Compressive Sensing Based Signal Processing in Wireless Sensor Networks: A Survey
Thakshila Wimalajeewa, Member IEEE, and Pramod K. Varshney, Fellow IEEE

Abstract—Compressive sensing (CS) has been shown to be promising in a wide variety of applications including compressive imaging, video processing, communication, and radar to name a few. Out of the many potential applications, in this survey paper, our goal is to discuss recent advances of CS based signal processing in wireless sensor networks (WSNs) including the main ongoing/recent research efforts, challenges and research trends in this area. In a WSN, CS based techniques are well motivated by not only the sparsity prior but also by the requirement of energy efficient processing even with nonsparse signals. First, we address different forms of sparsity including temporal and spatial that are exhibited in WSNs. We discuss sparsity aware signal processing techniques exploiting such forms of sparsity in centralized as well as distributed/decentralized settings from signal reconstruction perspective. We then review recent research efforts that exploit CS to solve different types of inference problems such as detection, classification, and estimation where signal reconstruction is not necessary. We further review the work that takes practical considerations such as quantization, physical layer secrecy constraints, and imperfect channel conditions into account in the CS based signal processing framework. Finally, open research issues and challenges are discussed in order to provide perspectives for future research directions. With this paper, the readers are expected to gain a thorough understanding of the potential of CS in solving distributed inference problems in WSNs with high dimensional data under practical constraints.

Index Terms—Wireless sensor networks, Distributed inference, Compressive sensing (CS), Distributed/decentralized CS, Compressive detection, Compressive classification, Quantized CS

I. INTRODUCTION

Over the last two decades, the wireless sensor network (WSN) technology has gained increasing attention by both the research community and actual users [1–8]. A typical sensor network consists of multiple sensors of the same or different modalities/types deployed over a geographical area for monitoring a phenomenon of interest (PoI). Once deployed, the distributed sensors are required to form a connected network without a backbone infrastructure as in cellular networks. Most of the sensors are power constrained since they are equipped with small sized batteries which are difficult or impossible to be replaced especially in hostile environments. At the same time, the available (limited) communication bandwidth needs to be efficiently used while exchanging information for efficient fusion. Thus, sensor networks are inherently resource constrained and they starve for energy and communication efficient protocols [1].

Applications of WSNs span a wide range including environmental monitoring [6], detection and classification [9–12], target/object tracking [13–17], industrial applications [18], [19], and health care [20] to name a few. The objective of WSN research is to design efficient protocols and algorithms to perform a given task for such applications taking the inherent resource constraints into account. There is abundant literature related to energy-savings in WSNs in the last several years. There are also several survey and tutorial papers that discuss energy efficiency in WSNs, e.g., [21]. However, there is still much ongoing research on investigating how to optimize power and communication bandwidth in resource constrained sensor networks since none of the standalone protocol is universally applicable.

Recent advances in compressive sensing (CS) have led to novel ways of thinking about approaches to design energy efficient WSNs. CS has emerged as a new paradigm for efficient sparse signal acquisition. Sparsity is one of the low dimensional structures exhibited in many signals of interest including audio, video and radar signals. A signal is said to be sparse if the coefficient vector, when represented in a known (orthogonal) basis, contains only a few significant elements. Sparsity has been exploited in signal processing and approximation theory for tasks such as compression, denoising, model selection, and image processing for a long time. The problem of sparse signal recovery has attracted much attention in the recent literature with advancements of the theory of CS. In the CS framework, a sparse signal can be reliably recovered with a small number of random projections under certain conditions [22]. Since its inception about a decade ago, the CS theory has been successfully applied to a wide variety of applications including compressive imaging [22], video processing [28], wireless channel estimation [29], sensor networks [31–39], cognitive radio networks [40], and radar signal processing [41] to name a few. Further, there exist several books, surveys and tutorial papers covering different aspects of CS such as theory and applications of CS [22], [24], [25], CS algorithms [43], [44], CS in cognitive radio networks [40], [45], [46], CS in communication networks [34], and CS in magnetic resonance imaging (MRI) [47].

CS is well motivated for WSN applications due to several reasons. Universal and agnostic nature of the CS measurement scheme is promising in acquiring signals for a variety of inference tasks. On the other hand, sparsity is a common characteristic that can be observed in various forms. In applications such as acoustic event monitoring, and temperature/pressure monitoring, the time samples collected at a given node can be represented in a sparse manner in a given basis [48]. When considering multiple measurement vectors (MMVs) collected at distributed nodes, different sparsity patterns with certain structures can be observed [48]. Joint processing of such MMVs using CS techniques by exploiting temporal sparsity...
along with different joint structures leads to energy efficient signal processing as desired by WSNs. Spatial sparsity of observations collected at distributed nodes is another form of sparsity that can be exploited in performing several inference tasks. For example, since not all the sensors gather informative observations at any given time, random projections can be employed for distributed compression [31]. Thus, efficient data gathering algorithms can be developed exploiting spatial sparsity. Spatial sparsity also can be leveraged by construction such as in source localization and sparse event detection [49]–[51]. Massive data streams produced by large scale sensor networks as in source localization and sparse event detection [49]–[51]. Thus, in order to process high dimensional data in the presence of inherent resource constraints, the spatio-temporal sparsity and correlation structures can be exploited. Further, the agnostic and universal nature of the CS mechanism motivates the use of CS in WSNs even without the sparsity prior due to the requirement of designing simple hardware for signal acquisition [55].

To employ CS based algorithms to solve different inference problems in resource constrained WSNs, further challenges need to be overcome compared to the standard CS framework. Some of these challenges arise due to

i) the requirement of distributed and decentralized processing

ii) the need to solve other inference problems, e.g., detection, classification, estimation and tracking, in addition to signal reconstruction as widely considered in the standard CS framework and

iii) the need to incorporate practical considerations; e.g., quantization and channel fading.

In this review paper, our goal is to provide an extensive review of CS based signal processing in WSNs focusing on the aforementioned factors. There are several survey/tutorial papers on CS based signal processing in WSNs in the literature. For example, in [33], the authors discuss as to how the certain sensor network parameters such as lifetime, delay, cost and power can be improved using CS. In [34], the applicability of CS for compressed data gathering and source localization is discussed under the general topic of CS in communications and networks. Compressed data gathering techniques, distributed compressive techniques and compressed data routing for WSNs are reviewed in [35]. However, a comprehensive review on CS methods used to perform fundamental inference tasks in WSNs overcoming the aforementioned challenges is missing from the literature. We discuss the extensions of the standard CS framework to solve different inference problems in distributed and decentralized settings exploiting different forms of sparsity under practical considerations.

The remainder of the paper is organized as follows. In Section II, an overview and background of CS is given. In Section III, different forms of sparsity exhibited in WSNs are discussed. The signal processing algorithms exploiting temporal and spatial sparsity in centralized as well as in distributed/decentralized settings are discussed in Section IV. CS based inference including detection, classification and tracking is reviewed in Section V. In Section VI, CS based signal processing under practical considerations such as channel fading and physical layer secrecy is discussed. CS based signal processing with quantization is discussed in Section VII. Finally, future research directions are discussed in Section VIII.

### Notation and abbreviations

Throughout the paper, we use the following notation. Scalars (in \(\mathbb{R}\)) are denoted by lower case letters and symbols, e.g., \(x\) and \(\theta\). Vectors (in \(\mathbb{R}^N\)) are written in boldface lower case letters and symbols, e.g., \(\mathbf{x}\) and \(\mathbf{\theta}\). Matrices are written in boldface upper case letters or symbols, e.g., \(\mathbf{A}\) and \(\mathbf{\Psi}\). By \(\mathbf{0}\) and \(\mathbf{1}\), we denote the vectors with appropriate dimension in which all elements are zeros and ones, respectively. The identity matrix with appropriate dimension is denoted by \(I\). The transpose of matrix \(\mathbf{A}\) is denoted by \(\mathbf{A}^T\). The \(i\)-th row vector and the \(j\)-th column vector of matrix \(\mathbf{A}\) are denoted by \(\mathbf{a}^i\) and \(\mathbf{a}_j\), respectively. The \((i,j)\)-th element of matrix \(\mathbf{A}\) is denoted by \(A_{i,j}\) or \(\mathbf{A}[i,j]\). The \(i\)-th element of vector \(\mathbf{x}\) is denoted by \(x[i]\) or \(x_i\). The \(l_p\) norm is denoted by \(||\cdot||_p\), while \(||\cdot||_2\) is used for both the cardinality (of a set) and the absolute value (of a scalar). The Frobenius norm of \(\mathbf{A}\) is denoted by \(||\mathbf{A}||_F\). The notation \(\mathbf{x} \sim \mathcal{N}(\mu, \Sigma)\) means that the random vector \(\mathbf{x}\) is distributed as multivariate Gaussian with mean \(\mu\) and covariance matrix \(\Sigma\). The abbreviations used in the paper are summarized in Table I.

### II. COMPRESSIVE SENSING (CS) BASICS

In this section, we briefly review the basics of CS. A length-\(N\) signal \(\mathbf{x}\) is said to be sparse if it contains only a few nonzero elements; i.e., \(||\mathbf{x}||_0 \ll N\). The support of a sparse signal (also known as the sparsity pattern/sparse support set) is defined as the set \(\mathcal{U} = \{1, \cdots, N\}\) such that

\[\mathcal{U} := \{i \in \{1, \cdots, N\} \mid x[i] \neq 0\}\]

where \(x[i]\) denotes the \(i\)-th element of \(\mathbf{x}\) for \(i = 1, \cdots, N\).

In the CS framework with a single measurement vector (SMV), one aims to solve for sparse \(\mathbf{x}\) from the following underdetermined linear system:

\[\mathbf{y} = \mathbf{A}\mathbf{x}\]

with no noise and

\[\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v}\]

with noise, where \(\mathbf{y}\) is the length-\(M\) compressed measurement vector, \(\mathbf{A}\) is the \(M \times N\) measurement matrix with \(M < N\), and \(\mathbf{v}\) is the \(M \times 1\) additive noise vector. While recovering \(\mathbf{x}\) from its compressed form \(\mathbf{y}\) in (1) (or (2)) is ill-conditioned when \(M < N\), it has been shown that it is possible to reconstruct \(\mathbf{x}\) exactly under certain conditions on the measurement matrix \(\mathbf{A}\) if \(\mathbf{x}\) is sufficiently sparse [22], [23], [25], [26].

