52nd IEEE Conference on Decision and Control
December 10-13, 2013. Florence, Italy

Opportunistic Sensor Activation in the Face of Data Deluge

Kamil Nar   Sourabh Bhattacharya   Tamer Başar

Abstract— **In this paper, we consider the problem of designing optimal measurement policies for a sensor that acquires sequential compressive measurements of a static vector of unknown sparsity as originally formulated in [3]. The scenario is modeled as a finite horizon sequential decision making problem when the number of samples is strictly restricted to be less than the overall horizon of the problem. We assume that at each instant of time the sensor can decide whether or not to take an observation, based on the quality of the sensing parameters. The objective of the sensor is to minimize the coherence of the final sensing matrix. We provide a closed-loop optimal measurement policy for a low-dimensional problem. We generalize the optimal policy to obtain a feasible policy for acquiring arbitrary length sparse vectors of unknown sparsity. Finally, we illustrate the performance of the proposed policy by providing simulation results.**

I. INTRODUCTION

Modern surveillance and monitoring networks are prone to the problem of *data deluge*. They are burdened with large amounts of data often containing very low information content. For example, a surveillance system placed in a remote area with the goal of detecting intruders will often observe an inactive scene, yet record or transmit the same amount of data during these periods as if the scene were active [26]. *Compressive Sensing* (CS) provides a promising solution to the problem of data deluge [2]. CS is an emerging field driven by the fact that a small group of non-adaptive linear projections of a compressible signal contains enough information for reconstruction and processing [9], [4]. The benefits of compressive sensing are clearly apparent in applications in which a high cost is associated with each measurement sample. Additional memory savings can be incurred if some prior temporal information about "events of interest" is used to activate the sensors. Such a strategy falls into the category of *opportunistic sensing* (OS) [1]: a methodology that aims to dynamically adjust sensing system parameters to the state of the environment. In this work, we address the problem of opportunistically activating sensors that use compressive sensing schemes to acquire measurements.

Kamil Nar is with the Department of Electrical and Computer Engineering at the University of Illinois at Urbana-Champaign, Urbana, IL 61801, nar2@illinois.edu. Sourabh Bhattacharya is with the Department of Mechanical Engineering at Iowa State University, Ames, IA 50011, sbhattac@iastate.edu. Tamer Başar is with the Department of Electrical and Computer Engineering and the Coordinated Science Laboratory at the University of Illinois at Urbana-Champaign, Urbana, IL 61801, basar1@illinois.edu.

This work was supported in part by an ISU research initiation grant, in part by the U.S. Army Research Laboratory (ARL) and the U.S. Army Research Office (ARO) under grant number W911NF-09-1-0383 and in part by the U.S. Air Force Office of Scientific Research (AFOSR) under grant number FA9550-10-1-0573.

In an opportunistic setting, it is apparent that the sensing system must make a dynamic decision based on the current state of the environment and the previous measurements. However, such a decision can be made only if either some statistics about the signal is available *a priori* or real-time information about the signal is available from a side channel [25]. The latter means that some secondary modality (e.g., a motion detector) must collect the actionable information. However, this involves extracting useful information at the expense of collecting additional data, which defeats the original goal of resolving the problem of data deluge unless the resolution of the side sensor is kept low. Hence a trade-off between the quantity and quality of data arises in the presence of a secondary modality. In order to avoid this dilemma, we consider the former case where a single sensor makes continuous observation in order to estimate the signal with some *a priori* knowledge about the sparsity of the signal. We assume that the signal is sparse, but the exact sparsity is unknown.

In order to modify the sensor activation policy based on prior measurements, we take observations sequentially. The sequential process of acquiring samples using CS schemes has garnered some attention in the recent past. An important application area is dynamic magnetic resonance imaging (MRI) for real-time medical applications such as MR image guided surgery, or functional MRI to track brain activation changes. Prior work in sequential compressive sensing has primarily focused on the problem of adaptively shaping the sensing mechanism based on previous measurements. In [21], the authors propose a measurement policy that maximizes the posterior variance of the expected measurement. A major drawback of the proposed technique is the lack of scalability with the problem dimension. In [6], [17], the authors propose computationally efficient Bayesian algorithms in which the adaptive projections can be computed in time linear in the length of the signal. Additionally, theoretical bounds on the performance of the adaptive scheme is presented for the first time in [17]. In [13], [14], the authors introduce a new adaptive sampling-and-refinement scheme called the *Distilled Sensing* for recovery of a sparse signal. *Compressive Distilled Sensing* (CDS) is an extension of the previous technique to compressive sampling regimes. In [16], the authors provide recovery results for CDS that are valid only for sparse signals with dynamic range on the order of a constant. In [15], the authors propose a sequential sensing-and-refinement approach based on their prior works in [14] that removes the dynamic range constraints present in [16]. They extend these ideas to compressive sensing and seek to identify conditions sufficient to enable a stronger exact support recovery result. For time varying sparsity in signals, there is some ongoing work in the area of recursive estimation of sparse sequences with structured temporal variation in the sparsity pattern. In [23], [24], the authors propose techniques for recursive sparse reconstruction of dynamic natural signal/image sequences for small changes in sparsity with time. In [20], the authors address the problem of stability of the aforementioned techniques. All the aforementioned work assume an *apriori* knowledge about the sparsity of the signal. [8] addresses the challenges associated with acquisition of signals of unknown sparsity. The authors propose a stopping rule for acquiring an adequate number of measurements when the sparsity of the inherent signal is unknown. They also address the scenario when the signals are near-sparse and corrupted with noise.

