Article

A Non-Convex Compressed Sensing Model Improving the Energy Efficiency of WSNs for Abnormal Events’ Monitoring

Yilin Huang, Haiyang Li* and Jigen Peng

School of Mathematics and Information Science, Guangzhou University, Guangzhou 510006, China
* Correspondence: fplihaiyang@126.com

Abstract: The wireless sensor network (WSN), a communication system widely used in the Internet of Things, usually collects physical data in a natural environment and monitors abnormal events. Because of the redundancy of natural data, a compressed-sensing-based model offers energy-efficient data processing to overcome the energy shortages and uneven consumption problems of a WSN. However, the convex relaxation method, which is widely used for a compressed-sensing-based WSN, is not sufficient for reducing data processing energy consumption. In addition, when abnormal events occur, the redundancy of the original data is destroyed, which makes the traditional compressed sensing methods ineffective. In this paper, we use a non-convex fraction function as the surrogate function of the $\ell_0$-norm, which achieves lower energy consumption of the sensor nodes. Moreover, considering abnormal event monitoring in a WSN, we propose a new data construction model and apply an alternate direction iterative thresholding algorithm, which avoids extra measurements, unlike previous algorithms. The results showed that our models and algorithms reduced the WSN’s energy consumption during abnormal events.

Keywords: wireless sensor network; monitoring abnormal events; energy-efficient compressed data gathering model; compressed sensing; alternating direction iteration fraction penalty thresholding algorithm

1. Introduction

Wireless sensor network (WSN), an important communication system in the Internet of Things, is widely used in fields such as the environment, medical care, transportation, and security [1,2]. Its basic structure consists of a large number of wireless sensor nodes and a sink: the nodes collect and transmit data to the sink, and the sink summarizes and processes the data. The energy of wireless sensor nodes usually depends on batteries or solar power, while the energy reserve of the sink is usually sufficient, which makes the energy consumption of wireless sensor nodes one of the key issues in a WSN [3]. In recent years, with the emergence of compressed sensing theory [4], signal processing methods in a WSN based on sparse representation have already attracted widespread attention [5–8]. Generally speaking, signal acquisition in a WSN can be expressed as a linear measurement:

$$y = \Phi d,$$

where $y$ is the measurement value obtained by the wireless sensor node sampling; $d$ is the original data; the measurement matrix $\Phi$ with a size of $M \times N$ is an encoding process of the original data $d$. It is assumed in compressed sensing theory that the data are redundant, which can be sparsely represented by a transform $\Psi$ satisfying $d = \Psi x$, where $x$ is a sparse vector. The compressed sensing theory shows that $d$ can be reconstructed by solving an optimization model. It is worth noting that sampling is operated by the nodes and that reconstruction is operated by the sink, which means that compressed sensing transfers some workload from the nodes to the sink, and it is suitable for the energy distribution.
characteristic of a WSN. Additionally, when the WSN is monitoring abnormal events, the sparsity of signals will be disturbed, i.e., signals are no longer sparse in the original transform domain, which makes it impossible to reconstruct signals using the original compressed sensing model. In general, the so-called abnormal events appear sparsely and sporadically; hence, they are naturally sparse in a WSN. In [6], a feasible approach was proposed that regards the original signal as the summation of the normal and abnormal events’ signals: the normal signal is sparse under the sparse transform domain, and the abnormal events’ signal is sparse under the identical transform domain, so the splicing signal in this kind of transform \((\Psi|I)\) is sparse.

The studies [9–12] discussed the feasibility and application of compressed sensing in signal acquisition and processing, demonstrating that a WSN can effectively reduce network energy consumption by using compressed sensing. The compressed sensing theory was applied to a WSN for the first time, and in [5], the concept of compressed wireless sensing (CWS) was proposed and found to be effective for energy efficiency in a WSN of a certain structure. In [6], a compressive data gathering (CDG) model was proposed, which opened a new path for data gathering. In [7], the authors discussed the feasibility of using a binary matrix as the measurement matrix to improve CDG, where the wireless sensor nodes did not need multiplication, further reducing computational consumption. Based on CDG, the researchers [8] developed a distributed compressed sensing method by dividing the data into several blocks for reconstruction with the idea of dividing-and-conquering and then summarizing the data to the sink.

Although compressed sensing theory and its improvements have achieved excellent results, there are still problems with the models in the above-mentioned research:

1. These studies of compressed sensing models usually use the \(\ell_1\)-norm as the convex surrogate form of the \(\ell_0\)-norm. Even though it is typical and popular [13,14], the \(\ell_1\)-norm often fails to produce the sparsest solution or select the position of the non-zero coefficient [15,16], which implies that it may not be the best choice for a WSN. In fact, many studies have shown that non-convex surrogate functions achieve a better approximation than the \(\ell_0\)-norm and a lower sampling rate than the \(\ell_1\)-norm [15–18].

2. A conventional and effective approach to focusing on abnormal events’ monitoring in a WSN is to treat abnormal events as sparse vectors under identity transformation, just as [6] did, to reconstruct additional abnormal event signals with the same length as normal signals. However, this method requires the reconstruction of a signal with twice the dimension of the original signal, which means more measurements and energy consumption.

This paper mainly discusses the two problems mentioned above. First, we adopted the non-convex fraction function \(P_a(x)\) as the surrogate form of the \(\ell_1\)-norm in the compressed sensing model, where

\[
P_a(x) = \sum_{i=1}^{N} p_a(x_i) = \sum_{i=1}^{N} \frac{a|x_i|}{1 + a|x_i|} (a \geq 0).
\]

Following estimates of the measurement number \(M\) in compressed sensing theory, we thought that the non-convex fraction function model could decrease the number of measurements significantly compared with the \(\ell_1\)-norm model, implying that the wireless sensor nodes in our model have a lighter workload and consume less energy. Second, regarding abnormal events’ monitoring, we propose an optimization model based on the non-convex fraction function and an alternate direction iterative fraction penalty thresholding (AD-IFPT) algorithm to solve it. In each iteration, the normal and abnormal events signals were all \(N\)-dimensional and alternately reconstructed. Meanwhile, we analyzed the AD-IFPT algorithm and verified its convergence. Based on the minimal number of measurements theory in compressed sensing and the results of our numerical experiment, we concluded that the AD-IFPT algorithm requires fewer measurements compared with other traditional approaches, meaning that its energy efficiency in abnormal events’ monitoring is better than that of other methods.
The rest of this paper is organized as follows. In Section 2, we briefly introduce the theory of compressed sensing and its applications in a WSN. In Section 3, we introduce a compressed sensing optimization model based on a specific non-convex surrogate function, the fraction function. Then, focusing on abnormal events’ monitoring, we designed an alternating direction algorithm, which has lower energy consumption and high efficiency in abnormal events’ monitoring. We also give several conclusions about the properties of the proposed algorithm. The results of the simulation experiments are presented in Section 4. Finally, we give our conclusion in Section 5.

