# Cost-Aware Activity Scheduling for Compressive Sleeping Wireless Sensor Networks

Wei Chen, *Member, IEEE* and Ian J. Wassell

***Abstract*—In this paper, we consider a compressive sleeping wireless sensor network (WSN) for monitoring parameters in the sensor field, where only a fraction of sensor nodes (SNs) are activated to perform the sensing task and their data are gathered at a fusion centre (FC) to estimate all the other SNs' data using the compressive sensing (CS) principle. Typically research published concerning CS implicitly assume the sampling costs for all samples are equal, and suggest random sampling as an appropriate approach to achieve good reconstruction accuracy. However, this assumption does not hold for compressive sleeping WSNs which have significant variability in sampling cost owing to the different physical conditions at particular SNs. To exploit this sampling cost non-uniformity, we propose a cost-aware activity scheduling approach that minimizes the sampling cost with constraints on the regularized mutual coherence of the equivalent sensing matrix. In addition, for the case with prior information about the signal support, we extend the proposed approach to incorporate the prior information by considering an additional constraint on the mean square error (MSE) of the oracle estimator for sparse recovery. Our numerical experiments demonstrate that in comparison with other designs in the literature the proposed activity scheduling approaches lead to improved trade-offs between reconstruction accuracy and sampling cost for compressive sleeping WSNs.**

***Index Terms*—Compressive sensing (CS), wireless sensor network (WSN), activity scheduling.**

## I. Introduction

WIRELESS sensor networks (WSNs) typically consist of a large number of spatially distributed sensor nodes (SNs) that measure, collect and process information of interest for some target area. With appropriate temporal and spatial scales, they have been used in many applications such as climate, habitat, and infrastructure monitoring. The successful deployment of WSNs has two main challenges. Firstly, as the number of SNs increases, a large amount of data needs to be processed, transported, and stored at the fusion centre (FC). Secondly, SNs are limited in terms of energy availability, computational capability and wireless bandwidth.

Compressive sensing (CS) is a sampling paradigm that takes advantage of the sparse characteristic of the natural physical signals of interest, and makes it possible to recover signals with a reduced number of random samples [1], [2]. In view of the temporal correlation and/or spatial correlation among the densely deployed SNs, CS can be used as a data acquisition technique to reduce the operating cost of WSNs, and various CS-based schemes have been proposed

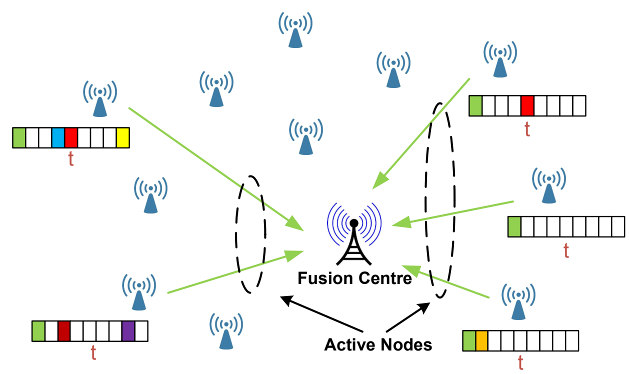


Fig. 1. Illustration of a compressive sleeping wireless sensor network where only a fraction of SNs are active in each time slot.

to implement this idea [3]–[9]. For example, Bajwa et al. propose compressive wireless sensing (CWS) where SNs coherently transmit weighted sample values in an analog fashion over the network-to-FC communication channel [3]. Wang et al. propose distributed sparse random projections which allows the collector to recover a sparse data approximation by querying a sufficient number of sensors from anywhere in the network [4]. Luo et al. propose an approach, namely compressive data gathering (CDG), to reduce the total number of message transmissions in a multi-hop routing scheme [5]. Quer et al. propose a combined CS and principal component analysis (PCA) technique to recover WSNs' signals using only a small number samples [6]. In [8], [9], Yang et al. and Chen et al. propose different CS schemes to achieve energy neutrality for energy harvesting (EH) WSNs where the sampling is constrained to suit the EH conditions.

The conventional CS framework considers the sparse characteristics of signals and employs a random sampling scheme. Built upon this foundation, various extensions of CS have been proposed to enhance the performance in terms of the reconstruction accuracy or the number of required samples for successful reconstruction. One type of extension is to consider a more restricted signal structure which goes beyond sparsity, e.g., the block sparsity and the wavelet tree model [10], or some prior knowledge about the signal [11], [12]. Another type of extension is to use some optimized sensing matrices instead of random ones. Recent works [13]–[17] have demonstrated that well-designed sensing matrices provide improved reconstruction performance in comparison to the random sensing matrix. In addition, prior knowledge about the signal can also be exploited in sensing matrix designs with further improved performance [18], [19].

In this paper, we consider a compressive sleeping WSN as

Wei Chen is with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China, and also with the Computer Laboratory, University of Cambridge, Cambridge CB3 0FD, United Kingdom (e-mail: wc253@cam.ac.uk).

Ian J. Wassell is with the Computer Laboratory, University of Cambridge, Cambridge CB3 0FD, United Kingdom (e-mail: ijw24@cam.ac.uk).

shown in Fig. 1, where only a fraction of SNs are active in each time slot to measure and transmit the information about the physical phenomena of interest [20], [21]. Owing to the spatial correlation of the monitored phenomena which leads to spatial sparsity under some appropriate basis, information of SNs in the sleep mode can be recovered following the CS principle. This CS-based data acquisition approach requires less wireless bandwidth resource and consumes less of the SNs' battery energy in comparison to the traditional approach that collects and sends all SNs' data. These active SNs can be randomly selected according to the CS principle. However, this random activity scheduling scheme has two drawbacks. Firstly, the random scheme is cost-blind, as it follows the conventional framework that implicitly assumes the same resource cost for every SN. However, this assumption is not necessarily realistic. For example, less energy and/or bandwidth resource are required to communicate messages from a SN which is close to the FC than for an SN that is far away from the FC. It is also desired to activate SNs having adequate energy reserves rather than these that are running out of energy. Secondly, the random activity scheduling scheme fails to exploit any prior knowledge about the signal beyond sparsity. For example, the support of the spatial signal of interest could change slowly in time owing to the temporal correlation of the physical phenomena.

The sensing matrix design approaches proposed in [13]–[16] cannot be directly applied to derive an appropriate activity scheduling scheme for compressive sleeping WSNs, which not only involves binary integer variables but also accounts for resource cost non-uniformity. In [22], Xu et al. propose a cost-aware compressive sensing approach that can be used for activity scheduling compressive sleeping WSNs. However, their approach is restricted to the Fourier basis and the assumption of a uniform distribution of the sparse signal representation among all sparse vectors. In [21], [23], Chen and Wassell consider the scenario of a slowly time-varying signal support and provide an activity scheduling design which is optimized according to the signal support of the previous time slot. The effectiveness of this approach is degraded when a fraction of the support is changed in two adjacent time slots.

In this paper, we propose a method for cost-aware activity scheduling in compressive sleeping WSNs, which is not restricted to a particular signal sparsifying basis and can incorporate prior information of the signal support. The proposed approach improves the trade-offs between reconstruction accuracy and sampling cost of the activated SNs. The contributions can be summarized as follows:

- We propose the regularized mutual coherence, i.e., an alternative measure that is used to characterize the upper bound of the CS reconstruction error. The use of the regularized mutual coherence facilitates our design of an appropriate scheme, in view of the fact that the signal reconstruction accuracy of CS systems is not tractable.
- We formulate the activity scheduling as an optimization problem that minimizes the sampling cost with constraints on the reconstruction accuracy indicators, and approximate the problem by a constrained convex relaxation plus a rounding scheme.
- We extend the proposed approach to the case where knowledge of the signal support probability is available. We characterize the quality of the activity scheduling matrix by using the mean square error (MSE) of the oracle estimator that performs ideal least squares (LS) estimation based on prior knowledge of the signal support. The MSE of this oracle LS estimator coincides with the unbiased Cramèr-Rao bound for sparse deterministic vectors [24], so that it represents the best achievable performance for any unbiased estimator.
- The superiority of the proposed approach is demonstrated by numerical experiments.

The rest of this paper is organized as follows. We begin by describing wireless sleeping WSNs and related works in Section II. Section III provides the rationale for the activity scheduling matrix designs, by highlighting the regularized mutual coherence as the performance indicator, and describes a convex relaxation approach to approximately solve the design problem. In Section IV, the proposed framework is extended to the case where knowledge of signal support probability is available and we use the oracle estimator MSE as a performance indicator. Section V presents numerical results that highlight the merits of our proposed designs in comparison to other designs in the literature. The main contributions of the article are summarized in Section VI.

