Optimized Node Selection for Compressive Sleeping Wireless Sensor Networks

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

***Abstract*—In this paper, we propose an active node selection framework for compressive sleeping wireless sensor networks (WSNs) in order to improve signal acquisition performance, network lifetime and the use of spectrum resources. While conventional compressive sleeping WSNs only exploit the spatial correlation of SNs, the proposed approach further exploits the temporal correlation by selecting active nodes using the support of the data reconstructed in the previous time instant. The node selection problem is framed as the design of a specialized sensing matrix, where the sensing matrix consists of selected rows of an identity matrix. By capitalizing on a genie-aided reconstruction procedure, we formulate the active node selection problem into an optimization problem, which is then approximated by a constrained convex relaxation plus a rounding scheme. Simulation results show that our proposed active node selection approach leads to an improved reconstruction performance, network lifetime and spectrum usage in comparison to various node selection schemes for compressive sleeping WSNs.**

I. Introduction

OVER the past two decades, the rapid development of technologies in sensing, computing and communication has made it possible to employ wireless sensor networks (WSNs) to continuously monitor physical phenomena in a variety of applications, for example air quality monitoring, wildlife tracking, biomedical monitoring and disaster detection [1], [2]. As the number and the resolution of the sensors grow, the performance bottleneck is the sensor node (SN), which usually has limited battery power, memory, computational capability, wireless bandwidth and physical size [1], [3].

Due to the temporal correlation and spatial correlation among the densely deployed SNs within the event area, it may not be necessary for every SN to report its data, which saves battery power and spectrum resource. Various sensor selection [4], [5] and data aggregation schemes [6] have been proposed to reduce the amount of redundant data transmitted in the network. For example, Vuran and Akyildiz have developed a medium access control (MAC) protocol, namely correlation-based collaborative medium access control (CC-MAC), that reduces the number of transmitted packets by restricting the reporting tasks to a small number of SNs [4]; Liu et al. propose an optimal node-selection algorithm to select a subset of camera sensors for estimating the location of a target while minimizing the energy cost [5]; Fasolo et al. provide different in-network aggregation schemes that combines data coming from several SNs to reduce the overall network traffic [6]. Even though many sensor selection and data aggregation schemes have been investigated, they either compromise the fidelity for monitoring the specific physical phenomenon or require complicated in-network compression to be performed.

Compressive sensing (CS) [7], [8] provides a fresh perspective for efficient data acquisition without compromising data recovery. It enables a fusion center (FC) to reconstruct the physical phenomenon with a reduced amount of data, where knowledge of the characteristics of the data is exploited. The asymmetric characteristics of WSNs, which typically comprise a smart fusion center (FC) with high power and computational capability and many SNs with limited energy storage and computing capability, motivates the usage of CS [9]–[12], which trades-off the convenience of data acquisition against the computational complexity of data reconstruction by leveraging the compressibility of natural signals.

There have been various generalizations of the CS framework to scenarios where a large number of SNs are distributed in the field to collect information of interest. For example, the CS principle is used in [10], [13] as a compression and forwarding scheme to minimize the number of packets to transmit; A network compressive coding scheme for sensor networks is developed in [14]; A compressive sleeping strategy is applied in [9], [15] such that only a subset of SNs are active and all the others SNs are turned off in order to reduce the energy consumption; A Fréchet mean approach that estimates the signal support from multiple correlated signals and then leverages the support estimate to enhance the reconstruction is proposed in [11]; Distributed compressive sensing (DCS) [16], [17] that exploits a joint signal sparse model is proposed to jointly reconstruct multiple signals. However, DCS is not appropriate for many monitoring applications that are delay sensitive, and the computational complexity for DCS reconstruction is much higher than for CS.

In this paper, we focus on typical WSNs which consist of a large number of SNs distributed in the field to collect information of interest for applications such as geographical monitoring, industrial monitoring, security and climate monitoring. The physical phenomena have significant correlations in both the spatial and the temporal domains. We develop a novel active node selection framework in the presence data loss. In view of the fact that the signal support changes slowly with high temporal correlation [18], [19], the proposed approach selects SNs to be activated by exploiting the support information of the signal reconstructed in the previous time instant. In doing so, this improves reconstruction accuracy, spectrum usage and network lifetime compared with randomly

Wei Chen is with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University 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).

A part of this work has been submitted to the 2014 IEEE Global Communications Conference (GLOBECOM).

activated SNs that are employed in conventional compressive sleeping WSNs [9]. Our contributions can be summarized as follows:

- We propose a novel active node selection framework for compressive sleeping WSNs, where the FC performs an optimized selection of SNs. The node selection problem can be seen as a special sensing matrix design problem where the sensing matrix consists of selected rows of an identity matrix. We note that none of the existing approaches for sensing matrix optimization in the literature [20]–[22] can be directly applied to solve this problem.
- We formulate node selection as an optimization problem that minimizes the lower bound of the mean square error (MSE) of a genie-aided estimator, and approximate the problem of active node selection by a constrained convex relaxation plus a rounding scheme.
- We characterize the performance of the proposed approach via experiments with both synthetic data and real WSN data that embody the effect of various parameters such as signal sparsity, the number of active SNs, the size of the WSN, and the data loss probability, and shed light on the interplay among reconstruction accuracy, spectrum resource usage, and network lifetime performance for compressive sleeping WSNs.

The remainder of the paper is structured as follows. Section II describes in detail the proposed sensing approach including the data acquisition, transmission and reconstruction processes. In Section III, we formulate the active node selection problem that minimize the total number of SNs needed to be activated for compressive sleeping WSNs and derive the proposed approach. Then in Section IV we modify the proposed active node selection in order to take into consideration the network lifetime in addition to reconstruction accuracy. The overhead associated with the proposed centralized node selection scheme and its impact to the energy consumption and spectrum usage are discussed in Section V. Example results on both synthetic data and real WSN data are presented in Section VI, followed by conclusions in Section VII.

The following notation is used. Lower-case letters denote numbers, boldface upper-case letters denote matrices, boldface lower-case letters denote column vectors, and calligraphic upper-case letters denote sets. The superscripts $(\cdot)^T$ and $(\cdot)^{-1}$ denote the transpose and the inverse of a matrix, respectively. The trace of a matrix is denoted by $\mathtt{Tr}(\cdot)$. x_i denotes the ith element of $\mathbf{x}$, $X_{i,i}$ denotes the ith diagonal element of $\mathbf{X}$, and $\mathbf{X}_{\mathcal{J}}$ denotes the submatrix of $\mathbf{X}$ by selecting columns with indexes in the set $\mathcal{J}$. $\mathcal{J}^c$ denotes the complementary set of $\mathcal{J}$. $\mathbb{E}_{\mathbf{x}}(\cdot)$ denotes expectation with respect to the distribution of the random vector $\mathbf{x}$. $\binom{n}{m}$ denotes the number of m combinations from a given set of n elements. $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. $\mathbf{I}_n$ denotes the $n \times n$ identity matrix. The ℓ_0 norm, the ℓ_1 norm, and the ℓ_2 norm of vectors, are denoted by $\|\cdot\|_0$, $\|\cdot\|_1$, and $\|\cdot\|_2$, respectively.

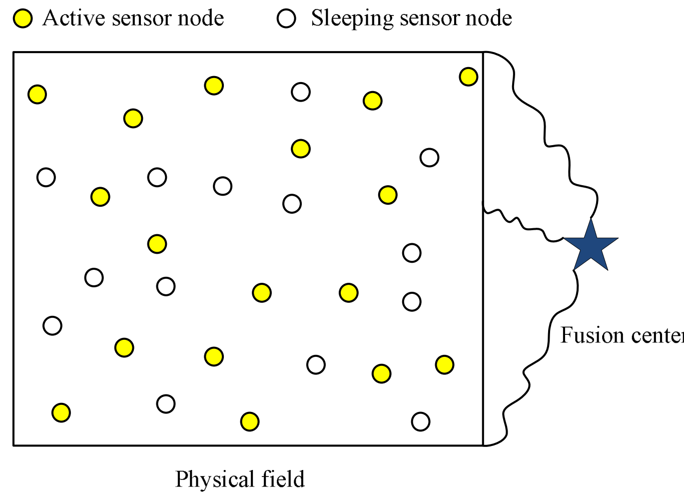

Fig. 1. A WSN with single hop communication.

