ADAPTIVE SEARCH FOR SPARSE TARGETS WITH INFORMATIVE PRIORS

Gregory Newstadt, Student Member, IEEE, Eran Bashan, and Alfred O. Hero III, Fellow, IEEE

Dept. of EECS, University of Michigan, Ann Arbor MI 48109-2122, USA
{newstage,bashan,hero}@umich.edu

ABSTRACT

This works considers the problem of energy constrained adaptive search for sparse targets given probabilistic prior knowledge of target locations. An Adaptive Resource Allocation Policy (ARAP) was introduced by Bashan (2008), showing significant gains over standard methods can be achieved without prior knowledge on the targets' locations. This work extends ARAP to account for non-uniform prior knowledge. It is shown that potential gains exist as compared to ARAP. Moreover, we show that by overestimating the true region of interest, the proposed search policy can always outperform ARAP in terms of worst-case gain. Lastly, results from an application involving estimating the approach of airplanes at an airport suggest that bi-level piecewise uniform priors are adequate approximations.

Index Terms— Adaptive sampling

1. INTRODUCTION

In this work, we consider the problem of estimating a sparse region of interest (ROI) with prior knowledge on the locations of the targets. In [1], a resource allocation approach called ARAP was developed for the detection and estimation of sparse signals containing targets in noise. ARAP was shown to asymptotically allocate resources in an optimal fashion when targets were assumed to be uniformly spread throughout the signal. We extend this analysis to non-uniform priors.

In many applications, prior information on target locations is highly non-uniform. This suggests that ARAP can be improved by taking advantage of this information. Moreover, in closed-form ARAP was limited to a 2-stage allocation policy, since measurements after the first stage yield a non-uniform prior for the next stage. Thus, developing the theoretical background for a 2-stage allocation using a non-uniform prior could provide invaluable insight into extending ARAP beyond 2-stages.

In this paper, we relax the assumption of a uniform prior in detecting/estimating the ROI using ARAP. We derive rules of thumb for when to use potentially inaccurate non-uniform prior knowledge as opposed to the uniform alternative. It is shown that potential gains over the uniform alternative exist as long as the non-uniform prior knowledge is reasonably accurate, where the gain depends on the individual application. Our analysis applies to simple bi-level piecewise uniform priors. The results indicate that a bi-level prior may be a sufficient approximation to the underlying model for many applications.

Much of the previous work in resource allocation has been directed towards inhomogeneous signals [2], [3]. Castro introduces a 2-stage algorithm to first coarsely search a signal or image, and then refine areas near boundary points. It was shown that fast convergence rates (in MSE) were possible for certain classes of inhomogeneous signals. In this work, we are interested in signals only restricted to having a small ROI.

Thus, our signals could be considered sparse in the sense that the ratio of the ROI to the entire signal is small. In compressed sensing [4], the goal is to reconstruct a sparse signal using many fewer samples than the size of the signal. Haupt, Castro and Nowak [5] introduce distilled sensing as an adaptive approach for recovering sparse signals in noise, and show that they can recover much weaker signals as compared to non-adaptive approaches. However, neither approach considers the best way to allocate resources within a stage, nor do they directly estimate the ROI.

The rest of this paper is organized as follows. In Section 2 we introduce notation and review ARAP. Section 3 presents a performance analysis, including uncertainty in the prior knowledge and an application where a non-uniform prior is likely. Finally, we conclude and point out future work in Section 4.

