

MULTISTAGE ADAPTIVE ESTIMATION OF SPARSE SIGNALS

Dennis Wei and Alfred O. Hero, III

University of Michigan, Department of Electrical Engineering and Computer Science
1301 Beal Avenue, Ann Arbor, MI 48109, USA

ABSTRACT

This paper considers sequential adaptive estimation of sparse signals under a constraint on the total sensing effort. A dynamic programming formulation is derived for the allocation of sensing resources to minimize a cost function related to mean squared estimation error. Allocation policies are developed based on the method of open-loop feedback control. These policies are optimal in the two-stage case and improve monotonically thereafter with the number of stages. Numerical simulations show gains up to several dB as compared to recently proposed adaptive methods, and dramatic gains approaching the oracle limit as compared to non-adaptive estimation.

Index Terms— Adaptive sensing, adaptive sampling, resource allocation, sparse signals.

1. INTRODUCTION

This work considers the estimation of sparse signals from observations that are taken sequentially and adaptively. It is now well-known that sparse signals can be efficiently acquired via compressive sensing (see e.g. [1]) using a relatively small number of observations that are incoherent with the basis in which the signal is sparse. However, when noise is present and sensing resources are limited, incoherent observations may not be the most efficient since a large fraction of the resources are allocated to dimensions where the signal is absent. Alternatively, by shaping future observations according to estimates of the signal support derived from past observations, better signal-to-noise ratios (SNR) are possible. Applications in which such adaptive sensing can be readily utilized include agile radars and medical imaging [2, 3].

Existing methods for adaptive sensing of sparse signals can be roughly grouped around two classes of models. In the first class, which is the focus of this paper, observations are restricted to single components in the basis that induces signal sparsity, while resources can be distributed arbitrarily over components and observation stages. An optimal two-stage resource allocation policy was developed in [2] for a cost function related to bounds on estimation and detection performance. The problem in [2] was simplified in [4] through Lagrangian constraint relaxation, while a multiscale approach that uses linear combinations in the first stage was suggested in [3] to reduce the number of measurements. Based on a similar model, a method known as distilled sensing was proposed in [5] for signal support identification and was shown to be asymptotically reliable at SNR levels significantly lower than non-adaptive limits. In the second class of models, the observations can consist of arbitrary linear combinations as in compressive sensing, but in most cases the resource budget is assumed to be discrete, measured

in units of normalized observations. Methods in this second category include Bayesian approaches based on maximizing information gain [6, 7], a bisection algorithm [8], and a generalization of distilled sensing [9].

In this paper, we extend the two-stage allocation policy in [2] to an arbitrary number of stages, focusing on mean squared estimation error directly as opposed to performance bounds in [2]. It is shown that the problem can be formulated as a dynamic program, a framework that facilitates the development of allocation policies. An approximate dynamic programming solution is proposed based on *open-loop feedback control* (OLFC). The performance of these OLFC policies improves monotonically with the number of stages, and in particular upon the optimal two-stage policy in [2]. Simulations show MSE reductions up to 3 dB relative to the optimal two-stage policy and dramatic reductions relative to non-adaptive policies, approaching the oracle limit at high SNR. The OLFC policies also outperform distilled sensing [5] at all SNR and most significantly at higher SNR.

In Section 2, we specify the signal and observation models and formulate a problem of resource-constrained sequential estimation, which is then recast as a dynamic program. In Section 3, optimal and OLFC approaches to the resource allocation problem are discussed and a family of OLFC policies is proposed. Numerical simulations comparing our OLFC policies to other policies are presented in Section 4. Future directions are discussed in Section 5.

2. PROBLEM FORMULATION

We consider signals that are sparse with respect to the standard basis (without loss of generality) in a Q -dimensional space. The subset of indices corresponding to non-zero signal components is referred to as the region of interest (ROI). Let I_i be an indicator such that $I_i = 1$ if i is in the ROI and $I_i = 0$ otherwise. We use a probabilistic signal model in which $I_i = 1$ with prior probability $p_i(0)$, independently of the other indicators. The non-zero signal amplitudes are modelled as independent Gaussian random variables θ_i with prior means $\mu_i(0)$ and variances $\sigma_i^2(0)$. As in [2, 3], a non-informative uniform prior is assumed with $p_i(0) = p_0$, $\mu_i(0) = \mu_0$ and $\sigma_i^2(0) = \sigma_0^2$ for all i , although non-uniform priors could also be accommodated.

