Balanced Spatio-Temporal Compressive Sensing for Multi-hop Wireless Sensor Networks

Mohammadreza Mahmudimanesh, Abdelmajid Khelil, Neeraj Suri
Department of Computer Science, Technical University of Darmstadt
{reza, khelil, suri}@deeds.informatik.tu-darmstadt.de

Abstract—**Compressive Sampling (CS) is a powerful sampling technique that allows accurately reconstructing a compressible signal from a few random linear measurements. CS theory has applications in sensory systems where acquiring individual samples is either expensive or infeasible. A Wireless Sensor Network (WSN) is a distributed sensory system comprised of resource-limited sensor nodes. Transferring all the recorded samples in a WSN can easily result in data traffic that can exceed the network capacity. There are ongoing attempts to devise efficient and accurate compression schemes for WSNs and CS has proved to be a key sampling method compared to many other existing techniques. In this paper, specifically targeting the dominant WSN deployments of multi-hop WSNs, we develop a novel CS-based concept of sampling window as an efficient spatio-temporal signal acquisition/compression technique. We show that much higher energy-efficient signal acquisition is possible, if composite temporal and spatial correlations are considered. Our model is also capable of abnormal event detection which is a crucial feature in WSNs. It guarantees balanced energy consumption by the sensor nodes in a multi-hop topology to prevent overloaded nodes and network partitioning.**

I. INTRODUCTION

Wireless Sensor Networks (WSNs) constitute a bridge between information systems and the physical world [1] and are seeing varied applications such as surveillance, monitoring, industrial automation, home control, etc [2]. A WSN is comprised of low-cost resource-constrained Sensor Nodes (SNs) that result in a multi-hop wireless network via low-power and low-range ad-hoc communication. Each SN is typically equipped with a sensing device that records a physical parameter such as light intensity, motion, temperature and humidity, radiation level, etc. These fine-grained sensor acquired values are quantized and converted to digital data and transmitted over the wireless network to a dedicated node called *sink*. The sink is a sufficiently powerful computing unit that processes the sensor data for the end user.

A. Motivation

The SNs are either battery-powered or self-powered, thus their energy capability is often limited. The hardware capabilities (processing and memory) of off-the-shelf SN platforms are also basic. Therefore, the computation and data transmission is highly constrained by the scarce resources of SNs. Naturally, transferring all the obtained raw sensor readings overloads the

network and also leads to energy-drained nodes to result in network partitions and WSN degradation [3]. In particular, nodes closer to the sink forward the traffic of all other nodes and are critically endangered by a rapid battery depletion.

Several studies have shown that the sensor observations are both spatially and temporally correlated [4]. Such spatio-temporal data correlation encourages us to improve the efficiency of data gathering schemes for WSNs by incorporating compression techniques. In several applications like micro-climate monitoring, environmental protection, distributed acoustic measurement, etc. more efficient *lossy* compression methods are preferred to resource-intensive *lossless* compression techniques [5]. In such cases, data compression in WSNs is closely related to distributed signal compression [6]. From signal processing point of view, the data recorded by the SNs are modeled as a distributed spatio-temporal signal. Signal compression usually reduces the amount of measurements that are transmitted to the sink. Signal *reconstruction* or *recovery* estimates the original signal from the received measurements. The key challenge to design a signal compression technique for WSNs is to make it efficiently implementable in a distributed manner.

B. State of the Art

There are ongoing efforts to achieve energy-efficient data gathering without sacrificing the required accuracy [7]. In-network compression [8] is an efficient technique especially when there is an optimal coordination between routing and compression algorithms. Such an optimal coordination is a hard task that relatively degrades the practical value of those methods. Distributed source coding [9] achieves a good performance when the statistical attributes of the spatio-temporal signal do not change drastically. However, its accuracy is greatly degraded when unexpected patterns occur.

Compressed Sensing or Compressive Sampling (CS) [10] has opened a new avenue in a magnitude of research areas where a sparse or compressible signal is to be acquired from fewest possible measurements. Distributed sampling systems like WSNs are one of the most important fields in which CS can be employed. CS-based spatial sampling introduces four advantages to the traditional methods: First, the data acquisition is performed by random linear measurements, hence no centralized control is needed. This significantly reduces the coordination overhead. Second, the energy consumption is bal-

anced across SNs to avoid the occurrence of exhausted nodes. Third, SNs perform simple data processing operations while the complex signal reconstruction is conducted on the sink. Finally, CS is robust against communication failures since the messages transmitted through the network carry equally valuable information pieces. Therefore, infrequent message drops will not lead to significant losses. These properties make CS a very attractive technique for efficient, balanced and robust data collection in distributed sensor networks [11].

Spatial CS for WSNs was proposed in Compressive Wireless Sensing (CWS) [12] and developed to Compressive Data Gathering (CDG) [13], [14]. Distributed Compressive Sensing (DCS) [15] extends CS-based spatial sampling techniques for WSNs to temporal domain by considering temporal as well as spatial correlations between sensor observations. DCS model is best suited for WSNs with star topology.

C. Paper Contributions

On this background, this paper addresses two key shortcomings of the existing CS-based spatio-temporal sampling techniques, namely: a) Implementing DCS on multi-hop WSNs with tree topology leads to unbalanced energy consumption and exhausted nodes, and b) Existing spatio-temporal CS techniques for WSN usually consider discontinuous sampling periods. Consequent intervals are treated separately when acquiring measurements and recovering the signal. As we show in this paper, temporal correlation of overlapping sampling periods allows for non-interruptive and more efficient measurement and signal recovery. In this paper, we significantly enhance CS-based spatio-temporal sampling in multi-hop WSNs by: 1) A new spatio-temporal CS model that puts a balanced computation and communication load on all SNs, and 2) Applying the concept of *sampling window* to streamline the process of compressive measurement and signal reconstruction without imposing data acquisition delay.

D. Paper Organization

Section II describes our system model and reviews the concepts of CS theory as relevant for the paper. Section III lists the related work and briefly discusses their features. Section IV presents our spatio-temporal model for multi-hop WSNs. Section V evaluates our proposed models.

II. PRELIMINARIES

This section details the WSN models where CS is applied as well as topics of CS utilized in this work.

A. WSN System Model and Requirements

We consider multi-hop WSNs consisting of static homogeneous SNs. We assume that SNs accurately record the physical parameter of interest at the position they are located and the quantization error of their sensors is negligible. A WSN consists of n battery-powered SNs with limited memory and processing capabilities. Data collected from the WSN are transmitted to the sink. The sink is a high performance computation unit. It is responsible for reconstructing the state

of the environment from the measurements collected and transmitted through the network. The WSN should be able to represent the current state of the environment periodically on a regular basis. This representation should meet a certain level of accuracy. Besides the regular monitoring, WSN should be able to detect abnormal events like extremely high temperature at a certain point. Since SNs have very scarce resources, it is crucial to transmit as few data as possible. We are especially interested in methods that reconstruct the state of the environment at the sink from fewest number of measurements collected from each SN.