In contradistinction with the aforementioned work on sequential compressed sensing that deals with asymptotic performance of the sensor, we address here a finite horizon problem. This problem formulation appeals to a large variety of practical problems. For example, in tracking applications involving a static sensor, a mobile target remains in the sensing range for a limited amount of time [7], [26]. Accordingly, in this work we address the problem of optimal measurement policies for a sensor that is constrained to acquire compressed measurements (for example, a single pixel camera [11]) of a static sparse vector in finite time. The problem is formulated as a sequential decision problem, and the objective of the sensor is to maximize the *incoherence* [10], [22] of the final sensing matrix, which in turn extends our technique to scenarios in which the sparsity of the data is unknown.

In addition to addressing the problem of data deluge, optimal activation policies for sensors also appeals to the paradigm of energy-efficient sensing. Energy consumption is an important factor that governs the overall life time of a sensor network. Apart from the vast literature in wireless networks that deals with efficient communication and routing protocols for conserving energy, there has been some recent work that proposes energy efficient sensor scheduling schemes using ideas that lie at the interface of estimation theory and control. [18] considers a remote estimation problem with an energy-harvesting sensor and a remote estimator. Due to the randomness of energy available for communication, the problem of finding a communication scheduling strategy for the sensor is a challenge. The estimator relies on messages communicated by the sensor to produce real-time estimates of the sensors observations. The paper employs dynamic programming to characterize the optimal strategies for finding a communication scheduling strategy for the sensor and an estimation strategy for the estimator that jointly minimize an expected sum of communication and distortion costs over a finite time horizon. [12] and [19] introduce the problem of recursive estimation with *limited* information in order to study the limited battery power of wireless devices in sensor networks. This is modeled by imposing a hard constraint on the number of available transmissions that are possible between an encoder-decoder pair that communicates using a wireless channel for estimating a random process. These papers provide, under different settings, the optimal structure of the encoder-decoder pair and the corresponding optimal transmission policy of the encoder. In [3], the previous formulation on limited sensing has been extended to sensors that use compressive sensing techniques to acquire sparse signals. The problem has been formulated as a sequential decision making problem in which the objective of the sensor is to minimize the final coherence matrix. A greedy approach is provided to the aforementioned problem which is suboptimal. In this work, we revisit this problem, and provide optimal acquisition strategies for low dimensional data.

The organization of the paper is as follows. Section II introduces the problem. Section III presents the encoder structure when the sensing mechanism is based on random Bernoulli trials. Section IV gives the optimal policy for $N = 2$. Section V presents an extension of the optimal policy proposed in Section IV to arbitrary N and provides simulation results to illustrate the performance of the proposed policy. Section VI presents our conclusions.

II. Problem Formulation

In this section, we present the formal problem statement. Consider a sensor taking measurements of a vector $x \in \mathbb{R}^N$. For example, x can be obtained by vectorizing a matrix comprising of pixel intensities in an image. The sensor is comprised of an encoder-decoder pair, denoted by (Φ, Δ). In this paper, we assume Φ to be a matrix A of size $M \times N$, and thereby restrict ourselves to linear measurements. Let a_{ij} denote the element in row i and column j of matrix A.

Now we describe our sensing mechanism. We assume that the measurement process consists of L sequential stages, where $L \in \mathbb{R}_+$. At each stage, prior to acquiring a measurement, the sensor generates a row vector $a^k = [a_{k1}, \ldots, a_{kN}] \in \mathbb{R}^N$, where $k \in [1, \ldots, L]$. Based on prior measurements, the sensor either takes a measurement at the current stage or decides otherwise. Therefore, at stage k, the set of actions available to the sensor is defined as follows:

$$\sigma_k = \begin{cases} 0 & \text{No Observation (NO)} \\ 1 & \text{Observation (O)} \end{cases}$$

