



2011

Decentralized Turbo Bayesian ompressed Sensing with application to UWB Systems

Depeng Yang

Husheng Li

Gregory D. Peterson

University of Tennessee - Knoxville

Follow this and additional works at: https://trace.tennessee.edu/utk_elecpubs

 Part of the [Electrical and Computer Engineering Commons](#)

Recommended Citation

EURASIP Journal on Advances in Signal Processing 2011, 2011:817947 doi:10.1155/2011/817947

This Article is brought to you for free and open access by the Engineering -- Faculty Publications and Other Works at TRACE: Tennessee Research and Creative Exchange. It has been accepted for inclusion in Faculty Publications and Other Works -- EECS by an authorized administrator of TRACE: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.

Research Article

Decentralized Turbo Bayesian Compressed Sensing with Application to UWB Systems

Depeng Yang, Husheng Li, and Gregory D. Peterson

Department of Electrical Engineering and Computer Science, The University of Tennessee, Knoxville, TN 37996, USA

Correspondence should be addressed to Depeng Yang, dyang7@utk.edu

Received 19 July 2010; Revised 1 February 2011; Accepted 28 February 2011

Academic Editor: Dirk T. M. Slock

Copyright © 2011 Depeng Yang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In many situations, there exist plenty of spatial and temporal redundancies in original signals. Based on this observation, a novel Turbo Bayesian Compressed Sensing (TBCS) algorithm is proposed to provide an efficient approach to transfer and incorporate this redundant information for joint sparse signal reconstruction. As a case study, the TBCS algorithm is applied in Ultra-Wideband (UWB) systems. A space-time TBCS structure is developed for exploiting and incorporating the spatial and temporal *a priori* information for space-time signal reconstruction. Simulation results demonstrate that the proposed TBCS algorithm achieves much better performance with only a few measurements in the presence of noise, compared with the traditional Bayesian Compressed Sensing (BCS) and multitask BCS algorithms.

1. Introduction

Compressed sensing (CS) theory [1, 2] is blooming in recent years. In the CS theory, the original signal is not directly acquired but reconstructed based on the measurements obtained from projecting the signal using a random sensing matrix. It is well known that most natural signals are sparse, that is, in a certain transform domain, most elements are zeros or have very small amplitudes. Taking advantage of such sparsity, various CS signal reconstruction algorithms are developed to recover the original signal from a few observations and measurements [3–5].

In many situations, there are multiple copies of signals that are correlated in space and time, thus providing spatial and temporal redundancies. Take the CS-based Ultra-Wideband (UWB) system as an example (A UWB system utilizes a short-range, high-bandwidth pulse without carrier frequency for communication, positioning, and radar imaging. One challenge is the acquisition of the high-resolution ultrashort duration pulses. The emergence of CS theory provides an approach to acquiring UWB pulses, possibly under the Nyquist sampling rate [6, 7].) [8, 9]. In a typical UWB system as shown in Figure 1, one transmitter periodically sends out ultrashort pulses (typically nano- or sub-nanosecond Gaussian pulses). Surrounding the transmitter,

several UWB receivers are receiving the pulses. The received echo signals at one receiver are similar to those received at other receivers in both space and time for the following reasons: (1) at the same time slot, the received UWB signals are similar to each other because they share the same source, which leads to spatial redundancy; (2) at the same receiver, the received signals are also similar in consecutive time slots because the pulses are periodically transmitted and propagating channels are assumed to change very slowly. Hence, the UWB echo signals are correlated both in space and time, which provides spatial and temporal redundancies and helpful information. Such *a priori* information can be exploited and utilized in the joint CS signal reconstruction to improve performance. On the other hand, our work is also motivated to reduce the number of necessary measurements and improve the capability of combating noise. For successful CS signal reconstruction, a certain number of measurements are needed. In the presence of noise, the number of measurement may be greatly increased. However, more measurements lead to more expensive and complex hardware and software in the system [6]. In such a situation, a question arises: can we develop a joint CS signal reconstruction algorithm to exploit temporal and spatial *a priori* information for improving performance in terms of less measurements, more noise tolerance, and better quality of reconstructed signal?

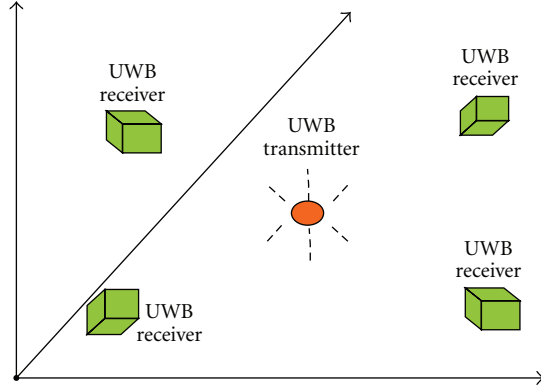


FIGURE 1: A typical UWB system with one transmitter and several receivers.