2. Preliminary Knowledge

As described above, compressed sensing theory can reduce the energy consumption of signal acquisition and transmission, but increase the energy consumption of signal decompression, which just meets the requirements of low-energy consumption at wireless sensor nodes and unlimited energy consumption at the sink in a WSN. In this section, we introduce the compressed sensing theory and the definition of abnormal events to be able to discuss the proposed model clearly in the next section and to choose two applications of compressed sensing theory to evaluate the performance of the proposed model in the experiment section.

2.1. Compressed Sensing Theory

In traditional signal processing, obtaining an $N$-dimensional signal requires at least $N$ measurements. However, compressed sensing theory shows that, if an original signal $d$ can be sparsely represented, that is, if there exists a sparse transform $\Psi$ such that

$$d = \Psi x,$$

where $x$ is a sparse vector, meaning that most entries of $x$ are zero, then we can reconstruct the original signal $d$ with fewer measurement times $M$ by solving the following optimization problems:

$$(P_0) \quad \min_x \|x\|_0 \text{ s. t. } y = \Phi \Psi x,$$

where $y$ is the measurement value and $\Phi$ is the measurement matrix with the size of $M \times N$ and $M \ll N$. Because of the discrete and discontinuous properties of the $\ell_0$-norm, the $P_0$ problem is actually nondeterministic polynomial time (NP) hard [19]. Moreover, the $\ell_0$-norm is very sensitive to noise, which makes the model $(P_0)$ problematic for solving this problem. To overcome these difficulties, we replaced the highly discontinuous $\ell_0$-norm with a continuous surrogate function [13,14,18,20–23], in which the $\ell_1$-norm is typical and popularly proposed. The constrained problem corresponding to the $\ell_1$-norm is

$$(P_1) \quad \min_x \|x\|_1 \text{ s. t. } y = \Phi \Psi x.$$

Given a signal with noise, the corresponding regularization problem is

$$(P_{\lambda 1}) \quad \min_x \lambda \|x\|_1 + \|\Phi \Psi x - y\|_2^2,$$

where $\lambda > 0$, the regularized parameter, represents a trade-off between reconstruction error and sparsity. It is worth noting that $(P_1)$ and $(P_{\lambda 1})$ are convex optimization problems, which can be solved by many methods such as basis pursuit [24], $\ell_1$-magic [25], an iterative soft thresholding algorithm [4,26], and split Bregman methods [27,28].

In fact, some studies have shown that non-convex surrogate functions can obtain a lower sampling rate than the $\ell_1$-norm [29,30], which is extremely important in a WSN because it means that there is a lower energy consumption of signal acquisition and transmission at the nodes. In Section 3, we discuss the energy consumption of signal
acquisition and transmission at the nodes and the abnormal events’ monitoring of a WSN by taking a specific non-convex surrogate function, the fraction function.

2.2. Analog Compressed Sensing

Using compressed sensing to measure analog signals has been applied in many fields [9–12]. The main idea is to obtain $M$ measured values of an $N$-dimensional signal in an analog domain based on the random demodulator (RD) architecture (see Figure 1). The RD consists of three stages: mixer, integrator, and the analog-to-digital converter (ADC). In this case, there are three steps to signal acquisition:

1. The mixer multiplies an input signal by a random sequence.
2. The integrator accumulates the output voltage of the mixer.
3. The ADC samples it at a rate of $1/M$.

![Figure 1](image1.png)

Figure 1. A classical random demodulator (RD) architecture in an analog compressed sensing application.

Step 1 corresponds to the linear measurement represented by the inner product of the original signal $x$ with each row of the measurement matrix $\Phi$. Step 2 gives the measurement values, and the sampling rate of Step 3 is lower than the Nyquist rate when $M \ll N$, which implies that it will use less energy compared to traditional signal sampling methods. To implement compressed sensing, a prerequisite is that analog signals be sparsely represented as (3). In applications, for example, natural images can usually be sparsely represented under the discrete cosine transform (DCT) (see Figure 2), and [2] showed that the ocean temperature measured by a WSN can be sparsely represented under a discrete wavelet transform (DWT).

![Figure 2](image2.png)

Figure 2. (A) Original image; (B) coefficient matrix under discrete cosine transform (DCT); (C) compressed image using 5% largest coefficient; (D) the 5% largest coefficient matrix.

2.3. Compressed Data Gathering in a Multi-Hop WSN

In [6], a CDG model was proposed, which applies compressed sensing to a multi-hop WSN. In the traditional wireless sensor signal transmission model, each sensor node collects
signals and transmits them to the nearest node in the direction of the sink. If the sensor receives signals from other sensors, those will also be sent to the next node. With the increase in sensor network scale, the power consumption of this data transmission mode will increase greatly. Moreover, the energy consumption between sensors is extremely uneven because, the closer a sensor is to the sink, the more signals it receives and the more transmission energy it consumes.

To overcome the poor performance of the traditional multi-hop WSN architecture, the CDG model was proposed based on compressed sensing theory [6]. In this model, each node only needs to multiply its reading by a random number, add it to the upstream node reading, and transmit the result to the next node, which avoids extra transmitting of data in the multi-hop architecture. During data sampling, the entire wireless sensor network needs to repeat the above steps \( M \) times and the data received at the sink can be represented by

\[
\begin{pmatrix}
y_1 \\
y_2 \\
\vdots \\
y_M
\end{pmatrix} = \begin{pmatrix}
\Phi_{11} & \Phi_{12} & \cdots & \Phi_{1N} \\
\Phi_{21} & \Phi_{22} & \cdots & \Phi_{2N} \\
\vdots & \vdots & \ddots & \vdots \\
\Phi_{M1} & \Phi_{M2} & \cdots & \Phi_{MN}
\end{pmatrix} \begin{pmatrix}
d_1 \\
d_2 \\
\vdots \\
d_M
\end{pmatrix},
\]

where \( y_i \) represents the \( i \)th measurement readings of the sink with \( M \ll N \); \( \Phi_{ij} \) represents the random coefficient of the \( j \) node in the \( i \)th measurement; \( d_j \) represents the readings of the \( j \) node for \( i = 1, 2, \ldots, M \). We set \( d = \Psi x \) as the sparse representation of \( d \), which made this model the same as the compressed sensing model (4). Its working model is shown in Figure 3.

