

# Two-stage Multi-Scale Search for Sparse Targets

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

January 25, 2011

**Abstract**—We consider the problem of energy constrained and noise-limited search for targets that are sparsely distributed over a large area. We propose a multi-scale search algorithm that significantly reduces the search time of the adaptive resource allocation policy (ARAP) introduced in [Bashan et al., 2008]. Similarly to ARAP, the proposed approach scans a  $Q$ -cell partition of the search area in two stages: first the entire domain is scanned and second a subset of the domain, suspected of containing targets, is re-scanned. The search strategy of the proposed algorithm is driven by maximization of a modified version of the previously introduced ARAP objective function, which is a surrogate for energy constrained target detection performance. We analyze the performance of the proposed multi-stage ARAP approach and show that it can reduce mean search time with respect to ARAP for equivalent energy constrained detection performance. To illustrate the potential gains of M-ARAP, we simulate a moving target indicator (MTI) radar system and show that M-ARAP achieves an estimation performance gain of 7 dB and a 85% reduction in scan time as compared to an exhaustive search. This comes within 1 dB of the previously introduced ARAP algorithm at a fraction of its required scan time.

## I. INTRODUCTION

This work considers the problem of localizing and estimating targets in noise. We are specifically interested in cases where targets occupy only a small fraction of the scanned domain, which we refer to as the region of interest (ROI). Related problems include detection of tumors in early cancer detection and surveillance systems using agile radars. In [1], a novel cost function was introduced, and the solution of a related minimization problem yielded an asymptotically optimal adaptive resource allocation policy, namely ARAP. In this work we introduce a multi-scale modification of ARAP (M-ARAP) and show that it leads to additional performance gains in search complexity, target localization and target amplitude estimation.

Bandiera et. al., showed significant performance gains in detecting targets in heterogeneous noise when one knows the locations of targets within a scene [2]. They propose an adaptive detection scheme that estimates the ROI in an intermediate step to the target detection problem. Their work suggests that taking advantage of a probabilistic characterization of the measurements can lead to significant detection gains. In this paper, we propose a method for estimating the ROI using a two-stage resource allocation policy.

The notion that one can reduce the number of measurements when sampling sparse signals has been developed by many researchers in recent years. These include compressive sensing, [3], [4], adaptive testing [5], [6], and adaptive sampling [7], [8]. The contribution of this paper is most closely related to the adaptive sampling problem.

Abdel-Samad and Tewfik [8] propose an adaptive sampling solution to the problem of how to best allocate  $N$  measurements to find a single target hidden in  $Q$  cells for radar target localization when  $N < Q$ . They propose a hierarchical approach recursively grouping the  $Q$  cells into  $q < Q$  groups in a tree like structure. They assume that the radar beam pattern can be shaped accordingly and that signal to noise ratio (SNR) decreases as the group size increases. Their multiple hypothesis testing approach is computationally intense and does not easily scale to large  $N$  and  $Q$  ( $Q = 64$  is used in their example). The high solution complexity limits the number of measurements they can allocate at each stage of the sequential search. On the other hand, our proposed M-ARAP search strategy is explicitly designed to detect and localize multiple targets and has lower solution complexity than the multi-hypothesis testing approach of [8].

Like the M-ARAP approach proposed in this paper, the adaptive sampling procedure proposed by Haupt, Castro and Nowak [9], called distilled sensing, is a general sequential multiple hypothesis testing approach that simultaneously seeks to localize the target and to test for presence of targets in the scene under a sensing energy constraint. Like ARAP [1], the generalized version of the distilled sensing procedure [7] applies adaptive coordinate-wise thresholding to each pixel, retaining only those pixels that exceed the threshold for consideration at the next stage. They show that there exists a sequence of thresholds and sensing energy parameters that asymptotically guarantee perfect recovery of the targets at arbitrarily small false discovery rate. While they have similarities, ARAP approaches introduced here and in [1] differ from distilled sensing approaches of [7], [9] in important ways. (1) ARAP approaches adopt a Bayesian framework and generate a posterior probability of target presence at each pixel. (2) ARAP approaches use an optimization procedure to select the sensing energy parameters that are functions of the posterior probabilities. (3) The ARAP optimization is simply performed on a surrogate convex performance metric. M-ARAP can be interpreted as an extension of ARAP that incorporates beam-forming, i.e. formation of linear combinations of pixels, at the first stage. This results in improved search performance with respect to the previous coordinate-wise thresholding method.

To the best of our knowledge, the first multi-scale search approach was the adaptive pooled blood sample algorithm introduced in the early 1940's. Dorfman considered the prob-

Copyright (c) 2011 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

Eran Bashan, Gregory Newstadt, and Alfred Hero are with the Dept. of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, E-mail: ({bashan},{newstage},{hero}@umich.edu).

lem of detecting defective members of a large population in the context of weeding out all syphilitic men called up for military service [10]. The test was so sensitive and accurate that Dorfman suggests the following procedure: (1) draw blood from each candidate, (2) use half of each sample to create a pool containing a mixture of  $n$  individual subjects, (3) test the pool. If a pool tested positive, the other half sample of each pool member was individually tested to detect the defective member. In the case of low disease prevalence rates, Dorfman showed that one can save a great amount of time by averaging (pooling) measurements at a first coarse scale. Dorfman procedures use a binary model (B-model) and do not account for false alarms or missed detections, which in our setting is equivalent to an infinite SNR. Therefore they do not require the additional degree of freedom of resource allocation and are only concerned with minimizing the total number of samples required. An optimal group size,  $n$ , can be analytically evaluated for each disease prevalence rate by optimizing the ratio of the expected number of tests using the Dorfman procedure and prevalence rate. Dorfman procedures enjoyed great success due to their simplicity and effectiveness. Pfeifer modifies the binomial model (M-model) and considers test values that were either zero (for negative) or greater than zero (for the degree of contamination) [11]. This way, when a pool is tested positive, each sample of a subgroup from the pool reveals information regarding the other pool members; thus even greater savings are achieved. Nonetheless, the modified model still does not account for false alarms or missed detections. Although M-ARAP uses a similar sampling structure to the Dorfman procedure, we account for the presence of noise, which arises in most signal processing applications, and this constitutes a major difference.

In ARAP [1], we assumed independent and identical distribution (i.i.d.) of targets among pixels and presented a two stage approach for resource allocation under a fixed energy constraint. In this paper we look at detecting and estimating targets under additional sampling constraints. Since targets are assumed to be sparse, a logical approach is to use a coarse measurement scale to reduce the number of samples used at the first stage. The modified version of ARAP (M-ARAP) is introduced to select locations and energy levels with which areas in the scanned domain are re-sampled on a fine grid. The two measurements can be later combined to both detect and estimate the region of interest (ROI) and its content. In [1], we showed ARAP to be asymptotically optimal and the estimation gains were inversely proportional to the signal sparsity. In the modified version, performance continues to depend on sparsity, as well as the extent of the multi-scaling (which is directly related to the number of samples taken) and the inherent contrast of the signal.

The first contribution of this work is the specification of a multi-scale algorithm that adaptively allocates search resources according to the posterior probability of target presence at each pixel and an overall sensing energy constraint. The second contribution of this work is a mathematical analysis of the asymptotic behavior of M-ARAP under the assumption that there is at most one target per region and that the target amplitudes are non-random. This analysis establishes:

convergence to the true ROI, specification of bounds on the expected number of samples and on ‘wasted’ resources due to multi-scaling, specification of a lower bound on the asymptotic gain in estimation in terms of the multi-scale support size, and identification of a ‘detectability index’,  $d$  that can be used to predict performance across contrast level, scale, and SNR. Our simulations show that the predicted gains continue to hold under less restrictive assumptions. In particular, we show that with sufficient ‘detectability’ and a sparsely distributed set of targets (sparsity 0.1%), we improve estimation accuracy by more than 9 [dB] while using less than 20% of the number of measurements used by an exhaustive sampling scheme.

The rest of this paper is organized as follows. In Section II, we introduce our notation. In Section III, we introduce the search problem and define the cost function. Section IV discusses the proposed M-ARAP search method and analyzes its properties. In Section V, we compare the performance of our adaptive multi-scale approach to exhaustive search. In Section VI, we apply M-ARAP to an MTI radar simulation. We conclude and point to future work in Section VII.

## II. NOTATION

The following is a list of notation used throughout the paper:

- $Q$  - Number of cells in search space.
- $L$  - Scale factor between stages, with  $L$  a factor of  $Q$ .
- $N_t$  - # of measurements at stage  $t$ .
- $N$  - # of total measurements, with  $N = \sum_{t=1}^2 N_t$ .
- $\mathcal{X} = \{1, 2, \dots, Q\}$  - Discrete space of  $Q$  cells.
- $\Psi \subseteq \mathcal{X}$  - Subset of  $\mathcal{X}$  referred to as the ROI.
- $I_i$  - indicator function of the ROI such that

$$I_i = \begin{cases} 1, & i \in \Psi \\ 0, & \text{Otherwise} \end{cases} \quad \text{for } i \in \{1, 2, \dots, Q\}. \quad (1)$$

- $H_j$  - Indicator function of the  $j$ -th support region.
- $\hat{\Psi} \subseteq \mathcal{X}$  - Estimated ROI, learned from measurements.
- $\lambda(i, t) \in [0, \lambda_T]$  - Search effort allocated to cell  $i$  at stage  $t$ .
- $\lambda_T = \sum_{t=1}^2 \sum_{i=1}^Q \lambda(i, t)$  - Total search effort.
- $\mathbf{Y} = \{\mathbf{y}(1), \mathbf{y}(2)\} \in \mathbf{R}^N$  - Set of measurements.
- $\tilde{y}_j(1)$  - Measurement of the  $j$ -th support region for  $t = 1$ .
- $\tilde{p}_{H_j|\mathbf{y}(1)}$  - Posterior probability for target existence in  $j$ -th support region.
- $\nu$  - A selectable parameter (see [1] for details).
- $\tilde{w}_j$  - A quantity that is a function of  $\tilde{p}_{H_j|\mathbf{y}(1)}$  and  $\nu$ , used in the allocation of resources to cells at stage 2.
- $\tau(\cdot)$  - A sorting operator for  $\tilde{w}_j$ 's, so that

$$\tilde{w}_{\tau(1)} \leq \tilde{w}_{\tau(2)} \leq \dots \leq \tilde{w}_{\tau(N_1)}.$$

- $u(\cdot)$  - The discrete unit step function.
- $\tilde{\cdot}$  - Tilde notation, used to denote the energy allocation or measurement applied to the support region (as opposed to an individual cell.)