2. BACKGROUND OF ARAP

We will follow the notation laid out in [1]. Consider a discrete space $\mathcal{X} = \{1, 2, \ldots, Q\}$ containing Q cells. Let Ψ denote a region of interest (ROI) in $\mathcal{X}$, i.e., $\Psi \subseteq \mathcal{X}$. In the sequel Ψ will be a randomly selected small subset of $\mathcal{X}$, $|\Psi| \ll |\Psi^c|$, where $|\Psi|$ equals the number of elements in Ψ and Ψ^c is the relative complement, $\mathcal{X} \setminus \Psi$, of Ψ. Let I_i be an indicator function of the ROI such that $I_i = 1$ if $i \in \Psi$ and 0 otherwise, and $\{p_i = \Pr(I_i = 1)\}_{i=1}^{Q}$ is an associated set of prior probabilities. Define

$$y_i(t) = \sqrt{\lambda(i,t)}\theta_i I_i + n_i(t), \quad i = 1, 2, \ldots, Q \qquad (1)$$

be the measurement of the i-th cell at time t, where $\lambda(i,t) \geqslant 0$ is the search effort, e.g. energy, allocated to cell i at time t, $\theta_i \sim \mathcal{N}(\mu_\theta, \sigma_\theta^2)$ is the random target return, and $n_i(t) \sim \mathcal{N}(0, \sigma^2)$ is additive Gaussian noise. In [1] the following cost function was introduced

$$J(\Lambda) = \sum_{i=1}^{Q} \frac{\nu I_i + (1-\nu)(1-I_i)}{\Lambda(i)}, \qquad (2)$$

with $\Lambda(i) = \sum_{t=1}^{2} \lambda(i,t)$. and $\nu \in [\frac{1}{2}, 1]$. When $p_i = p$ for all i (i.e., uniform prior knowledge), the expected value of (2) can be directly minimized subject to a total energy constraint λ_T. The solution of the minimization problem yields a search policy, $\{\lambda(i,t)\}$, which we refer to as ARAP.

For a single stage, it was shown in [1] that the optimal allocation is given by

$$\lambda(i,1) = \frac{\alpha \lambda_T \sqrt{p_i}}{\sum\limits_{j=1}^{Q} \sqrt{p_j}} \qquad (3)$$

where α is the percentage of total energy used in the first stage (in the case of a 1-stage algorith, $\alpha = 1$). Moreover, it was shown that given $\{\lambda(i,1)\}_{i=1}^{Q}$ and the measurements at the first stage $Y(1)$, the optimal allocation at the second stage is given by

$$\lambda(i,2) = \left(\frac{\lambda_T - \sum\limits_{j=1}^{Q} \lambda(j,1)}{\sum\limits_{j=k_0+1}^{Q} \sqrt{w_{(j)}}} \sqrt{w_{(i)}} - \lambda(i,1) \right) I(i > k_0) \quad (4)$$

where $p_{I_i|\boldsymbol{y}(1)} \triangleq \Pr(I_i = 1|\boldsymbol{y}(1))$, $w_i = \nu p_{I_i|\boldsymbol{y}(1)} + (1-\nu)(1 - p_{I_i|\boldsymbol{y}(1)})$, $w_{(i)}$ is an ordered version of w_i, and k_0 defines a cutoff point based upon that ordering. Proof of existence of a unique k_0 is given in [1]. However, since $\{\lambda(i,1)\}_{i=1}^{Q}$ can take on Q values in general, minimizing the expected value of (2) becomes a combinatorially complex problem. Therefore, [1] is restricted to the case of uniform prior knowledge on target locations, i.e., $p_i = p$ for all i, reducing the optimization to a single parameter grid search.

In this work, we elaborate on the myopic approach suggested in [1] for non-uniform priors. Rather than solve the minimization problem directly, resources are allocated optimally within each stage as follows:

Algorithm 1: Two stage Non-uniform Adaptive Resource Allocation Policy (NU-ARAP)

Step 1: Allocate first stage according to (3) and measure $\boldsymbol{y}(1)$ in (1).
Step 2: Compute posterior probabilities $p_{I_i|\boldsymbol{y}(1)} \triangleq \Pr(I_i = 1|\boldsymbol{y}(1))$ and $w_i = \nu p_{I_i|\boldsymbol{y}(1)} + (1-\nu)(1 - p_{I_i|\boldsymbol{y}(1)})$.
Step 3: Rank order the w_i's using the permutation operator τ, then use λ_1^* and the ordered statistic $w_{\tau(i)}$ to find a threshold k_0.
Step 4: Given k_0, apply $\lambda(i,2)$, the energy allocation, to cell i as in 4 and measure $\boldsymbol{y}(2)$