A sequence of T observations are made with effort levels $\lambda_i(t)$ that vary with index i and time $t = 0, \dots, T-1$. Depending on the application, the effort $\lambda_i(t)$ might represent observation time, number of samples, energy, cost, or computation. It is assumed that the precision (inverse variance) of an observation increases with effort: given $\lambda_i(t-1)$, the corresponding observation $y_i(t)$ takes the form

$$y_i(t) = I_i\theta_i + \lambda_i(t-1)^{-1/2}n_i(t), \quad t = 1, \dots, T, \quad (1)$$

where $n_i(t)$ represents i.i.d. zero-mean Gaussian noise with variance σ^2 . We restrict attention to static signals so that the signal component $I_i\theta_i$ in (1) does not change with time. For convenience, we

This research was partially supported by Army Research Office grant W911NF-11-1-0391.

write $\mathbf{y}(t) = [y_1(t) \dots y_Q(t)]^T$ (similarly for other indexed quantities) and denote by $\mathbf{Y}(t) = \{\mathbf{y}(1), \dots, \mathbf{y}(t)\}$ the history of observations up to time t . In adaptive sensing, the effort allocation $\boldsymbol{\lambda}(t)$ at time t can depend on the observations $\mathbf{Y}(t)$ collected up to that point, thus incorporating current knowledge. The task is to determine the mapping from $\mathbf{Y}(t)$ to $\boldsymbol{\lambda}(t)$, referred to as an effort or resource allocation policy, subject to a total budget constraint:

$$\sum_{t=0}^{T-1} \sum_{i=1}^Q \lambda_i(t) = Q. \quad (2)$$

For notational brevity, we suppress the dependence of $\boldsymbol{\lambda}(t)$ on $\mathbf{Y}(t)$.

To guide the selection of an effort allocation policy, we seek to minimize the mean squared error (MSE) associated with estimates $\hat{\theta}_i$ of the amplitudes θ_i , based on all observations up to time T and summed over the ROI:

$$\mathbf{E} \left\{ \sum_{i=1}^Q I_i (\hat{\theta}_i - \theta_i)^2 \right\}, \quad (3)$$

where the expectation is taken over I_i , θ_i , and $\mathbf{Y}(T)$. It can be shown [10] that $\hat{\theta}_i$ should be chosen as the conditional mean $\mathbf{E}[\theta_i | I_i = 1, \mathbf{Y}(T)]$. Define $\mu_i(t)$ and $\sigma_i^2(t)$ to be the conditional mean and variance of $\theta_i | I_i = 1, \mathbf{Y}(t)$, and $p_i(t) = \Pr(I_i = 1 | \mathbf{Y}(t))$. Substituting $\hat{\theta}_i = \mu_i(T)$ in (3) and simplifying, we obtain

$$\sigma^2 \mathbf{E}_{\mathbf{Y}(T)} \left\{ \sum_{i=1}^Q \frac{p_i(T)}{\sigma^2/\sigma_0^2 + \sum_{t=0}^{T-1} \lambda_i(t)} \right\}, \quad (4)$$

where the form of the denominator can be derived from (5c) below. In the sequel, we focus on minimizing the expected cost in (4) with respect to the effort allocation policy $\boldsymbol{\lambda}(0), \dots, \boldsymbol{\lambda}(T-1)$, subject to the total effort constraint (2). We note that (4) is similar but not identical to the cost function in [2].

2.1. Formulation as a dynamic program

The determination of an optimal effort allocation policy according to (4) can be formulated as a dynamic program. The dynamic programming viewpoint makes available a well-developed set of approaches, some of which are considered in Section 3. Further background in dynamic programming can be found in [11].

