

Article



Applicability of Compressive Sensing for Wireless Energy Harvesting Nodes

Thu L. N. Nguyen ¹, Yoan Shin ^{1,*}, Jin Young Kim ² and Dong In Kim ³

- ¹ School of Electronic Engineering, Soongsil University, Seoul 06978, Korea; thunguyen@ssu.ac.kr
- ² Department of Wireless Communications Engineering, Kwangwoon University, Seoul 01897, Korea; jinyoung@kw.ac.kr
- ³ College of Information and Communication Engineering, Sungkyunkwan University, Suwon 16419, Gyeonggi-do, Korea; dikim@skku.ac.kr
- * Correspondence: yashin@ssu.ac.kr

Received: 30 September 2017; Accepted: 2 November 2017; Published: 3 November 2017

Abstract: This paper proposes an approach toward solving an issue pertaining to measuring compressible data in large-scale energy-harvesting wireless sensor networks with channel fading. We consider a scenario in which N sensors observe hidden phenomenon values, transmit their observations using amplify-and-forward protocol over fading channels to a fusion center (FC), and the FC needs to choose a number of sensors to collect data and recover them according to the desired approximation error using the compressive sensing. In order to reduce the communication cost, sparse random matrices are exploited in the pre-processing procedure. We first investigate the sparse representation for sensors with regard to recovery accuracy. Then, we present the construction of sparse random projection matrices based on the fact that the energy consumption can vary across the energy harvesting sensor nodes. The key ingredient is the sparsity level of the random projection, which can greatly reduce the communication costs. The corresponding number of measurements is chosen according to the desired approximation error. Analysis and simulation results validate the potential of the proposed approach.

Keywords: compressive sensing; energy harvesting; sparse random projection

1. Introduction

A wireless sensor network (WSN) is an intelligent system with data collection, data fusion and independent transmission, which involves in many applications such as military surveillance, embedded systems, computer networks and communications. It consists of several sensors and each node is generally small in size and has a battery of limited capacity and energy. The lifetimes of WSNs thus are extremely limited by the total energy available in the batteries. Thus, using optimal techniques for energy management such as energy harvesting (EH), we can prolong the lifetime and duration of maintenance-free operation of WSNs. For instance, the energy existing in our environment from solar, wind, and thermal sources is converted into that can be used electrically. The advantages of EH-WSN solutions include high reliability, low energy needs, time savings, ecological compatibility and cost benefits [1–3]. In EH-WSN, each sensor node provides two functionalities: sensing, transmitting data to the fusion center (FC), and harvesting energy from ambient energy sources. The FC collects and reconstructs the observed signal by querying only a subset of sensors [4,5]. In order to reduce the energy consumption while forwarding observations to FC, we consider an innovative data gathering and reconstruction process based on three key subproblems: (i) compressive sensing (CS) based data acquisition; (ii) transmission of sparse random projection under fading for adapting random energy availability in EH systems; and (iii) CS based data reconstruction.

Data collected from wireless sensors are typically correlated, and are thus compressible in some appropriate domains. According to the CS theory, if a signal $\mathbf{x} \in \mathbb{C}^N$ is compressible, it can be well approximated using a small number $k \ll N$ of orthogonal transform coefficients [6–8]. Based on the CS model, the FC receives a compressed approximation of the original signal at multiple nodes by exploiting dense random matrix, i.e., all of the EH sensors in the networks participate in forwarding observations, and the FC randomly chooses them. However, in order to avoid this situation, which consumes a large amount of energy, we first have to build a sparse random projection such that the information can be extracted from any k-sparse signal. Second, we need to design a suitable recovery algorithm to reconstruct the original signal with good accuracy for given energy neutral conditions. Therefore, it would be good for EH sensors to prolong their lifetime and for FC to query an appropriate number of random projections and to still reconstruct a good approximation. Regarding sparse random projections, a good random projection will preserve all pairwise distances with a high probability. Thus, it can be used as a reliable estimator of distances in the original space. In [9], the authors proposed a distributed compressive sensing scheme for WSNs, where the sparsity of the random projections is used consistently to reduce the computational complexity and the communication cost. They also proved that the sparse random projections are sufficient to recover a data approximation that is comparable to the optimal k-term approximation with a high probability. With the fading channel and energy-harvesting constraints, the problems regarding sparsity of random projections are studied in [10,11]. In [10], the authors only considered additive white Gaussian noise (AWGN) channels, while, in [11], the authors focused on Rayleigh fading channels and investigated sufficient conditions for guaranteeing a reliable and computationally-efficient data approximation for the sparse random projections. Due to the harvesting conditions, the sensors typically have different energy harvesting rates that lead to different available energy constraints. However, the sparsity factors in the aforementioned works have been assumed to be homogeneous for all sensors and were kept fixed for entire transmission states. Thus, they cannot be responsive to battery dynamics and channel conditions. To overcome those issues, we consider a dynamic sparsity factor that relates to the available energy constraints of wireless sensors and transmission between them, and then build a sparse random projection matrix that is stable and robust under channel fading effects and CS recovery.

The main purpose of this paper is to study sparse representation and sparse random projection for EH WSNs under fading channels. We consider a problem of data transmission in EH WSNs where multiple sensors send spatially-correlated data to a fusion center using amplify-and-forward (AF) protocol over independent Rayleigh fading channels with additive noise. Supposing that the measured data are compressible under an approximate orthogonal transform, our task is first choosing a certain number of sensors to query according to the desired approximation error by designing a sparse random projection matrix, and then exploiting the CS recovery algorithm to obtain an optimal approximation. Inspired by the work in [12] on sparse random projections for heavy-tailed data, we propose a random projection-based CS scheme where the sparsity factor is dynamic due to energy constraints. We also prove that, under the fading channel condition, our projection matrix still satisfies the restricted isometry property (RIP) for successful recovery in CS.

The organization of this paper is as follows. In Section 2, we introduce the problem of recovering a signal observed by an EH WSN under channel fading, and briefly introduce the compressive sensing. In Section 3, we present our construction on basis representations for compressible data and sparse random projection design. Section 4 proves that our sparse random matrices preserve the pairwise distance under the fading and guarantee the reconstruction accuracy subject to the energy constraints. The simulation results and conclusion are presented in Sections 5 and 6, respectively.

Notations: We denote $\mathbf{A} = [a_{ij}]$ as a matrix whose entries are a_{ij} , $(\cdot)^{-1}$ as the matrix inverse operation, $(\cdot)^*$ as the conjugate transpose, $\lfloor \cdot \rfloor$ as the floor operation, $|T| = \operatorname{supp}(T)$ as the number of elements in a given set T, $\mathbb{E}(\cdot)$ and $\operatorname{Var}(\cdot)$ as the expectation and the variance operators, respectively. The ℓ_p norm of a vector $\mathbf{x} = [x_1 \cdots x_n]^T$ is defined as $||\mathbf{x}||_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ for a positive integer p. We call a signal \mathbf{x} is a k-sparse vector if $||\mathbf{x}||_0 \triangleq |\operatorname{supp}(\mathbf{x})| \leq k$. The notation $\mathcal{O}(\cdot)$ denotes the

complexity operation, $\mathcal{CN}(\mu, \Sigma)$ denotes the circularly symmetric complex Gaussian distribution with mean μ and covariance Σ , and $\mathbf{w} \sim \mathcal{CN}(\mu, \Sigma)$ means that \mathbf{w} is distributed according to $\mathcal{CN}(\mu, \Sigma)$, alternatively, $\mathbf{w} \sim \mathcal{N}\left(\mu, \frac{1}{2}\begin{bmatrix} \operatorname{Re}(\Sigma) & -\operatorname{Im}(\Sigma) \\ \operatorname{Im}(\Sigma) & \operatorname{Re}(\Sigma) \end{bmatrix} \right)$.