## III. PROBLEM FORMULATION

Consider a space  $\mathcal{X}$  containing  $Q$  cells equipped with a probability measure  $P$ , and a region of interest (ROI),  $\Psi \subseteq \mathcal{X}$ .

In the sequel,  $\Psi$  will be a randomly selected small subset of  $\mathcal{X}$ . Exact definition of the ROI is application dependent. In radar target localization, the ROI is the collection of all cells containing targets and target related phenomena, e.g., target shadows. In a medical imaging application, such as early detection of breast cancer, where tumor boundaries are poorly defined, the ROI may be defined as the collection of all cells containing targets (a tumor) plus some neighboring cells.

Let  $I_i$  be an indicator function defined in (1) and  $\{p_i = \Pr(I_i = 1)\}_{i=1}^Q$  be an associated set of prior probabilities. As in [1] we consider the case of non-informative priors, so that the  $I_i$ 's are i.i.d. and  $p_i = p = \mathbb{E}\{\frac{|\Psi|}{Q}\}$  for all  $i$ . Let  $\mathbf{I}_\Psi = [I_1, \dots, I_Q]^T$  be a vector corresponding to the set of all indicators and  $(\cdot)^T$  denote the transpose operator. We say that the presence of a target affects cell  $i$  if  $i \in \Psi$ . As in [1], targets are assumed static, so  $I_i$  is constant over time. We define the random vector of  $N$  observations,  $\mathbf{Y} : \mathcal{X} \rightarrow \mathbb{R}^N$  and consider the conditional probability  $p(\mathbf{Y}|\mathbf{I}_\Psi)$ .

Consider a sequential experiment where at the first stage we observe  $\mathbf{y}(1) : \mathcal{X} \rightarrow \mathbb{R}^{N_1}$  with  $N_1 < Q$ , and at the second stage we observe a selected subset of  $\mathbf{y}(2) : \mathcal{X} \rightarrow \mathbb{R}^Q$ , i.e., we observe  $y_i(2)$  for  $i \in \hat{\Psi}$ , defined in (2). This formulation allows us to limit the total number of observations to  $N = N_1 + |\hat{\Psi}|$ , where  $|\hat{\Psi}|$  is the number of elements of  $\hat{\Psi}$ . Let  $\lambda(i, t) \geq 0$  denote the search effort allocated to cell  $i$  at time  $t$ , under total energy constraint  $\lambda_T$ . Let  $\{\lambda(i, 2)\}_{i=1}^Q$  be a mapping from past observations  $\mathbf{y}(1)$  to the probability simplex that we call an effort allocation policy. The set  $\hat{\Psi}$  is formally defined as

$$\hat{\Psi} = \{i \in \mathcal{X} : \lambda(i, 2) > 0\}. \quad (2)$$

The combination of  $\{\lambda(i, t)\}$ ,  $N_1$ , and  $\hat{\Psi}$  is termed a search policy. We focus here on deterministic mappings  $\lambda$ , although a more general random mapping could also be incorporated into our framework. We assume that a sample's 'quality' is an increasing function of the allocated effort to the associated cell, e.g. measured in terms of information or inverse variance, called a precision parameter in [7]. In general, effort might be computing power, complexity, cost, or energy that is allocated to acquiring a particular cell location. In [1] we introduced the following cost function

$$J(\lambda) = \sum_{i=1}^Q \frac{\nu I_i + (1 - \nu)(1 - I_i)}{\lambda(i, 1) + \lambda(i, 2)}, \quad (3)$$

with  $\nu \in [\frac{1}{2}, 1]$ . Minimizing (3) subject to a total energy constraint  $\lambda_T$  for the case of  $N_1 = Q$  yielded ARAP, which is summarized below<sup>1</sup>.

*Algorithm 1: Two stage Adaptive Resource Allocation Policy (ARAP)  $\lambda_{ARAP}$*

Step 1: Allocate $\lambda_{ARAP}(i, 1) = \lambda_1^*$ to each cell and measure $\mathbf{y}(1)$ .
Step 2: Given $\mathbf{y}(1)$ compute probabilities, $p_{I_i \mathbf{y}(1)} \triangleq \Pr(I_i = 1 \mathbf{y}(1))$ , and $w_i \triangleq \nu p_{I_i \mathbf{y}(1)} + (1 - \nu)(1 - p_{I_i \mathbf{y}(1)})$ , then rank order the $w_i$ 's.
Step 3: Use $\lambda_1^*$ and the ordered statistics $w_{\tau(i)}$ to find an optimal threshold $k_0$ .
Step 4: Given $k_0$ , apply $\lambda_{ARAP}(i, 2)$ , the energy allocation, to cell $i$ as
$\lambda_{ARAP}(\tau(i), 2) = \lambda(\tau(i), 2) = \left( \frac{\lambda_T - k_0 \lambda_1^*}{\sum_{j=k_0+1}^Q \sqrt{w_{\tau(j)}}} \sqrt{w_{\tau(i)}} - \lambda_1^* \right) u(\tau(i) - k_0), \quad (4)$
and measure $\mathbf{y}(2)$ .

Define  $|\Psi|$  to be the number of elements in  $\Psi$ , and  $\Psi^c \triangleq \mathcal{X} \setminus \Psi$ , the relative complement of  $\Psi$ . As in our previous work, we make the sparsity assumption  $|\Psi| \ll |\Psi^c|$ . In this case, the bounds on the cost function and the gain over an exhaustive search described in [1] are valid. For ARAP the normalized number of observations,  $N^* = \frac{Q + (Q - k_0)}{Q} = 1 + (1 - \frac{k_0}{Q})$ , is a random variable that was shown to converge in probability to  $1 + p$  at high SNR and large  $Q$ , where  $p = \mathbb{E}\{\frac{|\Psi|}{Q}\}$ . Our goal in this work is to add a multi-scale sampling constraint on top of the fixed energy constraint to limit the total search time and the number of measurements. Thus, we modify ARAP as follows: Keep  $N_1 < Q$  and let  $N^*$  be a random variable with bounded mean,  $\mathbb{E}\{N^*\}$ . Note that we do not intend to solve a new optimization problem but rather to modify an existing optimal solution, namely ARAP, in a manner that suits the additional constraints.

#### IV. SEARCH POLICY UNDER TOTAL EFFORT AND MULTI-SCALE SAMPLING CONSTRAINTS

Let  $\Theta = [\Theta_1, \dots, \Theta_Q]^T$  be a random vector where  $\Theta_i \sim \mathcal{N}(\mu_\theta, \sigma_\theta^2)$  are i.i.d. random variables (r.v.) corresponding to the target amplitudes, and  $(\cdot)^T$  denotes vector transpose. Let  $\mathbf{I}_\Psi$  be a vector of indicators marking whether or not cell  $i$  contains a target. Consider the following measurement model:

$$\tilde{\mathbf{y}}(t) = \mathbf{H}(t) \text{diag} \left\{ \sqrt{\tilde{\lambda}(t)} \right\} \text{diag} \{ \Theta \} \mathbf{I}_\Psi + \mathbf{n}(t), \quad (5)$$

where  $\mathbf{H}(t)$  is an  $(N_t \times Q)$  matrix describing the "beamforming" measurement operator,  $\tilde{\lambda}(t)$  is a vector describing resource allocation at time  $t$ ,  $[\sqrt{\tilde{\lambda}}]_i$  denotes  $\sqrt{\tilde{\lambda}_i}$ ,  $\text{diag} \{ \mathbf{x} \}$  is a square matrix with  $[\text{diag} \{ \mathbf{x} \}]_{ii} = x_i$  and  $[\text{diag} \{ \mathbf{x} \}]_{ij} = 0$  for  $i \neq j$ , and  $\mathbf{n}(t) \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_{N_t \times N_t})$  where  $\mathbf{I}_{N_t \times N_t}$  is an  $(N_t \times N_t)$  identity matrix. The beamforming operator  $\mathbf{H}(t)$  simply forms linear combinations of neighboring pixels and is what distinguishes M-ARAP from ARAP. In our model (5), both  $\mathbf{H}(t)$  and  $\tilde{\lambda}(t)$  are design parameters that satisfy the user-defined constraints. We propose the following simple design.

<sup>1</sup> $\lambda_1^* \leq \lambda_T/Q$  and can be selected, based on criteria specified in [1].

