ADAPTIVE COMPRESSIVE SAMPLING USING PARTIALLY OBSERVABLE MARKOV DECISION PROCESSES

Ramin Zahedi, Lucas W. Krakow, Edwin K. P. Chong, and Ali Pezeshki

Colorado State University, ECE Department
1373 Campus Delivery, Fort Collins, CO 80523-1373

ABSTRACT

We present an approach to adaptive measurement selection in compressive sensing for estimating sparse signals. Given a fixed number of measurements, we consider the sequential selection of the rows of a compressive measurement matrix to maximize the mutual information between the measurements and the sparse signal's support. We formulate this problem as a partially observable Markov decision process (POMDP), which enables the application of principled reasoning for sequential measurement selection based on Bellman's optimality condition.

Index Terms— Compressive sensing, POMDP, rollout, Q-value approximation, adaptive sensing

1. INTRODUCTION

The notion of adaptively scheduling compressive measurements in a closed-loop fashion for estimating sparse signals has been discussed in a number of recent papers; e.g., [1]–[6]. The basic idea is to select, at each measurement step, the "best" measurement vector (perhaps from a predefined library of vectors) given the measurements that have been obtained thus far. The objective function for the optimization and the assumptions about the sparse signal to be estimated can take a number of possible forms. For example, in [3], the objective is to maximize the posterior variance of the expected measurement.

In this paper, we consider the problem of sequentially scheduling compressive measurements so that at the end, we optimize a measure of performance for estimating an s-sparse signal in $\mathbb{R}^N$. We assume that the number of measurements m is fixed. Moreover, the performance measure used here is the mutual information between the measurements and the sparse signal's support set, i.e., the set containing locations of all the nonzero entries of the signal, which we consider to be random. We further assume that given the support, the signal amplitudes follow a Gaussian distribution. The noise is taken to be additive white Gaussian, added at the output of the compressive sampler. Under these assumptions, maximizing the mutual information is equivalent to minimizing the conditional signal's support entropy given all the measurements.

With this objective, we present a finite-horizon partially observable Markov decision process (POMDP) formulation (see [7]), whose solution gives the optimal policy for sequentially selecting compressive measurements. One of the main problems in using the POMDP formulation is that finding the optimal policy in general is computationally prohibitive. Fortunately, several methods exist for approximating optimal policies in POMDP (see [8] for a review of such methods). In our numerical examples, we have used one of these approximation methods, known as *rollout*, and have compared our solution with a class of nonadaptive methods described in Section 4.

2. POMDP FORMULATION

Let $\mathbf{x} \in \mathbb{R}^N$ represent the s-sparse signal,[1] which we wish to estimate from a fixed number m of compressive measurements. The kth compressive measurement is given by

$$y_k = \mathbf{a}_k^T \mathbf{x} + w_k, \tag{1}$$

where $\mathbf{a}_k^T$ is the kth row of an $m \times N$ compressive measurement matrix to be designed, and $w_k \sim \mathcal{N}(0, \sigma_w^2)$. For the s-sparse signal $\mathbf{x}$, there are $\binom{N}{s}$ possibilities for the locations of the nonzero entries, called the support of the signal $\mathbf{x}$. We consider a prior probability distribution over such possibilities for the signal's support. Given the support, the values of the nonzero entries of this support have a multivariate Gaussian distribution.

Our goal is to sequentially design the row vectors $\mathbf{a}_k^T$, $k = 1, 2, \ldots, m$, to optimize the mutual information between the observation vector $\mathbf{y} = [y_1, y_2, \ldots, y_m]^T$ and the support of $\mathbf{x}$. This is a multistep scheduling problem, which can be formulated as a POMDP. To formulate this problem as a POMDP, we need to specify the following components:

States and State Transition Law: Let $\mathbf{d}$ be an $(s \times 1)$ discrete random vector whose entries represent locations of

This work is supported in part by DARPA/DSO contract N66001-11-C-4023, ONR contract N00014-08-1-110, and AFOSR contract FA9550-09-1-0518.