Throughout this text, we may use the terms *signal* and *vector* interchangeably. For example, a WSN consisting of n SNs each of which recording r samples in every T time units, produces a discrete spatio-temporal signal that can be represented by a vector $\mathbf{f} \in \mathbb{R}^N$ where $N = nr$. In case of sole spatial sampling, N and n are equal.

B. Basics of Compressive Sampling

Vector $\mathbf{x} \in \mathbb{R}^N$ is said to be S-sparse if $\|\mathbf{x}\|_0 = S$, i.e., $\mathbf{x}$ has only S nonzero entries and its all other $N - S$ entries are zero. S-sparse vector $\mathbf{x}_S \in \mathbb{R}^N$ is made from non-sparse vector $\mathbf{x} \in \mathbb{R}^N$ by keeping S largest entries of $\mathbf{x}$ and zeroing its all other $N - S$ entries. Signal $\mathbf{f}$ is compressible under orthonormal basis Ψ when $\mathbf{f} = \Psi\mathbf{x}$ and $\|\mathbf{x} - \mathbf{x}_S\|_2$ is negligible for some $S \ll N$. The matrix Ψ is an orthonormal matrix with the basis vectors of the compressive system Ψ as its columns. Hereafter, the terms *system* and *domain* may also refer to either a measurement or a compressive basis. Most signals recorded from natural phenomena are compressible under Fourier transform, Discrete Cosine Transform (DCT) and the family of wavelet transforms [16]. This is the fundamental fact behind every traditional compression technique. CS is distinguished from traditional compression techniques in signal acquisition method. It combines compression into the sampling layer and tries to recover the original signal from fewest possible measurements.

Definition 1. *[17] Measurement matrix Φ is an $m \times N$ real matrix consisting of $m < N$ vectors randomly selected from measurement basis Φ. It is used to produce a measurement vector $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{y} = \Phi\mathbf{f}$.*

Definition 2. *[18] Coherence between the measurement basis Φ and the compressive basis Ψ is denoted by $\mu(\Phi, \Psi)$ and is equal to $\max_{1 \le i,j \le N} |\phi_i \cdot \psi_j|$ where for each $1 \le i,j \le N$, ϕ_i and ψ_j are basis vectors of Φ- and Ψ-domain respectively.*

Theorem 1. *[17] Suppose signal $\mathbf{f} \in \mathbb{R}^N$ is S-sparse in Ψ-domain, i.e., $\mathbf{f} = \Psi\mathbf{x}$ and $\mathbf{x}$ is S-sparse. We acquire m linear random measurements by projecting $\mathbf{f}$ on m randomly selected basis vectors of the measurement system Φ. Assume $\mathbf{y} \in \mathbb{R}^m$ represents these measurements such that $\mathbf{y} = \Phi\mathbf{f}$ where Φ is the measurement matrix. Then it is possible to recover $\mathbf{f}$ from $\mathbf{y}$ by solving the convex optimization problem:*

$$\hat{\mathbf{x}} = \operatorname*{argmin}_{\tilde{\mathbf{x}} \in \mathbb{R}^N} \|\tilde{\mathbf{x}}\|_1 \quad subject\ to \quad \mathbf{y} = \Phi\Psi\tilde{\mathbf{x}} \qquad (1)$$

Recovered signal will be $\hat{\mathbf{f}} = \Psi\hat{\mathbf{x}}$.

Signal recovery is possible when the number of measurements follows

$$m > C \cdot S \cdot \log N \cdot \mu^2(\Phi, \Psi) \qquad (2)$$

where $C > 1$ is a small real constant [18].

CS is specially targeted at *incoherent* Φ and Ψ bases such that signal $\mathbf{f}$ can be compressively projected on Ψ [18]. From Equation (2), it is clear why compressibility and incoherence are crucial for the practicality/utility of CS. In order to efficiently incorporate CS theory in a specific sampling scenario, we need the measurement and compressive bases to be as incoherent as possible to decrease parameter μ in Equation (2). Moreover, compressive basis must be able to effectively compress the signal $\mathbf{f}$ to decrease S in Equation (2). When these two preconditions hold for a certain sampling configuration, it is possible to recover the signal $\mathbf{f}$ from m measurements where m can be much smaller than the dimension of the original signal [19].

Interestingly, certain random matrices such as a Gaussian matrix with independent and identically distributed entries from a normal distribution $\mathcal{N}(0,1)$ have low coherence with any fixed orthonormal basis [17]. The elements of such a random matrix can be calculated on the fly using a pseudo-random number generator which is common between SNs and the sink. When the Gaussian random number generator at every SN is initialized by the id-number of that SN, the sink can exactly reproduce the measurement matrix. Note that in this case, the measurement matrix does not need to be stored on the SNs. Therefore, using random measurement matrices allows for more flexibility and requires less memory on the SNs. The reproduced measurement matrix at the sink is used to recover the signal according to Theorem 1.

In practice, the signal $\mathbf{f}$ cannot be always perfectly transformed into a strictly sparse projection. Instead, it is always transformed into a compressible form with many near zero entries and few relatively large values. Candès [20] showed that if $\|\mathbf{x} - \mathbf{x}_S\|_2 < \varepsilon$ for some integer $S \ll N$ and a small real constant ε, then the recovery error by solving the optimization problem in Equation (1) is bounded to $O(\varepsilon)$.

III. RELATED WORK

There are a variety of compression techniques for WSNs ranging from distributed source modelling [21] to distributed transform coding [22] and distributed source coding [23]. A comprehensive comparison of data compression techniques for WSNs can be found in [24]. Duarte et al. has also conducted a survey on existing signal compression techniques from a signal processing perspective in [6]. This survey provides an in-depth view of several spatial and spatio-temporal signal compression techniques for WSNs. For each class of WSN applications, a specific compression scheme may be chosen. We are not going to compare various techniques, as such detailed comparative surveys can be found in other works such as the papers by Duarte et al. [6] and Sirooksai et al. [24]. These consider CS as a key technique for distributed signal acquisition.
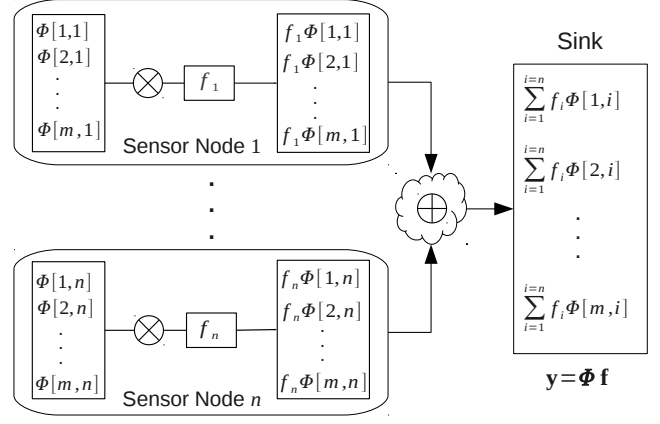

Fig. 1. Distributed spatial sampling using CWS

We categorize related research on applications of CS in WSNs into three subsections. First, spatial sampling without considering temporal correlations is discussed. Second, ongoing attempts to implement CS on common hardware platforms of WSNs is studied. Finally, spatio-temporal CS-based distributed signal acquisition for WSNs is investigated.