Overall, the sensor is restricted to take M measurements within L stages, where $M \leq L$. In case the sensor decides to take the jth measurement at stage k, the measurement sample y_j is given by the following expression:

$$y_j = \sum_{i=1}^{N} a_{ki} x_i, \quad j \in [1, \cdots, M] \tag{1}$$

Therefore, each measurement corresponds to the projection of x along a^k. Let $y = [y_1, \cdots, y_M] \in \mathbb{R}^M$ denote the final measurement vector. We obtain the following relationship between x and y,

$$y = Ax, \tag{2}$$

where A is a matrix formed by stacking the row vectors a^k corresponding to y_k based on the relation (1). Let s_k denote the number of measurements taken by the sensor till the k^{th} stage; s_k satisfies the following equation:

$$s_{k+1} = s_k + \sigma_k$$

The decision of the sensor whether or not to take a measurement at any stage depends on the information available to the sensor. Assuming perfect recall, we have the following information structure available to the sensor:

$$I_k = \{(s_k, k); a^1, \ldots, a^k\}, \quad 1 \leq k \leq L$$

Let the sensor's decision at time k be denoted by $\mu_k(\cdot)$. σ_k, $\mu_k(\cdot)$ and I_k are related in the following manner:

$$\sigma_k = \mu_k(I_k)$$

The decision of the sensor at stage k is a function of $\{a^1, \ldots, a^k\}$, and $k-1$ past decisions, i.e.,

$$\mu_k(\cdot) : \mathbb{Z}^+_{M+1} \times \mathbb{Z}^+_{L+1} \times \underbrace{\mathbb{R}^N \times \cdots \times \mathbb{R}^N}_{k \text{ times}} \to \{\mathrm{O}, \mathrm{NO}\},$$

where $\mathbb{Z}^+_n = \{1, ..., n-1\}$. The policy of the sensor, Π, can be defined as follows:

$$\Pi = \{\mu_1, \ldots, \mu_L\}$$

A policy Π is admissible if μ_k maps to O M times. So as to maximize the success rate of signal recovery algorithms in CS literature, the objective is to find an admissible policy π^* for the sensor that minimizes the *coherence* (or maximizes the *incoherence*) of the final sensing matrix, which is given by the following expression,

$$C^{\pi}_{(M,L)} = \max_{l,m \in [1,N]} E[\frac{|a_l \cdot a_m|}{\|a_l\|_2 \|a_m\|_2}], \tag{3}$$

where a_l and a_m denote the columns of matrix $A = [a_1 \cdots a_N]^T$.

From the CS literature [4], it is well known that the incoherence of the final sensing matrix A plays an important role in the design of the decoder when the original vector x is sparse. For perfect recovery, sensing matrices are generated from specific random processes, for example, from Bernoulli trials, Gaussian ensembles or random DFT matrices [4], which ensures that the resulting sensing matrix is incoherent with high probability. However, in a standard CS setting, the complete sensing matrix is generated *a priori*. Therefore, this work explores the scenario in which the sensor can adapt its sensing mechanism sequentially in order to improve the incoherence of the final sensing matrix. Moreover, since the incoherence of the final sensing matrix does not depend on the sparsity of the x, the resulting technique can be extended to signals of unknown sparsity, which is the problem under consideration.

In the next section, we analyze the above problem when a_{ij}'s are an outcome of i.i.d. Bernoulli trials.

III. Bernoulli trials

In this section, we address the problem of minimizing (3), stated in the previous section, when the elements of Φ are obtained from Bernoulli trials, and Δ is chosen as the Basis Pursuit decoder. The aforementioned forms of the encoder-decoder pair have been found to be efficient for reconstruction of sparse vectors in undersampled signals [5].

To be more specific, we precisely define the following structure for the encoder-decoder pair.

A. Encoder (Random Symmetric Signs Ensemble)

If $\sigma_k = 1$, then the encoder transmits $y_j = \sum_{i=1}^{N} a_{ki} x_i$ to the decoder, where each a_{ki} is drawn from an i.i.d. Bernoulli process with the following distribution:

$$a_{ki} = \begin{cases} \frac{1}{\sqrt{M}} & w.p. \quad \frac{1}{2} \\ -\frac{1}{\sqrt{M}} & w.p. \quad \frac{1}{2} \end{cases} \tag{4}$$

B. Decoder (Basis Pursuit)

We choose the decoder to be the basis pursuit decoder that has the following form,

$$\min \|\hat{x}\|_1 \text{ subject to } A\hat{x} = y,$$

where $\hat{x}$ is the reconstruction of x from the observation y. Based on the sampling distribution, we conclude that $\|a_i\| = 1 \quad \forall i$, irrespective of the outcome of the individual trials. From the definition of σ_k, (3) can be reformulated as follows:

$$\pi^* \quad = \quad \arg\min_{\pi} \max_{l,m \in [1,N]} E[|\sum_{k=1}^{L} \sigma_k a_{kl} a_{km}|] \tag{5}$$

Finally, the outcome of the Bernoulli trials can be redefined as follows without any change in the problem statement:

$$a_{ij} = \begin{cases} +1 & w.p. \quad \frac{1}{2} \\ -1 & w.p. \quad \frac{1}{2} \end{cases} \tag{6}$$

In [3], optimal policies for the problem under open-loop information structure were given. It was shown that all open-loop admissible policies lead to the same cost. In the next section, we consider the class of closed-loop admissible policies Π and provide the optimal strategies for the case $N = 2$.