Throughout this paper, lower-case letters denote scalars, boldface upper-case letters denote matrices, bold face lower-case letters denote column vectors, and calligraphic upper-case letters denote support sets. $\mathbf{0}_{m\times n}$ and $\mathbf{1}_{m\times n}$ denote an $m \times n$ matrix with all zeros and all ones, respectively. The superscripts $(\cdot)^T$ and $(\cdot)^{-1}$ denote matrix transpose and matrix inverse, respectively. The $\ell_0$ norm, the $\ell_1$ norm, and the $\ell_2$ norm of vectors, are denoted by $\|\cdot\|_0$, $\|\cdot\|_1$, and $\|\cdot\|_2$, respectively. The trace of a matrix is denoted by $\mathtt{Tr}(\cdot)$. $\mathtt{diagm}(\mathbf{A})$ denotes a diagonal matrix corresponding to matrix $\mathbf{A}$, and $\mathtt{diag}(\mathbf{A})$ denotes a vector composed of the diagonal elements of $\mathbf{A}$. $\mathtt{nondiagm}(\mathbf{A})$ denotes a matrix such that $\mathtt{nondiagm}(\mathbf{A}) = \mathbf{A} - \mathtt{diagm}(\mathbf{A})$. The element corresponding to the *i*th row and *j*th column of the matrix $\mathbf{A}$ is denoted by $a_{i,j}$, and $\mathbf{a}_i$ denotes the *i*th column of the matrix $\mathbf{A}$. $\mathbf{I}_n$ denotes the $n \times n$ identity matrix. For a vector $\mathbf{x}$, the notation $\mathbf{x}_{\mathcal{J}}$ refer to a sub-vector that contains the elements with indexes in $\mathcal{J}$. $\mathcal{J}^c$ denotes the complementary set of $\mathcal{J}$. $\mathbf{E}_{\mathcal{J}}$ denotes the matrix that results from the identity matrix by deleting the set of columns out of the support $\mathcal{J}$. $\mathbb{E}(\cdot)$ denotes the expectation, $\mathbb{E}_{\mathbf{x}}(\cdot)$ and $\mathbb{E}_{\mathcal{J}}(\cdot)$ denote expectation with respect to the distribution of the random vector $\mathbf{x}$, and the random support $\mathcal{J}$, respectively. $\binom{n}{m}$ denotes the number of $m$ combinations from a given set of $n$ elements. Finally, $\Pr(\cdot)$ denotes the probability.

## II. System Description and Background

In this section, we provide a system description of compressive sleepling WSNs and related work in CS sensing matrix design.

### A. CS Reconstruction

For a data vector $\mathbf{f} \in \mathbb{R}^n$ that can be represented by a sparse vector $\mathbf{x} \in \mathbb{R}^{\hat{n}}$ $(n \leq \hat{n})$ by $\mathbf{f} = \mathbf{\Psi}\mathbf{x}$, the CS measurement vector is given by

$$\mathbf{y} = \mathbf{\Phi}\mathbf{f} + \mathbf{z} = \mathbf{A}\mathbf{x} + \mathbf{z}, \tag{1}$$

where $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}_m)$ represents the sensing noise, $\mathbf{\Phi} \in \mathbb{R}^{m\times n}$ and $\mathbf{A} = \mathbf{\Phi}\mathbf{\Psi} \in \mathbb{R}^{m\times \hat{n}}$ denotes the sensing matrix and the equivalent sensing matrix, respectively, and $m < n$. One typical strategy for recovering the sparse signal representation $\mathbf{x}$ is to cast the problem as an optimization problem, and the sparsity level $\|\mathbf{x}\|_0$, i.e., a nonconvex term, is relaxed by the $\ell_1$ norm of $\mathbf{x}$. Thus, we consider the following optimization problem:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathbf{x}\|_1 \\ \text{s.t.} \quad & \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 \leq \epsilon, \end{aligned} \tag{2}$$

where $\epsilon$ is an estimate of the noise level. This optimization problem is also known as basis pursuit de-noise (BPDN). Define $s = \|\mathbf{x}\|_0$ as the sparsity level of the true signal. It has been demonstrated that only $\mathcal{O}(s \log \frac{\hat{n}}{s})$ measurements are required for robust CS reconstruction with randomly generated sensing matrices [25].

The conventional CS framework only exploits the sparse characteristics of the signal in the reconstruction. A recent growing trend relates to the use of additional signal structures that go beyond the simple sparsity model to further enhance the performance of CS. For example, Vaswani and Lu consider a time sequence of signals, and exploit the support of the signal in a previous time slot as side information to enhance the reconstruction of the current signal [11] via solving the following optimization problem

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathbf{x}_{\mathcal{J}^c}\|_1 \\ \text{s.t.} \quad & \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 \leq \epsilon. \end{aligned} \tag{3}$$

In [12], prior information of the signal support is exploited to assist the reconstruction. To be specific, defining $p_i$ as the probability that the entry $x_i$ is non-zero, Scarlett et al. propose to solve the following optimization problem

$$\begin{aligned} \min_{\mathbf{x}} \quad & \sum_{i=1}^{\hat{n}} (-\log p_i)|x_i| \\ \text{s.t.} \quad & \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 \leq \epsilon, \end{aligned} \tag{4}$$

and demonstrate that significantly fewer samples are required for successful reconstruction if the prior distribution is sufficiently non-uniform [12].

### B. Compressive Data Gathering

We consider a typical WSN architecture with $n$ SNs and a FC. Assume the coherence time of the monitored physical phenomena is $t_{\text{coh}}$ so that the $i$th SN's readings satisfy $f_i(t_1) \approx f_i(t_2)$ for two distinct time instances $t_1$ and $t_2$ with $|t_1 - t_2| \leq t_{\text{coh}}$. We focus on an observation window of $t_{\text{obs}}$ which satisfies $t_{\text{obs}} < t_{\text{coh}}$, and drop the time index. Owing to the spatial correlation, the data vector $\mathbf{f} \in \mathbb{R}^n$ can be represented by a sparse vector $\mathbf{x} \in \mathbb{R}^{\hat{n}}$ $(n \leq \hat{n})$ and $\mathbf{f} = \mathbf{\Psi}\mathbf{x}$, where the linear transform matrix $\mathbf{\Psi} \in \mathbb{R}^{n\times \hat{n}}$ can be determined a-priori or learned from some training data [26]. For example, in a dense sensor network, a smooth temperature signal or a humidity signal, i.e., $\mathbf{f}$, can be represented by only a few frequency components in the Fourier basis $\mathbf{\Psi}$, and the frequency components comprise the nonzero elements of $\mathbf{x}$. According to the theory of CS, only $m$ $(m < n)$ SNs are required to perform the sensing task and transmit their readings to the FC to reconstruct $\mathbf{f}$, while SNs that are not scheduled to perform the sensing task can go into the sleep mode to reduce resource use.

The activated SNs measure the signal of interest and encode their measurements into a packet which is then modulated and transmitted to the FC in a conflict-free manner, e.g., via time division multiplexing access (TDMA) or frequency division multiplexing access (FDMA). The received measurement vector $\mathbf{y} \in \mathbb{R}^m$ at the FC can be expressed as (1), where $\mathbf{\Phi} \in \mathbb{R}^{m\times n}$ denotes the activity matrix. The rows of the activity matrix $\mathbf{\Phi}$ can be regarded as $m$ rows of an $n \times n$ identity matrix, i.e., the entries are all zeros except for $m$ entries in $m$ different columns and rows, where the columns with 1s correspond to the active nodes.

### C. Cost-Aware Activity Scheduling: Motivations

Conventional CS implicitly assumes that all measurements have the same cost, which is suitable for many applications such as magnetic resonance imaging (MRI) and compressive radar. However, the cost for conducting measurements and communication at different SNs in a WSN can vary significantly. The sampling cost of a compressive sleeping WSN consists of various factors including but not limited to energy consumption, wireless spectrum resource and system considerations. For example, owing to the different channel conditions experienced at the SNs, the costs in terms of energy consumption and bandwidth are not uniform for the SNs. In addition, from a system point of view, to sustain the function of the network with a lifetime as long as possible, a higher priority of activity is desired for SNs with sufficient energy reserves, or those with greater energy harvesting capability. In this paper, we denote the sampling costs of different SNs by a vector $\mathbf{c} \in \mathbb{R}^n$, where various cost factors can be integrated into this vector. For example, if energy consumption is the main concern, the sampling cost of each node can be estimated by using an appropriate path loss model associated with the distance from the FC to each SN, or by using a feedback scheme where the frequency of feedback depends on the trade-off between overhead and the degree that the channel condition varies. However, if crowdsourced signal strength data for the monitored area is available, the dynamic sampling cost can be obtained at almost zero cost.

The sampling costs that are owing to the limitations of the devices and the physical environment, will often in practice have spatial and temporal correlations. It has been shown in [22] that activating SNs having the lowest costs does not necessarily improve the performance trade-off between reconstruction accuracy and sampling cost. Thus, it is worthwhile to pursue cost-aware activity scheduling approaches in order to provide a good performance trade-off for wireless sleeping WSNs.