[1] In this paper, we call an N-dimensional signal s-sparse if it has exactly $s \ll N$ nonzero entries.

the nonzero entries of the signal $\mathbf{x}$. Each entry of the vector $\mathbf{d}$ can take a value from $\Omega = \{1, 2, \ldots, N\}$. Therefore, the vector $\mathbf{d}$ takes values from the set $\Omega_s = \{\mathcal{A} : \mathcal{A} \subset \Omega, |\mathcal{A}| = s\}$. Let $\mathbf{v}$ be a continuous random vector in $\mathbb{R}^s$ whose entries are the s nonzero elements of $\mathbf{x}$. The state space of the POMDP model is defined as the Cartesian product of two components: the first is the set Ω_s and the second is the set $\mathbb{R}^s$.

Let $\mathbf{s}_k = (\mathbf{d}, \mathbf{v})$ represent the state of the POMDP at time k. Since the signal $\mathbf{x}$ does not change over time, the state transition law is trivial: $\mathbf{s}_k = \mathbf{s}_T$ for $k = 1, \ldots, m$, where $\mathbf{s}_T$ is the true state of the system.

Define H_k as $H_k = \{\mathbf{a}_1, y_1, \ldots, \mathbf{a}_k, y_k\}$, i.e., the information available at time k. Let $f_{\mathbf{v},\mathbf{d}|H_k}(\mathbf{z}, \mathbf{q}|H_k)$ be the conditional joint probability distribution function of random vectors $\mathbf{v}$ and $\mathbf{d}$ at time k given the information H_k. Also, let $P_{\mathbf{d}|H_k}(\mathbf{q}|H_k)$ be the conditional probability mass function for the random vector $\mathbf{d}$ defined over the state space Ω_s given H_k. The conditional probability distribution function $f_{\mathbf{v}|H_k}(\mathbf{z}|H_k)$ is a Gaussian mixture, i.e.,

$$f_{\mathbf{v}|H_k}(\mathbf{z}|H_k) = \sum_{\mathbf{q}\in\Omega_s} P_{\mathbf{d}|H_k}(\mathbf{q}|H_k) f_{\mathbf{v}|\mathbf{d},H_k}(\mathbf{z}|\mathbf{q}, H_k),$$

where each component $f_{\mathbf{v}|\mathbf{d},H_k}(\mathbf{z}|\mathbf{q}, H_k)$ is a multivariate Gaussian distribution with $(s \times 1)$ mean vector $\boldsymbol{\mu}_{\mathbf{d},k}$ and $(s \times s)$ covariance matrix $\underline{\mathbf{C}}_{\mathbf{d},k}$.

The function $f_{\mathbf{v},\mathbf{d}|H_k}(\mathbf{z}, \mathbf{q}|H_k)$ is the *belief state* $\mathbf{b}_k$ of the POMDP at time k. The transition probability for $\mathbf{b}_k$ is obtained from Bayes' rule applied to the state transition law and the observation law. We will derive the belief state update equation later.

Actions: In this model, each action $\mathbf{a}_k$ is a real N-dimensional vector. Thus, the action space is $\mathcal{A} = \mathbb{R}^N$.

Observations and Observation Law: The set of possible observations is the real line $\mathbb{R}$. These observations are collected using the linear model (1). Given $\mathbf{s}_k = (\mathbf{d}, \mathbf{v})$ and $\mathbf{a}_k = \mathbf{a}$ at time k, then $y_k|(\mathbf{d}, \mathbf{v}, \mathbf{a}) \sim \mathcal{N}(\mathbf{a}_{\mathbf{d}}^T \mathbf{v}, \sigma_w^2)$, where the vector $\mathbf{a}_{\mathbf{d}}$ is an $(s \times 1)$ vector whose ith entry is the $\mathbf{d}(i)$th entry of the vector $\mathbf{a}$ (where $\mathbf{d}(i)$ is the ith entry of the vector $\mathbf{d}$).

