Article

A Compressed Sensing Recovery Algorithm Based on Support Set Selection

Wandi Liang ¹, Zixiong Wang ^{1,*}, Guangyu Lu ¹ and Yang Jiang ²

¹ School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China; liang_wandi@tju.edu.cn (W.L.); luy_2020@163.com (G.L.)

² College of Physics, Guizhou University, Guiyang 550025, China; jiangy@gzu.edu.cn

* Correspondence: wangzixiong@tju.edu.cn

Abstract: The theory of compressed sensing (CS) has shown tremendous potential in many fields, especially in the signal processing area, due to its utility in recovering unknown signals with far lower sampling rates than the Nyquist frequency. In this paper, we present a novel, optimized recovery algorithm named supp-BPDN. The proposed algorithm executes a step of selecting and recording the support set of original signals before using the traditional recovery algorithm mostly used in signal processing called basis pursuit denoising (BPDN). We proved mathematically that even in a noise-affected CS system, the probability of selecting the support set of signals still approaches 1, which means supp-BPDN can maintain good performance in systems in which noise exists. Recovery results are demonstrated to verify the effectiveness and superiority of supp-BPDN. Besides, we set up a photonic-enabled CS system realizing the reconstruction of a two-tone signal with a peak frequency of 350 MHz through a 200 MHz analog-to-digital converter (ADC) and a signal with a peak frequency of 1 GHz by a 500 MHz ADC. Similarly, supp-BPDN showed better reconstruction results than BPDN.

Citation: Liang, W.; Wang, Z.; Lu, G.; Jiang, Y. A Compressed Sensing Recovery Algorithm Based on Support Set Selection. *Electronics* **2021**, *10*, 1544. <https://doi.org/10.3390/electronics10131544>

Academic Editor: George A. Papakostas

Received: 16 May 2021
Accepted: 21 June 2021
Published: 25 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 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 (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: compressed sensing; microwave photonics; signal recovery algorithm; basis pursuit denoising; sub-Nyquist sampling

1. Introduction

Compressed sensing (CS) is a theory that describes an unknown high-dimensional signal with few non-zero components that can be recovered from extremely incomplete information, which could reduce data redundancy to a considerable extent. Since CS theory was conceived [1,2], it attracted great attention all over the world for its applicability to many fields, such as signal processing, radar systems [3], imaging [4,5], optical microscopy [6], direction of arrival (DOA) [7], and cognitive radio (CR) [8]. In the signal processing area in particular, the rapid development of communication technologies, the utilization of high-speed computers, and high-frequency or ultrawide band signals have caused explosive growth of data, along with serious challenges to digital devices such as the analog-to-digital converter (ADC) using the Nyquist sampling theory [9,10], which limits the sampling rate to twice or above the highest frequency of signals. Nevertheless, CS systems can greatly overcome above difficulties.

CS systems are composed of two parts: measurement and reconstruction. In the former part, signals are compressively and randomly measured by the measurement matrix (MM), which usually consists of groups of pseudo random bit sequence (PRBS) in communication systems, by mixing the signal with PRBS with mixers or modulators. In other words, high-dimensional signals are projected randomly and linearly to a much lower-dimensional space, making it possible to reduce the requirement of sampling rate. Many works have been done in this area during the past decades and have made great progress. In the very beginning, an CS system was implemented in electronic systems called the random demodulator (RD) [11,12]. In the RD, a microwave signal is directly

mixed with PRBS via a mixer and accumulated by an accumulator or a low-pass filter (LPF); this is done in awareness of the sampling rate being exponentially lower than the Nyquist rate, which was pioneering at that time. Recent years, researchers turned their eyes to photonics-enabled CS systems due to the advantages of low loss, large bandwidth, and high speed offered by photonics [13–24]. In [13], a CS system in a photonic link was demonstrated for the first time, which successfully recovered a 1 GHz harmonic signal with 500 MS/s sampling rate. Afterwards the researcher introduced the time stretch technique to broaden the pulse width in order to lower the sampling rate even further, and realized the integration process by compressing the optical pulses before photodiode (PD) [17,21,22]. Under this scheme, signals spanning from 900 MHz to 14.76 GHz can be recovered by a 500 MHz ADC. In contrast to the above systems using optical pulses generated by lasers, a CS system can be accomplished by using a broadband light source instead [23]. Moreover, [24] mixed a signal with PRBS using spatial light modulator (SLM), rather than Mach–Zehnder modulator (MZM), as in the above schemes. In the signal reconstruction part, until now recovery algorithms have mostly been divided into three categories, which are convex optimization methods [25,26], greedy algorithms [27–32], and Bayesian compressive sensing (BCS) approaches [33–35]. Convex optimization methods are basically composed of basis pursuit (BP), BPDN, and extended algorithms, and explicit details are discussed in the following sections. The main form of greedy algorithms is iteration and residual updating. As a fundamental method in greedy algorithms, the matching pursuit (MP) method chooses an optimal value, and only columns related to the chosen value in MM will be updated in the measurements during each iteration, whereas in the orthogonal matching pursuit (OMP) method, residuals are updated and iterated in the global range. Additionally, optimized algorithms have been formulated that choose more than one value each time [29], have dynamic selecting numbers [30], have different application situations [31,32] and so on. Moreover, based on the Bayesian formalism, BCS algorithm came about and is a promising research direction as well. Differently from convex optimization methods and greedy algorithms, which are non-adaptive, BCS provides a framework for sequential adaptive sensing, i.e., automatically adjusting the recovery process. Besides, under BCS, other observations such as posterior density function, confidence, and uncertainty estimation are obtained. Among these algorithms, the greedy algorithm is fastest and easiest; the convex optimization algorithm is the most accurate and is suitable for large signals. Additionally, although BCS acquires the smallest reconstruction errors, its complexity is a barrier. Thus, in practical CS communication systems, convex optimization algorithms such as BPDN are more popular choices.

In this paper, based on support set selection, we present a novel algorithm which has better reconstruction performance than that of the BPDN algorithm. In our optimized algorithm, finding where the non-zero components are is the first thing to do, and it is named the supp-BPDN algorithm. We proved that not only in ideal CS systems, but also in the more common situation with noise influence, the support set of the original signal can be found from the maximum value of an inner product of measurements and a column vector of MM at high probability.

We combined the support set selection and BPDN innovatively. Additionally, supp-BPDN recovers signals by relying on BPDN rather than the least squares method; thus, a more accurate reconstruction can be obtained. Furthermore, differently from the support set selection of MP and relative algorithms, it is unnecessary to update measurements and get residuals during each iteration during the mathematical analysis. Overall, supp-BPDN has a better performance in recovering signals.

