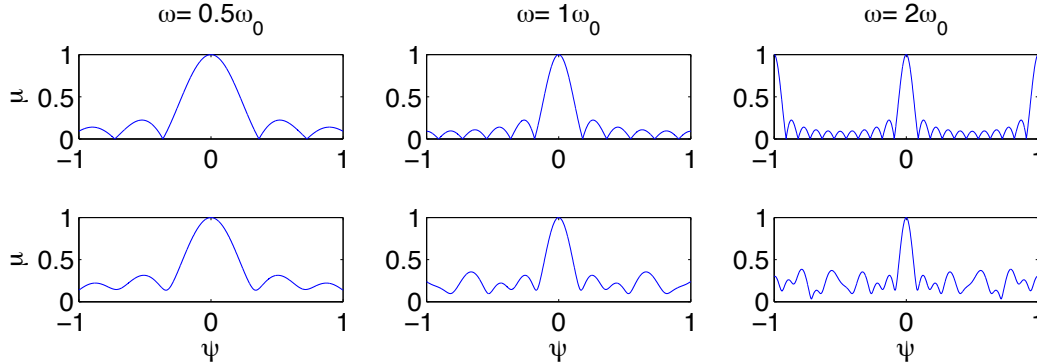
Figure 2. Example beampatterns (or coherence) for uniform (top) and random (bottom) arrays of 11 elements and same aperture size, at frequencies $\omega = \{0.5, 1 \text{ and } 2\}\omega_0$. The coherence is measured against the $\psi = 0$ direction, and the width of the main lobe determines the resolution.

which for a uniform array takes the form of a discrete Fourier transform (DFT). Thus, it can be efficiently applied using an FFT-based implementation and classical signal processing theory can be immediately used to analyze it.

4.2 Linear Array Coherence Properties

To understand the sparse reconstruction properties of $\mathbf{A}(\omega)$ we first examine it's coherence for different frequencies. Specifically, in Figure 2 we plot the coherence of the column of $\mathbf{A}(\omega)$ corresponding to $\psi = 0$ with respect to all other columns, i.e., $|\langle \mathbf{a}_0, \mathbf{a}_n \rangle| / \|\mathbf{a}_0\|_2 \|\mathbf{a}_n\|_2$. The coherence measure of the whole matrix is the maximum of this value away from the angle $\psi = 0$. High coherence between two locations indicates that these locations are easily confusable in narrowband localization problems and source recovery is very badly conditioned (or impossible, if the coherence is 1). Plotting the coherence with respect to the broadside ($\psi = 0$) column of the matrix provides us with a complete picture since this plot is symmetric around the selected direction and shift-invariant. In other words, plotting with respect another direction just shifts the plot. We should note that although we use sparsity-motivated terminology, such plots,referred to as beampattern plots, are widely used in the array processing community.

The figure demonstrates three different frequency selections, both for uniform and random arrays. The frequency ω increases from left to right as a multiple of $\omega_0 = 2\pi cM/L$, at factors of 0.5, 1, and 2. This frequency is selected so that at ω_0 the average distance of array sensors is one wavelength—equal to $\lambda = 2\pi c/\omega_0$. The width of the center lobe in the beampattern determines the resolution of the array. The manifold vectors in this center lobe have high coherence with each other, and therefore two sources coming from these direction cannot be easily discerned. As shown in the figure this width decreases as the frequency increases, i.e., higher frequencies improve the resolution. Unfortunately, for uniform arrays, if the distance of array elements remains constant, higher frequency introduces spatial aliasing. For example, at $\omega = 2\omega_0$ the broadside direction ($\theta = 0$) of the array is coherent with the endfire directions ($\theta = \pm\pi/2$), and therefore sources from any of these directions cannot be distinguished. Using randomized spacing on the array, at the same average sensor spacing, eliminates those ambiguities without increasing the beamwidth, at the expense of increasing the coherence with respect to some of the formerly low-coherence directions.

