

ADAPTIVE SEARCH FOR DYNAMIC TARGETS UNDER RESOURCE CONSTRAINTS

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,hero}@umich.edu

ABSTRACT

Previous work on resource constrained adaptive search for sparse static targets has produced two-stage allocation policies with desirable properties. For example, for large asymptotic SNR, such policies converge to the true region of interest (ROI) and attain optimal energy allocations relative to exhaustive search. This work investigates the problem of extending previous allocation policies to $T \gg 2$ stages, with particular emphasis on cases where the SNR for any particular stage is considerably less than the asymptotic SNR. Furthermore, a new formulation is given that can account for non-static targets, including a dynamic transition model for target location and a population model to account for targets that leave or enter the scene. Under this formulation, a dynamic adaptive resource allocation policy (D-ARAP) is proposed that performs well and has low computational cost. It is shown that this policy provides significant gains over an exhaustive search policy in both static and dynamic target cases with near optimal performance as $T \rightarrow \infty$. Moreover, D-ARAP is shown to be more robust than a greedy (myopic) policy when there are outliers or when targets may be obscured for periods of time.

I. MOTIVATION

In recent years, there has been significant interest in learning the sparse support of a signal β by intelligently choosing the measurement matrix \mathbf{X} given noisy measurements of the form

$$y = \mathbf{X}\beta + n \quad (1)$$

Given conditions on \mathbf{X} , compressive sensing and sparse approximation methods have shown that one can recover the sparse vector β with probability 1 by using many fewer measurements than the size of β . When the user has the additional degree of freedom to choose \mathbf{X} adaptively as measurements are acquired, it has been shown that one can perform significantly better compared to static policies. Benefits include near-optimal gains in estimation error and related cost functions [1], provable convergence to the true support of β [1] often at rates significantly faster than non-adaptive policies [2] and reliably with significantly smaller

minimum amplitudes below which signal detection is impossible [3]. However, the provable gains of the adaptive algorithms are usually restricted to the asymptotic regime, either in high signal-to-noise ratio (SNR), or in the large size of the state, e.g., number of possible target locations. This work considers applications that only have low to medium SNR levels, but the scene may be repeatedly probed over time. For example, consider air traffic control (ATC) that employs an adaptive radar system to continuously monitor a scene. Another relevant application consists of real-time detection systems that may be limited by a finite amount of computational resources per iteration. In this situation, it would be desirable to have an adaptive allocation policy for $T \gg 2$ stages that

- optimizes an appropriate time-varying cost function subject to a per-stage budget. In an ATC system, the allocation budget is the total time it takes for the radar to complete one complete cycle.
- accounts for dynamic targets.
- provides significant gains over an exhaustive policy that allocates the budget evenly over all potential target locations.
- accomplishes these goals with low computational cost.

This work provides the Dynamic Adaptive Resource Allocation Policy (D-ARAP) that accomplishes these goals by using a heuristic approximation to a complicated optimization problem. We show that D-ARAP performs close to the optimal level as $T \rightarrow \infty$, has considerable gains over an exhaustive search, even for low values of T , and outperforms a greedy (myopic) policy in terms of robustness and convergence.

The rest of the paper is organized as follows. We formalize the problem in Section II and provide the adaptive sensing policy in Section III. Performance analysis is given in Section IV. Finally, in Section V, we conclude and point to future work.

II. PROBLEM FORMULATION

Consider a space $\mathcal{X} = \{1, 2, \dots, Q\}$ containing Q cells equipped with a probability measure P , and a region of interest (ROI), $\Psi \subset \mathcal{X}$. We assume that Ψ will be a randomly selected small subset of \mathcal{X} , though its definition is application specific. 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. In this work, we generalize the formulation from [1], [4] to account for a time-varying ROI so that that $\Psi = \Psi(t)$ is a function of time. Define indicator functions:

$$I_i(t) = \begin{cases} 1, & i \in \Psi(t) \\ 0, & \text{else} \end{cases} \quad (2)$$