Define  $L$  to be a factor of  $Q$ , and define the column vectors  $\mathbf{g}_L, \mathbf{0}_L \in \mathbb{R}^L$ :

$$\mathbf{g}_L = L \text{ times } \left\{ \begin{bmatrix} \frac{1}{L} \\ \vdots \\ \frac{1}{L} \end{bmatrix} \right\}, \quad \mathbf{0}_L = L \text{ times } \left\{ \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \right\} \quad (6)$$

Then, we consider the beamforming matrix for the first stage measurements (defined through the transpose)

$$\mathbf{H}(1) = \underbrace{\begin{bmatrix} \mathbf{g}_L & \mathbf{0}_L & \mathbf{0}_L & \cdots & \mathbf{0}_L \\ \mathbf{0}_L & \mathbf{g}_L & \mathbf{0}_L & \cdots & \mathbf{0}_L \\ \mathbf{0}_L & \mathbf{0}_L & \mathbf{g}_L & \cdots & \mathbf{0}_L \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_L & \mathbf{0}_L & \mathbf{0}_L & \cdots & \mathbf{g}_L \end{bmatrix}}_{\frac{Q}{L} \text{ times}} \in \mathbf{R}^{(Q/L) \times Q} \quad (7)$$

Hence, an element-wise version of the first stage measurement model (5) is given by

$$\tilde{y}_j(1) = \frac{\sqrt{\tilde{\lambda}_1}}{L} \sum_{i=(j-1)L+1}^{jL} \theta_i I_i + n_j(1), \quad (8)$$

$$j = 1, 2, \dots, N_1,$$

where  $\tilde{y}_j(1)$  is the  $j$ -th element of  $\tilde{\mathbf{y}}(1)$ . Let  $\mathcal{X}_j = \{(j-1)L+1, \dots, jL-1, jL\}$  denote the support of the  $j$ -th row of  $\mathbf{H}(t)$ , and note that  $|\mathcal{X}_j| = L$  for all  $j$ . With small  $p$ , large  $Q$ , and  $L \ll Q$  the probability that  $\mathcal{X}_j$  contains more than one target is negligible. Let  $H_j$  denote the indicator function of the event, “ $I_i = 1$  for some  $i \in \mathcal{X}_j$ .” Then we know that

$$\tilde{y}_j(1) | (H_j = 0) \sim \mathcal{N}(0, \sigma^2), \quad (9)$$

$$\tilde{y}_j(1) | (H_j = 1) \sim \mathcal{N}\left(\frac{\sqrt{\tilde{\lambda}_1}}{L} \mu_\theta, \sigma^2 + \frac{\tilde{\lambda}_1}{L^2} \sigma_\theta^2\right)$$

As in [1], the knowledge of these two distributions will be used to focus measurements at the second stage based on the received measurements at the first stage, using posterior probabilities derived from the first stage measurements. Second stage measurements are defined with  $\mathbf{H}(2) = \mathbf{I}_{Q \times Q}$ , which gives the element-wise version

$$y_j(2) = \sqrt{\lambda(j, 2)} \theta_j I_j + n_j(2), \quad j = 1, 2, \dots, Q. \quad (10)$$

#### A. Multi-scale version of ARAP

We define the following resource allocation policy based on ARAP:

*Algorithm 2: Two stage Multi-scale Adaptive Resource Allocation Policy (M-ARAP)  $\lambda_M$*

Step 1: Allocate  $\tilde{\lambda}_{M-ARAP}(j, 1) = \tilde{\lambda}_1^*$  to each support  $\mathcal{X}_j$  and measure  $\tilde{\mathbf{y}}(1)$  in (5).

Step 2: Compute probabilities  $\tilde{p}_{H_j|\tilde{\mathbf{y}}(1)} \triangleq \Pr(H_j = 1 | \tilde{\mathbf{y}}(1))$  and  $\tilde{w}_j \triangleq \nu \tilde{p}_{H_j|\tilde{\mathbf{y}}(1)} + (1 - \nu)(1 - \tilde{p}_{H_j|\tilde{\mathbf{y}}(1)})$  over each support region.

Step 3: Rank order the  $\tilde{w}_j$ 's using (13), then use  $\tilde{\lambda}_1^*$  and the ordered statistic  $\tilde{w}_{\tau(j)}$  to find a threshold  $\tilde{k}_0$  via (14) and (15).

Step 4: Given  $\tilde{k}_0$ , define the energy allocation to support region  $j$  as

$$\tilde{\lambda}(\tau(j), 2) = \left( \frac{\lambda_T - \tilde{k}_0 \tilde{\lambda}_1^*}{\sum_{l=\tilde{k}_0+1}^Q \sqrt{\tilde{w}_{\tau(l)}}} \sqrt{\tilde{w}_{\tau(j)}} - \tilde{\lambda}_1^* \right) u(\tau(j) - \tilde{k}_0), \quad (11)$$

Step 5: Define the energy allocation to cell  $i$  in support region  $j$  as

$$\lambda_M(i, 2) = \frac{\tilde{\lambda}(j, 2)}{L} \quad (12)$$

for  $j = 1, 2, \dots, N_1$ , and measure  $\mathbf{y}(2)$  using  $\mathbf{H}(2) = \mathbf{I}_{Q \times Q}$  and  $[\boldsymbol{\lambda}(2)]_i = \lambda_M(i, 2)$ .

To complete the definition of M-ARAP, define the permutation operator  $\tau$  corresponding to the rank ordering of the  $\tilde{w}_j$ 's as

$$\tau(j) = \arg \min_{i=1, \dots, N_1} \{\tilde{w}_i : \tilde{w}_i \geq \tilde{w}_{\tau(j-1)}\}, \quad j \in \{1, 2, \dots, N_1\}, \quad (13)$$

with  $\tilde{w}_{\tau(0)} \triangleq 0$ . Whenever the r.h.s. of (13) is not unique, we select an arbitrary  $i$  satisfying  $\tilde{w}_{\tau(1)} \leq \tilde{w}_{\tau(2)} \leq \dots \leq \tilde{w}_{\tau(N_1)}$ . Then, assuming  $\tilde{w}_{\tau(1)} > 0$ , define  $\tilde{k}_0$ , the threshold parameter, as  $\tilde{k}_0 = 0$  if

$$\frac{\lambda_T}{\tilde{\lambda}_1^*} > \frac{\sum_{j=1}^{N_1} \sqrt{\tilde{w}_{\tau(j)}}}{\sqrt{\tilde{w}_{\tau(1)}}}, \quad (14)$$

otherwise  $\tilde{k}_0 \in \{1, \dots, N_1 - 1\}$  is the integer satisfying

$$\frac{\sum_{j=\tilde{k}_0+1}^{N_1} \sqrt{\tilde{w}_{\tau(j)}}}{\sqrt{\tilde{w}_{\tau(\tilde{k}_0+1)}}} < \frac{\lambda_T}{\tilde{\lambda}_1^*} - \tilde{k}_0 \leq \frac{\sum_{j=\tilde{k}_0+1}^{N_1} \sqrt{\tilde{w}_{\tau(j)}}}{\sqrt{\tilde{w}_{\tau(\tilde{k}_0)}}} \quad (15)$$

A proof of the existence and uniqueness of  $\tilde{k}_0$  in (15), as well as a discussion of its properties, is given in [1]. Note that the definition of  $\tilde{k}_0$  is identical to ARAP for a search space of size  $N_1 = Q/L$ . For  $N_1 = Q$ , M-ARAP is completely equivalent to ARAP, provided that  $\tilde{\lambda}_1^*$  is correctly defined. To define  $\tilde{\lambda}_1^*$ , let

$$\tilde{\lambda}_1^* = \arg \min_{\tilde{\lambda}_1 \in (0, \frac{\lambda_T}{N_1})} \mathbb{E} \left\{ \sum_{i=1}^Q \frac{\nu I_i + (1 - \nu)(1 - I_i)}{\frac{\tilde{\lambda}_1}{L} + \lambda(i, 2)} \right\}, \quad (16)$$

where  $\lambda(i, 2)$  is defined in (12) substituting  $\tilde{\lambda}_1$  for  $\tilde{\lambda}_1^*$ . Note that  $\lambda(i, 2)$  depends on  $\tilde{p}_{H_j|\tilde{\mathbf{y}}(1)}$ , which, in turns depends

on the distribution of target amplitudes,  $\Theta$ . To analyze the performance of M-ARAP, we first establish properties of  $\tilde{p}_{H_j|\tilde{y}(1)}$ .

**B. Detectability index and asymptotic properties of  $\tilde{p}_{H_j|\tilde{y}(1)}$  when  $\nu = 1$**

Let  $|\mathcal{X}_j| = L$  denote an observed support size for the first stage in M-ARAP. Let the true mean sparsity of the observed signal be  $p = \frac{E\{|\Psi|\}}{Q}$ . Define the *detectability index* as

$$d = \frac{|E\{\tilde{y}_j(1)|H_j = 1\} - E\{\tilde{y}_j(1)|H_j = 0\}|}{\sqrt{\text{var}(\tilde{y}_j(1)|H_j = 0)}}. \quad (17)$$

Substituting (9) into (17) yields

$$d = \frac{\mu_\theta}{L} \sqrt{\frac{\tilde{\lambda}_1}{\sigma^2}}, \quad (18)$$

which is proportional to  $\mu_\theta$  and to the square root of effective SNR  $\frac{\tilde{\lambda}_1}{L\sigma^2}$ . Therefore, for a specified false alarm rate, we expect the power of the likelihood ratio test (LRT) performed on  $\mathbf{y}(1)$  to increase as either the inherent contrast  $\mu_\theta$  or effective SNR increase. In this section, we analyze asymptotic properties of M-ARAP, where by asymptotic we mean high SNR and large  $Q$ . We further assume that  $\mathcal{X}_j$  contains at most one target, and that the target amplitude variance,  $\sigma_\theta^2 = 0$ . Under these assumptions, we can establish Claim 1, Claim 2, and Claim 3 given below.