Related research about joint CS signal reconstruction has been developed in the literature recently. Distributed compressed sensing (DCS) [10, 11] studies joint sparsity and joint signal reconstruction. Simultaneous Orthogonal Matching Pursuit (SOMP) [12, 13] for simultaneous signal reconstruction is developed by extending the traditional Orthogonal Matching Pursuit (OMP) algorithm. Serial OMP [14] studies time sequence signal reconstruction. The joint sparse recovery algorithm [15] is developed in association with the basis pursuit (BP) algorithm. These algorithms focus on either temporal or spatial joint signal reconstruction. They are developed by extending convex optimization and linear programming algorithms but ignore the impact of possible noise in the measurements.

Other work on sparse signal reconstruction is based on a statistical Bayesian framework. In [16, 17], the authors developed a sparse signal reconstruction algorithm based on the belief propagation framework for the signal reconstruction. The information is exchanged among different elements in the signal vector in a way similar to the decoding of low-density parity check (LDPC) codes. In [18], the LDPC coding/decoding algorithm has been extended for real number CS signal reconstruction. Other Bayesian CS algorithms also have been developed in [3, 4, 19, 20]. In [3], a pursuit method in the Bernoulli-Gaussian model is proposed to search for the nonzero signal elements. A Bayesian approach for Sparse Component Analysis for the noisy case is presented in [4]. In [19], a Gaussian mixture is adopted as the prior distribution in the Bayesian model, which has similar performance as the algorithm in [21]. In [20], using a Laplace prior distribution in the hierarchical Bayesian model can reduce reconstruction errors than using the Gaussian prior distribution [21]. However, all these algorithms are designed only for a single signal reconstruction and are not applied for multiple simultaneous signal reconstruction. We are looking for a suitable prior distribution for mutual information transfer. The prior distributions proposed in [3, 19, 20] are too complex for exploiting redundancy information for joint signal reconstruction. In [22], the redundancies of UWB signals are incorporated into the framework of Bayesian Compressed Sensing (BCS) [5, 21] and have achieved

good performance. However, only a heuristic approach is proposed to utilize the redundancy in [22].

More related work for the joint sparse signal reconstruction includes [23], in which the authors proposed multitask Bayesian compressed sensing (MBCS) for simultaneous joint signal reconstruction by sharing the same set of hyperparameters for the signals. The mutual information is directly transferred over multiple simultaneous signal reconstruction tasks. The mechanism of sharing mutual information in [24] is similar to the MBCS [23]. This sharing scheme is effective and straightforward. For the signals with high similarity, it has a much better performance than the original BCS algorithm. However, for a low level of similarity, *a priori* information may adversely affect the signal reconstruction, resulting in much worse performance than the original BCS. In the situation where there exist lots of low-similarity signals, this disadvantage could be unacceptable.

Our work and MBCS [23] are both focused on reconstructing multiple signal frames. However, MBCS cannot perform simultaneous multitask signal reconstruction until all measurements have been collected, which is purely in a batch mode and cannot be performed in an online manner. Moreover, MBCS is centralized and is hard to decentralize. Our proposed incremental and decentralized TBCS has a more flexible structure, which can reconstruct multiple signal frames sequentially in time and/or in parallel in space through transferring mutual *a priori* information.

In this paper, we propose a novel and flexible Turbo Bayesian Compressed Sensing (TBCS) algorithm for sparse signal reconstruction through exploiting and integrating spatial and temporal redundancies in multiple signal reconstruction procedures performed in parallel, in serial, or both. Note the BCS algorithm has an excellent capability of combating noise by employing a statistically hierarchical structure, which is very suitable for transferring *a priori* information. Based on the BCS algorithm, we propose an *a priori* information-based iterative mechanism for information exchange among different reconstruction processes, motivated by the Turbo decoding structure, which is denoted as *Turbo BCS*. To the authors' best knowledge, there has not been any work applying the Turbo scheme in the BCS framework. Moreover, in the case study, we apply our TBCS algorithm in UWB systems to develop a *Space-Time Turbo Bayesian Compressed Sensing (STTBCS)* algorithm for space-time joint signal reconstruction. A key contribution is the space-time structure to exploit and utilize the temporal and spatial redundancies.

A primary challenge in the proposed framework is *how to yield and fuse a priori information in the signal reconstruction procedure in order to utilize spatial and temporal redundancies*. A mathematically elegant framework is proposed to impose an exponentially distributed hyperparameter on the existing hyperparameter α of the signal elements. This exponential distribution for the hyperparameter provides an approach to generate and fuse *a priori* information with measurements in the signal reconstruction procedure. An incremental method [25] is developed to find the limited nonzero signal elements, which reduces the computational complexity compared with

the expectation maximization (EM) method. A detailed STTBCS algorithm procedure in the case study of UWB systems is also provided to illustrate that our algorithm is universal and robust: when the signals have low similarities, the performance of STTBCS will automatically equal that of the original BCS; on the other hand, when the similarity is high, the performance of STTBCS is much better than the original BCS.