A. Spatial CS for WSNs

CWS [12] is a special implementation of CS for distributed spatial sampling in WSNs. CWS models a WSN as a distributed sampling system in which every single SN records the local value of the physical phenomenon of interest (such as radiation level, temperature, humidity, light intensity, etc.). Putting samples from each SN together, a discrete spatial signal is composed that can be represented by a vector $\mathbf{f} \in \mathbb{R}^n$ where n is the number of SNs. CWS has introduced a distributed method for computing the matrix multiplication of Definition 1 in WSNs [12]. This is possible if the SN that records the ith entry of the signal vector $\mathbf{f}$, can access the ith column of the matrix Φ in Definition 1 for all $1 \leq i \leq n$.

Figure 1 shows a simple configuration in which n SNs are directly connected to the sink. The ith SN can access the ith column of the matrix Φ. This column might be embedded into the SN before deployment or calculated on the fly when using random measurement matrices. Note that since addition is a commutative operation, the summation can be calculated in any order. Therefore, CWS can be applied to any topology like star topology, or multi-hop WSNs pertaining tree or chain topologies. The cloud shape enclosing the summation operator signifies that any tree-based topology can compute the measurement vector $\mathbf{y}$ and deliver it to the sink.

B. Practical Implementation of CS in WSNs

CDG is one of the first detailed implementations of CWS for large-scale WSNs introduced by Luo et al. [13], [14]. The authors presented a comprehensive comparative discussion showing that CS leads to a more efficient and stable signal acquisition technique compared to some traditional methods

such as in-network compression [8] and distributed source coding [25]. Using over-complete dictionaries [26] in solving the convex optimization in Equation (1), CDG is also able to tackle abnormal sensor readings and detect events. There is a continuing research to incorporate CS theory in environment monitoring as well as event detection [27]. In addition to emerging theoretical concepts of CS, there is also ongoing research to implement CS practically on WSNs of different configurations and scales such as ZigBee networks [28]. Our work constitutes a key advancement to realize the abstract idea of distributed CS in WSN. As we approach a more realistic implementation of CWS, more and more challenges pop up [29], [30]. We aim to recognize those problems and systematically solve them.

C. Spatio-Temporal CS for WSNs

DCS [15], [31] is an extension of the pure spatial CS-based data acquisition techniques for WSNs to the temporal domain. DCS equips each SN with its own independent random measurement matrix. The signal reconstruction takes place using a joint recovery algorithm. Duarte et al. in [15] have shown that when the number of SNs is very large, their method requires much less measurements than when applying CS individually to every single SN. DCS exploits not only spatial but also temporal compressibility. It is excellent for applications where SNs can communicate directly to the sink and the temporal sampling rate is relatively high, for example body-area sensor networks [32]. Unfortunately, DCS does not preserve the key property of CS, i.e., balancing the overhead across all SNs as explained in more detail in the next section, where we investigate DCS shortcomings for multi-hop topologies and propose our new model preserving the balancing property.

Shen et al. [33] studied the concept of CS-based spatio-temporal sampling in WSN from a different point of view. They assume a heterogeneous WSN where every SN has its own power profile. In their non-uniform CS measurement technique, the available energy plays a more essential role than the signal reconstruction accuracy required by the user. In this work, we assume homogeneous SNs in a balanced power consumption mode and try to maintain the energy consumption balanced during the operation of the WSN. Our goal is to achieve better accuracy with lower measurement ratios.

This paper is closely related to CWS and DCS. We try to keep the measurement method computationally simple with balanced energy consumption across the network. Similar to CWS, we aim to maintain balanced energy consumption by all SNs. However we propose a model for a general spatio-temporal sampling scheme rather than plain spatial sampling. Like DCS we consider a spatio-temporal model for the recorded signal and try to exploit temporal as well as spatial correlations between sensor observations. Next section describes our spatio-temporal CS model as well as the concept of sampling window used in this paper. We will discuss the efficiency and accuracy of our proposed model. Then, we examine how event detection can be applied to a streamlined measurement model.

IV. SPATIO-TEMPORAL COMPRESSIVE SAMPLING

If we take another look at Equation (2), we realize that CS is also very efficient in number of required measurements. The number of required compressive measurements, namely the parameter m in Equation (2), grows logarithmically with the dimension of the signal $\mathbf{f}$. This means that the real power of CS is intrinsically visible for high dimensional signals when N is large enough. In plain spatial CS methods like CWS, the dimension of the spatial signal is equal to the number of SNs. Therefore, simple spatial CS might be less useful for small- to medium-scale WSNs. If we extend our model to the temporal domain, we can exploit the desirable logarithmic cost growth even in small- and medium-scale WSNs by increasing the temporal sampling rate of individual SNs. This is the chief motive for our spatio-temporal method.

A. Unbalanced Spatio-temporal CS for Multi-hop WSN

The DCS technique models jointly sparse signals in a distributed system and introduces a new algorithm suited for recovering jointly-sparse signals. In DCS, the measurement matrix is a block-diagonal matrix composed of several *temporal* measurement sub-matrices. Assume that the ith SN is recording r_i samples every T time units and build up a vector $\mathbf{f}'_i \in \mathbb{R}^{r_i}$. Temporal values of each SN produce such a discrete temporal signal and all of them together form a discrete spatio-temporal signal $\mathbf{f}' = [\mathbf{f}'_1{}^{tr} \ \mathbf{f}'_2{}^{tr} \ \cdots \ \mathbf{f}'_n{}^{tr}]^{tr}$ of size $N = r_1 + r_2 + ... + r_n$ where $[\cdot]^{tr}$ is the transpose operator. The measurement vector $\mathbf{y}' \in \mathbb{R}^m$ is also composed of n subvectors such that:

$$\mathbf{y}' := \begin{pmatrix} \mathbf{y}'_1 \\ \mathbf{y}'_2 \\ \vdots \\ \mathbf{y}'_n \end{pmatrix} = \begin{pmatrix} \mathbf{\Phi}'_1 & & & \\ & \mathbf{\Phi}'_2 & & \\ & & \ddots & \\ & & & \mathbf{\Phi}'_n \end{pmatrix} \begin{pmatrix} \mathbf{f}'_1 \\ \mathbf{f}'_2 \\ \vdots \\ \mathbf{f}'_n \end{pmatrix} \quad (3)$$

where each $\mathbf{\Phi}'_i \ i \in \{1, 2, \cdots, n\}$ has r_i columns. Having

$$\mathbf{\Phi}' := \begin{pmatrix} \mathbf{\Phi}'_1 & & & \\ & \mathbf{\Phi}'_2 & & \\ & & \ddots & \\ & & & \mathbf{\Phi}'_n \end{pmatrix} \quad \text{and} \quad \mathbf{f}' = \begin{pmatrix} \mathbf{f}'_1 \\ \mathbf{f}'_2 \\ \vdots \\ \mathbf{f}'_n \end{pmatrix}, \quad (4)$$

Equation (3) can be written in the form of $\mathbf{y}' = \mathbf{\Phi}'\mathbf{f}'$. Remembering Definition 1, we observe that $\mathbf{\Phi}'$ is a block-diagonal measurement matrix. The spatio-temporal signal $\mathbf{f}'$ can be recovered from measurement vector $\mathbf{y}'$ using Theorem 1. Duarte et al. [31] have discussed joint sparsity models and provided an algorithm to efficiently recover $\mathbf{f}'$ from $\mathbf{y}'$.