**Figure 3.** (A) Traditional transmitting method of a multi-hop WSN; (B) CS method of a multi-hop WSN.

### 2.4. Abnormal Events’ Monitoring

WSNs are widely used for monitoring abnormal events. For example, when a WSN collects temperature information in a forest, abnormally high readings may indicate a fire. In this case, when abnormal readings appear on nodes, the sparsity of the data will disappear, and the data can no longer be sparsely represented under the original sparse transform (see Figure 4). If it is determined that abnormal events do not appear on a large
scale, then the abnormal readings are considered to be sparse in all node readings in the WSN. For this reason, assuming the original signal contains both normal and abnormal events, we express the sensor readings as

\[ d = d_0 + d_s, \]  

(7)

where \( d_0 \) represents the original readings and \( d_s \) represents the abnormal readings. Since abnormal readings are sparse and the original signal is sparse under the sparse transform \( \Psi \), then

\[ y = \Phi d, \quad d = d_0 + d_s = \Psi x_0 + Ix_s, \]  

(8)

where \( d_0 = \Psi x_0; d_s = x_s; x_0 \) and \( x_s \) are sparse vectors. A reasonable choice was given in [6], which combines the matrices \( \Phi, \Psi, \) and \( I \) as a new measurement matrix of the following form:

\[ \min \| x \|_0 \quad \text{s. t.} \quad y = \Phi d, d = (\Psi I)x, \]

\[ x = \begin{pmatrix} x_0^T \\ x_s^T \end{pmatrix}^T. \]  

(9)

If \( \Psi = (\Psi I) \), then it is the original compressed sensing model as in (4). However, the dimension of the signal is doubled, which presents a challenge for reconstructing a signal with abnormal readings using fewer measurements. In the following discussion, we propose an alternative reconstruction method and an algorithm for abnormal event monitoring that does not need to solve the original signal with twice the length in the reconstruction process.

3. Non-Convex Compressed Sensing Model in a WSN and Alternating Direction Reconstruction Algorithm

In this section, we first introduce a compressed sensing optimization model based on a specific non-convex surrogate function, the fraction function. Second, focusing on abnormal events’ monitoring, we propose an alternating direction algorithm that has lower energy consumption and high efficiency.
3.1. Fraction Function Abnormal Events Monitoring Model in WSN

In this part, we mainly introduce our proposed model, but first, we discuss the estimation of the minimum number of measurements in different compressed sensing models: the number $M$ of rows of the measurement matrix. The convex relaxation model (6) is the most widely used and derives its optimal solution easily because of the excellent properties of the convex function. According to the existing conclusions of compressed sensing theory, for the estimation of the minimum number of measurements in (6), the following inequalities hold:

$$M \geq CK \ln\left(\frac{N}{K}\right),$$

where $N$ is the number of columns of the measurement matrix; $K$ is the sparsity level of the original data; $C$ is a constant number [31]. In applications, $M$ is directly related to the sampling times of the sensor networks. To achieve better energy consumption, a lower bound on $M$ is required. Since the energy at the sink is sufficient, the additional effort for the reconfiguration is negligible. Therefore, we sacrificed the excellent performance of the convex function in optimization to choose the fraction non-convex function as a surrogate function to achieve better performance for the minimum number of measurements.

Let the fraction penalty compressed sensing model be set $(P_{\alpha})$ as

$$\min_x P_{\alpha}(x) \quad \text{s. t.} \quad y = \Phi d, \quad d = \Psi x,$$  \hspace{1cm} (11)

where $P_{\alpha}(x)$ is defined in (2). Although there is no rigorous theoretical proof, numerical experiments showed that the minimum number of measurements required for a non-convex fraction function is lower than for convex functions in compressed sensing, which implies that it can recover data with less sampling. The comparison of the fraction penalty function with another non-convex function the $\ell_{0.5}$-norm was explained in [17] in detail, which showed that the fraction penalty function has better sparse approximation and a lower number of measurements than the $\ell_{0.5}$-norm. The better properties of the fraction penalty function are helpful to achieve better energy efficiency in a WSN.

After having introduced abnormal events’ monitoring in a WSN, let us now give its fraction penalty form for abnormal events’ reconstruction. We assumed that the WSN signal $d$ could be split into a normal signal $d_0$ and an abnormal events–readings vector $d_s$ as in (7). $d_0$ has its sparse representation $d_0 = \Psi x_0$, and $d_s$ sparsely appears in the WSN such that $d_s = I x_s$ as (8), where $x_0$ and $x_s$ are $K_1$-sparse and $K_2$-sparse vectors, respectively. Thus, the abnormal events’ monitoring based on fraction compressed sensing is

$$\min_{x_0, x_s} P_{\alpha}(x_0) + P_{\alpha}(x_s) \quad \text{s. t.} \quad y = \Phi d, \quad d = \Psi x_0 + x_s,$$  \hspace{1cm} (12)

Next, we focus on the algorithm design of Problem (12).

3.2. Algorithm Design

Before we discuss the solution of the problem (14), we first discuss the shortcomings of the solution of such problems in previous studies. In [6], an available method was to splice the sparse transform and identity matrices into an $M \times 2N$ matrix, which regards the abnormal readings as a sparse vector under the identity transform. Hence, within abnormal events’ monitoring, it took the form:

$$\min_x \left\{ \| \Phi (\Psi I) x - y \|_2^2 + \lambda P_{\alpha}(x) \right\},$$

where $x = (x_0^T \quad x_s^T)^T$. It is worth noting that $x$ in this model is a $2N$-dimensional signal, which implied that we needed to reconstruct a $2N$-dimensional signal with $M$ measurements. In addition, the sparsity of the reconstructed vector $x$ became $K = K_1 + K_2$. Based on
increased signal dimension and sparsity led to a higher measurement number and more WSN energy consumption.

Now, let us introduce our approach. To solve (12), we considered its corresponding regularization problem:

\[
\min_x \left\{ \| \Phi \Psi x_0 + \Phi x_s - y \|^2_2 + \lambda_1 P_\mu(x_0) + \lambda_2 P_\mu(x_s) \right\},
\]

where \( \lambda_1 \) and \( \lambda_2 \) are regularization parameters, which are used to trade off the sparsity weight. Inspired by the idea of the alternating direction iteration algorithm, we propose an approach to reconstruct the original signal and abnormal events' vectors alternately. In each iteration, the iterative fraction penalty thresholding algorithm is used to solve the optimization problem (15) or (16). More specifically, we considered the following two optimization problems:

\[
\min_{x_0} \left\{ \| \Phi \Psi x_0 + \Phi x_s - y \|^2_2 + \lambda_1 P_\mu(x_0) \right\},
\]

and

\[
\min_{x_s} \left\{ \| \Phi \Psi x_0 + \Phi x_s - y \|^2_2 + \lambda_2 P_\mu(x_s) \right\}.
\]

To solve (15), we set \( x_s \) to be fixed and did the same to solve (16). Notice that (15) and (16) have a similar form, which could be made equivalent by a simple substitution. To solve (15), we set \( \hat{y} = y - \Phi x_S \); similarly, to solve (16), we set \( y_0 = y - \Phi \Psi x_0 \). Therefore, we only needed to consider the following suboptimization problem:

\[
(P_{\lambda_1}) \quad \min_x \lambda P_\mu(x) + \| \Phi \Psi x - y \|^2_2
\]

and then alternate iterations through the above substitution form.