As in previous work, let $p = \Pr(I_i(1) = 1)$ be the uniform prior probability that a target is in any one cell (i.e., the non-informative case). Consider a sequential experiment where we observe measurements $\mathbf{y}(t) : \mathcal{X} \rightarrow \mathbf{R}^Q$ at time $t \in \{1, 2, \dots, T\}$. Let $\lambda(i, t) \geq 0$ denote the search effort allocated to cell i at time t . For $t > 1$, we consider mappings of $\{\lambda(i, t)\}_{i=1}^Q$ from past observations $\mathbf{Y}(t-1) \triangleq \{\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(t-1)\}$. The choice of $\{\lambda(i, t)\}_{i,t}$ is called a search policy. We focus here on deterministic mappings of λ , 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. In general, effort might be computing power, complexity, cost, or energy that is allocated to probing a particular cell location. For this work, we assume the measurement model

$$y_i(t) = \sqrt{\lambda(i, t)}\theta_i(t)I_i(t) + n_i(t) \quad (3)$$

where $\theta_i(t) \sim \mathcal{N}(\mu, \sigma_\theta^2)$ are i.i.d. random variables corresponding to the target amplitudes and $n_i(t) \sim \mathcal{N}(0, 1)$ is i.i.d. zero-mean Gaussian noise. Moreover, it is assumed that the posterior probabilities $\Pr(\Psi(t)|\mathbf{Y}(t-1))$ can be computed, e.g., from some signal processing algorithm such as a Kalman or particle filter.

II-A. Dynamic state model

In this work, we propose a simple dynamic state model to incorporate targets that may move. In particular, we assume

- 1) A target in cell i can transition to a neighboring cell with probability q or stay in the same cell with probability, $1 - |W_i^{(t)}|q$ where $W_i^{(t)}$ is the set of neighbors to cell i at time t and $|W_i^{(t)}|q < 1$.
- 2) The target amplitudes follow a random walk, $\theta_i(t) = \theta_i(t-1) + \mathcal{N}(0, \Delta^2)$

For small $p = \Pr(I_i(1) = 1)$, it can be shown that

$$\pi_i^{t|t-1} = \sum_{\Psi_{W_i}^{t-1}} \Pr(I_i(t) = 1|\Psi_{W_i}^{t-1}) \Pr(\Psi_{W_i}^{t-1}|\mathbf{Y}(t-1)) \quad (4)$$

$$\pi_i^{t|t} = \left[1 + \frac{1 - \pi_i^{t|t-1} \Pr(y_i(t)|I_i(t) = 0)}{\pi_i^{t|t-1} \Pr(y_i(t)|I_i(t) = 1)} \right]^{-1}, \quad (5)$$

where $\pi_i^{t|t'} = \Pr(I_i(t) = 1|Y(t'))$, $\Psi_{W_i}^{t-1} = \{I_j(t-1)\}_{j \in W_i^{(t-1)}}$, and $\Pr(I_i(t) = 1|\Psi_{W_i}^{t-1})$ is a function of $\{\pi_j^{t-1|t-1}\}_{j \in W_i^{(t-1)}}$. Note that this can be efficiently computed in an iterative fashion (over t) with $O(Q)$ calculations per stage.

III. ADAPTIVE SEARCH POLICY

III-A. Cost function

In [1] we introduced the following cost function

$$J_{ARAP}(\lambda) = E \left[\sum_{i=1}^Q \frac{I_i}{\lambda(i, 1) + \lambda(i, 2)} \right]. \quad (6)$$

Minimizing (6) over $\lambda(i, t)$ under a total energy constraint $\sum_i^Q \sum_t^T \lambda(i, t) = \lambda_T$ yielded the search policy ARAP, in the case where $T = 2$ and targets were static, i.e., $I_i(t) = I_i$. As discussed before, this work assumes that the between-stages allocation is set a priori. Thus, we consider a separable cost function for $T \gg 2$ as

$$J(\lambda) = E \left[\sum_{t=1}^T \sum_{i=1}^Q \frac{I_i}{\lambda(i, t)} \right]. \quad (7)$$

The goal is to minimize $J(\lambda)$ subject to the constraints, $\sum_{i=1}^Q \lambda(i, t) = \lambda_{total}(t)$, for known $\{\lambda_{total}(t)\}_t$.