To formulate a sequential decision problem as a dynamic program, the cost function must be expressible as a sum of terms indexed by time t , where each term depends only on the current system state $\mathbf{x}(t)$ and the current effort allocation $\boldsymbol{\lambda}(t)$. While at first glance the cost function (4) does not appear to have such a time-separable property, it can be recast in the required form by defining the state $\mathbf{x}(t)$ as $\mathbf{x}(t) = (\mathbf{p}(t), \boldsymbol{\mu}(t), \boldsymbol{\sigma}^2(t), \Lambda(t))$, where $\Lambda(t)$ represents the effort budget remaining at time t . The state variables are initialized as $p_i(0) = p_0$, $\mu_i(0) = \mu_0$, $\sigma_i^2(0) = \sigma_0^2$, and $\Lambda(0) = Q$, and evolve according to the following recursions derived in [10]:

$$p_i(t+1) = \frac{p_i(t)\phi_1}{p_i(t)\phi_1 + (1-p_i(t))\phi_0}, \quad (5a)$$

$$\mu_i(t+1) = \frac{\sigma^2 \mu_i(t) + \lambda_i(t) \sigma_i^2(t) y_i(t+1)}{\sigma^2 + \lambda_i(t) \sigma_i^2(t)}, \quad (5b)$$

$$\sigma_i^2(t+1) = \frac{\sigma^2 \sigma_i^2(t)}{\sigma^2 + \lambda_i(t) \sigma_i^2(t)}, \quad (5c)$$

$$\Lambda(t+1) = \Lambda(t) - \sum_{i=1}^Q \lambda_i(t), \quad (5d)$$

where

$$\begin{aligned} \phi_0 &= \phi(y_i(t+1); 0, \sigma^2/\lambda_i(t)), \\ \phi_1 &= \phi(y_i(t+1); \mu_i(t), \sigma_i^2(t) + \sigma^2/\lambda_i(t)), \end{aligned}$$

and $\phi(\cdot; \mu, \sigma^2)$ denotes a Gaussian probability density function with mean μ and variance σ^2 . It can also be shown that the distribution for $y_i(t+1) | \mathbf{Y}(t)$ is given by the denominator in (5a), from which it follows that

$$\mathbf{E}_{\mathbf{Y}(t+1)}\{p_i(t+1) | \mathbf{Y}(t)\} = p_i(t), \quad t = 0, \dots, T-1. \quad (6)$$

Using (5c) and (6), the allocation problem may be stated as

$$\begin{aligned} \min_{\boldsymbol{\lambda}(0), \dots, \boldsymbol{\lambda}(T-1)} \quad & \mathbf{E}_{\mathbf{Y}(T-1)} \left\{ \sum_{i=1}^Q \frac{p_i(T-1)}{\sigma^2/\sigma_i^2(T-1) + \lambda_i(T-1)} \right\} \\ \text{s.t.} \quad & \sum_{t=0}^{T-1} \sum_{i=1}^Q \lambda_i(t) = Q, \end{aligned} \quad (7)$$

where the cost function is now of the desired form with a single non-zero term at time $T-1$. The cost function depends implicitly on $\boldsymbol{\mu}(t)$ and $\boldsymbol{\lambda}(t)$, $t = 0, \dots, T-2$ through the distribution of $\mathbf{Y}(T-1)$.

3. EFFORT ALLOCATION POLICIES

3.1. Optimal policies

In principle, it is possible to employ exact dynamic programming to obtain an optimal policy for (7). We decompose (7) into a sequence of optimizations proceeding backward in time, starting with

$$\begin{aligned} J_{T-1}^*(\mathbf{x}(T-1)) &= \min_{\boldsymbol{\lambda}(T-1)} \sum_{i=1}^Q \frac{p_i(T-1)}{\sigma^2/\sigma_i^2(T-1) + \lambda_i(T-1)} \\ \text{s.t.} \quad & \sum_{i=1}^Q \lambda_i(T-1) = \Lambda(T-1), \end{aligned} \quad (8)$$

and defining recursively

$$\begin{aligned} J_t^*(\mathbf{x}(t)) &= \min_{\boldsymbol{\lambda}(t)} \mathbf{E}_{\mathbf{Y}(t+1)} \{J_{t+1}^*(\mathbf{x}(t+1)) | \mathbf{x}(t), \boldsymbol{\lambda}(t)\} \\ \text{s.t.} \quad & \sum_{i=1}^Q \lambda_i(t) \leq \Lambda(t) \end{aligned} \quad (9)$$

for $t = T-2, T-3, \dots, 0$. The desired optimal cost in (7) is $J_0^*(\mathbf{x}(0))$. The notation in (9) reflects the fact that the distribution of $\mathbf{y}(t+1) | \mathbf{Y}(t)$ is completely specified by $\mathbf{x}(t)$ and $\boldsymbol{\lambda}(t)$, and the state $\mathbf{x}(t+1)$ is specified by $\mathbf{x}(t)$, $\boldsymbol{\lambda}(t)$, and $\mathbf{y}(t+1)$ through (5).