### D. Previous Work on Sensing Matrix Design

A theoretical question in CS is what conditions should the equivalent sensing matrix $\mathbf{A}$ satisfy in order to guarantee the success of reconstruction. The most widely used conditions in the literatures include the null space property (NSP), the restricted isometry property (RIP) and mutual coherence. Although it has been proved that the NSP and the RIP can be satisfied for randomly generated matrices with a high probability [25], these two conditions are not computationally tractable and thus not suitable for evaluating the quality of a given sensing matrix design. The mutual coherence is defined as follows

$$\mu(\mathbf{A}) = \max_{1 \leq i,j \leq \hat{n}, i \neq j} \frac{|\mathbf{a}_i^T \mathbf{a}_j|}{\|\mathbf{a}_i^T\|_2 \ \|\mathbf{a}_j\|_2}, \tag{5}$$

which is straightforward to compute and thus is a convenient way to evaluate the quality of a given sensing matrix. In [27], Donoho et al. demonstrate that the $s$-sparse signal can be exactly recovered from the measurements in the noiseless case as long as

$$\mu(\mathbf{A}) < \frac{1}{2s-1}. \tag{6}$$

The earliest work in optimizing sensing matrix design is Elad's method given in [14]. In view of the fact that iteratively reducing the mutual coherence by adjusting the related pair of columns is not an efficient approach since it only improve the worst pair of columns in each iteration, Elad considers a different coherence indicator, called $t$-averaged mutual coherence, which is defined as the average of all normalized absolute inner products between different columns in the equivalent sensing matrix that are not smaller than a positive number $t$. Later in [15], [16], it is proposed to make the Gram matrix of the equivalent sensing matrix, defined as $\mathbf{A}^T\mathbf{A}$, as close as possible to an identity matrix, and the superiority of the designs is witnessed in their experimental results. In [13], [28], [29], it is proposed to force the equivalent sensing matrix to be a tight frame or close to an equiangular tight frame where the coherence value between any two columns of $\mathbf{A}$ are equal. These approaches implicitly assume a uniform distribution of the sparse signal representation among all sparse vectors. In [18], Zelnik-Manor et al. consider optimizing a sensing matrix for block sparse signals, and propose to minimize a weighted sum of the inter-block column coherence and the intra-block column coherence of $\mathbf{A}$. It is worth noting that these approaches aim to optimize sensing matrices with the assumption of equal sensing cost and no restrictions on the entries, while in our activity scheduling problem, the entries of the sensing matrix are binary values and the sampling costs of the SNs are non-uniform.

Several methods have been proposed for the activity scheduling [21]–[23], which are able to balance sampling cost and reconstruction accuracy. In [22], Xu et al. observe that the recovery accuracy with a partial Fourier matrix can be approximated by the regularized column sum (RCS), which is defined as $\max_{j=2,\ldots,\hat{n}} \log|\sum_i a_{i,j}|^2$, and propose to minimize the total sampling cost with constraints on the RCS. The connection between the recovery accuracy and the RCS is only demonstrated for the Fourier basis and the case that all sparse vectors are uniformly drawn from the space. In [21], [23], Chen and Wassell propose an optimized node selection (ONS) scheme to predict the reconstruction accuracy using the MSE of the LS estimation of the signal with an estimated signal support. Although improved performance is observed, this approach requires a relatively accurate knowledge of the signal support that is not always available.

## III. Proposed Activity Scheduling Approach

In this section, we first provide a rationale for the proposed activity scheduling design, and then describe a heuristic approach, based on convex relaxation, to approximately solve the design problem.

### A. Design Rationale

The sampling cost is straight-forward to compute for a given activity pattern, while it is difficult to efficiently and effectively quantify the reconstruction accuracy, since the signal reconstruction accuracy of CS systems is not tractable. In the literature [13]–[16], alternative measures such as those based upon mutual coherence are used to indicate the reconstruction performance for sensing matrix design. These mutual-coherence-based approaches assume that the columns of the equivalent sensing matrix can be normalized to a unit $\ell_2$-norm, and thus, instead of dealing with the mutual coherence, these approaches actually only need to consider the column correlation. However, the equivalent sensing matrix in a compressive sleeping WSN does not have equal-norm columns, as the activity scheduling imposes a restricted sampling operator $\mathbf{\Phi}$ which makes the equivalent sensing matrix comprise of a selection of rows of the sparsifying basis.

To obtain an appropriate indicator which facilitates the design of activity scheduling, we define the regularized mutual coherence by

$$\hat{\mu}(\mathbf{\Phi}, \mathbf{\Psi}) = \max_{1 \leq i,j \leq \hat{n}, i \neq j} |\hat{\mathbf{a}}_i^T \hat{\mathbf{a}}_j|, \tag{7}$$

where $\hat{\mathbf{A}} = \mathbf{\Phi}\hat{\mathbf{\Psi}}$, and $\hat{\mathbf{\Psi}}$ denotes a matrix that comprises columns of $\mathbf{\Psi}$ normalized to unit $\ell_2$-norm. The following theorem establishes the bound of the reconstruction accuracy of (2) using the newly defined regularized mutual coherence.

*Theorem 1:* Let $\hat{\mu}(\mathbf{\Phi}, \mathbf{\Psi})$ be the the regularized mutual coherence corresponding to the activity scheduling matrix $\mathbf{\Phi}$ and the sparsifying basis $\mathbf{\Psi}$, $\hat{\mathbf{\Psi}}$ be a matrix that comprises columns of $\mathbf{\Psi}$ normalized to unit $\ell_2$-norm, $d_{\min} = \min_i \|\boldsymbol{\psi}_i\|_2$ and $d_{\max} = \max_i \|\boldsymbol{\psi}_i\|_2$. If some $s$-sparse representation $\mathbf{x}_0$ of the noiseless signal satisfies $\mathbf{\Phi}\mathbf{\Psi}\mathbf{x}_0 = \mathbf{\Phi}\mathbf{f}$ and the regularized mutual coherence satisfies

$$\hat{\mu}(\mathbf{\Phi}, \mathbf{\Psi}) \leq \frac{d_{\min}^2 \rho}{4s d_{\max}^2 - d_{\min}^2}, \tag{8}$$

where $\rho = \min_i \|\hat{\mathbf{a}}_i\|_2^2$ and $\hat{\mathbf{A}} = \mathbf{\Phi}\hat{\mathbf{\Psi}}$, then the reconstruction error of (2) is bounded by

$$\|\hat{\mathbf{x}} - \mathbf{x}_0\|_2^2 \leq \frac{2\epsilon}{d_{\min}^2\rho + d_{\min}^2\hat{\mu}(\mathbf{\Phi},\mathbf{\Psi}) - 4sd_{\max}^2\hat{\mu}(\mathbf{\Phi},\mathbf{\Psi})}, \quad (9)$$

where $\hat{\mathbf{x}}$ is the solution of (2).

*Proof:* Letting $\Delta\mathbf{x} = \mathbf{x} - \mathbf{x}_0$ and $\mathbf{v} = \mathbf{y} - \mathbf{\Phi\Psi x}_0$, we can rewrite the optimization problem in (2) as

$$\begin{aligned} \min_{\Delta\mathbf{x}} \quad & \|\mathbf{x}_0 + \Delta\mathbf{x}\|_1 \\ \text{s.t.} \quad & \|\mathbf{\Phi\Psi}\Delta\mathbf{x} - \mathbf{v}\|_2^2 \leq \epsilon, \end{aligned} \quad (10)$$

where $\epsilon \geq \|\mathbf{v}\|_2^2$ by definition. Define $\Delta\hat{\mathbf{x}}$ as the solution of (10). According to the triangle inequality, we have

$$\|\mathbf{\Phi\Psi}\Delta\hat{\mathbf{x}}\|_2^2 \leq \|\mathbf{\Phi\Psi}\Delta\hat{\mathbf{x}} - \mathbf{v}\|_2^2 + \|\mathbf{v}\|_2^2 \leq 2\epsilon. \quad (11)$$

Now, define a diagonal scaling matrix $\mathbf{D}$ such that $\mathbf{\Psi} = \hat{\mathbf{\Psi}}\mathbf{D}$, where columns of $\hat{\mathbf{\Psi}}$ have unit $\ell_2$ norm, and the minimal and maximal diagonal elements of $\mathbf{D}$ are $d_{\min}$ and $d_{\max}$, respectively. Then the left-hand-side of (11) can be lower bounded by

$$\begin{aligned} &\|\mathbf{\Phi\Psi}\Delta\hat{\mathbf{x}}\|_2^2 = \|\mathbf{\Phi}\hat{\mathbf{\Psi}}\mathbf{D}\Delta\hat{\mathbf{x}}\|_2^2 \\ =&\Delta\hat{\mathbf{x}}^T\mathbf{D}^T \mathtt{diagm}\left(\hat{\mathbf{\Psi}}^T\mathbf{\Phi}^T\mathbf{\Phi}\hat{\mathbf{\Psi}}\right)\mathbf{D}\Delta\hat{\mathbf{x}} \\ &\quad + \Delta\hat{\mathbf{x}}^T\mathbf{D}^T \mathtt{nondiagm}\left(\hat{\mathbf{\Psi}}^T\mathbf{\Phi}^T\mathbf{\Phi}\hat{\mathbf{\Psi}}\right)\mathbf{D}\Delta\hat{\mathbf{x}} \\ \geq&\rho\|\mathbf{D}\Delta\hat{\mathbf{x}}\|_2^2 + \hat{\mu}(\mathbf{\Phi},\mathbf{\Psi})\Delta\hat{\mathbf{x}}^T\mathbf{D}^T\left(\mathbf{1}_{\hat{n}\times\hat{n}} - \mathbf{I}_{\hat{n}\times\hat{n}}\right)\mathbf{D}\Delta\hat{\mathbf{x}} \\ \geq&\rho\|\mathbf{D}\Delta\hat{\mathbf{x}}\|_2^2 - \hat{\mu}(\mathbf{\Phi},\mathbf{\Psi})|\Delta\hat{\mathbf{x}}|^T\mathbf{D}^T\left(\mathbf{1}_{\hat{n}\times\hat{n}} - \mathbf{I}_{\hat{n}\times\hat{n}}\right)\mathbf{D}|\Delta\hat{\mathbf{x}}| \\ \geq&(\rho + \hat{\mu}(\mathbf{\Phi},\mathbf{\Psi}))\|\mathbf{D}\Delta\hat{\mathbf{x}}\|_2^2 - \hat{\mu}(\mathbf{\Phi},\mathbf{\Psi})\|\mathbf{D}\Delta\hat{\mathbf{x}}\|_1^2 \\ \geq&d_{\min}^2(\rho + \hat{\mu}(\mathbf{\Phi},\mathbf{\Psi}))\|\Delta\hat{\mathbf{x}}\|_2^2 - d_{\max}^2\hat{\mu}(\mathbf{\Phi},\mathbf{\Psi})\|\Delta\hat{\mathbf{x}}\|_1^2. \end{aligned} \quad (12)$$

