IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 14, 2017, accepted April 13, 2017, date of publication April 26, 2017, date of current version June 7, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2696745

Unbalanced Expander Based Compressive Data Gathering in Clustered Wireless Sensor Networks

XIANGLING LI[1], (Student Member, IEEE), XIAOFENG TAO[1], (Senior Member, IEEE), AND GUOQIANG MAO[2], (Senior Member, IEEE)

[1]National Engineering Laboratory for Mobile Network Technologies, Beijing University of Posts and Telecommunications, Beijing 100876, China
[2]School of Computing and Communications, University of Technology Sydney and National ICT Australia, Sydney, NSW 1466, Australia

Corresponding author: Xiaofeng Tao (taoxf@bupt.edu.cn)

This work was supported in part by the Natural Science Foundation of China under Grant 61325006 and Grant 61461136002, in part by the Beijing Training Project for the Leading Talents in S & T under Grant Z141101001514026, and in part by the 111 Project of China under Grant B16006.

ABSTRACT Conventional compressive sensing-based data gathering (CS-DG) algorithms require a large number of sensors for each compressive sensing measurement, thereby resulting in high energy consumption in clustered wireless sensor networks (WSNs). To solve this problem, we propose a novel energy-efficient CS-DG algorithm, which exploits the better reconstruction accuracy of the adjacency matrix of an unbalanced expander graph. In the proposed CS-DG algorithm, each measurement is the sum of a few sensory data, which are jointly determined by random sampling and random walks. Through theoretical analysis, we prove that the constructed $M \times N$ sparse binary sensing matrix is the adjacency matrix of a (k, ε) unbalanced expander graph when $M = O(k \log N/k)$ and $t = O(N_c/(kq))$ for WSNs with N_c clusters, where $0 \leq q \leq 1$ and $N_c > k$. Simulation results show our proposed CS-DG has better performance than existing algorithms in terms of reconstruction accuracy and energy consumption. When hybrid energy-efficient distributed clustering algorithm is used, to achieve the same reconstruction accuracy, our proposed CS-DG can save energy by at least 27.8%.

INDEX TERMS Compressive sensing, data gathering, unbalanced expander graph, sparse binary matrix, wireless sensor networks.

I. INTRODUCTION

Reducing energy consumption during data gathering is important in wireless sensor networks (WSNs). To achieve this objective, energy-efficient data gathering (DG) has been widely studied in the past several years, by which the sensory data can be transmitted to the sink. Compressive sensing (CS) [1] is an important technique which has been employed in DG. It can decrease the energy consumption by reducing the number of redundant transmissions and the amount of transmitted data. CS utilizes a random matrix (named as the sensing matrix) to map the high-dimensional sensory data into the low-dimensional compressed data (named as measurements). The measurements are the weighted sums of the sensory data of sensors collected along energy-efficient paths to the sink. An accurate reconstruction of the sensory data can be obtained at the sink as long as the sensory data are sparse or compressible in some transform domain.

Various approaches were applied into the compressive sensing based data gathering (CS-DG) to save energy, such as the improvement of reconstruction accuracy [2], [3], the employment of cooperative transmission [4], [5] and the construction of the sensing matrix [8]–[14], [17]–[22]. The sensing matrix is closely related to the reconstruction accuracy and the routes to obtain the measurements. The non-zeros in the sensing matrix indicate which sensors should be on the routes and whose sensory data are added into the measurements. In this paper, to reduce the network energy consumption and improve the CS reconstruction accuracy, we develop the CS-DG algorithm based on the construction of the sensing matrix.

The theoretical analysis results in [6] and [7] show that either Gaussian random matrix or Bernoulli random matrix can be used as a good candidate for the dense sensing matrix. In [8] and [9], the authors proposed the compressive data gathering schemes by jointly utilising the dense sensing matrix and the traditional routing tree structure. In [8], the CS coding was employed by all sensors in the network. Each sensor delivered the same number of measurements to the

sink which was large enough for an accurate reconstruction at the sink. To further save energy for sensors far away from the sink, the authors in [9] proposed the hybrid CS aggregation, in which the CS coding was conducted only when the number of the received sensory data was larger than the number of measurements. Otherwise, the original sensory data would be routed along the paths to the sink.

Compared with the dense sensing matrix, the sparse sensing matrix requires a smaller number of sensors to obtain the measurements and generates a smaller number of data transmissions between sensors. Wang *et al.* [10], Sartipi and Fletcher [11], Chou *et al.* [12], Ebrahimi and Assi [13], [14] focused on the CS-DG based on the sparse sensing matrix to save energy.

Wang *et al.* [10] proposed a novel distributed algorithm based on a sparse random sensing matrix, and stated that the sparse random sensing matrix can perform as well as a dense one, if designed correctly. In [12], the sparse sensing matrix and the energy-efficient routing paths were jointly optimized utilising the Bayesian compressive sensing algorithm [15]. Each row vector of the sparse sensing matrix was chosen iteratively with the aim of maximizing the information gain per energy expenditure. In [13], a sparse sensing matrix, in which each row vector had the same number of non-zeros and none of the column vectors had all zero entries, was applied. To save energy and distribute the energy load, a number of minimum spanning trees were built, each rooted at a randomly chosen encoding node, based on which the sensory data were separatively collected and compressed according to the relative row of the sparse sensing matrix.

The above works were proposed for WSNs adopting tree-based routing algorithms (tree-type WSNs). It is shown in [16] that clustering has better performance than the tree-based routing algorithms in terms of energy consumption and traffic load balancing. So far, in clustered WSNs, there exists a few works on the CS-DG based on the construction of the dense or sparse sensing matrix [17]–[22].

Xie and Jia [17] proposed a clustering method that used hybrid CS with a dense sensing matrix. Within a cluster, the sensors transmitted the original sensory data to the cluster head (CH). The CHs obtained the measurements by using CS coding and then sent the measurements to sink. The routing trees of CHs are similar to the ones in [8]. Each CH also transmitted the same number of measurements. The theoretical analysis results in [18] and [19] show that block diagonal matrix (BDM) can be used as a good candidate for the sparse sensing matrix. The CS-DG algorithms based on BDM were proposed in [20] and [21]. In [21], the sparse one considered in [13] was employed as the sub-matrix of BDM, and the energy-efficient routing trees were formed to gather the sensory data in each cluster. In [22], the combination of clusters was used in the recovery process, i.e., each measurement was obtained with the uniformly selected w clusters. The sensing matrix can be a BDM or a full Gaussian matrix by separatively setting $w = 1$ or $w = N_c$, where N_c denotes the number of clusters.

It is shown in [23] that if a sparse binary sensing matrix is the adjacency matrix of an unbalanced expander graph, then linear programming (LP) decoding procedure can be used for reconstructing sparse approximations. In [25] and [26], the sparse binary sensing matrix, which could be the adjacency matrix of an unbalanced expander graph, was constructed with a flexible data gathering algorithm based on random walks[1] in the tree-type WSNs. In this paper, we extend the works for the tree-type WSNs into the ones for clustered WSNs leveraging on the significant advantage of unbalanced expander graph in terms of the reconstruction accuracy. We propose a novel energy-efficient CS-DG algorithm based on random sampling and random walks and prove that the constructed sparse binary sensing matrix is the adjacency matrix of an unbalanced expander graph under certain conditions.

The main contributions of this paper are summarized as follows:

1) A novel energy-efficient CS-DG algorithm is proposed for clustered WSNs, in which random sampling and random walks are jointly applied to construct a sparse binary sensing matrix. During the data gathering process, the CS coding is performed on the sensory data of partial sensors. Random walks are implemented to visit some clusters and the partial sensors are randomly sampled from the visited clusters under a certain probability.
2) The theoretical analysis shows that the sparse binary sensing matrix constructed by our proposed CS-DG is the adjacency matrix of an unbalanced expander graph under certain conditions. The conditions are closely related to such parameters as the length of random walk and the probability of random sampling, the number of measurements, the number of clusters and the number of sensors.
3) The simulation results show that the reconstruction accuracy can be improved by increasing the length of walks and appropriately selecting the probability of random sampling. Our proposed CS-DG has better performance than existing algorithms based on dense random matrix and BDM, in terms of reconstruction accuracy and energy consumption. When the well known Hybrid Energy-Efficient Distributed (HEED) [27] clustering algorithm is used, to achieve the same reconstruction accuracy, our proposed CS-DG can save energy by at least 27.8% for a network with 1000 sensors organized into 160 clusters.

A. OUTLINE OF THE PAPER

The rest of this paper is organized as follows. Section II introduces the basic concepts of CS theory and unbalanced expander graph. In Section III, the CS-DG with random

[1] A random walk of length t is started at any node which is selected uniformly at random from nodes in the network. The next hop of each random walk is selected randomly from the set of its neighbors. After $(t-1)$ hops, the random walk is stopped.

TABLE 1. Main notations for Section III and Appendix.

Notation	Meaning	Notation	Meaning
M	Number of measurements in the network	S	Subset of sensors, $S \subset V$ and $\|S\| \leq k$
N	Number of sensors in the network	$N(S)$	Set of nodes in Y that are connected to the nodes in S
N_c	Number of clusters in the network	$E(S)$	Set of links connecting from S to the nodes in Y
k	Sparsity of the sensors in the network	n_i	Number of sensors in the i-th cluster
t	Length of random walks	q	Probability of random sampling
$\mathbf{L}$	Laplacian Matrix of the graph $G=(V,E)$	l	Number of clusters containing the sensors in S
$\mathbf{B}$	Sensing matrix	$\mathbf{B}_S$	Submatrix of $\mathbf{B}$ drawn from the corresponding columns of S
$\boldsymbol{b}_i^m$	Vector where $b_{i,g}^m$ indicates whether $x_{i,g}$ are sampled with probability q when the m-th random walk visits the i-th cluster	$\boldsymbol{x}_i$	Vector of sensory data in the i-th cluster, where $x_{i,g}$ are the g-th sensory data
X	Set of sensory data in the network	$\boldsymbol{x}$	Vector of sensory data in the network
Y	Set of measurements in the network	$\boldsymbol{y}$	Vector of measurements, where y_m is the m-th entry
E, E'	Direct communication links between two sensors and between two CHs, respectively	$\mathcal{A}_t(v')$	Event that the walk visits v' by time t when $t \geq T$, where T denotes the mixing time of a random walk
V, V'	Set of sensors and CHs in the network, respectively	V_i	Set of sensors in the i-th cluster, where $v_{i,g}$ is the g-th sensor
H	Total number of random walks whose measurements are the sums of at least one sensory data in S	H_i	Variable indicating whether the measurement obtained by the i-th random walk is the sum of at least one sensory data in S
Z	Total times of sampling the sensory data from the set S via M independent random walks	Z_i	Total times of sampling the sensory data from the i-th subset S_j via M independent random walks
r,r_c	Transmission range for the sensors and the CHs, respectively	$E(H)$	Expect value of H
$\Gamma(\cdot)$	Set of neighbors	$E(Z)$	Expect value of Z
$\tau(\cdot)$	Node degree	$\Pr(\cdot)$	Probability of the event
Ω_m	Set of cluster IDs where the CHs are visited by the m-th walk	$d(v,u)$	Euclidean distance between v and u, where $v,u \in V$

sampling and random walks is described, and then the constructed sparse binary sensing matrix and its connection with unbalanced expander graph are analyzed. The proofs are included in Section IV. Section V reports our experiment and simulation results. In Section VI, the conclusions are drawn.