We considered using the thresholding algorithm to solve the regularization problem, because in [17], it was proven that the optimal solution \( x^* = \{ x_1^*, x_2^*, \ldots, x_N^* \} \) to \( P_{\lambda_1} \) could be represented by the following threshold operator form:

\[
x_i^* = H_{\lambda,\mu}(z_\mu(x^*)),
\]

where \( t \) is a positive thresholding parameter; \( \mu \) satisfies \( 0 < \mu < \| \Phi \Psi \|^2_2 \); \( z_\mu(x^*) = x^* + \mu (\Phi \Psi)^T (y - \Phi \Psi x^*) \); the thresholding operator \( H \) has the following form:

\[
H_{\lambda,\mu}(x) = (H_{\lambda,\mu}(x_1), H_{\lambda,\mu}(x_2), \ldots, H_{\lambda,\mu}(x_N)),
\]

\[
H_{\lambda,\mu}(x_i) = \begin{cases} h_\lambda(x_i), & |x_i| > t \\ 0, & \text{otherwise} \end{cases}
\]

\[
h_\lambda(x_i) = \text{sgn}(x_i) \left( \frac{1 + a |x_i|}{3} \left( 1 + 2 \cos \left( \frac{G(x_i)}{3} - \frac{\pi}{3} \right) \right) - 1 \right),
\]

and

\[
G(x_i) = \arccos \left( \frac{27 \lambda a^2}{4(1 + a|x_i|)^3} - 1 \right).
\]

Hence, in the algorithm design, we had the following iteration form of \( x \):

\[
x^{k+1} = H_{\lambda,\mu}(z_\mu(x^k)).
\]
To sum up, we set the iterative scheme in the thresholding algorithm as follows:

Update $x_{0}^{k+1}$:

set $y_{0}^{k} = y - \Phi x_{0}^{k}$

$$
(x_{0}^{k+1})_{i} = H_{\lambda_{1}u_{1},t}(\{z_{u_{1}}(x_{0}^{k})\})
$$

(24)

Update $x_{s}^{k+1}$:

set $y_{s}^{k} = y - \Phi \Psi x_{0}^{k+1}$

$$
(x_{s}^{k+1})_{i} = H_{\lambda_{2}u_{2},t}(\{z_{u_{2}}(x_{s}^{k})\})
$$

(25)

Based on the above iterative scheme, the alternating direction fraction penalty algorithm (AD-IFPT) is given as Algorithm 1. In this alternating direction iteration model, each step solves a compressed sensing optimization problem of reconstructing an $N$-dimensional vector from $M$ measurements. This makes up the deficiency of (9), which doubles the signal dimension because of matrix splicing. In the subsequent simulation experiment, we saw that the AD-IFPT was effective and decreased the measurements more compared to the IFPT or traditional compressed sensing methods.

Algorithm 1 Alternating direction iteration fraction penalty thresholding algorithm.

Require: $A$, $y$, $\lambda_{1}$, $\lambda_{2}$, $x_{0}^{0}$, $x_{s}^{0}$, $u_{1}$, $u_{2}$, $t$, and $a$

If no convergence, do:

$y_{0}^{n} = y - \Phi x_{0}$

$z_{n}^{n} = x_{0}^{n} + u_{1}(\Phi \Psi)^{T}(y_{0}^{n} - \Phi \Psi x_{0}^{n})$

for $i = 1 : \text{length}(x_{0})$ do

if $|z_{n}^{n}| > t$ then

$x_{0}^{n+1} = g_{\lambda_{1}u_{1}}(z_{n}^{n})$

else

$x_{0}^{n+1} = 0$

end if

end for

$y_{s}^{n} = y - \Phi \Psi x_{0}$

$z_{n}^{n} = x_{s}^{n} + u_{2}\Phi^{T}(y_{s}^{n} - \Phi x_{s}^{n})$

for $i = 1 : \text{length}(x_{s})$ do

if $|z_{n}^{n}| > t$ then

$x_{s}^{n+1} = g_{\lambda_{2}u_{2}}(z_{n}^{n})$

else

$x_{s}^{n+1} = 0$

end if

end for

return $x_{0}^{n+1}, x_{s}^{n+1}$

3.3. Parameter Setting and Convergence Analysis

Next, we analyzed the convergence of the AD-IFPT. Beforehand, we needed to make use of a few auxiliary functions and introduce several lemmas that will be helpful in proving convergence.

Lemma 1 ([17]). Given $v \in \mathbb{R}^{N}$, the following equation holds:

$$
H_{\lambda_{t}}(v) = \arg \min_{x} \|x - v\|_{2}^{2} + \lambda P_{t}(x)
$$

(26)
where $t$ satisfies

$$
t = \begin{cases} 
\frac{\lambda}{2}, & \text{if } \lambda \leq \frac{1}{\sigma^2} \\
\sqrt{\lambda - \frac{1}{\sigma^2}}, & \text{if } \lambda > \frac{1}{\sigma^2}.
\end{cases}
$$

(27)

In our algorithm, we set the threshold parameter $t$ following the above condition.

**Lemma 2** ([17]). Given $\lambda > 0$, $u \in \left(0, \frac{1}{\sqrt{\Phi \Psi} + 1} \right)$, if $x^*$ is an optimal solution of ($P_{\lambda u}^*$), then $x^*$ is a fixed point of the threshold operator as (18).

The above two lemmas will be used in the convergence proof of the algorithm. Next, we defined several auxiliary functions to prove convergence.

$$R(x_0, x_0) := \|\Phi \Psi x_0 + \Phi x_0 - y\|_2^2 + \lambda_1 P_a(x_0) + \lambda_2 P_d(x_0)$$

(28)

$$B_{x_0}(x_0, v) := \left. \frac{d}{dx} \right|_{x=x_0} \left[ \|\Phi \Psi x_0 - (y - \Phi x_0)\|_2^2 + \lambda_1 P_a(x_0) + \lambda_2 P_d(x_0) 
- \|\Phi \Psi x_0 - \Phi \Psi v\|_2^2 \right] + \|x_0 - v\|_2^2$$

(29)

$$C_{x_0}(x_0, v) := \left. \frac{d}{dx} \right|_{x=x_0} \left[ \|\Phi x_0 - (y - \Phi x_0)\|_2^2 + \lambda_1 P_a(x_0) + \lambda_2 P_d(x_0) 
- \|\Phi x_0 - \Phi v\|_2^2 \right] + \|x_0 - v\|_2^2$$

(30)