Necessary and sufficient conditions satisfied by the matrix \(\mathbf{A}\) so that \(\mathbf{x}\) can be recovered from \(\mathbf{y}\) have been discussed focusing on several properties [22], [23], [25], [26]. One of the commonly used such properties is the restricted isometry property (RIP) property. The matrix \(\mathbf{A}\) is said to satisfy RIP of order \(k\) if there exists a \(\delta_k \in (0, 1)\) such that

\[(1 - \delta_k)||\mathbf{x}||_2^2 \leq ||\mathbf{A}\mathbf{x}||_2^2 \leq (1 + \delta_k)||\mathbf{x}||_2^2\]

(3)
for all $x \in \Sigma_k$ where $\Sigma_k = \{x : ||x||_0 \leq k\}$. It has been shown that when the entries of $A$ are chosen according to a Gaussian, Bernoulli or in general from a sub-Gaussian distribution, $A$ satisfies RIP with high probability when $M = \mathcal{O}(k \log(N/k))$ [22].

In order to recover $x$ from $y$, the natural choice is to solve the following optimization problem (with no noise) [22], [23], [25], [26]:

$$\min_{x} ||x||_0 \text{ such that } y = Ax. \quad (4)$$

Unfortunately, this $l_0$ norm minimization problem is generally computationally intractable. In order to approximately solve (4), several approaches have been proposed. One of the commonly used approaches is to replace $l_0$ norm in (4) with convex $l_1$ norm. Greedy pursuit and iterative algorithms are also promising in approximately solving (4). In the following, we briefly discuss these approaches.

- Convex relaxation: A fundamental approach for signal reconstruction proposed in the CS theory is the so-called basis pursuit (BP) [50] in which the $l_0$ term in (4) is replaced by the $l_1$ norm to get

$$\min_{x} ||x||_1 \text{ such that } y = Ax. \quad (5)$$

Under some favorable conditions, the solution to (4) coincides with that in (5). In the presence of noise, basis pursuit denoising (BPDN) [50] aims at solving

$$\min_{x} ||x||_1 \text{ such that } ||y - Ax||_2 \leq \epsilon_1 \quad (6)$$

and least absolute shrinkage and selection operator (LASSO) [57] - [59] solves

$$\min_{x} ||y - Ax||_2 \text{ such that } ||x||_1 \leq \epsilon_2 \quad (7)$$

or equivalently

$$\min_{x} \lambda ||x||_1 + \frac{1}{2} ||y - Ax||_2 \quad (8)$$

where $\lambda$ is the penalty parameter and $\epsilon_1, \epsilon_2 > 0$. 

### TABLE I: Abbreviations used throughout the paper

| Abbreviation | Description                                      | Abbreviation | Description                                      |
|--------------|--------------------------------------------------|--------------|--------------------------------------------------|
| ADMM         | Alternating direction method of multipliers      | JSM          | Joint sparsity model                             |
| AFC          | Analog fountain codes                            | LASSO        | Least absolute shrinkage and selection operator  |
| AMP          | Approximate message passing                      | LMS          | Least-mean squares                               |
| AWGN         | Additive white Gaussian noise                    | MAC          | Multiple access channel                         |
| BCD          | Block-coordinate descent                         | MAP          | Maximum a posteriori                            |
| BCS          | Bayesian CS                                      | MMV          | Multiple measurement vector                     |
| BHIT         | Binary IHT                                       | MSP          | Matching sign pursuit                            |
| BP           | Basis pursuit                                    | MT-BCS       | Multitask BCS                                   |
| BPDN         | Basis pursuit denoising                         | NHTP         | Normalized BCS                                  |
| BBSL         | Block BLS                                        | NIHT         | Normalized IHT                                  |
| CB-DIHT      | Consensus based distributed IHT                  | NP           | Neyman Pearson                                  |
| CB-DSBL      | Consensus based distributed SBL                  | OMP          | Orthogonal matching pursuit                      |
| CoSaMP       | Compressive sampling matching pursuit            | PoI          | Phenomenon of interest                           |
| CS           | Compressive sensing                              | RIP          | Restricted isometry property                     |
| CWs          | Compressive wireless sensing                     | RLS          | Recursive least squares                          |
| DBs          | Distributed BP                                   | RSSI         | Received signal strength indicator               |
| DC-OMP       | Distributed and collaborative OMP                | RVM          | Relevance vector machines                        |
| DCSP         | Distributed and collaborative SP                 | SBL          | Sparse Bayesian learning                         |
| DCT          | Discrete cosine transform                        | SCoSaMP      | Simultaneous CoSaAMP                             |
| DIHT         | Distributed IHT                                  | SCS          | Sequential CS                                   |
| DiOMP        | Distributed OMP                                  | SDP          | Semidefinite programming problem                 |
| DiSP         | Distributed SP                                   | SHTP         | Simultaneous HTP                                |
| DOI          | Difference of innovations                        | SIHT         | Simultaneous IHT                                |
| DR-LASSO     | Decentralized row-based LASSO                    | SIOMP        | Side information based OMP                      |
| DWT          | Discrete wavelet transform                       | SMV          | Single measurement vector                        |
| FOCUSS       | FOCal Underdetermined System Solver             | SNHTP        | Simultaneous NHTP                               |
| FPC          | Fixed point continuation                         | SNIHT        | Simultaneous NHT                                |
| GAMP         | Generalized AMP                                  | S-OPT        | Simultaneous OMP                                |
| GLRT         | Generalized likelihood ratio test                | SOCP         | Second-order cone programming                    |
| GMM          | Gaussian mixture model                           | S-OMP        | Simultaneous OMP                                |
| GMP          | Greedy matching pursuit                          | SP           | Subspace pursuit                                |
| GSM          | Gaussian scale mixture                           | SR           | Sparse representation                           |
| GSP          | Generalized SP                                   | SSP          | Simultaneous SP                                 |
| HTP          | Hard thresholding pursuit                        | VQ           | Vector quantizers                               |
| IHT          | Iterative hard thresholding                      | WSN          | Wireless sensor network                         |
Greedy and iterative algorithms: Greedy/iterative algorithms aim to solve (4) (or its noise resistant extension) in a greedy/iterative manner which are in general less computationally complex than the optimization based approaches. Examples of such algorithms include orthogonal matching pursuit (OMP) [60] and its variants such as regularized OMP [61], and stagewise OMP [62]. Compressive sampling matching pursuit (CoSAMP) [63], subspace pursuit (SP) [64], iterative hard thresholding (IHT) [65], [66] and its variants such as Normalized IHT (NiHT) [67] and Hardthresholding pursuit (HTP) [68].

Bayesian algorithms: Another class of sparse recovery algorithms falls under the Bayesian formulation. In the Bayesian framework, the sparse signal reconstruction problem is formulated as a random signal estimation problem after imposing a sparsity promoting pdf on x in (2). A widely used sparsity promoting pdf is the Laplace pdf [69]. With Laplace prior, x is imposed with the pdf,

\[ p(x|\rho) = \left(\frac{\rho}{2}\right)^{\rho/2} e^{-\sum_{i=1}^{N} |x(i)|} \]

where \( \rho > 0 \). When the noise \( v \) in (2) is modeled as Gaussian with mean 0 and covariance matrix \( \sigma^2 I \), the solution in (8) corresponds to a maximum a posteriori (MAP) estimate of x with the prior [9]. Computation of the MAP estimator in closed-form with the Laplace prior is computationally intractable, and several computationally tractable algorithms have been proposed in the literature. Sparse signal recovery using sparse Bayesian learning (SBL) algorithms has been proposed in [70]. In [69], a Bayesian CS framework has been proposed where relevance vector machines (RVM) [71] are used for signal estimation after introducing a hierarchical prior which shares similar properties as the Laplace prior, yet, providing tractable computation. Babacan et.al. in [72] also have considered a hierarchical form of the Laplace prior of which RVM is a special case. CS via belief propagation has been considered in [73]. Bayesian CS by means of expectation propagation has been considered in [74]. An interesting characteristic of the Bayesian formulation is that it lets one exploit the statistical dependencies of the signal or dictionary atoms while developing sparse signal recovery algorithms. The authors in [75] have considered the problem of sparse signal recovery in the presence of correlated dictionary atoms in a Bayesian framework.

A review on some of the convex relaxation-based and greedy sparse recovery algorithms can be found in [44]. Most of these recovery techniques are extended for the multiple sensor setting by various authors so that they are applicable for WSNs as will be discussed in Section IV. Before discussing the specific sparsity aware algorithms developed for WSN applications, in the next section, we motivate the use of CS techniques discussing several forms of sparsity exhibited in WSN applications.

Fig. 1: Acquisition of compressed measurements of observations with temporal sparsity

III. SPARSITY IN DISTRIBUTED WIRELESS SENSOR NETWORKS

Sparsity is one of the common characteristics observed in WSNs in different forms. In this section, we review different forms of sparsity and motivate the use of CS based signal processing techniques in WSNs.

A. Temporal sparsity

Consider a WSN with L sensor nodes observing a PoI. The time samples collected at the j-th node are represented by the vector \( \mathbf{x}_j \in \mathbb{R}^N \). Further, let \( \mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_L] \in \mathbb{R}^{N \times L} \).

Sensor readings, \( \mathbf{x}_j \)'s, with many types of sensors such as acoustic, seismic, IR, pressure, temperature, etc., can have sparse representation in a certain basis. For example, audio signals collected by acoustic sensors (microphones) can be sparsely represented in DCT and DWT [76]. Formally, with an orthonormal basis \( \Psi_j \in \mathbb{R}^{N \times N} \), when \( \mathbf{x}_j \) is represented as \( \mathbf{x}_j = \Psi_j s_j \) for \( j = 1, \ldots, L \), \( \mathbf{x}_j \) is said to be sparse when \( s_j \) contains only a few nonzero elements compared to \( N \). This motivates each sensor to employ the CS measurement scheme to compress its time samples before transmission/further processing (Fig. 1). The compressed measurement vector at the j-th node is given by \( \mathbf{y}_j = \mathbf{A}_j \mathbf{x}_j \) where \( \mathbf{A}_j \in \mathbb{R}^{M \times N} \) with \( M < N \) is the measurement matrix.