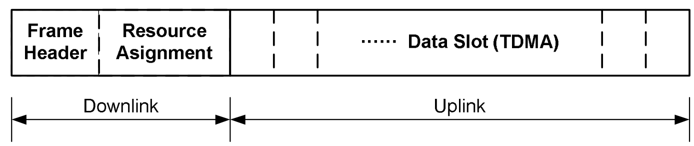

Fig. 2. Frame structure.

II. System Model

In this paper, we consider a typical WSN architecture for collecting physical field information with n SNs and a fusion center (FC) as shown in Fig. 1. While a baseline data gathering approach requires the collection of all SNs' readings, a compressive sleeping WSN only requires a portion of the SNs to be activated to monitor and transmit data [9], [15]. To support such a CS based scheme, we consider the usage of a time division multiplexing access (TDMA) based transmission as shown in Fig. 2. The FC selects SNs to be activated for gathering and reporting data, and coordinates channel access among these active SNs via the downlink. Active SNs' readings are reported to the FC via the uplink. By using such a MAC protocol, collision-free operation is achieved and thus energy wasted due to collisions can be eliminated. In addition, SNs can be turned off in unassigned slots, thus saving energy expenditure due to idle sensing and overhearing. Perfect synchronization is assumed in this paper.

A. Data Acquisition and Transmission

Assume each SN has a monitored parameter f_i $(i = 1, \ldots, n)$ to report to the FC. These data gathered by different SNs have high spatial correlations and thus we can represent the spatial signal $\mathbf{f} \in \mathbb{R}^n$ as a sparse vector $\mathbf{x} \in \mathbb{R}^n$ under some basis $\boldsymbol{\Psi} \in \mathbb{R}^{n \times n}$ such that

$$\mathbf{f} = \boldsymbol{\Psi}\mathbf{x}, \tag{1}$$

where the sparsifying basis $\boldsymbol{\Psi}$ can be determined a-priori, or by using an adaptive approach updated via principal component analysis (PCA) [23] or dictionary learning [24]. Here, sparse means that only s ($s \ll n$) elements in vector $\mathbf{x}$ are nonzeros while all the other elements are zeros, i.e., $\|\mathbf{x}\|_0 = s$.

For a compressive sleeping WSN as in [9], [15], only m ($m < n$) SNs are active for transmitting their monitored physical parameter within different time slots, and all the others SNs are turned off in order to save energy and spectrum resources. As data loss in wireless sensing applications is inevitable owing to transmission medium impairments, the received signal vector $\mathbf{y} \in \mathbb{R}^k$ at the FC can be expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{\Phi}\mathbf{f} + \mathbf{z}, \tag{2}$$

where $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}_k)$ denotes the noise term for the measuring process, $\mathbf{\Phi} \in \mathbb{R}^{m\times n}$ denotes the projection matrix where the entries are all zeros except for m entries in m different columns and rows, and $\mathbf{H} \in \mathbb{R}^{k\times m}$ ($k < m$) is the packet loss matrix where the entries are all zeros except for k entries in k different columns and rows. The entries of the projection matrix $\mathbf{\Phi}$ correspond to the SNs' states, i.e., sleeping or active. For example, if SN i is active, then there is a unity element in the ith column of $\mathbf{\Phi}$, otherwise all the elements of the ith column are zeros. Similarly, the entries of the packet loss matrix $\mathbf{H}$ correspond to the successfully collected measurements of different SNs. For example, if there is a unity element in the ith column of $\mathbf{H}$, then the data packet of SN i is received, otherwise it is lost in transmission. The FC can determine if an SN's message is lost via a cyclic redundancy code (CRC) check. Therefore, according to (1) and (2), we can rewrite the received signal vector as

$$\mathbf{y} = \mathbf{H}\mathbf{\Phi}\mathbf{\Psi}\mathbf{x} + \mathbf{z} = \mathbf{A}\mathbf{x} + \mathbf{z}, \tag{3}$$

where $\mathbf{A} = \mathbf{H}\mathbf{\Phi}\mathbf{\Psi} \in \mathbb{R}^{k\times n}$ denotes the equivalent sensing matrix.

Note that the sensing noise $\mathbf{z}$ is caused by the limitations of the sensing devices, while transmission medium impairments such as fading and additive channel noise lead to packet errors and its impact is captured by the matrix $\mathbf{H}$.

Without loss of generality, we assume measurements are in the form of equal-sized packets and take one time slot for transmission. Thus, the number of active SNs, i.e., m, which determines the number of required time slots for transmission, can be used to indicate the efficiency of spectrum usage.

B. Data Reconstruction

The typical signal reconstruction process behind conventional CS approaches involves solving the following optimization problem to recover the original signal:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathbf{x}\|_1 \\ \text{s.t.} \quad & \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 \le \epsilon, \end{aligned} \tag{4}$$

where ϵ is an estimate of the noise level. It has been demonstrated that only $\mathcal{O}(s \log \frac{n}{s})$ measurements [25] are required for robust reconstruction in the CS framework.

In WSNs, the data measurement acquired from sensing natural phenomena have significant correlations in both the spatial and the temporal domain. For example, the data of the sensor readings provided by the Intel Berkeley Research lab [26] exhibits both high spatial correlations by examining the data from all the SNs at one time instant and high temporal correlations by examining the data from one SN in consecutive time intervals as shown in [11]. To reduce the number of projections required to recover the signal, it is desired to exploit both spatial correlations and temporal correlations. While compressive sleeping WSNs only leverage spatial correlations, temporal correlation is exploited in the proposed selection of active SNs. The signal support is assumed to change slowly in time owing to the temporal correlations, and thus the signal support of previous time instant can be used as a rough estimate of the signal support of current time instant [18], [19]. In [19], the following optimization problem is proposed for recovering the original signal:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathbf{x}_{\mathcal{J}^c}\|_1 \\ \text{s.t.} \quad & \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 \le \epsilon, \end{aligned} \tag{5}$$

where $\mathcal{J}$ denote the estimate of the signal support.

III. The Active Node Selection Framework

In this section, we provide the active node selection framework for compressive sleeping WSNs in order to minimize the total number of SNs needed be activated. The proposed approach, in conjunction with the consideration of network characteristics, can be applied to improve the network lifetime, which will be presented in Section IV.

A. Problem Formulation

For a conventional compressive sleeping WSN as given in [9], [15], SNs are randomly activated with a predetermined probability in a distributed manner. However, this approach has the following drawbacks: i) it lacks flexibility to vary the number of active SNs according to the signal sparsity or the channel condition which could be time-varying; ii) the projection matrix is random rather than an optimized one, and optimized projection matrices have been shown to offer superior performance in various applications [20]–[22]. Existing optimized projection matrix designs in the literature cannot be applied to the compressive sleeping WSN whose projection matrix is restricted to m unity entries.