B. NOTATION

Throughout this paper, normal letters are used for scalars. The finite-dimensional vectors and matrices are indicated with lowercase and uppercase respectively by boldface. $[\cdot]^T$ indicates the transpose operator and $[\cdot]^\dagger$ is the Moore-Penrose pseudoinverse. For a set, $|\cdot|$ denotes the number of elements in the set. In the vector domain, the concept of ℓ_p-norm is defined as $\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$, where $|x_i|$ is the absolute value of x_i. $\mathbb{R}^n$ means the n-dimensional real coordinate space. The main notations used in Section III and Appendix are listed in Table 1.

II. PRELIMINARIES

A. CS BASICS

In this section, we review some fundamental principles of CS. The theory of CS states that, under certain conditions, instead of directly obtaining the signal $\boldsymbol{x} \in \mathbb{R}^N$, we collect the measurements $\boldsymbol{y} \in \mathbb{R}^M$ by CS coding, and then reconstruct the signal with the measurements, where $M \ll N$. The measurements are the weighted sums of the signal. The matrix form of CS coding can be represented as $\boldsymbol{y} = \mathbf{B}\boldsymbol{x}$, where $\mathbf{B}$ denotes the sensing matrix.

In order to accurately reconstruct the signal using the collected fewer measurements, the signal should be sparse or compressible (approximate sparse). If the number of nonzeros in $\boldsymbol{x}$ is smaller than k ($k \ll N$), the signal is k-sparse. If the number of large coefficients in $\boldsymbol{x}$ is than k and the remaining coefficients are small, the signal is compressible.

The other key factor in CS is the sensing matrix $\mathbf{B}$. The Restricted Isometry Property (RIP) and the mutual coherence are two features that must be satisfied by the sensing matrix. Both are used extensively in formulating the performance guarantees of the CS reconstruction performance. For the definition of $\text{RIP}_{p,k,\delta}$ [23], it is said that $\mathbf{B}$ satisfies the $\text{RIP}_{p,k,\delta}$ condition with parameter δ if, for any k-sparse vector $\boldsymbol{x}$, we have

$$\|\boldsymbol{x}\|_p \leq \|\mathbf{B}\boldsymbol{x}\|_p \leq (1+\delta)\|\boldsymbol{x}\|_p . \tag{1}$$

The signal could be sparse or compressible in some transform domain. That is, if $\mathbf{D}$ is a transform matrix and $\boldsymbol{\theta} = \mathbf{D}^\dagger \boldsymbol{x}$ is sparse, the measurements can be represented as $\boldsymbol{y} = \mathbf{B}\boldsymbol{x} = \mathbf{B}\mathbf{D}\boldsymbol{\theta}$. If $\mathbf{B}$ is a random sensing matrix satisfying the RIP in (1), then in many cases so is the product matrix $\boldsymbol{\Phi} = \mathbf{B}\mathbf{D}$ [24]. Consequently, CS does not rely on assumptions about the domain in which the signal is compressible.

The mutual coherence is defined as

$$\mu(\boldsymbol{\Phi}) = \max_{i \neq j, 1 \leq i,j \leq N} \frac{\left|\phi_i^T \phi_j\right|}{\|\phi_i\|_2 \left\|\phi_j\right\|_2}, \tag{2}$$

where $\phi_i \in \mathbb{R}^M$ is the column vector in $\boldsymbol{\Phi}$ [28]. Both Gaussian random matrix and Bernoulli random matrix satisfy RIP and have a large incoherence with any transform basis, so either of them can be used as a good candidate for the sensing matrix $\mathbf{B}$ [6], [7].

Given a sufficient number of measurements, the sparse or compressive signal can be reconstructed by using the ℓ_1-minimization decoding algorithms, such as Dantzig selector (DS) [29] and basis pursuit denoising (BPDN) [30]. We briefly review the ℓ_1-minimization decoding algorithms. DS finds a reconstructed signal $\hat{\boldsymbol{x}}$ of $\boldsymbol{x}$ to the ℓ_1-minimization problem

$$\min_{\boldsymbol{x},\boldsymbol{r}} \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \|\mathbf{B}^T\boldsymbol{r}\|_\infty < \eta, \boldsymbol{r} = \boldsymbol{y} - \mathbf{B}\boldsymbol{x}. \tag{3}$$

BPDN can recover $\boldsymbol{x}$ from the measurements $\boldsymbol{y}$ by solving

$$\min_{\boldsymbol{x},\boldsymbol{r}} \gamma\|\boldsymbol{x}\|_1 + \frac{1}{2}\|\boldsymbol{r}\|_2^2 \quad \text{subject to} \quad \boldsymbol{r} = \boldsymbol{y} - \mathbf{B}\boldsymbol{x}. \tag{4}$$

η and γ are both the user-selected parameters related to the noise power. It has been shown that the above ℓ_1-minimization problem can be solved with the LP techniques [1].

B. EXPANDER GRAPH

In this section, the definition of an unbalanced expander graph is given and the relationships between an unbalanced expander graph and CS are described in details. According to the descriptions in this section, we show that the adjacency matrix of an unbalanced expander can provide another convenient way to obtain the sensing matrix.

1) DEFINITION (UNBALANCED EXPANDER)

A (k, ϵ) unbalanced expander is a bipartite graph $G = (X, Y, E)$ with $|X| = N$ left nodes, $|Y| = M$ right nodes and regular left degree τ (i.e., each node in X is connected to τ nodes in Y) such that for any set $S \subset X$ with $|S| \leq k$, the set of neighbors $N(S)$ of S (i.e., the nodes in Y that are connected to nodes in S) has size $|N(S)| \geq (1-\epsilon)\tau\,|S|$, for all $0 < \epsilon < 1$.

$|E(S)| = \tau|S|$ is the number of links that go from S to the nodes in Y. E denotes the set of edges where an edge connects a node in X with a node in Y. No edges exist between two nodes in X or two nodes in Y. The structure of the graph G can be equivalently represented by the $M \times N$ adjacency matrix $\mathbf{B}$, where the entry $b_{i,j} = 1$ if the edge exists and $b_{i,j} = 0$ otherwise. It is shown in [23] that $\mathbf{B}$ satisfies $\text{RIP}_{p,k,\delta}$ property for $1 \leq p \leq 1 + O(1)/\log N$ if the matrix $\mathbf{B}$ is the adjacency matrix of an unbalanced expander graph. It also showed that a $\text{RIP}_{p,k,\delta}$ matrix and the adjacency matrix of an unbalanced expander are essentially equivalent. Thus, the adjacency matrix of an unbalanced expander provides another convenient way to obtain the sensing matrix.

The analytical results in [25] show that if $\mathbf{B}$ is the adjacency matrix of an irregular unbalanced expander graph (the degree τ of the nodes on the left-hand side does not need to be constant), then LP decoding procedure can be used for recovering sparse approximations. Let $\tau_{\min}$ and $\tau_{\max}$ denote the minimum and maximum of the left degree (i.e. the degrees of nodes in X), respectively. If $\mathbf{B}$ is the adjacency matrix of a $(2k, \epsilon)$ irregular unbalanced expander, the viability of using the ℓ_1-minimization decoding algorithms for reconstruction is shown below.

Theorem 1([25]): Given $\boldsymbol{y} = \mathbf{B}\boldsymbol{x}$, we can recover $\hat{\boldsymbol{x}}$ from $\boldsymbol{y}$ using the ℓ_1-minimization decoding algorithms such that

$$\left\|\boldsymbol{x} - \hat{\boldsymbol{x}}\right\|_1 \leq c\,(\epsilon)\,\|\boldsymbol{x} - \boldsymbol{x}_S\|_1 \tag{5}$$

as long as $\tau_{\min} > 6\epsilon\tau_{\max}$, where $c(\epsilon) = \frac{2\tau_{\min} - 4\epsilon\tau_{\max}}{\tau_{\min} - 6\epsilon\tau_{\max}}$ and S is the set of k largest coefficients of $\boldsymbol{x}$.

In this paper, we utilize the adjacency matrix of an unbalanced expander as a new feature to characterize the sensing matrix. We design a novel CS-DG and prove that the constructed sparse binary sensing matrix with the novel CS-DG is the adjacency matrix of an unbalanced expander.

III. DESIGN

In this section, the network model is given. Then, the reasons that HEED can be used as the clustering algorithm and the eigenvectors of the graph Laplacian can be used to construct an orthonormal basis are analyzed. After that, the proposed CS-DG based on random sampling and random walks is introduced in details and then the constructed sparse binary sensing matrix and its connection with unbalanced expander graph are analyzed.

A. NETWORK MODEL

Assume that there are a large number of sensors uniformly and densely deployed in clustered WSNs and the sensory data are spatially correlated. The sensors are organized into the non-overlapping clusters using the traditional clustering algorithms, such as HEED [27]. In this paper, we develop a routing algorithm based on random walks implemented on a graph consisting of the CHs. The analytical result for HEED [27] shows that the probability that two sensors within each other's cluster radius are both CHs is small, i.e., CHs are even-distributed, where sensors are the members of the same cluster when the distances between the sensors and its CH are smaller than the cluster radius. Thus, we model the network consisting of the CHs as a random geometric graph (RGG) $G(2, N_c, r_c)$,[2] where N_c represents the number of clusters and r_c denotes the transmission range of the CHs used for the intercluster communication. Our analytical result is obtained according to the property of the random walks implemented on a RGG. Set $r_c = \sqrt{c\log N_c/(\pi N_c)}$, where $c > 1$, so that the graph $G(2, N_c, r_c)$ is connected with high probability [31], [32].