Employing DCS in a multi-hop WSN, leads to unbalanced communication overhead that eventually causes network partitioning and or coverage drops due to the depletion of the batteries of more active nodes. Figure 2 shows what happens when transmitting vector $\mathbf{y}'$ in Equation (3) over a WSN with chain topology. Every SN calculates its own component of the measurement vector $\mathbf{y}'$. In a multi-hop topology each component is treated separately as a data packet. The message
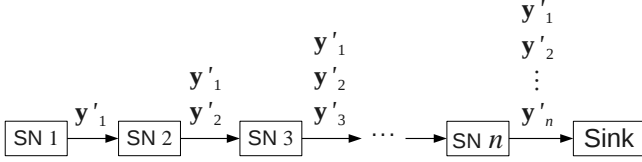
Fig. 2.   Measurement mechanism of DCS in multi-hop WSNs

Fig. 3.   Spatio-temporal compressibility over longer periods using block-diagonal measurement matrix

length increases as the data packets approach the sink. New sub-vector components must be attached to the messages received from previous hops. Apparently, DCS does not lead to a balanced data acquisition mechanism for multi-hop WSNs.

B. Balanced Spatio-temporal CS for Multi-hop WSN

Yap et al. [34] have shown that block-diagonal random measurement matrices can perform as good as dense random measurement matrices in CS signal acquisition and recovery. Similar to DCS, we also use a block-diagonal measurement matrix. However, we propose a different model for spatio-temporal signals and a new structure of the measurement matrix. Let $\mathbf{\Phi}_t$ and $\mathbf{f}_t \in \mathbb{R}^n$ denote the measurement matrix and the spatial signal at time t, respectively. The measurement vector at time t will be $\mathbf{y}_t = \mathbf{\Phi}_t \mathbf{f}_t$. We perform the measurement process T times and then recover the spatio-temporal signal $[\mathbf{f}_1^{tr} \mathbf{f}_2^{tr} \cdots \mathbf{f}_T^{tr}]^{tr}$ from υ such that:

$$\upsilon := \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_T \end{pmatrix} = \begin{pmatrix} \mathbf{\Phi}_1 & & & \\ & \mathbf{\Phi}_2 & & \\ & & \ddots & \\ & & & \mathbf{\Phi}_T \end{pmatrix} \begin{pmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \\ \vdots \\ \mathbf{f}_T \end{pmatrix} \quad (5)$$

where for every $1 \leq t \leq T$, $\mathbf{\Phi}_t$ has m_t rows and n columns.

There is a fundamental difference between measurement matrices in Equation (4) and Equation (5). In Equation (4), each block $\mathbf{\Phi}'_i$ $i \in \{1, 2, \cdots, n\}$ corresponds to a single SN sampling over a period T. In Equation (5), each block $\mathbf{\Phi}_t$ $t \in \{1, 2, \cdots, T\}$ corresponds to spatial samples acquired by all n SNs at time instance t. Note that in contrast to DCS, each SN in our model transmits $m = m_1 + m_2 + \cdots + m_T$ measurements to deliver the measurement vector υ to the sink. Therefore, our model still benefits from the balanced energy consumption like CWS. Remember from Figure 2 that in the extreme case of chain topology, the number of transmissions in DCS increases in order of $O(n^2)$ as the measurements traverse the network hop by hop to approach the sink.

Evaluations show that our model for the spatio-temporal signal leads to a more compressible representation of the signal especially when the spatio-temporal signal is acquired over longer periods. As mentioned earlier, the logarithmic growth of sampling cost in CS encourages us to try acquiring more samples over longer periods. First, we formally define the efficiency of a CS-based signal acquisition method in WSNs. This efficiency is directly related to the compressibility of the signal. We will assess our model of spatio-temporal signals by investigating the level of compressibility that we achieve by extending CS to the temporal domain over longer sampling periods. We also need to exactly define *sampling period* and *sampling round*.
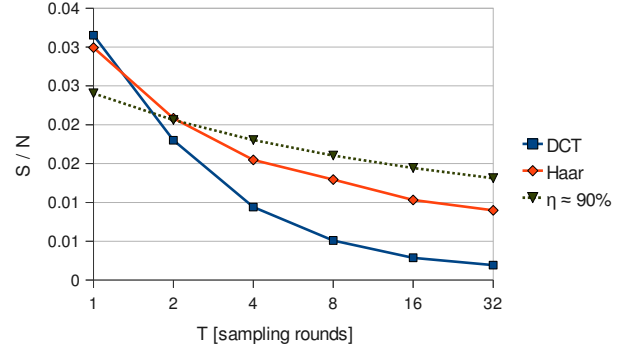
Definition 3. *Sampling period of length T is composed of T sampling rounds. Each sampling round occurs at a discrete time instance on a regular basis. During a sampling round, all SNs record the sensed value at that time instance. Compressive measurements are then calculated distributively from the recorded values and transmitted over the multi-hop WSN to the sink.*

Definition 4. *The efficiency of the signal acquisition in a CS-based system is denoted by η defined as $\eta = (N-m)/N$ where m is the number of measurements according to Equation (2) and N is the dimension of spatio-temporal signal.*

When more measurements need to be acquired for a fixed N, the efficiency will be lower. If the signal is recoverable from fewer measurements, the efficiency increases. According to Equation (2), for an efficiency of η, we require S/N to be less than $(1 - \eta)/(C[\mu^2(\Phi, \Psi)] \log N)$. We have investigated how this efficiency can be achieved by running numerical experiments on real-world data sets. Evaluations are done using real-world data of air temperature values collected by the LUCE (Lausanne Urban Canopy Experiment) WSN deployment at EPFL [35]. The dashed curve in Figure 3 shows S/N when efficiency is roughly equal to 90% and recovery error is bounded to $\pm 1^\circ C$ per SN. We have calculated S/N under Haar wavelet and DCT transformation for spatio-temporal signal with different sampling periods. We see that when the sampling period T increases, the S/N drops quickly under the limit that satisfies $\eta = 90\%$. In Section V, we will see how higher compressibility (i.e., lower S/N) leads to higher efficiency.