2. System Model

2.1. Problem Formulation

We consider an EH-WSN consisting of N sensors, each of which observes a single value $x_j \in \mathbb{C}$ then transmits it to the FC with the AF protocol. Note that the decode-and-forward (DF) approach can be used for sensor transmissions, where we apply digital modulation schemes to transmit the data. However, as shown in [13–15], for a simple distributed sensor network, the AF approach over multiple-access channel (MAC) is optimal for signal detection and estimation, as well as saving energy in relaying data. Thus, we restrict our analysis in this paper to analog transmission suitable for energy-constrained EH-WSNs, while digital modulated signals will be considered in the future work The FC collects the received signals in M (M < N) time slots and recovers the original signal based on this measurement as

$$\mathbf{y} = \mathbf{B}\mathbf{x} + \mathbf{w},\tag{1}$$

where $\mathbf{B} = \mathbf{H} \odot \mathbf{A}$ with \odot representing the Hadamard product. The matrix $\mathbf{H} = [h_{ij}] \in \mathbb{C}^{M \times N}$ represents the flat fading channels between the sensors and the FC, which is a random matrix having independent and identically distributed (i.i.d.) complex circular Gaussian entries with zero-mean and unit variance, i.e., $h_{ij} = h_{ij}^R + jh_{ij}^I$, where $h_{ij}^R \sim \mathcal{N}(0, \frac{1}{2})$ and $h_{ij}^I \sim \mathcal{N}(0, \frac{1}{2})$. The matrix $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{M \times N}$ represents the random projection with energy constraints, which will be described later, $\mathbf{x} = [x_1, \cdots, x_N]^T$ is the transmitted vector, and $\mathbf{w} = [w_1, \cdots, w_M]$ is the additive noise where each $w_j \sim \mathcal{CN}(0, \sigma_w^2)$. The real-value of Equation (1) can be written as

$$\widehat{\mathbf{y}} = \begin{bmatrix} \operatorname{Re}(\mathbf{H} \odot \mathbf{A}) & -\operatorname{Im}(\mathbf{H} \odot \mathbf{A}) \\ \operatorname{Im}(\mathbf{H} \odot \mathbf{A}) & \operatorname{Re}(\mathbf{H} \odot \mathbf{A}) \end{bmatrix} \begin{bmatrix} \operatorname{Re}(\mathbf{x}) \\ \operatorname{Im}(\mathbf{x}) \end{bmatrix} + \begin{bmatrix} \operatorname{Re}(\mathbf{w}) \\ \operatorname{Im}(\mathbf{w}) \end{bmatrix}$$

$$= \widehat{\mathbf{B}} \begin{bmatrix} \operatorname{Re}(\mathbf{x}) \\ \operatorname{Im}(\mathbf{x}) \end{bmatrix} + \begin{bmatrix} \operatorname{Re}(\mathbf{w}) \\ \operatorname{Im}(\mathbf{w}) \end{bmatrix},$$

$$(2)$$

where

$$\widehat{\mathbf{B}} = \begin{bmatrix} \operatorname{Re}(\mathbf{H}) \odot \mathbf{A} & -\operatorname{Im}(\mathbf{H}) \odot \mathbf{A} \\ \operatorname{Im}(\mathbf{H}) \odot \mathbf{A} & \operatorname{Re}(\mathbf{H}) \odot \mathbf{A} \end{bmatrix} \in \mathbb{R}^{2M \times 2N}.$$
(3)

Our goal is to find a good approximation for **x** given **y** and **B**. According to the CS theory, for a given upper bound of error ϵ , the FC can recover **x** by solving the following optimization problem [6]:

$$\widehat{\mathbf{x}} = \min_{\mathbf{z} \in \mathbb{C}^N} ||\mathbf{z}||_1 \quad \text{subject to } ||\mathbf{y} - \mathbf{B}\mathbf{z}||_2^2 \le \epsilon.$$
(4)

However, in order to obtain recovery guarantees of a given x based on Equation (4), some essential conditions for x and B will be considered in the next section.

2.2. Signal Reconstruction with Compressive Sensing

We define Σ_k by the set of all *k*-sparse signals as $\Sigma_k = \{\mathbf{x} \in \mathbb{C}^N, ||\mathbf{x}||_0 \le k\}$. We say that the matrix **B** satisfies the restricted isometry property (RIP) of order *k* if there exists a number $\delta \in (0, 1)$ such that $(1 - \delta_k) ||\mathbf{x}||_2^2 \le ||\mathbf{B}\mathbf{x}||_2^2 \le (1 + \delta_k) ||\mathbf{x}||_2^2$, $\forall \mathbf{x} \in \Sigma_k$. The best *k*-term approximation denoted by \mathbf{x}_k , can be obtained by

$$\mathbf{x}_{k} = \arg\min_{\mathbf{z}\in\Sigma_{k}}||\mathbf{x}-\mathbf{z}||_{1}.$$
(5)

We suppose that **B** satisfies the RIP of order 2k, i.e., $\mathbf{B}\mathbf{x}_1 \neq \mathbf{B}\mathbf{x}_2$ for any pair $\mathbf{x}_1, \mathbf{x}_2 \in \Sigma_k$, where $\mathbf{x}_1 \neq \mathbf{x}_2$, and the corresponding restricted isometry constant $\delta_{2k} < \sqrt{2} - 1$. For a given measurement **y** in Equation (1), where $||\mathbf{w}||_2 \leq \epsilon$, and the solution of Equation (4) obeys

$$||\widehat{\mathbf{x}} - \mathbf{x}||_2 \le c_0 \epsilon + c_1 \frac{||\mathbf{x} - \mathbf{x}_k||_2}{\sqrt{k}},\tag{6}$$

where $c_0 = 4 \frac{\sqrt{1+\delta_{2k}}}{1-(1+\sqrt{2})\delta_{2k}}$, $c_1 = 2 \frac{1-(1-\sqrt{2})\delta_{2k}}{1-(1+\sqrt{2})\delta_{2k}}$, according to the results in [8].

3. Compressive Sensing for Wireless Energy Harvesting Nodes

3.1. Motivation

In this section, we present the idea of constructing a sparse random projection and basis representation for achieving a significant speed up with little loss in accuracy recovery. First, we introduce an appropriate sparse representation basis for complex data, which takes into account data transmission cost and data recovery quality. The goal of this step is to obtain the sparse representation learned from the sensor data; thus, it has the ability to adapt the signal under fading. Second, we provide the concept of the sparse random projection, which is used for measurement matrix design. This design ensures that each sampling value represents one CS measurement guaranteeing a successful recovery, and satisfies the assumptions of energy on sensing. Finally, we verify the sparsity level and the RIP condition for the sensing matrix.

3.2. Basis Representation for Compressible Data

In practice, the signal **x** may not be sparse, but we want to resolve it in a certain sparse basis Ψ , i.e., $\mathbf{x} = \Psi \alpha$, where Ψ is a unitary $N \times N$ matrix and $\alpha \in \mathbb{C}^N$ has at most $k < \lfloor \frac{N}{2} \rfloor$ non-zero components. The matrix Ψ can be obtained either from an appropriate transform (e.g., wavelets transform, discrete Fourier transform, etc.) or from learning a dictionary to perform best based on a training set [16]. In this case, we require that $\mathbf{B}\Psi$ satisfies the RIP and the performance will depend on $||\hat{\alpha} - \alpha||_2$. We can sort elements of the vector \mathbf{x} in the decreasing order of magnitude, where the *i*-th largest coefficient satisfies $|x_{\mathcal{I}(j)}| \leq Gi^{-1/r}, j = 1, \dots, N$. Here, $\mathcal{I} \subset \{1, \dots, N\}$ represents the index set of the sorted elements, where $|\mathcal{I}(j)| \leq M \ll N$. For a rate of decay 0 < r < 2, the approximation error in ℓ_2 -norm can be obtained by taking the *k* largest coefficients as

$$||\widehat{\mathbf{x}} - \mathbf{x}||_2 = ||\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}||_2 \le (rs)^{-\frac{1}{2}} Gk^{-s},\tag{7}$$

where *G* is a constant and $s = \frac{1}{r} - \frac{1}{2}$ [7].