IV. Optimal policy for $N = 2$

For the case of $N = 2$, we define the state of the system as follows. Let $w_k = a_{k1} a_{k2}$, and let $V_k(n_k^+, n_k^-)$ denote the value function at the k^{th} step for $0 \leq k \leq L$, where n_k^+ and n_k^- are the numbers of $+1$'s and -1's, respectively, that have been accepted until step k. The number of measurements remaining at step k is given by $(M - n_k^+ - n_k^-)$. Since the decision at step k is made after $w_k \in \{-1, +1\}$ is observed, the dynamic programming equation and the optimal decision are as follows:

$$V_k(n_k^+, n_k^-) = E_{w_k}\left[\min_{\sigma_k} V_{k+1}\left(n_k^+ + \sigma_k \frac{1+w_k}{2}, n_k^- + \sigma_k \frac{1-w_k}{2}\right)\right] \tag{7}$$

$$\sigma_k^* = \underset{\sigma_k}{\operatorname{argmin}} \left[V_{k+1}\left(n_k^+ + \sigma_k \frac{1+w_k}{2}, n_k^- + \sigma_k \frac{1-w_k}{2} \right) \right]$$

The boundary condition is given as

$$V_L(n_L^+, n_L^-) = \left| n_L^+ - n_L^- \right|. \tag{8}$$

Since an admissible policy accepts only M rows, (8) is equivalent to

$$V_k(n_k^+, M - n_k^+) = \left| M - 2n_k^+ \right| = \left| M - 2n_k^- \right|. \tag{9}$$

Theorem 1: For even M, the optimal policy is given by

$$\mu_k^* = \begin{cases} \sigma_k = 1 & \text{if } n_k^+ < \frac{M}{2}, w_k = +1, n_k^+ + n_k^- < M \\ \sigma_k = 1 & \text{if } n_k^- < \frac{M}{2}, w_k = -1, n_k^+ + n_k^- < M \\ \sigma_k = 1 & \text{if } M - n_k^+ - n_k^- = L - k \\ \sigma_k = 0 & \text{otherwise} \end{cases}$$

Proof: In order to prove optimality, we explicitly provide an expression for the value function. We prove that the function satisfies the dynamic programming equation (7), along with the boundary conditions. Finally, we compute the optimal policy based on the proposed value function from the dynamic programming equation,

The value function for the problem is given by the following,

$$V_k(n_k^+, n_k^-) =$$

$$\begin{cases} p_k \sum_{i=0}^{\frac{M}{2}-n_k^+} \binom{L-k}{i} (M - 2n_k^+ - 2i) & \text{if } n_k^+ \le \frac{M}{2}, \\ +p_k \sum_{i=0}^{\frac{M}{2}-n_k^-} \binom{L-k}{i} (M - 2n_k^- - 2i) & n_k^- \le \frac{M}{2} \\ & \\ p_k \sum_{i=0}^{M-n_k^+-n_k^-} \binom{L-k}{i} (M - 2n_k^- - 2i) & \text{if } n_k^+ \ge \frac{M}{2}, \\ +p_k \sum_{i=M-n_k^+-n_k^-+1}^{L-k} \binom{L-k}{i} (2n_k^+ - M) & n_k^- \le \frac{M}{2} \\ & \\ p_k \sum_{i=0}^{M-n_k^+-n_k^-} \binom{L-k}{i} (M - 2n_k^+ - 2i) & \text{if } n_k^+ \le \frac{M}{2}, \\ +p_k \sum_{i=M-n_k^+-n_k^-+1}^{L-k} \binom{L-k}{i} (2n_k^- - M) & n_k^- \ge \frac{M}{2} \end{cases}$$

where $p_k = 2^{-(L-k)}$ for even M values.

To verify that $V_k(n_k^+, n_k^-)$ is indeed the value function, first note that at the boundary, $n_k^+ + n_k^- = M$. Thus, either $n_k^+ \ge \frac{M}{2}$ or $n_k^- \ge \frac{M}{2}$. Assume we have the former. Then, as $M - 2n_k^- = 2n_k^+ - M$,

$$\begin{aligned} V(n_k^+, M - n_k^+) &= 2^{-(L-k)} \sum_{i=0}^{L-k} \binom{L-k}{i} (2n_k^+ - M) \\ &= \left| 2n_k^+ - M \right| \end{aligned}$$

and (9) is obtained. The latter, $n_k^- \ge \frac{M}{2}$, is the symmetric case and again yields (9), showing that the boundary condition is satisfied.