*Claim 1:*  $\tilde{p}_{H_j|\tilde{y}(1)} \rightarrow H_j$  in probability as  $\text{SNR} \rightarrow \infty$ .

*Derivation of Claim 1:* Under the assumptions that  $\mathcal{X}_j$  contains at most one target and  $\sigma_\theta^2 = 0$ , the first stage measurement reduces to

$$\tilde{y}_j(1) = \frac{\sqrt{\tilde{\lambda}_1} H_j \mu_\theta}{L} + n_j(1) = \sqrt{\gamma_1} H_j \mu_\theta + n_j(1) \quad (19)$$

where  $\gamma_1 = \tilde{\lambda}_1/L^2$ . In [1] we proved that  $\tilde{p}_{H_j|\mathbf{Y}^{\tilde{y}(1)}} \rightarrow H_j$  in probability as  $\gamma_1 \rightarrow \infty$ . Thus, Claim 1 follows directly from this result, noting that  $\text{SNR} \rightarrow \infty$  implies  $\gamma_1 \rightarrow \infty$ .

In [1] we used the asymptotic consistency property of  $\tilde{p}_{H_j|\mathbf{Y}^{\tilde{y}(1)}}$  to show that the threshold parameter  $\tilde{k}_0$  converges to the true number of empty search cells  $(1-p)Q$  of the scanned domain (recall that  $|\hat{\Psi}| = Q - \tilde{k}_0$ ). For  $\lambda_M$  we can provide an asymptotic bound on  $\tilde{k}_0$  that holds in probability. The logic behind the bound is that if  $p$  can be used to bound  $K$ , the number of support regions  $\mathcal{X}_j$  that contain targets, then  $\tilde{k}_0 \geq Q - KL$ . Therefore, we have the following:

*Claim 2:* The normalized number of samples  $N^*$  used by M-ARAP, defined as

$$N^* = \frac{N_1 + |\hat{\Psi}|}{Q}, \quad (20)$$

is bounded between  $\frac{1}{L}$  and  $\frac{1}{L} + p^*L$ , where  $p^*$  is the true sparsity of the underlying domain, in the sense that

$$\lim_{\text{SNR} \rightarrow \infty} \Pr \left( \frac{1}{L} \leq N^* \leq \frac{1}{L} + p^*L \right) = 1, \quad (21)$$

for sufficiently large  $Q$ .

*Derivation of Claim 2:* To prove (21) note first that

$$N^* = \frac{N_1 + |\hat{\Psi}|}{Q} = \frac{1}{L} + \frac{|\hat{\Psi}|}{Q} \geq \frac{1}{L}. \quad (22)$$

It suffices to show

$$\lim_{\text{SNR} \rightarrow \infty} \Pr \left( \frac{|\hat{\Psi}|}{Q} \leq pL \right) = 1. \quad (23)$$

In [1], it was shown that if  $p_{I_i|\mathbf{y}(1)} \rightarrow I_i$  in probability as  $\text{SNR} \rightarrow \infty$ , then

$$\tilde{k}_0 \rightarrow Z \sim \text{Bin}(S, 1-r) \quad (24)$$

in probability as  $\text{SNR} \rightarrow \infty$  for a search space of size  $S$  and sparsity  $r$ . Recalling that  $\tilde{k}_0$  is calculated using ARAP for  $S = Q/L$  and  $r = pL$ , it follows that  $\tilde{k}_0 \rightarrow Z \sim \text{Bin}(Q/L, 1-pL)$  as  $\text{SNR} \rightarrow \infty$ . Using the definition of  $\hat{\Psi}$  from (2), we have

$$\begin{aligned} |\hat{\Psi}| &= |\{i \in \mathcal{X} : \lambda(i, 2) > 0\}| \\ &= \left| \bigcup_{j=1}^{Q/L} \{i \in \mathcal{X}_j : H_j = 1\} \right| \quad (\text{as } \text{SNR} \rightarrow \infty) \\ &= \sum_{j=1}^{Q/L} |\{i \in \mathcal{X}_j : H_j = 1\}| \quad (\text{for disjoint } \mathcal{X}_j) \\ &= \sum_{j=1}^{Q/L} L H_j = L \sum_{j=1}^{Q/L} H_j = L(Q/L - \tilde{k}_0) \end{aligned} \quad (25)$$

Since,  $\tilde{k}_0 \sim \text{Bin}(Q/L, 1-pL)$  as  $\text{SNR} \rightarrow \infty$ , we have  $E\left[\frac{|\hat{\Psi}|}{Q}\right] \rightarrow pL$  and  $\text{var}\left(\frac{|\hat{\Psi}|}{Q}\right) \rightarrow \frac{pL(1-pL)}{Q}$  as  $\text{SNR} \rightarrow \infty$ . Moreover,  $\text{var}\left(\frac{|\hat{\Psi}|}{Q}\right) \rightarrow 0$  as  $Q \rightarrow \infty$ . Thus,  $\frac{|\hat{\Psi}|}{Q} \rightarrow pL$  in the mean square sense. This establishes Claim 2.

*Claim 3:* The expected proportion of the scanned area that is empty of targets at the second stage is bounded with probability, in the sense that

$$\Pr \left( \frac{|\Psi \Delta \hat{\Psi}|}{Q} \leq p(L-1) \right) = 1. \quad (26)$$

*Derivation of Claim 3:* This result is established similarly to Claim 2. In particular, note that when  $\text{SNR} \rightarrow \infty$

$$\begin{aligned} |\Psi \Delta \hat{\Psi}| &= \left| \bigcup_{j=1}^{Q/L} \{i \in \mathcal{X}_j, i \notin \Psi : H_j = 1\} \right| \\ &= \sum_{j=1}^{Q/L} |\{i \in \mathcal{X}_j : H_j = 1\} \Delta \{i \in \mathcal{X}_j, i \in \Psi : H_j = 1\}| \\ &= \sum_{j=1}^{Q/L} |\{i \in \mathcal{X}_j : H_j = 1\}| - |\{i \in \mathcal{X}_j, i \in \Psi : H_j = 1\}| \\ &= \sum_{j=1}^{Q/L} (L-1)H_j = (L-1)(Q/L - \tilde{k}_0) \end{aligned} \quad (27)$$

This leads to  $E\left[\frac{|\Psi \Delta \hat{\Psi}|}{Q}\right] = p(L-1)$  and  $\text{var}\left(\frac{|\Psi \Delta \hat{\Psi}|}{Q}\right) = \frac{(L-1)^2}{LQ} p(1-pL) \rightarrow 0$  as  $\text{SNR}, Q \rightarrow \infty$ . Thus,  $\frac{|\Psi \Delta \hat{\Psi}|}{Q} \rightarrow$

$p(L-1)$  in the mean square sense (which implies convergence in probability as required). This establishes Claim 3.

Claim 3 provides a bound that can be used to evaluate the tradeoff between reducing the number of measurements and increasing the expected estimation gains. Specifically, use the cost function (3) and assume that asymptotically we learn  $\widehat{\Psi}$  at almost no cost. Then an optimal allocation policy is given by

$$\lambda_{N_o}(i, 2) = \begin{cases} \frac{\lambda_T}{|\widehat{\Psi}|}, & i \in \widehat{\Psi} \\ 0, & \text{otherwise} \end{cases}. \quad (28)$$

Hence,  $\lambda_{N_o}(i) = \frac{\lambda_T}{|\widehat{\Psi}|} I_{i \in \widehat{\Psi}}$  and with  $\nu = 1$  the expected cost is

$$J(\lambda_{N_o}) = \sum_{i \in \widehat{\Psi}} \frac{1}{\lambda_T / |\widehat{\Psi}|} = \frac{|\widehat{\Psi}| |\Psi|}{\lambda_T}. \quad (29)$$

At the same time, the cost associated with an exhaustive search policy  $\lambda_U(i) = \frac{\lambda_T}{Q}$  is given by  $J(\lambda_U) = \frac{Q|\Psi|}{\lambda_T}$ . Define the gain function

$$G(\lambda) = 10 \log \frac{J(\lambda_U)}{J(\lambda)}, \quad (30)$$

then

$$G(\lambda_{N_o}) = 10 \log \frac{Q}{|\widehat{\Psi}|} = 10 \log \frac{Q}{|\Psi| + |\Psi \Delta \widehat{\Psi}|}. \quad (31)$$

Using Claim 3 we obtain

$$G(\lambda_{N_o}) \geq 10 \log \frac{Q}{pQ(1+L-1)} = 10 \log \frac{1}{p} - 10 \log L, \quad (32)$$

where  $10 \log L$  is the gain penalty that we pay due to multi-scale search.

### C. Discussion of performance for clustered targets

The assumption that only a single target may appear in a support region may not hold when apparent targets clump together. This scenario arises in radar target detection where large scattering objects may occupy consecutive pixels on the radar screen and appear as a cluster of targets. Nevertheless, the overall area occupied by targets is small compared to the area a scanning radar system covers. As another example, in early detection of cancer tumors such as in breast cancer, the diameter of a typical tumor is a few millimeters to 1.5 centimeters. Hence, on a fine grid a tumor may appear as a cluster of targets, yet, its overall volume is very small compared to the volume of the entire breast.