III-B. Discussion of optimal policies

Similar to previous work, the omniscient allocation at time t is given by

$$\lambda_{omn}(i, t) = \frac{\lambda_{total}(t)}{|\Psi(t)|} I_i(t) \quad (8)$$

yielding an optimal cost of $J(\lambda_{omn}) = (pQ)^2$. At the other extreme, define the exhaustive policy as $\lambda_E(i, t) = \lambda_{total}(t)/Q$ for all i . This yields the cost, $J(\lambda_E) = pQ^2$. Thus the gain of the omniscient policy compared to an exhaustive search is

$$G(\lambda_{omn}) = -10 \log(pQ)^2 / (pQ^2) = -10 \log p \text{ (dB)}, \quad (9)$$

which is identical to the omniscient gain in [1]. Consider also the policy that knows the target locations at the last state, $\Psi(t-1)$. When targets can transition between cells, $q > 0$, then a proportion of the budget must be used to search neighboring cells. We term this the 'optimal' policy since any policy that is based solely on previous measurements cannot perform better than this policy. The optimal policy is defined as

$$\lambda_{opt}(i, t) = \frac{\lambda_{total}(t)}{|\Psi(t)|} \frac{w_{ij}(t)}{\sum_{k \in W_j^{(t-1)}} w_{kj}(t)} I_j(t-1). \quad (10)$$

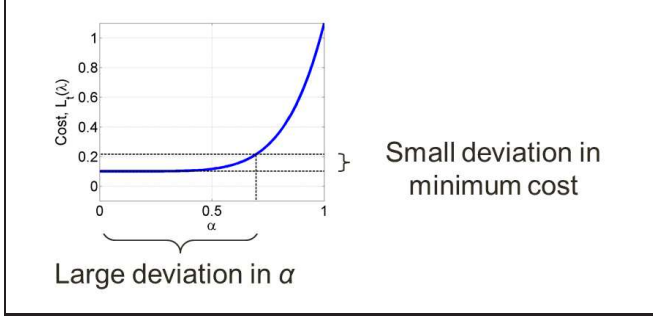


Fig. 1. The myopic cost (14) is plotted as a function of the floor percentage, α . It can be seen that in this case, a large deviation in α leads to a small deviation from the minimum cost ($\alpha = 0$).

where $w_{ij}(t) = \sqrt{\Pr(I_i(t) = 1 | I_j(t-1) = 1)}$. With this policy, one can analyze the loss incurred by searching the neighborhood cells. This can be shown to be

$$\text{Loss} = 10 \log \frac{w_{jj}(t)}{\sum_{k \in W_j^{(t-1)}} w_{kj}(t)} \quad (\text{dB}) \quad (11)$$

III-C. Myopic approach

In general, optimizing (7) with the given budget constraints is combinatorially hard as T increases. If we consider a suboptimal approach by choosing the allocation to the t -th stage given all previous allocations and measurements, the $\{\lambda(i, t)\}_i$ that minimizes (7) is

$$\lambda(i, t) = \frac{\lambda_{total}(t) \sqrt{\Pr(I_i(t) | \mathbf{Y}(t-1))}}{\sum_{j=1}^Q \sqrt{\Pr(I_j(t) | \mathbf{Y}(t-1))}} \quad (12)$$

This myopic approach will tend to track well the targets that have already been detected, but have little ability to recover missed targets.