Let \( \mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_L] \). The MMV model in matrix form with the same projection matrix so that \( \mathbf{A}_j = \mathbf{A} \) for \( j = 1, \ldots, L \) is given by

\[ \mathbf{Y} = \mathbf{AX}. \]

For the more general case with different projection matrices, the observation matrix can be expressed as

\[ \mathbf{Y} = [\mathbf{A}_1 \mathbf{x}_1, \ldots, \mathbf{A}_L \mathbf{x}_L]. \]

In the simultaneous sparse approximation framework, the goal is to recover \( \mathbf{X} \) given \( \mathbf{Y} \) (or its noisy version) with known \( \mathbf{A} \) (with the same projection matrix) or \( \mathbf{A}_1, \ldots, \mathbf{A}_L \) (with different matrices). When processing \( \mathbf{Y} \) efficiently under resource
constraints, there are several important factors to be focused on. One of the factors is to exploit joint structures of the sparse data vectors in $\mathbf{X}$. Another important factor is the communication architecture. In centralized processing, it is assumed that $\mathbf{Y}$ is available at a central fusion center. However, such a scheme may be prohibitive in many resource constrained WSNs. With a distributed or decentralized communication model, inference regarding the PoI needs to be made without having access to complete knowledge of $\mathbf{Y}$. There have been recent research efforts that focus on processing temporal sparse signals jointly under communication constraints which will be discussed in Section IV.

B. Spatial sparsity

Spatial sparsity can be exploited for efficient signal processing in WSNs in several ways. Spatial sparsity (mostly in an approximate sense) can be observed in the data collected at distributed sensors at a given time. By employing proper communication architectures, a compressed version of the data vector (over all the nodes) can be made available at the fusion center. Further spatial sparsity can be exploited for sparse event/source monitoring and localization. In the following, we discuss how CS can be applied while exploiting spatial sparsity focusing on data gathering and sparse event detection/localization.

1) Data gathering: Consider a WSN with a large number of distributed sensors, i.e., $L$ is large. The observation vector collected at all the nodes at a given time can be sparse or compressible since not all the nodes will have significant observations. For example, consider the problem of aggregation of signals emitted by an acoustic source by a network of sensors. Due to the signal decay being a function of the distance between the source and a given sensor, the signal amplitude at the sensors located far from the source will be very small. Thus, at any given time instant, the concatenated observation vector (over all the nodes) will have only a certain number of elements with significant amplitudes resulting an approximately sparse vector. This results in spatial sparsity of networked data. Spatial sparsity has motivated researchers to develop efficient distributed compression techniques to solve a variety of inference problems in WSNs.

To exploit CS techniques via spatial sparsity, several techniques have been proposed in the recent literature. One of the architectures, known as compressive wireless sensing (CWS), has been proposed in [31]. [77]. The CWS framework enables distributed compression employing multiple access channels (MACs) where individual nodes transmit their scaled observations to the fusion center coherently via a MAC. This architecture is depicted in Fig. 2 for one MAC transmission.

With this model, the observation at the $k$-th sensor, denoted by $x_k$, is multiplied by a scalar factor, and transmitted to the fusion center via a MAC channel. The received signal at the fusion center after the $i$-th MAC transmission is given by,

$$ y_i = \sum_{k=1}^{L} A_{i,k} x_k + v_i $$  \hspace{1cm} (12)

where $A_{i,k}$ is the scalar factor, and $v_i$ is the receiver noise. After $M < L$ such MAC transmissions, the observation vector at the fusion center can be expressed as

$$ y = A\mathbf{x} + \mathbf{v} \hspace{1cm} (13) $$

where the $(i,k)$-th element of $A$ is given by $A_{i,k}$ for $i = 1, \ldots, M$ and $k = 1, \ldots, L$, $\mathbf{x} = [x_1, \ldots, x_L]^T$ and $\mathbf{v} = [v_1, \ldots, v_M]^T$. With this model, the fusion center receives a compressed version of the sparse (or compressible) signal $\mathbf{x}$. In the CWS framework, one of the main concerns which gained a considerable attention was the selection of the scalar parameters in $A$. As in CS theory, if dense matrices such as Gaussian or Bernoulli with entries $[-1, +1]$ are used, all the nodes have to transmit their multiplied observations to the fusion center. However, in resource constrained WSNs, this may be not desirable due to the consumption of a large amount of transmit power. The use of sparse random projections to minimize the transmit power and for sensor management has gained a considerable attention in the CWS related works. Further, in many practical situations, the sparse signal recovery problem can be different from what is widely considered in the standard CS framework. For example, in the presence of fading, the effective measurement matrix at the fusion center may violate many desirable properties as discussed in the CS theory. This has led to new research directions on establishing recovery guarantees under practical conditions.

Another architecture to exploit spatial sparsity in efficient data gathering has been proposed in [78]. In this framework, the data aggregation process is implemented with multi-hop routing as illustrated in Fig. 2. A tree based architecture is used to implement the multi-hop routing protocol so that the fusion center receives the observation vector $\mathbf{y} = A\mathbf{x}$ where $A$ and $\mathbf{x}$ have the same notational meaning as in the previous model.

2) Sparse events monitoring: Another application of CS in WSNs which exploits spatial sparsity is multiple event detection/monitoring. There are several ways that spatial sparsity can be exploited.

Model 1: The works in [79], [80] model the sparse event monitoring task as a non-negative least squares problem as follows. Let there be $L$ distributed sensors each having a position denoted by $r_i$ for $i = 1, \ldots, L$. The sources of events are confined to sensor points; i.e., an event occurs only at a sensor point as shown in Fig. 3.

The magnitude of the event at the $r_i$-th location is denoted by a positive scalar $s_i$. If no event occurred at the location $r_i$, $s_i = 0$, then those sensors are called inactive sensors. Forming
where \( \Psi \) the position \( r \)

\[
\lambda \text{ where } K
\]

considers the sparse event detection problem in the following

\[
M < L \]

The observation vector at \( M \) < \( L \) active sensors in vector-

matrix notation can be expressed as,

\[
y = \mathbf{A} \Psi \mathbf{s} + \mathbf{v}
\]

where \( y = [y_1, \cdots, y_M]^T \), \( \Psi \) is an \( L \times L \) matrix with \( \Psi_{i,j} \) being the \((i,j)\)-th element, \( \mathbf{A} \) is an \( M \times L \) matrix which selects the \( M \) rows of \( \Psi \) corresponding to active sensors and \( \mathbf{v} = [v_1, \cdots, v_M]^T \). In this formulation, the problem of detecting the active sources is formulated as

\[
\arg \min_\mathbf{s} \frac{\lambda}{2} || \mathbf{A} \Psi \mathbf{s} - \mathbf{y} ||_2^2 + || \mathbf{s} ||_1 \text{ such that } \mathbf{s} \geq \mathbf{0}
\]

where \( \lambda \) is a penalty parameter.

**Model 2:** In contrast to Model 1, the authors in [49] considers the sparse event detection problem in the following sense. Let there be altogether \( K \) sources in which \( k \) (out of \( K \)) are active. In particular, the \( k \) events are assumed to occur simultaneously.

\[
\mathbf{y} = \Psi(\mathbf{r})\mathbf{s} + \mathbf{v}
\]
where the \((i, j)\)-th element of the \(M \times K\) array manifold matrix \(\Psi(r)\) contains the delay and gain information from the \(j\)-th source located at \(r_j\) to the \(i\)-th sensor. In the SR framework, \([19]\) is represented as an overcomplete representation by constructing a fine grid, to account for all possible source locations, over the region of interest. Let \(\tilde{r} = [\tilde{r}_1, \ldots, \tilde{r}_L]^{T}\) denote the grid locations where \(L >> K\). Then, the matrix \(\Psi(r)\) is redefined as a \(M \times L\) matrix whose \((i, j)\)-th element corresponds to the gain and delay information between the source location \(\tilde{r}_j\) and the \(i\)-th sensor. The signal field is given in \(\tilde{s}\) where the \(n\)-th element of \(\tilde{s}\) is nonzero if the \(n\)-th location has an active source while it is zero otherwise. With this representation, \([19]\) can be rewritten as

\[
y = \Psi(\tilde{r})\tilde{s} + v
\]

where \(\tilde{s}\) is a sparse vector. In this formulation, the sparse localization problem reduces to estimating a sparse signal \(\tilde{s}\) based on \(y\) where the elements of the \(\Psi(\tilde{r})\) are determined by the particular signal model under consideration.

When exploiting spatial sparsity as discussed above, the effective measurement vector to be processed \([13, 15, 17, 20]\) has a similar form as that in the standard CS framework. However, due to additional constraints, different properties of the effective measurement matrix depending on the particular applications, and the impact of imperfect channels such as channel fading, the existing work on the standard CS model cannot be directly applied. Related work on specific algorithms and theoretical recovery guarantees exploiting spatial sparsity under these conditions will be discussed in Section IV in detail.

IV. Sparse Signal Recovery Techniques Exploiting Temporal and Spatial Sparsity

In this section, we discuss energy efficient signal processing techniques developed for sparse signal recovery exploiting temporal and spatial sparsity in sensor networks. We discuss the algorithms focusing on two main categories; centralized and distributed/decentralized solutions.

A. Centralized solutions for simultaneous sparse signal recovery exploiting temporal sparsity

First, we focus on sparse signal recovery algorithms developed exploiting temporal sparsity. In simultaneous sparse approximation, the goal is to reconstruct a set of sparse vectors jointly based on compressed observations obtained at individual nodes. In the centralized setting, it is assumed that the compressed observations \(Y\) as given in \([10]\) or \([11]\) are available at a central processing unit without any errors. In order to jointly estimate \(X\) (or the support of \(X\)), joint sparse structures of \(X\) can be exploited. There are several such structures proposed in the literature \([43, 82, 83]\).

1) Common support set model: The widely considered joint sparse model for \(X\) is the common support set model which is termed as JSM-2 model in \([83]\). In this model, the sparse signals observed at multiple nodes, \(x_j\)'s, have the same but unknown sparsity pattern. However, the corresponding amplitudes can be different in general. The common support set model with the same measurement matrix as in \([10]\) is commonly termed as the MMV model \([84, 85]\). This model has been investigated by many authors \([84]–[89]\) from different perspectives be it performing theoretical recovery guarantees and/or algorithm development.