We define an $n \times n$ diagonal matrix $\tilde{\mathbf{\Phi}}$ such that

$$\tilde{\Phi}_{i,i} = \begin{cases} 1, & \text{SN } i \text{ is activated} \\ 0, & \text{otherwise} \end{cases}. \tag{6}$$

Then, $\tilde{\mathbf{\Phi}}$ can be written as a row-permutation of the concatenation of the sensing matrix $\mathbf{\Phi}$ and a matrix of $n - m$ rows of zeros, which is given by

$$\tilde{\mathbf{\Phi}} = \mathbf{\Pi} \begin{bmatrix} \mathbf{\Phi} \\ \mathbf{0}_{(n-m)\times n} \end{bmatrix}, \tag{7}$$

where $\mathbf{\Pi}$ is a row-permutation matrix, and $\mathbf{\Phi}^T\mathbf{\Phi} = \tilde{\mathbf{\Phi}}$.

We now provide a rationale for the proposed framework. The goal relates to the minimization of the MSE and penalty of the node selection subject to appropriate constraints, which is given as follows

$$\begin{aligned}\min_{\tilde{\Phi}_{i,i}} \quad & \mathbb{E}_{\mathbf{z},\mathbf{H}}\big(\|\mathcal{F}(\mathbf{y},\mathbf{H},\tilde{\mathbf{\Phi}},\mathbf{\Psi})-\mathbf{x}\|_2^2\big) && \text{(8a)}\\ \text{s.t.} \quad & \tilde{\Phi}_{i,i}\in\{0,1\},\ i=1,\ldots,n, && \text{(8b)}\\ & \mathtt{Tr}(\tilde{\mathbf{\Phi}})=m, && \text{(8c)}\end{aligned}$$

where $\mathcal{F}(\cdot)$ denotes an estimator of the sparse signal representation.

The derivation of such a node selection design is very difficult though, because the squared reconstruction error term in (8a) depends upon the actual estimator. In view of the lack of closed-form tractable squared reconstruction error expressions for actual estimators, we use a genie-aided reconstruction procedure that is assumed to know the real sparse representation support and performs least squares (LS) estimation based on prior knowledge of the support. The genie-aided reconstruction has also been used for the design of projection matrices [22] and in the analysis of the performance of various reconstruction approaches [27], [28]. Here we assume that the estimated signal representation support $\mathcal{J}$ is the actual support and is used in the genie-aided reconstruction.

With the use of the signal representation support $\mathcal{J}$, the solution of the LS estimation can be given by

$$\begin{aligned}\hat{\mathbf{x}} =&(\mathbf{\Psi}_{\mathcal{J}}^T\mathbf{\Phi}^T\mathbf{H}^T\mathbf{H}\mathbf{\Phi}\mathbf{\Psi}_{\mathcal{J}})^{-1}\mathbf{\Psi}_{\mathcal{J}}^T\mathbf{\Phi}^T\mathbf{H}^T\mathbf{y}\\ =&(\mathbf{\Psi}_{\mathcal{J}}^T\mathbf{\Phi}^T\mathbf{H}^T\mathbf{H}\mathbf{\Phi}\mathbf{\Psi}_{\mathcal{J}})^{-1}\mathbf{\Psi}_{\mathcal{J}}^T\mathbf{\Phi}^T\mathbf{H}^T\mathbf{H}\mathbf{\Phi}\mathbf{\Psi}\mathbf{x}\\ &+(\mathbf{\Psi}_{\mathcal{J}}^T\mathbf{\Phi}^T\mathbf{H}^T\mathbf{H}\mathbf{\Phi}\mathbf{\Psi}_{\mathcal{J}})^{-1}\mathbf{\Psi}_{\mathcal{J}}^T\mathbf{\Phi}^T\mathbf{H}^T\mathbf{z},\end{aligned} \tag{9}$$

and the MSE of the LS estimation is

$$\mathbb{E}_{\mathbf{z},\mathbf{H}}\big(\|\hat{\mathbf{x}}-\mathbf{x}\|_2^2\big)=\sigma^2\mathbb{E}_{\mathbf{H}}\Big(\mathtt{Tr}\big((\mathbf{\Psi}_{\mathcal{J}}^T\mathbf{\Phi}^T\mathbf{H}^T\mathbf{H}\mathbf{\Phi}\mathbf{\Psi}_{\mathcal{J}})^{-1}\big)\Big). \tag{10}$$

We note that the reconstruction performance in (10) is affected by the interplay among the sparsifying matrix $\mathbf{\Psi}$, the node selection matrix $\mathbf{\Phi}$ and the packet loss matrix $\mathbf{H}$. However, the node selection matrix has to be determined before data transmission, and thus $\mathbf{H}$ is unknown. While it is difficult to predict instantaneous $\mathbf{H}$ for a time varying wireless channel before transmission, it is possible to estimate the expected packet reception success probability given knowledge of the modulation type, the signal-to-noise ratio (SNR), the packet length and the channel type, e.g., additive white Gaussian noise (AWGN) channel or Rayleigh channel [29].

Assuming that packet loss of different SNs are independently distributed, we have

$$\mathbb{E}_{\mathbf{H}}\left(\mathbf{H}^T\mathbf{H}\right)=\mathbf{Q}, \tag{11}$$

where $\mathbf{Q}\in\mathbb{R}^{m\times m}$ is a diagonal matrix with diagonal elements q_i $(i=1,\ldots,m)$ denoting the probability that a packet is successfully received. The following Proposition is applied to enable a node selection design without the explicit usage of $\mathbf{H}$.

Proposition 1: Let $\mathbf{P}$ be an $a\times b$ matrix with rank b and $\mathbf{Q}$ be an $a\times a$ diagonal matrix with nonnegative numbers. Then $\mathtt{Tr}\big((\mathbf{P}^T\mathbf{Q}\mathbf{P})^{-1}\big)$ is convex in $q_{i,i}$ $(i=1,\ldots,a)$.

Proof: We first prove that the expression $\mathtt{Tr}\big((\mathbf{X})^{-1}\big)$ is convex in $\mathbf{X}$ if $\mathbf{X}$ is a positive definite symmetric matrix. Let

$$g(t)=\mathtt{Tr}\big((\mathbf{X}+t\mathbf{V})^{-1}\big), \tag{12}$$

where $\mathbf{X}\succ 0$ and $\mathbf{V}$ is a symmetric matrix. Now we can rewrite $g(t)$ as

$$\begin{aligned}g(t)&=\mathtt{Tr}\big((\mathbf{X}+t\mathbf{V})^{-1}\big)\\ &=\mathtt{Tr}\big(\mathbf{X}^{-1}-t(\mathbf{X}+t\mathbf{V})^{-1}\mathbf{V}\mathbf{X}^{-1}\big)\\ &=\mathtt{Tr}\big(\mathbf{X}^{-1}-t\mathbf{X}^{-1}\mathbf{V}\mathbf{X}^{-1}+t^2\mathbf{X}^{-1}\mathbf{V}\mathbf{X}^{-1}\mathbf{V}\mathbf{X}^{-1}\\ &\quad-t^3\mathbf{X}^{-1}\mathbf{V}\mathbf{X}^{-1}\mathbf{V}\mathbf{X}^{-1}\mathbf{V}\mathbf{X}^{-1}+\ldots\big),\end{aligned} \tag{13}$$

where the first equality can be proved by using the Searle set of identities [30]. Then we have the second derivative that can be derived as

$$\lim_{t\to 0}\frac{\partial^2 g(t)}{\partial t^2}=\mathtt{Tr}\big(\mathbf{X}^{-1}\mathbf{V}\mathbf{X}^{-1}\mathbf{V}\mathbf{X}^{-1}\big)=\mathtt{Tr}\big(\mathbf{W}^T\mathbf{X}^{-1}\mathbf{W}\big), \tag{14}$$

where $\mathbf{W}=\mathbf{V}\mathbf{X}^{-1}$. As $\mathbf{X}$ is positive definite, then we have $\mathbf{W}^T\mathbf{X}^{-1}\mathbf{W}\succ 0$. Therefore, $\mathtt{Tr}\big((\mathbf{X})^{-1}\big)$ is convex in $\mathbf{X}$.