III-D. Exploitation vs. Exploration: A Heuristic Approach

We suggest a simple (but still suboptimal) approach to allocate a percentage of the budget $\lambda_{total}(t)$ to all cells, yielding the allocation policy that we term D-ARAP:

$$\tilde{\lambda}_{\alpha}(i, t) = \frac{\alpha(t) \lambda_{total}(t)}{Q} + (1 - \alpha(t)) \lambda(i, t), \quad (13)$$

where $\alpha(t) \in (0, 1)$ and $\lambda(i, t)$ is given by (12). In general, choosing $\alpha(t)$ is still a difficult problem. This work proposes a heuristic approximation to optimizing equation (7) that provides significant gains over an exhaustive search at a low computational cost. To understand this heuristic, it is illustrative to look at Figure 1 which plots the myopic cost

$$C_{t,\alpha}(\lambda) = E \left[\sum_{i=1}^Q \frac{I_i(t)}{\tilde{\lambda}_{\alpha}(i, t)} | \mathbf{Y}(t-1) \right] \quad (14)$$

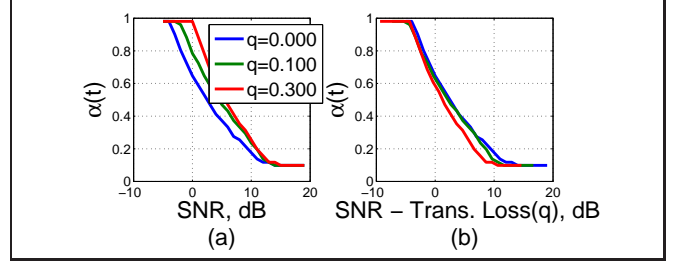


Fig. 2. The left hand plot shows the selection of $\tilde{\alpha}$ within $\gamma = 0.1$ of the minimum myopic cost. Each curve represents a different transition model, with $q = 1$ (blue), $q = 0.8$ (green), and $q = 0.4$ (red). We see that, as expected, $\tilde{\alpha}$ is high for low SNR and decays to a value near zero for high SNR. Simulation was limited to $\alpha > 0.1$ which accounts for the apparent lower bound in the results. The right hand plot shows the selection of $\tilde{\alpha}$ within $\gamma = 0.1$ of the minimum myopic cost, with a SNR loss that depends on q , defined in equation (11).

for varying $\alpha(t)$ when $\tilde{\lambda}_{\alpha}(i, t)$ is chosen according to (13). By definition, $C_{t,\alpha}$ is optimized when $\alpha(t) = 0$. However, as seen in the illustration, large deviations in $\alpha(t)$ do not always lead to large deviations in the cost. Let γ control the deviations that we are willing to tolerate, so that

$$\tilde{\alpha}(t) = \arg \max_{\alpha} \{ \alpha : C_{t,\alpha}(\lambda) < (1 + \gamma) C_{t,0}(\lambda) \} \quad (15)$$

This strategy allocates as much energy to the floor probability as we can, under the condition that the myopic cost is within $(100\gamma)\%$ of the minimum myopic cost. At low SNR, we would expect that $\tilde{\alpha}$ will be large and decay to zero as SNR improves. Figure 2 shows simulation-based results for selecting $\tilde{\alpha}(2)$ for $\gamma = 0.1$ and three transition probabilities to neighboring cells, $q = 1, 0.8, 0.4$. In the left-hand plot, we show the raw selections as a function of the total SNR for a 2-stage problem. Note that, as expected, $\tilde{\alpha}$ decays to near zero as SNR grows. The right-hand plot shows the curves shifted by the loss given in (11). Note that with this correction, the selections of $\tilde{\alpha}$ are nearly identical for all 3 values of q , suggesting that (11) may be useful as a performance index.

Note that choosing $\tilde{\alpha}(t)$ according to (15) still varies as a function of the stage t . Figure 3(a) shows the selection of $\tilde{\alpha}(t)$ as a function of $\lambda_{total}(t)$ (x-axis) and t (various curves). As expected, as either the SNR or t increases, the floor percentage decreases, reflecting the refinement in our estimate of the ROI, $\Psi(t)$. Moreover, Figure 3(b) plots the various selections of $\tilde{\alpha}(t)$ as a function of observed SNR, defined as $\sum_{t=1}^T \lambda_{total}(t)$. The excellent agreement among all of the curves suggests that we can use a functional approximation to selecting $\tilde{\alpha}(t)$ without having to explicitly calculate the functions for each t .