Cost: We define the POMDP cost $c_k(\mathbf{s}_k, \mathbf{a}_k)$ for each time step k to be the conditional mutual information between the random variable $\mathbf{d}$ and the observation y_k given the history H_k, i.e., $c_k(\mathbf{s}_k, \mathbf{a}_k) = I(y_k; \mathbf{d}|H_k)$. We also define the POMDP *belief cost* $r_k(\mathbf{b}_k, \mathbf{a}_k)$ in the following way:

$$r_k(\mathbf{b}_k, \mathbf{a}_k) = \mathbf{E}\left[c_k(\mathbf{s}_k, \mathbf{a}_k)|H_k\right].$$

Note that the expectation is with respect to the posterior distribution of the POMDP state $\mathbf{s}_k$ at time k, which is $\mathbf{b}_k$.

Having defined the POMDP components, we now describe how to update the belief state $\mathbf{b}_k$. Since

$$\mathbf{b}_k = f_{\mathbf{v}|\mathbf{d},H_k}(\mathbf{z}|\mathbf{q}, H_k) P_{\mathbf{d}|H_k}(\mathbf{q}|H_k),$$

we only have to show how values of $f_{\mathbf{v}|\mathbf{d},H_k}(\mathbf{z}|\mathbf{q}, H_k)$ and $P_{\mathbf{d}|H_k}(\mathbf{q}|H_k)$, the components of the Gaussian mixture, are updated once a new action $\mathbf{a}_k$ is chosen and a new observation y_k is observed. The idea behind doing this is very simple; the weights of the Gaussian mixture, $P_{\mathbf{d}|H_k}(\mathbf{q}|H_k)$, for each possible $\mathbf{q}$ can be updated using the following equation:

$$P_{\mathbf{d}|H_k}(\mathbf{q}|H_k) = \frac{P_{\mathbf{d}|H_{k-1}}(\mathbf{q}|H_{k-1}) f_{y_k|\mathbf{d},H_{k-1},\mathbf{a}_k}(y|\mathbf{q}, H_{k-1}, \mathbf{a}_k)}{\sum_{\mathbf{q}\in\Omega_s} P_{\mathbf{d}|H_{k-1}}(\mathbf{q}|H_{k-1}) f_{y_k|\mathbf{d},H_{k-1},\mathbf{a}_k}(y|\mathbf{q}, H_{k-1}, \mathbf{a}_k)},$$

where $y_k|\mathbf{d}, H_{k-1}, \mathbf{a}_k \sim \mathcal{N}(\mathbf{a}_{\mathbf{d},k}^T \boldsymbol{\mu}_{\mathbf{d},k}, \mathbf{a}_{\mathbf{d},k}^T \underline{\mathbf{C}}_{\mathbf{d},k} \mathbf{a}_{\mathbf{d},k} + \sigma_w^2)$. Given the vector $\mathbf{d}$, the following linear system describes the dynamics of our model:

$$\begin{cases} \mathbf{v}_{k+1} = \mathbf{v}_k, \\ y_k = \mathbf{a}_{\mathbf{d},k}^T \mathbf{v}_k + w_k, \end{cases}$$

where $\mathbf{a}_{\mathbf{d},k}$ is built from action $\mathbf{a}_k$ in the same way that the vector $\mathbf{a}_{\mathbf{d}}$ is built from the vector $\mathbf{a}$. The vector $\mathbf{v}_k$ has the conditional distribution $\mathcal{N}(\boldsymbol{\mu}_{\mathbf{d},k}, \underline{\mathbf{C}}_{\mathbf{d},k})$ for a given history H_k and $\mathbf{d}$. Thus, this distribution can be updated using a simple Kalman filter. Since there are $\binom{N}{s}$ possibilities for the vector $\mathbf{d}$, we have to use $\binom{N}{s}$ Kalman filters to keep track of values $\boldsymbol{\mu}_{\mathbf{d},k}$ and $\underline{\mathbf{C}}_{\mathbf{d},k}$ after getting a new observation y_k at each time k.

Next, we will briefly discuss optimal policies and Q-value approximation.

3. Q-VALUE APPROXIMATION

The main goal is to find, for each time step k, an optimal mapping π_k^* from the set of distributions $\mathcal{B}$ over the state space $\mathcal{S}$ to the actions space $\mathcal{A}$, i.e., $\pi_k^* : \mathcal{B} \to \mathcal{A}$, such that if the action $\mathbf{a}_k$ is chosen based on this mapping, then over the time horizon m, a predefined objective function is maximized. We refer to $\pi^* = \{\pi_1^*, \ldots, \pi_m^*\}$ as the optimal policy.