3.3. Measurement Matrix: Sparse Random Projection Design

We first introduce the Johnson–Lindenstrauss (JL) lemma and its connection with the random matrix constructions of CS with regard to the stable embedding of a finite set of points under a random dimensionality-reducing projection. This lemma is stated as follows.

Lemma 1. Let Q be a collection of finite points in \mathbb{R}^N . Given $0 < \epsilon < 1$ and $\beta > 0$, let **A** be a random orthogonal projection from \mathbb{R}^N to \mathbb{R}^M with $M \ll N$ and

$$M \ge \left(\frac{4+2\beta}{\epsilon^2/2 - \epsilon^3/3}\right) \log(|Q|).$$
(8)

According to [17], with probability at least $1 - |Q|^{-\beta}$, for all $\mathbf{q}_i, \mathbf{q}_j \in Q$, $i \neq j$, the following statement holds

$$(1-\epsilon)\sqrt{\frac{M}{N}}||\mathbf{q}_i-\mathbf{q}_j||_2^2 \le ||\mathbf{A}\mathbf{q}_i-\mathbf{A}\mathbf{q}_j||_2^2 \le (1+\epsilon)\sqrt{\frac{M}{N}}||\mathbf{q}_i-\mathbf{q}_j||_2^2.$$
(9)

With respect to the JL lemma, the authors in [9,10] consider the AWGN channel, where the random projection matrix **A** with i.i.d. entries is defined as

$$a_{ij} = \frac{1}{\sqrt{\rho}} \begin{cases} +1, & \text{with probability (w. p.)} \ \frac{\rho}{2}, \\ 0, & \text{w. p. } 1 - \rho, \\ -1, & \text{w. p. } \frac{\rho}{2}, \end{cases}$$
(10)

where ρ is a factor that provides the probability of measurement and controls the sparsity level of **A**. For example, if $\rho = 1$, the random matrix has no sparsity, and if $\rho = \frac{\log N}{N}$, the expected number of non-zeros in each row is log *N*. Moreover, these authors proved that the a_{ij} are four-wise independent in rows and independent across rows, i.e.,

$$\mathbb{E}[a_{ij}] = 0, \ \mathbb{E}[a_{ij}^2] = 1, \ \mathbb{E}[a_{ij}^4] = \frac{1}{\rho}.$$
(11)

Note that the energy consumed for wireless transmission cannot exceed the energy available in each slot. Thus, it is reasonable to take into account both the energy constraint and the sparse random projection for reducing data transmission cost and improving data recovery quality. With regard to fading channels, Ref. [11] gave us an improvement of Equation (10), with which the projected data matrix **A** is associated with a squared-amplitude $b_i > 0$. Each entry is

$$a_{ij} = \sqrt{b_j} \begin{cases} +1, & \text{w. p. } \frac{\rho}{2}, \\ 0, & \text{w. p. } 1 - \rho, \\ -1, & \text{w. p. } \frac{\rho}{2}, \end{cases}$$
(12)

for $j = 1, \dots, M$. Given an available energy E_j , the value of b_j is chosen such that

$$\rho b_j \le E_j, \ j = 1, \cdots, N. \tag{13}$$

It leads to $E(a_{ij}^2) = \frac{pb_j}{\sqrt{N}} \le E_j$, that is, the energy can be saved for the future transmission. In this paper, we define our sparse measurement matrix as $\mathbf{A} = [a_{ij}]$, where each entry a_{ij} is defined by

$$a_{ij} = \frac{\sqrt[4]{N}}{\sqrt{\rho_{ij}}} \begin{cases} +1, & \text{w. p. } \frac{\rho_{ij}}{2\sqrt{N}}, \\ 0, & \text{w. p. } 1 - \rho_{ij}, \\ -1, & \text{w. p. } \frac{\rho_{ij}}{2\sqrt{N}}. \end{cases}$$
(14)

Based on the definition of \mathbf{A} in Equation (14), the measurement vector \mathbf{y} can be expressed as in Equation (1). In the sequel, we explain the reasons for choosing this sparse random projection.

(i) It has been shown that the conventional random projections A ~ N(0,1) are appropriate only for the l₂ norm, while, in many applications, there is greater concern for the inner product [12]. Moreover, one can use 1/ρ ≫ 3 in Equation (10) to speed up the computation process [12].

(ii) The projection given in Equation (14) selects the sensor to transmit and assigns weights to the data according to the harvested energy at the sensor. For instance, the sparsity of random projection given E_{ij} , can be defined as

$$\rho_{ij} = \frac{p_{ij}^{opi}}{E_{ij}},\tag{15}$$

where p_{ij}^{opt} is the optimal power allocation and E_{ij} is the available energy of node *j* during the *i*-th slot. Therefore, we obtain

$$\mathbb{E}(a_{ij}) = 0, \ \mathbb{E}(a_{ij}^2) = \operatorname{Var}(a_{ij}) = 1, \ \mathbb{E}(a_{ij}^4) = \frac{\sqrt{N}}{\rho_{ij}}.$$
 (16)

4. Proposed Distributed Algorithm and Analysis

4.1. Sparse Random Projection with Fading Channels

Suppose that we have two input vectors $\mathbf{x}_1 = [x_1^{(1)}, \dots, x_N^{(1)}]^T$, $\mathbf{x}_2 = [x_1^{(2)}, \dots, x_N^{(2)}]^T \in \mathbb{C}^N$ (alternatively, $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^{2N}$) and the random matrix $\hat{\mathbf{B}} = [\hat{b}_{ij}]$ given in Equation (3). The corresponding projections of \mathbf{x}_1 and \mathbf{x}_2 are defined by

$$\mathbf{u} = \frac{1}{\sqrt{M}} \widehat{\mathbf{B}} \mathbf{x}_1, \ \mathbf{v} = \frac{1}{\sqrt{M}} \widehat{\mathbf{B}} \mathbf{x}_2 \in \mathbb{R}^{2M}.$$
 (17)

We also assume that under fading the channel matrix **H** is independent of the random matrix **A**. Thus, we have

$$\mathbb{E}(\hat{b}_{ij}) = 0, \ \mathbb{E}(\hat{b}_{ij}^2) = \text{Var}(\hat{b}_{ij}) = 1, \text{ and } \mathbb{E}(\hat{b}_{ij}^4) = \frac{3\sqrt{N}}{\rho_{ij}}.$$
 (18)