As $\mathbf{P}^T\mathbf{Q}\mathbf{P}$ is invertible and $q_{i,i}$ $(i=1,\ldots,a)$ are nonnegative, we have $\mathbf{P}^T\mathbf{Q}\mathbf{P}\succ 0$. Thus, we can conclude that $\mathtt{Tr}\big((\mathbf{P}^T\mathbf{Q}\mathbf{P})^{-1}\big)$ is convex in $q_{i,i}$ $(i=1,\ldots,a)$ in view of the fact that $\mathtt{Tr}\big((\mathbf{X})^{-1}\big)$ is convex in $\mathbf{X}$ if $\mathbf{X}$ is positive definite and the concavity of a function is preserved under an affine transformation. ■

From Proposition 1 and Jensen's inequality, the MSE of the LS estimation given in (10) can be lower bounded by

$$\begin{aligned}\mathbb{E}_{\mathbf{z},\mathbf{H}}\big(\|\hat{\mathbf{x}}-\mathbf{x}\|_2^2\big)&\geq\sigma^2\mathtt{Tr}\Big(\Big(\mathbf{\Psi}_{\mathcal{J}}^T\mathbf{\Phi}^T\mathbb{E}_{\mathbf{H}}\big(\mathbf{H}^T\mathbf{H}\big)\mathbf{\Phi}\mathbf{\Psi}_{\mathcal{J}}\Big)^{-1}\Big)\\ &=\sigma^2\mathtt{Tr}\Big(\Big(\mathbf{\Psi}_{\mathcal{J}}^T\mathbf{Q}^{1/2}\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{Q}^{1/2}\mathbf{\Psi}_{\mathcal{J}}\Big)^{-1}\Big)\\ &=\sigma^2\mathtt{Tr}\Big(\Big(\mathbf{\Psi}_{\mathcal{J}}^T\mathbf{Q}^{1/2}\tilde{\mathbf{\Phi}}\mathbf{Q}^{1/2}\mathbf{\Psi}_{\mathcal{J}}\Big)^{-1}\Big).\end{aligned} \tag{15}$$

Then, according to (8) and (15), we put forth the following optimization problem:

$$\begin{aligned}\min_{\tilde{\Phi}_{i,i}} \quad & \mathtt{Tr}\big((\mathbf{\Psi}_{\mathcal{J}}^T\mathbf{Q}^{1/2}\tilde{\mathbf{\Phi}}\mathbf{Q}^{1/2}\mathbf{\Psi}_{\mathcal{J}})^{-1}\big)\\ \text{s.t.} \quad & \tilde{\Phi}_{i,i}\in\{0,1\},\ i=1,\ldots,n,\\ & \mathtt{Tr}(\tilde{\mathbf{\Phi}})=m.\end{aligned} \tag{16}$$

This optimization problem defines the node selection matrix that minimizes the lower bound of the oracle MSE regarding data loss. Unfortunately, the optimization problem in (16) is non-convex. As the variables $\tilde{\Phi}_{i,i}$ are binary integers, a straightforward way to solve (16) is to perform an exhaustive search over $\binom{n}{m}$ different combinations of m active nodes. However, the complexity of the exhaustive search is impractical for a network having a large number of SNs.

B. Active Node Selection Via Convex Relaxation

In this subsection, we formulate the active node selection problem in (16) as a relaxed convex optimization problem that can be solved efficiently using numerical methods such as interior-point algorithms. By relaxing the binary integer constraints so that $\tilde{\Phi}_{i,i}$ can be in the range from 0 to 1, we can express the active node selection problem as follows

$$\min_{\tilde{\Phi}_{i,i}} \quad \mathtt{Tr}\big((\mathbf{\Psi}_{\mathcal{J}}^T\mathbf{Q}^{1/2}\tilde{\mathbf{\Phi}}\mathbf{Q}^{1/2}\mathbf{\Psi}_{\mathcal{J}})^{-1}\big) \tag{17a}$$

$$\text{s.t.} \quad 0 \leq \tilde{\Phi}_{i,i} \leq 1,\ i = 1, \ldots, n, \tag{17b}$$

$$\mathtt{Tr}(\tilde{\mathbf{\Phi}}) = m. \tag{17c}$$

The concavity of the optimization problem (17) can be unveiled by Proposition 1.

This relaxation of an optimization problem with binary integers to a convex form makes the problem much easier to solve than the original integer program. After solving (17), the m largest $\tilde{\Phi}_{i,i}$ can be chosen and the corresponding indexes relate to the selected nodes. This relaxation for binary integer constraints has also been used for antenna selection in multi-antenna wireless communication systems [31], and for sensor selection in parameter estimation [32].

Note that the proposed node selection framework for compressive sleeping WSNs capitalizes on the genie-aided reconstruction procedure which knows the accurate signal support. The use of genie-aided reconstruction aims to unveil the rationale of the proposed approach, and it does not imply that our method is bound to fail when the support information is only partially correct. Also, we observe that the signal support is likely to change slowly with time for practical applications. The effect of the estimation accuracy regarding the actual support is investigated in Section V, and performance gains are observed both in experiments using synthetic data with partially correct support estimation, and in experiments using real WSN data.

IV. Prolonging The Network Lifetime

The proposed active node selection procedure can be modified in order to satisfy some additional requirements. In this section, we apply the proposed active node selection approach in order to improve the lifetime of compressive sleeping WSNs.

Network lifetime is determined by the time instant when the network cannot support application-specific functions. For example, it can be the instant when the first SN runs out of energy, or a specified fraction of SNs run out of energy, or the network partitions. In [33], Chen and Zhao investigate an approach to network lifetime which can be applied to any definition of the network lifetime and holds independently of the underlying network model. They derive the average network lifetime $\mathbb{E}(L)$ as

$$\mathbb{E}(L) = \frac{\xi_0 - \mathbb{E}(E_u)}{P_c + \rho\mathbb{E}(E_t)}, \tag{18}$$

where ξ_0 is the initial total network energy, P_c is the constant continuous power consumption over the whole network, $\mathbb{E}(E_u)$is the average total unused energy in the network when it dies, ρ is the average data reporting rate, and $\mathbb{E}(E_t)$ is the average transmission energy consumed by all sensors. According to (18), in order to prolong the network lifetime, it is desired to reduce the unused energy $\mathbb{E}(E_u)$ and the transmission energy consumption $\mathbb{E}(E_t)$. By defining ξ_i and η_i as the required energy for transmission of a message and the node energy storage for SN i, respectively, Chen and Zhao propose a greedy node selection approach that recursively activates the SN with the maximum value of $\eta_i - \xi_i$.

Chen and Zhao's approach can be directly applied to compressive sleeping WSNs to prolong the network lifetime. However, this approach has the drawback that it fails to guarantee the reconstruction accuracy and the characteristics of the monitored signal are not exploited. Our proposed active node selection given in Section III aims to reduce the number of SNs, while it does not involve other factors which also have an influence on the network lifetime. In order to modify our active node selection approach to explicitly address the problem of prolonging network lifetime, we put forth the following optimization problem

$$\begin{aligned} \min_{\tilde{\Phi}_{i,i}} \quad & \mathtt{Tr}\big((\mathbf{\Psi}_{\mathcal{J}}^T\mathbf{Q}^{1/2}\tilde{\mathbf{\Phi}}\mathbf{Q}^{1/2}\mathbf{\Psi}_{\mathcal{J}})^{-1}\big) + \beta\mathtt{Tr}(\tilde{\mathbf{\Phi}}\mathbf{P}) \\ \text{s.t.} \quad & 0 \leq \tilde{\Phi}_{i,i} \leq 1,\ i = 1, \ldots, n, \\ & \mathtt{Tr}(\tilde{\mathbf{\Phi}}) = m, \end{aligned} \tag{19}$$

where $\mathbf{P} \in \mathbb{R}^{n\times n}$ is a diagonal matrix with $P_{i,i} \geq 0$ and $\beta > 0$ denotes the weight for the influence of $\mathbf{P}$. The term $\mathtt{Tr}(\tilde{\mathbf{\Phi}}\mathbf{P})$ denotes the penalty function of the weighted active node selection matrix $\tilde{\mathbf{\Phi}}\mathbf{P}$ where $P_{i,i}$ $(i = 1, \ldots, n)$ is the penalty for selecting SN i. Various design targets could result in different interpretations of the penalty term. For example, SNs could consume different amounts of energy to communicate their readings to the FC owing to their distinct transmission pathloss values, where the penalty term $\tilde{\mathbf{\Phi}}\mathbf{P}$ could reflect the constraint on the total energy consumed by the selected SNs by setting $P_{i,i}$ to be the energy consumption of the ith SN. Alternatively, it could be desirable to wake up only those SNs having adequate levels of stored energy rather than those SNs already low in stored energy, so in this case an SN with inadequate energy is associated with a relatively large penalty.