In these cases we speculate that the performance of M-ARAP with clustered targets is lower bounded by the performance in the case of single-pixel targets. Indeed, if we let  $\Xi_j$  be the total number of targets in support  $j$ , with  $E|\Xi_j| = E|\Xi| \geq 1$ , and Claim 1 still holds, then it can be shown that

$$G(\lambda_{N_o}) \geq 10 \log \frac{1}{p} - 10 \log(1+L-E|\Xi|), \quad (33)$$

where  $10 \log(1+L-E|\Xi|) \leq 10 \log L$  is the gain penalty that we pay due to multi-scaling. In other words, having clustered targets within a support region tends to reduce the gain penalty due to multi-scale processing.

Despite the limitations of the single deterministic target-per-region assumption in Section IV-B to obtain asymptotic limits, we believe that M-ARAP's predicted performance gains will hold under broader conditions. This belief is supported by our numerical results shown in the next section. Note that the gains established in Section IV-B require only that the posterior probabilities,  $\tilde{p}_{H_j|\mathbf{Y}^{(1)}}$  converge to the indicator function  $H_j$ , which may not require the assumption of a single target per multi-resolution cell.

## V. PERFORMANCE COMPARISONS

### A. Estimation

Assume that cell  $l$  belongs to the ROI, i.e.,  $I_l = 1$ . Here we introduce an estimate of target amplitude  $\theta_l$  using the measurement pair  $(\tilde{\mathbf{y}}(1), \mathbf{y}(2))$ . Assuming these amplitudes are independent and obey a Gaussian prior distribution,  $\theta_l \sim \mathcal{N}(\mu_\theta, \sigma_\theta^2)$ , we can derive the conditional mean estimator (CME)  $\theta_l = E[\theta_l | \tilde{\mathbf{y}}(1), \mathbf{y}(2)]$ , which is the minimum mean squared error (MSE) estimator. As a baseline, we will compare this estimator to the CME  $E\{\theta_l | \mathbf{y}(0)\}$  under an exhaustive search policy, where

$$y_i(0) = \sqrt{\lambda_0} \theta_i I_i + n_i(0), \quad n_i(0) \sim \mathcal{N}(0, \sigma^2) \quad (34)$$

and  $\lambda_0 = \frac{\lambda_T}{Q}$ . The MSE of the CME for an exhaustive search policy is given by

$$\text{var}\{\theta_l | y_l(0)\} = \sigma_\theta^2 - \frac{\lambda_0 \sigma_\theta^4}{\sigma^2 + \lambda_0 \sigma_\theta^2} = \frac{\sigma_\theta^2}{1 + \lambda_0 \frac{\sigma_\theta^2}{\sigma^2}}. \quad (35)$$

Recall from equation (9) that  $\tilde{y}_j(1) | H_j = r$  is Gaussian for  $r = 0, 1$ . Thus, we know for  $I_l = 1$  and  $l \in \mathcal{X}_j$  that

$$\tilde{y}_j(1) | I_l = 1 \sim \mathcal{N}\left(\frac{\sqrt{\tilde{\lambda}_1}}{L} \mu_\theta, \sigma^2 + \frac{\tilde{\lambda}_1}{L^2} \sigma_\theta^2\right). \quad (36)$$

The conditional distribution of  $y_l(2)$  given  $\tilde{\mathbf{y}}(1)$  is also Gaussian and defined as

$$y_l(2) | I_l = 1, \tilde{\mathbf{y}}(1) \sim \mathcal{N}\left(\sqrt{\lambda_M(i, 2)} \mu_\theta, \sigma^2 + \lambda_M(i, 2) \sigma_\theta^2\right). \quad (37)$$

The Naive Bayes approximation [12] to  $E\{\theta_l | \mathbf{y}(1), y_l(2)\}$  is derived under the assumption that  $(\mathbf{y}(1), y_l(2))$  defined in (36) and (37), respectively, are independent. This is suboptimal but serves as a good comparison benchmark. Let  $l \in \mathcal{X}_j$ ,  $\mathbf{v}_l = [\tilde{y}_j(1) \ y_l(2)]^T$ , then the Naive Bayes estimator is given by

$$\begin{aligned} \hat{\theta}_l &\triangleq E\{\theta_l | \mathbf{v}_l, I_l = 1\} \\ &= \mu_\theta + \text{cov}(\theta_l, \mathbf{v}_l)^T \text{cov}^{-1}(\mathbf{v}_l) (\mathbf{v}_l - E\{\mathbf{v}_l\}). \end{aligned} \quad (38)$$

where

$$\text{cov}(\theta_l, \mathbf{v}_l) = \sigma_\theta^2 \left[ \frac{\sqrt{\tilde{\lambda}_1 \lambda_M(l, 2)}}{L} \right]^T, \quad (39)$$

$$\text{cov}(\mathbf{v}_l) = \begin{bmatrix} \sigma^2 + \frac{\tilde{\lambda}_1}{L^2} \sigma_\theta^2 & \sigma_\theta^2 \frac{\sqrt{\tilde{\lambda}_1 \lambda_M(l, 2)}}{L} \\ \sigma_\theta^2 \frac{\sqrt{\tilde{\lambda}_1 \lambda_M(l, 2)}}{L} & \sigma^2 + \lambda_M(l, 2) \sigma_\theta^2 \end{bmatrix}, \quad (40)$$

$$E\{\mathbf{v}_l\} = \mu_\theta \left[ \frac{\sqrt{\tilde{\lambda}_1 \lambda_M(l, 2)}}{L} \right]^T. \quad (41)$$

Simulation results are shown in Figs. 1-2. In both figures we plot the MSE performance gain  $g(\lambda)$ , defined as

$$g(\lambda) = 10 \log \frac{\text{var}(\theta_l | y_l(0))}{\text{MSE}(\hat{\theta}_l)} \quad (42)$$

as a function of SNR (Fig. 1) and the detectability index given in (18) (Fig. 2). Monte-Carlo simulations were used to estimate the MSE of (38). We chose  $Q = 12,000$ ,  $p = \frac{1}{1000}$ , and each point on the figure represents the median MSE based on 500 realizations. We let signal to noise ratio, defined as  $10 \log \frac{\lambda_T/Q}{\sigma^2}$ , vary from 0 to 40 [dB], used contrast level  $\mu_\theta \in \{2, 4, 8\}$ , and set  $\sigma_\theta^2 = \frac{1}{16}$ . Different lengths  $L$  were simulated for the first stage, but we present here the cases of  $L = 8$  and  $L = 32$  since it is enough to understand the general trends. Curves with different markers and colors represent different contrast levels  $\mu_\theta$ .

Note that, in contrast to ARAP, we do not claim optimality of the naive Bayes amplitude estimator that we proposed for M-ARAP. Indeed the optimal gain of 30 [dB] is not attained. Moreover, asymptotic gains decrease as  $L$  increases. This is natural since the posterior probabilities  $\tilde{p}_{H_j | \tilde{\mathbf{y}}(1)}$  are identical within each support. Hence, if the resource allocation scheme  $\lambda_M$  suspects that a target exists in  $\mathcal{X}_j$ , all cells within this support receive the same effort allocation for the second stage. As  $L$  increases, this translates to wasted resources according to *Claim 3*. Fig. 1 shows asymptotic gain of 21 [dB] for  $L = 8$  and 15 [dB] for  $L = 32$ , both agreeing with (32).

In Fig. 2 we plot estimation gains vs. the detectability index since it incorporates both the contrast level and the scale in a single parameter. We see that the detectability index is a reasonable predictor of the performance of M-ARAP across contrast levels and scales. In the transition region between no gain (i.e., index values less than 2.5) and asymptotic gain (index values greater than 10), we see some discrepancies among the curves. These discrepancies are greatest when  $L$  is smallest or  $\mu_\theta$  is largest. This corresponds to situations where the assumption of independent measurements in the naive Bayes estimator breaks down. It is hypothesized that with a better estimator that accounts for dependence between  $\mathbf{Y}(1)$  and  $\mathbf{Y}(2)$ , the aforementioned discrepancies would be less apparent.

### B. Normalized number of samples, $N^*$

The normalized number of samples  $N^*$  is lower bounded by  $\frac{1}{L}$ , and hence there is a tradeoff between estimation gain and reduction in number of measurements. We show here that if the detectability index is sufficiently high ( $d > 5$  in our case) we can both save measurements (according to *Claim 2*) and enjoy significant estimation gain within the ROI. Fig. 3 shows the expected saving in measurements or  $N^*$  for the scenario depicted in Fig. 2. Black curves represent  $L = 8$ , blue curves represent  $L = 16$ , and purple curves represent  $L = 32$ . Combining the information on both figures shows measurement saving relative to estimation gain. For example,

for  $d = 5$ , M-ARAP with  $L = 8$  yields about 10 [dB] performance gain in estimation while using only 18% of the number of samples used by an exhaustive search. Similar performance gain is achieved by M-ARAP for  $L = 32$  with  $d = 5$  and about 8% of the samples. Note that in all three cases,  $N^*$  converges to the upper bound in *Claim 2* (0.133, 0.0785, and 0.06325 for  $L = 8, 16, 32$  respectively).