There are different ways to define the objective function. We choose the objective function for a policy $\pi = \{\pi_1, \pi_2, \ldots, \pi_m\}$ to be the *expected cumulative reward* defined as

$$V_m^{\pi}(\mathbf{b}_1) = \mathbf{E}\left[\sum_{k=1}^{m} r_k(\mathbf{b}_k, \pi_k(\mathbf{b}_k)) \middle| \mathbf{b}_1\right].$$

Let $V_m^{\pi^*}$ be the optimal objective function value over the horizon m when an optimal policy π^* is used. In general, finding an optimal policy for such an objective function is very hard. However, *Bellman's principle*, which states that

$$V_m^{\pi^*}(\mathbf{b}_1) = \max_{\mathbf{a}} (r_1(\mathbf{b}_1, \mathbf{a}) + \mathbf{E}[V_{m-1}^{\pi^*}(\mathbf{b}_2)|\mathbf{b}_1, \mathbf{a}]),$$

and also,

$$\pi_1^*(\mathbf{b}_1) = \arg\max_{\mathbf{a}} (r_1(\mathbf{b}_1, \mathbf{a}) + \mathbf{E}[V_{m-1}^{\pi^*}(\mathbf{b}_2)|\mathbf{b}_1, \mathbf{a}]),$$

suggests an alternative way to find an optimal policy. Define

$$Q_{m-k}(\mathbf{b}_k, \mathbf{a}) = r_k(\mathbf{b}_k, \mathbf{a}) + \mathbf{E}[V^{\pi^*}_{m-k}(\mathbf{b}_{k+1})|\mathbf{b}_k, \mathbf{a}]$$

to be the *Q-value* of taking action $\mathbf{a}$ given the belief state $\mathbf{b}_k$ at time k. Then, by Bellman's principle,

$$\pi^*_k(\mathbf{b}_k) = \arg\max_{\mathbf{a}} Q_{m-k}(\mathbf{b}_k, \mathbf{a}).$$

This means that finding the optimal action at time k for the state $\mathbf{b}_k$ is equal to finding the action with the largest Q-value for state $\mathbf{b}_k$ at time k.

Although Bellman's principle provides a method to find an optimal solution, for many problems, because of the size of the state space, the action space, and also the observation space, finding such a solution is computationally prohibitive. Therefore, we resort to estimating the Q-value of each action at each belief state. There are different ways to do this approximation. In this paper, we describe a method known as *rollout*.

In rollout, for a given state $\mathbf{b}_k$ and action $\mathbf{a}$, we replace the value $\mathbf{E}[V^{\pi^*}_{m-k}(\mathbf{b}_{k+1})|\mathbf{b}_k, \mathbf{a}]$ in $Q_{m-k}(\mathbf{b}_k, \mathbf{a})$ with $\mathbf{E}[V^{\pi^{\text{base}}}_{m-k}(\mathbf{b}_{k+1})|\mathbf{b}_k, \mathbf{a}]$. The term $V^{\pi^{\text{base}}}_{m-k}(\mathbf{b}_{k+1})$ is the expected cumulative reward for the $m-k$ remaining time steps when actions are chosen from a predefined policy known as the *base* policy π^{base}. In other words, for finding an estimate for $Q_{m-k}(\mathbf{b}_k, \mathbf{a})$, instead of solving optimization problems nested in layers one over another, we only solve one optimization problem for finding the optimal immediate action $\mathbf{a}^*_k$ and we choose the remaining future actions from the base policy π^{base}. It is shown in [9] that the resulting policy from rollout performs at least as well as the base policy used in the rollout.