4.3 Joint Sparsity and Linear Arrays

The coherence properties described above focus on the properties of the manifold matrix for a single frequency ω, i.e., for a narrow band array. While these properties are helpful in analyzing broadband array properties as well, they do not paint the whole picture. To investigate these further, we plot the beampattern of a uniform array at frequencies $\omega = \{1, 2, 3, 4\}\omega_0$ in Figure 3.

We consider a source emitting from the broadside of the array, at $\psi = 0$. Using only the high frequency data, at $\omega = 4\omega_0$, this source cannot be accurately localized. In fact it could also be arriving from either side of the endfire, at $\psi = \pm 1$, or from $\psi = \pm 0.5$. Similar ambiguity exists at $\omega = 2\omega_0$ and $3\omega_0$, although with different
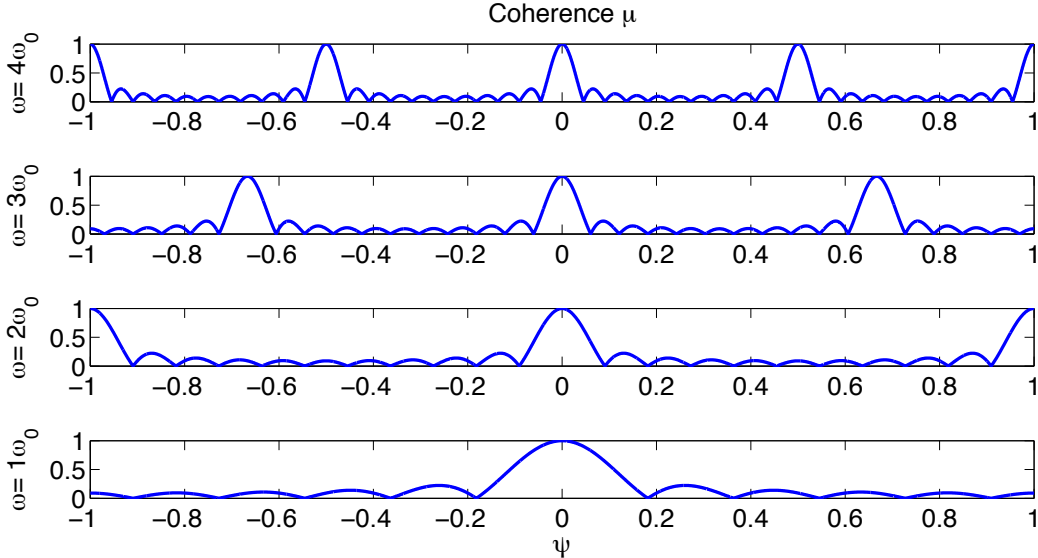
Figure 3. Coherence of a uniform 11-element linear array against the direction $\psi = 0$. At higher frequencies, resolution increases at the expense of significant ambiguity. Joint sparsity with the lower frequency information can resolve the ambiguity, while maintaining the high-frequency resolution.

directions. At $\omega = \omega_0$ that ambiguity does not exist, but the beamwidth is significantly higher, and therefore the resolution lower. It is, however, evident that only the broadside direction is consistent with all frequency bins. Joint sparsity models are designed to explicitly take this information into account. Thus, a broadband source emitting at all 4 frequencies can be uniquely resolved at the resolution implied by the high-frequency beampattern.

The source can often be uniquely localized even if it emits in a subset of the bandwidth, e.g., only at the highest two frequencies $3\omega_0$ and $4\omega_0$. However, if the source only emits at $2\omega_0$ and $4\omega_0$, the joint sparsity does not uniquely resolve the ambiguity. This is the reason that the classical worst-case analysis of joint-sparsity models is often very pessimistic in this case, whereas a probabilistic analysis[21, 23] is more appropriate.

Localization is not the same as reconstruction. In the example in the figure, consider a second source, at $\psi = 0.5$. While joint sparsity models will actually detect that source because of the additional content in the lower three frequencies, reconstruction is not possible at all frequencies. Specifically, it is fundamentally impossible to recover the content of the sources at $\omega = 4\omega_0$ because these directions are perfectly coherent. Without further information on the sources, it is impossible to invert the resulting system at that frequency.