C. The Concept of Sampling Window

So far, we have seen improvements in compressibility of spatio-temporal signals when using block-diagonal random measurement matrices over longer sampling periods. The trade-off is a delay equal to T sampling rounds. Increasing T usually leads to better compressibility of the spatio-temporal signal and improves efficiency. However, the signal

reconstruction is delayed by T sampling rounds. Here, we introduce the concept of *sampling window* and show that the delay affects measurements acquisition and signal recovery only at initialization. Assume for any time instance $\tau \geq T$, the measurement vectors $\mathbf{y}_{\tau-T+1}$, $\mathbf{y}_{\tau-T+2}$, ... , $\mathbf{y}_\tau$ are delivered to the sink. Therefore, all of the vectors $\mathbf{f}_{\tau-T+1}$, $\mathbf{f}_{\tau-T+2}$, ... , $\mathbf{f}_\tau$ which describe the sate of the environment during the interval $[\tau - T + 1, \tau]$ are recoverable from υ_τ where:

$$
\upsilon_\tau := \begin{pmatrix} \mathbf{y}_{\tau-T+1} \\ \mathbf{y}_{\tau-T+2} \\ \vdots \\ \mathbf{y}_\tau \end{pmatrix}
$$
$$
= \begin{pmatrix} \boldsymbol{\Phi}_{\tau-T+1} & & & \\ & \boldsymbol{\Phi}_{\tau-T+2} & & \\ & & \ddots & \\ & & & \boldsymbol{\Phi}_\tau \end{pmatrix} \begin{pmatrix} \mathbf{f}_{\tau-T+1} \\ \mathbf{f}_{\tau-T+2} \\ \vdots \\ \mathbf{f}_\tau \end{pmatrix}. \quad (6)
$$

In the next sampling round, $\mathbf{f}_{\tau+1}$ is sensed by the SNs and should be reconstructed at the sink. Of course we do not want to perform typical CWS from scratch at sampling round $\tau+1$. Having previous measurements using block-diagonal measurement matrix in Equation (6), we can recover the signal at the next sampling round from fewer measurements compared with typical spatial CWS. Our spatio-temporal measurement method should calculate the new measurement vector:

$$
\upsilon_{\tau+1} := \begin{pmatrix} \mathbf{y}_{\tau-T+2} \\ \mathbf{y}_{\tau-T+3} \\ \vdots \\ \mathbf{y}_{\tau+1} \end{pmatrix}
$$
$$
= \begin{pmatrix} \boldsymbol{\Phi}_{\tau-T+2} & & & \\ & \boldsymbol{\Phi}_{\tau-T+3} & & \\ & & \ddots & \\ & & & \boldsymbol{\Phi}_{\tau+1} \end{pmatrix} \begin{pmatrix} \mathbf{f}_{\tau-T+2} \\ \mathbf{f}_{\tau-T+3} \\ \vdots \\ \mathbf{f}_{\tau+1} \end{pmatrix}. \quad (7)
$$

Comparing Equations (6) and (7), we observe that if υ_τ is already present at the sink, then we only need $\mathbf{y}_{\tau+1}$ to be delivered to the sink in order to make $\upsilon_{\tau+1}$. In fact, a *sampling window* of T allows us to efficiently recover the spatio-temporal signal at any sampling round. Note that in this case, to have $\mathbf{y}_{\tau+1}$ (and consequently $\upsilon_{\tau+1}$) at the sink, only $m_{\tau+1}$ more measurements are needed which would be lower than when running CWS on sampling round $\tau + 1$. Moreover, all SNs transmit an equal amount of data proportional to $m_{\tau+1}$.

D. Benefits of Sampling Window

Here, we clarify the advantage of our sampling window mechanism compared with the state of the art CS-based spatio-temporal signal acquisition techniques discussed in this paper. Sampling window as described above allows for recovering the spatio-temporal signal from a history of the current and previously acquired measurements. We recall that CWS only considers instantaneous spatial measurements and DCS operates over disjunct intervals. The advantages of our

proposed sampling window are twofold: First, total number of measurements, namely $\sum_{i=1}^{i=T} m_{\tau-T+i}$ is much less than when running CWS separately for T sampling rounds because considering temporal as well as spatial correlations leads to better compressibility and hence more efficient signal acquisition. Second, acquiring measurements and recovering the signal are done seamlessly with a much less delay. Remember that DCS recovers spatio-temporal signals after a delay proportional to T when all measurements of the last T sampling rounds are received at the sink. In particular, for an extensive multi-hop WSN, our model decreases the delay by a factor of T. Note that here, *delay* refers to the time required to acquire measurements from the network and not the time required by the recovery algorithm to reconstruct the spatio-temporal signal. Many efficient CS reconstruction algorithms such as orthogonal matching pursuit [36] are developed to recover the signal in a timely manner.

E. Detecting Events

Luo et al. [13], [14] have proposed an effective method to detect and handle sparse abnormal sensor readings using overcomplete dictionaries [26]. We apply a similar method for detecting abnormal events in order to trigger notification about potentially harmful situations. What we are especially interested in, is how our sliding sampling window can detect the new events occurring in the most recent sampling round. Assume few SNs record unexpected values at time u where $u \geq T$. We can decompose $\mathbf{f}_u$ into two vectors $\mathbf{f}_c$ and $\mathbf{f}_e$ such that $\mathbf{f}_u = \mathbf{f}_c + \mathbf{f}_e$ where $\mathbf{f}_e$ is the sparse abnormal innovation vector. We know that $[\mathbf{f}_{u-T+1}^{tr} \quad \mathbf{f}_{u-T+2}^{tr} \quad \cdots \quad \mathbf{f}_{u-1}^{tr} \quad \mathbf{f}_c^{tr}]^{tr} = \boldsymbol{\Psi}\xi$ is compressible in the $\boldsymbol{\Psi}$-domain. Substituting $\mathbf{f}_u$ with $\mathbf{f}_c + \mathbf{f}_e$ we will have:

$$
\mathbf{g} := \begin{pmatrix} \mathbf{f}_{u-T+1} \\ \mathbf{f}_{u-T+2} \\ \vdots \\ \mathbf{f}_u \end{pmatrix} = \begin{pmatrix} \mathbf{f}_{u-T+1} \\ \mathbf{f}_{u-T+2} \\ \vdots \\ \mathbf{f}_c + \mathbf{f}_e \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Psi} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \xi \\ \epsilon \end{pmatrix} \quad (8)
$$

such that $\epsilon = [\mathbf{0}^{tr} \quad \mathbf{f}_e^{tr}]^{tr}$, where $\mathbf{0}$ means a zero vector of size $nT - n$. For recovering the original signal as well as detecting abnormal events, we solve a convex optimization problem similar to that in Theorem 1. From Equation (8), we see that $[\xi^{tr} \epsilon^{tr}]^{tr}$ is compressible under overcomplete system $\boldsymbol{\Psi}' = [\boldsymbol{\Psi}\mathbf{I}]$. Assume:

$$
\boldsymbol{\Lambda} = \begin{pmatrix} \boldsymbol{\Phi}_{u-T+1} & & & \\ & \boldsymbol{\Phi}_{u-T+2} & & \\ & & \ddots & \\ & & & \boldsymbol{\Phi}_u \end{pmatrix} \quad (9)
$$

and $\mathbf{z} = \boldsymbol{\Lambda}\mathbf{g}$ is the spatio-temporal measurement vector. If we solve the l_1 minimization problem:

$$
\hat{\mathbf{x}}' = \operatorname*{argmin}_{\tilde{\mathbf{x}}' \in \mathbb{R}^N} \|\tilde{\mathbf{x}}'\|_1 \quad \text{subject to} \quad \mathbf{z} = \boldsymbol{\Lambda}\boldsymbol{\Psi}'\tilde{\mathbf{x}}', \quad (10)
$$

then it is possible to recover the original spatio-temporal signal $\mathbf{g}$ (including the abnormal samples). The recovered signal of $\mathbf{g}$ will be $\hat{\mathbf{g}} = \boldsymbol{\Psi}'\hat{\mathbf{x}}'$. Here, $\hat{\mathbf{x}}' \in \mathbb{R}^{2N}$ and $\boldsymbol{\Psi}'$ has $2N$ columns.