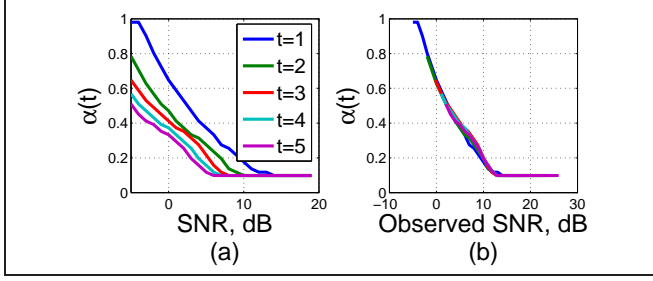


Fig. 3. In general, $\alpha(t, \lambda_{total}(t))$ as computed in (15) depends on both the stage, t and the SNR for that stage, $\lambda_{total}(t)$. This plot shows the selection of α for various values, showing that we can approximate $\alpha(t, \lambda_{total}(t))$ by a simpler function of the ‘observed SNR’, defined as $\sum_t \lambda_{total}(t)$.

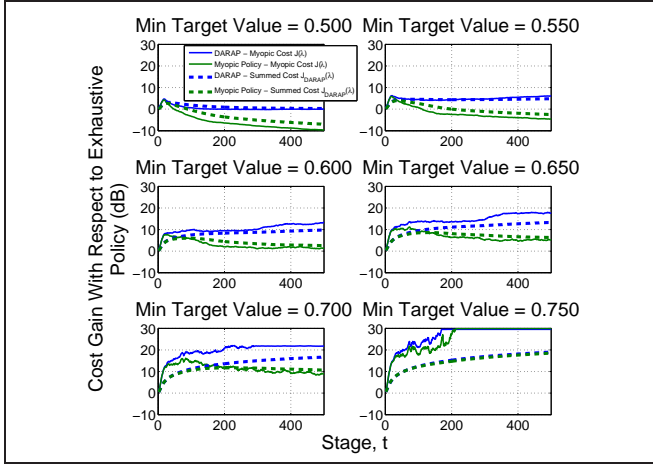


Fig. 4. This figure shows a comparison of the proposed policy (D-ARAP, blue) with the myopic policy (green) as a function of gains in cost over an exhaustive search in a worst-case analysis (static, $q = 0$), where the target returns $\theta_i(t)$ are set to various values, $\theta_0 < \mu_\theta = 1$.

IV. PERFORMANCE ANALYSIS

IV-A. Comparison to the Myopic Policy

In this section, D-ARAP is compared to the myopic policy that sets $\alpha(t) = 0$ for all t . In Figure 4, we compare the policies to an exhaustive search in the static case ($q = 0$) for a long horizon, $T = 500$. In this analysis, we set $\theta_i(t) = \theta_0$ for various values of $\theta_0 = (0.50, 0.55, 0.60, 0.65, 0.70, 0.75)$. Note that for low values of θ_0 , noisy measurements may cause missed detections that could adversely affect the myopic policy allocations. Indeed, for low values of θ_0 , noisy measurements can cause missed targets that are never recovered by the myopic policy for $\theta_0 < 0.75$. On the other hand, D-ARAP is monotonically increasing for all $\theta_0 > 0.5$, suggesting greater robustness than the myopic policy. Moreover, even when $\theta_0 = 0.75$, D-ARAP converges

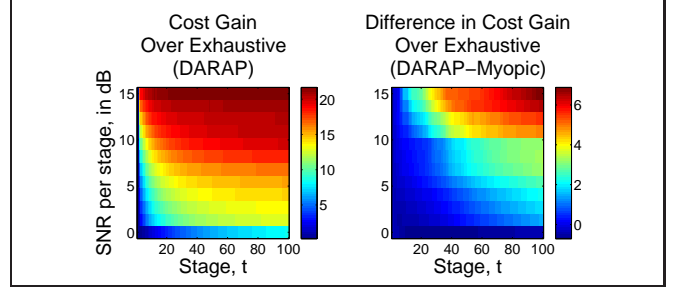


Fig. 5. This figure shows the performance gain in cost for D-ARAP (left) as well as the difference in gains compared to the myopic policy (right) as a function of SNR and stage number in the case where 25% of the scene is obscured for 3 consecutive stages and $q = 0.3$.