When $n_k^+ + n_k^- = M$, accepting a new row renders the policy inadmissible. On the other hand, when $M - n_k^+ - n_k^- = L - k$, i.e., the number of remaining steps is equal to the number of measurements to be accepted, the policy becomes inadmissible if any new measurement is rejected. Therefore, to show that $V_k(n_k^+, n_k^-)$ obeys the dynamic programming equation (7), it is sufficient to consider the following nontrivial cases:

Case (i): $n_k^+ < \frac{M}{2}, n_k^- < \frac{M}{2}$, which imply $n_k^+ + 1 \le \frac{M}{2}$ and $n_k^- + 1 \le \frac{M}{2}$, and $n_k^+ + n_k^- < M$:

(a) $w_k = +1$

$$V_{k+1}(n_k^+, n_k^-) - V_{k+1}(n_k^+ + 1, n_k^-) = 2^{-(L-k-2)} \sum_{i=0}^{\frac{M}{2}-1-n_k^+} \binom{L-k-1}{i} > 0$$

$$V_{k+1}(n_k^+ + 1, n_k^-) < V_{k+1}(n_k^+, n_k^-) \implies \sigma_k^* = 1 \tag{10}$$

(b) $w_k = -1$

$$V_{k+1}(n_k^+, n_k^-) - V_{k+1}(n_k^+, n_k^- + 1) = 2^{-(L-k-2)} \sum_{i=0}^{\frac{M}{2}-1-n_k^-} \binom{L-k-1}{i} > 0$$

$$V_{k+1}(n_k^+, n_k^- + 1) < V_{k+1}(n_k^+, n_k^-) \implies \sigma_k^* = 1 \tag{11}$$

Inserting (10) and (11) into the right hand side of (7):

$$\begin{aligned} &E_{w_k} \left[\min_{\sigma_k} V_{k+1} \left(n_k^+ + \sigma_k \tfrac{1+w_k}{2}, n_k^- + \sigma_k \tfrac{1-w_k}{2} \right) \right] \\ &= \tfrac{1}{2} V_{k+1}(n_k^+ + 1, n_k^-) + \tfrac{1}{2} V_{k+1}(n_k^+, n_k^- + 1) \end{aligned} \tag{12}$$

The summations containing n_k^+ in (12) yield

$$\begin{aligned} &\tfrac{1}{2} 2^{-(L-k-1)} \textstyle\sum_{i=0}^{\frac{M}{2}-n_k^+-1} \binom{L-k-1}{i} (M - 2n_k^+ - 2 - 2i) \\ &\quad + \tfrac{1}{2} 2^{-(L-k-1)} \textstyle\sum_{i=0}^{\frac{M}{2}-n_k^+} \binom{L-k-1}{i} (M - 2n_k^+ - 2i) \end{aligned}$$

$$\begin{aligned} &= 2^{-(L-k)} \textstyle\sum_{i'=1}^{\frac{M}{2}-n_k^+} \binom{L-k-1}{i'-1} (M - 2n_k^+ - 2i') \\ &\quad + 2^{-(L-k)} \textstyle\sum_{i=0}^{\frac{M}{2}-n_k^+} \binom{L-k-1}{i} (M - 2n_k^+ - 2i) \end{aligned} \tag{13}$$

$$\begin{aligned} &= 2^{-(L-k)} \textstyle\sum_{i=1}^{\frac{M}{2}-n_k^+} \binom{L-k}{i} (M - 2n_k^+ - 2i) \\ &\quad + 2^{-(L-k)} \binom{L-k-1}{0} (M - 2n_k^+) \end{aligned} \tag{14}$$

$$= 2^{-(L-k)} \textstyle\sum_{i=0}^{\frac{M}{2}-n_k^+} \binom{L-k}{i} (M - 2n_k^+ - 2i) \tag{15}$$

where the relation between (13) and (14) follows from the identity

$$\binom{L-k-1}{i} + \binom{L-k-1}{i-1} = \binom{L-k}{i} \tag{16}$$

and that between (14) and (15) from

$$\binom{L-k-1}{0} = \binom{L-k}{0}. \tag{17}$$

(15) is identical to the summation with n_k^+ in $V_k(n_k^+, n_k^-)$. Similarly, the summations with n_k^- in (12) and that in $V_k(n_k^+, n_k^-)$ are found to be equal, verifying (7) for case (i).