As $\Delta\hat{\mathbf{x}}$ is the solution of (10), we have

$$0 \geq \|\mathbf{x}_0 + \Delta\hat{\mathbf{x}}\|_1 - \|\mathbf{x}_0\|_1 \geq \|\Delta\hat{\mathbf{x}}\|_1 - 2\|\Delta\hat{\mathbf{x}}_{\mathcal{J}}\|_1, \quad (13)$$

where $\mathcal{J}$ denotes the support of $\mathbf{x}_0$. According to the inequalities (11), (12) and (13), it follows that

$$\begin{aligned} 2\epsilon &\geq d_{\min}^2(\rho + \hat{\mu}(\mathbf{\Phi},\mathbf{\Psi}))\|\Delta\hat{\mathbf{x}}\|_2^2 - d_{\max}^2\hat{\mu}(\mathbf{\Phi},\mathbf{\Psi})\|\Delta\hat{\mathbf{x}}\|_1^2 \\ &\geq d_{\min}^2(\rho + \hat{\mu}(\mathbf{\Phi},\mathbf{\Psi}))\|\Delta\hat{\mathbf{x}}\|_2^2 - 4d_{\max}^2\hat{\mu}(\mathbf{\Phi},\mathbf{\Psi})\|\Delta\hat{\mathbf{x}}_{\mathcal{J}}\|_1^2 \\ &\geq d_{\min}^2(\rho + \hat{\mu}(\mathbf{\Phi},\mathbf{\Psi}))\|\Delta\hat{\mathbf{x}}\|_2^2 - 4sd_{\max}^2\hat{\mu}(\mathbf{\Phi},\mathbf{\Psi})\|\Delta\hat{\mathbf{x}}_{\mathcal{J}}\|_2^2 \\ &\geq (d_{\min}^2\rho + d_{\min}^2\hat{\mu}(\mathbf{\Phi},\mathbf{\Psi}) - 4sd_{\max}^2\hat{\mu}(\mathbf{\Phi},\mathbf{\Psi}))\|\Delta\hat{\mathbf{x}}\|_2^2, \end{aligned} \quad (14)$$

which leads to the reconstruction error bound in (9) if

$$\hat{\mu}(\mathbf{\Phi},\mathbf{\Psi}) \leq \frac{d_{\min}^2\rho}{4sd_{\max}^2 - d_{\min}^2}. \quad (15)$$

■

*Remark 1:* As $d_{\min}^2 - 4sd_{\max}^2 < 0$, the upper bound of the reconstruction error in (9) tends to increase with the growth of the regularized mutual coherence, and tends to decrease with the growth of the minimum $\ell_2$-norm of all columns. This observation encourages our design for the activity scheduling to promote a low regularized mutual coherence and a large value of $\rho$.

*Remark 2:* In comparison to the mutual coherence defined in (5) which requires normalization of the columns of the equivalent sensing matrix, the use of regularized mutual coherence has the attraction that it only needs the computation of column coherence and the $\ell_2$-norm of columns, which will facilitate the derivation of the activity scheduling approach to be presented shortly. Note that $d_{\max}$ and $d_{\min}$ are fixed values, which are determined by the sparsifying basis $\mathbf{\Psi}$.

*Remark 3:* The proposed regularized mutual coherence extends the previous work of [27], where the mutual coherence of sensing matrices requires each column to have uniform $\ell_2$-norm. However, in our case, the sensing matrices have non-uniform $\ell_2$-norm, which leads to a distinct proof presented (12) and (14) that incorporates the effect of non-uniform $\ell_2$-norm.

## *B. The Proposed Design via Convex Optimization*

Building upon the previous analysis, we now propose an activity scheduling problem for cost-aware compressive sleeping WSNs, that performs a balance between the CS reconstruction accuracy and the total cost imposed by the activated SNs.

To facilitate the design, we define $\tilde{\mathbf{\Phi}} = \mathbf{\Phi}^T\mathbf{\Phi}$, which is an $n \times n$ diagonal matrix such that

$$\tilde{\Phi}_{i,i} = \begin{cases} 1, & \text{SN } i \text{ is activated} \\ 0, & \text{otherwise} \end{cases}. \quad (16)$$

The activity matrix $\mathbf{\Phi}$ can be obtained directly from $\tilde{\mathbf{\Phi}}$. Defining $\mathcal{D}$ as the set of all diagonal matrices, we now put forth the following optimization problem

$$\begin{aligned} \min_{\tilde{\mathbf{\Phi}}\in\mathcal{D}} \quad & \mathbf{c}^T\mathtt{diag}(\tilde{\mathbf{\Phi}}) \\ \text{s.t.} \quad & \mathtt{nondiagm}(\hat{\mathbf{\Psi}}^T\tilde{\mathbf{\Phi}}\hat{\mathbf{\Psi}}) \leq \eta, \\ & \mathtt{diag}(\hat{\mathbf{\Psi}}^T\tilde{\mathbf{\Phi}}\hat{\mathbf{\Psi}}) \geq \rho, \\ & \tilde{\Phi}_{i,i} \in \{0,1\},\ i = 1,\ldots,n, \\ & \mathtt{Tr}(\tilde{\mathbf{\Phi}}) = m. \end{aligned} \quad (17)$$

where $\mathbf{c} \in \mathbb{R}^n$ denotes the cost profile for activating different SNs, $\eta$ $(0 < \eta < 1)$ relates to the required regularized mutual coherence, and $\rho$ $(0 < \rho < 1)$ denotes the threshold for the $\ell_2$-norm of the columns of $\tilde{\mathbf{\Phi}}\hat{\mathbf{\Psi}}$. Note that the first two constraints guarantee that the reconstruction error is bounded according to Theorem 1. The value of $\eta$ and $\rho$ can be tuned if some training data is available or chosen as $\eta = \hat{\mu}(\mathbf{\Phi}^{rand},\mathbf{\Psi})$ and $\rho = \min_i \|\hat{\mathbf{a}}_i^{rand}\|_2^2$, where $\mathbf{\Phi}^{rand} \in \mathbb{R}^{m\times n}$ is a randomly selected activity matrix and $\hat{\mathbf{A}}_i^{rand} = \mathbf{\Phi}^{rand}\hat{\mathbf{\Psi}}$, which makes the worst-case error bound of the optimized activity matrix no worse than that of a random one. Although more strict constraints that involve a lower value of $\eta$ and a larger value of $\rho$ could improve the worst-case performance, they reduce the size of the feasible solution set and could rule out good solutions with a low sampling cost, which reflects the trade-off between the reconstruction accuracy and the sampling cost.

This optimization problem is 0-1 integer linear programming (ILP), and provided the number of SNs is small, a variety of algorithms such as the cutting-plane method and the branch-and-bound method can be used to solve the problem exactly [30]. However, as ILP is in general NP-hard, the computational complexity for solving this problem increases exponentially with a growing number of SNs. To deal with the computational complexity issue, a widely used heuristic approach for solving the 0-1 ILP problem with a low computational complexity is to ignore the integer constraints and solve the resulting relaxed convex problem, which leads to

$$\begin{aligned}\min_{\tilde{\boldsymbol{\Phi}}\in\mathcal{D}} \quad & \mathbf{c}^T \mathtt{diag}(\tilde{\boldsymbol{\Phi}}) \\ \text{s.t.} \quad & \mathtt{nondiagm}(\hat{\boldsymbol{\Psi}}^T\tilde{\boldsymbol{\Phi}}\hat{\boldsymbol{\Psi}}) \le \eta, \\ & \mathtt{diag}(\hat{\boldsymbol{\Psi}}^T\tilde{\boldsymbol{\Phi}}\hat{\boldsymbol{\Psi}}) \ge \rho, \\ & 0 \le \tilde{\Phi}_{i,i} \le 1,\ i=1,\ldots,n, \\ & \mathtt{Tr}(\tilde{\boldsymbol{\Phi}}) = m.\end{aligned} \tag{18}$$

Then the $m$ largest $\tilde{\Phi}_{i,i}$ are rounded up and the others are rounded down. This heuristic approach is also used in the related work on activity scheduling designs [21]–[23] and many other applications such as antenna selection in multi-antenna wireless communication systems [31], and sensor selection for parameter estimation [32].