While developing algorithms and evaluating performance with the common support set model, several measures have been defined. To measure the number of nonzero elements of \(X\), \(\|\cdot\|_{row-0}\) norm is widely used where

\[
\|X\|_{row-0} = |\text{rowsupp}(X)|
\]

with

\[
\text{rowsupp}(X) = \{i \in [1, \cdots, N] : X_{i,j} \neq 0 \text{ for some } j\}.
\]

The natural approach to solve for \(X\) from \(Y\) in \([10]\) is to solve the following optimization problem:

\[
\min_X \|X\|_{row-0} \text{ such that } Y = AX
\]

with no noise or

\[
\min_X \|X\|_{row-0} \text{ such that } \frac{1}{2}\|Y - AX\|_F \leq \epsilon
\]

in the presence of noise. Since the problem \([23]\) (and \([24]\)) is NP hard, one often solves \([23]\) (and \([24]\)) by using the mixed norm approach which is an extension of the convex relaxation method in SMV to the MMV case.

Convex relaxation: A large class of relaxation versions of \(\|X\|_{row-0}\) aims at solving the optimization problem of the following form \([86]\):

\[
\min_X J_{p,q}(X) \text{ such that } Y = AX
\]

in the noiseless case and

\[
\min_X \frac{1}{2}\|Y - AX\|_F^2 + \lambda J_{p,q}(X)
\]

in the noisy case where \(\lambda\) is a penalty parameter and

\[
J_{p,q}(X) = \sum_i (\|x_i\|_p)^q
\]

where typically \(p \geq 1\) and \(q \leq 1\). Different approaches have been proposed in the literature to solve the optimization problems in \([23]\) and \([24]\). The case where \(p = 2\) and \(q = 1\) is commonly known as M-BP for which recovery guarantees and algorithm development have been discussed in \([44]\), \([81]\), \([84]\), \([86]\), \([87]\), \([90]–[93]\). This case is also dubbed as the mixed \(l_2/l_1\) norm minimization approach. The properties of the M-BP cost function and algorithms for its minimization have been explored in \([81]\), \([84]\). The average case analysis on recovery guarantees using multichannel BP has been discussed in \([87]\). In \([91]\), alternating direction methods (ADM) based approach has been proposed to solve \([26]\) which is called MMV-ADM. The M-BP problem as a special case of group LASSO, with the block coordinate descent algorithm called M-BCD has been discussed in \([86]\), \([92]\), \([93]\). In \([90]\), the mixed \(l_2/l_1\) type norm minimization problem has been solved via a semi-definite program which is shown to reduce to solving a second-order cone program (SOCP)-a special type of semi-definite program. In \([83]\), the case where \(p = \infty\) and
q = 1 has been considered and the authors have discussed the conditions under which the convex relaxation is capable of ensuring recovery guarantees. Some known theoretical results on recovery guarantees of SMV have been generalized to MMV in [85] considering \( q = 1 \) and arbitrary \( p \). In [84], M-FOCUSS (FOCal Underdetermined System Solver) is developed taking \( p = 2 \) and \( q \leq 1 \) for the noiseless case, and the regularized M-FOCUSS for the noisy case. M-FOCUSS is an iterative algorithm that uses Lagrange multipliers. Recovery guarantees using fast thresholded Landweber algorithms considering \( p = 1, 2, \infty \), and \( q = 1 \) have been discussed in [94]. A comparison of different simultaneous sparse approximation methods with different values for \( p \) and \( q \) can be found in [86], [88]. In Table [11] we summarize the relaxations considered in the literature along with specific algorithms developed (if any) with different values for \( p \) and \( q \).

**Greedy algorithms:** In order to solve (23), several greedy and iterative algorithms have been proposed which typically have less computational complexity than convex relaxation based approaches. In particular, most of these approaches are extensions of their SMV counterparts. The extension of the OMP algorithm to the MMV case with the common support set model, S-OMP, has been considered in [89]. Performance analysis of the S-OMP algorithm including noise robustness has been presented in a recent paper [110]. An MMV extension of the IHT algorithm, SIHT, has been considered in [90], [91]. Some variants of SIHT such as SNIHT, and extensions of other greedy algorithms such as HTP, NHTP, and CoSaMP have also been discussed in [90], [92]. Generalized SP as an extension of the SP algorithm with MMV has been considered in [99] which is shown to outperform the natural extension of SP, i.e., simultaneous SP (S-SP).

**Bayesian algorithms:** Solutions to the MMV problem with the common support set model in the Bayesian setting have been discussed in [100]–[106], [109]. The MSBL algorithm, which is an extension of the SBL algorithm with SMV, has been developed in [100]. Multitask BCS (MT-BCS) has been developed in [104] in which applications of multitask learning to solve the MMV problem in the Bayesian framework have been discussed. In [105], the authors have extended the MT-BCS framework to take the intra-task dependency into account. In [106], the MMV problem in a Bayesian setting has been considered focusing on DOA estimation where the prior probability distribution is modeled with a Gaussian scale mixture (GSM) model. The proposed approach is dubbed M-BCS-GSM. Belief propagation based methods have been proposed in [101], [102] in which the approximate message passing (AMP) framework is used to jointly estimate the sparse signals. In [103], the MMV problem has been solved when the nonzero elements of a given column of \( X \) are correlated and the two Bayesian learning algorithms, called T-SBL and T-MSBL, have been developed.

**Other approaches:** The MMV problem is treated as a block sparse signal recovery problem after vectorizing \( X \) to a block sparse signal in [108]. Then, the algorithms developed for block sparse signal recovery with SMV can be used to solve the MMV problem. In [109], block SBL (BSBL) algorithms, which take intra-block correlations into account, have been developed where the MMV problem is treated as a block sparse signal recovery problem. In [107], the authors have shown that most simultaneous sparse recovery algorithms discussed above are rank blind. They have proposed rank aware algorithms with mixed norm minimization as well as greedy algorithms which show better performance than rank blind algorithms under certain conditions.

With different measurement matrices, the goal is to recover \( X \) from \( Y \) where \( Y \) is as in (11). Joint recovery of \( X \) for this case is considered in [111] which employs \( l_2/l_1 \) norm minimization. The S-OMP algorithm to account for different measurement matrices is extended in [83].

2) **Common support set + innovation model:** Another widely considered joint sparse support set model with many WSN applications is the common support set + innovation model, which is termed as the JSM-1 model in [83]. It is assumed that each sparse signal \( x_j \) can be expressed as \( x_j = x_{Cj} + \tilde{x}_j \) where \( x_{Cj} \) is a common (to all \( x_j \)'s) component which is sparse and \( \tilde{x}_j \) is called the innovation component which is also sparse but different from \( x_{Cj} \)'s sparsity pattern. Recovery of \( X \) with the JSM-1 model has been discussed in [83]. In [83], weighted \( l_1 \) norm minimization for jointly estimating \( X \) is discussed. Considering the signal vector at one node as side information, the authors in [112] have developed the Difference-Of-Innovations (DOI) and Texas DOI algorithms for simultaneous sparse signal approximation considering the JSM-1 model. In [113], centralized (as well as distributed) implementation of the ADMM algorithm with MMV has been presented.

Although not as interesting as the JSM-1 and JSM-2 models, some works can be found for the JSM-3 model discussed in [83]. In the JSM-3 model, the signal observed at each node is assumed to be composed of a nonsparse common component + sparse innovation component. Applications of this model in sensor network settings have been discussed in [83]. The algorithms developed in [112] for the JSM-1 model are applicable to the JSM-3 model as well.

B. **Centralized solutions for sparse signal recovery exploiting spatial sparsity**

As discussed in Section [118], the problems of compressive data gathering and sparse event detection can be formulated as CS recovery problems. However, in contrast to the standard CS framework, this problem needs further attention when solving under resource constraints. Further, additional factors such as channel fading need to be taken into account in establishing recovery guarantees.

1) **Compressive data gathering:** Exploitation of spatial sparsity in data gathering and sparse event detection has been addressed by many authors in different contexts. In [114], the authors have exploited adaptive CS for efficient data gathering exploiting spatial sparsity. In the CWS framework proposed by [81], [77], the projection operation can be treated as distributed since nodes participate in a distributed manner via MAC to form the compressed observation vector in [13]. The authors in [77] analyze power, distortion, and latency relationships as a function of the number of sensors in the network.
the CWS framework, the designer has the choice to design the projection matrix so that the energy and the bandwidth requirements are optimized. One direction of research in exploiting spatial sparsity is the design of projection matrices such that the communication overhead is minimized. To that end, the authors in [115–118] have used sparse random projection matrices under which, only a subset of sensors transmit during a given MAC transmission.

2) Sparse random projections for CWS: The use of sparse random projections has gained much interest in the recent literature [3, 42, 116–121]. Some widely considered sparse random projections are summarized below. In particular, the (i, k)-th element of \( A \) is drawn from [3]

\[
A_{i,k} = \begin{cases} 
1 & \text{with prob } \frac{\gamma}{2} \\
0 & \text{with prob } 1 - \gamma \\
-1 & \text{with prob } \frac{\gamma}{2} 
\end{cases}
\]  

or \[A_{i,k} = \begin{cases} 
\mathcal{N}(0, \frac{1}{\gamma}) & \text{with prob } \gamma \\
0 & \text{with prob } 1 - \gamma 
\end{cases}
\]

with \( 0 < \gamma < 1 \). The CWS scheme with sparse random projections is depicted in Fig. 6.

Using the sparse matrix [43], in [3], the authors have shown that reliable sparse signal recovery can be performed querying only a small number of nodes anywhere in the network. They further investigate the trade-off, which can be controlled by the sparsity of the matrix, between the communication cost to pre-process the data in the network, and the query latency to obtain the desired approximation error. In [117, 118], recovery guarantees have been established with sparse matrices when the channels between the sensors and the fusion center undergo fading.