The organization of the paper is as follows. Section 1 is the the background and introduction of CS theory. Section 2 outlines the principle of CS theory and briefly analyzes the convex optimization algorithm. In Section 3, supp-BPDN is put forward and the principle of it is analyzed in detail. Section 4 demonstrates a series of numerical simulations to support our mathematical proof and show the superiority of supp-BPDN. In the follow-

ing section, a CS system is proposed and simulated results are presented, verifying the effectiveness of supp-BPDN. Conclusions are drawn in the final section.

2. Principles

2.1. CS Theory

Consider an unknown vector $\mathbf{x} \in \mathbb{R}^N$ of N -length, $\mathbf{x}$ only containing k non-zero elements, and $k \ll N$, the index set of these non-zero elements is defined as the support set $\text{supp}(\mathbf{x})$, which is

$$\text{supp}(\mathbf{x}) = \{i \in [N], x_i \neq 0\}, \quad (1)$$

where $[N] = \{1, 2, \dots, N\}$ is the index set of all elements in $\mathbf{x}$. In other words, $[N]$ is all the subscripts of elements in $\mathbf{x}$. Equation (1) means the other $(N - k)$ components are totally equal to zero. It is the strict definition of “sparsity”, which is the precondition of CS theory. However, there are few signals totally meeting the requirement that other components are all zero. To expand the application range of CS, luckily, the condition can be relaxed. At this time $\mathbf{x}$ is called “compressible”. The $(N - k)$ components in compressible signal are really small values, not strictly equal to zero. The sparsity or compressibility of a signal is the premise of CS.

Under such conditions, CS states that

$$\mathbf{y} = \Phi \mathbf{x}, \quad (2)$$

in which $\mathbf{y} \in \mathbb{R}^M$ is the measurement vector, composed of measurements. Besides, $\Phi \in \mathbb{R}^{M \times N}$ is the MM, and they hold the relation that $k < M \ll N$. MM must obey the restricted isometry property (RIP) [36], which defines the restricted isometry constant (RIC) to be the smallest value δ_k that makes

$$(1 - \delta_k) \|\mathbf{x}\|_2^2 \leq \|\Phi \mathbf{x}\|_2^2 \leq (1 + \delta_k) \|\mathbf{x}\|_2^2 \quad (3)$$

hold for all k -sparse signals. RIP guarantees that the mapping relationship from $\mathbf{x}$ to $\mathbf{y}$ is unique and determined; i.e., different k -sparse signals cannot be mapped to the same $\mathbf{y}$ through Φ which obeys RIP.

The process of signal reconstruction is actually solving Equation (2), and the original unknown signal can be recovered by reconstruction algorithms. Besides, it is worth noticing that as long as the signal is sparse in any domain, such as the frequency domain or wavelet domain, the signal can be regarded as sparse and recovered as well. In a mathematical expression, it is

$$s_i = \sum_{j=1}^N \psi_{ij} x_j \quad (i, j = 1, 2, \dots, N), \quad (4)$$

or in a matrix expression, $\mathbf{s} = \Psi \mathbf{x}$, among which $\Psi \in \mathbb{R}^N$ is the orthogonal transformation basis. $\mathbf{x}$ is still the sparse vector. It is known that when M is not very small, $\Theta = \Phi \Psi$ also abides by the RIP at high probability according to the universality of CS [37,38], making the application range of CS much wider as well. Therefore, although $\mathbf{s}$ is not sparse, it can be recovered in CS system. It is more suitable for communication systems when Ψ stands for the Fourier transform matrix.

2.2. Problem Statement

We mean to outline signal reconstruction in communication systems under CS theory. The original unknown signal, denoted as $\mathbf{s}$, is the sum of different cosine signals, and obviously it is sparse in frequency domain. In consideration of noise in actual systems, such as shot noise and thermal noise in PD, and quantization noise in ADC, from Equation (2) and Equation (4), the model of CS-enabled communication system is

$$\mathbf{y}_{\text{noise}} = \Theta \mathbf{x} + \mathbf{n} = \Phi \mathbf{s} + \mathbf{n}, \quad (5)$$

where $\mathbf{n}$ is additive noise, and $\mathbf{n} \in \mathbb{R}^M$. In this specific situation, MM is usually used as the matrix composed of M groups of PRBS consisting of 0 and 1, denoted as $\mathbf{r}_1$ until $\mathbf{r}_M$. Hence, the detailed expression of MM is

$$\Phi = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_M \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1N} \\ & & \ddots & \\ & & & \\ r_{M1} & r_{M2} & \dots & r_{MN} \end{pmatrix} = (\boldsymbol{\varphi}_1 \quad \boldsymbol{\varphi}_2 \quad \dots \quad \boldsymbol{\varphi}_N), \quad (6)$$

in which $\boldsymbol{\varphi}_i = (r_{1i} \quad r_{2i} \quad \dots \quad r_{Mi})^T$ is the i^{th} column vector of Φ .

2.3. Convex Optimization Recovery Algorithm

Under the condition that $\mathbf{y}_{\text{noise}}$ and Φ are known, one can intuitively derive Equation (2) to an optimization problem, which is

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathbf{x}\|_0 \\ \text{s.t.} \quad & \Phi\mathbf{x} = \mathbf{y}_{\text{noise}}. \end{aligned} \quad (7)$$

However, the model established by ℓ_0 -norm is not differentiable. In order to solve Equation (7), enumerating all the possibilities of $\mathbf{x}$ is necessary—up to C_N^k kinds. Therefore, it is unrealistic to directly solve such a non-deterministic polynomial hard (NPH) problem [39]. Fortunately, because of the sparse signal, Equation (7) can be approximated to

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathbf{x}\|_1 \\ \text{s.t.} \quad & \Phi\mathbf{x} = \mathbf{y}_{\text{noise}}, \end{aligned} \quad (8)$$

which is a convex problem; therefore, Equation (8) can be resolved directly by the BP algorithm.

More practically, the BP algorithm should be adapted to a noise-affected case, as mentioned in Equation (5). Then Equation (8) is evolved to a BPDN algorithm which demands

$$\min_{\mathbf{x}} \quad \frac{1}{2} \|\mathbf{y}_{\text{noise}} - \Phi\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1, \quad (9)$$

where λ is a selected parameter to make results optimal or near-optimal, and

$$\lambda = \sigma \sqrt{2 \log(N)}, \quad (10)$$

and $\sigma > 0$ is the noise level [25,40].

3. Recovery Algorithm Based on Support Set Selection

3.1. Analysis of Ideal System

BPDN is an overly pessimistic algorithm, and recovery errors are often caused by clutter frequencies. In this paper we figure this problem out by finding the support set of sparse signal $\mathbf{x}$, $\text{supp}(\mathbf{x})$, and only recovering the information which corresponds to $\text{supp}(\mathbf{x})$.