HEED periodically selects CHs considering both the sensor residual energy and the average minimum reachability power (AMRP). Assume that each sensor within a cluster can use power control mechanism to adjust the transmission power according to the transmission distance. Set the transmission range of the sensors used for the intracluster

[2] A random geometric graph $G = G(2, N, r_0)$ in two dimensions is defined as follows: Sample N nodes V independently and uniformly at random from a unit square area. The node set $|V| = N$. r_0 denotes the transmission range. For $v, u \in V$, the edge $e = (v, u) \in E$ if and only if the Euclidean distance $d\,(v, u) \leq r_0$.

communication to be equal to the cluster radius. Thus, the cluster members (CMs) can directly communicate with their CH. Before a node starts executing HEED, each sensor selects itself as a tentative CH with the probability, expressed as $CH_p = C_p \frac{E_{res}}{E_0}$, where C_p is the initial probability of becoming a tentative CH, E_{res} is the estimated residual energy and E_0 is the initial energy. The AMRP of a tentative CH is defined as the mean square distance from all nodes within its cluster radius to the tentative CH, expressed as $AMRP = \frac{\sum_{i=1}^{Q} d_i^2}{Q}$, where Q is the number of sensors within the cluster radius and d_i denotes the distance between the tentative CH and the i-th sensor within its cluster radius. Each sensor select its CH to be the tentative CH with the smallest AMRP in the set of the tentative CHs. The number of selected CHs by HEED varies according to the cluster radius.

We construct a graph $G = (V, E)$ representing spatial correlations. The node set V contains all the sensors. For $v, u \in V$, the edge $e = (v, u) \in E$ if and only if the distance $d(v, u) \leq r$. If the range r is too large, spatial correlations will be reduced. In this paper, the range r is set to be equal to the cluster radius. Because of the spatial correlations, the sensory data can be compressible in some transform domain. Although the deployment of the sensors is irregular, the sensors, which are geographically close to each other, are more likely to have similar sensory data. The sensory data are smooth on the graph $G = (V, E)$ and can be represented sparsely in a Laplacian eigenvector basis [33]. We construct an orthonormal basis using the eigenvectors of the graph Laplacian (EGL) to approximately sparsify the sensory data. Let N represent the number of sensors in clustered WSNs. Let n_i denote the number of sensors in the i-th cluster and $N = \sum_{i=1}^{N_c} n_i$. Let $V_i = \{v_{i,1}, v_{i,2}, \cdots, v_{i,n_i}\}$ represent the set of sensors in the i-th cluster, where $v_{i,1}$ denotes the CH and $v_{i,g}$ denotes the CMs, $1 \leq i \leq N_c$ and $2 \leq g \leq n_i$. The Laplacian Matrix L of the graph $G = (V, E)$ is defined as follow [34]:

$$L = \begin{cases} -1, & \text{if } (v_{i,g}, v_{j,h}) \in E \\ \tau(v_{i,g}), & \text{if } i = j, g = h \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

where $\tau(v_{i,g})$ is the degree of $v_{i,g}$, for $v_{i,g}, v_{j,h} \in V$ $(1 \leq i, j \leq N_c, 1 \leq g \leq n_i$ and $1 \leq h \leq n_j)$.

B. CS-DG WITH RANDOM SAMPLING AND RANDOM WALKS

The CS-DG with sparse sensing matrix generates the CS measurements which are the weighted sums of the sensory data of a small fraction of sensors. The generated measurements are transmitted to the sink and used to reconstruct all the original sensory data. We propose a novel CS-DG with sparse binary sensing matrix, where random sampling and random walks are jointly employed to determine a small fraction of sensors in clustered WSNs. An illustration of our proposed CS-DG is depicted in Fig. 1, which consists of two layers, i.e., sensor layer and cluster head layer.

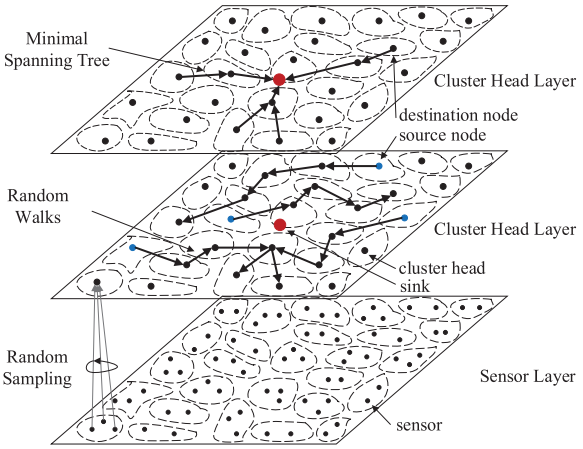

FIGURE 1. Illustration of the CS-DG with random sampling and random walks.

In the cluster head layer, the routing strategies based on random walks and minimal spanning trees are respectively implemented to generate the measurements and to send the measurements to the sink. We consider a network which consists of the CHs, and represent the network with a graph $G_{CH} = (V', E')$. $V' = \{v_{1,1}, v_{2,1}, \cdots, v_{N_c,1}\}$ denotes the set of CHs and E' denotes the set of all direct communication links between two CHs, respectively. Both random walks and minimal spanning trees are constructed in the graph $G_{CH} = (V', E')$. Each sensor can obtain its own location information by GPS or other localization techniques [35], [36], the sensors list within its cluster radius r, and their locations by periodically exchanging beacon messages [37]. Specially, each CH can obtain the CHs list within its transmission range r_c, and their locations.

To generate a measurement with a random walk constructed in the graph $G_{CH} = (V', E')$, a random walk of length t is started from a randomly selected CH, named as the source node, and stopped at the CH, named as the destination node, after $t - 1$ hops. The next hop of each CH is randomly selected from the CHs in its range r_c. In this paper, if a random walk visits a CH, we also say that the random walk visits its cluster. Meanwhile, a measurement will be initialized at the beginning of each walk, be transmitted and be updated on the walk. To initialize the measurement, the source node randomly samples the partial sensors in its cluster with probability q and the initial value of the measurement is the sum of their sensory data. Similarly, to update the measurement, when a CH is visited by the random walk for the first time, the partial sensors in its cluster are randomly sampled with probability q and their sensory data are added into the measurement.

When the random walks are stopped, the obtained measurements are the sums of the sensory data of the sampled sensors in the visited clusters. The destination nodes of random walks transmit the measurements to the sink via the routing strategy based on minimal spanning tree. At sink, the original sensory data can be reconstructed with sufficient

Algorithm 1 CS-DG with random sampling and random walks

1. Sensors in each cluster transmit their sensory data to the CH
2. Initialization
   FOR $m = 1$ **TO** $m = M$
   1) select one source node $v_{i,1} \in V'$
   2) $v_{i,1}$ performs random sampling with probability q
   3) $v_{i,1}$ generates the packet
      $r(m).a1 = t$
      $r(m).a2 = i$
      $r(m).a3 = \boldsymbol{b}_i^m \boldsymbol{x}_i$

   END
3. Synchronously data gathering with random sampling and random walks
   FOR $m = 1$ **TO** $m = M$
   1) $v_{i,1}$ randomly chooses $v_{j,1} \in V'$ from $\Gamma(v_{i,1})$ and transmits its packet to $v_{j,1}$
   2) $v_{j,1}$ performs random sampling with probability q
   3) $v_{j,1}$ updates the achieved packet
      $r(m).a1 = r(m).a1 - 1$
      IF $v_{j,1}$ is visited by the m-th random walk for the first time
      $r(m).a3 = r(m).a3 + \boldsymbol{b}_j^m \boldsymbol{x}_j$
      END
   4) replace $v_{i,1}$ with $v_{j,1}$, then repeat step 1)-3) and update the packet
   5) stop the iteration from 1)-4) when $r(m).a1 = 1$ and obtain $y_m = r(m).a3$
   6) send the packet containing y_m to the sink via the routing strategy based on minimal spanning tree

END

number of measurements by solving the ℓ_1-minimization problem, such as DS and BPDN.

Random sampling with probability q can be performed before or after data transmission to the CH. Both cases are considered in this paper. In Case 1, all the sensory data are directly transmitted to the CHs before the implementation of random walks. Once a CH is visited by one random walk for the first time, the CH will randomly sample the sensory data from the received data with probability q. In Algorithm 1, the proposed CS-DG with Case 1 is described in details. In this case, the energy consumption will not be varied with probability q.

In Case 2, the sampled sensory data are transmitted to the CH during the implementation of random walks. Once a CH is visited by one random walk for the first time, the CH samples partial sensors with probability q and sends the requests to the sampled sensors. The sampled sensors receive the requests, and then directly send their sensory data to the CH. In this case, the sensory data may be transmitted more than once because a CH may be visited by multiple random walks. The energy consumption will be affected by probability q.