5. EXPERIMENTAL EXAMPLES

To understand the performance of this algorithm we performed localization experiments for a variety of source and sensor configuration. For brevity of presentation, we only present the results a single array configuration here. Our experiments with other configurations demonstrated similar trends.

We consider a 1m×1m square scene with a sensor placed at each of the four corners, i.e, $M = 4$. We assumed acoustic sources with bandwidth 4KHz and used a grid resolution of 5cm. We used snapshots of length $F = 512$, 1024, and 2048, i.e., we solve (12) for F distinct frequencies ω. We examine the localization performance as the number of source increases, i.e., $K = 1, \ldots, 15$. The sources are randomly placed in the grid, and generate random while signals of the same variance. The left-hand side of Figure 4 plots the results of this experiment. Specifically, it plots the fraction of sources correctly localized, on average, as a function of the number of sources in the scene for the three different snapshot lengths. The right-hand side shows an example instance of the problem with 8 sources, indicated with a green circle, and the localization output of the algorithm, indicated with a red cross.
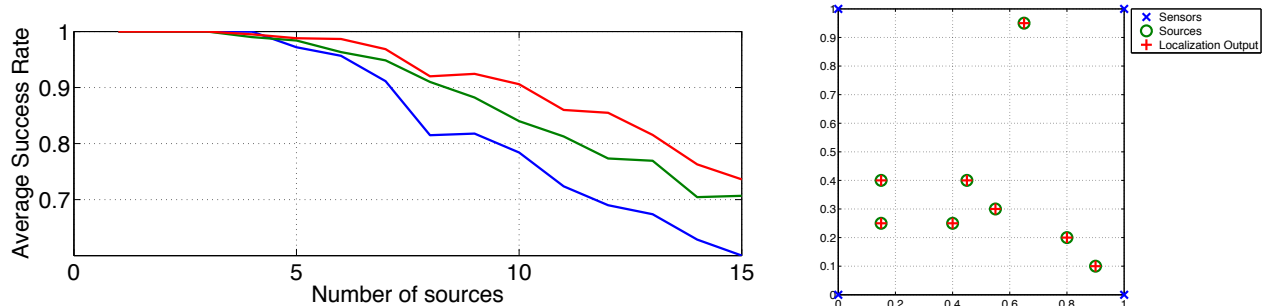
Figure 4. Experimental Results. The fraction of sources correctly localized, on average, is plotted as a function of the number of sources for a variety of snapshot lengths F (left). An example instance of the problem is shown with 8 sources plotted using green circles, and the localization output, plotted using red crosses (right). Blue 'x' marks the location of the sensors.

As is evident in the figure, localization is quite accurate, even as the number of sources exceeds the number of sensors $M = 4$. As expected, using longer snapshots provides more information and diversity in the problem and the localization performance improves at the expense of longer buffering and time delay, and a larger system to compute.

6. DISCUSSION

Joint/group sparsity models can be very useful in the array processing literature. In this work we attempted to examine how such models can be used in broadband array localization problems and improve localization and reconstruction accuracy. The tools developed in the recent literature—thanks to the advent of compressive sensing—provide a very useful algorithmic and theoretical framework to understand and solve such problems.

While we demonstrated how localization is possible, we also showed some fundamental limitations in reconstruction capabilities. Specifically the number of sources that can be reconstructed with such models cannot be greater than the number of sensors, unless further assumptions on the structure of the sources is assumed. Investigating algorithms and models that provide such structure, depending on the application, is a very promising area of further research.

ACKNOWLEDGMENTS

REFERENCES

[1] Malioutov, D., *A Sparse Signal Reconstruction Perspective for Source Localization with Sensor Arrays*, Master's thesis, MIT, Cambridge, MA (July 2003).

[2] Malioutov, D., Cetin, M., and Willsky, A., "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Processing* **53**(8), 3010–3022 (2005).