**Theorem 1.** Suppose that $x_0^*$ is a minimizer of $B_{x_0}(x_0, v)$ for any fixed $\lambda_1$, $\lambda_2$, $u_1$, $x_0$, and $v$. Then,

$$\left( x_0^* \right)_i = H_{\lambda_1 u_1, 1} \left( z_{u_1} \left( v_i \right) \right).$$

(31)

**Proof.** Note that $B_{x_0}(x_0, v)$ can be written as

$$B_{x_0}(x_0, v) = u_1 \left( \|\Phi \Psi x_0 - (y - \Phi x_0)\|_2^2 + \lambda_1 P_a(x_0) + \lambda_2 P_d(x_0) 
- \|\Phi \Psi x_0 - \Phi \Psi v\|_2^2 \right] + \|x_0 - v\|_2^2$$

$$= u_1 \|y - \Phi x_0\|_2^2 + 2 u_1 \langle \Phi \Psi x_0, y - \Phi x_0 \rangle + \lambda_1 P_a(x_0) + \lambda_2 P_d(x_0) 
- u_1 \|\Phi \Psi v\|_2^2 + 2 u_1 \langle \Phi \Psi x_0, \Phi \Psi v \rangle + \|x_0\|_2^2 + \|v\|_2^2 - 2 \langle x_0, v \rangle$$

$$= \|x_0\|_2^2 - 2 \langle x_0, z_{u_1} \left( v \right) \rangle + \lambda_1 P_a(x_0) + \lambda_2 P_d(x_0) 
+ \|v\|_2^2 + u_1 \|y - \Phi x_0\|_2^2 - u_1 \|\Phi \Psi v\|_2^2$$

$$= \|x_0 - z_{u_1} \left( v \right)\|_2^2 + \lambda_1 P_a(x_0) + \lambda_2 P_d(x_0) + \|v\|_2^2 + u_1 \|y - \Phi x_0\|_2^2 - u_1 \|\Phi \Psi v\|_2^2 - u_1 \|z_{u_1} \left( v \right)\|_2^2.$$
Because \(\lambda_1, \lambda_2, u_1, x_s, \) and \(v\) are fixed, using Lemma 1 gives us
\[
(x_0^i) = H_{\lambda_1 u_1, l}(z_{u_1}(v_i)).
\]

\[\square\]

**Theorem 2.** Suppose that \(x_s^*\) is a minimizer of \(C_{x_0}(x, v)\) for any fixed \(\lambda_1, \lambda_2, u_2, x_0, \) and \(v\). Then,
\[
(x_s^*) = H_{\lambda_2 u_2, l}(z_{u_2}(v_i)).
\]

**Proof.** It is similar to Theorem 1, so it was omitted. \[\square\]

**Theorem 3.** Let \(u_1 \in \left(0, \frac{1}{\|\Phi\|_2}\right)\). Given that \(\lambda_1, \lambda_2 > 0\) for any fixed \(x_s\), suppose that \(x_0^*\) is the global minimizer of \(R(x_0, x_s)\). Then, for any \(x \in \mathbb{R}^N\), we have
\[
B_{x_s}(x_0^*, x_s^*) \leq B_{x_s}(x, x_s^*). \tag{33}
\]

**Proof.** Because of the properties of the operator norm, for any \(x \in \mathbb{R}^N\), the following inequality holds:
\[
\|\Phi\Psi x_0 - \Phi\Psi x_0^*\| \leq \|\Phi\Psi\|^2_2 \|x - x_0^*\|_2 \leq \frac{1}{u_1} \|x - x_0^*\|_2
\]

Hence, for any \(x\),
\[
B_{x_s}(x_0^*, x_s^*) = u_1 R(x_0^*, x_s) \leq u_1 R(x, x_s)
\]
\[
= u_1 \left(\|\Phi\Psi x - (y - \Phi x_s)\|_2^2 + \lambda_1 P_0(x) + \lambda_2 P_0(x_s)\right)
\]
\[
\leq u_1 \left(\|\Phi\Psi x - (y - \Phi x_s)\|_2^2 + \lambda_1 P_0(x) + \lambda_2 P_0(x_s)\right)
\]
\[
- u_1 \|\Phi\Psi x - \Phi\Psi x_0^*\|_2^2 + \|x - x_0^*\|_2^2
\]
\[
= B_{x_s}(x, x_s^*)
\]

\[\square\]

**Theorem 4.** Let \(u_2 \in \left(0, \frac{1}{\|\Phi\|_2}\right)\). Given that \(\lambda_1, \lambda_2 > 0\), for any fixed \(x_0\), suppose that \(x_s^*\) is the global minimizer of \(R(x_0, x_s)\). Then, for any \(x \in \mathbb{R}^N\), we have
\[
C_{x_0}(x_s^*, x_s^*) \leq C_{x_0}(x, x_s^*). \tag{34}
\]

**Proof.** It is similar to Theorem 3, so it was omitted. \[\square\]

Now, we can derive conclusions about the convergence of the AD-IFPT algorithm.

**Theorem 5.** Given \(\lambda_1, \lambda_2 > 0\); \(u_1 \in \left(0, \frac{1}{\|\Phi\|_2}\right)\), and \(u_2 \in \left(0, \frac{1}{\|\Phi\|_2}\right)\), let sequence \(\{x_0^k\}, \{x_s^k\}\) be generated by the form (24) and (25). Then, the following propositions hold:

(i) \(\{\mathcal{R}(x_0^k, x_s^k)\}\) is a monotonically decreasing sequence associated with \(k\).

(ii) \(\lim_{k \to \infty} \|x_0^{k+1} - x_0^k\|_2 = 0\) and \(\lim_{k \to \infty} \|x_s^{k+1} - x_s^k\|_2 = 0\).
Proof. (i) According to Theorem 1 and 3, we derived
\[
\min_x B_{\mathcal{A}}^k(x, x_0^k) = B_{\mathcal{A}}^k\left(H_{A_1U_1}^k(x_0^k + U_1(\Phi \Psi)^T(y - \Phi x_0^k - \Phi \Psi x_0^k)), x_0^k\right)
\]
\[
= B_{\mathcal{A}}^k(x_0^k, x_0^k).
\]