To obtain the exact $\text{supp}(\mathbf{x})$, we start with analyzing the ideal CS system. Normally each row vector of Φ is a different and random PRBS denoted as $\mathbf{r}$ whose length is the same as that of the measured signal, and each column of Φ is denoted as $\boldsymbol{\varphi}$ as Equation (6) shows. Therefore, from expression

$$\mathbf{y} = \begin{pmatrix} r_{11}x_1 + r_{12}x_2 + \cdots + r_{1N}x_N \\ r_{21}x_1 + r_{22}x_2 + \cdots + r_{2N}x_N \\ \vdots \\ r_{M1}x_1 + r_{M2}x_2 + \cdots + r_{MN}x_N \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^N r_{1i}x_i \\ \sum_{i=1}^N r_{2i}x_i \\ \vdots \\ \sum_{i=1}^N r_{Mi}x_i \end{pmatrix}, \quad (11)$$

one could realize that $\mathbf{y}$ just relates to x_i and $\boldsymbol{\varphi}_i$, in which $i \in \text{supp}(\mathbf{x})$. It means only column vectors of Φ which corresponding to $\text{supp}(\mathbf{x})$ make contributions to $\mathbf{y}$. Therefore, by means of solving

$$\text{argmax} |\langle \mathbf{y}, \boldsymbol{\varphi}_i \rangle|, \quad (12)$$

$\text{supp}(\mathbf{x})$ could be inferred. In this case, $\langle \cdot \rangle$ means inner product, denoted as p in the following parts.

Suppose that $\mathbf{x}$ is a 1-sparse signal and $\text{supp}(\mathbf{x}) = \{1\}$; i.e., x_1 is the element which carries effective information, and the remaining elements $x_j = 0$ ($j \neq 1$) or x_j can be regarded as zero. Under this circumstance, measurement vector $\mathbf{y}$ is

$$\mathbf{y} = (r_{11}x_1, r_{21}x_1, \dots, r_{M1}x_1)^T. \quad (13)$$

By substituting Equation (13), $\boldsymbol{\varphi}_1$ and $\boldsymbol{\varphi}_j$ ($j \neq 1$), respectively, into Equation (12), one can get the inner products

$$p_1 = |\langle \mathbf{y}, \boldsymbol{\varphi}_1 \rangle| = \left| \sum_{i=1}^M r_{i1}^2 x_1 \right| \quad (14)$$

and

$$p_j = |\langle \mathbf{y}, \boldsymbol{\varphi}_j \rangle| = \left| \sum_{i=1}^M r_{i1} r_{ij} x_1 \right|. \quad (15)$$

As Φ is composed of 0 and 1 due to PRBS, and x_1 is a positive number, actually we have

$$|\langle \mathbf{y}, \boldsymbol{\varphi}_1 \rangle| = \langle \mathbf{y}, \boldsymbol{\varphi}_1 \rangle = \sum_{i=1}^M r_{i1}^2 x_1 \quad (16)$$

and

$$|\langle \mathbf{y}, \boldsymbol{\varphi}_j \rangle| = \langle \mathbf{y}, \boldsymbol{\varphi}_j \rangle = \sum_{i=1}^M r_{i1} r_{ij} x_1. \quad (17)$$

Besides, on account of PRBS characteristics, r_{i1} and r_{ij} each holds the probability of $\frac{1}{2}$ being 1 or 0, which is

$$P(r_{i1} = 0) = P(r_{i1} = 1) = P(r_{ij} = 0) = P(r_{ij} = 1) = \frac{1}{2}, \quad (18)$$

where $P(\cdot)$ means probability. Therefore, there is

$$\begin{aligned} P(r_{i1} = 0, r_{ij} = 0) &= P(r_{i1} = 0, r_{ij} = 1) \\ &= P(r_{i1} = 1, r_{ij} = 0) = P(r_{i1} = 1, r_{ij} = 1) = \frac{1}{4}, \end{aligned} \quad (19)$$

Substituting Equation (19) into Equations (16) and (17), for one single term in inner products, gives the following probabilities:

$$P(r_{i1}^2 x_1 = r_{i1} r_{ij} x_1) = \frac{3}{4}, \quad (20)$$

and

$$P(r_{i1}^2 x_1 \gg r_{i1} r_{ij} x_1) = \frac{1}{4}. \quad (21)$$

It can be known from above formulae that the relationship

$$r_{i1}^2 x_1 \geq r_{i1} r_{ij} x_1 \quad (22)$$

of each term of Equations (16) and (17) always holds. Therefore, the summations of Equation (22), which are

$$\langle \mathbf{y}, \boldsymbol{\varphi}_1 \rangle \geq \langle \mathbf{y}, \boldsymbol{\varphi}_j \rangle, \quad (23)$$

hold as well. It means that in this example, p_1 is the maximum of all the inner products, and the subscript of p_1 is at least one of the elements in $\text{supp}(\mathbf{x})$. Therefore, finding the support set can be equivalently transformed to obtaining the position of maximum of all the inner products.

Furthermore, for the sum of all M terms in inner products, there is

$$P(p_1 = p_j) = P\left(\sum_{i=1}^M r_{i1}^2 x_1 = \sum_{i=1}^M r_{i1} r_{ij} x_1\right) = \left(\frac{3}{4}\right)^M. \quad (24)$$

Generally in CS system, M is a number that makes $\left(\frac{3}{4}\right)^M$ small enough. Meanwhile, if $p_1 \neq p_j$, and because $x_1 \gg 0$, leading to the fact that even there is only one term being $r_{i1}^2 x_1 \gg r_{i1} r_{ij} x_1$, $p_1 \gg p_j$ holds. Additionally, the probability of $p_1 \gg p_j$ is

$$P(p_1 \gg p_j) = P\left(\left|\sum_{i=1}^M r_{i1}^2 x_1\right| \gg \left|\sum_{i=1}^M r_{i1} r_{ij} x_1\right|\right) = 1 - \left(\frac{3}{4}\right)^M. \quad (25)$$

Equation (25) is very useful for the analysis of the following noise-affected CS system.