Note that when using a progressive sampling window, we are mainly interested in detecting events at the most recent sampling round. In this case, we see that the abnormal sensor readings are not uniformly distributed over the spatio-temporal signal. The abnormal observations only affect the last chunk of the spatio-temporal signal vector. Detection capability of CS using overcomplete dictionaries has been discussed in CDG [13] for plain spatial signals with a uniformly random distribution of event occurrence. However, this specific distribution of abnormal events that only occur at the very end of the spatio-temporal vector needs to be elaborated in more detail as a future work.

V. EVALUATION

We have evaluated our proposed methods using real-world data collected by the LUCE WSN deployment at EPFL [35]. The ambient temperature values of 64 SNs are used as the physical parameter for evaluating our model. In the LUCE dataset, some records were missing or too desynchronized, i.e. the sampling rounds of the SNs were not aligned. Therefore, we have preprocessed the dataset while preserving the attributes of the spatio-temporal signal to have a synchronized data set which is suitable for testing our model. We assume that measurements are calculated while transferring to the sink using a reliable hop-by-hop transport protocol [37].

Figure 4 shows the accuracy of signal recovery using a block-diagonal measurement matrix as described in Equation (5). Figure 4(a) and Figure 4(b) illustrate the results using DCT and Haar wavelet as the the compressive basis respectively. This means that the Ψ-domain in Figure 4(a) and Figure 4(b) is DCT and Haar respectively. For each compressive domain, we have tested the measurement and recovery for different sampling window lengths. Parameter T in Equation (5) represents the initialization delay as well as the width of the sampling window of our spatio-temporal sampling model for multi-hop WSNs. The X-axis represents the ratio of the number of measurements to the number of all spatio-temporal samples, namely m/N as in Equation (2). The accuracy of the signal reconstruction is measured by the Signal to Noise Ratio (SNR). The Y-axis represents the SNR in decibels (dB). We observe that in all cases, the quality of signal reconstruction generally improves as the ratio m/N increases.

The evaluations are first done for $T = 1$ which is basically equivalent to the plain spatial sampling case. Then, the width of the sampling window is increased and the evaluation is repeated for $T = 2$, $T = 4$ and $T = 8$. For larger T, we see that higher signal reconstruction accuracy is possible for lower m/N. This means that a higher reconstruction accuracy can be achieved more efficiently if the width of the sampling window T increases.

Rare abnormal readings discussed in Section IV.D are simulated by deliberately modifying the values recorded by 3 SNs from the LUCE dataset. We have selected these SNs at random and increased their recorded value at the last sampling round to above 100 degrees Celsius. This may resemble a fire starting in the environment that is sensed by three SNs. The
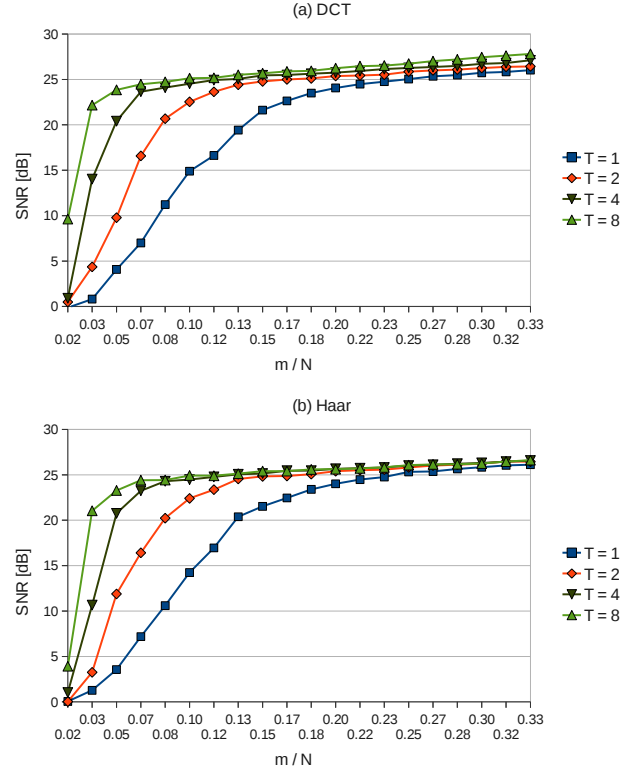

Fig. 4. Multi-hop spatio-temporal CS using block-diagonal measurement matrix

sampling window during this simulation was set to $T = 8$. Figure 5 illustrates how the compressive presentation of the spatio-temporal signal is distorted in presence of abnormal sensor readings. Figure 5(a) and Figure 5(b) show the projection of the signal contaminated with abnormal readings on DCT and Haar wavelet respectively. The DCT projection is not compressible any more. Haar wavelet projection leads to a more compressible representation, since it can preserve hard edges in the signal better than DCT. Now, we use an overcomplete system $\Psi' = [\Psi \mathbf{I}]$ as discussed in Section IV.D to recover the compressible projection. We use two systems $\Psi'_{\mathrm{DCT}} = [\Psi_{\mathrm{DCT}} \mathbf{I}]$ and $\Psi'_{\mathrm{Haar}} = [\Psi_{\mathrm{Haar}} \mathbf{I}]$ where Ψ_{DCT} and Ψ_{Haar} represent DCT and Haar wavelet bases respectively. Figure 5(c) and Figure 5(d) show the recovered compressive projections on Ψ'_{DCT} and Ψ'_{Haar} respectively. Naturally, we observe that the projections of the spatio-temporal signal on the overcomplete bases are much more compressible that typical DCT and Haar bases.