Case (ii): $n_k^+ \ge \frac{M}{2}, n_k^- \le \frac{M}{2}$, $n_k^+ + n_k^- < M$ and $M - n_k^+ - n_k^- < L - k$:

(a) $w_k = +1$

$$V_{k+1}(n_k^+ + 1, n_k^-) - V_{k+1}(n_k^+, n_k^-) = 2^{-(L-k-2)} \sum_{i=M-n_k^+-n_k^-}^{L-k-1} \binom{L-k-1}{i} > 0$$

$$V_{k+1}(n_k^+, n_k^-) < V_{k+1}(n_k^+ + 1, n_k^-) \implies \sigma_k^* = 0 \quad (18)$$

(b) $w_k = -1$

$$V_{k+1}(n_k^+, n_k^-) - V_{k+1}(n_k^+, n_k^- + 1) =$$
$$2^{-(L-k-2)} \sum_{i=0}^{M-n_k^+-n_k^--1} \binom{L-k-1}{i} > 0$$

$$V_{k+1}(n_k^+, n_k^- + 1) < V_{k+1}(n_k^+, n_k^-) \implies \sigma_k^* = 1 \quad (19)$$

Employing (18) and (19) in (7), we have

$$\begin{aligned} &E_{w_k}\left[\min_{\sigma_k} V_{k+1}\left(n_k^+ + \sigma_k \tfrac{1+w_k}{2}, n_k^- + \sigma_k \tfrac{1-w_k}{2}\right)\right] \\ &= \tfrac{1}{2} V_{k+1}(n_k^+, n_k^-) + \tfrac{1}{2} V_{k+1}(n_k^+, n_k^- + 1) \\ &= V_k(n_k^+, n_k^-) \end{aligned}$$

where, again, the identities (16) and (17) are used in the last equality.

Case (iii): $n_k^+ \leq \frac{M}{2}, n_k^- \geq \frac{M}{2}$, $n_k^+ + n_k^- < M$ and $M - n_k^+ - n_k^- < L - k + 1$:

As this case is symmetric to the case (ii),

(a) $w_k = +1 \implies \sigma_k^* = 1$

(b) $w_k = -1 \implies \sigma_k^* = 0$

and (7) holds.

It has been shown that $V_k(n_k^+, n_k^-)$ satisfies (7) as well as (9) for all (nontrivial) cases along with the optimal decisions forming the decision rule μ_k^* of the Theorem 1. ■

Theorem 2: The optimal policy for odd M is given by

$$\mu_k^* = \begin{cases} \sigma_k = 1 & \text{if } n_{k-1}^+ < \frac{1}{2}(M-1), w_k = +1 \\ \sigma_k = 1 & \text{if } n_{k-1}^- < \frac{1}{2}(M-1), w_k = -1 \\ \sigma_k \text{ is arbitrary} & \text{if } n_{k-1}^+ = \frac{1}{2}(M-1),\ w_k = +1 \\ \sigma_k \text{ is arbitrary} & \text{if } n_{k-1}^- = \frac{1}{2}(M-1),\ w_k = -1 \\ \sigma_k = 0 & \text{otherwise} \end{cases}$$

Proof: The following cost function

$$V_k(n_k^+, n_k^-) =$$
$$\begin{cases} p_k \sum_{i=0}^{\frac{M-1}{2}-n_k^+} \binom{L-k}{i}(M - 2n_k^+ - 2i) & \text{if } n_k^+ \leq \frac{M-1}{2} \\ +p_k \sum_{i=\frac{M+1}{2}-n_k^+}^{L-k-\frac{M+1}{2}+n_k^-} \binom{L-k}{i} & n_k^- \leq \frac{M-1}{2} \\ +p_k \sum_{i=0}^{\frac{M-1}{2}-n_k^-} \binom{L-k}{i}(M - 2n_k^- - 2i) & \\ & \\ p_k \sum_{i=0}^{M-n_k^+-n_k^-} \binom{L-k}{i}(M - 2n_k^- - 2i) & \text{if } n_k^+ \geq \frac{M+1}{2} \\ +p_k \sum_{i=M-n_k^+-n_k^-+1}^{L-k} \binom{L-k}{i}(2n_k^+ - M) & n_k^- \leq \frac{M-1}{2} \\ & \\ p_k \sum_{i=0}^{M-n_k^+-n_k^-} \binom{L-k}{i}(M - 2n_k^+ - 2i) & \text{if } n_k^+ \leq \frac{M-1}{2} \\ +p_k \sum_{i=M-n_k^+-n_k^-+1}^{L-k} \binom{L-k}{i}(2n_k^- - M) & n_k^- \geq \frac{M+1}{2} \end{cases}$$

where $p_k = 2^{-(L-k)}$, satisfies (7) and (9) for odd M along with μ_k^* given in Theorem 2. This can be verified similar to the value function for the previous case when M is even. ■

For even M, with the policy provided in Theorem 1, the resulting correlation of the columns becomes zero if at least $\frac{M}{2}$ $+1$'s and $\frac{M}{2}$ -1's appear in the sequence. If, for example, enough -1's do not appear, the policy accepts all -1's that come up, keeping the resulting correlation as close to 0 as possible. Therefore, the policy ensures the smallest correlation that can be obtained from any given sequence. The same holds when M is odd, but in this case the minimum correlation that can be achieved is 1.