An optimal policy can be obtained by first solving (8) for $\boldsymbol{\lambda}(T-1)$ and then using the result in (9) to solve for $\boldsymbol{\lambda}(T-2)$. The remaining allocations are determined in the same recursive way. This exact procedure is tractable in a few cases. For $T = 1$, it suffices to solve (8), a convex optimization problem whose solution is derived in [10]. For $T = 2$ and a uniform prior ($p_i(0) = p_0$, $\mu_i(0) = \mu_0$, $\sigma_i^2(0) = \sigma_0^2$), symmetry allows the initial allocation $\boldsymbol{\lambda}(0)$ to be restricted to the form $\boldsymbol{\lambda}(0) = \beta^{(2)}(0)\mathbf{1}$, where $\mathbf{1}$ denotes a vector with unit entries. Thus (9) becomes a one-dimensional optimization with respect to the multiplier $\beta^{(2)}(0)$. For fixed $\beta^{(2)}(0)$, the expectation in (9) can be evaluated by sampling from the distribution of $\mathbf{y}(1)$ and then solving (8) for the resulting values of the state $\mathbf{x}(1)$. In other cases however, an exact solution to (7) is very difficult. Thus for $T > 2$ we consider an approximate method as described next.

3.2. Open-loop feedback control policies

A well-known approach to approximate dynamic programming is that of open-loop feedback control (OLFC) [11]. To determine the allocation $\lambda(t)$ at time t in an OLFC policy, we make the simplifying assumption that future allocations $\lambda(t+1), \dots, \lambda(T-1)$ can depend only on the current set of observations $\mathbf{Y}(t)$ and not future observations. In light of this assumption, the expectations $\mathbf{E}_{\mathbf{y}(t+1)|\mathbf{Y}(t)} \dots \mathbf{E}_{\mathbf{y}(T-1)|\mathbf{Y}(T-2)}$ may be applied to the numerator only in (7), yielding $p_i(t)$ using (6). Conditioning on $\mathbf{Y}(t)$ or equivalently the state $\mathbf{x}(t)$, the problem becomes

$$\begin{aligned} \min_{\lambda(t), \dots, \lambda(T-1)} \quad & \sum_{i=1}^Q \frac{p_i(t)}{\sigma^2/\sigma_i^2(t) + \sum_{\tau=t}^{T-1} \lambda_i(\tau)} \\ \text{s.t.} \quad & \sum_{i=1}^Q \sum_{\tau=t}^{T-1} \lambda_i(\tau) = \Lambda(t), \end{aligned} \quad (10)$$

which is an optimization in the variables $\bar{\lambda}_i(t) = \sum_{\tau=t}^{T-1} \lambda_i(\tau)$ and is of the same form as the last-stage optimization (8). To state the optimal solution to (10), we define π to be an index permutation that sorts the quantities $p_i(t)\sigma_i^4(t)$ in non-increasing order:

$$p_{\pi(1)}(t)\sigma_{\pi(1)}^4(t) \geq p_{\pi(2)}(t)\sigma_{\pi(2)}^4(t) \geq \dots \geq p_{\pi(Q)}(t)\sigma_{\pi(Q)}^4(t). \quad (11)$$

Next define $g(k)$ to be the monotonically non-decreasing function of $k = 0, 1, \dots, Q$ with $g(0) = 0$, $g(Q) = \infty$, and

$$g(k) = \frac{\sigma^2}{\sqrt{p_{\pi(k+1)}(t)\sigma_{\pi(k+1)}^2(t)}} \sum_{i=1}^k \sqrt{p_{\pi(i)}(t)} - \sum_{i=1}^k \frac{\sigma^2}{\sigma_{\pi(i)}^2(t)} \quad (12)$$

for $k = 1, \dots, Q-1$. Then the solution $\bar{\lambda}^*(t)$ to (10) is given by