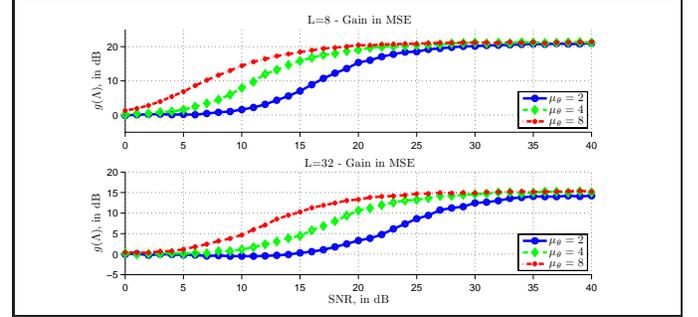


Fig. 1. We plot estimation gains as a function of SNR for different contrast levels. The upper plot show gains for  $L = 8$  while the lower plot show gains for  $L = 32$ . In the upper plot, significant gains of 10 [dB] are achieved for all contrasts at SNR values less than 17 [dB]. In the lower plot, 10 [dB] gains occur at high contrasts at SNR less than 20 [dB]. Note that the asymptotic lower bound on the gain (32) yields 21.0 [dB] and 15.0 [dB] for  $L = 8$  and  $L = 32$  respectively, which agree well with the gains in these plots.

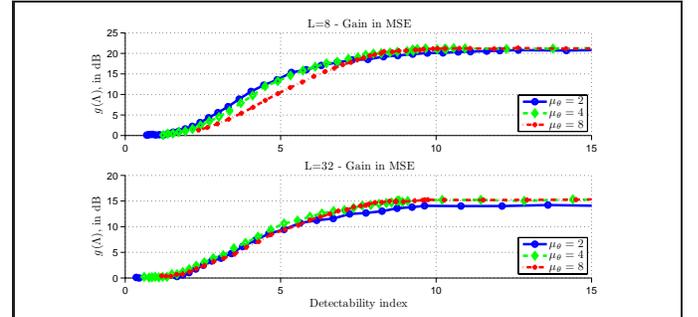


Fig. 2. Estimation gains (in MSE) are plotted against detectability index for  $L = 8$  and  $L = 32$ . Note that the detectability index can be used as a reasonable predictor of MSE gain, regardless of the actual contrast, SNR, or scale.

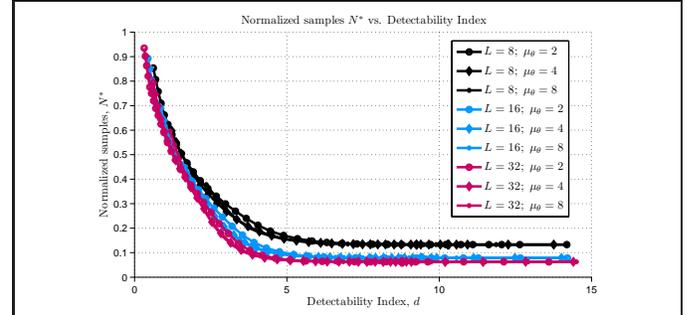


Fig. 3. We plot the normalized number of samples  $N^*$  as a function of detectability index for  $L = 8, 16, 32$ , and different contrast levels  $\mu_\theta \in \{2, 4, 8\}$ . These  $N^*$  values are associated with estimation gains seen in Fig. 2. For example for a relatively low detectability index of  $d = 5$  and  $L = 8$ , estimation performance gain of 10 [dB] is achieved with less than 18% of the sampling used by exhaustive search. Similar gains are achieved for  $d = 5$ ,  $L = 32$ , and less than 8% of the samples.

Our previous experience with ARAP and M-ARAP [13]

in the context of detection tasks suggest that optimizing the energy allocation between the two stages is difficult and very much application dependent. This will negatively affect performance if noise variance  $\sigma^2$  is unknown and must be estimated from the data. Since M-ARAP is a coarser version of ARAP, we present detection performance only in the context of the moving target indicator example given in Section VI-A, where we compute false discovery rate test for target detections using M-ARAP as compared to exhaustive search (see Fig. 8).

### C. Computational complexity comparison

The computation of the search policy in M-ARAP requires computing  $N_1 = Q(1/L)$  posterior probabilities ( $\mathcal{O}(N_1)$ ), sorting the probabilities ( $\mathcal{O}(\log N_1)$  operation), and computation of the second stage allocations ( $\mathcal{O}(|\hat{\Psi}|)$ ). Thus, the complexity of M-ARAP is

$$\mathcal{C}_{\text{M-ARAP}} = \mathcal{O}(N_1) + \mathcal{O}(\log N_1) + \mathcal{O}(|\hat{\Psi}|) = \mathcal{O}(N) \quad (43)$$

where  $N$  is the number of total measurements used by M-ARAP. Assuming that the asymptotic properties of the previous section hold,  $N \rightarrow Q(1/L + pL)$  so that the complexity of M-ARAP is  $\mathcal{O}(Q(1/L + pL))$ . Note that ARAP is just a special case of M-ARAP for  $L=1$ . Moreover, the computational complexity of the Abdel-Samad and Tewfik algorithm (henceforth referred to as AS-T) is given by the authors as  $\mathcal{O}(tN^2)$ , where  $t = \log_2(Q)$  is the number of stages in their hierarchical binary search [14]. Thus, the ratio of computational complexity between M-ARAP and AS-T is  $\mathcal{O}(N \log_2 Q)$ . Table I shows the dB gain in number of measurements between the AS-T algorithm and M-ARAP for  $p = 0.01$  and various values of  $Q$ , and  $L$ . This comparison highlights the computational burden of AS-T, which requires significantly more computations than M-ARAP in all cases studied.

Distilled sensing only requires calculation of an allocation once for each stage in its implementation. The authors state that the number of stages should be chosen to be  $K = 1 + \left\lceil \frac{\log_2 \log_2 Q}{\log_2(2-\Delta)} \right\rceil$  [7], where we choose  $\Delta = 0.9$  for our comparison. Then the ratio of computational complexity to M-ARAP can be calculated as  $\mathcal{O}(K/N)$ . However, it should be noted that although DS has lower computational complexity, the number of measurements is generally larger than M-ARAP. Under the assumption that exactly one half of the cells with  $I_i = 0$  are removed at each stage of DS, the expected number of measurements can be derived as

$$E[N_{DS}] = pQK + 2(1-p)Q(1-2^{-K}) \quad (44)$$

We set  $Q = 12000$ ,  $p = 0.001$  and compare DS to M-ARAP for  $L = 1, 8, 32$  over a range of SNR. Figure 4(a) plots the loss in computational complexity between M-ARAP and DS. In the studied case, DS requires 26 dB, 17 dB, and 14 dB fewer computations asymptotically than M-ARAP for  $L = 1, 8, 32$ , respectively. Figure 4(b) plots the gain in the cost function over exhaustive search as a function of SNR for all four algorithms. It is seen that DS outperforms M-ARAP for low SNR values, but its asymptotic performance is significantly lower (which is to be expected, considering

TABLE I  
COMPUTATIONAL COMPLEXITY COMPARISON BETWEEN M-ARAP AND AS-T FOR  $m=2$  IN DB

Q	ARAP (M-ARAP, L=1)	M-ARAP L=8	M-ARAP L=32
128	30.0	23.1	20.8
1024	40.5	33.6	31.3
8192	50.7	43.8	41.5

that ARAP optimizes this cost function). Figure 4(c) plots the gain as a function of the detectability index, and shows that for  $d > 5$ , M-ARAP outperforms DS at all given scales. Finally, Figure 4(d) plots the percentage of measurements used by M-ARAP compared to DS as a function of detectability index. In (a) and (d), yellow markers indicate the points on the curve where the performance of DS is approximately equal to that of M-ARAP. It is seen that M-ARAP saves considerably on measurements at all observed scales.

## VI. APPLICATION: MOVING TARGET INDICATION/DETECTOR

Moving target indication (MTI) radars provide the capability to detect targets reflections having differential radial motion with respect to an interfering background (called clutter) that might typically consist of reflections from terrain, sea, weather, or chaff. A typical application of such radar is surveillance, e.g., to detect low-flying aircraft moving over terrain through possible weather disturbances. The function of the MTI radar is to reject returns from terrain and weather so to maintain adequate target detection [15]. In many cases MTI radar suppresses clutter by more than 20 [dB].

In the following section we use a simplified MTI simulation to illustrate potential benefits yielded by M-ARAP. A full and realistic emulation of MTI is outside the scope of this paper. We simulate a field of view (FOV) about 66 [km<sup>2</sup>] with pixel dimensions of  $20 \times 20$  [m<sup>2</sup>] and radar resolution cell of  $100 \times 100 \times 150$  [m<sup>3</sup>]. A sparsity level of  $p = 0.0007$  was selected and  $Q = 408^2$ . We chose identical targets with target reflection coefficient (per pixel) representing an aircraft similar to an Airbus A-320. An operating point of 0.1 was selected for the Radar Cross Section (RCS) and this parameter was varied in simulation from 0.01 to 1. Target velocities were normally distributed with  $v_t \sim \mathcal{N}(200, 49)$  centered at 200 [m/s]. Rain intensity was random between 0-6 [mm/hr] and spatial correlation on the order of 1 [km<sup>2</sup>]. Maximal clutter velocity was 30 [m/s] and targets were randomly placed within the clutter regions. A simple second order FIR line canceler was implemented (frequency response of  $\sin^2 \frac{\omega T}{2}$ ) with the pass band centered at 200 [m/s]. The Swerling II noise model (Exponential) was used as a measurement noise model and the total energy budget was  $\lambda_T = NQ$  pulses, i.e.,  $N$  pulses transmitted at each grid stop.