Similarly, we had
\[
\min_x C_{\mathcal{A}}^k(x, x_0^k) = C_{\mathcal{A}}^k(x_0^k, x_0^k);
\]

hence,
\[
R(x_0^{k+1}, x_0^{k+1}) = \frac{1}{u_2} \left( C_{\mathcal{A}}^{k+1}(x_0^{k+1}, x_0^{k}) - \|x_0^{k+1} - x_0^{k}\|^2\right)
\]
\[
+ \|\Phi x_0^{k+1} - \Phi x_0^{k}\|^2
\]
\[
\leq \frac{1}{u_2} C_{\mathcal{A}}^{k+1}(x_0^{k}, x_0^{k}) + \|\Phi x_0^{k+1} - \Phi x_0^{k}\|^2
\]
\[
- \frac{1}{u_2} \|x_0^{k+1} - x_0^{k}\|^2
\]
\[
\leq \frac{1}{u_2} C_{\mathcal{A}}^{k+1}(x_0^{k}, x_0^{k})
\]
\[
= R(x_0^{k+1}, x_0^{k})
\]
\[
= \frac{1}{u_1} \left( B_{\mathcal{A}}^k(x_0^{k+1}, x_0^k) - \|x_0^{k+1} - x_0^k\|^2\right)
\]
\[
+ \|\Phi \Psi x_0^{k+1} - \Phi \Psi x_0^k\|^2
\]
\[
\leq \frac{1}{u_1} B_{\mathcal{A}}^k(x_0^k, x_0^k) - \|\Phi \Psi x_0^{k+1} - \Phi \Psi x_0^k\|^2
\]
\[
- \frac{1}{u_1} \|x_0^{k+1} - x_0^k\|^2
\]
\[
\leq \frac{1}{u_1} B_{\mathcal{A}}^k(x_0^k, x_0^k)
\]
\[
= R(x_0^k, x_0^k),
\]

which shows that \( \{ R(x_0^k, x_0^k) \} \) is a monotonically decreasing sequence associated with \( k \).
(ii) Using the properties of the operator norm and Theorem 3, we had

$$
\begin{align*}
\sum_{k=0}^{N} \left\| x_{0}^{k+1} - x_{0}^{k} \right\|_2 & \leq \frac{u_1 \left( R \left( x_{0}^{k}, x_{s}^{k} \right) - R \left( x_{0}^{k-1}, x_{s}^{k-1} \right) \right)}{1 - u_1 \left\| \Phi \| \right\|_2^2} \\
& \leq \frac{u_1 R \left( x_{0}^{0}, x_{s}^{0} \right)}{1 - u_1 \left\| \Phi \| \right\|_2^2} \\
& = \frac{u_1 \left\| y \right\|_2^2}{1 - u_1 \left\| \Phi \| \right\|_2^2}
\end{align*}
$$

and

$$
\begin{align*}
\sum_{k=0}^{N} \left\| x_{s}^{k+1} - x_{s}^{k} \right\|_2 & \leq \frac{u_2 \left( R \left( x_{0}^{0}, x_{s}^{0} \right) - R \left( x_{0}^{N}, x_{s}^{N} \right) \right)}{1 - u_2 \left\| \Phi \| \right\|_2^2} \\
& \leq \frac{u_2 R \left( x_{0}^{0}, x_{s}^{0} \right)}{1 - u_2 \left\| \Phi \| \right\|_2^2} \\
& = \frac{u_2 \left\| y \right\|_2^2}{1 - u_2 \left\| \Phi \| \right\|_2^2}
\end{align*}
$$

Notice that to the right of the above two inequalities are constant values. Let $N \to +\infty$; it followed immediately that these two series were convergent, and this gave us

$$
\lim_{k \to \infty} \left\| x_{0}^{k+1} - x_{0}^{k} \right\|_2 = 0
$$

and

$$
\lim_{k \to \infty} \left\| x_{s}^{k+1} - x_{s}^{k} \right\|_2 = 0
$$
4. Performance Evaluation and Experimental Results

In this part, we first introduce an energy consumption model of the WSN, which evaluated power consumption by the number of instructions executed on the hardware. Then, we evaluated the energy efficiency performance of the proposed non-convex fraction function compressed sensing model. Furthermore, we compared the energy consumption of the IFPT and AD-IFPT with other traditional methods on the WSN during abnormal events’ monitoring.

4.1. Energy Consumption Estimation Model of the WSN

In our analysis, we used the energy consumption model of Mica2 [32] to describe all the energy consumption of our models. Mica2 is equipped with an Atmel Atmega 128L processor and a Chipcon CC1000 radio. We took into consideration the energy consumption caused by the hardware instruction execution times of different methods under a fixed hardware setup and transmission distance. Assuming that the distance between each node was 100 m, the hardware parameters of the sensor energy consumption are shown in Table 1, as given by [32]. The energy consumption at the node was expressed as follows:

\[ E = E_{REC} + E_{SMP} + E_{BG} + E_{COM} + E_{TSM}, \]

where \( E_{REC} \) is the energy consumption of receiving signal; \( E_{SMP} \) is the energy consumption of signal sampling; \( E_{BG} \) is the energy consumption of background power consumption during hardware operation; \( E_{BG} \) is the background energy consumption of the hardware; \( E_{TSM} \) is the energy consumption of a transmitting signal; \( E_{COM} \) is the energy of computation in the hardware, which depends on the kind of processing and operating carried out for the signal on the nodes.

In this paper, two working modes of compressed sensing were analyzed: analog compressed sensing and compressed data gathering in a multi-hop WSN.

Table 1. Hardware parameters.

| Parameter | Value     | Description                        |
|-----------|-----------|------------------------------------|
| \( P_{acq} \) | 15.01 mW | Signal sampling power consumption |
| \( E_{rx} \) | 0.923 \( \mu \)J | Energy consumption for receiving signals |
| \( E_{tx} \) | 1.984 \( \mu \)J | Energy consumption for transmitting signals |
| \( D_{op} \) | 0.14 \( \mu \)s | Duration of a CPU operation |
| \( P_{bg} \) | 9.6 mW | Hardware background power consumption |
| \( \epsilon_{add} \) | 3.30 nJ | Energy consumption for addition operation |
| \( \epsilon_{mul} \) | 9.90 nJ | Energy consumption for multiplication operation |
| \( \epsilon_{mrd} \) | 0.26 nJ | Energy consumption for memory reading |
| \( \epsilon_{mwr} \) | 4.30 nJ | Energy consumption for memory writing |

4.1.1. Energy Consumption of Analog Compressed Sensing

In this case, we assumed that the length of the signal was \( N \) under Nyquist sampling. When using compressed sensing, we sampled the original signal \( M \) times with \( M \ll N \). For
processor instruction, there were $M$ times of inner product computing of $N$-dimensional vectors under the analog compressed sensing mode we introduced in Section 2. We set the random sequence to satisfy the Bernoulli distribution, which takes the value of 0 or 1 with a 0.5 probability, respectively. Therefore, the energy consumption calculation formula of this case is given by

$$E_{(trad)} = E_{REC} + E_{SMP} + E_{BG} + E_{COM} + E_{TSM}$$

$$= (NE_{rx}) + (NP_{acq}D_{op}) + (2NP_{bg}D_{op})$$

$$+ (N\varepsilon_{mrd} + N\varepsilon_{mwr}) + (NE_{tx}),$$

$$E_{(CS)} = E_{REC} + E_{SMP} + E_{BG} + E_{COM} + E_{TSM}$$

$$= (NE_{rx}) + (NP_{acq}D_{op})$$

$$+ M[(M + N + MN)P_{bg}D_{op}]$$

$$+ (N\varepsilon_{mrd} + MN\varepsilon_{add} + M\varepsilon_{mwr})$$

$$+ (ME_{tx}),$$

where $E_{(trad)}$ represents the energy consumption of the WSN in traditional mode and $E_{(CS)}$ represents its energy consumption in compressed sensing mode.