To complete the definition of NU-ARAP, one must perform a line search over α, the percentage allocated at each stage to determine the best allocation. Due to the suboptimality that is inherent when a myopic approach is used, this work is concerned with answering the following questions: (1) Are there benefits to using a suboptimal approach with a non-uniform prior, as compared to an asymptotically optimal solution (i.e., ARAP)? In other words, do the gains of using a non-uniform prior outweigh the loss in optimality? And (2), how robust is NU-ARAP to mismatches in the prior distribution of target locations?

It should be noted that this work faces a fundamental tradeoff in potential gains versus practicality. On the one hand, if an underlying model is highly complex and non-uniform, there is a clear advantage to using this model in creating allocation policies, and larger gains would be expected. Conversely, as the underlying model becomes more intricate, it becomes more likely that there will be inaccuracies and/or mismatches in the assumed models. For practical reasons, this work considers the case of two-level piecewise uniform priors that may capture the essence of a general non-uniform prior, while maintaining analytical simplicity.

3. PERFORMANCE ANALYSIS

This section presents performance analysis of NU-ARAP in several fashions. Section 3.1 gives asymptotic properties of NU-ARAP. Sections 3.2 and 3.3 present a basis for analysis and corresponding simulations for the robustness of the algorithm. Lastly, Section 3.4 considers an air traffic control (ATC) radar application that fits well with the presented framework.

3.1. Asymptotic properties for $\nu = 1$

In the current section the following asymptotic properties of NU-ARAP are presented (where by asymptotic we mean high SNR and large Q):

1) Consistency - Given that a target is present in cell i, the posterior probabilities $p_{I_i|\boldsymbol{Y}(1)} \rightarrow 1$ in probability for all $i \in 1, \ldots, Q$ as long as $\hat{p}_i > 0$. For the complement case where $I_i = 0$, we have $p_{I_i|\boldsymbol{Y}(1)} \rightarrow 0$.
2) Asymptotic Gain - The gain of using the suboptimal approach described in the previous section over using an exhaustive search approaches the optimal gain, $-10 \log p$ for $p = E|\Psi|/Q$ as SNR $\rightarrow \infty$, as long as $I_i = 1 \rightarrow p_i > 0$.

Consistency follows from the fact that $p_{I_i|y_j}$ is proportional to the LRT, which is known to be the uniformly most powerful test for comparing Gaussians with different means (as is the case here). A full proof is given in [6]. The asymptotic gain follows directly from [1] where for $p_{I_i|\boldsymbol{y}(1)} \rightarrow I_i$, $\lambda(i,1) \rightarrow 0$ and the gain approaches the optimal gain over an exhaustive search.

3.2. Robustness analysis

In this section, the performance of NU-ARAP is analyzed in the context of its robustness to mismatches in the assumed prior as compared to the underlying model. For these purposes, the underlying model will be denoted as $\boldsymbol{g}^* = \{p_i\}_{i=1}^{Q}$, and the assumed prior as $\hat{\boldsymbol{g}} = \{\hat{p}_i\}_{i=1}^{Q}$. Note that in general, allocations will depend on both the assumed and underlying priors, since the first stage allocation depends only on the assumed prior, while subsequent allocations depend on both. Moreover, $\boldsymbol{u} := \{\hat{p}_i = p\}_{i=1}^{Q}$ will be referred to as a uniform prior and will be used as a base for comparison. The expected cost minimized by our algorithm given the underlying model, assumed model, and percentage of energy allocated at each stage, $E[J(\Lambda)|\boldsymbol{g}^*, \hat{\boldsymbol{g}}, \alpha]$, is derived in [6]. For sake of space, the equations are not reproduced here.