The sparse random projection **A** is desired to have the properties of length, distance, and inner product preservation. We need to check that those properties are still preserved under fading channel conditions. In order to check the *length* preservation of the sparse random matrix $\hat{\mathbf{B}}$, we first express $\mathbb{E}(||\mathbf{u}||_2^2) = \sum_{i=1}^{2M} \mathbb{E}(u_i^2)$, where

$$\begin{split} \mathbb{E}(u_i^2) &= \frac{1}{2M} \mathbb{E}\left[\sum_{j=1}^{2N} \{x_j^{(1)}\}^2 \hat{b}_{ij}^2 + \sum_{l \neq m} x_l^{(1)} x_m^{(1)} \hat{b}_{il} \hat{b}_{im}\right] \\ &= \frac{1}{2M} \left[\sum_{j=1}^{2N} \{x_j^{(1)}\}^2 \mathbb{E}(\hat{b}_{ij}^2) + \sum_{l \neq m} x_l^{(1)} x_m^{(1)} \mathbb{E}(\hat{b}_{il}) \mathbb{E}(\hat{b}_{im})\right] \\ &= \frac{1}{2M} \left[\sum_{j=1}^{2N} \{x_j^{(1)}\}^2\right] = \frac{1}{2M} ||\mathbf{x}_1||_2^2. \end{split}$$

Thus, $\mathbb{E}(||\mathbf{u}||_2^2) = \sum_{i=1}^{2M} \frac{1}{2M} ||\mathbf{x}_1||_2^2 = ||\mathbf{x}_1||_2^2$. For the *distance* preservation, we have

$$\begin{split} \mathbb{E}(||\mathbf{u} - \mathbf{v}||_{2}^{2}) &= \sum_{i=1}^{2M} \mathbb{E}[(u_{i} - v_{i})^{2}] \\ &= \sum_{i=1}^{2M} \frac{1}{2M} \left[\sum_{j=1}^{2N} \{x_{j}^{(1)} - x_{j}^{(2)}\}^{2} \mathbb{E}(\hat{b}_{ij}^{2}) \right] \\ &= \sum_{i=1}^{2M} \frac{1}{2M} ||\mathbf{x}_{1} - \mathbf{x}_{2}||_{2}^{2} = ||\mathbf{x}_{1} - \mathbf{x}_{2}||_{2}^{2}. \end{split}$$

Similarly, we can compute the *inner product* as $\mathbb{E}(\mathbf{u} \cdot \mathbf{v}) = \mathbb{E}(\mathbf{u}^T \mathbf{v}) = \sum_{i=1}^{2M} \mathbb{E}(u_i v_i)$, where

$$\begin{split} \mathbb{E}(u_i v_i) &= \mathbb{E}\left[\frac{1}{2M} \left(\sum_{j=1}^{2N} x_j^{(1)} \hat{b}_{ij}\right) \left(\sum_{j=1}^{2N} x_j^{(2)} \hat{b}_{ij}\right)\right] \\ &= \frac{1}{2M} \sum_{j=1}^{2N} \left[x_j^{(1)} x_j^{(2)} \mathbb{E}(\hat{b}_{ij}^2) + \sum_{l \neq m} x_l^{(1)} x_m^{(2)} \mathbb{E}(\hat{b}_{il}) \mathbb{E}(\hat{b}_{im})\right] \\ &= \frac{1}{2M} \sum_{j=1}^{2N} x_j^{(1)} x_j^{(2)} = \frac{1}{2M} \mathbf{x}_1^T \mathbf{x}_2. \end{split}$$

Thus, the inner product is still preserved by applying \widehat{B} since we have $\mathbb{E}(\mathbf{u} \cdot \mathbf{v}) = \mathbf{x}_1 \cdot \mathbf{x}_2$.

4.2. Stability and Robustness of Sparse Random Projections

By partitioning the sparse random matrix $\widehat{\mathbf{B}}$ into $\widehat{\mathbf{B}}_{(1)}, \dots, \widehat{\mathbf{B}}_{M_2}$, where each $\widehat{\mathbf{B}}_{(\ell)}(\ell = 1, \dots, M_2)$ has size $M_1 \times N$ and $2M = M_1M_2$ (M_1 and M_2 will be determined later), the corresponding measurement $\mathbf{y} \in \mathbb{R}^{2M}$ can be split into M_2 vectors $\{\mathbf{y}_{(1)}, \dots, \mathbf{y}_{(M_2)}\}$. Each $\mathbf{y}_{(\ell)} \in \mathbb{R}^{M_1}$ is defined as

$$\mathbf{y}_{(\ell)} = \widehat{\mathbf{B}}_{(\ell)} \mathbf{x} + \mathbf{w}_{(\ell)}.$$
(19)

We let $\mathbf{z}_{(\ell)} = \widehat{\mathbf{B}}_{(\ell)} \boldsymbol{\psi}$, where $\boldsymbol{\psi} \in \mathbb{R}^{2N}$ and $||\boldsymbol{\psi}||_2^2 = 1$. Thus, we can perform

$$\mathbf{z}_{(\ell)}^{T}\mathbf{y}_{(\ell)} = \boldsymbol{\psi}^{T}\widehat{\mathbf{B}}_{(\ell)}^{T}\widehat{\mathbf{B}}_{(\ell)}\mathbf{x} + \boldsymbol{\psi}^{T}\widehat{\mathbf{B}}_{(\ell)}^{T}\mathbf{w}_{(\ell)} = \sum_{i_{0}=1}^{M_{1}}u_{i_{0}}^{(\ell)} + \sum_{i_{0}=1}^{M_{1}}v_{i_{0}}^{(\ell)},$$
(20)

where $u_{i_0}^{(\ell)} = \left(\sum_{j=1}^{2N} \hat{b}_{i_0j}^{(\ell)} \psi_j\right) \left(\sum_{j=1}^{2N} \hat{b}_{i_0j}^{(\ell)} x_j\right)$ and $v_{i_0}^{(\ell)} = \left(\sum_{j=1}^{2N} \hat{b}_{i_0j}^{(\ell)} \psi_j\right) w_{i_0}$. The corresponding means, variances of u_i , v_i , and their covariances can be calculated as

$$\mathbb{E}[u_{i_0}^{(\ell)}] = \boldsymbol{\psi}^T \mathbf{x},\tag{21}$$

$$\operatorname{Var}[u_{i_0}^{(\ell)}] = (\boldsymbol{\psi}^T \mathbf{x})^2 + ||\boldsymbol{\psi}||_2^2 ||\mathbf{x}||_2^2 + \sum_{j=1}^{2N} \left(\frac{3\sqrt{N}}{\rho_{i_0j}^{(\ell)}} - 3\right) x_j^2 \psi_j^2, \tag{22}$$

$$\mathbb{E}[v_{i_0}^{(\ell)}] = 0, \tag{23}$$

$$\operatorname{Var}[v_{i_0}^{(\ell)}] = \sigma_w^2 ||\psi||_2^2, \tag{24}$$

$$\operatorname{Cov}[u_{i_0}^{(\ell)}, v_{\bar{i}_0}^{(\ell)}] = 0.$$
⁽²⁵⁾

The detailed derivations of Equations (21)–(25) are given in the Appendix A. Thus, we obtain

$$\mathbb{E}\left[\frac{1}{M_1}\mathbf{z}_{(\ell)}^T\mathbf{y}_{(\ell)}\right] = \frac{1}{M_1}\left\{\mathbb{E}\left[\sum_{i=1}^{M_1}u_i^{(\ell)}\right] + \mathbb{E}\left[\sum_{i=1}^{M_1}v_i^{(\ell)}\right]\right\} = \boldsymbol{\psi}^T\mathbf{x},\tag{26}$$

$$\operatorname{Var}\left[\frac{1}{M_{1}}\mathbf{z}_{(\ell)}^{T}\mathbf{y}_{(\ell)}\right] = \frac{1}{M_{1}^{2}}\left[M_{1}\left\{(\boldsymbol{\psi}^{T}\mathbf{x})^{2} + (||\mathbf{x}||_{2}^{2} + \sigma_{w}^{2})||\boldsymbol{\psi}||_{2}^{2}\right\} + \sum_{i_{0}=1}^{M_{1}}\sum_{j=1}^{2N}\left(\frac{3\sqrt{N}}{\rho_{i_{0}j}^{(\ell)}} - 3\right)x_{j}^{2}\psi_{j}^{2}\right].$$
 (27)