$$\bar{\lambda}_{\pi(i)}^*(t) = \left(\Lambda(t) + \sum_{j=1}^k \frac{\sigma^2}{\sigma_{\pi(j)}^2(t)} \right) \frac{\sqrt{p_{\pi(i)}(t)}}{\sum_{j=1}^k \sqrt{p_{\pi(j)}(t)}} - \frac{\sigma^2}{\sigma_{\pi(i)}^2(t)} \quad (13)$$

for $i = 1, \dots, k$ and $\bar{\lambda}_{\pi(i)}^*(t) = 0$ otherwise, where the number of non-zero components k is determined by the interval $(g(k-1), g(k)]$ to which the budget parameter $\Lambda(t)$ belongs. The monotonicity of $g(k)$ ensures that the mapping from $\Lambda(t)$ to k is well-defined.

Equations (11)–(13) specify the optimal values of $\bar{\lambda}_i(t) = \sum_{\tau=t}^{T-1} \lambda_i(\tau)$, but they do not specify how much effort should be put into the present allocation $\lambda(t)$. In the remainder of the section, we restrict attention to $\lambda(t)$ of the form $\lambda(t) = \beta^{(T)}(t)\bar{\lambda}^*(t)$, where $\beta^{(T)}(t) \in [0, 1]$ is the fraction of the remaining budget used at time t and the superscript T denotes the number of stages.

The multipliers $\beta^{(T)}(t)$ are chosen based on a generalization of the optimal policies for $T = 1, 2$ discussed in Section 3.1. It can be seen that both of these optimal policies belong to the OLFC class with $\beta^{(1)}(0) = \beta^{(2)}(1) = 1$ and $\beta^{(2)}(0)$ determined according to the exact procedure in (9). Note that the second stage in the $T = 2$ policy is identical to the $T = 1$ policy. For $T > 2$, we follow the same strategy of nesting the $(T-1)$ -stage policy within the T -stage policy, setting $\beta^{(T)}(t) = \beta^{(T-1)}(t-1)$ for $t = 1, 2, \dots, T-1$. We then optimize over the first-stage multiplier $\beta^{(T)}(0)$. Define $J_t^{(T)}(\mathbf{x}(t))$ to be the cost-to-go of a T -stage policy in this family starting from time t and state $\mathbf{x}(t)$. Then $\beta^{(T)}(0)$ is given by

$$\beta^{(T)}(0) = \arg \min_{0 \leq \beta(0) \leq 1} \mathbf{E}_{\mathbf{y}(1)} \left\{ J_0^{(T-1)}(\mathbf{x}(1)) \mid \mathbf{x}(0), \beta(0)\bar{\lambda}^*(0) \right\}, \quad (14)$$

where $\bar{\lambda}^*(0) = \mathbf{1}$ under a uniform prior. Policies with an increasing number of stages are determined recursively using (14) starting from $\beta^{(1)}(0) = 1$ for $T = 1$. The expectation in (14) can be computed by sampling from the distribution of $\mathbf{y}(1)$ and then simulating the $(T-1)$ -stage policy starting from state $\mathbf{x}(1)$. These computations can be done offline since they depend only on the initial state $\mathbf{x}(0)$ and the previously determined policy of $T-1$ stages.

The cost of the nested OLFC policies defined above improves with the number of stages. The cost of the T -stage policy is

$$J_0^{(T)}(\mathbf{x}(0)) = \min_{0 \leq \beta(0) \leq 1} \mathbf{E}_{\mathbf{y}(1)} \left\{ J_0^{(T-1)}(\mathbf{x}(1)) \mid \mathbf{x}(0), \beta(0)\bar{\lambda}^*(0) \right\}, \quad (15)$$

using the fact that $J_1^{(T)}(\mathbf{x}(1)) = J_0^{(T-1)}(\mathbf{x}(1))$ by construction. When $\beta(0) = 0$, the observations $\mathbf{y}(1)$ are not taken, $\mathbf{x}(1) = \mathbf{x}(0)$, and $J_0^{(T)}(\mathbf{x}(0)) = J_0^{(T-1)}(\mathbf{x}(0))$. It follows from (15) that

$$J_0^{(T)}(\mathbf{x}(0)) \leq J_0^{(T-1)}(\mathbf{x}(0)), \quad T = 2, 3, \dots, \quad (16)$$

implying in particular that the nested policies for $T > 2$ improve upon the optimal policy for $T = 2$.