In this work, mismatches are considered by making the assumption that both $\boldsymbol{g}^*$ and $\hat{\boldsymbol{g}}$ are elements of a class of priors, $\boldsymbol{G}$ (note that these are priors for each cell i, and not over all cells; i.e., they do not integrate to one over all i). Thus, for any assumed model $\hat{\boldsymbol{g}}$, two criteria are considered: (1) the expected cost, $C(\hat{\boldsymbol{g}}; \boldsymbol{g}^*)$, and (2) the worst-case gain when compared with using a uniform prior, $K(\hat{\boldsymbol{g}})$. Expressions for these quantities are given below.

$$C(\hat{\boldsymbol{g}}; \boldsymbol{g}^*) = E[J(\Lambda)|\boldsymbol{g}^*, \hat{\boldsymbol{g}}, \alpha] \quad (5)$$

$$K(\hat{\boldsymbol{g}}) = \max_{\boldsymbol{g}^* \in \boldsymbol{G}} -10 \log \frac{C(\hat{\boldsymbol{g}}; \boldsymbol{g}^*)}{C(\boldsymbol{u}; \boldsymbol{g}^*)} \quad (6)$$

3.3. Robustness simulations

The remainder of this section considers a specific class of prior models in order to form intuition on the robustness of NU-ARAP. Let us consider a set of priors $\boldsymbol{G} = \{\hat{\boldsymbol{g}}_1, \hat{\boldsymbol{g}}_2, \ldots, \hat{\boldsymbol{g}}_m\}$ and $\hat{\boldsymbol{g}}_i = \{p_j^i\}_{j=1}^{Q}$. As mentioned previously, there is a tradeoff between the complexity of the assumed prior versus the potential gains over the uniform alternative. For this reason, then, the considered set $\boldsymbol{G}$ contains only 2-level piecewise uniform priors; i.e., each $p_j^i \in \hat{\boldsymbol{g}}_i$ can take one of two values. For valid comparisons between priors, the number of expected targets, $E|\Psi| = \sum_{j=1}^{Q} p_j = Q p_{unif}$ is set to a single constant. Define $H_i = \{j : p_j^i = p_H^i\}$ to be the set corresponding to the high probability region, and L_i to be its complement set. For simplicity of analysis, the low-level probability will be constant across all priors in $\boldsymbol{G}$; i.e., $p_j^i \in L_i = p_L$ for all i, j. In this way, each $\hat{\boldsymbol{g}}_i$ can be described completely by the set H_i. Lastly, it will be assumed that the indices are ordered, so that

(a) Worst-Case Gain
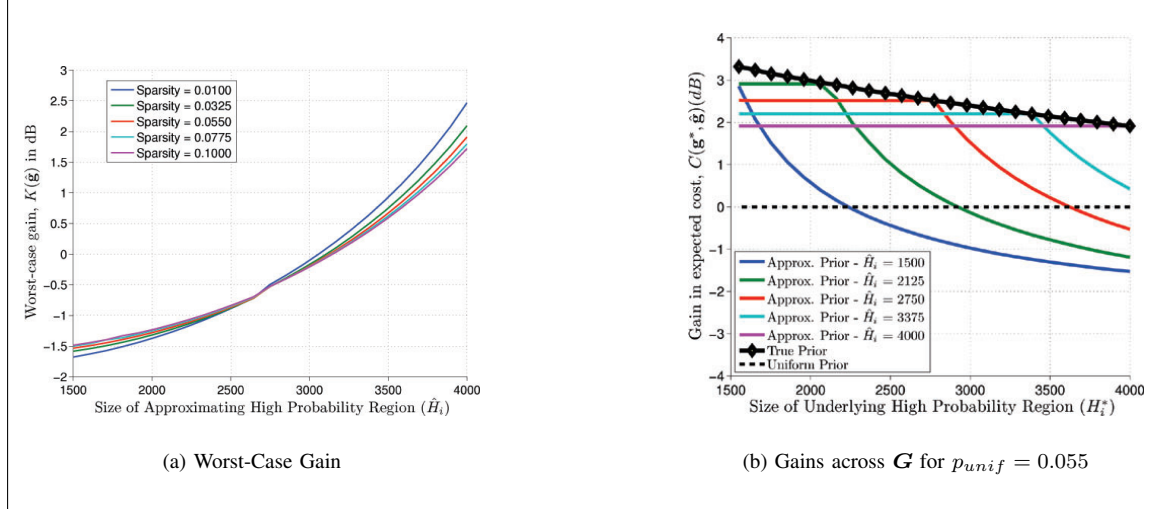
(b) Gains across G for $p_{unif} = 0.055$