Higher compressibility means that less measurements are required to recover the signal with comparable accuracy, or in other words, it is possible to achieve better signal reconstruction quality from the same number of random measurements. Figure 6 verifies this statement. With low level of compressibility that typical DCT and Haar projections provide, it is impossible to recover the signal and detect abnormal sensor readings. Using overcomplete dictionaries, the signal could be recovered and the abnormal events are localized. Figure 6(a)
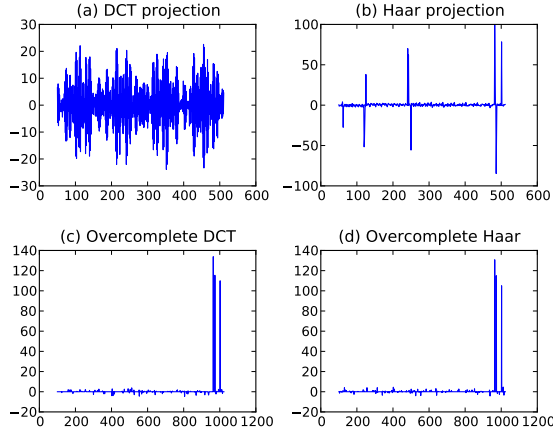
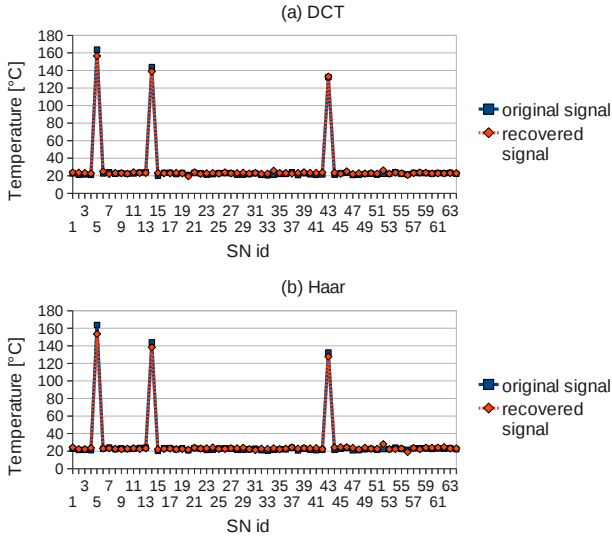Fig. 5. Compressive projection of the spatio-temporal signal contaminated with abnormal readings

Fig. 6. Event detection in multi-hop spatio-temporal CS using block-diagonal measurement matrix and overcomplete compressive systems
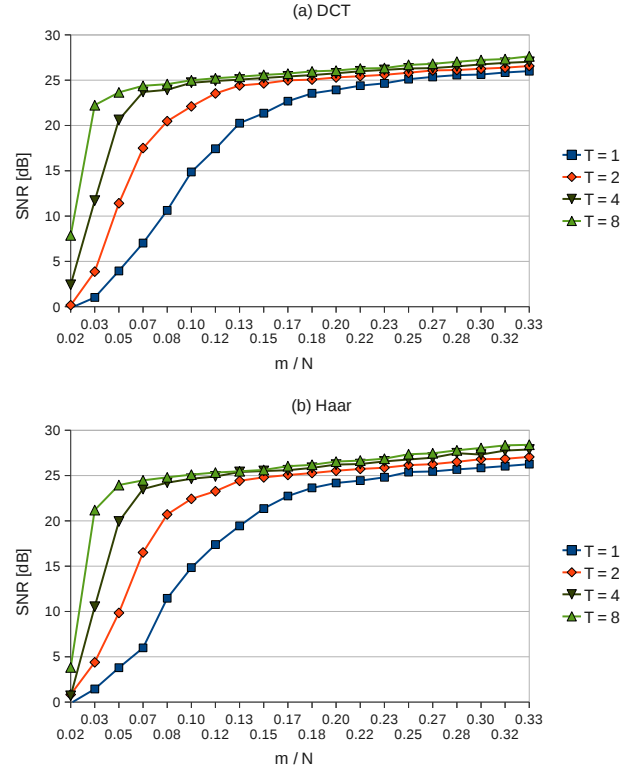

Fig. 7. Reconstruction from DCS measurements

by dense Gaussian random measurement matrices. The SNR is closely comparable to schemes with block-diagonal measurement matrices like DCS and our spatio-temporal block-diagonal model for CS measurements in multi-hop WSNs.

VI. CONCLUSION

In this paper, we have presented a generalized framework for implementing spatio-temporal CS for multi-hop WSNs. We have proposed a new model for spatio-temporal signals in WSNs and introduced a CS measurement mechanism suited for multi-hop WSNs. By incorporating the concept of sampling window, we could streamline the process of spatio-temporal CS measurement and signal recovery. Using overcomplete dictionaries, abnormal sensor readings can be detected. When the sampling window is filled and the measurement process is fully streamlined, then only in-network communication delays affect the timeliness of our composite spatio-temporal data gathering and event detection technique. Evaluations show that the performance of our method is comparable to that of DCS measurement and even dense spatio-temporal Gaussian measurement matrices. The advantage of our model over the state of the art methods is balanced energy consumption and streamlined sampling. This work helps to design practical implementation of spatio-temporal CS for multi-hop WSNs.

REFERENCES

[1] J. Elson and D. Estrin, "Sensor networks: A bridge to the physical world," in *Wireless Sensor Networks*, C. S. Raghavendra, K. M. Sivalingam, and T. Znati, Eds. Springer US, 2004, pp. 3–20.
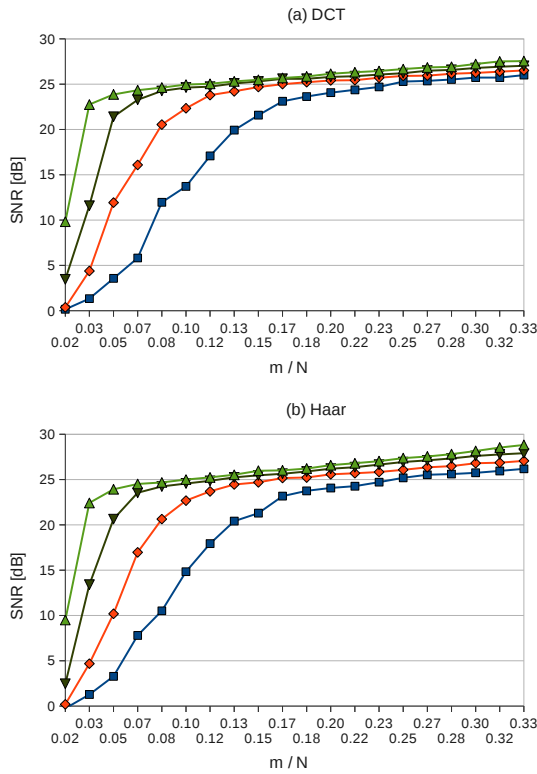
and Figure 6(b) show the recovered signal at the last (8th) sampling round using Ψ'_{DCT} and Ψ'_{Haar} overcomplete systems respectively. Note that when the sampling window is filled and the measurement process is streamlined, such abnormal events can be detected very shortly as only the last measurement vector is needed to be reported to the sink.

We have also compared our sampling method with that of DCS. Figure 7 shows the same results as discussed in explanation of Figure 4 for the DCS measurements. The signal reconstruction quality using our proposed block-diagonal spatio-temporal measurement matrix is comparable to signal reconstruction using spatio-temporal model of DCS. More interestingly, we have observed that block-diagonal matrices lead to almost the same performance of dense measurement matrices as expected in [34]. Figure 8 shows the signal reconstruction quality from measurement vectors produced

Fig. 8. Signal recovery using dense Gaussian measurement matrix

[2] K. Sohraby, D. Minoli, and T. Znati, *Wireless Sensor Networks, Technology, Protocols and Applications*. Wiley-Interscience, 2007.