4. NUMERICAL SIMULATIONS

We use numerical simulations to quantify the performance gains predicted by (16) for OLFC policies. We set $Q = 1000$ and generate signals and observations according to the model in Section 2. The signal mean μ_0 is normalized to 1 and the signal standard deviation σ_0 is set to $1/4$. The OLFC policies are compared to the optimal non-adaptive policy, which under a uniform prior allocates one unit of sensing effort to all components, and to distilled sensing (DS) [5]. For both OLFC and DS, the number of stages T is varied from 2 to 10 and the final estimate is $\mu(T)$. For DS, we use the allocation of effort over stages suggested in [5] with equal first and last stages and a geometric decrease by a factor of $3/4$ for the intervening stages.

In Fig. 1, we plot estimation gains (i.e. reductions in MSE) relative to the non-adaptive policy as a function of SNR, defined as $10 \log_{10}(\mu_0^2/\sigma^2)$ in dB. Each point represents the average of 4000 simulations. For context, we also plot the gain of the oracle policy, which has perfect knowledge of the ROI and distributes resources uniformly over the ROI. Higher gains are achieved in general for $p_0 = 0.01$ since resources can be concentrated on fewer components once the ROI is identified. The 10-stage OLFC policy improves upon the 2-stage OLFC policy as expected with the largest gains at intermediate SNR, although the gains are proportionally larger at low SNR. Recall that the 2-stage OLFC policy is optimal for $T = 2$. At high SNR, the OLFC policies approach the oracle gain, which in turn approaches the sparsity factor $1/p_0$. In contrast, the DS policies saturate at significantly lower gains since they are not designed with estimation performance in mind. While the 10-stage DS policy outperforms the optimal two-stage policy at lower SNR, the 10-stage OLFC policy has the best performance at all SNR.

Fig. 2 shows increases in estimation gains with the number of stages T . The incremental gains predicted by (16) diminish as T increases. Allowing more stages is more beneficial at lower SNR and higher sparsity, whereas at higher SNR most of the signal components can be located in a single step and a two-stage policy performs almost as well as a policy with many more stages. In all cases shown, a 5-stage OLFC policy performs better than a 10-stage DS policy.

5. CONCLUSIONS AND FUTURE WORK

We have presented multistage resource allocation policies that improve upon the optimal two-stage policy in [2] for sequential esti-

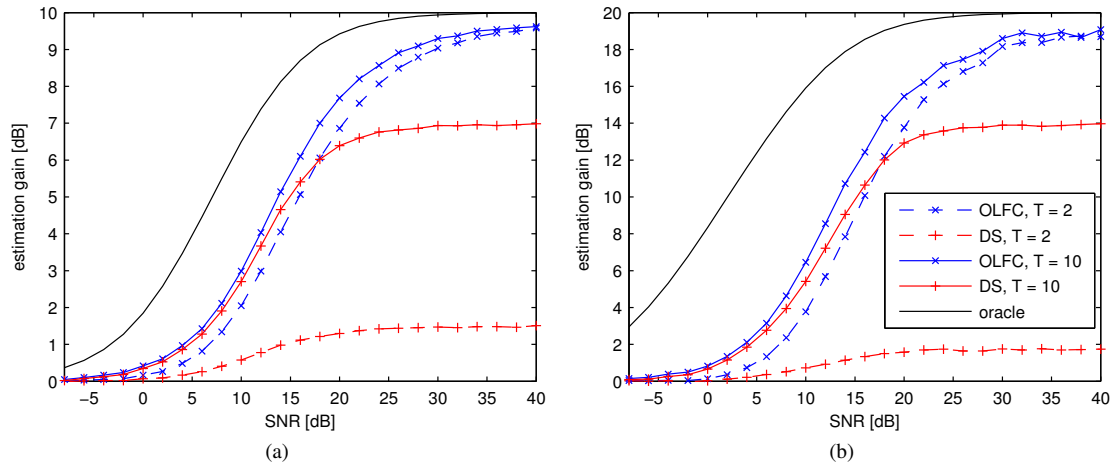


Fig. 1. Estimation gain relative to non-adaptive allocation as a function of SNR for (a) $p_0 = 0.1$ and (b) $p_0 = 0.01$. The 10-stage open-loop feedback control (OLFC) policy improves upon the 2-stage OLFC policy with maximum gains around 1 dB for $p_0 = 0.1$ and SNR = 10–20 dB, and 2.5–3 dB for $p_0 = 0.01$ and SNR = 10–15 dB. Note that the 2-stage OLFC policy is optimal for $T = 2$. As the SNR increases, the proposed OLFC policies approach the oracle gain of $1/p_0$ and outperform distilled sensing (DS) by several dB.