3) Sparse event monitoring: Centralized algorithms for sparse event monitoring and source localization exploiting spatial sparsity have been developed in [49, 112–126]. In [49], Bayesian algorithms exploiting marginal likelihood maximization have been developed for sparse event detection. The sparse event detection problem is formulated from the perspective of coding theory by modeling the detection problem as a decoding of analog fountain codes (AFCs) in [126]. In [122], the problem of sparse event detection has been addressed in an adaptive manner using sequential compressive sensing (SCS). In [123], a CS-based source localization approach has been proposed with multiple sensors using RSSI measurements. Constructing a multisolution dictionary, the authors in [124] have proposed a \( l_1 \)-norm minimization based source localization scheme which again uses RSSI values. In [125], the authors have discussed a Bayesian formulation of the localization problem and have posed it as a sparse recovery problem.

C. Algorithms for sparse signal recovery in distributed/decentralized settings

In decentralized solutions, it is assumed that the distributed nodes are not connected with a central fusion center. Decentralized solutions are attractive and are sometimes necessary in resource contained sensor networks. These algorithms can be classified into multiple classes depending on the objective (estimation of a single or multiple sparse signals jointly), and how sparsity is exploited. To make the presentation clear, we categorize decentralized algorithms into two classes: (i) estimation of a single sparse vector in a decentralized manner and (ii) estimation of multiple sparse signals jointly. It is further noted that some of these algorithms are developed focusing on a specific task which adds more constraints compared to the standard CS problem.

1) Decentralized implementation for estimating a single sparse signal using partially available data at each node: As mentioned under the first category above, one approach for decentralized estimation is to estimate a single sparse vector in a distributed/decentralized manner when a given node has access to some partial information, e.g., partial information on \( A \). For example, in [127], distributed basis pursuit (DBS) has been developed based on ADMM in which the BP problem [5] is solved in a distributed fashion. In particular, the authors consider two distributed frameworks. In one framework, \( A \) is partitioned by rows, with its rows distributed over a network with an arbitrary number of nodes. In the other framework, it is the columns of the matrix that are distributed. In [128, 129], the distributed LASSO (D-LASSO) algorithm has been employed to compute the common LASSO estimator by collaboration among nodes. In [80], the \( l_1 \) regularized least squares problem with an additional (nonnegative) constraint as given in [10] has been solved in a decentralized manner after reformulating it as an equivalent consensus optimization problem. Several distributed sparse recursive least squares (RLS) algorithms have been proposed in [130] by imposing \( l_1 \) norm sparsity penalties. Sparsity promoting adaptive learning has been exploited for distributed estimation in [131]. In [132–134], distributed least-mean squares (LMS) estimation is considered by taking the sparse property into account. These

| Approach | Algorithms and References |
|----------|---------------------------|
| Convex relaxation, \( p = 2, q = 1 \) | Group LASSO (with BCD) [56, 92, 93], SDP-SOCP [90], MMV-ADM [21], MMV_prox [95] |
| Convex relaxation, \( p = 2, q \leq 1 \) | M-FOCUSS [84] |
| Convex relaxation, \( p = 1, 2, \infty, q = 1 \) | Landweber algorithms [24] |
| Convex relaxation, \( p = \infty, q = 1 \) | [85] |
| Greedy and iterative | S-OMP [79], SIHT, SNIHT [95, 97], SHTP [98], SNHTP [96], SCoSaMP [96], Generalized SP [99] |
| Bayesian | MSBL [103], MMV-AMP [104, 105], T-MSBL [103], MT-BCS [104, 105], M-BCS-GSM [106] |
| Other approaches | Rank-aware algorithms [107], block sparse signal recovery based algorithms [108], BSBL [109] |
sparser algorithms can benefit from the underlying sparsity to reduce the estimation error.

2) Simultaneous sparse approximation in a decentralized manner: In the other class of distributed/decentralized processing, simultaneous estimation of a set of jointly sparse signals (or the support of the signals) based on MMV is carried out in a decentralized manner. In particular, each node compresses its temporally sparse observation vector with a measurement matrix (which can be the same or different among multiple nodes) and joint sparse signal/support recovery is performed when all the compressed observations are not available at a central processing unit. The decentralized extensions of widely considered simultaneous sparse recovery algorithms for the common sparse support set model can be found in [79], [113], [135]–[142]. In [136], a non-convex optimization model has been discussed using reweighted $l_q$ norm minimization for simultaneous sparse signal recovery in a decentralized manner. In [135], a distributed soft-thresholding algorithm (DJ-IST) via reweighted $l_1$ norm minimization technique has been proposed. In [137], a decentralized row-based LASSO (DR-LASSO) algorithm has been proposed to reconstruct the common sparse support with MMV. In [142], a gossip based algorithm has been developed to implement the $l_1$ norm minimization approach in a decentralized manner. Formulating the $l_1$ norm minimization problem as a bound-constrained quadratic program, the gradient computation at each step of the coordinate descent algorithm is expressed as a sum of quantities computed at each node applying distributed consensus algorithm. In [113], the authors have proposed a distributed version of the ADMM algorithm (dubbed as DADMM) considering the JSM-1 model. In [79], the problem of multiple event detection has been formulated as a $l_1$ regularized nonnegative least squares problem and solved in a decentralized manner using two approaches; partial consensus based and the Jacobi approach.

Distributed and decentralized versions of greedy/iterative algorithms have been discussed in [133]–[134], [143]–[146]. In [139], [146], distributed SP (DiSP) and Distributed OMP (DiOMP) have been developed which are applicable to both JSM-1 (common support + innovation) and JSM-2 (common support set) models. In [147], the distributed parallel pursuit algorithm is proposed for the common support + innovation model. Distributed IHT and a consensus based distributed IHT named (CB-DIHT) have been proposed in [138], [148]. In [143], side information based OMP (SiOMP) is proposed. In SiOMP, the distributed CS is performed for the JSM-1 model where the estimate at one node via OMP is considered to be the initial value for the estimate at the next node. Embedding fusion within OMP iterations, in [145], the authors have proposed two decentralized versions of the OMP algorithm, called DC-OMP 1 and DC-OMP 2 for the JSM-2 model. In [149], [150], the distributed and collaborative subspace pursuit (DCSP) algorithm has been developed for the JSM-2 model. A review on optimization based and greedy based distributed compressive sensing algorithms is provided in [140] focusing on signal models, network topologies and some distributed versions of sparse recovery algorithms.

Distributed and decentralized Bayesian algorithms for sparse signal recovery have been proposed in [151]–[153]. In [152], a decentralized Bayesian CS framework has been proposed for the common support set + innovation model in which multiple sparse signals are constructed in a decentralized manner. In [151], an approximate message passing (AMP) based decentralized algorithm, AMP-DCS, is developed to reconstruct a set of jointly sparse signals. In [153], a Consensus Based Distributed Sparse Bayesian Learning (C-DSBL) algorithm for decentralized estimation of joint sparse signals is proposed.

In Table III, we summarize the distributed and decentralized sparse recovery algorithms discussed above.

V. CS BASED DISTRIBUTED INFERENCE

In WSNs, there are applications where complete signal reconstruction is not necessary to make an inference decision. For example, in order to perform detection, classification and tracking, it is sufficient to estimate a reliable decision statistic without completely recovering the signal. Investigation of the capability of the CS mechanism to solve such inference problems under resource constraints is attractive in WSNs. In the following, we discuss the use of CS techniques to solve detection, classification and tracking problems with/without exploiting the sparsity prior.

A. Compressive detection

Theories and concepts developed in CS for sparse signal recovery have been exploited in the recent literature for signal detection problems [49], [55], [144], [154]–[170]. These works include deriving performance bounds for CS based signal detection [55], [155]–[157], [161], [162], [164], [165], [167], [169], developing algorithms [144], [154], [160] and designing low dimensional projection matrices [158], [159], [166]. While some of the works, such as [144], [154], [155], [160], [163]–[165], [168], [169] have focused on sparse signal detection when the underlying subspace where the signal lies

Fig. 6: CWS with sparse random projections; only few sensors transmit during a given MAC transmission
TABLE III: Decentralized algorithms for joint sparse signal recovery

| Approach               | Algorithms and References |
|------------------------|---------------------------|
| Optimization based     | $l_1$ regularized least square [80], [143] |
|                        | Distributed BP [127], Distributed LASSO [128], [129] |
|                        | DR-LASSO [117], DADMM [113] |
| Greedy and iterative   | reweighted $l_q$ norm minimization [156] |
|                        | reweighted $l_1$ norm minimization [155] |
| Bayesian               | DCS-AMP [153], CB-DSBL [153] |

is unknown, some other works [55], [157], [158], [161], [162] have considered the problem of detecting signals which are not necessarily sparse. CS based detection approaches can be categorized into two classes. In one class, the detection problem is solved after estimating, partially or completely, the signal of interest exploiting the sparsity prior. In the other class, the detection problem is solved completely in the compressed domain where, the signal of interest is not estimated to perform detection.

1) Detection via partially/completely reconstructing the signals: In CS based sparse signal detection, the main objective is to utilize a small number of compressed measurements to extract decision statistics to make a reliable detection decision. Let the goal be to solve the following binary hypothesis testing problem:

$$
\mathcal{H}_1 : y_j = \theta_j + v_j \\
\mathcal{H}_0 : y_j = v_j
$$

(30)

where $\theta_j$ is the (unknown) signal observed by the $j$-th node and $v_j$ is the additive noise for $j = 1, \cdots, L$. Let the signal to be detected be sparse in the basis $\Psi$ so that $\theta_j = s_j(u)$ with $s_j$ having only a few nonzero elements. With the common support set model, the coefficients $s_j$ for $j = 1, \cdots, L$ share the same support. Now assume that this problem is solved after compression via random projections:

$$
y_j = A_j x_j
$$

(31)

for $j = 1, \cdots, L$ where $A_j$ is an $M \times N$ ($M < N$) matrix. Let $B_j = A_j \Psi$. When the elements of $A_j$ are selected so that $A_j$ is an orthoprojector, $A_j A_j^T = I_M$. Then, $B_j B_j^T = I_M$ for $j = 1, \cdots, L$ when $\Psi$ is an orthonormal basis. The goal is to decide between hypotheses $\mathcal{H}_1$ and $\mathcal{H}_0$ based on $\tilde{y}_j$.