For any $\epsilon > 0$, using the Chebychev's inequality and the fact that $||\psi||_2^2 = 1$, we have

$$P\left\{ \left| \frac{1}{M_{1}} \mathbf{z}_{(\ell)}^{T} \mathbf{y}_{(\ell)} - \boldsymbol{\psi}^{T} \mathbf{x} \right| \geq \epsilon ||\mathbf{x}||_{2} \right\} \leq \frac{\operatorname{Var} \left[\frac{1}{M_{1}} \mathbf{z}_{(\ell)}^{T} \mathbf{y}_{(\ell)} \right]}{\epsilon^{2} ||\mathbf{x}||_{2}^{2}}$$

$$= \frac{1}{\epsilon^{2} M_{1}^{2}} \left\{ M_{1} \left[\frac{(\boldsymbol{\psi}^{T} \mathbf{x})^{2}}{||\mathbf{x}||_{2}^{2}} + \frac{||\mathbf{x}||_{2}^{2} + \sigma_{w}^{2}}{||\mathbf{x}||_{2}^{2}} \right] + \frac{\sum_{i_{0}=1}^{M_{1}} \sum_{j=1}^{2N} \left(\frac{3\sqrt{N}}{\rho_{i_{0}j}^{(\ell)}} - 3 \right) x_{j}^{2} \psi_{j}^{2}}{||\mathbf{x}||_{2}^{2}} \right\} \quad (28)$$

$$\leq \frac{1}{\epsilon^{2} M_{1}} \left\{ 2 + \frac{\sigma_{w}^{2}}{||\mathbf{x}||_{2}^{2}} + \sum_{j=1}^{2N} \frac{3\sqrt{N}}{\min_{i_{0}} \rho_{i_{0}j}^{(\ell)}} \mu^{2} \right\} \triangleq \delta.$$

We have also used the fact that for any data vector $\mathbf{u} \in \mathbb{R}^{2N}$, it satisfies the peak-to-total energy condition, i.e., $\frac{||\mathbf{u}||_{\infty}}{||\mathbf{u}||_2} \leq \mu$ [10]. Following the approach given in [9], the probability that an estimate lies outside the tolerable approximation interval cannot exceed $e^{-c^2M_2/12}$, where 0 < c < 1. Setting $M_1 = \mathcal{O}\left[\frac{1}{\epsilon^2}\left\{2 + \frac{\sigma_w^2}{||\mathbf{x}||_2^2} + \sum_{j=1}^{2N} \frac{3\sqrt{N}}{\min_0 \rho_{i_0j}^{(\ell)}}\mu^2\right\}\right]$ yields $\delta = \frac{1}{4}$, and setting $M_2 = \mathcal{O}[(1+\eta)\log 2N]$ gives $p_e \leq (2N)^{-\eta}$ for some constant $\eta > 0$. Finally, for $M = \frac{1}{2}M_1M_2 = \frac{1}{2}\mathcal{O}\left[\frac{\mu^2(1+\eta)}{\epsilon^2}\left(2 + \frac{\sigma_w^2}{||\mathbf{x}||_2^2} + \sum_{j=1}^{2N} \frac{3\sqrt{N}}{\min_0 \rho_{i_0j}^{(\ell)}}\right)\log 2N\right]$, the sparse random matrix $\hat{\mathbf{B}}$ can preserve all the pairwise inner products within an approximation error ϵ with the probability of at least $1 - (2N)^{-\eta}$.

Remark 1 (Complexity Analysis). According to CS model with a dense random matrix [6,8], it requires at least $O(k \log \frac{N}{\eta})$ measurements for obtaining an approximation via ℓ_1 -minimization problem in Equation (4) with probability exceeding $1 - \eta$, and the CS decoding has computational complexity $O(N^3)$. On the other hand, the sparse random projection scheme requires at least $O(k^2 \log N)$ random projections and the corresponding decoding process takes $O(MN \log N)$, where M is the number of measurements. Since $k \ll N$, using sparse random projection attains low decoding complexity, which makes it applicable for EH sensors, while the FC can request a little more measurements from sensors and recover the signal with a better approximation. Our proposed scheme has inherited this advantage and optimized the sparsity level, which adapts to channel conditions and energy constraints.

4.3. Sparsity Level and RIP Verification

Following the signal model in Equation (1), i.e., $\mathbf{y} = \mathbf{\hat{B}} \Psi \mathbf{\alpha} + \mathbf{w}$, the decoding process is to recover the sparse signal $\mathbf{\alpha}$ instead of recovering the sensor data \mathbf{x} by using Equation (4). However, we must verify that Ψ satisfies the sparse basis representation in \mathbb{R}^N and the matrix $\mathbf{\bar{Z}} = \mathbf{\widehat{B}} \Psi$ obeys the RIP condition to guarantee successful recovery via ℓ_1 -minimization.

To analyze the feasibility of the measurement matrix and the sparse basis design, we have to answer the following two questions:

- (1) Is it reasonable to select Ψ obtained from Section 3.2 as an orthogonal basis for x?
- (2) For the matrix **A** obtained from Section 3.3, does $\overline{\mathbf{Z}} = \widehat{\mathbf{B}} \Psi$ obey the RIP condition?

First, as we demonstrated in Section 3.2, the matrix Ψ is obviously an orthogonal basis in \mathbb{C}^N from an appropriate transformation. Otherwise, it can be an overcomplete dictionary from data learning approach, which promised to represent a wider range of signal phenomena [16].

Second, in order to show that the random variable $||\mathbf{Z}\alpha||_2$ is highly concentrated about $||\alpha||_2$, we can assume that the row of $\hat{\mathbf{B}}$ is independent of Ψ . Fixing $\epsilon \in (0, 1)$, and with each row of $\bar{\mathbf{Z}}$

satisfying the sub-Gaussian distribution, we prove that $\mathbf{Z} = [\mathbf{z}_i]_{i=1}^M = \frac{1}{\sqrt{M}} [\bar{\mathbf{z}}_1, \cdots, \bar{\mathbf{z}}_M]^T$ will satisfy the RIP with high probability, i.e.,

$$(1-\epsilon) \leq |\frac{||\mathbf{Z}\boldsymbol{\alpha}||_2^2}{||\boldsymbol{\alpha}||_2^2} \leq (1+\epsilon), \ \forall \boldsymbol{\alpha} \in \Sigma_k.$$
(29)

To do that, we prove that each part of the matrix $\mathbf{Z} = \mathbf{Z}_R + j\mathbf{Z}_I$ satisfies the RIP for complex data α , i.e.,

$$\frac{1}{2}(1-\epsilon) \le |\frac{||\mathbf{Z}_R \boldsymbol{\alpha}||_2^2}{||\boldsymbol{\alpha}||_2^2} \le \frac{1}{2}(1+\epsilon),\tag{30}$$

and
$$\frac{1}{2}(1-\epsilon) \le |\frac{||\mathbf{Z}_{I}\boldsymbol{\alpha}||_{2}^{2}}{||\boldsymbol{\alpha}||_{2}^{2}} \le \frac{1}{2}(1+\epsilon), \ \forall \boldsymbol{\alpha} \in \Sigma_{k}.$$
 (31)