Fig. 1. In (a), the worst-case gain is plotted as a function of the approximating prior, $\hat{g}_i$ which is characterized completely by the size of the high probability region, $\hat{H}_i$. Across all sparsity levels, the worst-case gain is monotonically increasing in $\hat{H}_i$. In (b), the gains in expected cost over using the uniform alternative for several possible approximating priors is plotted as a function of the underlying prior.

$j > k \rightarrow p_i^j \geqslant p_i^k$. For a discussion of these constraints, see [6]. Figure 2 plots two possible priors in the described G.
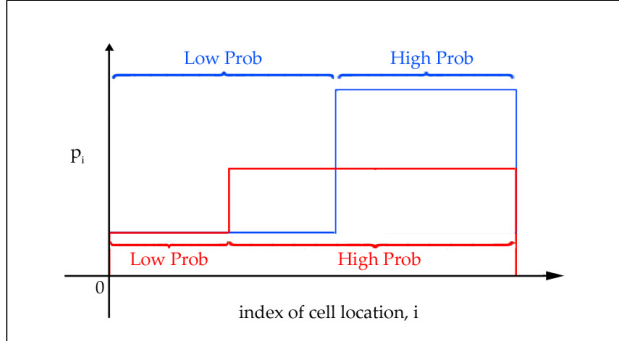

Fig. 2. This plot shows two possible priors in the considered class G that share the same number of expected targets. Each curve represents a single prior parameterization, $g_i \in G$.

For the simulations, we let $Q = 10000$, $|H_i| \in [1500, 4000]$, $p_L = 0.002$, and vary the number of expected targets across five sparsity levels, $p_{unif} = \{0.01, 0.0325, 0.055, 0.0755, 0.1\}$. Figure 1(a) plots $K(\hat{g})$ versus $\hat{g}$. Clearly, the figure suggests that the worst-case gain is monotonically increasing in $|H_i|$. Intuitively, this indicates that in terms of worst-case gain, it is better to oversample the region where targets are likely to be.

Figure 1(b) presents the gain in expected cost over the uniform alternative as a function of the underlying model for a single sparsity level ($p_{unif} = 0.055$). The performance of several approximating priors ($|\hat{H}_i| \in \{1500, 2125, 2750, 3375, 4000\}$) is compared to the underlying model, as well as the uniform alternative. It should be noted that the gain in expected cost decays rapidly when the high probability region is underestimated (i.e., $|\hat{H}_i| < |H^*|$), indicating that it is better to be conservative in the estimation of H_i^*, agreeing with the analysis in the previous paragraph. Note that as long as the reverse is true (i.e., the high probability region is overestimated), then the gain is nearly constant and equal to a point on the underlying model curve. This suggests that as long as the assumed prior is sufficiently conservative in

estimating the high probability region, then NU-ARAP will always perform at least as well, and typically much better, as ARAP with a uniform prior.