In this paper, inspired by the average network lifetime formula (18) given in [33], we formulate the penalty $P_{i,i}$ as

$$P_{i,i} = -\frac{\eta_i}{\xi_i}. \tag{20}$$

This penalty promotes the usage of an SN with large energy storage η_i and/or small transmission cost ξ_i. This penalty expression, in conjunction with the proposed framework (16), aims to prolong the network lifetime by avoiding excessive use of any particular SN, and the active nodes are selected by solving the following problem:

$$\begin{aligned} \min_{\tilde{\Phi}_{i,i}} \quad & \mathtt{Tr}\big((\mathbf{\Psi}_{\mathcal{J}}^T\mathbf{Q}^{1/2}\tilde{\mathbf{\Phi}}\mathbf{Q}^{1/2}\mathbf{\Psi}_{\mathcal{J}})^{-1}\big) - \beta\sum_{i=1}^{n}\frac{\eta_i\tilde{\Phi}_{i,i}}{\xi_i} \\ \text{s.t.} \quad & 0 \leq \tilde{\Phi}_{i,i} \leq 1,\ i = 1, \ldots, n, \\ & \mathtt{Tr}(\tilde{\mathbf{\Phi}}) = m. \end{aligned} \tag{21}$$

The m largest $\tilde{\Phi}_{i,i}$ of the solution are chosen and the corresponding indexes relate to the selected nodes.

V. Discussion: Centralized Node Selection vs. Distributed Node Selection

In the previous work [9], [15], sensor nodes are activated randomly, which can be achieved either in a centralized manner, i.e., nodes are selected by the fusion center as in the proposed approach, or in a distributed manner, for example, the states of nodes in different slots could be predetermined randomly using their MAC address as seeds. The distributed scheme has the following drawbacks: i) the node selection is random rather than an optimized one, ii) it lacks flexibility to vary the number of active SNs according to the signal sparsity or the channel condition which could be time-varying, while the centralized scheme has the drawback that extra bandwidth resource and energy are required to communicate active nodes IDs via the downlink.

In order to compare the proposed centralized node selection scheme against the distributed random selection scheme, one needs to consider the overhead associated with the proposed scheme. We assume that the length of each sensing data packet to be b^{tx} bits and the overhead to be b^{rx} bits. We use the energy model given in [34], where the energy required by the ith SN to send b^{tx} bits to the FC separated by a distance of $\hat{r}_i$ is

$$E_i^{tx} = E^e b_i^{tx} + \bar{E}^{tx} b^{tx} \hat{r}_i^{\alpha}, \tag{22}$$

where E^e denotes the energy associated with the radio electronics, $\bar{E}^{tx}$ denotes the energy consumption rate in transmission to achieve a target signal strength with a unit distance between the FC and the SN, and α represents the path loss exponent. The energy required by the ith SN to receive b^{rx} bits from the FC is

$$E_i^{rx} = E^e b^{rx}. \tag{23}$$

Therefore, the energy consumed by the ith SN in the sleep mode and active mode for the transmission of one packet via the proposed scheme can be modelled as

$$E_i = \begin{cases} E^c + E^e b^{rx} & \texttt{sleep mode} \\ E^e(b^{rx} + b^{tx}) + \bar{E}^{tx} b^{tx} \hat{r}_i^{\alpha} & \texttt{active mode} \end{cases}, \tag{24}$$

where E^c denotes the continuous level of energy consumption needed to sustain the network without data collection in the sleep mode.

Here we present a first-order comparison of the proposed node selection scheme against the distributed random node selection scheme. We assume m^p SNs need to be activated in the proposed scheme to achieve a target reconstruction accuracy while the random scheme requires m^r active SNs. We also assume the distances between the FC and all the SNs are equal to $\hat{r}$, and both the downlink and the uplink use the same modulation scheme. Then the energy consumption ratio and spectrum usage ratio of the proposed scheme against the random scheme can be expressed by

$$G^e = \frac{m^p(E^e(b^{rx} + b^{tx}) + \bar{E}^{tx} b^{tx} \hat{r}^{\alpha}) + (n - m^p)(E^c + E^e b^{rx})}{m^r(E^e b^{tx} + \bar{E}^{tx} b^{tx} \hat{r}^{\alpha}) + (n - m^r) E^c}, \tag{25}$$

and

$$G^s = \frac{b^{rx} + m^p b^{tx}}{m^r b^{tx}}, \tag{26}$$

respectively. According to (25) and (26), the gain brought by the proposed approach in terms of energy consumption and spectrum usage, depends upon the interplay among the number of active SNs, the network size, the data packet size in the downlink and uplink, and the parameters associated with the energy model. As illustrated by Fig. 3 and Fig. 4, both the energy consumption performance and spectrum usage performance of the proposed approach against the distributed random scheme tend to worsen with increasing overhead, i.e., the size of the message sent to active SNs, and also with the number of active SNs in the proposed scheme. More comprehensive evaluation of the performance of the proposed approach is given in Section VI.

Note that in this paper we only investigate the benefits brought by the proposed node selection optimization, although the centralized scheme facilitates the implement of other techniques to further enhance the performance. For example, the number of active SNs can be adjusted according to the signal sparsity to improve the reconstruction performance and

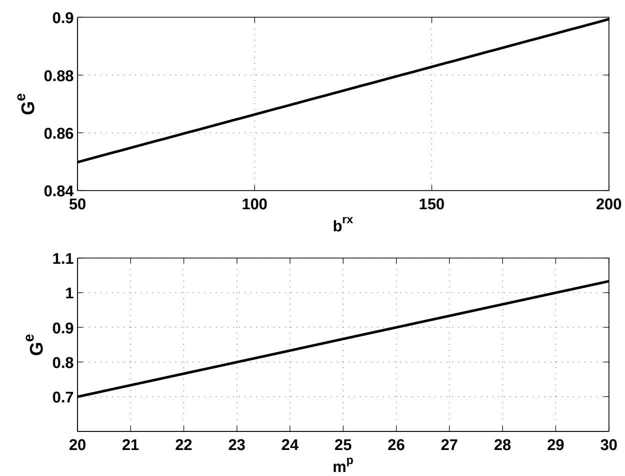

Fig. 3. Energy consumption performance of the proposed approach against the distributed random scheme ($n = 100$, $E^e = \bar{E}^{tx} = 1$, $E^c = 0.1$, $\hat{r} = 10$, $\alpha = 2$, $m^r = 30$, $b^{tx} = 100$, and $m^p = 25$ for the upper sub-figure and $b^{rx} = 100$ for the lower sub-figure).