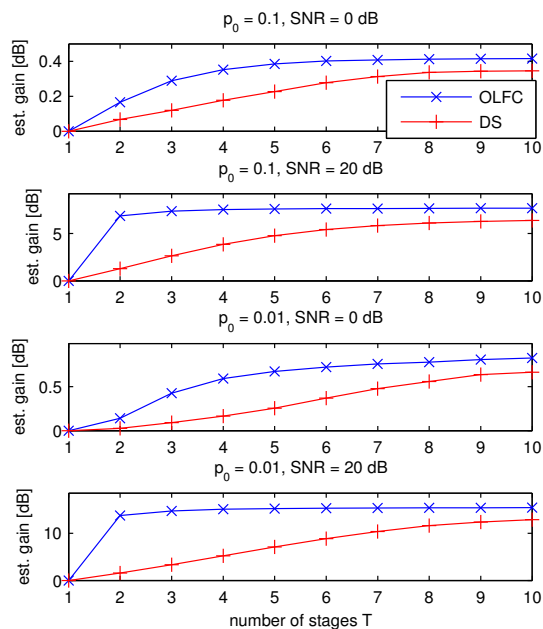


Fig. 2. Estimation gain as a function of the number of stages T . Gains diminish as T increases but less quickly at lower SNR and higher sparsity. In all cases shown, our proposed open-loop feedback control (OLFC) policy with 5 stages performs better than a 10-stage distilled sensing (DS) policy.

mation of sparse signals. Demonstrating gains in a more realistic example is part of ongoing work. The dynamic programming framework introduced in this paper can potentially be leveraged to develop tractable policies for other inference tasks such as detection or a combination of detection and estimation. More general observation models involving linear combinations may also be incorporated.

6. REFERENCES

- [1] E. J. Candes, J. Romberg, and T. Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” *IEEE Trans. Inf. Theory*, vol. 52, pp. 489–509, Feb. 2006.
- [2] E. Bashan, R. Raich, and A. O. Hero, “Optimal two-stage search for sparse targets using convex criteria,” *IEEE Trans. Signal Process.*, vol. 56, pp. 5389–5402, Nov. 2008.
- [3] E. Bashan, G. Newstadt, and A. O. Hero, “Two-stage multi-scale search for sparse targets,” *IEEE Trans. Signal Process.*, vol. 59, pp. 2331–2341, May 2011.
- [4] D. Hitchings and D. A. Castanon, “Adaptive sensing for search with continuous actions and observations,” in *Proc. IEEE Conf. Decision and Control (CDC)*, Dec. 2010, pp. 7443–7448.
- [5] J. Haupt, R. M. Castro, and R. Nowak, “Distilled sensing: Adaptive sampling for sparse detection and estimation,” *IEEE Trans. Inf. Theory*, vol. 57, pp. 6222–6235, Sept. 2011.
- [6] S. Ji, Y. Xue, and L. Carin, “Bayesian compressive sensing,” *IEEE Trans. Signal Process.*, vol. 56, pp. 2346–2356, June 2008.
- [7] R. M. Castro, J. Haupt, R. Nowak, and G. M. Raz, “Finding needles in noisy haystacks,” in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Apr. 2008, pp. 5133–5136.
- [8] M. Iwen and A. H. Tewfik, “Adaptive group testing strategies for target detection and localization in noisy environments,” Tech. Rep. 2311, IMA Preprint Series, June 2010.
- [9] J. D. Haupt, R. G. Baraniuk, R. M. Castro, and R. D. Nowak, “Compressive distilled sensing: Sparse recovery using adaptivity in compressive measurements,” in *Conf. Rec. Asilomar Conf. Signals Syst. Comput.*, Nov. 2009, pp. 1551–1555.
- [10] D. Wei and A. O. Hero, “Multistage adaptive estimation of sparse signals,” in preparation, 2012.
- [11] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 1, Athena Scientific, Nashua, NH, 3rd edition, 2005.