Assume that M measurements are large enough for an accurate reconstruction of the sensory data. Because a measurement is obtained by one random walk, M independent random walks are performed to generate M measurements. We take the m-th random walk of length t as an example to explain the details of the random walk constructed in the graph G_{CH}, the selection of the source node, the initialization and update of the packet, where $1 \le m \le M$. Let $\boldsymbol{b}_i^m$ be the row vector indicating whether the sensory data in the i-th cluster are sampled with probability q. The entry $b_{i,g}^m \in \boldsymbol{b}_i^m$ is an i.i.d binary variable indicating whether the sensory data of $v_{i,g} \in V_i$ are sampled with probability q, where $i \le N_c$ and $1 \le g \le n_i$. If the sensory data of $v_{i,g}$ are sampled, $b_{i,g}^m = 1$. Otherwise, $b_{i,g}^m = 0$. Obviously, $b_{i,g}^m$ can be expressed as

$$b_{i,g}^m = \begin{cases} +1, & \text{with probability } q \\ 0, & \text{with probability } 1-q \end{cases} \tag{7}$$

1) A random walk constructed in the graph $G_{CH} = (V', E')$. A random walk of length t is started at a source node $v_{i,1}$, $1 \le i \le N_c$. $v_{i,1}$ generates a packet and injects the values into its packet for the initialization. Let $\Gamma(v_{i,1}) = \{v_{j,1} \in V' : (v_{i,1}, v_{j,1}) \in E'\}$ denote the neighbor set of $v_{i,1}$, consisting of the CHs within its transmission range r_c. For one hop, $v_{i,1}$ randomly selects a CH $v_{j,1}$ from its neighbor set $\Gamma(v_{i,1})$, and sends the packet to $v_{j,1}$. After the reception of the packet, $v_{j,1}$ updates the packet and selects a CH from the set $\Gamma(v_{j,1})$ for the next hop. After $(t-1)$ hops, the random walk is stopped. If r_c is large enough to ensure the connectivity of $G_{CH} = (V', E')$, the neighbor sets consisting of the CHs can not be empty. If the neighbor set $\Gamma(v_{i,1})$ is empty, the random walk will be stopped at $v_{i,1}$ and the obtained measurement will be discarded and not be sent to the sink.
2) Selection of the source node. Because M may be larger than N_c, we have $M = M_i N_c + M_r$, where M_r denotes the remainder after division of M by N_c. M_i and M_r are the positive integer, and $0 \le M_r \le N_c - 1$. The selection methods of the source nodes for $M_i N_c$ random walks and M_r random walks are different. When $M_i \ne 0$, each CH is selected as the source node by M_i independent random walks. The source nodes for M_r random walks are randomly selected from the set V'.
3) Initialization of the packet. Before the implementation of random walk, the source node $v_{i,1}$ generates a packet, denoted as $r(m)$. In the packet, $r(m).a1 = t$ denotes the hop counter, $r(m).a2 = i$ denotes the identification (ID) of the cluster whose CH is selected as the source node and $r(m).a3$ denotes the initial value of a measurement. Let the column vector $\boldsymbol{x}_i$ denote all the sensory data in the i-th cluster. The initial value of a measurement is the sum of the sensory data of the

sampled sensors, denoted as

$$r(m).a3 = \boldsymbol{b}_i^m \boldsymbol{x}_i. \tag{8}$$

4) Update of the packet. Once $v_{j,1}$ receives the packet, two values in the packet will be updated. The value of the hop counter is decreased as $r(m).a1 = r(m).a1 - 1$. The update of the packet and the selection of the neighbor are repeated until $r(m).a1 = 1$. The random walk and the update of the measurement should be stopped when $r(m).a1 = 1$. Let the column vector $\boldsymbol{x}_j$ denote all the sensory data in the j-th cluster. The sampled sensory data in the j-th cluster are added into the measurement, expressed as

$$r(m).a3 = r(m).a3 + \boldsymbol{b}_j^m \boldsymbol{x}_j. \tag{9}$$

If $v_{j,1}$ is revisited by the same random walk, the measurement $r(m).a3$ should not be updated. If $v_{j,1}$ is visited by different random walks, $v_{j,1}$ generates $\boldsymbol{b}_j^m$ and selects the neighbor for each random walk, separately.

For a full and accurate reconstruction, the measurements should contain the information of all the sensory data. That is to say, the measurements should be obtained by linear combinations of all the sensory data. The obtained sensing matrix should have no column vectors with all zeros. Thus, all the clusters should be visited by the random walks and all the sensors should be sampled at least once, which can be obtained by increasing the values of parameters, i.e., M, t and q. However, the energy consumption rises as the values of parameters are increased. The values of parameters should be set to obtain the accurate reconstruction and the low energy consumption. If the sensing matrix is the adjacency matrix of an unbalanced expander graph, then LP decoding procedure can be used for recovering sparse approximations. In this paper, the conditions about the parameters, i.e., M, t and q, will be given, under which the obtained sensing matrix is the adjacency matrix of an unbalanced expander graph.

C. THE CONSTRUCTED SPARSE BINARY MATRIX

The measurements obtained by the proposed CS-DG are the sums of the sensory data of a small fraction of sensors. The partial sensors are determined by random sampling and random walks. Only when the CH of the i-th cluster is visited by the m-th random walk and the g-th sensory data $x_{i,g}$ in the i-th cluster are randomly sampled by the CH, $x_{i,g}$ can be added into the measurement y_m. Let Ω_m denote the set of cluster IDs whose CHs are visited by the m-th walk, where $|\Omega_m| \leq t$. The measurement obtained by the m-th walk is

$$\begin{aligned} y_m &= \sum\nolimits_{i\in\Omega_m} \boldsymbol{b}_i^m \boldsymbol{x}_i \\ &= \sum\nolimits_{i\in\Omega_m} \sum\nolimits_{g=1}^{Q} b_{i,g}^m x_{i,g}. \end{aligned} \tag{10}$$

Let $\boldsymbol{x} = \left[\boldsymbol{x}_1^T, \boldsymbol{x}_2^T, \cdots, \boldsymbol{x}_{N_c}^T\right]^T \in \mathbb{R}^N$ denote the sensory data in clustered WSNs. By performing M independent random walks of length t constructed in the graph G_{CH}, the CS measurement vector $\boldsymbol{y} = [y_1, \ldots y_M]^T \in \mathbb{R}^M$ is given by

$$\boldsymbol{y} = \mathbf{B}\boldsymbol{x}. \tag{11}$$

The achieved sensing matrix $\mathbf{B} \in \mathbb{R}^{M\times N}$ consists of the random row vectors, in which the entries satisfy (7). So, the constructed sensing matrix $\mathbf{B}$ is the sparse binary matrix.

For example, $\mathbf{B}$ shown in (12) is generated by random walks with $t = 2$, $M = 6$ and $N_c = 4$, where the row vectors generated by the source nodes are marked in red. Due to $M_i = 1$ and $N_c = 4$, the source nodes for the first four random walks are four CHs, respectively. Due to $M_r = 2$, the last two random walks are started from randomly selected CHs.

$$\mathbf{B} = \begin{bmatrix} \boldsymbol{b}_1^1 & \mathbf{0} & \mathbf{0} & \boldsymbol{b}_4^1 \\ \mathbf{0} & \boldsymbol{b}_2^2 & \boldsymbol{b}_3^2 & \mathbf{0} \\ \boldsymbol{b}_1^3 & \mathbf{0} & \boldsymbol{b}_3^3 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{b}_2^4 & \mathbf{0} & \boldsymbol{b}_4^4 \\ \mathbf{0} & \mathbf{0} & \boldsymbol{b}_3^5 & \boldsymbol{b}_4^5 \\ \boldsymbol{b}_1^6 & \mathbf{0} & \boldsymbol{b}_3^6 & \mathbf{0} \end{bmatrix}, \tag{12}$$

where $\mathbf{0}$ denotes zero vector where each component is zero. Assume that the number of sensors in each cluster is 3 and N=12. The entry $b_{i,g}^m \in \boldsymbol{b}_i^m$ is an i.d.d binary variable satisfying (7), where $1 \leq i \leq 4$, $1 \leq g \leq 3$ and $1 \leq m \leq 6$. The details of (12) are given as follow.

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{13}$$

When the length t and the probability q is large enough, the CH in each cluster can be visited by random walks and more sensors can be sampled. So there will be no column vectors with all zeros. When $t = 1$, the measurements are provided by the randomly selected CHs. When $M \geq N_c$ and $t = 1$, the achieved sensing matrix $\mathbf{B}$ will be similar to BDM. The algorithm in [21] can be seen as a special case of our proposed CS-DG. When $N_c = N$, the tree-type WSNs can be obtained where only one sensor exits in each cluster. When $N_c = N$ and $q = 1$, the sensor must be sampled when it is visited by one random walk for the first time. The algorithm in [25] can be seen as a special case of our proposed CS-DG with $N_c = N$ and $q = 1$.

To reconstruct the original sensory data, the sink needs to know the sensing matrix $\mathbf{B}$, which is determined by random sampling and random walks. If the IDs of sensors selected via random sampling and the clusters IDs of the CHs visited by random walks are injected into the packets, the length of the packet will be increased, leading to additional energy consumption. To solve this problem, we can adopt the following approach. Before the implementation of the CS-DG algorithm, each CH generates two random seeds and sends

them to the sink. One random seed can be used to perform random sampling through a pseudo-random number generator. The other random seed can be used to determine the IDs of clusters visited by random walks. Finally, the sink can know the IDs of sensors selected via random sampling and the clusters IDs of the CHs visited by random walks according to the seeds and the cluster ID of source node in the packet. Then, the sink can obtain the sensing matrix $\mathbf{B}$.

D. CONNECTION BETWEEN UNBALANCED EXPANDER AND SPARSE BINARY MATRIX

According to Theorem 1, if the sensing matrix is the adjacency matrix of an unbalanced expander, the sparse signal can be reconstructed with ℓ_1 minimization decoding algorithm. Thus, we utilize the adjacency matrix of an unbalanced expander as a new feature to characterize the sensing matrix. We prove that the sensing matrix constructed by random sampling and random walks is the adjacency matrix of a (k, ϵ) unbalanced expander. The result is shown below.

Theorem 2: Performing random sampling with probability q and taking $M = O(k \log(N/k))$ independent random walks of length t satisfying

$$t = \begin{cases} O\left(\frac{N_c}{kq}\right), & N_c > k \\ O\left(\frac{1}{\left[1-(1-q)^{k/N_c}\right]}\right), & N_c \le k, \end{cases} \tag{14}$$

the sensing matrix $\mathbf{B}$ constructed from random sampling and random walks will be a (k, ϵ) unbalanced expander with $\varepsilon > 1 - \frac{1-\varepsilon'}{\beta(1+\varepsilon')}$, where $0 \le q \le 1$, $\epsilon' > 0$ and $0 < \beta < 1$.

The method, which proves the expansion property of random matrix according to the definition of unbalanced expander in [25], is adapted to obtain Theorem 2. It showed that the random matrix was an adjacency matrix of a bipartite graph, and then proved that an unbalanced expander was the bipartite graph with $|N(S)| \ge (1-\varepsilon)|E(S)|$.