Simulation results have demonstrated that our TBCS significantly improves performance. We first use spike signals to illustrate the performance which can be achieved at each iteration employing the original BCS, MBCS, and our TBCS algorithms. It shows that our TBCS outperforms the original BCS and MBCS algorithms at each iteration for different similarity levels. We also choose IEEE802.15a [26] UWB echo signals for performance simulation. For the same number of measurements, the reconstructed signal using TBCS is much better compared with the original BCS and MBCS. To achieve the same reconstruction percentage, our proposed scheme needs significantly fewer measurements and is able to tolerate more noise, compared with the original BCS and MBCS algorithms. A distinctive advantage of TBCS is that when the similarity is low, MBCS performance is worse than the original BCS while our TBCS is close to the original BCS and much better than MBCS.

The remainder of this paper is organized as follows. The problem formulation is introduced in Section 2. Based on the BCS framework, *a priori* information is integrated into signal reconstruction in Section 3. A fast incremental optimization method is detailed in Section 4 for the *posterior* function. Taking UWB systems as a case study, Section 5 develops a space-time TBCS algorithm by applying our TBCS into the UWB system. The space-time TBCS algorithm is summarized in Section 5. Numerical simulation results are provided in Section 6. The conclusions are in Section 7.

2. Problem Formulation

Figure 2 shows a typical decentralized CS signal reconstruction model. We assume that the signals received at the receiver sides and the received signal are sparse. And we ignore any other effects such as propagation channel and additive noise on the original signal. We assume the received signals are sparse. Taking the UWB system as an example, all those original UWB echo signals, $s^{11}, s^{12}, s^{21}, \dots$, are naturally sparse in the time domain. These signals can be reconstructed in high resolution from a limited number of measurements using low sampling rate ADCs by taking advantage of CS theory. We define a *procedure* as a signal reconstruction process from measurements to recover the signal vector. Signal reconstruction procedures are performed distributively. We will develop a decentralized TBCS reconstruction algorithm to exploit and transfer mutual *a priori* information among multiple signal reconstruction procedures in time sequence and/or in parallel.

We assume that the time is divided into K frames. Temporally, a series of K original signal vectors at the first procedure is denoted as, s^{11}, s^{12}, \dots , and s^{1k} ($s^{1k} \in \mathbb{R}^N$), which can be correspondingly recovered from the

measurements y^{11}, y^{12}, \dots , and y^{1k} ($y^{1k} \in \mathbb{R}^N$) by using the projection matrix Φ^1 . All the measurement vectors are collected in time sequence. Spatially, at the same time slot, for example, the k th frame, a set of I original signal vectors, denoted as s^{1k}, s^{2k}, \dots , and s^{Ik} ($s^{ik} \in \mathbb{R}^N$), are needed to be reconstructed from the M -vector measurements, correspondingly y^{1k}, y^{2k}, \dots , and y^{Ik} ($y^{ik} \in \mathbb{R}^M$) by using the different projection matrix $\Phi^1, \Phi^2, \dots, \Phi^I$. All the spatial measurement vectors are collected at the same time.

The measurements are linear transforms of the original signals, contaminated by noise, which are given by

$$y^{ik} = \Phi^i s^{ik} + \epsilon^{ik}, \quad (1)$$

with $k = \{1, 2, \dots, K\}$ and $i = \{1, 2, \dots, I\}$; the matrix Φ^i , ($\Phi^i \in \mathbb{R}^{M \times N}$) is the projection matrix with $M \ll N$. The ϵ^{ik} are additive white Gaussian noise with unknown but stationary power β^{ik} . The noise level for different i and k may be different; however, the stationary noise variance can be integrated out in BCS and does not affect the signal reconstruction [5, 21, 25]. For mathematical convenience, we assume that the β^{ik} are identical for all i and k and denote it by β . Without loss of generality, we assume that s^{ik} is sparse, that is, most elements in s^{ik} are zero.

Signal reconstruction is performed among different BCS procedures in parallel and in time sequence. Information is transferred in parallel and serially. Note that the original signals, $s^{11}, s^{12}, s^{22}, \dots$, may be correlated with each other because of the spatial and temporal redundancies. However, without loss of generality, we do not specify the correlation model among the signals at different BCS procedures. This similarity leads to *a priori* information which can be introduced into decentralized TBCS signal reconstruction for improving performance in terms of reducing the number of measurements and improving the capability of combating noise.