3.2. Analysis of a Noise-Affected System

Similarly, we analyze the noise-affected CS system. The noise model is $\mathbf{n} = (n_1, n_2, \dots, n_M)^T$, considered as absolute random. Likewise, with regard to the sparse vector $\mathbf{x}$, we suppose that x_1 carries information while other elements are equal to zero. Therefore, under this circumstance, Equation (11) is

$$\mathbf{y}_{\text{noise}} = \left(\sum_{i=1}^N r_{1i} x_i + n_1, \sum_{i=1}^N r_{2i} x_i + n_2, \dots, \sum_{i=1}^N r_{Mi} x_i + n_M \right)^T. \quad (26)$$

Additionally, the inner products are

$$|\langle \mathbf{y}_{\text{noise}}, \boldsymbol{\varphi}_1 \rangle| = \left| \sum_{i=1}^M (r_{i1}^2 x_1 + r_{i1} n_i) \right| \quad (27)$$

and

$$|\langle \mathbf{y}_{\text{noise}}, \boldsymbol{\varphi}_j \rangle| = \left| \sum_{i=1}^M (r_{i1} r_{ij} x_1 + r_{ij} n_i) \right|. \quad (28)$$

Due to the influence of noise, finding the support set of signal is not a certain event. The subscript of the maximum inner product is no longer precisely equivalent to $\text{supp}(\mathbf{x})$. Therefore, we analyze the probability of equivalent transformation holds.

From Cauchy inequality $|\langle \boldsymbol{\alpha}, \boldsymbol{\beta} \rangle| \leq \|\boldsymbol{\alpha}\| \|\boldsymbol{\beta}\|$, there are

$$p_1 = |\langle \mathbf{y}_{\text{noise}}, \boldsymbol{\varphi}_1 \rangle| \leq \|\mathbf{y}_{\text{noise}}\| \|\boldsymbol{\varphi}_1\| \quad (29)$$

and

$$p_j = |\langle \mathbf{y}_{\text{noise}}, \boldsymbol{\varphi}_j \rangle| \leq \|\mathbf{y}_{\text{noise}}\| \|\boldsymbol{\varphi}_j\|. \quad (30)$$

According to the random characteristic of PRBS, in the case of an extremely large sample size, $|\varphi_1|$ should be equal to $|\varphi_j|$ because φ_1 and φ_j have the same number of 1s and 0s. Therefore,

$$\begin{cases} |\varphi_1| = |\varphi_j| = \sqrt{\frac{M}{2}}, & M \text{ is an even number} \\ |\varphi_1| = |\varphi_j| = \sqrt{\frac{M+1}{2}}, & M \text{ is an odd number} \end{cases} \quad (31)$$

Substitute Equation (31) to Equations (29) and (30); then one can know that p_1 and p_j have the same upper bound, which is

$$\sqrt{\frac{M}{2}} |y_{\text{noise}}| \quad \text{or} \quad \sqrt{\frac{M+1}{2}} |y_{\text{noise}}|. \quad (32)$$

From the triangle inequality $|a + b| \leq |a| + |b|$, there are

$$p_1 = \left| \sum_{i=1}^M (r_{i1}^2 x_1 + r_{i1} n_i) \right| = \left| \sum_{i=1}^M r_{i1}^2 x_1 + \sum_{i=1}^M r_{i1} n_i \right| \leq \left| \sum_{i=1}^M r_{i1}^2 x_1 \right| + \left| \sum_{i=1}^M r_{i1} n_i \right| \quad (33)$$

and

$$p_j = \left| \sum_{i=1}^M (r_{i1} r_{ij} x_1 + r_{ij} n_i) \right| = \left| \sum_{i=1}^M r_{i1} r_{ij} x_1 + \sum_{i=1}^M r_{ij} n_i \right| \leq \left| \sum_{i=1}^M r_{i1} r_{ij} x_1 \right| + \left| \sum_{i=1}^M r_{ij} n_i \right|. \quad (34)$$

In the preceding part of the text, the large probability of the following was proven:

$$\left| \sum_{i=1}^M r_{i1}^2 x_1 \right| \gg \left| \sum_{i=1}^M r_{i1} r_{ij} x_1 \right|. \quad (35)$$

Meanwhile, in a extremely large sample size, $\left| \sum_{i=1}^M r_{i1} n_i \right|$ should be equal to $\left| \sum_{i=1}^M r_{ij} n_i \right|$. Therefore,

$$\left| \sum_{i=1}^M r_{i1}^2 x_1 \right| + \left| \sum_{i=1}^M r_{i1} n_i \right| \gg \left| \sum_{i=1}^M r_{i1} r_{ij} x_1 \right| + \left| \sum_{i=1}^M r_{ij} n_i \right|. \quad (36)$$

On the basis of the magnitude of the noise power being less than that of the original signal power, and the fact that $r_{ij} = 1$ or 0,

$$\left| \sum_{i=1}^M r_{i1}^2 x_1 \right| = \left| \sum_{i=1}^M r_{i1} x_1 \right| \approx |y_{\text{noise}}| < \sqrt{\frac{M}{2}} \cdot |y_{\text{noise}}|. \quad (37)$$

By joining Equations (32), (36), and (37), one can figure out a more accurate interval of p_1 and p_j , which is

$$\begin{cases} 0 \leq p_1 \leq \left| \sum_{i=1}^M r_{i1}^2 x_1 \right| + \left| \sum_{i=1}^M r_{i1} n_i \right| = a_{\max} \\ 0 \leq p_j \leq \left| \sum_{i=1}^M r_{i1} r_{ij} x_1 \right| + \left| \sum_{i=1}^M r_{ij} n_i \right| = b_{\max} \end{cases} \quad (38)$$

According to the central-limit theorem, p_1 and p_j are both approximate Gaussian distributions in their intervals. For convenience, we denote the upper bounds as $a_{\max}$ and $b_{\max}$, respectively. On account of $\left| \sum_{i=1}^M r_{i1}^2 x_1 \right| \gg \left| \sum_{i=1}^M r_{i1} r_{ij} x_1 \right|$, one can regard the relationship of $a_{\max}$ and $b_{\max}$ as

$$a_{\max} \gg b_{\max} \quad (39)$$

as well.

On the basis of determining the intervals and distributions, one can analyze the probability of $p_1 \geq p_j$. $P(p_1 \geq p_j)$ is actually the probability of satisfying $p_1 \geq p_j$ in a universal set Ω , which is

$$P(p_1 \geq p_j) = P(p_1 \geq p_j, \Omega) = \sum_{i=1}^{\max} P(p_1 \geq p_j, p_j = b_i). \quad (40)$$

In accordance with the contingent probability formula $P(AB) = P(A|B)P(B)$, each term is

$$P(p_1 \geq p_j, p_j = b_i) = P(p_1 \geq b_i)P(p_j = b_i). \quad (41)$$

Therefore, the probability that $p_1 \geq p_j$ is

$$P(p_1 \geq p_j) = \sum_{i=1}^{\max} P(p_1 \geq b_i)P(p_j = b_i), \quad (42)$$

where p_1 and p_j are both approximately subject to Gaussian distribution; i.e.,

$$p_1 \sim \mathcal{N}\left(\frac{a_{\max}}{2}, 1\right), \quad p_j \sim \mathcal{N}\left(\frac{b_{\max}}{2}, 1\right). \quad (43)$$

Accordingly, for p_1 and p_j , the probability density functions are denoted as $g(a)$ and $f(b)$; the distribution functions are written as $G(a)$ and $F(b)$, respectively.