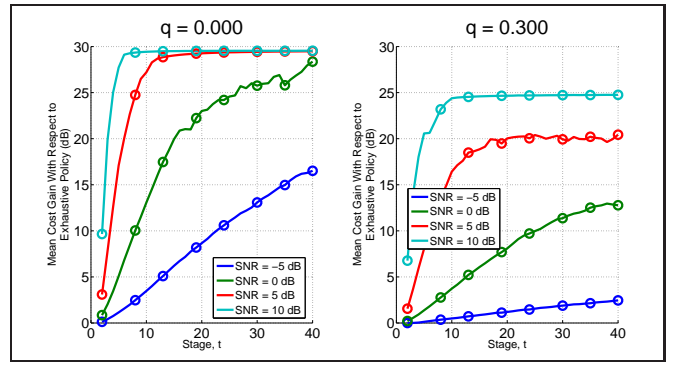


Fig. 6. These plots show the gains in cost over an exhaustive search for various values of λ_{total} (curves) and transition probabilities, $q = 0$ (left) and $q = 0.3$ (right).

to the optimal gain in fewer stages than the myopic policy.

In Figure 5, D-ARAP is also compared to the myopic policy in the unobservable case where 25% of the scene is obscured at each stage for 3 consecutive stages (rotating cells after 3 stages, $T = 40$) and targets can move with transition probability, $q = 0.3$. A non-myopic approach would likely spend additional resources on locations that are going to be obscured in the near future. Although D-ARAP does not explicitly allocate resources in this manner, the results presented here show that D-ARAP significantly outperforms the myopic policy for larger SNR values. Moreover, D-ARAP has monotonically increasing behavior as t or SNR $\rightarrow \infty$, in contrast to the myopic policy, which performs significantly worse in the asymptotic SNR regions.

IV-B. Comparison to Optimal/Exhaustive Policies

This section now considers the case where targets are completely observable and have few outliers ($\mu_\theta = 1$, $\sigma_\theta^2 = 1/36$, $\Delta^2 = 1/64$). We let $Q = 10,000$, $p = 0.001$, and $T = 40$, and compare performance in terms of gains in cost and MSE over an exhaustive policy and probability of

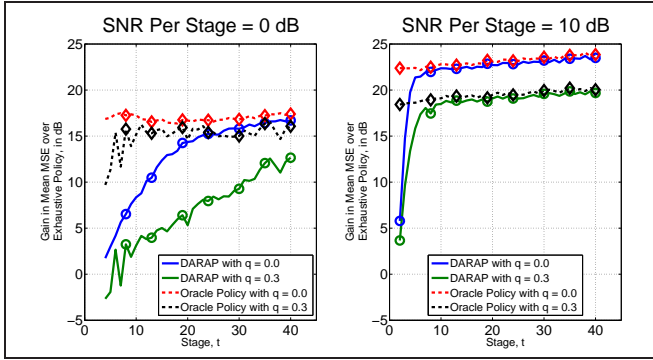


Fig. 7. This plot shows the MSE gains with respect to an exhaustive search for estimating θ_i as function of stage number, t . The solid lines show the results of D-ARAP for the static case (blue, $q = 0$) and dynamic case (green, $q = 0.3$). The dashed lines show the MSE gains using the optimal policy (red, $q = 0$, and black, $q = 0.3$)