We note that extra cost is required to communicate the IDs of the active nodes (that are obtained from the sampling matrix) via the downlink from the FC. However, i) it is not usually necessary to change the node activation matrix very frequently (e.g., changing the activation pattern for every 100 rounds of data collection from SNs), and ii) the energy consumed by the SN to receive a message from the FC is often much lower than the energy used to transmit its message to the FC (if the distance between the FC and the SN is large). In these cases, the cost of transmitting the information about the SN activation could be a relatively minor factor.

## IV. Extensions With Prior Information of the Signal Support

In many cases it has been observed that various support patterns occur with different probabilities. For example, as the physical phenomena have correlation in both the spatial and the temporal domains, it is likely that the actual support has some overlap with the signal support in a previous time slot. This prior information concerning the signal support is not utilized in the activity matrix design presented previously, and the regularized mutual coherence employed in the previous section actually characterizes the quality of an activity matrix in terms of the worst case performance. In this section, we consider extending the proposed activity scheduling approach to the case where prior information about the signal support is available. This additional knowledge is exploited to improve the trade-off between the CS reconstruction accuracy and the sampling cost.

To indicate the CS reconstruction accuracy of the equivalent sensing matrix with non-uniform signal support probability, we adopt the oracle estimator MSE in the evaluation. The oracle estimator performs ideal LS estimation based on prior knowledge of the sparse vector support. The MSE of this oracle LS estimator coincides with the unbiased Cramèr-Rao bound for exactly $s$-sparse deterministic vectors [24], so that it represents the best achievable performance for any unbiased estimator. It has been shown that the oracle estimator MSE performance acts as a performance benchmark for the key sparse recovery algorithms such as the BPDN, the Dantzig selector, the orthogonal matching pursuit (OMP) and thresholding algorithms [33].

In the presence of a standard Gaussian noise vector, the oracle estimator MSE can be written as

$$\begin{aligned}&\mathbb{E}_{\mathbf{x},\mathbf{z}}\left(\left\|\mathcal{F}^{\text{oracle}}(\mathbf{y},\boldsymbol{\Phi}\boldsymbol{\Psi})-\mathbf{x}\right\|_2^2\right) \\ =&\sigma^2\mathbb{E}_{\mathcal{J}}\left(\mathtt{Tr}\big((\mathbf{E}_{\mathcal{J}}^T\boldsymbol{\Psi}^T\boldsymbol{\Phi}^T\boldsymbol{\Phi}\boldsymbol{\Psi}\mathbf{E}_{\mathcal{J}})^{-1}\big)\right) \\ =&\sigma^2\sum_{\mathcal{J}\in\mathcal{S}}\Pr(\mathcal{J})\,\mathtt{Tr}\big((\mathbf{E}_{\mathcal{J}}^T\boldsymbol{\Psi}^T\tilde{\boldsymbol{\Phi}}\boldsymbol{\Psi}\mathbf{E}_{\mathcal{J}})^{-1}\big),\end{aligned} \tag{19}$$

where $\mathcal{F}(\cdot)$ denotes the oracle estimator, $\mathcal{J}\subset\{1,\ldots,\hat{n}\}$ denotes the support of $\mathbf{x}$, and $\mathcal{S}$ is the set of all support patterns. For a fixed sparsity level $s$, the size of $\mathcal{S}$ is $\binom{\hat{n}}{s}$.

Building upon the previous analysis and the proposed activity scheduling framework in Section III, we will henceforth be concerned with a new activity scheduling problem as follows

$$\begin{aligned}\min_{\tilde{\boldsymbol{\Phi}}\in\mathcal{D},\mathcal{J}\in\mathcal{S}} \quad & \mathbf{c}^T \mathtt{diag}(\tilde{\boldsymbol{\Phi}}) \\ \text{s.t.} \quad & \sum_{\mathcal{J}\in\mathcal{S}}\Pr(\mathcal{J})\,\mathtt{Tr}\big((\mathbf{E}_{\mathcal{J}}^T\boldsymbol{\Psi}^T\tilde{\boldsymbol{\Phi}}\boldsymbol{\Psi}\mathbf{E}_{\mathcal{J}})^{-1}\big)\le\alpha, \\ & \mathtt{nondiagm}(\hat{\boldsymbol{\Psi}}^T\tilde{\boldsymbol{\Phi}}\hat{\boldsymbol{\Psi}}) \le \eta, \\ & \mathtt{diag}(\hat{\boldsymbol{\Psi}}^T\tilde{\boldsymbol{\Phi}}\hat{\boldsymbol{\Psi}}) \ge \rho, \\ & 0 \le \tilde{\Phi}_{i,i} \le 1,\ i=1,\ldots,n, \\ & \mathtt{Tr}(\tilde{\boldsymbol{\Phi}}) = m,\end{aligned} \tag{20}$$

where $\alpha > 0$. The first constraint in (20) guarantee that the oracle estimator MSE is upper bounded by $\frac{\alpha}{\sigma^2}$. The value of $\alpha$ can be tuned with training data or selected by using a random sampling scheme, i.e., $\alpha = \sum_{\mathcal{J}\in\mathcal{S}}\Pr(\mathcal{J})\,\mathtt{Tr}\big((\mathbf{E}_{\mathcal{J}}^T\boldsymbol{\Psi}^T(\boldsymbol{\Phi}^{rand})^T\boldsymbol{\Phi}^{rand}\boldsymbol{\Psi}\mathbf{E}_{\mathcal{J}})^{-1}\big)$. As $\mathbf{E}_{\mathcal{J}}^T\boldsymbol{\Psi}^T\tilde{\boldsymbol{\Phi}}\boldsymbol{\Psi}\mathbf{E}_{\mathcal{J}}$ is a positive semidefinite matrix, according to the Schurs complement in A.5.5 of [34] we can rewrite (20) as

$$\begin{aligned}\min_{\tilde{\boldsymbol{\Phi}}\in\mathcal{D},\mathbf{Q}_{\mathcal{J}},\mathcal{J}\in\mathcal{S}} \quad & \mathbf{c}^T \mathtt{diag}(\tilde{\boldsymbol{\Phi}}) \\ \text{s.t.} \quad & \sum_{\mathcal{J}\in\mathcal{S}}\Pr(\mathcal{J})\,\mathtt{Tr}\big(\mathbf{Q}_{\mathcal{J}}\big)\le\alpha, \\ & \begin{bmatrix}\mathbf{E}_{\mathcal{J}}^T\boldsymbol{\Psi}^T\tilde{\boldsymbol{\Phi}}\boldsymbol{\Psi}\mathbf{E}_{\mathcal{J}} & \mathbf{I}_s \\ \mathbf{I}_s & \mathbf{Q}_{\mathcal{J}}\end{bmatrix}\succeq\mathbf{0},\ \forall\mathcal{J}\in\mathcal{S}, \\ & \mathtt{nondiagm}(\hat{\boldsymbol{\Psi}}^T\tilde{\boldsymbol{\Phi}}\hat{\boldsymbol{\Psi}}) \le \eta, \\ & \mathtt{diag}(\hat{\boldsymbol{\Psi}}^T\tilde{\boldsymbol{\Phi}}\hat{\boldsymbol{\Psi}}) \ge \rho, \\ & 0 \le \tilde{\Phi}_{i,i} \le 1,\ i=1,\ldots,n, \\ & \mathtt{Tr}(\tilde{\boldsymbol{\Phi}}) = m.\end{aligned} \tag{21}$$

After solving (21), the $m$ largest $\tilde{\Phi}_{i,i}$ are chosen and the corresponding indexes relate to the selected nodes.

Generally, the computational complexity for solving (21) is high owing to large number of possible support patterns in $\mathcal{S}$. Specifically, as the semidefinite programming problem in (21) consists of one diagonal matrix variable $\tilde{\boldsymbol{\Phi}}$ of size $n\times n$ and $\binom{\hat{n}}{s}$ matrix variables $\mathbf{Q}_{\mathcal{J}}$ of size $s\times s$, it can be iteratively solved using interior point methods with computational complexity growing at most $\mathcal{O}(n^3+\binom{\hat{n}}{s}^3 s^6)$ arithmetic operations in each iteration [17]. Therefore, if the

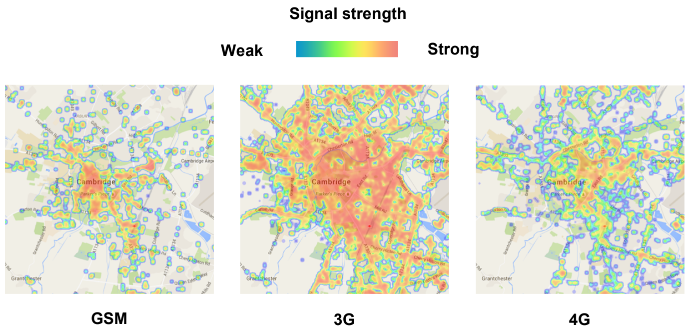


Fig. 2. The GSM, the 3G and the 4G signal strength maps of Cambridge.