From the total probability formula in the continuous interval, Equation (42) can be derived as

$$P(p_1 \geq p_j) = \int_0^{b_{\max}} P(p_1 \geq b_i)P(p_j = b_i)db_i = \int_0^{b_{\max}} [G(a_{\max}) - G(b_i)]f(b_i)db_i. \quad (44)$$

When $a_{\max} \gg b_{\max}$, among the interval $[0, b_{\max}]$, $[G(a_{\max}) - G(b_i)]$ can be approximate to 1, leading Equation (44) to

$$P(p_1 \geq p_j) \approx \int_0^{b_{\max}} f(b_i)db_i = 1. \quad (45)$$

Through the above analyses, for a 1-sparse signal, the support set of it is the maximum element in $\mathbf{p}$. Similarly, for a k -sparse signal, the support set contains top k maximum values.

In conclusion, even in a CS system affected by noise, acquiring $\text{supp}(\mathbf{x})$ via the maximum values of inner products is a high-probability event. Furthermore, the accuracy can be improved by increasing iteration times properly.

3.3. Algorithm Procedure

Procedures of the proposed algorithm are detailed in Algorithm 1. Before the process starts, input variables containing Φ , $\mathbf{y}_{\text{noise}}$ and iteration times t are required, among which $\Phi = [\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \dots, \boldsymbol{\varphi}_N]$. Besides, an initial zero vector denoted as $\boldsymbol{\xi}$ is demanded. $\boldsymbol{\xi}$ is an “assistant vector” used in supp-BPDN algorithm, which connects the estimated support set with the recovered signal information, and $\boldsymbol{\xi} = \mathbf{O} \in \mathbb{R}^{N \times 1}$, where $\mathbf{O}$ is the zero vector. Specific procedures are as follows: To begin with, calculate the inner products of $\mathbf{y}_{\text{noise}}$ and all initial $\boldsymbol{\varphi}$, finding the column vector $\boldsymbol{\varphi}_z^{(t)}$ corresponding to the maximum inner product, where z is the subscript and position of the column vector, and (t) means values in the t_{th} iteration. By means of this step one can get that $z^{(t)}$ belongs to the support set of $\mathbf{x}$. After that set $\boldsymbol{\varphi}_z^{(t)} = \mathbf{O}$, and put it back in Φ . Then repeat the above steps t times to get t values, forming the estimated support set written as $\text{supp}(\mathbf{x})$. Therefore, with a probability close to 1, $\text{supp}(\mathbf{x}) \subset \hat{\text{supp}}(\mathbf{x})$ holds. Next, set the elements corresponding to $\text{supp}(\mathbf{x})$ in $\boldsymbol{\xi}$ as 1, and the rests are still 0. In the following step, solve the convex optimization problem to obtain $\mathbf{x}^*$. Lastly, calculate $\hat{\mathbf{x}} = \mathbf{x}^* \odot \boldsymbol{\xi}$ to reconstruct the original sparse signal $\hat{\mathbf{x}}$ in frequency domain, where $\odot$ is the Hadamard product. If time domain of original signal is demanded, calculate $\hat{\mathbf{s}} = \mathcal{F}^{-1}\{\hat{\mathbf{x}}\}$.

Algorithm 1. The supp-BPDN algorithm procedure**Inputs:**Measurement matrix Φ , measurements y_{noise} , circulation times t **Initialization:**Assistant vector $\xi = \mathbf{O}$ ($\mathbf{O}$ denotes the zero vector), estimated support set $\text{supp}(\mathbf{x}) = \emptyset$ **Procedure :**1. $\varphi_z^{(t)} = \arg\max |\langle \varphi_i, y_{\text{noise}} \rangle| \ (i = 1, 2, \dots, N)$.2. $\text{supp}(\mathbf{x}) = \text{supp}(\mathbf{x}) \cup \{z^{(t)}\}$.3. $\varphi_z^{(t)} = \mathbf{O}$.4. Cycle step 1 to step 3 for t times.5. $\xi_{\text{supp}(\mathbf{x})} = 1$, the rest elements of ξ are still 0.

6. Solving

$$\min \quad \frac{1}{2} \|y_{\text{noise}} - \Phi \mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1$$

$$\text{s.t.} \quad \Phi \mathbf{x} = y_{\text{noise}}$$

to obtain $\mathbf{x}^*$.7. $\hat{\mathbf{x}} = \mathbf{x}^* \odot \xi$ ($\odot$ is the Hadamard product).**Output:**Estimation values $\hat{\mathbf{x}}$

For further detail, number of iterations t is a parameter which should be paid attention to. For every circulation of step 1 to step 3 in Algorithm 1, one element in $\text{supp}(\mathbf{x})$ will be generated. Thus the fundamental relationship that $k \leq t$ should be satisfied; otherwise, there will be a loss of effective information of $\mathbf{x}$. However, when we consider a CS system infected by noise, from Equations (25) and (45), the probability of selecting the support set exactly approaches 1 and is related to M . Therefore, t is usually an empirical parameter and there is a trade-off between accuracy and reconstruction performance. When t is a large number, redundant information will remain, leading to an increase in reconstruction error. However, if t is too small, perhaps the support set cannot be found completely. Generally we choose $t = 3k$ to make sure all the elements in the support set are selected.

In order not to be confused with greedy algorithms which aim to find the support set of sparse vectors as well, supp-BPDN algorithm will be analyzed further. The main distinction between supp-BPDN and greedy algorithms is the updating of measurements $\mathbf{y}$, which leads to the different construction of support set. In procedures of MP, by means of matrix multiplications involving Φ^T , residuals are calculated during evaluation, and this update focuses on the chosen column of Φ , making MP possibly choose the same column repeatedly [37]. In procedures of OMP, by solving the least squares method, residuals are updated to be orthogonal to all selected elements [28]. In supp-BPDN, through the mathematical analysis, it is not necessary to update and calculate values every iteration. Besides, ℓ_1 recovery algorithms have the advantage of high accuracy, especially when handling large signals. Additionally, supp-BPDN inherits this merit.