As defined before, let $U$ be the set which contains the indices of the locations of nonzero coefficients in $s_j$ so that $U := \{i \in \{1, \cdots, N\} \mid s_j[i] \neq 0\}$ where $s_j[i]$ denotes the $i$-th element of $s_j$ for $i = 1, \cdots, N$ and $j = 1, \cdots, L$. Then, $k = |U|$, where $|U|$ denotes the cardinality of $U$. It is noted that $U$ is the same for all the signals $s_j$ with the common sparse support model. Further let $\tilde{v}_j = A_j v_j$, where $\tilde{v}_j \sim (0, \sigma_v^2 I_M)$ when $AA^T = I_M$. Then, the detection problem in the compressed domain can be expressed as:

$$
\mathcal{H}_1 : y_j = B_j s_j + \tilde{v}_j \\
\mathcal{H}_0 : y_j = \tilde{v}_j
$$

(32)

for $j = 1, \cdots, L$.

When $U$ is exactly known, the detection problem in (32) reduces to

$$
\mathcal{H}_1 : y_j = B_j s_j(u) + \tilde{v}_j \\
\mathcal{H}_0 : y_j = \tilde{v}_j
$$

(33)

for $j = 1, \cdots, L$ where $B_j(u)$ denotes the $M \times k$ submatrix of $B_j = A_j \Psi$ in which columns are indexed by the ones in $u$, and $s_j(u)$ is a $k \times 1$ vector containing nonzero elements in $s_j$ indexed by $U$ for $j = 1, \cdots, L$. When $B_j(u)$ is known, $\mathcal{H}_j(u)$ is the subspace detection problem which has been addressed previously [171]–[173]. Depending on how the unknown coefficient vector $s_j(u)$ is modeled, different detectors have been proposed. In [171], a generalized likelihood ratio test (GLRT) based detector is proposed when $s_j(u)$ is assumed to be deterministic. In [172], the analysis has been extended to the case when $s_j(u)$ is modeled as random. The problem with multiple observation vectors has been addressed in [173] where the authors have proposed adaptive subspace detectors when the coefficients $\{s_j(u)\}_{j=1}^2$ follow first and second order Gaussian models.

In the case of sparse signal detection, it is unlikely that the exact knowledge of $U$ is available a priori. In other words, sparse signal detection needs to be performed when $U$ is unknown. With the advancements of CS, some algorithms have been developed to detect sparse signals based on exploiting algorithms developed for sparse signal recovery [144], [154], [155], [160]. In particular, the standard OMP algorithm is modified in [154] to detect the presence of a sparse signal based on a SMV. There, the detection decision is obtained after running a few iterations ($\leq k$) of the OMP algorithm. Let $\hat{s}_{t_0}$ be the estimated signal after running $t_0 \leq k$ number of iterations. Then, the decision statistic is taken to be the maximum absolute component of $\hat{s}_{t_0}$ ($||\hat{s}_{t_0}||_{\infty}$) in [154]. In [168], the sparse signal detection problem has been addressed with MMV. The authors have derived the minimum fraction of the support of the signal to be estimated to achieve a desired detection performance. Further, distributed algorithms for detection with only partial support set estimation via OMP are developed. In [144], a heuristic algorithm has been proposed for sparse signal detection in a decentralized manner based on partial support set estimation via OMP at individual nodes.

2) Detection in the compressed domain without reconstructing the signals: Let us revisit the detection problem presented
in (20) with a single sensor. When the signal to be detected, $\theta_j$ is known, the optimal detector which minimizes the average probability of error is given by the matched filter. Consider the same problem in the compressed domain via $y = Ax$ as addressed by [55]. When $\theta$ is known, the decision statistic of the matched filter, $\Lambda_c$, in the compressed domain is given by
\[
\Lambda_c = y^T(AA^T)^{-1}A\theta.
\] (34)
This results in the following probability of detection of the NP detector:
\[
P_d^e = Q\left(Q^{-1}(\alpha_0) - \frac{||PA^T\theta||^2}{\sigma^2_v}\right),
\] (35)
where $PA^T = AA^T(AA^T)^{-1}$ and $Q(\cdot)$ denotes the Gaussian $Q$-function. When $A$ is selected to be random and an orthoprojector so that $AA^T = I$, (35) can be approximated by
\[
P_d^e \approx Q\left(Q^{-1}(\alpha_0) - \sqrt{\frac{M}{N}}\text{SNR}\right)
\] (36)
where SNR = $\frac{||\theta||^2}{\sigma^2}$. With uncompressed observations, the matched filter results in the following probability of detection, $P_d^u$:
\[
P_d^u = Q\left(Q^{-1}(\alpha_0) - \sqrt{\text{SNR}}\right).
\] (37)
Thus, the impact of performing known signal detection in the compressed domain appears on the probability of detection via the argument of the $Q$ function. As discussed in [55], when SNR is large, compressive detection is capable of providing similar performance as that of the uncompressed detector. Thus, when the problem is detection (but not exact signal recovery), the CS measurement scheme can still be beneficial even without having the sparsity prior. To compensate for the loss in detection in the compressed domain, the authors in [167] have considered the multiple sensor case considering $\theta$ to be deterministic as well as random. In [165], the performance analysis of detection in a Bayesian framework with unequal probabilities for hypotheses has been presented. In a recent paper [174], [175], the authors have considered the problem of detection with multimodal independent data which analyzes the potential of CS in capturing the second order statistics of uncompressed data to compute decision statistics for detection.

In [169], the authors have considered the sparse signal detection problem in the compressed domain in which a sparsity promoting pdf is imposed on the sparse signal. Without complete signal reconstruction, the detection problem is solved in the compressed domain. In [164], the detection performance of random sparse signals has been derived assuming that the subspace in which the signal is sparse is known. Authors in [170] also have considered the detection of signals lying in a low dimensional known subspace in the presence of unknown noise variance. The authors in [156] have considered the sparse signal detection problem assuming the signal to be detected lies in a known subspace. Performance analysis of sequential detection in the compressed domain has been considered in [163] with a single as well as multiple sensors.

3) Design of measurement matrices for detection: In the CS literature, the properties of the measurement matrices are widely investigated focusing on complete signal reconstruction. However, when the problem is to perform detection, it is interesting to see if the same conditions as required for complete signal reconstruction are necessary for the measurement matrices. Design of measurement matrices for compressive detection so that a given objective function is optimized has been considered in [158], [159], [166]. In [158], measurement matrices are designed so that the worst case SNR and the average minimum SNR are maximized. In [166], the authors have considered the probability of detection of the NP detector as the objective function and shown that the optimal measurement matrices depend on the signal being detected. Detection with the designed measurement matrices in [158], [159], [166] is shown to outperform that with random measurement matrices (as widely considered in the CS literature).

B. Compressive classification

Extension of compressed detection problem in [55] to the classification in a parametric framework has been considered in [176]. The authors have established the relationships with several probabilistic distance measures with compressed as well as uncompressed data. Those distance measures can be used to evaluate how good the compressive classification is. In [174], classification is done in the compressed domain in which the projection matrix is optimized to improve the distinguishability among classes. Nonparametric approaches for classification exploiting CS have been proposed in [178–180].

C. Compressive estimation

In addition to detection and classification, another inference task where CS can be exploited is the estimation problem. An estimation problem with compressed measurements can be formulated in different ways. For example, assume that the signal of interest $x$ is parameterized by some $K$-dimensional parameter vector $\omega = [\omega_1, \ldots, \omega_K]^T$. With a SMV, the compressed observation vector (2) can be expressed as
\[
y = Ax(\omega) + v.
\] (38)
This model suits well for the fundamental problem of frequency estimation for a mixture of sinusoids [181–183]. Another interesting formulation of compressive estimation is to estimate a function of the data based on compressed measurements as discussed in [55] where the complete signal reconstruction is not necessary. In this framework with a SMV, the goal is to estimate a function of $x$, $f(x)$ based on $y$ in (2). In [55], the authors have considered the case where $f(x) = (g, x)$ where $g \in \mathbb{R}^N$. The existing compressive parameter estimation techniques mostly focus on the SMV case. However, extensions to the MMV case is worth investigating so that they are applicable for WSN applications.

D. Localization and tracking

As considered in Section III-B3 spatial sparsity can be exploited to solve the source localization problem. The works
in [81], [124] have used $l_1$ norm minimization to impose sparsity to solve the source localization problem with SMV as well as MMV. In [51], target counting and localization has been considered jointly where a greedy matching pursuit algorithm (GMP) is employed. In [50], the source localization problem has been considered using the same spatial grid model as discussed in Section III-B5 in which pre-processing is used to induce incoherence needed in the CS theory, and post-processing is included to compensate for the spatial concretization caused by the grid assumption. In [123], a CS based source localization approach has been proposed with multiple sensors using RSSI measurements where an OMP based algorithm is used. In [125], [184], distributed and adaptive source localization schemes have been proposed using BCS.

E. Sparsity aware sensor management

The problem of selecting the best subset of sensors that guarantees a certain inference performance is one of the fundamental issues addressed by many authors in WSN research. Finding the optimal solution for this problem is intractable in general and a wide class of sub optimal approaches, that fall into convex optimization, greedy and heuristic approaches, have been proposed in the literature [13], [185]–[187]. Sensor selection problem has gained much attention lately in the context of sparsity aware processing since efficient algorithms can be developed benefitting the inherent sparsity in many WSN applications. In this subsection, we review the sensor management work that exploits sparsity aware techniques. Centralized and distributed algorithms for sparsity aware sensor selection for the problem of estimation with a linear model have been proposed in [188]. In [189], the sensor selection problem for linear estimation has been formulated as a sparsity aware optimization problem considering two types of sensor collaboration; information constrained and energy constrained collaborations. The authors have employed reweighted $l_1$ norm based ADMM and the bisection algorithm to solve the optimization problem. In [190], the sensor selection problem has been formulated as the design of a sparse vector considering a nonlinear estimation framework. In [191], sensor selection for target tracking has been considered in which the selection is performed by designing a sparse gain matrix. A probabilistic sensor management scheme for target tracking has been proposed in [192] where the MAC model with distributed sparse random projections as discussed in IV-B2 is used to compress the spatially sparse data.

VI. CS BASED INFERENCE UNDER PRACTICAL CONSIDERATIONS

As discussed in the Introduction section, another factor that needs to be taken into account while developing CS based techniques for WSNs is the physical layer security. In WSNs, transmissions by distributed nodes may be observed by eavesdroppers. Further, the network may operate under malicious/byzantine attacks. Thus, the secrecy of a detection system against eavesdropping attacks is of utmost importance [194]. In a fundamental sense, eavesdropper can be selfish and malicious, to compromise the secrecy of a given inference network. In the recent past, there has been a significant interest in the research community in addressing eavesdropping attacks on distributed inference networks. However, while there a are few recent works [160], [167], [195]–[198], a detailed study with respect to CS based techniques has not yet been done thus far.