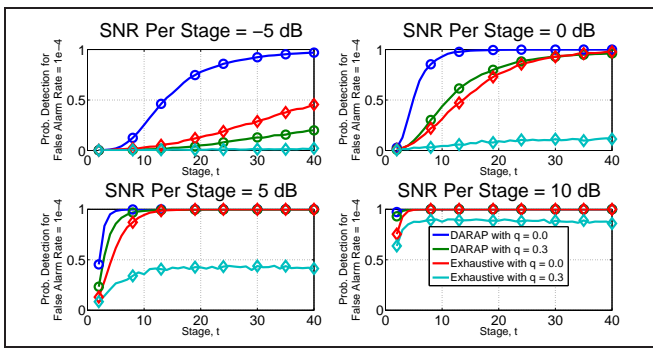


Fig. 8. These plots compare the probability of detection for a fixed probability of false alarm (10^{-4}) as function of the stage number, t . The four subplots show different values of $\lambda_{total}(t)$ in increasing order. Within each subplot, the two solid lines show the results of D-ARAP for $q = 0$ and $q = 0.3$, while the dashed lines show the results from an exhaustive search.

detection for a small fixed false alarm rate. Figure 6 shows gains in the cost given by equation 6 for various values of $\lambda_{total}(t)$ (curves) for the static case ($q = 0$, left) and the dynamic case ($q = 0.3$, right). In both cases, the cost gains are approximately monotonically increasing in t . For the static case, the actual gains approach the optimal gain, regardless of the per-stage-budget λ_{total} . For the dynamic case, the upper bound on performance depends on λ_{total} , but reaches the optimal gain (taking into account the loss in equation (11)) for SNR=10dB.

Figure 7 shows the MSE gains of targets tracked using a modified Kalman filter with respect to an exhaustive search for estimating $\theta_i(t)$ as a function of t . The blue and green curves ($q = 0, 0.3$, respectively) show the results using D-ARAP, while the red and black curves ($q = 0, 0.3$, respec-

tively) show the results using the optimal policy. We see D-ARAP performs favorably in comparison to the optimal policy for both the low SNR case (left plot, SNR=0 dB per stage) after $T = 40$ stages and the higher SNR case (right plot, SNR=10 dB per stage) after $T \ll 40$ stages.

Finally, Figure 8 shows the probability of detection for a fixed probability of false alarm (10^{-4}) as a function of t . Note that the probability of detection for D-ARAP consistently approaches 1 regardless of transition probabilities, as t and $\lambda_{total}(t)$ approach infinity. Moreover, they approach 1 significantly faster than the exhaustive search, which doesn't always have consistent results for $q > 0$ (in particular for SNR < 10 dB). Finally, note that for the larger values of $\lambda_{total}(t)$, D-ARAP approaches $P_d = 1$ within very few stages.

V. CONCLUSIONS AND FUTURE WORK

This work extended previous work to a multiple stage allocation policy by collapsing the previous 2-stage ARAP into a sequence of 1-stage allocations. A heuristic approach to approximating the solution of a complicated optimization problem was provided that performed well in simulation at low computational cost. These results showed asymptotic consistency in time, significant gains over an exhaustive policy alternative, and increased robustness over a myopic (greedy) policy.

Future work plans to compare to existing work in the field of partially observable Markov decision processes, to derive analytical results such as provable consistency as $t \rightarrow \infty$ and/or performance bounds, to consider alternate allocation strategies for $\alpha(t)$ (i.e., strategies that save measurements and/or consider regions that will be obscured in the near future), and to apply our policy to an appropriate real-world problem, such as allocating dwell times for an air traffic control radar system.

VI. 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] J. Haupt, R. Castro, and R. Nowak, "Distilled Sensing: Adaptive Sampling for Sparse Detection and Estimation," *Arxiv preprint arXiv:1001.5311*, 2010.
- [3] J. Haupt, R. Baraniuk, R. Castro, and R. Nowak, "Compressive distilled sensing: Sparse recovery using adaptivity in compressive measurements," in *Signals, Systems and Computers, 2009 Conference Record of the Forty-Third Asilomar Conference on*. IEEE, 2009, pp. 1551–1555.
- [4] E. Bashan, G. Newstadt, and A. Hero, "Two-Stage Multi-Scale Search for Sparse Targets," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2331–2341, 2011.