4. SIMULATION RESULTS

As a simple example, we consider the problem of estimating a 1-sparse signal in $\mathbb{R}^{75}$ using $m = 8$ measurements. Recall from Section 2 that $\mathbf{d}$ and $\mathbf{v}$ are two random vectors representing locations and values of the nonzero entries of the signal $\mathbf{x}$. We use a specific prior for the probability mass function $P_{\mathbf{d}}$, shown in Fig. 1, which suggests that the signal's nonzero entry is located somewhere in the first 16 indices of the signal, which covers about 20% of the range of possible locations. Therefore, when designing rows of the measurement matrix, each measurement row $\mathbf{a}^T_k$ can be divided into two parts: the first part is a 16-dimensional vector with norm 1, and the second part is a 59-dimensional vector of zeros. Moreover, our actions are chosen from a static library of 50 measurement vectors that together, build a Grassmannian line packing (see [10] and [11]) in $\mathbb{R}^{16}$. We also assume that $\mathbf{v} \sim \mathcal{N}(0, \sigma^2)$. For our simulations, we generated 100 signal samples and for each signal sample, we repeated our simulations 50 times.

We have implemented two variations of our method. In the first variation (Greedy), we set the decision-making horizon to 1, i.e., at each time step k, we choose the action that maximizes the one step ahead belief cost $r_k(\mathbf{b}_k, \mathbf{a}_k)$. In other words, this method is a greedy method for choosing actions. In the second variation (Rollout), at each time step and for each action candidate, a rollout method looks 4 steps ahead and chooses the action with the best Q-value. The base policy we use for rollout randomly chooses one of the 50 measurement vectors in the above library and adds a 59-dimensional vector of zeros to it. To estimate the Q-value for each candidate action, we took an average over 50 Q-value samples.

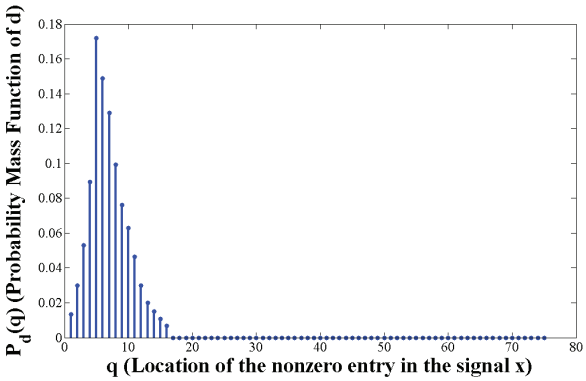

Fig. 1. Prior structure used for $P_{\mathbf{d}}$.

We compare the performance of the measurement matrices obtained from the variations Greedy and Rollout of our method with three other methods. In the first method (Random), the measurement rows are unit norm vectors in $\mathbb{R}^{75}$ with i.i.d. $\mathcal{N}(0, 1/m)$ values. This method does not use the prior in designing the measurement rows. In the second and third methods, measurement rows have a similar structure to those used in variations of our method. In other words, they are divided into two parts where the first part is of dimension 16 and has a unit norm and the second part contains zeros. In the second method (Limited Random), the first part of the vector contains i.i.d. $\mathcal{N}(0, 1/m)$ values, and in the third method (Random from Library), the first part of the vector is one of the measurement vectors from the library introduced above that is chosen randomly for each time step.

Fig. 2 shows the performance of the five methods introduced above. The metric used in this figure for comparing these methods is the posterior probability of the true support, i.e., the value of $P_{\mathbf{d}|H_8}(\mathbf{d}_T|H_8)$ where $\mathbf{d}_T$ is the true location of the nonzero entry of the signal. We have shown the performance of these methods for different values of signal-to-noise-ratio (SNR), which is defined as SNR $= \sigma^2/\sigma^2_w$. This figure shows that both variations of our method, i.e., Greedy and Rollout, perform similar to each other but better than the other three methods as the SNR increases.