V. Extension to Higher Dimensions

In this section, we extend the optimal policy proposed in the previous section to arbitrary N. Based on the notations introduced in Section II, we can write the following equation

$$y^k = A_k x$$

where $y^k = [y_1, \ldots, y_k]'$, and A_k is obtained by stacking the vectors a^i as rows for $i = 1, \ldots, k$. Define g_{lm}^j as follows:

$$g_{lm}^j = \sum_{r=1}^{j} a_{rl} a_{rm}$$

If we consider two specific columns, i and j in A_k, then we can define μ_{ij}^{k*} and σ_{ij*}^{k*} to be the optimal policy and the optimal action, respectively, corresponding to the two columns based on the analysis in the previous section. Let

$$S^k = \{(i,j) | (i,j) = \arg\max_{m,n} g_{mn}^k\}$$

$$S_0^k = \{(i,j) | [(i,j) \in S^k] \bigwedge [\sigma_{ij}^{k*} = 0]\}$$

$$S_1^k = \{(i,j) | [(i,j) \in S^k] \bigwedge [\sigma_{ij}^{k*} = 1]\}$$

The policy we propose is as follows:

$$\mu_k = \begin{cases} \mu_{(\arg\max_{ij} g_{ij}^k)}^{k*}, & |S^k| = 1 \\ 0, & |S^k| > 1, |S_0^k| > \frac{S^k}{2} \\ 1, & |S^k| > 1, |S_1^k| \geq \frac{S^k}{2} \end{cases},$$

where μ_k^* is the policy at stage k, and $|\cdot|$ denotes the cardinality of the set. For $N = 2$, μ_k coincides with μ_k^*. For the case of general N, tight bounds on the performance of μ_k is a topic of our ongoing research. In this paper, we illustrate the performance of μ_k by presenting some simulation results.

Figure 1 shows the simulation of the implementation of μ_k for a scenario in which $N = 1000, L = 500$ and $M = 400$. The plot on the left shows the sequential variation of $|g^k|_\infty$ of A_k. The plot on the right shows σ_k as a function of time. The plot on the right in Figure 2 shows the histogram of $|g^L|_\infty$ for 100 simulation runs when one implements the open loop optimal policy proposed in [3], whereas the plot on the left is the histogram obtained when μ_k is used as the activation policy. Since the weights of large correlations values have decreased, μ_k shows some improvement over the open-loop optimal policy. However, for very small values of M compared to N, the following lemma shows that the final coherence is unaffected by the activation policy.

Lemma 1: If $N > 2^{M-1}$, the coherence of the final matrix is 1 irrespective of the activation policy.

Proof: Consider the matrix $A \in \{-1, +1\}^{M \times N}$ that is obtained at the end of the process. With M rows, only 2^M

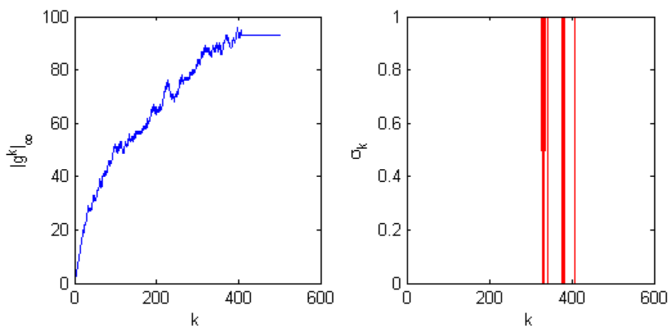

Fig. 1. The variation of $|g^k|_\infty$ [left] and σ_k [right] with k for $N = 1000, L = 500, M = 400$

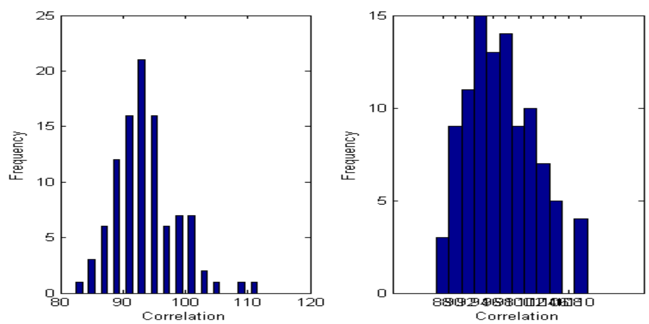

Fig. 2. The histogram of the coherence of the final sensing matrix for μ_k [left] and for the open-loop policy proposed in [3] [right] as a result of 100 runs with $N = 1000$ and $M = 400$

distinct columns can be generated; however, half of these columns are negative of the other half. Hence, if $N > 2^{M-1}$, there is at least one pair of identical columns or one column is negative of another. ■

VI. Conclusion

In this paper, we have considered the problem of optimal measurement policies for a sensor that acquires sequential compressive measurements of a static vector as originally formulated in [3]. We obtained closed-loop optimal measurement policies for low-dimensional problems in contrast to the previous work [3] that had proposed a greedy strategy. Finally, we have generalized the optimal policy to obtain a feasible policy for acquiring arbitrary length sparse vectors of unknown sparsity. The performance of the proposed policy has been illustrated through simulation results. In the future, we plan to extend this work to dynamic signals with bounded variations in sparsity.