First, in order to to prove Equation (30), by letting $\mathbf{Z}_{R} = [\mathbf{z}_{ij}^{R}]$, we have

$$\mathbb{E}(\mathbf{z}_{i}^{R} \cdot \boldsymbol{\alpha}) = \mathbb{E}\left(\sum_{j=1}^{N} z_{ij}^{R} \alpha_{j}\right) = \sum_{j=1}^{N} \mathbb{E}(z_{ij})^{R} \alpha_{j} = 0,$$

$$\operatorname{Var}(\mathbf{z}_{i}^{R} \cdot \boldsymbol{\alpha}) = \operatorname{Var}\left(\sum_{j=1}^{N} z_{ij}^{R} \alpha_{j}\right) = \sum_{j=1}^{N} \operatorname{Var}(z_{ij}^{R}) \alpha_{j}^{2} = \frac{||\boldsymbol{\alpha}||_{2}^{2}}{M},$$

$$\mathbb{E}(||\mathbf{Z}^{R} \boldsymbol{\alpha}||_{2}^{2}) = \mathbb{E}\left[\sum_{i=1}^{M} (\mathbf{z}_{i}^{R} \cdot \boldsymbol{\alpha})^{2}\right] = \sum_{i=1}^{M} \mathbb{E}(\mathbf{z}_{i}^{R} \cdot \boldsymbol{\alpha})^{2}$$

$$= \sum_{i=1}^{M} \operatorname{Var}(\mathbf{z}_{i}^{R} \cdot \boldsymbol{\alpha}) = \sum_{i=1}^{M} \frac{||\boldsymbol{\alpha}||_{2}^{2}}{M} = ||\boldsymbol{\alpha}||_{2}^{2}.$$

Then, following Theorem 4.2 of [18], we obtain $P\left[\frac{||\mathbf{Z}^R \boldsymbol{\alpha}||_2^2}{||\boldsymbol{\alpha}||_2^2} \leq \frac{1}{2}(1-\epsilon)\right] = e^{-M\epsilon^2/4c^2}$ and $P\left[\frac{||\mathbf{Z}^R \boldsymbol{\alpha}||_2^2}{||\boldsymbol{\alpha}||_2^2} \geq \frac{1}{2}(1+\epsilon)\right] = e^{-M\epsilon^2/4c^2}$. Thus,

$$P\left(\left|\frac{||\mathbf{Z}^{R}\boldsymbol{\alpha}||_{2}^{2}}{||\boldsymbol{\alpha}||_{2}^{2}} \leq \frac{1}{2}\right| \geq \frac{\epsilon}{2}\right) \leq 2e^{-M\epsilon^{2}/4c^{2}}.$$
(32)

Fixing an index set $\mathcal{I} \subset \{1, \dots, N\}$ with $|\mathcal{I}| = k$, there are $\binom{N}{k}$ possible *k*-dimensional subspaces of \mathbb{Z}^R and the probability of a *k*-sparse vector $\boldsymbol{\alpha}$ satisfying $\left(\left| \frac{||\mathbf{Z}\boldsymbol{\alpha}||_2^2}{||\boldsymbol{\alpha}||_2^2} \leq \frac{1}{2} \right| \geq \frac{\epsilon}{2} \right)$ is given by $2(eN/k)^k e^{-M^2\epsilon^2/4c^2} = \mathcal{O}(k\log(N/k))$. Here, we use the Sterling's approximation, which states that $k! \geq (k/e)^k$. Thus, it leads to $\binom{N}{k} \leq (eN/k)^k$. Finally, we conclude that the probability of \mathbb{Z}^R satisfying the RIP for all *k*-sparse vector $\boldsymbol{\alpha}$ approaches $\frac{1}{2}$. Similarly, we obtain the same result for \mathbb{Z}^I .

Remark 2 (trade-off between the MSE and the system delay). There exists a trade-off between the system delay and the approximation error, which is described as follows. For an allowable mean-square error (MSE) $\xi > 0$, the achievable system delay $\mathcal{D}(\xi)$ is defined as

$$\mathcal{D}(\xi) \triangleq \min_{M} M \quad s.t. \quad \mathbb{E}(||\widehat{\alpha} - \alpha||_{2}^{2}) \le \xi,$$
(33)

where ξ relates to the bounded error in Equation (6).

Thus, the total energy consumption for all sensors is $D(\xi) \times \sum_{j=1}^{N} \frac{\sqrt{N}}{\rho_{ij}}$. In order to minimize the total network energy consumption, ρ_{ij} should be chosen to be as large as possible, or maximizing p_{ij}^{opt} as shown in Equation (15), which leads to the following problem.

Remark 3 (throughput maximum problem). The optimal power allocation p_{ij}^{opt} in Equation (15) can be obtained by solving the maximum output problem [19], which is given by

$$\max_{p_{ij}} \sum_{i=1}^{M} C_{ij}$$

s.t. $\sum_{k=1}^{i} p_{kj} \leq \sum_{k=1}^{i-1} E_{kj}, k = 1, \cdots, M,$
 $\sum_{k=0}^{i} E_{kj} - \sum_{k=1}^{i} p_{kj} \leq P_{\max}, k = 1, \cdots, M-1,$
 $p_{ij} \geq 0.$ (34)

Here, $C_{ij} = \sum_{i=1}^{M} \log_2 \left[1 + \frac{p_{ij}|h_{ij}|^2}{\sum_{l=1,l\neq j}^n |h_l|^2 E_{il} + \sigma_w^2} \right]$ and the value P_{\max} is a constant that depends on the hardware limitations. The above problem can be efficiently solved by using the iterative resource allocation algorithm method [19].

5. Simulation Results

We now present the results of a number of numerical simulations that illustrate the effectiveness of our approach. All simulations are performed in MATLAB R2015a (version 8.5.0.197613 (R2015a), The MathWorks Inc., Seoul, Korea) on a 3.60 GHz Intel Core i7 machine with 8 GB of RAM. We use MATLAB codes of the competing algorithms for our numerical studies. The vector **x** was assumed to be uniformly distributed in the interval [1, 10]. In our work, we used the basis pursuit de-noising algorithm [20] to compute the sparse solution in Equation (4). We evaluate the performance based on the MSE, which is given by

$$MSE = \mathbb{E}\left\{\frac{||\widehat{\mathbf{x}} - \mathbf{x}||_2}{||\mathbf{x}||_2}\right\}.$$
(35)

Figure 1 plots the MSE versus the compression ratio M/N for support cardinality k and fading channels plotted for N = 100, $\sigma_w^2 = 1$, and E_{ij} is uniformly distributed in the interval [0 dB, E_{max}], where $E_{max} = 2$ dB and $M \ge O(k \log N)$ to guarantee a stable recovery [6]. When $E_{ij} = 0$ dB, ρ_{ij} was set at 0.25 as the conventional baseline. In order to minimize the total energy consumption, we can perform the power allocation among different transmission time slots subject to the causality of the harvested energy, which refers to the resource allocation problem with energy constraint. Note that the sparsity level in Equation (15) is still adaptive since the optimal power allocations p_{ij}^{opt} obtained by solving the maximum output problem [1] are dynamic. The MSE values decrease as k decreases, as expected. We observed that the proposed scheme performs well compared to the conventional ones with AWGN and Rayleigh fading channels. The performance gap between those schemes is getting smaller when the ratio M/N increases. This is because when M is large enough, the MSE may not achieve any improvements. We notice that the sparsity level of the random projection determines the amount of communication. Increasing the sparsity level yields to decrease the preprocessing cost but unfortunately increase the latency to recover a CS approximation. We will show this trade-off in the simulation results.



(b) k = 5

Figure 1. MSE versus different compressed ratios with fading channels.