Firstly, we show the sensing matrix $\mathbf{B}$ is the $M \times N$ adjacency matrix of a bipartite graph. In the bipartite graph, the sensory data form the left node set X with $|X| = N$, and are divided into N_c clusters. The measurements form the right node set Y with $|Y| = M$. Obviously, no edges exist between two nodes in X or two nodes in Y. A node in X connects to a node in Y if and only if a random walk visits a cluster and a sensory data in the visited cluster are sampled, which is different from the work in [25]. Figure 2 illustrates an example, which is a bipartite graph represented by the sensing matrix (13).

For simplicity of the analysis, we assume that any set $S \subset X$ with $|S| = k$. In this paper, we analyze a special case of unbalanced expander, where the nodes in the set S are equally distributed in l clusters, where $1 \le l \le \min\{k, N_c\}$, which is also different from the work in [25]. The set S can be denoted as $S = \{S_1, S_2, \cdots, S_l\}$, where S_j represents the j-th subset contained in a cluster. In Fig.2, the subsets in $S = \{S_1, S_2, S_3\}$ are indicated by the dotted oval.

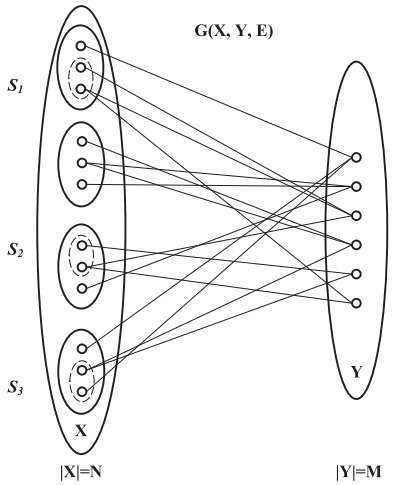

FIGURE 2. The bipartite graph corresponding to the sensing matrix B in (13). The nodes on the lefthand side are divided into 4 clusters which are indicated by the medium sized oval. The subsets in $S = \{S_1, S_2, S_3\}$ are indicated by the dotted oval.

Let $\mathbf{B}_S$ denote the $M \times k$ submatrix of $\mathbf{B}$ drawn from the corresponding columns of S. $|E(S)|$ represents the total number of links between nodes in S and nodes in Y, which also represents the total number of non-zeros in $\mathbf{B}_S$. $|N(S)|$ represents the total number of nodes in Y that are connected to any nodes in S, which also represents the number of non-all-zero rows in $\mathbf{B}_S$. According to the definition of unbalanced expander, we prove that a (k, ε) unbalanced expander can be constructed from random sampling and random walks, which is a bipartite graph with $|N(S)| \ge (1-\varepsilon)|E(S)|$. The proofs of Theorem 2 are included in Appendix.

IV. EXPERIMENTAL RESULTS

In this section, experiments are implemented to validate the performance of our proposed CS-DG, in terms of reconstruction quality and energy consumption. Firstly, the analytical results are verified using the sparse data sets by experiment, where orthogonal matrices (the identity matrix (IM) and EGL) are employed as the transform matrices. Then, the performance of our proposed CS-DG is evaluated with the real data sets. The impacts of different parameters are estimated with different types of data sets and the comparisons are made with other approaches.

The experiments are conducted using MATLAB. For signal reconstruction, we use cvx package to solve ℓ_1 programming [38]. Assume that the data gathering process is divided into rounds. In each round, all the sensory data in the network are compressed and delivered to the sink, and reconstructed at the sink. Each experiment is performed 1000 rounds for the average performance.

A. METRICS

1) RECONSTRUCTION QUALITY

To evaluate the reconstruction quality, we define the normalized reconstruction error as

$$error_x = \frac{\|\hat{\boldsymbol{x}} - \boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2}, \tag{15}$$

where $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$ represents the original sensory data and the reconstructed sensory data, respectively.

2) PROBABILITY OF SUCCESSFUL RECONSTRUCTION

We define the probability of successful reconstruction to measure the reconstruction performance of our algorithm. Assume that the data are reconstructed successfully if $error_x \leq 10^{-3}$. Let n denote the number of rounds that the experiment is executed. n_0 denotes the number of rounds of successful reconstruction. So, the probability of successful reconstruction is defined as n_0/n.

3) ENERGY CONSUMPTION

We calculate energy consumption with the model as follows [39].

$$E_t(L,d) = E_{elec} \times L + \epsilon_{amp} \times L \times d^2 \tag{16}$$

$$E_r(L) = E_{elec} \times L \tag{17}$$

where d denotes the Euclidean distance between two connected nodes. $E_t(L,d)$ represents the energy consumption for transmitting a L-bit packet over distance d. $E_r(L)$ represents the energy consumption for receiving a L-bit packet. E_{elec} is the energy consumption for transmitting or receiving one bit message, and ϵ_{amp} is the transmission amplifier. We set $E_{elec} = 50\,nJ/bit$ and $\epsilon_{amp} = 100\,pJ/bit/m^2$ [39], $L = 160\,bit$ [8]. We consider that the wireless channel obeys a square-law path loss.[3] The energy consumptions of baseband signal processing blocks such as source coding and pulse-shaping are ignored, which are quite small compared to that of the RF circuitry [40].

B. EXPERIMENTS ON SPARSE DATA SETS

In this section, the experiments are performed to verify the analytical results. The k-sparse data are generated, for which the non-zeros (i.e., +1) are randomly distributed in the coefficients. The orthogonal matrices (IM and EGL) are employed as the transform matrices. In clustered WSNs, $N = 1000$ sensors are randomly deployed in a square area with the size $2000 \times 2000\ m^2$, the sink being at the center. The sensors are divided into clusters using HEED algorithms. In each cluster, the transmission range r of the CMs is set to be equal to the cluster radius so that the CMs can directly communicate with the CH. The transmission range r_c for the inter-cluster communication is set to be $r_c = \sqrt{1.5 \log N_c/(\pi N_c)}$ [31], [32].

1) DETERMINATION OF THE PROBABILITY q

$M = 300$ random walks are implemented, where $t = 15$ for IM and $t = 10$ for EGL. The number of selected CHs by HEED varies according to the cluster radius. The average number of clusters is 100 by setting $r = 134.5\ m$. The probability of successful reconstruction over the probability q for random sampling is shown in Fig. 3, where $q \in [0,\ 1.0]$ and $k = 60$. We found that a high probability of successful reconstruction can be obtained by setting q to be the middle value of $[0,\ 1.0]$. When EGL is used as the transform basis, the wide range of q is available for a high probability of successful reconstruction.

[3]The channel path loss exponent κ could usually lie in the range $2-4$ for wireless communications channels, with $\kappa = 2$ corresponding to free space propagation.

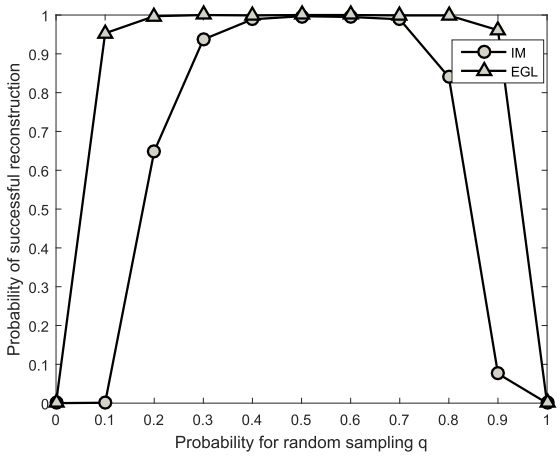

FIGURE 3. The probability of successful reconstruction with the different probability q for random sampling. The identity matrix (IM) and the eigenvectors of the graph Laplacian (EGL) are employed as the transform matrices.

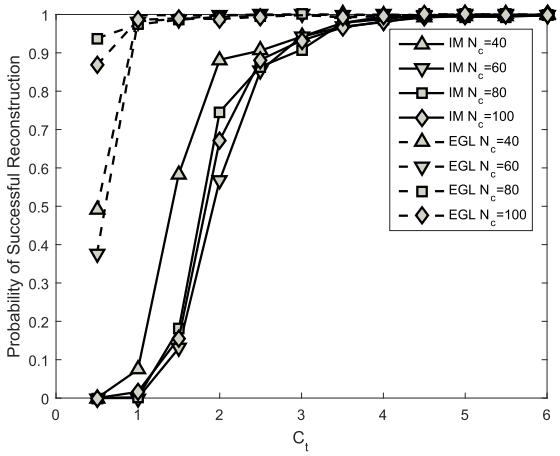

FIGURE 4. The probability of successful reconstruction with different constants C_t for the length of random walks and the number of clusters N_c.

2) RELATIONSHIP BETWEEN THE PROBABILITY OF SUCCESSFUL RECONSTRUCTION AND PARAMETERS

To verify Theorem 2, $M = 1.7k\log(N/k)$ random walks of length t satisfying

$$t = \begin{cases} C_t \frac{N_c}{kq}, & N_c > k \\ C_t \frac{1}{\left[1-(1-q)^{k/N_c}\right]}, & N_c \leq k, \end{cases}$$

are implemented, where $C_t \in [0.5,\ 6]$, $k = 60$ and $q = 0.5$. Figure 4 shows the probability of successful reconstruction over the constant C_t. The average number of clusters is 40, 60, 80, 100 obtained by varying r in the

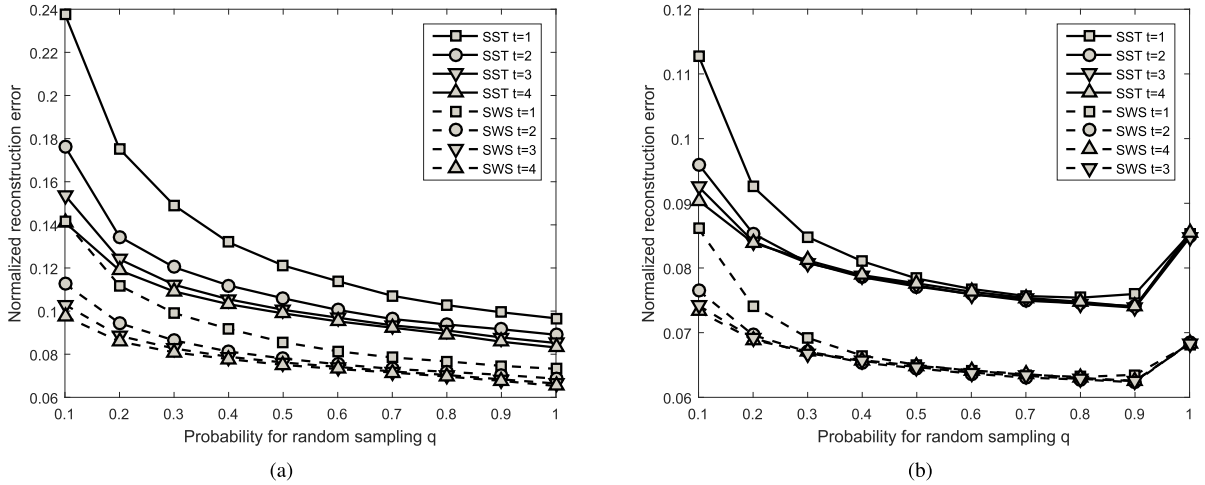

FIGURE 5. The reconstruction accuracy obtained by varying the probability for random sampling q and the length of random walks t (a) $M = 100$, $N_c = 160$. (b) $M = 200$, $N_c = 80$.