number of SNs, i.e., $n$, is large, the computational complexity grows exponentially owing to the term $\binom{\hat{n}}{s}$ where $\hat{n} \geq n$. To reduce the computational complexity, one can randomly select a subset of support patterns $\mathcal{S}' \subset \mathcal{S}$ based upon the signal support probability $\Pr(\mathcal{J})$, and approximately calculate the oracle estimator MSE by

$$\begin{aligned}&\mathbb{E}_{\mathbf{x},\mathbf{z}}\left(\left\|\mathcal{F}^{\text{oracle}}(\mathbf{y},\boldsymbol{\Phi}\boldsymbol{\Psi})-\mathbf{x}\right\|_2^2\right)\\ \approx&\sigma^2\sum_{\mathcal{J}\in\mathcal{S}'}\Pr(\mathcal{J})\,\mathtt{Tr}\big((\mathbf{E}_{\mathcal{J}}^T\boldsymbol{\Psi}^T\tilde{\boldsymbol{\Phi}}\boldsymbol{\Psi}\mathbf{E}_{\mathcal{J}})^{-1}\big).\end{aligned} \tag{22}$$

This computational complexity reduction method has also been used in [17], where a uniform support distribution is assumed. For the case that there is a support pattern $\mathcal{J}'$ with a dominant probability, one can ignore the other support patterns and solve the following activity scheduling problem instead

$$\begin{aligned}\min_{\tilde{\boldsymbol{\Phi}}\in\mathcal{D},\mathbf{Q}_{\mathcal{J}'}}\quad & \mathbf{c}^T\mathtt{diag}(\tilde{\boldsymbol{\Phi}})\\ \text{s.t.}\quad & \mathtt{Tr}\big(\mathbf{Q}_{\mathcal{J}'}\big)\leq\alpha,\\ & \begin{bmatrix}\mathbf{E}_{\mathcal{J}'}^T\boldsymbol{\Psi}^T\tilde{\boldsymbol{\Phi}}\boldsymbol{\Psi}\mathbf{E}_{\mathcal{J}'} & \mathbf{I}_s\\ \mathbf{I}_s & \mathbf{Q}_{\mathcal{J}'}\end{bmatrix}\succeq\mathbf{0},\\ & \mathtt{nondiagm}(\hat{\boldsymbol{\Psi}}^T\tilde{\boldsymbol{\Phi}}\hat{\boldsymbol{\Psi}})\leq\eta,\\ & \mathtt{diag}(\hat{\boldsymbol{\Psi}}^T\tilde{\boldsymbol{\Phi}}\hat{\boldsymbol{\Psi}})\geq\rho,\\ & 0\leq\tilde{\Phi}_{i,i}\leq 1,\ i=1,\ldots,n,\\ & \mathtt{Tr}(\tilde{\boldsymbol{\Phi}})=m.\end{aligned} \tag{23}$$

For a typical environmental monitoring application, in view of the high temporal correlation of the spatial signal, the support of the signal in the previous time slot can be chosen as the support pattern $\mathcal{J}'$ in (23).

## V. Experimental Results

In this section, we assess how the proposed activity scheduling approaches perform in comparison with other approaches including the random scheme, the greedy scheme, the RCS scheme proposed in [22], and the ONS scheme proposed in [21], where the ONS scheme is only suitable for the scenario where prior information of the signal support is available.

It is worth mentioning that for a random Gaussian sparsifying basis or a random sampling cost map, the greedy scheme not only has the lowest sampling cost but also maintains good reconstruction accuracy, and thus outperforms the other approaches. However, in practical applications, the sparsifying basis often has a structure and the sampling cost map has a high spatial correlation. In this case, the greedy scheme may not perform well, which motivates our work. In order to generate correlated cost maps in the experiments, we use maps showing the mean signal strength for the global system for mobile communications (GSM), 3G and 4G mobile systems in Cambridge that we assume to be nonchanging as shown in Fig. 2, and we also assume that the sampling cost is inversely proportional to the signal strength. We consider the cost ratio, i.e., the cost owing to the active SNs over the cost if all SNs are activated. Consequently it is only the relative signal strength that matters, which is indicated by the signal strength bar in Fig. 2. We assume that the SNs are placed at a $16 \times 16$ grid placed over the maps in Fig. 2 and are used for monitoring environmental parameters. We generate synthetic data to evaluate the CS reconstruction accuracy, which will be described in related subsections in the sequel.

In addition, in the evaluation, we consider two different sparsifying bases, i.e., the discrete Fourier transform (DFT) basis and the discrete cosine transform (DCT) basis. We use CVX [35], a package for specifying and solving convex programs, to solve the proposed convex-relaxed activity scheduling problems and to reconstruct the signals. The reconstruction accuracy performance is evaluated by using the averaged relative error which is the average of $\frac{\|\hat{\mathbf{f}}-\mathbf{f}\|_2^2}{\|\mathbf{f}\|_2^2}$, where $\hat{\mathbf{f}}$ denotes a reconstructed vector including all the SNs' data. For our proposed approaches, the value of $\eta$, $\rho$ and $\alpha$ are chosen by using a random activity scheduling matrix. All the experiments consist of 5000 independent trials.

### A. Experiments With a Uniform Signal Support Probability

We now consider the case with no knowledge about the signal support, and generate the signal support with a uniform distribution. In each trial, $s$ non-zero components of the signal representation $\mathbf{x}$ are drawn from an independent and identically distributed (i.i.d.) zero mean and unit variance Gaussian distribution, and the sensor measurements are corrupted by additive zero-mean Gaussian noise with variance $\sigma^2$.

Fig. 3 and Fig. 4 show the performance trade-off between the reconstruction accuracy and the sampling cost for the various approaches using the DFT and the DCT basis, respectively. In these experiments, we fix the signal sparsity level to be $s = 5$, and vary the number of active SNs $m$, which results in different sampling costs and reconstruction accuracy. We can observe that the averaged relative error of the CS reconstruction decreases with an increasing sampling cost ratio for all the approaches. Interestingly, with the same settings, the greedy approach performs worse than the random approach for the DCT basis, while the reverse is observed for the DFT basis. In addition, the performance gap between the greedy approach and the random approach varies for different cost maps. These observations demonstrate the limitations of the random and the greedy approaches, i.e., their performance is significantly affected by the spatial correlation cost map and the signal sparsifying basis. We note that the proposed approach is superior to the other approaches for all the spatial correlation cost maps and the signal sparsifying basis.

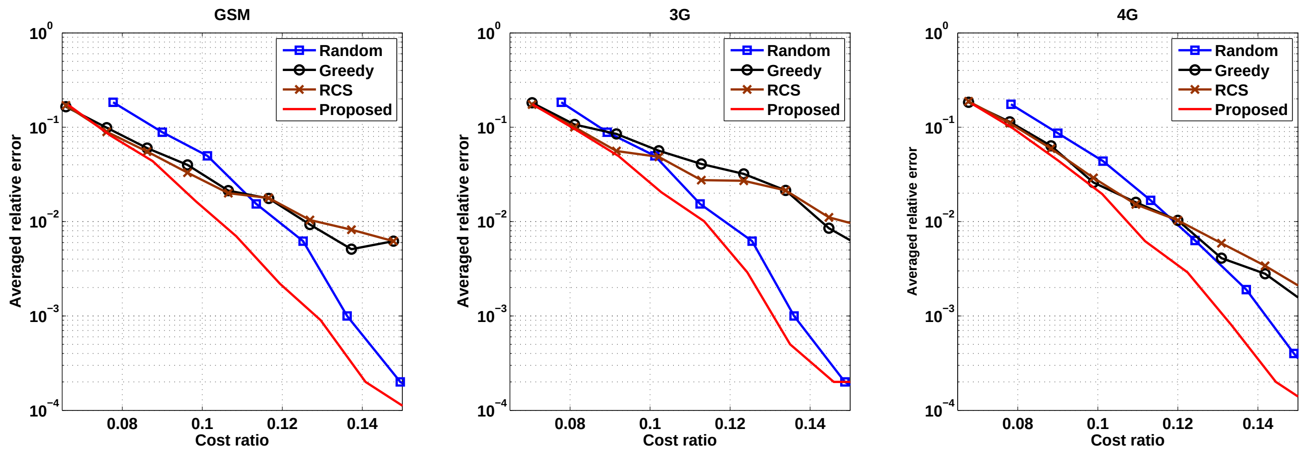


Fig. 3. The accuracy-cost trade-off performance with the DFT basis and uniformly generated signal supports. ($\sigma^2 = 10^{-2}$ and $s = 5$)

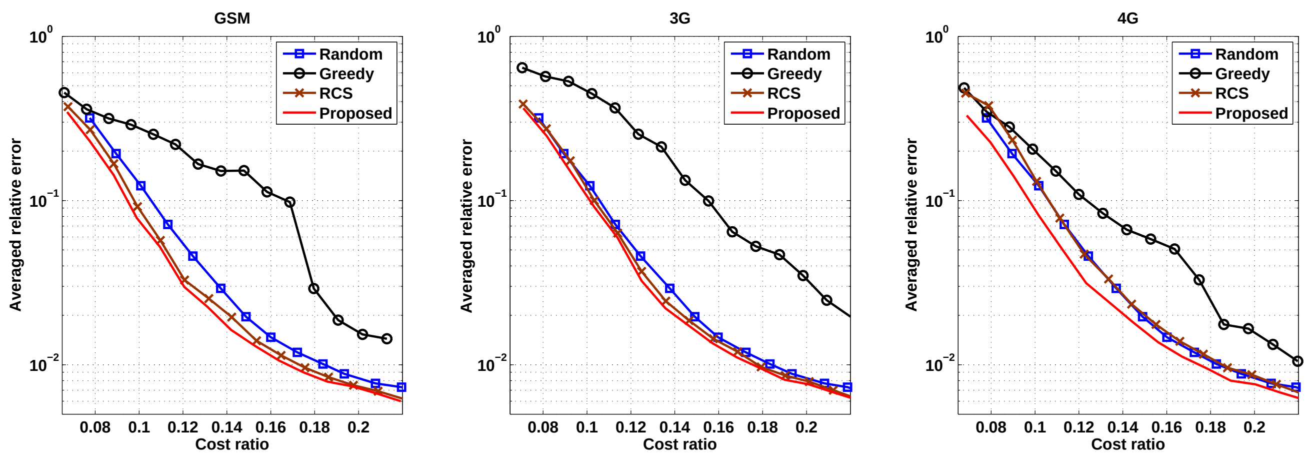


Fig. 4. The accuracy-cost trade-off performance with the DCT basis and uniformly generated signal supports. ($\sigma^2 = 10^{-2}$ and $s = 5$)