Figure 2 shows the outage probability with several ρ_{ij} intervals when k = 5 and N = 100. The outage probability is defined as the probability that the matrix **Z** does not satisfy the RIP, which scales as $\exp\{-M\epsilon^2/4c^2\}$ as shown in Equation (32). For sufficiently large number M, we observe that the optimal compressed rate M/N decreases as the sparsity level ρ_{ij} increases. This means the number of measurements M must approximately obey O(klog(N/k)) for effective CS recovery, while it is large enough for minimizing the outage probability. Moreover, from the result in Section 4.2, since M is proportional to ρ_{ij} , the larger value of ρ_{ij} leads to a smaller compressed ratio M/N for a fixed N.



Figure 2. Outage probability with different sparsity levels.

Figure 3 illustrates a trade-off between the system delay and the MSE threshold ξ for the proposed approach as we have discussed in Remark 2 when k = 5 and N = 100. We observed that the proposed scheme achieves a better trade-off when either SNR or ξ increases as expected. This is because higher SNR means the signal is more clearly readable, the CS recovery procedure will be much easier. Moreover, for a tight MSE threshold, the procedure of choosing the estimate to minimize the expected MSE will take longer, since the best MSE scaling depends on the value of *M*.



Figure 3. System delay for several allowable MSE thresholds.

Remark 4 (sparsity level option). This scheme is developed for transformative sensing mechanisms, which can be used in conjunction with current or upcoming EH capabilities in order to enable the deployment of energy neutral EH WSNs with practical network lifetime and improve data gathering rates. However, the

sparsity level in Equation (15) should be carefully chosen to maintain a good trade-off between the MSE and the system complexity. For example, when the channel condition is not good, we should select ρ_{ij} large enough (e.g., $\rho_{ij} \ge 1/4$) to guarantee an acceptable MSE.

6. Conclusions

In this paper, we have aimed to address the problem of recovering a sparse signal observed by a resource constrained in EH-WSNs for optimal data transmission strategy. By exploiting sparse random projections, there are significant reductions of the data measurements to be made. First, we studied a basis representation that can make the measurement matrix sufficiently sparse. The EH sensors store the sparse random projections of data, and thus the FC can estimate using compressive sensing with a sufficient number of measurements of sensors. Due to fading channels, the sparsity level can be adaptively chosen according to the available harvested energy at each EH sensor. This approach provides a better trade-off of the query latency and the desired approximation error, and also speeds up the processing time. We plan to generalize this concept in future work to incorporate sparsity of user activity and imperfect channel information as well. In addition, we would to like to emphasize that there are many ideas in the literature that would certainly enhanced our proposed scheme. We mentioned a few such possibilities as the following. First, we limited our analysis on AF protocol while different approach such as applying channel coding, and then using a modulation scheme for data transmission can be a huge open field. Second, approximation recovery for imperfect data via different norm, e.g., ℓ_q -norm ($0 < q \le 1$), can be promising due to its high quality of solutions and various types of sensing matrices that can be used in the CS reconstruction algorithms. Third, it needs to further discuss how to design the sparsity parameter of the random projection matrix based on different channel fading statistics so that the number of measurements required for signal recovery at the FC is minimized. Finally, for many application of interests, we often have prior information on addition constraints, e.g., rate-energy trade-off for simultaneous information and power transfer in EH-WSNs. Thus, other sparse random projections can be provided according to those constraints.

Acknowledgments: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (2014R1A5A1011478).

Author Contributions: The work was realized with the collaboration of all of the authors. Thu L. N. Nguyen contributed to the main results and code implementation. Yoan Shin, Jin Young Kim, and Dong In Kim organized the work, provided the funding, supervised the research and reviewed the draft of the paper. All authors discussed the results, approved the final version, and wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AWGN	Additive white Gaussian noise
AF	Amplify-and-forward
CS	Compressive sensing
DF	Decode-and-forward
EH	Energy harvesting
FC	Fusion center
i.i.d.	Independent and identically distributed
MAC	multiple-access channel
MSE	Mean-square error
JL	Johnson–Lindenstrauss
RIP	Restricted isometry property
SNR	Signal-to-noise ratio
s.t.	Subject to
WSN	Wireless sensor network

Appendix A. Proofs of Equations (21)–(25)

First, the mean and the variance of $u_{i_0}^{(\ell)}$ are calculated as follows:

$$\mathbb{E}[u_{i_{0}}^{(\ell)}] = \mathbb{E}\left[\sum_{j=1}^{2N} (\hat{b}_{i_{0}j}^{(\ell)})^{2} \psi_{j} x_{j} + \sum_{j \neq j_{0}} \hat{b}_{i_{0}j}^{(\ell)} \hat{b}_{i_{0}j_{0}}^{(\ell)} \psi_{j} x_{j_{0}}\right]$$

$$= \sum_{j=1}^{2N} \mathbb{E}\left[(\hat{b}_{i_{0}j}^{(\ell)})^{2} \right] \psi_{j} x_{j} + \sum_{j \neq j_{0}} \mathbb{E}\left[\hat{b}_{i_{0}j}^{(\ell)} \right] \mathbb{E}\left[\hat{b}_{i_{0}j_{0}}^{(\ell)} \right] \psi_{j} x_{j_{0}}$$

$$= \boldsymbol{\psi}^{T} \mathbf{x};$$
(A1)

$$\mathbb{E}\left[(u_{i_{0}}^{(\ell)})^{2}\right] = \mathbb{E}\left[\left(\sum_{j=1}^{2N} (\hat{b}_{i_{0}j}^{(\ell)})^{2}\psi_{j}x_{j}\right)^{2} + \left(\sum_{j\neq j_{0}} \hat{b}_{i_{0}j}^{(\ell)} \hat{b}_{i_{0}j}^{(\ell)}\psi_{j}x_{j_{0}}\right)^{2} + 2\left(\sum_{j=1}^{2N} (\hat{b}_{i_{0}j}^{(\ell)})^{2}\psi_{j}x_{j}\right)\left(\sum_{j\neq j_{0}} \hat{b}_{i_{0}j}^{(\ell)} \hat{b}_{i_{0}j}^{(\ell)}\psi_{j}x_{j_{0}}\right)\right]$$
$$= \sum_{j=1}^{2N} \mathbb{E}\left[(\hat{b}_{i_{0}j}^{(\ell)})^{4}\right]\psi_{j}^{2}x_{j}^{2} + 2\sum_{j
$$+ 2\sum_{j(A2)$$$$

$$=\sum_{j=1}^{2N} \frac{3\sqrt{N}}{\rho_{i_0j^{(\ell)}}} \psi_j^2 x_j^2 + 2\sum_{j \neq j_0} \psi_j x_j \psi_{j_0} x_{j_0} + \sum_{j \neq j_0} \psi_j^2 x_{j_0}^2$$

$$= 2\left(\sum_{j=1}^{2N} \psi_j^2 x_j^2 + \sum_{j \neq j_0} \psi_j x_j \psi_{j_0} x_{j_0}\right) + \left(\sum_{j=1}^{2N} \psi_j^2 x_j^2 + \sum_{j \neq j_0} \psi_j^2 x_{j_0}^2\right) + \sum_{j=1}^{2N} \left(\frac{3\sqrt{N}}{\rho_{i_0j^{(\ell)}}} - 3\right) \psi_j^2 x_j^2$$

$$= 2(\psi^T \mathbf{x})^2 + ||\psi||_2^2 ||\mathbf{x}||_2^2 + \sum_{j=1}^{2N} \left(\frac{3\sqrt{N}}{\rho_{i_0j}^{(\ell)}} - 3\right) x_j^2 \psi_j^2.$$

$$\operatorname{Var}[u_{i_0}^{(\ell)}] = \mathbb{E}\left[(u_{i_0}^{(\ell)})^2\right] - \left(\mathbb{E}[u_{i_0}^{(\ell)}]\right)^2$$

$$= (\psi^T \mathbf{x})^2 + ||\psi||_2^2 ||\mathbf{x}||_2^2 + \sum_{j=1}^{2N} \left(\frac{3\sqrt{N}}{\rho_{i_0j}^{(\ell)}} - 3\right) x_j^2 \psi_j^2.$$
(A3)