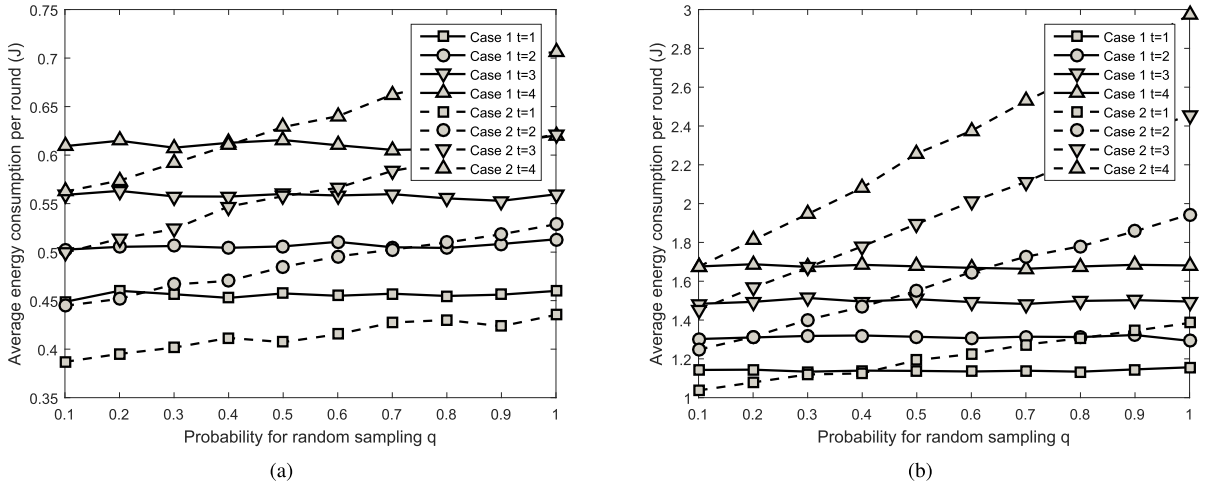

FIGURE 6. The average energy consumption per round obtained by varying the probability for random sampling q and the length of random walks t (a) $M = 100$, $N_c = 160$. (b) $M = 200$, $N_c = 80$.

set 228.5 m, 181 m, 153.5 m, 134.5 m. Both IM and EGL are considered as the transform matrices. We find that the k-sparse data can be exactly reconstructed with a high probability using ℓ_1 minimization decoding algorithm when the parameters are set according to Theorem 2. The best values for C_t should be larger than 2 for EGL and 5.5 for IM, respectively. When EGL is employed as the transform basis, a high probability of successful reconstruction can be obtained with random walks with the short length t.

C. EXPERIMENTS ON REAL DATA SETS

To evaluate the performance of the proposed CS-DG for the compressible signals through simulation, real data sets are used in this paper, which are obtained from a remote sensing system [41] to provide an uninterrupted view of ocean measurements. The real data sets consist of Sea Surface Temperature (SST) in *deg C*, Surface Wind Speeds (SWS) in m/s. The employed data set of size 80×80 are measured in a region centred at (-47.5S, 10.0E), sensed from July. 07, 2012 to Nov. 12, 2016. Assume that $N = 1000$ nodes are randomly deployed in a square area with size 2000×2000 m^2, the sink being at the center. Their sensory data are extracted from the real data set. The EGL is employed as the transform basis.

1) IMPACT OF DIFFERENT PARAMETERS

The performance of two cases of our proposed CS-DG is checked by considering different types of signals (SST and SWS), where $t \in [1, 4]$ and $q \in [0, 1.0]$. The sensors are divided into N_c clusters using the HEED algorithm.

In Fig. 5, the reconstruction quality for SST and SWS is shown by varying t and q, respectively. The reconstruction quality can be improved by increasing t or q. When q is small,

the reconstruction quality can be obviously improved by increasing t. Because a small number of sensors are sampled with small q in each hop of each walk. To obtain the sensing matrix with no column vectors with all zeros, the length of random walks t should be increased to sample a sufficient number of sensors. When q is large, the reconstruction quality cannot be obviously improved if $M \leq N_c$ and $t \geq 3$ or if $M > N_c$ and $t \geq 2$. So, when $M \leq N_c$, the best value for t is 3 and the best value for q is 1.0. When $M > N_c$, the best value for t is 2 and the best value for q is 0.9. Because different random walks start from the same cluster and sample all the sensory data in the cluster by setting $M > N_c$ and $q = 1.0$. The measurements generated by different random walks may be the same. Thus, the reconstruction quality is reduced.

In Fig. 6, the average energy consumption per round for two cases of our proposed CS-DG is shown by varying t and q, respectively. For Case 1, the average energy consumption is not affected by the parameter q, and can be increased when t is larger. For Case 2, the average energy consumption can be more when t or q is larger. In Case 2, the random sampling can be performed only when a random walk visits its CH for the first time. Specially, only the sampled sensory data are transmitted to its CH. Thus, when q is small, a small number of sensory data can be delivered and a low energy consumption is required. However, when M and t are increased, each CH may be visited more than once, causing the energy consumption for Case 2 to be larger than that of Case 1, as shown in Fig.6(b). Thus, when M is small enough, by properly selecting q and t, the energy for Case 2 can be saved as compared to that for Case 1.

2) COMPARISON WITH DIFFERENT ALGORITHMS

The comparisons with other algorithms are implemented, where $M \in [100, 420]$. We name the CS-DG algorithms according to the property of the sensing matrix, i.e., Dense-DG [17], dBDM-DG [20], sBDM-DG [21], wBDM-DG [22], and the non-zeros in the sensing matrices satisfy Gaussian distribution. In [20] and [21], the sub-matrices of the BDM are dense and sparse, respectively. In [22], each measurement is obtained from uniformly and randomly selected w clusters. (In this paper, w is set to be 3 which can obtain high reconstruction quality and save energy.) The sensing matrix is similar to the BDM. Because the sparse binary matrix (SBM) is used as the sensing matrix, our proposed CS-DG is named as SBM-DG.

The data set of SST is employed for the comparison of reconstruction quality. The sensors are divided into $N_c = 160$ clusters. To achieve high reconstruction quality, we set that $q = 0.9$ for $M > N_c$ and $q = 1.0$ for $M \leq N_c$, $t = 2$.

In Fig. 7, the reconstruction quality over the number of measurements M is given. The reconstruction accuracy of our proposed CS-DG is the best of all, which can be improved by increasing M. In Fig. 8, the average energy consumption per round over the number of measurements M is given. When the parameter M is the same, the energy consumption of our proposed CS-DG is slightly larger than that of sBDM-DG,

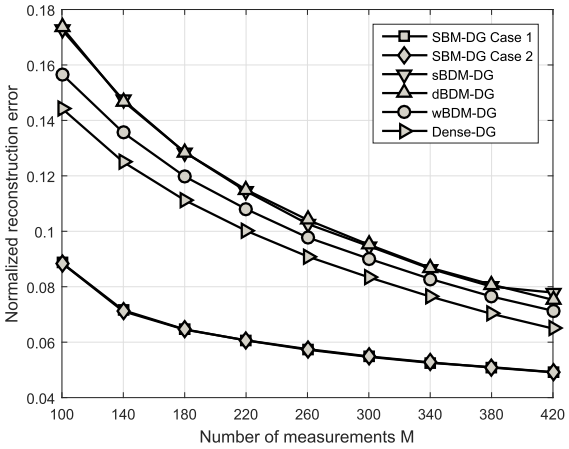

FIGURE 7. The reconstruction accuracy obtained by varying the number of measurements M.

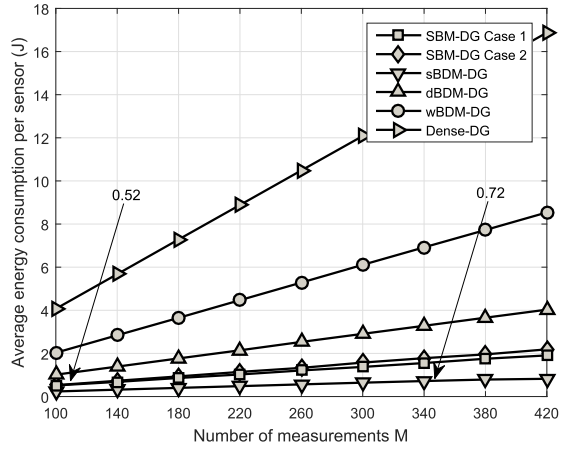

FIGURE 8. The average energy consumption per round obtained by varying the number of measurements M.

lower than that of other three algorithms. If HEED is used as the clustering algorithm, when $M = 100$ for SBM-DG and $M = 340$ for sBDM-DG, $error_x = 0.088$ can be obtained and SBM-DG can save energy by 27.8% as compared to sBDM-DG. Thus, to obtain the same reconstruction accuracy, the energy consumption of our proposed CS-DG is the lowest of all.

V. CONCLUSION

We investigated the compressive sensing based data gathering (CS-DG) in clustered WSNs and proposed a novel CS-DG based on random sampling and random walks to save energy. A sparse binary sensing matrix is constructed, in which the non-zeros are jointly determined by random sampling and random walks. The sensory data can be added into the measurements only when their cluster heads are visited by random walks and the sensors are randomly sampled with a certain probability. Theoretical analysis results show that the constructed sparse binary matrix is the adjacency matrix of an unbalanced expander graph when the parameters are appropriately chosen, i.e., the length of random walk and the

probability of random sampling, the number of measurements and the number of clusters. Experimental results show that our proposed CS-DG has high reconstruction accuracy and low energy consumption as compared to existing algorithms using dense random matrix and block diagonal matrix. When HEED is used as the clustering algorithm, to achieve the same reconstruction accuracy, our proposed CS-DG can save energy by at least 27.8%.