4. Numerical Simulations

In this section, a series of numerical simulations performed via CVX toolbox in Matlab are reported. In the simulations, $N = 256$, (number of measurements) $M = 64$, (original

signal) s was the sum of three cosines of different frequencies, and the mathematical expression was

$$s(t) = a \cos(2\pi f_1 t) + b \cos(2\pi f_2 t) + c \cos(2\pi f_3 t), \quad (46)$$

where $f_1 = 50$ MHz; $f_2 = 150$ MHz; $f_3 = 350$ MHz; and a , b , and c are constants, leading to sparsity $k = 6$ in frequency domain. Meanwhile, as the original signals are measured by PRBS in communication systems, we adopted a matrix in which 0s and 1s are randomly generated, and multiplied it with Fourier transform matrix Ψ as MM.

The first simulation compared the single recovery results of BPDN and supp-BPDN algorithms, as shown in the Figure 1. Noise was present in this situation, and we set signal-to-noise ratio (SNR) of CS system to 5 dB; the parameters used in reconstruction algorithm were $\lambda = 0.05\sqrt{2\log N}$, $t = 3k$, and other parameters were the same as above settings. It can be seen that in Figure 1a, although frequencies of the original signal were recovered accurately, there was some wrong information recovered as well, which resulted in recovery errors. However, in Figure 1b, it is obvious that the recovered result is more “clear” outside the range of $\text{supp}(x)$ than with the supp-BPDN algorithm.

Figure 1. Spectra of original and reconstructed signals recovered by supp-BPDN and BPDN. (a) Reconstruction result of frequencies by BPDN. (b) Reconstruction result of frequencies by the proposed algorithm.

The second simulation exhibits the signal recovery performances of BP, BPDN, MP, OMP, and supp-BPDN algorithms when the system signal-to-noise ratio (SNR) changed. Quality of reconstruction result was measured by mean square error (MSE):

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (s_i - \hat{s}_i)^2, \quad (47)$$

where $\hat{s}$ is the reconstructed signal and s is the original signal.

In the simulation, SNR of the system increased from -10 to 40 dB and the step size was 1 dB. Under each variation, the algorithms were executed 100 times; we obtained recovery errors and averaged them. Results are shown in Figure 2; the performance of supp-BPDN was better than those of BPDN and BP algorithms in the entire SNR range. Curves of supp-BPDN and BPDN exhibit the same tendency in general; in the range of -5 to 25 dB, recovery errors of BP are less than those of BPDN, caused by the overly pessimistic characteristic of BPDN upon encountering noise mentioned in Section 2.

Figure 2. Signal reconstruction errors with respect to SNR among supp-BPDN, BPDN, BP, OMP, and MP algorithms using the logarithmic (base 10) scale for the y -axis.

In the next simulation, reconstruction performances among supp-BPDN, BPDN, BP, and OMP algorithms were obtained for numbers of measurements M ranging from 50 to 150. Likewise, we set $N = 256$, $\lambda = 0.05\sqrt{2\log N}$, iterations to 100, and SNR to 15 dB. The results are demonstrated in Figure 3. Similarly, supp-BPDN performed the best in the whole range. The curve of supp-BPDN shows a little decrease as M increases; there are small fluctuations during the variation. Additionally, a stable period with a little rise is shown in the curve of BPDN.

Figure 3. Signal reconstruction error with respect to different numbers of measurements M among supp-BPDN, BPDN, BP, and OMP algorithms.

The last simulation aimed to exhibit signal recovery performances among supp-BPDN, BPDN, and OMP algorithms. Distinctively, in this simulation, we used a different noise model:

$$\mathbf{y}_{\text{noise}} = \mathbf{y} + \sigma \mathbf{w}, \quad (48)$$

where $\mathbf{w}$ was the normalized Gaussian white noise. In this model, λ was no longer a constant, but a variable corresponding to the noise level σ , and the mapping relation was

$$\lambda(\sigma) = \sigma \sqrt{2\log N}.$$

Similarly, we set $M = 64$, $N = 256$, and circulation times to 3,000. For every noise level, we repeated these algorithms 100 times. Recovery results are shown in Figure 4.

When the noise level is from 0.01 to 0.1, curves of BPDN and OMP possess a steeper slope; recovery error is more steady in supp-BPDN.

Figure 4. Signal recovery errors with respect to noise level among supp-BPDN, BPDN, and OMP algorithms.

5. CS System for Microwave Photonics

In this section, we present the design of a photonic-assisted CS system to verify the feasibility and superiority of the supp-BPDN algorithm in practical situations.

5.1. System Setup

The layout of the proposed scheme is shown in Figure 5. Firstly, a continuous wave (CW) laser is used to generate continuous wave regarded as an optical carrier propagating in single mode fiber (SMF). Additionally, the light is modulated by the sparse-frequency radio frequency (RF) signal $s(t)$ during the first MZM. After that, the optical signal carrying information enters the second MZM to be modulated another time. During this step, known PRBS patterns are encoded on the optical wave, yielding a multiplied signal whose RF signal is mixed with PRBS patterns to propagate into PD, transforming it into an electrical signal. Both MZMs are worked at the quadrature bias point. After PD and the amplifier, the signal is integrated during each period by the integrator. Finally, an ADC with a much lower sampling rate than the Nyquist frequency is applied to digitalize the measurements y , and recovery algorithms are used to reconstruct signals.

Figure 5. Scheme of the proposed CS system. CW laser: continuous wave laser; MZM: Mach-Zehnder modulator; PD: photodiode; RFA: amplifier; $\int$: integrator; ADC: analog-to-digital converter; dashed line: optical link (single mode fiber); solid line: electrical circuit.

According to the mathematical model of CS theory, we can realize that there are two steps to establish a CS system, which are mixing and integrating. In the scheme shown in Figure 5, the original signal is mixed with MM via MZMs in the photonic link, and integration process is realized in an electrical circuit. It is noteworthy that if MM is an $(M \times N)$ matrix, it means M measurement times are needed using M groups of PRBS. It should be noted that the length of the signal during each measurement period must be

the same as that of PRBS; otherwise, aliasing and measuring confusion will occur, which brings errors to the reconstruction results. There is no need to use the system every single measurement; we can get M number of measurements during one execution.

5.2. Results and Discussion

We set up a CS system shown in Figure 5 by OptiSystem and recovered a two-tone signal of 350 and 150 MHz by a 200-MHz ADC, and a two-tone signal of 1 GHz and 600 MHz with a 500-MHz ADC.