As another way of comparing the above methods, we ran an experiment where we wanted to see on average how many more measurements the Random and Limited Random methods require in order to reach the performance of the Greedy

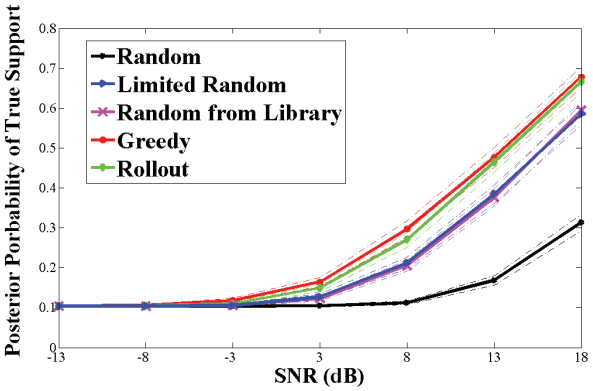

Fig. 2. Comparison of the five methods for $N = 75$ and $m = 8$. The dashed lines indicate the 95% confidence intervals.

method when the same metric (the posterior probability of the true support) is used for comparison. Fig. 3 shows the results for different values of SNR, suggesting that in this simple scenario with the particular prior used, simply knowing that the true support lies within the first 16 indices provides very significant performance gains over the random scheme. Moreover, further adaptation only provides slight marginal gains over the exploitation of the prior knowledge provided by our highly informative prior distribution.

5. CONCLUSIONS

In this paper, we have presented a principled approach to adaptive measurement selection in compressed sensing using POMDP theory. We have presented numerical examples to compare the performance of our design with alternative nonadaptive methods in estimating a 1-sparse signal, which indicate that at moderate to high SNR regimes, our design outperforms the nonadaptive methods. In the scenario presented here, the prior knowledge is found to be much more useful than "acquired" knowledge. This is perhaps unsurprising in view of the very specific nature of the prior used in our experiment, the simple static scenario being considered here, and the recent observations in [1] on the limits of adaptation under static scenarios. Currently, we are considering more dynamic scenarios, where we expect adaptation to play a more significant role.

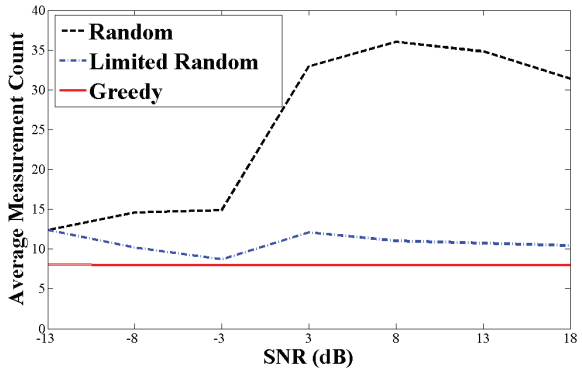

Fig. 3. Average number of measurements required for Random and Limited Random to reach the Greedy's performance.

6. REFERENCES

[1] E. Arias-Castro, E. J. Candès, and M. Davenport, "On the fundamental limits of adaptive sensing," *Arxiv preprint arXiv:1111.4646*, November 2011.

[2] M. Elad, "Optimized projections for compressed sensing," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5695–5702, December 2007.

[3] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2346–2356, June 2008.

[4] S. Ji, D. Dunson, and L. Carin, "Multitask compressive sensing," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 92–106, January 2009.

[5] R. M. Castro, J. Haupt, R. Nowak, and G. M. Raz, "Finding needles in noisy haystacks," in *Proc. IEEE ICASSP*, April 2008, pp. 5133–5136.

[6] J. Haupt, R. Castro, and R. Nowak, "Distilled sensing: Adaptive sampling for sparse detection and estimation," *IEEE Trans. Inform. Theory*, vol. 57, no. 9, pp. 6222–6235, September 2011.

[7] M. L. Littman, "A tutorial on partially observable Markov decision processes," *Journal of Mathematical Psychology*, vol. 53, no. 3, pp. 119–125, June 2009.

[8] E. K. P. Chong, C. Kreucher, and A. O. Hero III, "Partially observable Markov decision process approximations for adaptive sensing," *Discrete Event Dynamic Systems*, vol. 19, pp. 377–422, September 2009.

[9] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, Athena Scientific, 3 edition, January 2007.

[10] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes, etc.: Packings in Grassmannian spaces," *Experimental Mathematics*, vol. 5, no. 2, pp. 139–159, April 1996.

[11] R. Zahedi, A. Pezeshki, and E. K. P. Chong, "Measurement design for detecting sparse signals," *Physical Communication*, September 2011.