Note that this model violates the assumption of single-pixel targets that we have used for the performance prediction analysis. However, as hypothesized earlier, we believe that the performance with clustered targets will not be negatively impacted as compared to the single-target scenario. Indeed, for the case where target returns add constructively to the

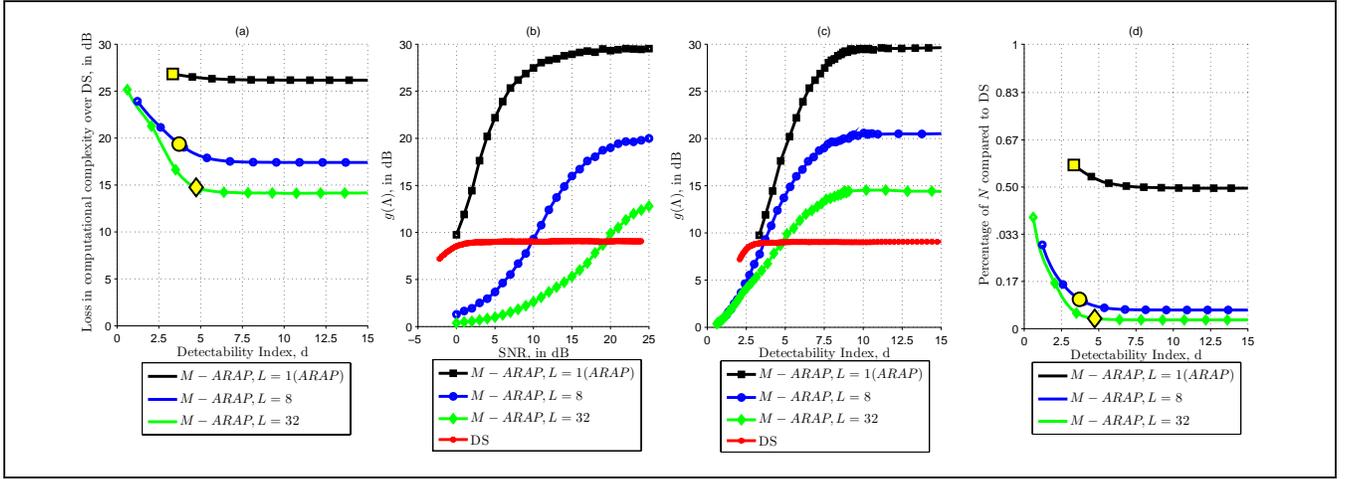


Fig. 4. In (a), we plot the loss in computational complexity of M-ARAP ( $L = 8, 32$ ) and ARAP ( $L = 1$ ) vs distilled sensing (DS). We see that DS requires significantly fewer computations than M-ARAP and ARAP. In (b), we plot the gain in cost function over an exhaustive search given by (30) for M-ARAP ( $L = 8, 32$ ), ARAP ( $L=1$ ), and DS. For lower values of SNR, DS outperforms all versions of M-ARAP. However, the asymptotic performance of DS is lower than M-ARAP. In (c), the same gains are plotted as a function of the detectability index. In (d), the percentage of total measurements between M-ARAP and DS is plotted. In (a) and (d), yellow markers indicate the points on the curve where the performance of DS equals M-ARAP. It is seen that in all cases, M-ARAP uses significantly fewer measurements to get similar performance to DS.

measurement, clustered targets will actually increase the detectability index.

Based on M-ARAP we suggest the following measurement scheme: (1) define a coarse grid pattern; (2) use  $N_1 < N$  pulses to measure each point on the coarse grid; (3) use M-ARAP to decide where and how to rescan the domain in a restricted fine grid; and (4) rescan the searched domain according to the pulse allocation of M-ARAP. An example of a single realization is given in Fig. 5.

#### A. MTI performance analysis

The first question addressed in implementing M-ARAP is how to choose  $N_1$ . One approach is to use the data shown in Fig. 5 in [1], where the optimal energy allocation at the first step is calculated for ARAP. In the simulations below we sweep over the range  $\{4, 8, 16\}$  of  $N$  values. Performance was evaluated in terms of estimation MSE gain and area under the curve (AUC) of the false discovery rate (FDR) Q-ROC curve. The results are shown in Figure 6. Note that the fewer pulses used (e.g.,  $N = 4$ ) the more crucial it is to select  $N_1$  appropriately.

We evaluate potential gains in both estimation MSE as well as performance of a false discovery rate type of target detection and localization. We used the estimator suggested in Section V-A and compared it to the CME implemented with exhaustive search, as given in (35). Results are seen in Fig. 7. Note that M-ARAP has nearly the same performance as measured by estimation gain as compared to ARAP (blue and green dotted lines), yet the number of measurements is significantly reduced (red solid line). For example, with an RCS of 1, the estimation gain differs by less than 1 dB, but M-ARAP uses only about 14% of the measurements.

We also compare false discovery rate and compare Q-ROC curves. Fig. 8 displays the two curves for both exhaustive and M-ARAP search. RCS value of 0.1 was selected and

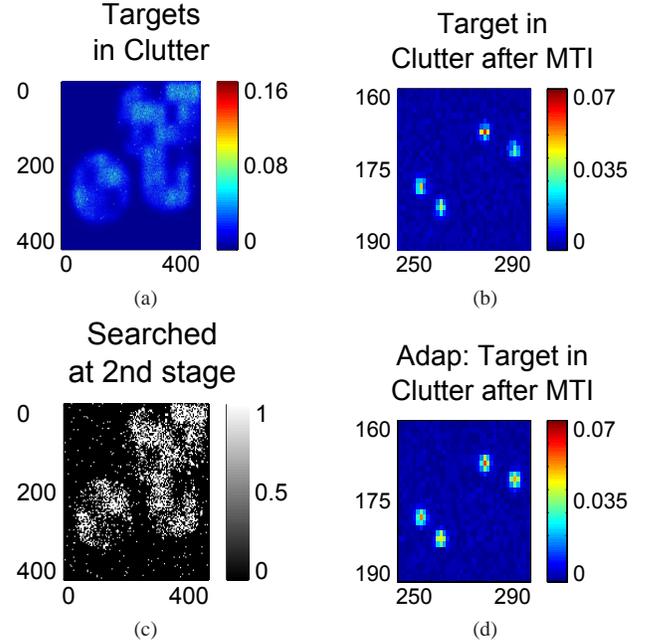


Fig. 5. Moving target indication example. We set targets RCS to 0.1 and chose  $N = 8$  and  $N_1 = 5$ . (a) A single realization of targets in clutter. Figures (b) and (d) zoom in on to the yellow rectangular to allow easier visualization of the improved estimation due to M-ARAP. (b) Portion of the estimated image when data was acquired using exhaustive search and MTI filtration. Figures (c) and (d) are due to M-ARAP search scheme where multi-scale was set to a coarse grid search of  $3 \times 3$  pixels at the first stage. (c) Estimated ROI  $\Psi$  that is searched on a fine resolution level on stage two. (d) Portion of the estimated image when data was acquired using M-ARAP.

it is clearly seen that M-ARAP provides better detection performance for equivalent false discovery rates.

Finally, we note that although the assumption of single-pixel targets may not be validated in this application, the performance in terms of estimation gain and probability detection is still significantly better than an exhaustive search with many

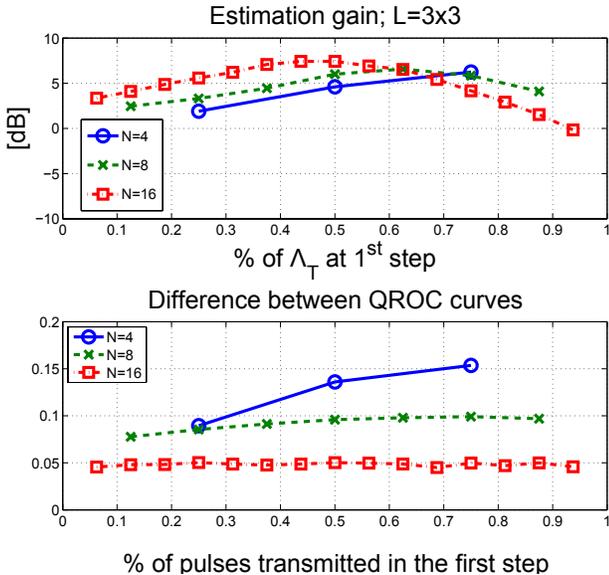


Fig. 6. Simulated gain in estimation and detection performances as a function of  $N_1$  the number of pulses used in the uniform search stage. The operating point of  $\text{RCS}=0.1$  was selected. The upper plot displays gains in estimation MSE. Note that with  $N = 16$  and  $N_1$  equals 7 or 8 yields almost 8 [dB] gains in MSE. The lower plot shows difference in the area under the curve of an FDR test as a function of  $N_1$ . For  $N = 8, 16$ , the exhaustive search yield an almost optimal curve and there is less room for improvement

fewer measurements than with ARAP.

## VII. CONCLUSIONS AND FUTURE WORK

We have proposed a modification, called M-ARAP, of our previously proposed two-stage adaptive resource allocation policy (ARAP) by incorporating multi-scale search. We have established that M-ARAP reduces the number of required measurements for equal gain in target search performance. Specific examples showed 9-15 [dB] gain in estimation performance, for detectability index  $d = 5$ , using less than 18% of the samples needed to perform an exhaustive search.

The M-ARAP method introduced here has not been optimized for detection, and this is a useful area for future work. Potential applications of M-ARAP include air-traffic control, airport security screening, and early detection of breast cancer tumors.

The research in this paper was partially supported by Air Force Office of Scientific Research award FA9550-06-1-0324 and by Air Force Research Laboratory award FA8650-07-D-1221-TO1.