4.1.2. Energy Consumption of CDG in Multi-Hop WSN

In this case, we assumed that there were $N$ nodes in a signal transmission chain of a multi-hop WSN. In the traditional mode, each node generates a reading from its sampling, receives the readings of all upstream nodes, and transmits them to the next node. Consequently, there were $N \times (N + 1) / 2$ times of data transmission in the overall process. In the compressed sensing mode, each node receives a binary number (0 or 1), which determines whether or not it was sampled. If a node received a 1, it generated a reading in the sampling, received a number from the nearest upstream node, and then, transmitted their sum to the next node. If a node received a 0, it only had to transmit a number from the nearest upstream node to the next one without sampling. This step was repeated $M$ times. Therefore, the energy consumption calculation formula of this case is given by:

$$E_{(trad)} = E_{REC} + E_{SMP} + E_{BG} + E_{COM} + E_{TSM}$$

$$= \left(\frac{1}{2}N(N-1)E_{rx}\right) + (NP_{acq}D_{op})$$

$$+ (N(N-1) + 2N)P_{bg}D_{op}$$

$$+ \left(\frac{1}{2}N(N+1)(\varepsilon_{mrd} + \varepsilon_{mwr})\right)$$

$$+ \left(\frac{1}{2}N(N+1)E_{tx}\right),$$

$$E_{(CS)} = E_{REC} + E_{SMP} + E_{BG} + E_{COM} + E_{TSM}$$

$$= (MNE_{rx}) + (NP_{acq}D_{op})$$

$$+ (4MN + M(N-1))P_{bg}D_{op}$$

$$+ (2MN\varepsilon_{mrd} + M(N-1)\varepsilon_{add} + 2MN\varepsilon_{mwr})$$

$$+ (MNE_{tx}),$$

where $E_{(trad)}$ represents the energy consumption of the WSN in traditional mode and $E_{(CS)}$ represents it in compressed sensing mode.
4.2. Performance of the Proposed Model and Energy Consumption Comparison

In this part, we first evaluated the availability of the proposed AD-IFPT algorithm in abnormal events’ monitoring. We randomly generated signals with a length of 512, sparsity level $K_1 = 64$, abnormal events $K_2 = 16$, and $-30$ dB of Gaussian noise, then reconstructed the signals using 250 measurements by the AD-IFPT (see Figure 5). The result showed that the AD-IFT algorithm could reconstruct the original signal with noise and abnormal events within a 0.1% error range. Therefore, it is advisable to use the fraction penalty function model and the AD-IFPT algorithm to reconstruct the signal.

The following is the comparison of the AD-IFPT, IFPT, and two traditional compressed sensing models in the experiment. The last three algorithms adopted the matrix splicing method in [6] when reconstructing abnormal event monitoring signals. The primitive dual-interior point method based on $\ell_1$-norm relaxation (the built-in compressed sensing algorithm in the L1-MAGIC package in MATLAB) and the greedy algorithm were selected as the contrast.

The result showed that the IFPT and AD-IFPT had better performance compared to the traditional compressed sensing methods; more specifically, the number of measurements of the proposed models was significantly reduced. The AD-IFPT had the best performance: it stably reconstructed a 512-dimensional signal in 170 measurements without abnormal events and in 225 measurements with abnormal events (see Figure 6).
In the simulation experiment, we determined that, when each algorithm reconstruction rate reached 100% stability, the number of its measurements could be accepted. Moreover, our experiment was carried out at different sparsity signal levels (0.05 to 0.275) and considered two scenarios: WSN monitoring with and without abnormal events. In addition, we calculated energy consumption in two cases: analog compressed sensing and compressed data gathering in the multi-hop WSN (see Figure 7). The results showed that, compared with transmitting the signal directly, the energy consumption of the compressed sensing method was significantly reduced when the signal sparsity decreased, and the energy savings reached more than 50% of the direct transmission when $K/N$ was below 0.05. In the comparison with the compressed sensing methods, the AD-IFPT reduced the energy consumption compared with other traditional compressed sensing models at the same sparsity level.

**Remark 1.** It is worth noting that, in each iteration, the time complexity of the IFPT algorithm was $O(MN)$, mainly caused by the multiplication of the matrix and vector. Assuming that it takes $T$ iterations to finish the algorithm, then the time complexity of our algorithm is $O(TMN)$, where $T$ is related to the parameters. The complexity of the algorithms compared in this paper, i.e., convex optimization and the greedy algorithm, was $O(N^3)$ and $O(KMN)$, respectively [33].
5. Conclusions

In this paper, we mainly discussed the energy consumption of the sensor nodes in a WSN by using a non-convex compressed sensing model, which is different from the current popular models where the $\ell_1$-norm is used as the surrogate function of the $\ell_0$-norm. However, we adopted a specific non-convex surrogate function (the fraction function) as the surrogate function of the $\ell_0$-norm. For abnormal events’ monitoring by the WSN, considering the sparse characteristics of abnormal events, we proposed an optimization model and designed an AD-IFPT algorithm to solve this optimization model, which avoided the extra measurements of previous algorithms. Finally, we used a WSN energy consumption model for the sensor nodes in the WSN under different methods. The results showed that our model and algorithm reduced the amount of data acquisition and transmission such that the energy consumption of the sensor nodes was effectively reduced.

Author Contributions: Conceptualization, Y.H., H.L. and J.P.; methodology, Y.H., H.L. and J.P.; software, Y.H.; validation, Y.H., H.L. and J.P.; formal analysis, Y.H.; investigation, Y.H.; resources, H.L. and J.P.; data curation, Y.H.; writing—original draft preparation, Y.H.; writing—review and editing, Y.H. and H.L.; visualization, Y.H.; supervision, H.L. and J.P.; project administration, H.L. and J.P.; funding acquisition, H.L. and J.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the National Science Foundation of China (No. 12271117, 12031003) and the basic research joint funding project of university and Guangzhou city (202102010434).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

1. Akylidiz, I.; Su, W.; Sankarasubramaniam, Y.; Cayirci, E. Wireless sensor networks: A survey. *Comput. Netw.* 2002, 38, 393–422. [CrossRef]

2. Freitas, D.M.D.; Kininmonth, S.; Woodley, S. Linking science and management in the adoption of sensor network technology in the Great Barrier Reef coast, Australia. *Comput. Environ. Urban Syst.* 2009, 33, 111–121. [CrossRef]

3. Ahlawat, M.K. Wireless Sensor Network—A Theoretical Review. *Int. J. Wired Wirel. Commun.* 2013, 1, 11–19.

4. Donoho, D.L. Compressed sensing. *IEEE Trans. Inf. Theory* 2006, 52, 1289–1306. [CrossRef]

5. Bajwa, W.U.Z.; Haupt, J.; Sayeed, A.M.; Nowak, R.D. Compressive wireless sensing. In Proceedings of the 2006 5th International Conference on Information Processing in Sensor Networks, Nashville, TN, USA, 19–21 April 2006; pp. 134–142.