In [195], [196], the performance limits of secrecy of CS based measurement schemes have been analyzed. The amplify-and-forward CS scheme is introduced in [197] as a physical layer secrecy solution for WSNs. The authors have studied the secrecy performance against a passive eavesdropper agent composed of several malicious and coordinated nodes. A
CS encryption framework for resource-constrained WSNs has been proposed in [198]. The authors establish a secure sensing matrix, which is used as a key, by utilizing the intrinsic randomness of the wireless channel. In [167], the authors have considered the distributed compressive detection problem in the presence of eavesdroppers. The authors have designed the optimal system parameters to maximize the detection performance at the fusion center while ensuring perfect secrecy at the eavesdropper. In [166], the problem of designing measurement matrices for compressive detection so that the detection performance of the network is maximized while guaranteeing a certain level of secrecy has been discussed.

Another important practical consideration in WSNs is the measurement quantization. We will discuss this issue in the next section.

VII. SPARSE SIGNAL PROCESSING USING QUANTIZED CS

In CS, compression is achieved via random projections at the sampling stage. However, further compression/quantization of the compressed measurements is desirable in many WSN applications due to inherent resource constraints. Coarse quantization reduces the bandwidth requirements and is well motivated for practical systems due to the easy implementation at the transmitter. To that end, sparse signal processing with quantized CS has attracted much attention in recent research. Consider the compressed observation model as given in (41) or (42). With quantized CS, one has access to \( z = \mathcal{Q}(y) \) instead of \( y \) where \( \mathcal{Q}(\cdot) \) is an entry-wise (scalar) quantizer which maps real valued measurements to a discrete quantized alphabet of size \( Q \). In particular, each element of \( y \) is quantized into \( Q \) levels so that

\[
z_i = \begin{cases} 
0, & \text{if } \tau_0 \leq y_i < \tau_1 \\
1, & \text{if } \tau_1 \leq y_i < \tau_2 \\
\vdots & \\
Q-1, & \text{if } \tau_{Q-1} \leq y_i < \tau_Q 
\end{cases} \quad (41)
\]

for \( i = 1, 2, \ldots, M \), where \( \tau_0, \tau_1, \ldots, \tau_Q \) represent the quantizer thresholds with \( \tau_0 = -\infty \) and \( \tau_Q = \infty \). With this approach, \( \lceil \log_2 Q \rceil \) bits per measurement are required to transmit each element of \( y \). In the special case with a 1-bit quantizer, the measurements are quantized into only two levels such that \( Q = 2 \). One example under this special case is to use only the sign information of the compressive measurements [199]-[203]. More specifically, the 1-bit CS scheme first proposed in [199] with sign measurements is given by

\[
z_i = \begin{cases} 
1, & \text{if } y_i \geq 0 \\
-1, & \text{otherwise} 
\end{cases} \quad (42)
\]

for \( i = 1, 2, \ldots, M \). Equivalently, we may express (42) by

\[
z = \text{sign}(y) \quad (43)
\]

where \( z = [z_1, \ldots, z_M]^T \), and \( \text{sign}(\cdot) \) denotes the element-wise sign operation. Sparse signal processing with 1-bit quantized CS is attractive since 1-bit CS techniques are robust under different kinds of non-linearities applied to measurements and have less sampling complexities at the hardware level because the quantizer takes the form of a comparator [199], [201].

Now, the goal is to perform sparse signal recovery or solve other inference tasks based on \( z \) instead of \( y \). There are several factors that need to be taken into account when evaluating the performance and developing algorithms with quantized CS:

- Quantization introduces nonlinearity. Thus, the algorithms and the recovery guarantees developed for sparse signal processing with real valued CS may not be directly applicable for quantized CS. This provides the motivation to develop quantization schemes and reconstruction algorithms so that the performance with quantized CS is very close to (or even better than) that with the real valued CS.
- Coarse quantization can reduce the ability of signal reconstruction or performing other inference tasks compared to real valued CS. This will provide the impetus to take more measurements to compensate for the loss due to quantization. Thus, investigation of the trade-off between the cost for quantization and the cost for sampling is important.

Over the past few years, there have been several research efforts that aim to consider quantized CS in different contexts. In the following, we first review the work on 1-bit CS.

A. Algorithm development and performance guarantees with 1-bit CS

Most of the early work related to 1-bit CS considers the SMV case. In [199], the authors have introduced the 1-bit CS problem with sign measurements [45] for the noiseless case (i.e., \( y \) is as in (1)). The authors have developed an optimization based algorithm for sparse signal recovery using a variation of the fixed point continuation (FPC) method [204]. In particular, they have considered solving

\[
\min_x ||x||_1 \text{ such that } ZAx \geq 0 \text{ and } ||x||_2 = 1 \quad (44)
\]

where \( Z = \text{diag}(z) \). It is noted that, for an \( N \times 1 \) vector \( x \), \( \text{diag}(x) \) denotes a \( N \times N \) diagonal matrix in which the main diagonal is composed of elements of \( x \). The authors have shown that the recovery performance can be significantly improved with the proposed algorithm compared to employing classical CS algorithms when the measurements are quantized to 1-bit. One problem in (44) is that it requires the solution of a non-convex optimization problem. In [205], the authors have presented a provable optimization algorithm to solve (44). A convex formulation of the signal recovery from 1-bit CS has been presented in [206] which solves for

\[
\min_x ||x||_1 \text{ such that } \text{sign}(Ax) = z \text{ and } ||Ax||_1 = M. \quad (45)
\]

It has been shown in [206] that, (45) can provide an arbitrarily accurate estimation of every \( k \)-sparse signal \( x \) with high probability when \( M = \mathcal{O}(k \log^2 (N/k)) \) 1-bit measurements. According to their results, the required number of CS measurements matches the known results with real valued CS up to the exponent of the logarithm and up to an absolute constant factor. In [207], [208], the log-sum penalty function, which has
the potential to be much more sparsity-encouraging than the $l_1$ norm, has been used for sparse signal recovery with 1-bit CS. They have developed an iterative reweighted algorithm which consists of solving a sequence of convex differentiable minimization problems for sparse signal recovery.

Greedy and iterative algorithms developed for classical CS have been extended to the 1-bit CS case in [202], [203], [209]. Modification of the CoSAMP algorithm for the 1-bit case to produce Matching Sign Pursuit (MSP) has been presented in [202]. In [203], an extension of the IHT algorithm with 1-bit CS, called binary IHT (BIHT) has been developed. The BIHT algorithm to incorporate additional information on the partial support has been considered in [209]. The authors in [210] have presented a Bayesian approach for signal reconstruction with 1-bit CS, and analyzed its typical performance using statistical mechanics. In [211], a Bayesian algorithm based on generalized approximate message passing (GAMP) has been developed.

Recovery guarantees and consistency of the estimator for both Gaussian and sub-Gaussian random measurements have been established in [212] for 1-bit CS using the recently proposed k-support norm [213]. In [214], the sample complexity of vector recovery using 1-bit CS has been discussed. In a recent work presented in [215], the authors have shown that 1-bit measurements allow for exponentially decreasing error with adaptive thresholds. This framework is slightly different from the 1-bit CS model discussed with sign measurements in [199]. More specifically, the $i$-th element of $z$ is given by

$$z_i = \text{sign}(y_i - \nu_i) = \begin{cases} 1, & \text{if } y_i \geq \nu_i \\ -1, & \text{if } y_i < \nu_i \end{cases} \quad (46)$$

for $i = 1, \cdots, M$. This scheme allows the quantizer to be adaptive, so that $\nu_i$ in (46) of the $i$-th entry may depend on the $1$st, 2nd, $\cdots$, $(i-1)$st quantized measurements. Adaptive one-bit quantization has also been considered in [216] where the authors have shown that when the number of one-bit measurements is sufficiently large, the sparse signal can be recovered with a high probability with a bounded error. The error bound is linearly proportional to the $l_2$ norm of the difference between the thresholds and the original unquantized measurements. In another recent work by Knudson et.al. in [217], it has been shown that the norm recovery is possible with sign measurements of the form $\text{sign}(Ax + b)$ for random $A$ and fixed $b$ which is impossible with $\text{sign}(Ax)$.

B. Algorithm development and performance guarantees with quantized CS

The authors in [218], [221] have considered the sparse signal/support recovery problem with a given quantizer (including 1-bit CS) with a SMV specifically focusing on support recovery. The effect of quantization has further been analyzed in [220]–[224]. The effects of precision in the measurements have been analyzed in [222] by considering the syndrome decoding algorithm for Reed-Solomon codes when applied in the context of compressed sensing as a reconstruction algorithm for Vandermonde measurement matrices. A universal scalar quantization with exponential decay of the quantization error as a function of the oversampling rate has been considered in [224]. In particular, the author has shown that non-monotonic quantizers achieve exponential error decay in the oversampling rate using consistent reconstruction. However, reconstruction from such a quantization method is not straightforward, and the same author has proposed a practical algorithm for reconstruction using a hierarchical quantization approach in [223]. In [220], the authors have presented a variant of basis pursuit denoising, based on $l_0$ norm rather than using the $l_2$ norm. They have proved that the performance of the proposed algorithm improves with larger $p$. In [221], an adaptation of basis pursuit denoising and subspace sampling has been proposed for dealing with quantized measurements.

Design and analysis of vector quantizers (VQ) with CS measurements have been considered in [225], [226]. In [225], the authors have addressed the design of VQ for block sparse signals using block sparse recovery algorithms. Inspired by Gaussian mixture model (GMM) for block sparse sources, optimal rate allocation has been designed for a GMM-VQ which aims to minimize quantization distortion. In [220], optimum joint source-channel VQ schemes have been developed for CS measurements. Necessary conditions for optimality of vector quantizer encoder-decoder pair with respect to end-to-end MSE have been derived.