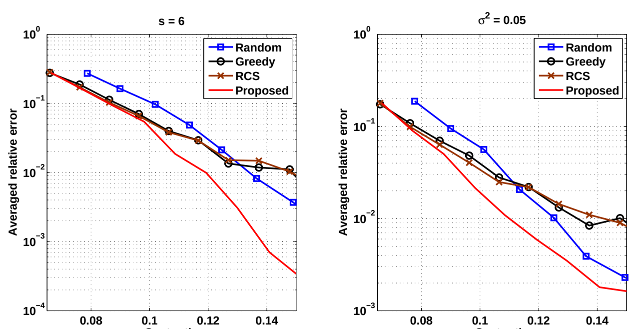


Fig. 5. The comparison of different approaches with various settings of parameters for the scenario of a uniform signal support probability. (DFT basis and the GSM cost map)

More performance comparisons for various parameter settings including the signal sparsity $s$ and the noise variance $\sigma^2$ are shown in Fig. 5, where the DFT basis and the GSM cost map of Cambridge are used. The proposed approach provides a good trade-off between the sampling cost and the reconstruction accuracy, and outperforms all the other competitors in all the scenarios.

### B. Experiments With Prior Information of the Signal Support

In each trial of the experiments, we first generate a non-uniform signal support probability, where $s$ randomly selected entries are associated with the probability $\frac{9}{10}\frac{\texttt{Constant}}{s}$, and the other entries are associated with the probability $\frac{1}{10}\frac{\texttt{Constant}}{n-s}$. Then we generate the sparse signal representation $\mathbf{x}$ with this signal support probability, and the amplitude of non-zero components are drawn from an i.i.d. zero mean and unit variance Gaussian distribution. The proposed optimization problem (23) is solved to obtain the activity scheduling pattern with the use of the prior information on the signal support.

Fig. 6 and Fig. 7 illustrate the performance trade-off for various approaches using the DFT and the DCT bases, respectively, where we set $\sigma^2 = 10^{-2}$ and $s = 5$. As with the previous experiments, by varying the number of active SNs $m$, we obtain the trade-off of the sampling cost versus the reconstruction accuracy for the various methods. It is observed again that for all the cost maps, the proposed approach provides the best performance trade-off. We note in Fig. 6 and Fig. 7 that the ONS approach doesn't provide a good activity scheduling in terms of the trade-off between sampling cost and reconstruction accuracy, since it requires a relatively accurate

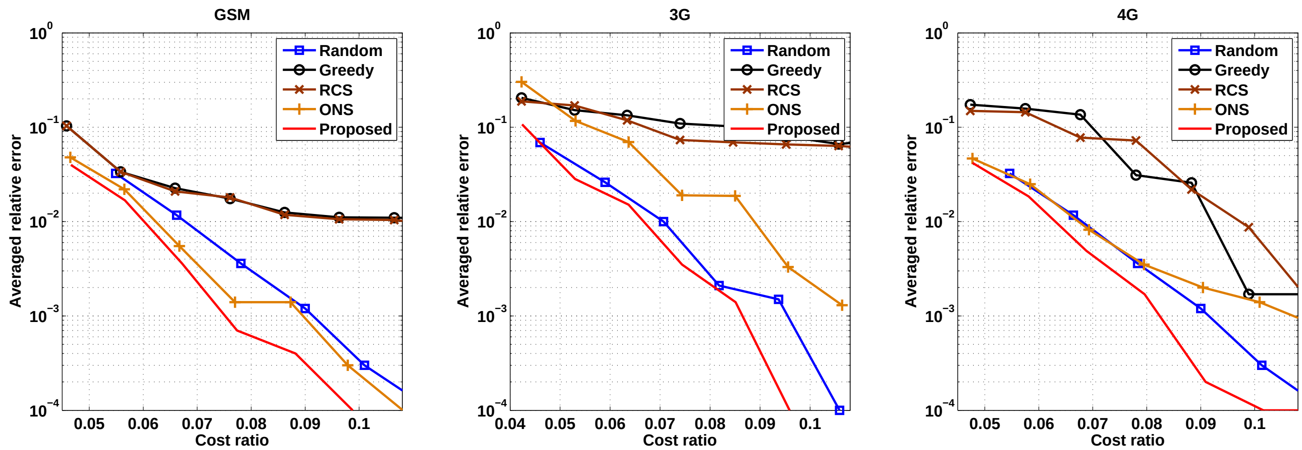


Fig. 6. The accuracy-cost trade-off performance with the DFT basis and prior information of the signal support. ($\sigma^2 = 10^{-2}$ and $s = 5$)

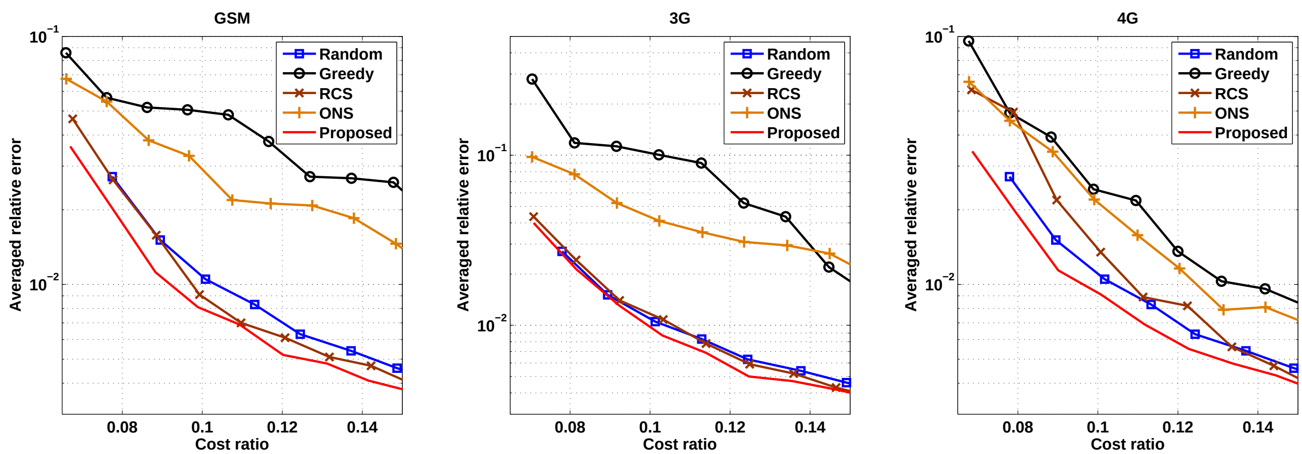


Fig. 7. The accuracy-cost trade-off performance with the DCT basis and prior information of the signal support. ($\sigma^2 = 10^{-2}$ and $s = 5$)

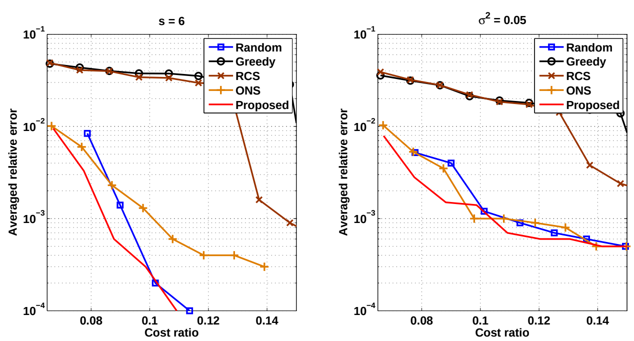


Fig. 8. The comparison of different approaches with various settings of parameters for the scenario with prior information of the signal support. (DFT basis and the GSM cost map)

knowledge of the signal support. In addition, in comparison to the results of the experiment in the previous subsection, all the designs work better with prior knowledge owing to the fact that prior knowledge is exploited in the CS reconstruction in (4).

Fig. 8 provides more comparison results for the various schemes for different settings including the signal sparsity $s$ and the noise variance $\sigma^2$, where the DFT basis and the GSM cost map of Cambridge are used. Again, we note that the proposed approach has a superior performance trade-off in comparison to other schemes.

### *C. Convex Relaxation vs. Exhaustive Search*

In this experiment, we investigate the performance degradation caused by convex relaxation in (18). Owing to the high computation cost in solving the original design problem in (17), we conduct simulations with only $n = 20$ SNs. We generate the elements of the dictionary $\mathbf{\Psi}$ and the sparse signal representation $\mathbf{x}$ from an i.i.d. zero mean and unit variance Gaussian distribution. The sampling cost of each SN is generated following an i.i.d. uniform distribution in the range from 0 to 1.