3.4. Application: ATC radar example

(a) Physical problem
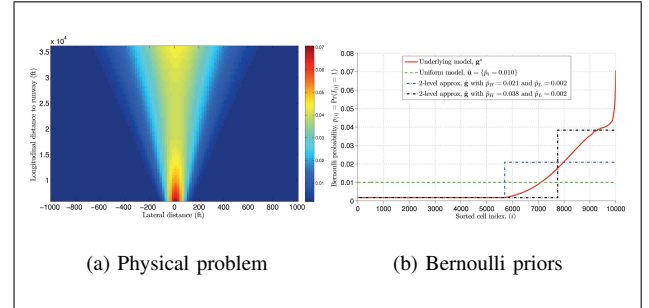
(b) Bernoulli priors

Fig. 3. In (a), we show a possible layout of a landing field based on the statistics given in [7]. In (b), we show four possible priors to represent the same problem, as a function of sorted cell index. The blue curve represents the underlying model and the green curve represents the equivalent uniform model (with a sparsity level of 0.01). The remaining two curves, which differ in the size of the high-probability regions, represent 2-level approximations to the underlying model. In all cases, the number of expected targets is kept constant.

Here we simulate an active air traffic control (ATC) radar system to detect and estimate locations of airplanes at a landing field. The baseline for comparison is an ATC that scans a 360 degree region at constant angular velocity, corresponding to an exhaustive resource allocation policy. An ARAP approach would acquire an initial scan at a constant angular velocity, but rotate the radar with a non-uniform angular velocity during the second scan as a function of the previous measurements. However, airplanes tend to approach from certain directions with much higher probability, due to geographic, political, and safety constraints. Thus, when determining a search

policy, we would expect a highly non-uniform prior. Therefore, it would be expected that one could do better by initially scanning the region at variable angular velocities corresponding to these probabilities. It is this NU-ARAP scenario that we consider here.

In [7], Shortle gives statistical characteristics of aircraft landing at a single runway at the Detroit Metropolitan Airport (DTW). In particular, the paper suggests that the lateral position of an aircraft landing at runway 21L is approximately Gaussian distributed given its longitudinal distance from the runway, though it differs from a standard Gaussian in the tails (which are exponentially distributed). Moreover, the variance was found to be a monotonically increasing function of distance.

In these simulations, the underlying model is chosen to roughly approximate aircrafts approaching runway 21L at the DTW airport. Thus, the field of view is set to size 30000×1000 ft^2. The underlying prior is assumed to be proportional to a Gaussian distribution along the lateral direction with a variance that is proportional to its longitudinal distance. It is further assumed that multiple targets are more likely for distant aircraft, but that these decrease as the aircraft approaches the airstrip. The prior probabilities are exponentially scaled correspondingly along the longitudinal direction to boost probabilities further from the runway. This simulation also considers two approximating priors, which differ only in the size of the high probability region, while keeping the number of expected targets constant. Lastly, for the sake of Monte-Carlo analysis, the minimum Bernoulli probability is set to 0.002 in order to guarantee a sufficient number of expected targets. The priors are shown in their physical representation in Figure 3(a), and as a set of Bernoulli priors in 3(b).
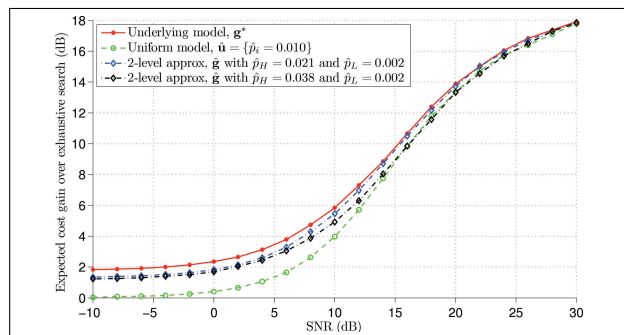

Fig. 4. This plot compares the performance of the four possible priors in terms of gains over an exhaustive search as a function of SNR. The largest gains occur when the underlying prior is known (solid red curve), while all non-uniform priors outperform the uniform alternative (green curve) by at least 1 dB for SNR values below 15 dB. Furthermore, the approximations (blue and black curves) may be more practical in a real application, yet they still maintain significant gains over the uniform model.