In the system, the output optical power of the first MZM is

$$P_{\text{out1}} = \frac{1}{2}[1 + s(t)], \quad (49)$$

where $s(t)$ is the normalized original signal, voltages of which range from -1 to 1 V. Additionally, after different PRBS modulate the optical wave, the output power of the second MZM is

$$P_{\text{out2}} = \frac{1}{2}r(t)[1 + s(t)], \quad (50)$$

in which $r(t)$ is the PRBS. Meanwhile, modulated waveforms of MZMs are shown in Figure 6, demonstrating a single period of signals in time domain. Figure 6a shows the original signal, and Figure 6b exhibits the original signal encoded by PRBS being measured.

Figure 6. Waveforms of modulated signals. (a) One period of original signal. (b) One period of mixed signal.

From Equation (50), after PD and the integrator, PRBS is integrated by a period as well. Therefore, data sampled by ADC are actually two integral values, which is

$$\hat{y}_m = \alpha \int_T \left\{ \frac{1}{2} r_m(t)[1 + s(t)] \right\} dt = \frac{\alpha}{2} \int_T [r_m(t)s(t)] dt + \frac{1}{2} \int_T r_m(t) dt, \quad (51)$$

where $\hat{y}_m$ is the m th measurement in practical systems, and $r_m(t)$ is the m th group of PRBS. α is a constant coefficient of the CS system, which is different when parameters change.

Figure 7 shows signals recovered by BPDN and supp-BPDN algorithms. The time interval of a period was 20 ns, the length of a group of PRBS was 256, measurement times $M = 64$, the system SNR was 15 dB, and the sampling rate of ADC was 200 MHz. We can figure out that signals recovered by supp-BPDN were the most coincident with the original signal.

Figure 7. Comparison of original and reconstructed signals under BPDN and supp-BPDN algorithms. Frequencies of signal were 350 MHz and 150 MHz; sampling rate was 200 MHz. (a) Signal recovered by BPDN. (b) Signal recovered by supp-BPDN.

In the next simulation, we changed the frequencies of signal to 1 GHz and 600 MHz, and the sampling rate of ADC was 500 MHz; recovery results are shown in Figure 8. Similarly, supp-BPDN demonstrated better performance at recovering signals.

Figure 8. Comparison of original and reconstructed signals using BPDN and supp-BPDN algorithms. Frequencies of signal were 1 GHz and 600 MHz; sampling rate was 500 MHz. (a) Signal recovered by BPDN. (b) Signal recovered by supp-BPDN.

6. Conclusions

A novel signal recovery algorithm named supp-BPDN is put forward in this paper based on the theoretical analysis of the signal support set. This algorithm executes the step of selecting the signal support set before the traditional recovery algorithm BPDN; in other words, it selects the position indexes corresponding to the top t values of the inner products and only retains the recovery information in said position indexes, which can effectively avoid errors from wrong information. In the proposed algorithm, it is not necessary to remove the contributions of local optimal values in measurements $\mathbf{y}$ during each iteration, which simplifies the procedure. By means of numerical simulations, we compared supp-BPDN and traditional algorithms, and supp-BPDN had the best performances. Moreover, based on the advantages of microwave photonic technologies, a photonic-enabled CS system was designed to verify the proposed algorithm. Additionally, a signal with a maximum frequency of 350 MHz can be recovered when digitalized at 200 MHz. Moreover, a signal with a peak of 1 GHz could be reconstructed when digitalized at 500 MHz, achieving a breakthrough in the Nyquist sampling rate. Similarly, signals recovered by supp-BPDN fit the original signals more closely.

Author Contributions: Conceptualization, W.L. and Z.W.; software, W.L.; validation, W.L., Z.W., G.L., and Y.J.; writing—original draft preparation, W.L.; writing—review and editing, Z.W.; supervision, Z.W. and Y.J. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China under grant 61835003.

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

Abbreviations

The following abbreviations are used in this manuscript:

CS	Compressive sensing
DOA	Direction of arrival
CR	Cognitive radio
MM	Measurement matrix
PRBS	Pseudo random bit sequence
ADC	Analog-to-digital converter
RIP	Restricted isometry property
PD	Photodiode
NPH	Non-deterministic polynomial hard
BP	Basis pursuit
BPDN	Basis pursuit denoising
SNR	Signal-to-noise ratio
MSE	Mean square error
RD	Random demodulator
SLM	Spatial light modulator
MZM	Mach–Zehnder modulator
MP	Matching pursuit
LPF	Low-pass filter
RIC	Restricted isometry constant
CW	Continuous wave
SMF	Single mode fiber
RF	Radio frequency
BCS	Bayesian compressive sensing
OMP	orthogonal matching pursuit

References

1. Emmanuel, J.C.; Romberg, J.; Tao, T. Robust Uncertainty Principles: Exact Signal Reconstruction From Highly Incomplete Frequency Information. *IEEE Trans. Inf. Theory* **2006**, *52*, 489–509.
2. David, L.D. Compressed Sensing. *IEEE Trans. Inf. Theory* **2006**, *52*, 1289–1306.
3. Cohen, D.; Eldar, Y.C. Sub-Nyquist Radar Systems. *IEEE Signal Process. Mag.* **2018**, *11*, 35–58. [[CrossRef](#)]
4. Marco F. Duarte; Mark A. Davenport; Dharmpal Takhar. Single-Pixel Imaging via Compressive Sampling. *IEEE Signal Process. Mag.* **2008**, *3*, 83–91.
5. Lustig, M.; Donoho, D.; Pauly, J.M. Sparse MRI: The Application of Compressed Sensing for Rapid MR Imaging. *Magn. Reson. Med.* **2007**, *58*, 1182–1195. [[CrossRef](#)]
6. Lei, C.; Wu, Y.; Sankaranarayanan, A.C. GHz Optical Time-Stretch Microscopy by Compressive Sensing. *IEEE Photon. J.* **2017**, *9*. [[CrossRef](#)]
7. Shen, Q.; Liu, W.; Cui, W.; Wu, S. Underdetermined DOA Estimation Under the Compressive Sensing Framework: A Review. *IEEE Access* **2017**, *4*, 8865–8878. [[CrossRef](#)]
8. Bazerque, J.A.; Giannakis, G.B. Distributed Spectrum Sensing for Cognitive Radio Networks by Exploiting Sparsity. *IEEE Trans. Signal Process.* **2010**, *58*, 1847–1862. [[CrossRef](#)]
9. Nyquist, H. Certain Factors Affecting Telegraph Speed. *Bell Syst. Tech. J.* **1928**, 124–130.
10. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
11. Tropp, J.A.; Laska, J.N.; Laska, J.N. Beyond Nyquist: Efficient Sampling of Sparse Bandlimited Signals. *IEEE Trans. Inf. Theory* **2010**, *27*, 520–544. [[CrossRef](#)]
12. Mishali, M.; Eldar, Y.C. From Theory to Practice: Sub-Nyquist Sampling of Sparse Wideband Analog Signals. *IEEE J-STSP* **2010**, *27*, 375–391. [[CrossRef](#)]
13. Nichols, J.M.; Bucholtz, F. Beating Nyquist with light: A compressively sampled photonic link. *Opt. Express* **2011**, *19*, 7339–7348. [[CrossRef](#)] [[PubMed](#)]
14. Chi, H.; Mei, Y.; Chen, Y. Microwave spectral analysis based on photonic compressive sampling with random demodulation. *Opt. Lett.* **2012**, *37*, 4636–4638. [[CrossRef](#)] [[PubMed](#)]
15. Chen, Y.; Chi, H.; Jin, T. Sub-Nyquist Sampled Analog-to-Digital Conversion Based on Photonic Time Stretch and Compressive Sensing With Optical Random Mixing. *J. Light. Technol.* **2013**, *31*, 3395–3401. [[CrossRef](#)]
16. Liang, Y.; Chen, M.; Chen, H. Photonic-assisted multi-channel compressive sampling based on effective time delay pattern. *Opt. Express* **2013**, *21*, 25700–25707. [[CrossRef](#)]