[3] G. Dini, M. Pelagatti, and I. Savino, "Repairing network partitions in wireless sensor networks," in *Proc. of IEEE Internatonal Conference on Mobile Adhoc and Sensor Systems (MASS)*, 2007, pp. 1 –3.

[4] M. C. Vuran, zgr B. Akan, and I. F. Akyildiz, "Spatio-temporal correlation: Theory and applications for wireless sensor networks," *Computer Networks*, vol. 45, no. 3, pp. 245 – 259, 2004.

[5] S. Arrabi and J. Lach, "Adaptive lossless compression in wireless body sensor networks," in *Proc. of the Fourth International Conference on Body Area Networks*, ser. BodyNets '09, 2009, pp. 19:1–19:8.

[6] A. O. Marco F. Duarte, Godwin Shen and R. G. Baraniuk, "Signal compression in wireless sensor networks," *Philosophical Trans. of the Royal Society*, vol. 370, no. 1958, pp. 118–135, 2012.

[7] G. Anastasi *et al.*, "Energy conservation in wireless sensor networks: A survey," *Ad Hoc Networks*, vol. 7, no. 3, pp. 537 – 568, 2009.

[8] A. Ciancio *et al.*, "Energy-efficient data representation and routing for wireless sensor networks based on a distributed wavelet compression algorithm," in *Proc. of the 5th international conference on Information processing in sensor networks (IPSN)*, 2006, pp. 309–316.

[9] Y.-W. Hong, Y.-R. Tsai, Y.-Y. Liao, C.-H. Lin, and K.-J. Yang, "On the throughput, delay, and energy efficiency of distributed source coding in random access sensor networks," *IEEE Trans. on Wireless Communications*, vol. 9, no. 6, pp. 1965 –1975, 2010.

[10] E. Candes *et al.*, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. on Information Theory*, vol. 52, no. 2, pp. 489 – 509, 2006.

[11] W. Bajwa, J. Haupt, A. Sayeed, and R. Nowak, "Joint source-channel communication for distributed estimation in sensor networks," *IEEE Trans. on Information Theory*, vol. 53, no. 10, pp. 3629 –3653, 2007.

[12] W. Bajwa *et al.*, "Compressive wireless sensing," in *Proc. of the 5th international conference on Information processing in sensor networks (IPSN)*, 2006, pp. 134–142.

[13] C. Luo *et al.*, "Compressive data gathering for large-scale wireless sensor networks," in *Proc. of the 15th annual international conference on Mobile computing and networking (Mobicom)*, 2009, pp. 145–156.

[14] C. Luo, F. Wu, J. Sun, and C. Chen, "Efficient measurement generation and pervasive sparsity for compressive data gathering," *IEEE Trans. on Wireless Communications*, no. 99, pp. 1 –11, 2010.

[15] M. F. Duarte *et al.*, "Universal distributed sensing via random projections," in *Proc. of the 5th international conference on Information processing in sensor networks (IPSN)*, 2006, pp. 177–185.

[16] A. Graps, "An introduction to wavelets," *IEEE Compututer Science Engineering.*, vol. 2, pp. 50–61, June 1995.

[17] E. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. on Information Theory*, vol. 52, no. 12, pp. 5406 –5425, 2006.

[18] E. Cands and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Problems*, vol. 23, no. 3, p. 969, 2007.

[19] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.

[20] E. J. Candes, "Compressive sampling," in *Proc. of the International Congress of Mathematicians*, vol. 3, 2006, pp. 1433–1452.

[21] A. Ali *et al.*, "An adaptive and composite spatio-temporal data compression approach for wireless sensor networks," in *Proc. of the 14th ACM international conference on Modeling, analysis and simulation of wireless and mobile systems (MSWiM)*, 2011, pp. 67–76.

[22] Ciancio *et al.*, "A distributed wavelet compression algorithm for wireless sensor networks using lifting," in *Proc. of the IEEE International Conference onAcoustics, Speech, and Signal Processing (ICASSP)*, vol. 4, 2004, pp. 633–636.

[23] Z. Xiong *et al.*, "Distributed source coding for sensor networks," *Signal Processing Magazine, IEEE*, vol. 21, no. 5, pp. 80 – 94, 2004.

[24] T. Srisooksai, K. Keamarungsi, P. Lamsrichan, and K. Araki, "Practical data compression in wireless sensor networks: A survey," *Journal of Network and Computer Applications*, vol. 35, no. 1, pp. 37 – 59, 2012.

[25] G. Hua and C. W. Chen, "Correlated data gathering in wireless sensor networks based on distributed source coding," *International Journal of Sensor Networks*, vol. 4, no. 1/2, pp. 13–22, 2008.

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

[27] J. Meng *et al.*, "Sparse event detection in wireless sensor networks using compressive sensing," in *Proc. of 43rd Annual Conference on Information Sciences and Systems (CISS)*, 2009, pp. 181 –185.

[28] C. Caione, D. Brunelli, and L. Benini, "Compressive sensing optimization over zigbee networks," in *Proc. of the International Symposium on Industrial Embedded Systems (SIES)*, 2010, pp. 36–44.

[29] M. Mahmudimanesh, A. Khelil, and N. Suri, "Reordering for better compressibility: Efficient spatial sampling in wireless sensor networks," *Proc. of the International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing 2010*, pp. 50–57, 2010.

[30] M. Mahmudimanesh *et al.*, "Map-based compressive sensing model for wireless sensor network architecture, a starting point," in *Proc. of the International Workshop on Wireless Sensor Networks Architectures, Simulation and Programming (WASP)*, 2009.

[31] M. Duarte *et al.*, "Distributed compressed sensing of jointly sparse signals," in *Conference Record of the 39th Asilomar Conference on Signals, Systems and Computers*, 2005, pp. 1537 – 1541.

[32] C. Seeger, A. Buchmann, and K. Van Laerhoven, "myhealthassistant: A phone-based body sensor network that captures the wearerś exercises throughout the day, to appear," in *Proc. of The 6th International Conference on Body Area Networks (BodyNets)*, 2011.

[33] Y. Shen *et al.*, "Non-uniform compressive sensing in wireless sensor networks: Feasibility and application," in *Proc. of the Seventh International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, 2011, pp. 271 –276.

[34] H. L. Yap *et al.*, "The restricted isometry property for block diagonal matrices," in *Proc. of the 45th Annual Conference on Information Sciences and Systems (CISS)*, 2011, pp. 1 –6.

[35] (2008) Sensorscope LUCE deployment at École Polytechnique Fédérale de Lausanne. [Online]. Available: http://sensorscope.epfl.ch/index.php/Environmental_Data

[36] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. on Information Theory*, vol. 53, no. 12, pp. 4655 –4666, 2007.

[37] F. Stann and J. Heidemann, "RMST: reliable data transport in sensor networks," in *Proc. of the First IEEE International Workshop on Sensor Network Protocols and Applications*, 2003, pp. 102 – 112.