References

[1] MURI - Opportunistic sensing. http://dsp.rice.edu/muri31.

[2] R. G. Baraniuk. More is less: Signal processing and the data deluge. *Science*, 331(6018):717–719, Feb 2011.

[3] S. Bhattacharya, A. Nayyar, and T.Başar. Optimal acquisition policy for compressed measurements with limited observations. In *Asilomar Conference on Signals, Systems and Computation*, Asilomar, California, November 2012.

[4] E. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, 2006.

[5] E. Candés and T. Tao. Decoding by linear programming. *Transaction on Information Theory*, 51(12):4203–4215, 2005.

[6] R. Castro, J. Haupt, R. Nowak, and G. Raz. Finding needles in noisy haystacks. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 5133 – 5136, Las Vegas, Nevada, March 2008.

[7] V. Cevher, A. Sankaranarayanan, M. Duarte, D. Reddy, R. Baraniuk, and R. Chellappa. Compressive sensing for background subtraction. In *European Conference on Computer Vision (ECCV)*, 2008.

[8] S. Sanghavi D. Malioutov and A. Willsky. Sequential compressed sensing. *IEEE Journal of Selected Topics in Signal Processing*, 4(2):435–444, 2010.

[9] D. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 6(4):1289–1306, 2006.

[10] D. Donoho and M. Elad. Optimally sparse representation in general (nonorthogonal) dictionaries via l1 minimization. *Proceedings of National Academy of Sciences*, 100:2197–2202, March 2003.

[11] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk. Single-pixel imaging via compressive sampling. *IEEE Signal Processing Magazine*, 25(2):83–91, March 2008.

[12] O. C. Imer and T. Başar. Optimal estimation with limited measurements. *IEEE Journal on Systems, Control and Communication*, 2(1/2/3):5–29, 2010.

[13] R. Castro J. Haupt and R. Nowak. Distilled sensing: Selective sampling for sparse signal recovery. In *Proc. 12th Conference on Artificial Intelligence and Statistics*, Clearwater Beach, Florida, April 2009.

[14] R. Castro J. Haupt and R. Nowak. Distilled sensing: Adaptive sampling for sparse detection and estimation. *IEEE Transactions on Information Theory*, 57(9):6222–6235, September 2011.

[15] R. Castro J. Haupt, R. Baraniuk and R. Nowak. Sequentially designed compressive sensing. In *IEEE Workshop on Statistical Signal Processing*, Ann Arbor, Michigan, August 2012.

[16] R. Castro J. Haupt, R.Baraniuk and R. Nowak. Compressed distilled sensing. In *43th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, California, October 2009.

[17] R. Nowak J. Haupt and R. Castro. Adaptive sensing for sparse recovery. In *Proc. IEEE Digital Signal Processing Workshop and Workshop on Signal Processing Education*, pages 702 – 707, Marco Island, FL, January 2009.

[18] A. Nayyar, T. Başar, D. Teneketzis, and V. V. Veeravalli. Communication scheduling and remote estimation with energy harvesting sensor. In *51st IEEE Conference on Decision and Control*, pages 843–848, December 2012.

[19] P.Bommannavar and T. Başar. Optimal estimation over channels of limited usage. In *Proceedings of the 17th IFAC World Congress*, Seoul, Korea, July 2008.

[20] F. Raisali and N. Vaswani. Stability (over time) of regularized modified-cs (noisy) for recursive causal sparse reconstruction. In *45th Annual Conference on Information Sciences and Systems, 2011*, Baltimore, March 2011.

[21] Y. Xue S. Ji and L. Carin. Bayesian compressive sensing. *IEEE Transactions on Signal Processing*, 56(6):2346–2356, June 2008.

[22] J. Tropp and A. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on Information Theory*, 53(12):4655–4666, December 2007.

[23] N. Vaswani. LS-CS-residual (LS-CS): Compressive sensing on least squares residual. *IEEE Transactions on Signal Processing*, 58(8):4108–4120, August 2010.

[24] N. Vaswani and W. Lu. Modified-CS: Modifying compressive sensing for problems with partially known support. *IEEE Transactions on Signal Processing*, 58(9):4595–4607, September 2010.

[25] R. Ward. Compressed sensing with cross validation. *IEEE Transactions on Information Theory*, 55(12):5773–5782, 2009.

[26] G. Warnell, D. Reddy, and R.Chellappa. Adaptive rate compressive sensing for background subtraction,. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 1477 – 1480, Kyoto, Japan, March 2012.