[3] Model, D. and Zibulevsky, M., "Signal reconstruction in sensor arrays using sparse representations," *Signal Processing* **86**(3), 624–638 (2006).

[4] Donoho, D. L., "Compressed sensing," *IEEE Trans. Inform. Theory* **52**(4), 1289–1306 (2006).

[5] Candès, E. J., Romberg, J. K., and Tao, T., "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math.* **59**(8), 1207–1223 (2006).

[6] Needell, D. and Tropp, J., "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis* **26**(3), 301–321 (2009).

[7] Dai, W. and Milenkovic, O., "Subspace pursuit for compressive sensing signal reconstruction," *Information Theory, IEEE Transactions on* **55**(5), 2230–2249 (2009).

[8] Blumensath, T. and Davies, M., "Iterative hard thresholding for compressed sensing," *Applied and Computational Harmonic Analysis* **27**(3), 265–274 (2009).

[9] Tropp, J. A., "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inform. Theory* **50**(10), 2331–2242 (2004).

[10] Zhang, Y., "A simple proof for recoverability of ℓ_1-minimization: Go over or under?." Rice University, Department of Computational and Applied Mathematics Technical Report TR05-09 (August 2005). `http://www.caam.rice.edu/~yzhang/reports/tr0509.pdf`.

[11] Tropp, J., Gilbert, A., and Strauss, M., "Algorithms for simultaneous sparse approximation. part i: Greedy pursuit," *Signal Processing* **86**(3), 572–588 (2006).

[12] Tropp, J., "Algorithms for simultaneous sparse approximation. part ii: Convex relaxation," *Signal Processing* **86**(3), 589–602 (2006).

[13] Peotta, L. and Vandergheynst, P., "Matching pursuit with block incoherent dictionaries," *Signal Processing, IEEE Transactions on* **55**, 4549 –4557 (Sept. 2007).

[14] Fornasier, M. and Rauhut, H., "Recovery algorithms for vector valued data with joint sparsity constraints," *SIAM J. Numer. Anal.* **46**(2), 577–613 (2008).

[15] Baraniuk, R., Cevher, V., Duarte, M., and Hedge, C., "Model-based compressive sensing," preprint (2008).

[16] Fornasier, M. and Rauhut, H., "Iterative thresholding algorithms," *Applied and Computational Harmonic Analysis* **25**(2), 187–208 (2008).

[17] Stojnic, M., Parvaresh, F., and Hassibi, B., "On the reconstruction of block-sparse signals with an optimal number of measurements," *Signal Processing, IEEE Transactions on* **57**, 3075 –3085 (aug. 2009).

[18] Eldar, Y. and Mishali, M., "Robust recovery of signals from a structured union of subspaces," *IEEE Trans. Info. Theory* **55**(11), 5302–5316 (2009).

[19] Kowalski, M. and Torrésani, B., "Sparsity and persistence: mixed norms provide simple signal models with dependent coefficients," *Signal, Image and Video Processing* **3**, 251–264 (Sept. 2009).

[20] Eldar, Y., Kuppinger, P., and Bolcskei, H., "Block-sparse signals: Uncertainty relations and efficient recovery," *Signal Processing, IEEE Transactions on* **58**(6), 3042 –3054 (2010).

[21] Boufounos, P., Kutyniok, G., and Rauhut, H., "Sparse recovery from combined fusion frame measurements," *IEEE Trans. Info. Theory* **57**(6), 3864–3876 (2011).

[22] Foucart, S., "Recovering jointly sparse vectors via hard thresholding pursuit," in [*Proc. Sampling Theory and Applications (SampTA)*], (May 2-6 2011).

[23] Eldar, Y. and Rauhut, H., "Average case analysis of multichannel sparse recovery using convex relaxation," *IEEE Trans. Inform. Theory* **56**(1), 505–519 (2010).

[24] Boyd, S. and Vandenberghe, L., [*Convex Optimization*], Cambridge University Press, New York, NY, USA (2004).