Fig. 9 illustrates the accuracy-cost trade-off performance for the random scheme, the proposed approach with convex relaxation and the original proposed design. Although the SNs selected by the proposed approach with convex relaxation and a rounding scheme are not guaranteed to satisfy all the constraints in the original design problem, we do not observe significant performance degradation (as least for the problem

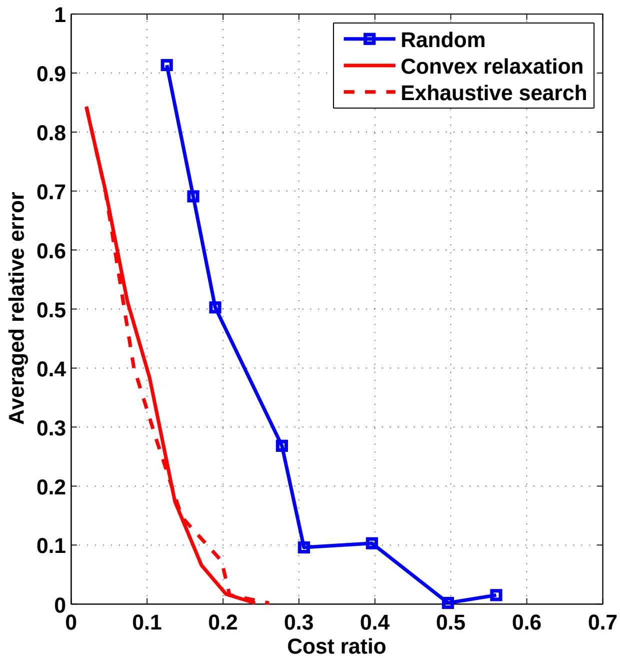


Fig. 9. Convex relaxation vs. exhaustive search. ($\sigma^2 = 10^{-2}$ and $s = 2$)

with a small ambient dimension). Both the original design by exhaustive search and the one with convex relaxation outperform the random scheme.

## VI. Conclusions

In this paper, we have considered the design of cost-aware activity scheduling for compressive sleeping WSNs that exploit the CS principle in data acquisition. In contrast to the conventional CS framework that implicitly assumes equal cost for all samples and is only concerned with the reconstruction accuracy, the proposed activity scheduling approach aims to improve the performance trade-off between the reconstruction accuracy and the sampling cost in view of the non-uniform sampling cost often experienced in compressive sleeping WSNs. We formulate the activity scheduling as an optimization problem which exploits the regularized mutual coherence of the equivalent sensing matrix as an indicator of the reconstruction accuracy, and further extend the approach to the case where prior information of the signal support is available. The proposed designs exhibit performance trade-off gains in relation to the conventional random activity scheme and the greedy activity scheme as well as to other optimized designs.

## References


[1] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[2] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[3] W. Bajwa, J. Haupt, A. Sayeed, and R. Nowak, "Compressive wireless sensing," in *The Fifth International Conference on Information Processing in Sensor Networks(IPSN)*, 2006, pp. 134–142.

[4] W. Wang, M. Garofalakis, and K. Ramchandran, "Distributed sparse random projections for refinable approximation," in *The 6th International Symposium on Information Processing in Sensor Networks (IPSN)*, April 2007, pp. 331–339.

[5] C. Luo, F. Wu, J. Sun, and C. W. Chen, "Efficient measurement generation and pervasive sparsity for compressive data gathering," *IEEE Transactions on Wireless Communications*, vol. 9, no. 12, pp. 3728–3738, December 2010.

[6] G. Quer, R. Masiero, G. Pillonetto, M. Rossi, and M. Zorzi, "Sensing, compression, and recovery for WSNs: Sparse signal modeling and monitoring framework," *IEEE Transactions on Wireless Communications*, vol. 11, no. 10, pp. 3447–3461, October 2012.

[7] W. Chen, M. Rodrigues, and I. Wassell, "A Fréchet mean approach for compressive sensing date acquisition and reconstruction in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 11, no. 10, pp. 3598 –3606, 2012.

[8] G. Yang, V. Tan, C. K. Ho, S. H. Ting, and Y. L. Guan, "Wireless compressive sensing for energy harvesting sensor nodes," *IEEE Transactions on Signal Processing*, vol. 61, no. 18, pp. 4491–4505, Sept 2013.

[9] W. Chen, Y. Andreopoulos, I. Wassell, and M. Rodrigues, "Towards energy neutrality in energy harvesting wireless sensor networks: A case for distributed compressive sensing?" in *IEEE Global Communications Conference (GLOBECOM)*, Dec 2013, pp. 474–479.

[10] R. Baraniuk, V. Cevher, M. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Transactions on Information Theory*, vol. 56, no. 4, pp. 1982–2001, April 2010.

[11] N. Vaswani and W. Lu, "Modified-CS: Modifying compressive sensing for problems with partially known support," *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4595–4607, 2010.

[12] J. Scarlett, J. Evans, and S. Dey, "Compressed sensing with prior information: Information-theoretic limits and practical decoders," *IEEE Transactions on Signal Processing*, vol. 61, no. 2, pp. 427–439, Jan 2013.

[13] W. Chen, M. Rodrigues, and I. Wassell, "Projection design for statistical compressive sensing: A tight frame based approach," *IEEE Transactions on Signal Processing*, vol. 61, no. 8, pp. 2016–2029, 2013.

[14] M. Elad, "Optimized projections for compressed sensing," *IEEE Transactions on Signal Processing*, vol. 55, no. 12, pp. 5695 –5702, 2007.

[15] J. Duarte-Carvajalino and G. Sapiro, "Learning to sense sparse signals: Simultaneous sensing matrix and sparsifying dictionary optimization," *IEEE Transactions on Image Processing*, vol. 18, no. 7, pp. 1395 –1408, 2009.

[16] G. Li, Z. Zhu, D. Yang, L. Chang, and H. Bai, "On projection matrix optimization for compressive sensing systems," *IEEE Transactions on Signal Processing*, vol. 61, no. 11, pp. 2887–2898, June 2013.

[17] A. Shirazinia and S. Dey, "Power-constrained sparse Gaussian linear dimensionality reduction over noisy channels," *IEEE Transactions on Signal Processing*, vol. PP, no. 99, pp. 1–1, 2015.

[18] L. Zelnik-Manor, K. Rosenblum, and Y. Eldar, "Sensing matrix optimization for block-sparse decoding," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4300–4312, Sept 2011.

[19] A. Jones, B. Adcock, and A. C. Hansen, "On asymptotic incoherence and its implications for compressed sensing of inverse problems," *arXiv preprint arXiv:1402.5324*, 2014.

[20] Q. Ling and Z. Tian, "Decentralized sparse signal recovery for compressive sleeping wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 58, no. 7, pp. 3816–3827, 2010.

[21] W. Chen and I. Wassell, "Optimized node selection for compressive sleeping wireless sensor networks," *IEEE Transactions on Vehicular Technology*, in preprint.

[22] L. Xu, X. Hao, N. D. Lane, X. Liu, and T. Moscibroda, "Cost-aware compressive sensing for networked sensing systems," in *Proceedings of the 14th International Conference on Information Processing in Sensor Networks*. ACM, 2015, pp. 130–141.

[23] W. Chen and I. Wassell, "Compressive sleeping wireless sensor networks with active node selection," in *IEEE Global Communications Conference (GLOBECOM)*, Dec 2014, pp. 7–12.

[24] Z. Ben-Haim and Y. Eldar, "The Cramèr-Rao bound for estimating a sparse parameter vector," *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3384–3389, June 2010.

[25] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, 2008.

[26] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3736–3745, 2006.

[27] D. Donoho, M. Elad, and V. Temlyakov, “Stable recovery of sparse overcomplete representations in the presence of noise,” *IEEE Transactions on Information Theory*, vol. 52, no. 1, pp. 6 – 18, Jan. 2006.

[28] W. Chen, M. Rodrigues, and I. Wassell, “On the use of unit-norm tight frames to improve the average mse performance in compressive sensing applications,” *IEEE Signal Processing Letters*, vol. 19, no. 1, pp. 8–11, Jan 2012.

[29] J. Xu, Y. Pi, and Z. Cao, “Optimized projection matrix for compressive sensing,” *EURASIP Journal on Advances in Signal Processing*, vol. 43, pp. 1–8, Feb. 2010.

[30] T. Achterberg, “Scip: solving constraint integer programs,” *Mathematical Programming Computation*, vol. 1, no. 1, pp. 1–41, 2009.

[31] M. Hanif, P. Smith, D. Taylor, and P. Martin, “MIMO cognitive radios with antenna selection,” *IEEE Transactions on Wireless Communications*, vol. 10, no. 11, pp. 3688–3699, 2011.

[32] S. Joshi and S. Boyd, “Sensor selection via convex optimization,” *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 451–462, 2009.

[33] Z. Ben-Haim, Y. Eldar, and M. Elad, “Coherence-based performance guarantees for estimating a sparse vector under random noise,” *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5030–5043, Oct 2010.

[34] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[35] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming, version 2.1,” http://cvxr.com/cvx, Mar. 2014.