APPENDIX
PROOF OF THEOREM 2

Firstly, the number of non-all-zero rows in $\mathbf{B}_S$, denoted as $N(S)$, is obtained. This is equal to the total number of random walks whose measurements are the sums of at least one sensory data in the set S, denoted as H. Let $H_i (1 \le i \le M)$ be an independent Bernoulli random variable indicating whether the measurement obtained by the i-th random walk is the sum of at least one sensory data in the set S. We have $H = \sum_{i=1}^{M} H_i$. By Chernoff bound, for any arbitrary small constant $\varepsilon' > 0$, we have

$$\Pr\left(H \le \left(1-\varepsilon'\right) E\left(H\right)\right) \le \exp\left(-\frac{\varepsilon'^2}{3} E\left(H\right)\right), \quad (18)$$

where $E(H)$ denotes the expected value of H.

The random walks are performed on the graph $G_{CH} = (V', E')$. Let $W_{u'}$ be the random walk starting at an arbitrary CH $u' \in V'$. Let $\mathcal{A}_t(v')$ be the event that a CH $v' \in V'$ is visited by the walk $W_{u'}$ by time t when t is larger than the mixing time.[4] According to the results in [25], if the graph G_{CH} can be modeled as a RGG, the probability of the event $\mathcal{A}_t(v')$, denoted as $\Pr\left(\mathcal{A}_t\left(v'\right)\right)$, can be expressed as

$$\frac{(1-\mu)\, t}{2\,(\eta+2)\, cN_c} \le \Pr\left(A_t\left(v'\right)\right) \le \frac{(1+\mu)\, t}{cN_c}, \quad (19)$$

where $\eta = \frac{1}{(1-\mu)\kappa}$ and c is a constant.

In the following, we derive the expected value $E(H)$, expressed with $\Pr\left(\mathcal{A}_t\left(v'\right)\right)$, to estimate the probability $\Pr\left(H \le \left(1-\varepsilon'\right) E\left(H\right)\right)$. We have $M = M_i N_c + M_r$, where M_i and M_r are the positive integer, and $0 \le M_r \le N_c - 1$. The selection method of the source nodes for $M_i N_c$ random walks is different from that for M_r random walks. The nodes in the set S are distributed in l clusters. Each CH in l clusters is the source node of M_i random walks and must be visited by M_i random walks. For the random walk starting from the cluster containing the subset S_i, let P_i denote the probability that the measurement is not the sum of any sensory data in the set S. For $M - lM_i$ random walks, l clusters are visited with the probability $\Pr\left(\mathcal{A}_t\left(v'\right)\right)$. Let P_r denote the probability that the measurement is not the sum of any sensory data in the set S for $M - lM_i$ random walks. We have

$$E(H) = M_i \sum_{i=1}^{l} (1-P_i) + (M - lM_i)(1-P_r). \quad (20)$$

[4]The mixing time of a graph is the time taken by a simple random walk on the graph to sample a node according to the steady state distribution of the graph, which means sampling uniformly at random if the graph is regular. If the mixing time is poly-logarithmic in the number of nodes, then we say that the graph is rapid mixing.

For lM_i random walks, the CH in one cluster must be visited as the source of the random walk, and the CHs in the other $l-1$ clusters can be visited by the random walk with probability $\Pr\left(\mathcal{A}_t\left(v'\right)\right)$. Assume that the subset S_i with $|S_i| = k_i$ is contained in one of l cluster. Only when the random walk does not visit the cluster or when the random walk visits the cluster and does not sample any sensory data in the subset S_i according to (7), the measurement obtained by the random walk is not the sum of any sensory data in the subset S_i. Due to the CHs independently visited by the random walk, we have

$$P_i = q_0^{k_i} \prod_{i \ne j} \left[1 - \Pr\left(A_t\left(v'\right)\right) + \Pr\left(A_t\left(v'\right)\right) q_0^{k_j}\right]. \quad (21)$$

where $q_0 = 1 - q$. For $M - lM_i$ random walks, the CHs in l clusters can be independently visited by the random walk with probability $\Pr\left(\mathcal{A}_t\left(v'\right)\right)$. We have

$$P_r = \prod_{i=1}^{l} \left[1 - \Pr\left(A_t\left(v'\right)\right) + \Pr\left(A_t\left(v'\right)\right) q_0^{k_i}\right]. \quad (22)$$

Then, according to (21) and (22), due to $q_0^{k_i} \le 1 - \Pr(A_t(v')) + \Pr(A_t(v'))q_0^{k_i}$, (20) can be rewritten as

$$E(H) \le M(1-P_r). \quad (23)$$

Suppose that each subset is composed of k/l nodes, (23) can be rewritten as

$$E(H) \le M\left\{1 - \left[1 - \Pr\left(A_t\left(v'\right)\right) + \Pr\left(A_t\left(v'\right)\right) q_0^{k/l}\right]^l\right\}$$

Combining (18), (19), (22) and (23), taking $M = C_m k \log(Ne/k)$ and $t = C_1 \frac{N_c}{l[1-(1-q)^{k/l}]}$ with an appropriate constant C_m and C_1, $H \le \left(1-\varepsilon'\right) M(1-P_r)$ with probability at most $O\left((eN/k)^{-k}\right)$, which can be arbitrary small when N is large. In other words, $|N(S)| \ge \left(1-\varepsilon'\right) M(1-P_r)$ with high probability when N is large.

Next, the number of nonzero elements in $\mathbf{B}_S$, i.e., $|E\,(S)\,|$, is obtained. That is the total times of sampling the sensory data from the set S via M independent random walks, denoted as Z. Let Z_i be the total times of sampling the sensory data from the i-th subset S_i via M independent random walks. We have $Z = \sum_{i=1}^{l} Z_i$.

The sampling of one sensory data is effective only when the CH in its cluster is visited by one random walk. One CH must be the source node of M_i random walks and be visited by $M - M_i$ random walks with the probability $\Pr\left(\mathcal{A}_t\left(v'\right)\right)$. Assume that the subset S_i is contained in one of l clusters. When the CH in the cluster is visited by the random walk, the k_i sensory data in the subset S_i are sampled with probability q. So, the expected value of Z_i is

$$E\,(Z_i) = M_i q k_i + (M - M_i) q k_i \Pr\left(\mathcal{A}_t\left(v'\right)\right) \quad (24)$$

and the expected value of Z is

$$\begin{aligned} E\,(Z) &= \sum_{i=1}^{l} \left\{M_i q k_i + (M - M_i) q k_i \Pr\left(\mathcal{A}_t\left(v'\right)\right)\right\} \\ &= M_i q k + (M - M_i) q k \Pr\left(\mathcal{A}_t\left(v'\right)\right) \\ &\ge M q k \Pr\left(\mathcal{A}_t\left(v'\right)\right). \end{aligned} \quad (25)$$

By Chernoff bound, we have

$$\Pr\left(Z \geq \left(1+\varepsilon'\right)E\left(Z\right)\right) \leq \exp\left(-\frac{\varepsilon'^2}{2}E\left(Z\right)\right). \quad (26)$$

We find that $t = C_1 \frac{N_c}{l\left[1-(1-q)^{k/l}\right]}$ is a monotone decreasing function of l, obtained by calculating the second derivative of t. The parameter l is unknown because the nodes in the set S are randomly distributed in l clusters, where $1 \leq l \leq \min\{k, N_c\}$. The expression about the parameter t can be rewritten as

$$t = \begin{cases} C_t \frac{N_c}{kq}, & N_c > k \\ C_t \frac{1}{\left[1-(1-q)^{k/N_c}\right]}, & N_c \leq k, \end{cases} \quad (27)$$

where C_t is a constant.

According to (19), (25), (26) and (27), with $M = O\left(k \log N/k\right)$, $Z \geq \left(1+\varepsilon'\right) Mqk \Pr\left(\mathcal{A}_t\left(v'\right)\right)$ with probability at most $O\left((eN/k)^{-k}\right)$. In other words, $|E\left(S\right)| \leq \left(1+\varepsilon'\right) Mqk \Pr\left(\mathcal{A}_t\left(v'\right)\right)$ with high probability when N is large.

According to the definition of unbalanced expander, if the bipartite graph represented by the sensing matrix $\mathbf{B}$ is an unbalanced expander graph, we should have $|N\left(S\right)| \geq \left(1-\varepsilon\right)|E\left(S\right)|$. Therefore, with high probability, we should have

$$\begin{aligned} |N\left(S\right)| &\geq \left(1-\varepsilon'\right) M\left(1-P_r\right) \\ &\geq \left(1-\varepsilon\right)\left(1+\varepsilon'\right) Mqk \Pr\left(\mathcal{A}_t\left(v'\right)\right) \\ &\geq \left(1-\varepsilon\right)|E\left(S\right)|. \end{aligned} \quad (28)$$

(28) can be further expressed as

$$\begin{aligned} \left(1-\varepsilon'\right)\left(1-P_r\right) &\geq \left(1-\varepsilon\right)\left(1+\varepsilon'\right) qk \Pr\left(\mathcal{A}_t\left(v'\right)\right) \\ \left(1-\varepsilon'\right) &> \left(1-\varepsilon\right)\left(1+\varepsilon'\right) qk \Pr\left(\mathcal{A}_t\left(v'\right)\right) \end{aligned} \quad (29)$$

According to (19), we know that the minimum value of $\Pr\left(\mathcal{A}_t\left(v'\right)\right)$ is $\frac{(1-\mu)t}{2(\eta+2)cN_c}$. Setting $C_t = \frac{2\beta c(\eta+2)}{1-\mu}$ $(0 < \beta < 1)$, we have

$$\begin{aligned} \left(1-\varepsilon'\right) &> \beta\left(1-\varepsilon\right)\left(1+\varepsilon'\right) \\ \varepsilon &> 1 - \frac{1-\varepsilon'}{\beta\left(1+\varepsilon'\right)}. \end{aligned} \quad (30)$$

From the above analysis, for a certain node set S with $|S| = k$, $|N\left(S\right)| \geq \left(1-\varepsilon\right)|E\left(S\right)|$ can be held with high probability. Since there are at most $\binom{N}{k} \leq \left(\frac{eN}{k}\right)^k$ node sets with $|S| = k$, for all the node sets with $|S| = k$, $|N\left(S\right)| \geq \left(1-\varepsilon\right)|E\left(S\right)|$ can be held with probability $1 - o\left(1\right)$, by a union bound. Thus, following the same argument, for all the node sets with $|S| \leq k$, we have that the bipartite graph can also hold $|N\left(S\right)| \geq \left(1-\varepsilon\right)|E\left(S\right)|$ when $M = O\left(k \log N/k\right)$ and t satisfying (27).