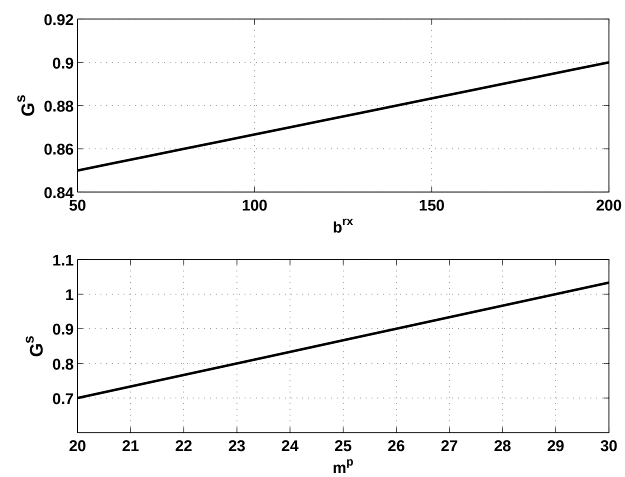

Fig. 4. Spectrum usage performance of the proposed approach against the distributed random scheme ($m^r = 30$, $b^{tx} = 100$, and $m^p = 25$ for the upper sub-figure and $b^{rx} = 100$ for the lower sub-figure).

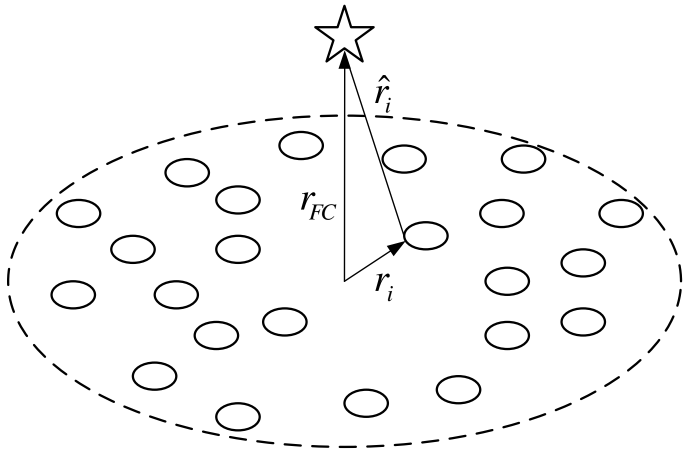

Fig. 5. WSN deployment in synthetic experiments.

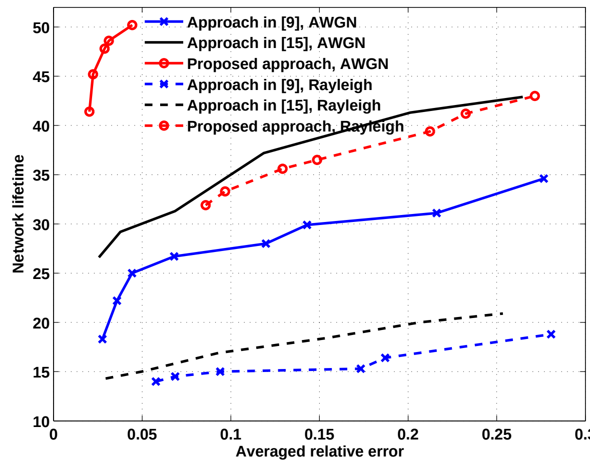

Fig. 6. The trade-off between the network lifetime and the reconstruction accuracy ($n = 50$, $s = 5$, $E_c = 1$, $E^e = \bar{E}^{tx} = 1$, $E_0 = 10^5$, $\alpha = 2$, $\beta = 10^{-3}$ and 20% incorrect support estimate for the proposed approach).

prolong network lifetime[1].

VI. Performance Results

In this section, by firstly using synthetic data and then using real data collected by the WSN located in the Intel Berkeley Research Lab [26], we demonstrate that the proposed active node selection approach achieves improved performance in terms of signal monitoring accuracy, network lifetime and spectrum resource compared with other approaches.

For all the experiments, we consider SNs powered by non-rechargeable batteries. We set the length of each sensing data packet and the downlink data packet to be 64 bits, and assume the modulation scheme is binary phase shift keying (BPSK). We consider an AWGN channel and also a Rayleigh fading channel, which leads to different data loss probabilities. We use CVX, a package for specifying and solving convex programs [35], to solve the CS inverse problems and to reconstruct signals. We evaluate the reconstruction performance by using averaged relative error which is the average of $\frac{\|\mathbf{x}\|_2^2}{\|\hat{\mathbf{x}}-\mathbf{x}\|_2^2}$ over 1000 trials. The network lifetime is determined as the time instant when the first SN dies[2]. As a comparison, we also provide the performance of the approach in [9], which only exploit the spatial correlations of SNs' messages, and the approach in [15], which exploits both spatial and temporal correlations[3].

A. Synthetic Experiments

In the synthetic experiments, n SNs are randomly placed in a disc of unit radius as shown in Fig. 5. Similar to [36], the FC is placed at a distance $r_{FC} = 3$ above the center of the disc, and r_i, i.e., the radial distance of SN i, is modeled as a random variable with the probability density function $\frac{3}{2*10^{3/2}}\sqrt{r_i}$ $(0 \leq r_i \leq 10)$[4]. The initial energy of all SNs is E_0. The sparse signal representations $\mathbf{x}$ is randomly generated with ambient dimension n for different time instants, where s non-zero components are drawn from an independent and identically distributed (i.i.d.) zero mean and unit variance Gaussian distribution. The dictionary $\mathbf{\Psi}$ is constructed by first creating a $n \times n$ matrix with i.i.d. draws of a Gaussian distribution $\mathcal{N}(0, 1)$, and then the columns of $\mathbf{\Psi}$ are normalized to a unit norm. The SNR of the wireless channel is set to be 20 dB, which together with the modulation type and packet length determines the data loss probability [29]. In addition, the sensor measurements are corrupted by additive zero-mean Gaussian noise yielding an SNR of 20 dB.

We first present a comparison of various approaches in Fig. 6, which sheds light on the trade-off between the network lifetime and the reconstruction accuracy for compressive sleeping WSNs. Each plotted point represents particular averaged relative error and network lifetime pair and is a result from the use of different numbers of activated SNs. To study how the proposed approach performs, we consider the non-ideal support estimate case where only 80% of the elements of the support estimate are correct. It can be seen that the network lifetime tends to increase with rising levels of averaged relative error, which shows that one has to pay the cost of reduced network lifetime in order to improve the signal reconstruction performance in compressive sleeping WSNs. It is observed that the proposed approach leads to a better lifetime-accuracy curve in comparison to the other approaches.

The reconstruction performance for different approaches using various numbers of time slots for data collection are given in Fig. 7, where we also consider a non-ideal case such

[1]Akin to the technique proposed in [12], by applying cross validation at the FC, the number of activated sensors can be adjusted to maintain an acceptable reconstruction performance while minimising the number of active sensors.

[2]This lifetime definition is somewhat strict and may not apply to some WSN applications. However, it sheds light on the unbalanced energy consumption of SNs and provides insights on node selection design.

[3]The approach in [15] uses spatial correlation and temporal correlation separately to reconstruct SNs' readings. The latency of that approach is the number of time slots taken to recover the data. Utilizing more time slots for reconstruction improves the performance of that approach, but also leads to longer latency, which could violate the required delay sensitivity of the specific application. We assume SNs' messages having a length of N time slots are reconstructed by the approach in [15].

[4]The constant term, $\frac{3}{2*10^{3/2}}$, ensures that probability density function integrates to one over the range 0 to 10.