For $Q = 10000$ and priors shown in Figure 3, the empirical expected cost from NU-ARAP is computed over 2000 Monte Carlo samples for each SNR point (SNR is defined as the signal to noise ratio per cell for an equal energy allocation, i.e., $SNR = \frac{\lambda_T/Q}{\sigma^2}$, where σ^2 is the noise variance). Recall that when a uniform prior is used, NU-ARAP reduces to the 'optimal' allocation described by ARAP. Figure 4 compares the gains with respect to an exhaustive search as a function of SNR and prior information. There are several notable characteristics in this plot.

First, although a myopic approach is used for all of the non-uniform priors, NU-ARAP yields improvement over ARAP at SNR values below 15 dB (which in many instances represent more realistic situations). As one may expect, the largest gains occur when the underlying model is known to NU-ARAP (red curve).

However, practicality limits access to complete (and accurate) priors. It is encouraging that the 2-level approximations give the most significant gain margin as compared to the uniform model. From these results, it is conjectured that only a few terms in a Haar basis for the prior probabilities may be required in order to retain the vast majority of the lost gain. Note also that in all cases, the asymptotic gains (i.e., for SNR $\geqslant$ 30 dB) approach the optimal value $-10 \log 0.01 = 20$ dB as discussed in Section 3.1. Lastly, it should be noted that Figure 4 agrees with the analysis in the previous section, as the more conservative approximation (blue curve) to the underlying model outperforms the other approximation (black curve).

4. CONCLUSIONS AND FUTURE WORK

In this work, we have extended previous work on adaptive resource allocation for search by exploring the case where prior knowledge is non-uniform. It was shown that there is room for performance gain when the underlying model is non-uniform and known, even when a suboptimal allocation policy is used. Moreover, even in the case where there might be uncertainty in the prior knowledge, we have provided a simple means for comparing mismatches through the analytical computation of the objective function minimized by NU-ARAP. Although the robustness analysis will be application-dependent, our results suggest that NU-ARAP will outperform the uniform alternative when we conservatively estimate the underlying prior model for target locations. Additionally, using NU-ARAP in the application of ATC radar, we find that 2-level approximations to the underlying prior still lead to significant gains.

Future work will consider more general classes of priors, and in particular, piecewise constant priors that may be represented by a Haar basis. Moreover, we plan to extend our analysis to more concrete measures, such as MSE or probability of detection.

5. REFERENCES

[1] E. Bashan, R. Raich, and A. O. H. III, "Optimal two-stage search for sparse targets using convex criteria," *IEEE Transaction Signal Processing*, vol. 56, no. 11, pp. 5389–5402, November 2008.

[2] R. Castro, J. Haupt, and R. Nowak, "Compressed sensing vs. active learning," in *Proceedings 2006 International Conference on Acoustics, Speech and Signal Processing*, vol. 3, May 2006, pp. III–III.

[3] R. Willett, A. Martin, and R. Nowak, "Backcasting: adaptive sampling for sensor networks," in *Third International Symposium on Information Processing in Sensor Networks*, April 2004, pp. 124–133.

[4] J. A. Tropp, "Just relax: convex programming methods for identifying sparse signals in noise," *IEEE Transactions on Information Theory*, vol. 52, no. 3, pp. 1030–1051, March 2006.

[5] J. Haupt, R. Castro, and R. Nowak, "Distilled sensing: Selective sampling for sparse signal recovery."

[6] G. Newstadt, E. Bashan, and A. O. Hero III, "Adaptive search for sparse targets with informative priors," University of Michigan, Communications and Signal Processing Lab., Tech. Rep. 393, September 2009.

[7] J. Shortle, Y. Zhang, and J. Wang, "Statistical characteristics of aircraft arrival tracks," 7 2009, submitted for publication.