C. Distributed and decentralized solutions with quantized CS

While most of the existing works on quantized CS are restricted to the SMV case, there are few recent works that have extended the quantized CS framework to the multiple sensor setting. Several existing CS algorithms developed for real valued CS have been extended to the quantized CS setting in [227]–[229] considering centralized as well as distributed/decentralized implementation. Solving the support recovery problem in a decentralized setting with 1-bit CS has been considered in [228], [229] where the authors have developed several decentralized versions of the BIHT algorithm.

VIII. Summary and Future Research Directions

In this paper, we have provided a comprehensive review on CS based signal processing in WSNs. By discussing the motivating factors, we have identified several challenges that need to be addressed to enable practical implementation of CS based techniques in WSNs. To that end, we have reviewed recent works that focus on developing centralized, distributed and decentralized solutions for data gathering and reconstruction exploiting temporal as well as spatial sparsity. While most of the CS based work in the literature focuses on complete signal reconstruction, in several WSN applications,
complete signal reconstruction is not necessary. We have discussed the benefits of using CS based techniques in solving several inference problems including detection, classification, estimation and tracking. We have further provided a discussion on incorporating practical considerations, such as channel fading, physical layer secrecy constraints and quantization, into the CS framework.

While CS has a relatively rich literature as of now, its implementation under practical considerations as desired by WSN has been discussed only during past few years. Thus, there is still a growing interest in further investigating several issues in detail. In the following, we summarize some of the potential areas worth investigating.

- Distributed/decentralized algorithms with quantized CS: While quantized CS (especially 1-bit CS) is appealing for WSN applications, most of the existing work on quantized CS is restricted to the SMV case. On the other hand, widely discussed distributed and decentralized solutions exploiting CS consider real valued measurements. Thus, further efforts are required to evaluate the benefits of quantized CS in WSNs.

- Developments under practical considerations: As discussed in the paper, one of the issues that needs further attention while making CS based techniques useful in WSNs is to incorporate the practical considerations such as channel fading, and non-Gaussian noise impairments into the CS framework. Most of the existing work with real valued as well as quantized CS that establishes theoretical guarantees needs to be extended to relax these assumptions. Further, the classical CS framework is restricted to linear estimation. When exploiting sparsity aware techniques for distributed estimation/tracking problems in WSNs, one may have to deal with nonlinear measurement models. Thus, there is a need to develop sparsity-aware inference techniques with nonlinear measurement models under resource constraints.

- Dependent data fusion exploiting CS: Dependence is one of the common characteristics exhibited in multiple sensor data. While there exist several recent works that exploit dependence in the Bayesian CS framework under restricted assumptions, CS based dependent data fusion especially in the presence of non-Gaussianity and the spatio-temporal dependence is not well understood. Thus, exploitation of higher order dependence and structured properties of high dimensional data in CS based fusion is worth investigating.

REFERENCES

[1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “Wireless sensor networks: a survey,” Computer Networks, vol. 38, no. 4, pp. 393–422, 2002.

[2] D. Puccinelli and M. Haenggi, “Wireless sensor networks: Applications and challenges of ubiquitous sensing,” IEEE Circuits Syst. Mag., vol. 5, no. 3, pp. 19–31, 2005.

[3] J. Yick, B. Muhukjee, and D. Ghosal, “Wireless sensor network survey,” Computer Networks, vol. 52, no. 12, pp. 2292–2303, 2008.

[4] L. Mainetti, L. Patrono, and A. Vilei, “Evolution of wireless sensor networks towards the internet of things: A survey,” in 19th International Conference on Software, Telecommunications and Computer Networks (SoftCOM), 2011, pp. 1–6.

[5] I. A. Stankovic, A. D. Wood, and T. He, “Realistic applications for wireless sensor networks,” Theoretical Aspects of Distributed Computing in Sensor Networks Springer, pp. 853–863, 2011.

[6] M. F. Othman and K. Shazali, “Wireless sensor network applications: A study in environment monitoring system,” Engineering Procedia, vol. 41, pp. 1204–1210, 2012.

[7] P. Rawat, K. D. Singh, H. Chauachi, and J. M. Bonnin, “Wireless sensor networks: a survey on recent developments and potential synergies,” Journal of supercomputing, vol. 68, no. 1, pp. 1–48, 2014.

[8] B. Rashid and M. H. Rehmani, “Applications of wireless sensor networks for urban areas: A survey,” Journal of Network and Computer Applications, vol. 60, pp. 192–219, 2016.

[9] J.-F. Chamberland and V. V. Veeravalli, “Wireless sensors in distributed detection applications,” IEEE Signal Process. Mag., vol. 24, no. 3, pp. 16–25, May 2007.

[10] W. Li and H. Dai, “Distributed detection in wireless sensor networks using a multiple access channel,” IEEE Trans. Signal Process., vol. 55, no. 3, pp. 822–833, March 2007.

[11] T. Wimalajeewa and S. K. Jayaweera, “Optimal power scheduling for correlated data fusion in wireless sensor networks via constrained PSO,” IEEE Trans. on Wireless commun., vol. 7, no. 9, pp. 3608–3618, Sept. 2008.

[12] M. Duarte and Y.-H. Hu, “Vehicle classification in distributed sensor networks,” Journal of Parallel and Distributed Computing, vol. 64, no. 7, pp. 826–838, 2004.

[13] L. Zuo, R. Niu, and P. K. Varshney, “Posterior CRLB based sensor selection for target tracking in sensor networks,” in Proc. Acoust., Speech, Signal Processing (ICASSP), vol. 2, Honolulu, Hawaii, USA, April 2007, pp. II–1041–II–1044.

[14] O. Ozbek, R. Niu, and P. K. Varshney, “Tracking in wireless sensor networks using particle filtering: physical layer considerations,” IEEE Trans. Signal Process., vol. 57, no. 5, pp. 1987–1999, May 2009.

[15] R. Olfati-Saber, “Distributed tracking for mobile sensor networks with information-driven mobility,” in Proc. of American Control Conf., New York City, NY, July 2007, pp. 4606–4612.

[16] Y. Zou and K. Chakrabarty, “Distributed mobility management for target tracking in mobile sensor networks,” IEEE. Trans. on Mobile Computing, vol. 6, no. 8, pp. 872–887, Aug. 2007.

[17] P. M. Djurić, M. Venu, and M. F. Bugallo, “Target tracking by particle filtering in binary sensor networks,” IEEE Trans. Signal Process., vol. 56, no. 6, pp. 2229–2238, June 2008.

[18] A. Flemmings, P. Ferrari, D. Marioli, E. Sisinini, and A. Taroni, “Wireless and sensor networks for industrial applications,” Microwaves Journal, vol. 40, pp. 1322–1336, Sept. 2009.

[19] V. C. Gungor and G. P. Hancke, “Industrial wireless sensor networks: Challenges, design principles, and technical approaches,” IEEE Trans. Ind. Electron., vol. 56, no. 10, pp. 4258–4265, Oct. 2009.

[20] H. Alemdar and C. Ersoy, “Wireless sensor networks for healthcare: a survey,” Computer Networks, vol. 54, no. 15, pp. 2688–2710, 2010.

[21] N. A. Pantazis, S. A. Nikolaidakis, and D. D. Vergados, “Energy-efficient routing protocols in wireless sensor networks: A survey,” IEEE Commun. Surveys Tuts., vol. 15, no. 2, pp. 551–591, 2013.

[22] Y. C. Eldar and G. Kutyniok, Compressed Sensing: Theory and Applications, Cambridge University Press, 2012.

[23] E. Candès and T. Tao, “Near-optimal signal recovery from random projections: Universal encoding strategies?” IEEE Trans. Inf. Theory, vol. 52, no. 12, pp. 5406 – 5425, Dec. 2006.

[24] E. Candès and M. B. Wakin, “An introduction to compressive sampling,” IEEE Signal Process. Mag., pp. 21–30, Mar. 2008.

[25] D. Donoho, “Compressed sensing,” IEEE Trans. Inf. Theory, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[26] E. Candès, J. Romberg, and T. Tao, “Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information,” IEEE Trans. Inf. Theory, vol. 52, no. 2, pp. 489 – 509, Feb. 2006.

[27] M. F. Duarte, M. A. Davenport, D. Takbar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, “Single-pixel imaging via compressive sampling,” IEEE Signal Process. Mag., vol. 25, no. 2, pp. 83–91, Mar. 2008.

[28] H. Jiang, S. Zhao, Z. Shen, W. Deng, P. A. Wilford, and R. Haimi-Cohen, “Surveillance video analysis using compressive sensing with low latency,” Bell Labs Technical Journal, vol. 18, no. 4, pp. 63–74, 2014.

[29] C. Berger, Z.Wang, J. Huang, and S. Zhou, “Application of compressive sensing to sparse channel estimation,” IEEE Signal Process. Mag., vol. 48, no. 11, pp. 164–174, Nov. 2010.
[223] P. T. Boufounos, “Hierarchical distributed scalar quantization,” in Proc. 9th Int. Conf. Sampling Theory Appl. (SampTA), 2011, pp. 2–6.

[224] ——, “Universal rate-efficient scalar quantization,” IEEE Trans. Inf. Theory, vol. 58, no. 3, pp. 1861–1872, Mar. 2012.

[225] A. Shirazinia, S. Chatterjee, and M. Skoglund, “Performance bounds for vector quantized compressive sensing,” in Int. Symp. Inf. Theory and App. (ISITA), Oct. 2012, pp. 289–293.

[226] ——, “Joint source-channel vector quantization for compressed sensing,” IEEE Trans. Signal Process., vol. 62, no. 14, pp. 3667–3681, July 2014.

[227] V. Gupta, B. Kailkhura, T. Wimalajeewa, S. Liu, and P. K. Varshney, “Joint sparsity pattern recovery with 1-bit compressive sensing in sensor networks,” in 49th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, Nov. 2015, pp. 1472–1476.

[228] S. Kafle, V. Gupta, B. Kailkhura, T. Wimalajeewa, and P. K. Varshney, “Joint sparsity pattern recovery with 1-bit compressive sensing in distributed sensor networks,” IEEE Trans. on Signal and Inf. Process. over Netw., submitted 2017.

[229] S. Kafle, B. Kailkhura, T. Wimalajeewa, and P. K. Varshney, “Decentralized joint sparsity pattern recovery using 1-bit compressive sensing,” in IEEE Conference on Global Signal and Information Processing (GlobalSIP), 2016, pp. 1354–1358.