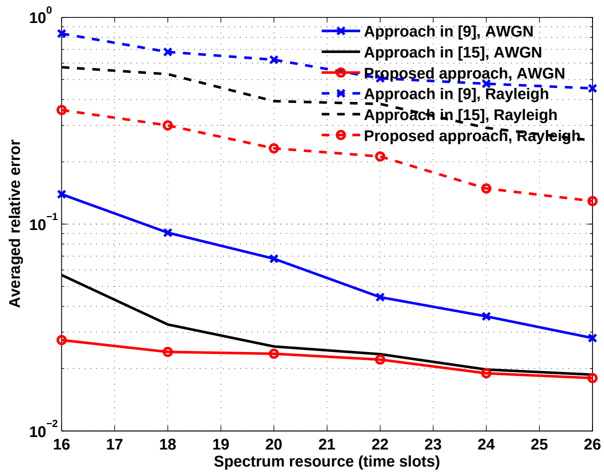

Fig. 7. Reconstrunction accuracy vs. spectrum resource ($n = 50$, $s = 5$, $E_c = 1$, $E^e = \bar{E}^{tx} = 1$, $E_0 = 10^5$, $\alpha = 2$, $\beta = 10^{-3}$ and 20% incorrect support estimate for the proposed approach).

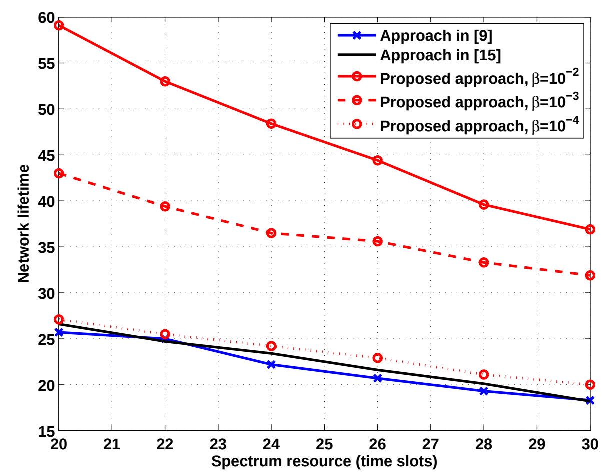

Fig. 8. Network lifetime vs. spectrum resource ($n = 50$, $s = 5$, $\alpha = 2$, $E_c = 1$, $E^e = \bar{E}^{tx} = 1$ and $E_0 = 10^5$).

that 20% of the estimated support is incorrect. We assume the spectrum resource used in the downlink is equal to the spectrum resource used in a time slot of the uplink. Fig. 7 demonstrates the superiority of the proposed approach in terms of the reconstruction accuracy as a function of spectrum resources. To achieve a target averaged relative error, fewer time slots are required for the proposed approach than for the other schemes.

The full reconstruction performance comparison with various settings of other parameters such as signal sparsity s, number of WSN nodes n, number of active SNs m and the accuracy of support estimation, are summarized in Table I. Again, it is demonstrated in Table I that the proposed approach outperforms the other approaches, and that the reconstruction accuracy tends to improve with a reduction in signal sparsity s and the size of WSN n. It is also clear that the reconstruction performance of the proposed approach is affected by the accuracy of the support estimate, and that more SNs are required to be activated to achieve a target reconstruction performance with incorrect support estimates.

We now investigate the network lifetime performance for different schemes in compressive sleeping WSNs. Fig. 8 shows that the network lifetime tends to decrease with an increased usage of spectrum resources. We note that in comparison to random node selection in [9] and [15], the network lifetime is significantly improved by using the proposed active node selection approach. It is also observed in Fig. 8 that the network lifetime gain for the proposed approach is enhanced by using a larger β that represents the level of penalty for selecting SNs with different levels of energy storage and energy consumption. Fig. 8 together with Fig. 7, indicates that by increasing the usage of spectrum resources, the reconstruction accuracy is improved at the cost of a shortened network lifetime, which is consistent with our observations from Fig. 6.

In the previous experiments, we fixed the size of the overhead for the proposed centralized node selection scheme, and we also fixed E^e, i.e., energy associated with the radio electronics, which affects the energy consumed in receiving the overhead data. To evaluate the impact of the overhead on the performance, we compare the proposed centralized node selection scheme with the distributed node selection schemes in [9], [15] with different parameter settings associated with the overhead in Table II. It is observed that the performance of the proposed scheme in terms of data acquisition accuracy and network lifetime deteriorates with increasing, i.e., b^{rx} and E^e. Thus, the proposed approach is more appropriate for data intensive applications where the energy and spectrum resource cost due to the overhead can be neglected in view of the large amount of data sent to the FC from each SN.

B. Experiments With Real Data

For the real data experiments, we use temperature data collected by a WSN located in the Intel Berkeley Research lab [26]. In this WSN, 54 Mica2Dot SNs with weather boards were deployed in one floor of the lab building.

We assume that the FC is placed at a height of $r_{FC} = 5$ (meters) at the center of the floor, and r_i, i.e., the distance between SN i and the center of the floor, is calculated by using deployment information given in [26]. The initial battery energy E_0 for various SNs is uniformly distributed within $[10^5, 2 \times 10^5]$, and we set $E_c = 1$, $E^e = \bar{E}^{tx} = 1$ and $\beta = 10^{-1}$. A learned sparsifying basis $\mathbf{\Psi}$ is used which captures the spatial and temporal characteristics from the data. The proposed approach selects active SNs for data collection by using the support of the reconstructed temperature signal of the previous time instant. The SNRs corresponding to the wireless channel for different SNs are drawn independently from a uniform distribution with the range $[0dB, 30dB]$.

Fig. 9 shows the reconstruction accuracy against spectrum resource usage for different approaches. The proposed node selection approach, which exploits the the temporal correlations of the temperature data at distinct time instants, achieves the best reconstruction performance in comparison to both the other approaches.

TABLE I
COMPARISON OF THE AVERAGED RELATIVE ERROR WITH VARIOUS SETTINGS OF PARAMETERS. ($\alpha = 2$, $E_c = 1$, $E^e = \bar{E}^{tx} = 1$, $E_0 = 10^5$ AND $\beta = 10^{-3}$)

	$s=4$ $m=20$ $n=40$ AWGN	$s=8$ $m=20$ $n=40$ AWGN	$s=8$ $m=30$ $n=40$ AWGN	$s=8$ $m=30$ $n=50$ AWGN	$s=8$ $m=30$ $n=50$ Rayleigh
Approach in [9]	0.0336	0.1330	0.0582	0.0612	0.5195
Approach in [15]	0.0227	0.0648	0.0343	0.0393	0.3794
Proposed approach (correct support estimate)	0.0184	0.0360	0.0217	0.0246	0.1329
Proposed approach (25% incorrect support estimate)	0.0198	0.0453	0.0270	0.0308	0.3269

TABLE II
COMPARISON OF DIFFERENT APPROACHES WITH VARIOUS SETTINGS OF PARAMETERS. ($n = 50$, $s = 6$, $\alpha = 2$, $E_c = 1$, $\bar{E}^{tx} = 1$, $E_0 = 10^5$, $\beta = 10^{-3}$ AND AWGN CHANNEL)

	Approach in [9]	Approach in [15]	Proposed $b^{rx}=64$ $E^e=\bar{E}^{tx}$	Proposed $b^{rx}=128$ $E^e=\bar{E}^{tx}$	Proposed $b^{rx}=256$ $E^e=\bar{E}^{tx}$	Proposed $b^{rx}=256$ $E^e=5\bar{E}^{tx}$	Proposed $b^{rx}=256$ $E^e=10\bar{E}^{tx}$
Average relative error	0.0973	0.0422	0.0147	0.0155	0.0164	0.0167	0.0163
Network lifetime	26.9	26.9	41.2	40.2	39.3	24.8	17
Spectrum resource (downlink, uplink)	(0, 20)	(0, 20)	(1, 19)	(2, 18)	(4, 16)	(4, 16)	(4, 16)

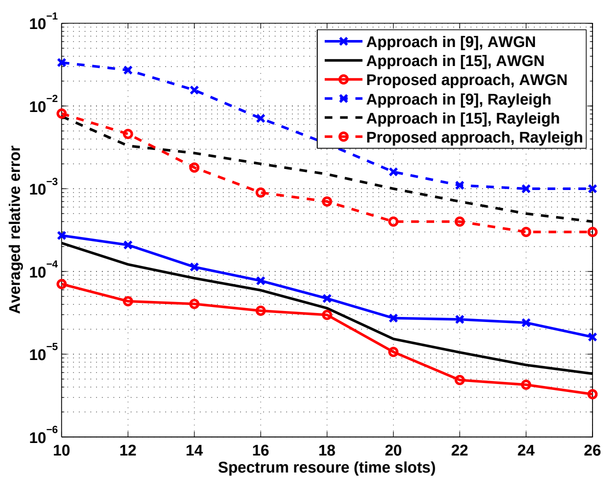

Fig. 9. Reconstrunction accuracy vs. spectrum resource for the Intel Berkeley Research lab data.