17. Bosworth, B.T.; Foster, M.A. High-speed ultrawideband photonic enabled compressed sensing of sparse radio frequency signals. *Opt. Lett.* **2013**, *38*, 4892–4895. [[CrossRef](#)]
18. Chi, H.; Chen, Y.; Mei, Y. Microwave spectrum sensing based on photonic time stretch and compressive sampling. *Opt. Lett.* **2013**, *38*, 136–138. [[CrossRef](#)]
19. Chen, Y.; Yu, X.; Chi, H. Compressive sensing in a photonic link with optical integration. *Opt. Lett.* **2014**, *39*, 2222–2224. [[CrossRef](#)]
20. Guo, Q.; Liang, Y.; Chen, M. Compressive spectrum sensing of radar pulses based on photonic techniques. *Opt. Express* **2015**, *23*, 4517–4522. [[CrossRef](#)]
21. Bosworth, B.T.; Stroud, J.R.; Tran, D.N. Ultrawideband compressed sensing of arbitrary multi-tone sparse radio frequencies using spectrally encoded ultrafast laser pulses. *Opt. Lett.* **2015**, *40*, 3045–3048. [[CrossRef](#)]
22. Chi, H.; Zhu, Z. Analytical Model for Photonic Compressive Sensing With Pulse Stretch and Compression. *IEEE Photon. J.* **2019**, *11*, 5500410. [[CrossRef](#)]
23. Zhu, Z.; Chi, H.; Jin, T. Photonics-enabled compressive sensing with spectral encoding using an incoherent broadband source. *Opt. Lett.* **2018**, *43*, 330–333. [[CrossRef](#)] [[PubMed](#)]
24. Valley, G.C.; Seffler, G.A.; Shaw, T.J. Compressive sensing of sparse radio frequency signals using optical mixing. *Opt. Lett.* **2012**, *37*, 4675–4677. [[CrossRef](#)] [[PubMed](#)]
25. Chen, S.S.; Donoho, D.L.; Saunders, M.A. Atomic Decomposition by Basis Pursuit. *SIAM Rev.* **2001**, *43*, 129–159. [[CrossRef](#)]
26. Hale, E.T.; Yin, W.; Zhang, Y. Fixed-Point Continuation for ℓ_1 -Minimization: Methodology and Convergence. *Siam J. Optim.* **2008**, *19*, 1107–1130. [[CrossRef](#)]
27. Mallat, S.G.; Zhang, Z. Matching pursuits with time-frequency dictionaries. *IEEE Trans. Signal Process.* **1993**, *41*, 3397–3415. [[CrossRef](#)]
28. Tropp, J.A.; Gilbert, A.C. Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit. *IEEE Trans. Inf. Theory* **2007**, *53*, 4655–4666. [[CrossRef](#)]
29. Needell, D.; Tropp, J.A. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *Appl. Comput. Harmon. Anal.* **2008**, *26*, 301–321. [[CrossRef](#)]
30. Lu, D.; Sun, G.; Li, Z.; Wang, S. Improved CoSaMP Reconstruction Algorithm Based on Residual Update. *J. Comput. Commun.* **2019**, *7*, 6–14. [[CrossRef](#)]
31. Dai, L.; Wang, Z.; Yang, Z. Compressive Sensing Based Time Domain Synchronous OFDM Transmission for Vehicular Communications. *IEEE J. Sel. Areas Commun.* **2013**, *31*, 460–469.
32. Zhang, W.; Li, H.; Kong, W.; Fan, Y.; Cheng, W. Structured Compressive Sensing Based Block-Sparse Channel Estimation for MIMO-OFDM Systems. *Wireless Pers. Commun.* **2019**, *108*, 2279–2309. [[CrossRef](#)]
33. Ji, S.; Xue, Y.; Carin, L. Bayesian Compressive Sensing. *IEEE Trans. Signal Process.* **2008**, *56*, 2346–2356. [[CrossRef](#)]
34. Wipf, D.P.; Rao, B.D. An Empirical Bayesian Strategy for Solving the Simultaneous Sparse Approximation Problem. *IEEE Trans. Signal Process.* **2007**, *55*, 3704–3716. [[CrossRef](#)]
35. Oikonomou, V.P.; Nikolopoulos, S.; Kompatsiaris, I. A Novel Compressive Sensing Scheme under the Variational Bayesian Framework. In Proceedings of the European Signal Processing Conference, 27th European Signal Processing Conference (EUSIPCO), A Coruna, Spain, 2–6 September 2019.
36. Candès, E.J. The restricted isometry property and its implications for compressed sensing. *CR Math.* **2008**, *346*, 589–592. [[CrossRef](#)]
37. Eldar, Y.C.; Kutyniok, G. *Compressed Sensing: Theory and Applications*, 3rd ed.; Cambridge University Press: Cambridge, UK, 2013; pp. 25–26, 349–351.
38. Baraniuk, R.; Davenport, M.; DeVore, R.; Wakin, M. A Simple Proof of the Restricted Isometry Property for Random Matrices. *Constr. Approx.* **2008**, *28*, 253–263. [[CrossRef](#)]
39. Cormen, T.H.; Leiserson, C.E.; Rivest, R.L.; Stein, C. *Introduction to Algorithms*, 3rd ed.; MIT Press: Cambridge, MA, USA, 2009; pp. 616–623.
40. Donoho, D.L. De-noising by soft-thresholding. *IEEE Trans. Inf. Theory* **1995**, *41*, 613–627. [[CrossRef](#)]