REFERENCES

[1] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[2] J. Wang, S. Tang, B. Yin, and X.-Y. Li, "Data gathering in wireless sensor networks through intelligent compressive sensing," in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 603–611.

[3] J. Cheng, Q. Ye, H. Jiang, D. Wang, and C. Wang, "STCDG: An efficient data gathering algorithm based on matrix completion for wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 850–861, Feb. 2013.

[4] B. Li, H. Li, W. Wang, Q. Yin, and H. Liu, "Performance analysis and optimization for energy-efficient cooperative transmission in random wireless sensor network," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4647–4657, Sep. 2013.

[5] H. Xu, L. Huang, C. Qiao, W. Dai, and Y.-E. Sun, "Joint virtual MIMO and data gathering for wireless sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 4, pp. 1034–1048, Apr. 2015.

[6] J. Haupt, W. U. Bajwa, M. Rabbat, and R. Nowak, "Compressed sensing for networked data," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 92–101, Mar. 2008.

[7] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[8] C. Luo, F. Wu, J. Sun, and C.-W. Chen, "Compressive data gathering for large-scale wireless sensor networks," in *Proc. Int. Conf. Mobile Comput. Netw. (MobiCom)*, Sep. 2009, pp. 145–156.

[9] L. Xiang, J. Luo, and A. Vasilakos, "Compressed data aggregation for energy efficient wireless sensor networks," in *Proc. 8th Annu. IEEE Commun. Soc. Conf. Sensor, Mesh Ad Hoc Commun. Netw. (SECON)*, Jun. 2011, pp. 46–54.

[10] W. Wang, M. Garofalakis, and K. Ramchandran, "Distributed sparse random projections for refinable approximation," in *Proc. Int. Symp. Inf. Process. Sensor Netw. (IPSN)*, Apr. 2007, pp. 331–339.

[11] M. Sartipi and R. Fletcher, "Energy-efficient data acquisition in wireless sensor networks using compressed sensing," in *Proc. Data Compress. Conf. (DCC)*, Mar. 2011, pp. 223–232.

[12] C. T. Chou, R. Rana, and W. Hu, "Energy efficient information collection in wireless sensor networks using adaptive compressive sensing," in *Proc. IEEE 34th Conf. Local Comput. Netw. (LCN)*, Oct. 2009, pp. 443–450.

[13] D. Ebrahimi and C. Assi, "Compressive data gathering using random projection for energy efficient wireless sensor networks," *Ad Hoc Netw.*, vol. 16, pp. 105–119, May 2014.

[14] D. Ebrahimi and C. Assi, "On the interaction between scheduling and compressive data gathering in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 4, pp. 2845–2858, Apr. 2016.

[15] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2346–2356, Jun. 2008.

[16] A. A. Abbasi and M. Younis, "A survey on clustering algorithms for wireless sensor networks," *Comput. Commun.*, vol. 30, nos. 14–15, pp. 2826–2841, Oct. 2007.

[17] R. Xie and X. Jia, "Transmission-efficient clustering method for wireless sensor networks using compressive sensing," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 3, pp. 806–815, Mar. 2014.

[18] H. L. Yap, A. Eftekhari, M. B. Wakin, and C. J. Rozell, "The restricted isometry property for block diagonal matrices," in *Proc. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2011, pp. 1–6.

[19] M. B. Wakin, J. Y. Park, H. L. Yap, and C. J. Rozell, "Concentration of measure for block diagonal measurement matrices," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Mar. 2010, pp. 3614–3617.

[20] M. T. Nguyen, K. A. Teague, and N. Rahnavard, "CCS: Energy-efficient data collection in clustered wireless sensor networks utilizing block-wise compressive sensing," *Comput. Netw.*, vol. 106, pp. 171–185, Sep. 2016.

[21] S. Abbasi-Daresari and J. Abouei, "Toward cluster-based weighted compressive data aggregation in wireless sensor networks," *Ad Hoc Netw.*, vol. 36, no. P1, pp. 368–385, Jan. 2016.

[22] M. T. Nguyen and K. A. Teague, "Compressive sensing based data gathering in clustered wireless sensor networks," in *Proc. IEEE Int. Conf. Distrib. Comput. Sensor Syst. (DCOSS)*, May 2014, pp. 187–192.

[23] R. Berinde, A. C. Gilbert, P. Indyk, H. Karloff, and M. J. Strauss. (2008). "Combining geometry and combinatorics: A unified approach to sparse signal recovery." [Online]. Available: https://arxiv.org/abs/0804.4666

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

[25] H. Zheng, F. Yang, X. Tian, X. Gan, X. Wang, and S. Xiao, "Data gathering with compressive sensing in wireless sensor networks: A random walk based approach," *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 1, pp. 35–44, Jan. 2015.

[26] S. Li and H. Qi, "Distributed data aggregation for sparse recovery in wireless sensor networks," in *Proc. IEEE Int. Conf. Distrib. Comput. Sensor Syst. (DCOSS)*, May 2013, pp. 62–69.
[27] O. Younis and S. Fahmy, "HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Trans. Mobile Comput.*, vol. 3, no. 4, pp. 366–379, Oct. 2004.
[28] R. Gribonval and M. Nielsen, "Sparse representations in unions of bases," *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 3320–3325, Dec. 2003.
[29] E. Candes and T. Tao, "The Dantzig selector: Statistical estimation when p is much larger than n," *Ann. Statist.*, vol. 35, no. 6, pp. 2313–2351, Dec. 2007.
[30] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1998.
[31] P. Gupta and P. R. Kumar, "Critical power for asymptotic connectivity in wireless networks," in *Stochastic Analysis, Control, Optimization and Applications*. Boston, MA, USA: Birkhaüser, 1999, pp. 547–566.
[32] M. D. Penrose, *Random Geometric Graphs*. London, U.K.: Oxford Univ. Press, 2003.
[33] D. Ustebay, R. Castro, and M. Rabbat, "Efficient decentralized approximation via selective gossip," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 4, pp. 805–816, Aug. 2011.
[34] F. Chung, *Spectral Graph Theory* (CBMS Regional Conference Series in Mathematics). Providence, RI, USA: American Mathematical Society, 1997.
[35] G. Mao, B. Fidan, and B. D. O. Anderson, "Wireless sensor network localization techniques," *Comput. Netw.*, vol. 51, no. 10, pp. 2529–2553, 2007.
[36] A. A. Kannan, B. Fidan, and G. Mao, "Robust distributed sensor network localization based on analysis of flip ambiguities," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov. 2008, pp. 1–6.
[37] J.-H. He and X.-H. Wu, "Variational iteration method: New development and applications," *Comput. Math. Appl.*, vol. 54, nos. 7–8, pp. 881–894, Oct. 2007.
[38] M. C. Grant and S. P. Boyd, *CVX: MATLAB Software for Disciplined Convex Programming*. [Online]. Available: http://cvxr.com/cvx/
[39] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. IEEE Annu. Hawaii Int. Conf. Syst. Sci.*, Jan. 2000, p. 10.
[40] H.-C. Liu, J. S. Min, and H. Samueli, "A low-power baseband receiver IC for frequency-hopped spread spectrum communications," *IEEE J. Solid-State Circuits*, vol. 31, no. 3, pp. 384–394, Mar. 1996.
[41] National Center for Atmospheric Research Staff (Eds). *The Climate Data Guide: SST (AMSRE): Sea Surface Temperature from Remote Sensing Systems*, accessed on Mar. 22, 2015. [Online]. Available: https://climatedataguide.ucar.edu/climate-data/sst-amsr-esea-surface-temperature-remote-sensing-systems

XIANGLING LI received the B.S. degree in information engineering and the M.S. degree in signal and information processing from Jilin University, Jilin, China, in 2008 and 2011, respectively. She is currently pursuing the Ph.D. degree with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China. Her research interests are in the area of wireless communications, with current emphasis on compressive sensing, data gathering in wireless sensor networks.

XIAOFENG TAO (SM'13) received the B.S. degree in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 1993, and the M.S.E.E. and Ph.D. degrees in telecommunication engineering from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 1999 and 2002, respectively.

He is currently a Professor with BUPT. He was the inventor or co-inventor of 80 patents and the author or co-author of 200 papers and two books, in wireless communication areas. He was a co-recipient of the Honored Mention Award at ACM MobiCom 2009 and best paper Awards at ISCIT 2012 and WCNC 2014, and also a Winner of Chinese National Invention Awards (2008 and 2013). He is the Chair of the IEEE ComSoc Beijing Chapter. His current research interests include 5G research.

GUOQIANG MAO (S'98–M'02–SM'08) received the Ph.D. degree in telecommunications engineering from Edith Cowan University in 2002. He was with the School of Electrical and Information Engineering, University of Sydney, between 2002 and 2014. He joined the University of Technology Sydney in 2014 as a Professor of wireless networking and the Director of the Center for Real-time Information Networks. The Center is among the largest university research centers in Australia in the field of wireless communications and networking. He has published about 200 papers in international conferences and journals, which have been cited more than 4000 times.

His research interest includes intelligent transport systems, applied graph theory, and its applications in telecommunications, Internet of Things, wireless sensor networks, wireless localization techniques, and network performance analysis. He has served as a Chair, a Co-Chair, and a TPC Member in a large number of international conferences. He is a Co-Chair of the IEEE Intelligent Transport Systems Society Technical Committee on Communication Networks. He received the Top Editor Award for outstanding contributions to the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY in 2011, 2014, and 2015. He has been an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS since 2014 and the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY since 2010.