Similarly, we have

$$\mathbb{E}[v_{i_0}^{(\ell)}] = \left(\sum_{j=1}^{2N} \mathbb{E}[\hat{b}_{i_0 j}^{(\ell)}]\psi_j\right) \mathbb{E}[w_{i_0}] = 0, \tag{A4}$$

$$\mathbb{E}[(v_{i_0}^{(\ell)})^2] = \mathbb{E}\left[\left(\sum_{j=1}^{2N} \hat{b}_{i_0j}^{(\ell)} \psi_j\right)^2 w_{i_0}^2\right] = \left(\sum_{j=1}^{2N} \mathbb{E}[(\hat{b}_{i_0j}^{(\ell)})^2] \psi_j^2 + 2\sum_{j \neq j_0} \mathbb{E}[\hat{b}_{i_0j}^{(\ell)}] \mathbb{E}[\hat{b}_{i_0j_0}^{(\ell)}] \psi_j \psi_{j_0}\right) \sigma_w^2$$

$$= \left(\sum_{j=1}^{2N} \psi_j^2\right) \sigma_w^2 = \sigma_w^2 ||\psi||^2,$$
(A5)

$$\operatorname{Var}[v_{i_0}^{(\ell)}] = \mathbb{E}[(v_{i_0}^{(\ell)})^2] - \left(\mathbb{E}[v_{i_0}^{(\ell)}]\right)^2 = \sigma_w^2 ||\psi||^2.$$
(A6)

The covariance of $u_{i_0}^{(\ell)}$ and $v_{\overline{i}_0}^{(\ell)}$ is obtained as

$$\begin{aligned} \operatorname{Cov}[u_{i_{0}}^{(\ell)}, v_{\bar{i}_{0}}^{(\ell)}] &= \mathbb{E}\left[u_{i_{0}}^{(\ell)}v_{\bar{i}_{0}}^{(\ell)}\right] - \mathbb{E}[u_{i_{0}}^{(\ell)}]\mathbb{E}[v_{\bar{i}_{0}}^{(\ell)}] \\ &= \mathbb{E}\left[\left(\sum_{j=1}^{2N}\hat{b}_{i_{0}j}^{(\ell)}\psi_{j}\right)\left(\sum_{j=1}^{2N}\hat{b}_{i_{0}j}^{(\ell)}x_{j}\right)\left(\sum_{j=1}^{2N}\hat{b}_{\bar{i}_{0}j}^{(\ell)}\psi_{j}\right)\right]\mathbb{E}[w_{\bar{i}_{0}}] - \mathbb{E}[u_{i_{0}}^{(\ell)}]\mathbb{E}[v_{\bar{i}_{0}}^{(\ell)}] \\ &= 0 \quad \text{since} \quad \mathbb{E}[w_{\bar{i}_{0}}] = 0 \quad \text{and} \quad \mathbb{E}[v_{\bar{i}_{0}}^{(\ell)}] = 0. \end{aligned}$$

References

- 1. Ho, C.-K.; Zhang, R. Optimal energy allocation for wireless communications with energy harvesting constraints. *IEEE Trans. Signal Process.* **2012**, *60*, 4808–4818.
- 2. Sharma, V.; Mukherji, U.; Joseph, V.; Gupta, S. Optimal energy management policies for energy harvesting sensor nodes. *IEEE Trans. Wirel. Commun.* **2010**, *9*, 1326–1336.
- 3. Lu, X.; Wang, P.; Niyato, P.; Kim, D.I.; Han, Z. Wireless networks with RF energy harvesting: A contemporary survey. *IEEE Commun. Surv. Tutor.* **2014**, *17*, 757–789.
- 4. Shaikh, F.K.; Zeadally, S. Energy harvesting in wireless sensor networks: A comprehensive review. *Renew. Sustain. Energy Rev.* 2016, *55*, 1041–1054.
- 5. Dong, M.; Ota, K.; Liu, A. RMER: Reliable and energy-efficient data collection for large-Scale wireless sensor networks. *IEEE Internet Things J.* 2016, *3*, 511–519.
- 6. Candès, E.J.; Romberg, J.; Tao, T. Stable signal recovery from incomplete and inaccurate measurements. *Commun. Pure Appl. Math.* **2006**, *59*, 1207–1223.
- 7. Baraniuk, R.G.; Cevher, V.; Duarte, M.F.; Hegde, C. Model-based compressive sensing. *IEEE Trans. Inf. Theory* **2010**, *56*, 1982–2001.
- 8. Davenport, M.A.; Boufounos, P.T.; Wakin, M.B.; Baraniuk, R.G. Signal processing with compressive measurements. *IEEE J. Sel. Top. Signal Process.* **2010**, *4*, 445–460.
- Wang, W.; Garofalakis, M.; Ramchandran, K. Distributed sparse random projections for refinable approximation. In Proceedings of the 6th International Conference on Information Processing in Sensor Networks, Cambridge, MA, USA, 25–27 April 2007; pp. 331–339.
- Rana, R.; Hu, W.; Chou, C. Energy-aware sparse approximation technique (EAST) for rechargeable wireless sensor networks. In Proceedings of the 7th European Conference on Wireless Sensor Networks, EWSN 2010, Coimbra, Portugal, 17–19 February 2010; pp. 306–321.
- 11. Yang, G.; Tan, V.Y.F.; Ho, C.-K.; Ting, S.H.; Guan, Y.L. Wireless compressive sensing for energy harvesting sensor nodes. *IEEE Trans. Signal Process.* **2013**, *61*, 4491–4505.
- Li, P.; Hastie, T.R.; Church, K.W. Very sparse random projections. In Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Philadelphia, PA, USA, 20–23 August 2006; ACM: New York, NY, USA, 2006; pp. 287–296.
- 13. Bajwa, W.U.; Haupt, J.D.; Sayeed, A.M.; Nowak, R.D. Joint source channel communication for distributed estimation in sensor networks. *IEEE Trans. Inf. Theory* **2007**, *53*, 3629–3653.
- 14. Gastpar, M. Uncoded transmission is exactly optimal for a simple Gaussian sensor network. *IEEE Trans. Inf. Theory* **2008**, *54*, 5247–5251.
- 15. Marano, S.; Matta, V.; Tong, L.; Willett, P. A likelihood-based multiple access for estimation in sensor networks. *IEEE Trans. Signal Process.* **2007**, *55*, 5155–5166.
- 16. Rubinstein, R.; Bruckstein, A.M.; Elad, M. Dictionaries for sparse representation modeling. *Proc. IEEE* **2010**, *98*, 1045–1057.
- 17. Dasgupta, S.; Gupta, A. An elementary proof of the Johnson-Lindenstrauss lemma. *Random Struct. Algorithms* **2003**, *22*, 60–65.
- Davenport, M. Random Observations on Random Observations: Sparse Signal Acquisition and Processing. Ph.D. Thesis, Rice University, Houston, TX, USA, 2010.
- 19. Ng, D.W.K.; Lo, E.S.; Schober, R. Wireless information and power transfer: Energy efficiency optimization in OFDMA systems. *IEEE Trans. Wirel. Commun.* **2013**, *12*, 6352–6370.
- Berg, E.V.D.; Friedlander, M.P. Probing the Pareto frontier for basis pursuit solutions. *Proc. Soc. Ind. Appl. Math.* 2008, 31, 890–912.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).