## REFERENCES

- [1] E. Bashan, R. Raich, and A. Hero, "Optimal Two-Stage Search for Sparse Targets Using Convex Criteria," *IEEE Transactions on Signal Processing*, vol. 56, pp. 5389–5402, 2008.
- [2] F. Bandiera, M. Guerriero, and G. Ricci, "Optimized Algorithms for Detection of Sparse Targets in Heterogeneous Gaussian Noise," in *Radar Conference-Surveillance for a Safer World, 2009. RADAR. International, 2009*, pp. 1–5.

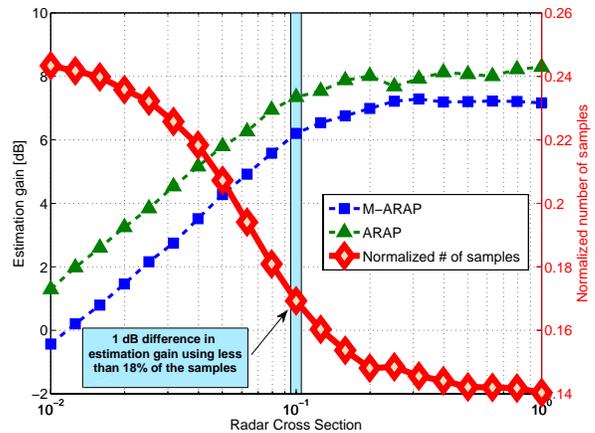


Fig. 7. Simulated gain in estimation and the normalized number of measurement used by M-ARAP vs. targets radar cross section (RCS) coefficient. RCS is alias to signal to noise ratio or contrast since background scatter level was kept fixed. The solid curve with square markers and dashed curve with triangular markers represent estimation gains of M-ARAP and ARAP compared to an exhaustive search, respectively. The dash-dotted curve with diamond markers represent  $N^*$  the number of measurements used by M-ARAP divided by  $Q$  with the corresponding Y-axis values on the right hand side of the figure. For both M-ARAP and ARAP a total of four pulses per cell ( $N = 4$ ) was selected as the energy budget of which three were used at the first stage ( $N_1 = 3$ ) for all RCS values. Recall that for ARAP we have  $N^* > 1$ . Our results clearly illustrate that significant estimation gains can be obtained using M-ARAP with a fraction of the number of measurement required by ARAP.

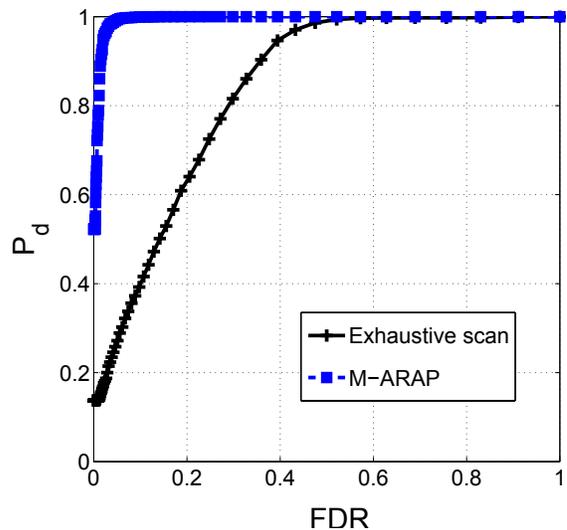


Fig. 8. The two curves on the above figure represent an FDR detection test. One hundred runs in a Monte-Carlo simulation were used to generate each point on the curves. Radar cross section coefficient of 0.1 was selected,  $N = 4$  (four pulses) was the overall energy budget, and  $N_1 = 3$  was used in the first scan for M-ARAP. It is clearly evident that M-ARAP yield significantly better detection performance for equivalent false discovery rate levels.

- [3] E. Candes and T. Tao, "The dantzig selector: statistical estimation when  $p$  is much larger than  $n$ ," *The Annals of Statistics*, vol. 35, no. 6, pp. 2313–2351, December 2007.
- [4] J. Haupt and R. Nowak, "Signal reconstruction from noisy random projections," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 4036–4048, 2006.
- [5] F. Abramovich, Y. Benjamini, D. Donoho, and I. Johnstone, "Special Invited Lecture: Adapting to Unknown Sparsity by Controlling the False

- Discovery Rate,” *The Annals of Statistics*, vol. 34, no. 2, pp. 584–653, 2006.
- [6] D. Donoho and J. Jin, “Asymptotic minimaxity of false discovery rate thresholding for sparse exponential data,” *The Annals of Statistics*, vol. 34, no. 6, pp. 2980–3018, 2006.
- [7] J. Haupt, R. Castro, and R. Nowak, “Distilled Sensing: Adaptive Sampling for Sparse Detection and Estimation,” *Arxiv preprint arXiv:1001.5311*, 2010.
- [8] A. A. Abdel-Samad and A. H. Tewfik, “Hierarchical radar target localization,” in *Proceeding 2000 International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, June 2000, pp. 3033–3036.
- [9] J. Haupt, R. Castro, and R. Nowak, “Distilled sensing: selective sampling for sparse signal recovery,” in *Proc. 12th International Conference on Artificial Intelligence and Statistics (AISTATS)*. Citeseer, 2009, pp. 216–223.
- [10] R. Dorfman, “The detection of defective members of large populations,” *The Annals of Mathematical Statistics*, vol. 14, no. 4, pp. 436–440, December 1943.
- [11] C. G. Pfeifer and P. Enis, “Dorfman-type group testing for a modified binomial model,” *Journal of the American Statistical Association*, vol. 73, no. 363, pp. 588–592, September 1978.
- [12] B. Wray, “Learning classification rules using Bayes,” in *Proceedings of the sixth international workshop on Machine learning*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1989, pp. 94–98.
- [13] E. Bashan, “Efficient resource allocation schemes for search,” Ph.D. dissertation, Univ. of Michigan, Ann Arbor, May 2008. [Online]. Available: [http://deepblue.lib.umich.edu/bitstream/2027.42/60698/1/bashan\\_1.pdf](http://deepblue.lib.umich.edu/bitstream/2027.42/60698/1/bashan_1.pdf)
- [14] A. A. Abdel-Samad and A. H. Tewfik, “Sequential techniques in hierarchical radar target localization,” in *Proceeding 2000 International Conference on Image Processing*, vol. 1, September 2000, pp. 697–700.
- [15] D. C. Schleher, *MTI and pulsed Doppler radar*. Norwood, MA: Artech House, Inc., 1991.



**Eran Bashan** was born in Tel Aviv, Israel, in 1971. He received the B.Sc. degree in Electrical and Computer Engineering from Ben-Gurion University, Beer-Sheva, Israel, in 2000, the M.Sc. degree from Tel-Aviv University, Tel Aviv, Israel, in 2004, and the Ph.D. degree from the University of Michigan, Ann Arbor, in 2008 both in Electrical Engineering (Systems).

He is currently the CEO of Hygieia, Inc. a medical device company focusing on optimization of insulin therapy for diabetes. His research focus is on human decision systems as applied to therapy adjustments. Other interests lie in adaptive and efficient resource-allocation policies with application for search, detection, and estimation theory.



**Gregory Newstadt** received the B.S. degrees (summa cum laude) from Miami University, Oxford, OH, in 2007 in Electrical Engineering and in Engineering Physics. He also received the Masters degree from the University of Michigan, Ann Arbor, MI, in 2009 in Electrical Engineering (Systems)

He is currently a Ph.D. student at the University of Michigan, Ann Arbor, MI, in Electrical Engineering (Systems). His research interests include detection, estimation theory, target tracking, sensor fusion, and statistical signal processing.



**Alfred O. Hero, III** received the B.S. (summa cum laude) from Boston University (1980) and the Ph.D from Princeton University (1984), both in Electrical Engineering. Since 1984 he has been with the University of Michigan, Ann Arbor, where he is the R. Jamison and Betty Professor of Engineering. His primary appointment is in the Department of Electrical Engineering and Computer Science and he also has appointments, by courtesy, in the Department of Biomedical Engineering and the Department of Statistics. In 2008 he was awarded the the Digiteo Chaire d’Excellence, sponsored by Digiteo Research Park in Paris, located at the Ecole Supérieure d’Electricité, Gif-sur-Yvette, France. He has held other visiting positions at LIDS Massachusetts Institute of Technology (2006), Boston University (2006), I3S University of Nice, Sophia-Antipolis, France (2001), Ecole Normale Supérieure de Lyon (1999), Ecole Nationale Supérieure des Télécommunications, Paris (1999), Lucent Bell Laboratories (1999), Scientific Research Labs of the Ford Motor Company, Dearborn, Michigan (1993), Ecole Nationale Supérieure des Techniques Avancées (ENSTA), Ecole Supérieure d’Electricité, Paris (1990), and M.I.T. Lincoln Laboratory (1987 - 1989).

Alfred Hero is a Fellow of the Institute of Electrical and Electronics Engineers (IEEE). He has been plenary and keynote speaker at major workshops and conferences. He has received several best paper awards including: a IEEE Signal Processing Society Best Paper Award (1998), the Best Original Paper Award from the Journal of Flow Cytometry (2008), and the Best Magazine Paper Award from the IEEE Signal Processing Society (2010). He received a IEEE Signal Processing Society Meritorious Service Award (1998), a IEEE Third Millennium Medal (2000) and a IEEE Signal Processing Society Distinguished Lecturership (2002). He was President of the IEEE Signal Processing Society (2006-2007). He sits on the Board of Directors of IEEE (2009-2011) where he is Director Division IX (Signals and Applications).

Alfred Hero’s recent research interests have been in detection, classification, pattern analysis, and adaptive sampling for spatio-temporal data. Of particular interest are applications to network security, multi-modal sensing and tracking, biomedical imaging, and genomic signal processing.