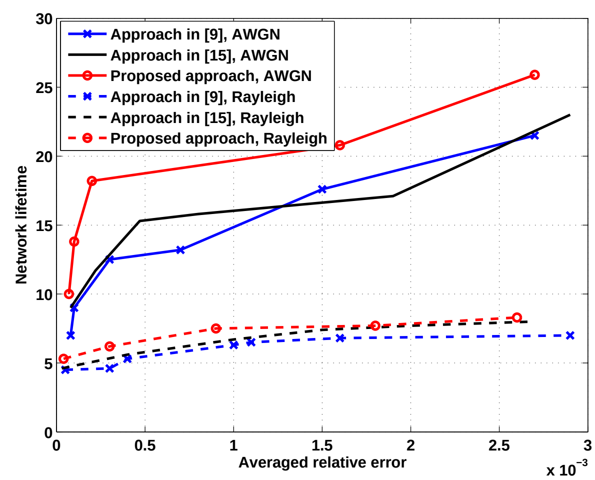

Fig. 10. The trade-off between the network lifetime and reconstruction accuracy for the Intel Berkeley Research lab data.

In the last experiment in Fig. 10, we compare the network lifetime against the reconstruction accuracy for different approaches. The proposed approach achieves the longest lifetime for various values of the averaged relative error, though the performance gain tends to diminish at low averaged relative error owing to the reconstruction performance floor caused by the non-ideal sparsity, i.e., nearly sparse characteristics of the data.

VII. CONCLUSION

In this paper, we propose a novel active node selection framework for compressive sleeping WSNs in order to improve the the performance of data reconstruction accuracy, network lifetime and spectrum resource usage. In addition to the spatial correlation of the data at various SNs, the proposed node selection exploits the temporal correlation, which is not utilised in conventional compressive sleeping WSNs. Capitalizing on the lower MSE bound of the oracle estimator and convex relaxation, the proposed node selection problem can be solved efficiently at the FC by using iterative algorithms. The superiority of our proposed approach in relation to the other approaches in the conventional CS framework is revealed by our experimental study.

Acknowledgment

The authors would like to thank Mr. Fan Xiao for his assistance in preprocessing the data collected by the WSN of the Intel Berkeley Research lab.

References

[1] R. Kulkarni, A. Forster, and G. Venayagamoorthy, "Computational intelligence in wireless sensor networks: A survey," *Communications Surveys Tutorials, IEEE*, vol. 13, no. 1, pp. 68–96, 2011.

[2] W.-L. Yeow, C.-K. Tham, and W.-C. Wong, "Energy efficient multiple target tracking in wireless sensor networks," *Vehicular Technology, IEEE Transactions on*, vol. 56, no. 2, pp. 918–928, March 2007.

[3] V. Gungor and G. Hancke, "Industrial wireless sensor networks: Challenges, design principles, and technical approaches," *Industrial Electronics, IEEE Transactions on*, vol. 56, no. 10, pp. 4258–4265, 2009.

[4] M. C. Vuran and I. Akyildiz, "Spatial correlation-based collaborative medium access control in wireless sensor networks," *Networking, IEEE/ACM Transactions on*, vol. 14, no. 2, pp. 316–329, April 2006.

[5] L. Liu, X. Zhang, and H. Ma, "Optimal node selection for target localization in wireless camera sensor networks," *Vehicular Technology, IEEE Transactions on*, vol. 59, no. 7, pp. 3562–3576, Sept 2010.

[6] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, "In-network aggregation techniques for wireless sensor networks: a survey," *Wireless Communications, IEEE*, vol. 14, no. 2, pp. 70–87, April 2007.

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

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

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

[10] C. Caione, D. Brunelli, and L. Benini, "Distributed compressive sampling for lifetime optimization in dense wireless sensor networks," *Industrial Informatics, IEEE Transactions on*, vol. 8, no. 1, pp. 30–40, 2012.

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

[12] W. Chen and I. Wassell, "Energy-efficient signal acquisition in wireless sensor networks: a compressive sensing framework," *Wireless Sensor Systems, IET*, vol. 2, no. 1, pp. 1–8, 2012.

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

[14] K. Rajawat, A. Cano, and G. Giannakis, "Network-compressive coding for wireless sensors with correlated data," *Wireless Communications, IEEE Transactions on*, vol. 11, no. 12, pp. 4264–4274, December 2012.

[15] T. Xue, X. Dong, and Y. Shi, "Multiple access and data reconstruction in wireless sensor networks based on compressed sensing," *Wireless Communications, IEEE Transactions on*, vol. 12, no. 7, pp. 3399–3411, July 2013.

[16] D. Baron, M. Wakin, M. Duarte, S. Sarvotham, and R. Baraniuk, "Distributed compressed sensing," *Technical Report ECE-0612, Electrical and Computer Engineering Department, Rice University*, Dec. 2006.

[17] M. Duarte, M. Wakin, D. Baron, S. Sarvotham, and R. Baraniuk, "Measurement bounds for sparse signal ensembles via graphical models," *Information Theory, IEEE Transactions on*, vol. 59, no. 7, pp. 4280–4289, 2013.

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

[19] W. Lu and N. Vaswani, "Regularized modified bpdn for noisy sparse reconstruction with partial erroneous support and signal value knowledge," *Signal Processing, IEEE Transactions on*, vol. 60, no. 1, pp. 182–196, 2012.

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

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

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

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

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

[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] P. Bodik, W. Hong, C. Guestrin, S. Madden, M. Paskin, and R. Thibaux. (2004, Feb.) Intel lab data. [Online]. Available: http://db.csail.mit.edu/labdata/labdata.html

[27] E. Candès and T. Tao, "The Dantzig selector: Statistical estimation when p is much larger than n," *The Annals of Statistics*, vol. 35, no. 6, pp. 2313–2351, 2007.

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

[29] J. G. Proakis, "Digital communications," 1995.

[30] S. R. Searle, "Matrix algebra useful for statistics (wiley series in probability and statistics)," 1982.

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

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

[33] Y. Chen and Q. Zhao, "On the lifetime of wireless sensor networks," *Communications Letters, IEEE*, vol. 9, no. 11, pp. 976–978, Nov 2005.

[34] M. Perillo and W. Heinzelman, "An integrated approach to sensor role selection," *Mobile Computing, IEEE Transactions on*, vol. 8, no. 5, pp. 709–720, May 2009.

[35] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.0 beta," http://cvxr.com/cvx, Sep. 2013.

[36] P. Venkitasubramaniam, S. Adireddy, and L. Tong, "Sensor networks with mobile access: optimal random access and coding," *Selected Areas in Communications, IEEE Journal on*, vol. 22, no. 6, pp. 1058–1068, Aug 2004.