6. Luo, C.; Wu, F.; Sun, J.; Chen, C. Compressive data gathering for large-scale wireless sensor networks. In Proceedings of the 15th Annual International Conference on Mobile Computing and Networking, MOBICOM 2009, Beijing, China, 20–25 September 2009.

7. Luo, C.; Wu, F.; Sun, J.; Chen, C. Efficient Measurement Generation and Pervasive Sparsity for Compressive Data Gathering. *IEEE Trans. Wirel. Commun.* 2010, 9, 3728–3738. [CrossRef]

8. Dandekar, D.R.; Deshmukh, P.R. Energy balancing multiple sink optimal deployment in multi-hop Wireless Sensor Networks. In Proceedings of the IEEE International Advance Computing Conference, Ghaziabad, India, 22–23 February 2013; pp. 408–412.

9. Tropp, J.A.; Laska, J.N.; Duarte, M.F.; Romberg, J.K.; Baraniuk, R.G. Beyond Nyquist: Efficient Sampling of Sparse Bandlimited Signals. *IEEE Trans. Inf. Theory* 2010, 56, 520–544. [CrossRef]

10. Polunin, M.; Bykova, A.; Losev, V.V. Analog to Information Converter with Random Demodulation Based on SAR ADC. In Proceedings of the 2011 IEEE Conference of Young Russian Researchers in Electrical and Electronic Engineering (ElConRus), St. Petersburg and Moscow, Russia, 26–28 January 2021; pp. 2734–2738.

11. Kirolos, S.; Laska, J.; Wakin, M.; Duarte, M.; Baron, D.; Ragheb, T.; Massoud, Y.; Baraniuk, R. Analog-to-Information Conversion via Random Demodulation. In Proceedings of the 2006 IEEE Dallas/CAS Workshop on Design, Applications, Integration and Software, Richardson, TX, USA, 29–30 October 2006; pp. 71–74.

12. Laska, J.N.; Kirolos, S.; Duarte, M.F.; Ragheb, T.S.; Baraniuk, R.G.; Massoud, Y. Theory and Implementation of an Analog-to-Information Converter using Random Demodulation. In Proceedings of the 2007 IEEE International Symposium on Circuits and Systems, New Orleans, LA, USA, 27–30 May 2007; pp. 1959–1962.

13. Donoho, D.L.; Tanner, J. Sparse nonnegative solution of underdetermined linear equations by linear programming. *Proc. Natl. Acad. Sci. USA* 2005, 102, 9446–9451. [CrossRef]

14. Candes, E.J.; Tao, T. Decoding by linear programming. *IEEE Trans. Inf. Theory* 2005, 51, 4203–4215. [CrossRef]

15. Fan, J.; Li, R. Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties. *J. Am. Stat. Assoc.* 2001, 96, 1348–1360. [CrossRef]

16. Zhang, C.H. Nearly unbiased variable selection under minimax concave penalty. *Ann. Stat.* 2010, 38, 894–942. [CrossRef]

17. Li, H.; Zhang, Q.; Cui, A.; Peng, J. Minimization of fraction function penalty in compressed sensing. *IEEE Trans. Neural Netw. Learn. Syst.* 2020, 31, 1626–1637. [CrossRef] [PubMed]

18. Chartrand, R. Exact Reconstruction of Sparse Signals via Nonconvex Minimization. *IEEE Signal Process. Lett.* 2007, 14, 707–710. [CrossRef]

19. Natarajan, B.K. Sparse approximate solutions to linear systems. *SIAM J. Comput.* 1995, 24, 227–234. [CrossRef]

20. Blumensath, T.; Davies, M.E. Iterative Thresholding for Sparse Approximations. *J. Fourier Anal. Appl.* 2008, 14, 629–654. [CrossRef]

21. Donoho, D.L.; Huo, X. Uncertainty principles and ideal atomic decomposition. *IEEE Trans. Inf. Theory* 2001, 47, 2845–2862.

22. Elad, M. Sparse and Redundant Representations—From Theory to Applications in Signal and Image Processing; Springer: New York, NY, USA, 2010.

23. Li, H.; Peng, J.; Yue, S. The Sparsity of Underdetermined Linear System via $\ell_p$ Minimization for $0 < p < 1$. *Math. Probl. Eng.* 2015, 2015, 1–6.

24. Chen, S.S.; Saunders, D. Atomic Decomposition by Basis Pursuit. *Siam Rev.* 2001, 43, 129–159. [CrossRef]

25. Candes, E.J.; Romberg, J.K.; Tao, T. Stable signal recovery from incomplete and inaccurate measurements. *Commun. Pure Appl. Math.* 2010, 59, 1207–1223. [CrossRef]

26. Donoho, D.L. De-noising by soft-thresholding. *IEEE Trans. Inf. Theory* 2002, 41, 613–627. [CrossRef]

27. Goldstein, T.; Osher, S. The Split Bregman Method for L1-Regularized Problems. *SIAM J. Imaging Sci.* 2009, 2, 323–343. [CrossRef]

28. Yin, W.; Osher, S.; Goldfarb, D.; Darbon, J. Bregman iterative algorithms for $\ell_1$-minimization with applications to compressed sensing. *SIAM J. Imaging Sci.* 2008, 1, 143–168. [CrossRef]

29. Foucart, S.; Lai, M. Sparsest solutions of underdetermined linear systems via $\ell_q$-minimization for $0 < q \leq 1$. *Appl. Comput. Harmon. Anal.* 2009, 26, 395–407.

30. Chartrand, R. Nonconvex Compressed Sensing and Error Correction. In Proceedings of the 2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP’07, Honolulu, HI, USA, 15–20 April 2007; Volume 3, pp. III-889–III-892.
31. Foucart, S.; Rauhut, H. *A Mathematical Introduction to Compressive Sensing*; Birkhäuser: New York, NY, USA, 2013.
32. Karakus, C.; Gurbuz, A.C.; Tavli, B. Analysis of energy efficiency of compressive sensing in wireless sensor networks. *IEEE Sens. J.* 2013, 13, 1999–2008. [CrossRef]
33. Qaisar, S.; Bilal, R.M.; Iqbal, W.; Naureen, M.; Lee, S. Compressive sensing: From theory to applications, a survey. *J. Commun. Netw.* 2013, 15, 443–456. [CrossRef]