For notational simplicity, we abbreviate s^{ik} into s^i to utilize one superscript representing either the temporal or spatial index, or both. We use the subscript to represent the element index in the vector. The main notation used throughout this paper is stated in Table 1.

3. Turbo Bayesian Compressed Sensing

In this section, we propose a Turbo BCS algorithm to provide a general framework for yielding and fusing *a priori* information from other parallel or serial reconstructed signals. We first introduce the standard BCS framework, in which selecting the hyperparameter α^i imposed on the signal element is the key issue. Then we impose an exponential prior distribution on the hyperparameter α^i with parameter λ^{fi} . The previous reconstructed signal element will impact the parameter λ^i to affect the α^i distribution, yielding *a priori* information. Next, *a priori* information will be integrated into the current signal estimation.

3.1. Bayesian Compressed Sensing Framework. Starting with Gaussian distributed noise, the BCS framework [5, 21] builds a Bayesian regression approach to reconstruct the

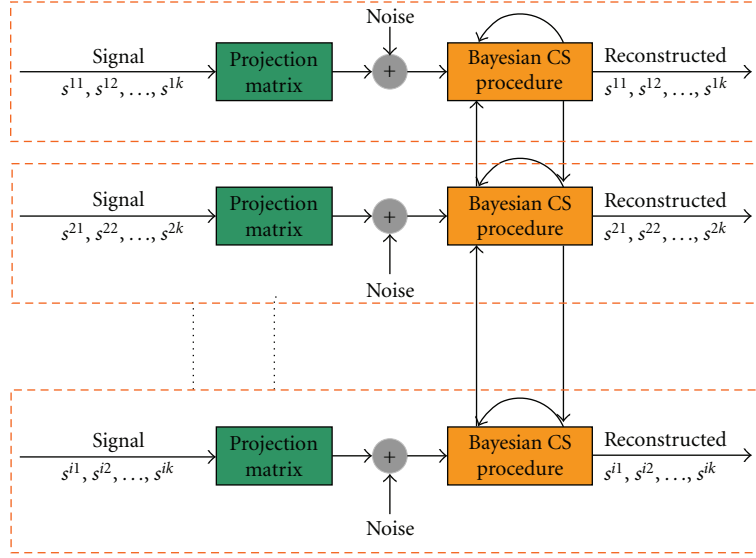


FIGURE 2: Block diagram of decentralized turbo Bayesian compressed sensing.

TABLE 1: Notation list.

$s_j^i, \mathbf{s}^i, \mathbf{s}$:	s_j^i is the j th signal element of the original signal vector \mathbf{s}^i at the i th spatial procedure or the i th time frame; the signal vector \mathbf{s}^i is $\mathbf{s}^i = \{s_j^i\}_{j=1}^N$, which can be abbreviated as \mathbf{s} .
$y_j^i, \mathbf{y}^i, \mathbf{y}$:	y_j^i is the j th element of the measurement vector for reconstructing the signal vector \mathbf{s}^i that is collected at either i th spatial procedure or i th time frame, which has $\mathbf{y}^i = \{y_j^i\}_{j=1}^M$; \mathbf{y}^i can be abbreviated as \mathbf{y} .
Φ^i :	The measurement matrix utilized for compressing the signal vector \mathbf{s}^i to yield \mathbf{y}^i .
β :	The noise variance.
$\alpha_j^i, \boldsymbol{\alpha}^i, \boldsymbol{\alpha}$:	α_j^i is the j th hyperparameter imposed on the corresponding signal element s_j^i ; it can be abbreviated as α_j , and it has $\boldsymbol{\alpha}^i = \{\alpha_j^i\}_{j=1}^N$; $\boldsymbol{\alpha}^i$ can be abbreviated as $\boldsymbol{\alpha}$.
$\lambda_j^i, \boldsymbol{\lambda}^i, \boldsymbol{\lambda}$:	λ_j^i is the parameter controlling the distribution of the corresponding hyperparameter α_j^i for mutual <i>a priori</i> information transfer, where $\boldsymbol{\lambda}^i = \{\lambda_j^i\}_{j=1}^N$ and it can be abbreviated as $\boldsymbol{\lambda}$.

original signal with additive noise from the compressed measurements. In the BCS framework, a Gaussian prior distribution is imposed over each signal element, which is given by

$$P(\mathbf{s}^i | \boldsymbol{\alpha}^i) = \prod_{j=1}^N \left(\frac{\alpha_j^i}{2\pi} \right)^{1/2} \exp \left(-\frac{(s_j^i)^2 \alpha_j^i}{2} \right) \quad (2)$$

$$\sim \prod_{j=1}^N \mathcal{N} \left(s_j^i | 0, (\alpha_j^i)^{-1} \right),$$

where α_j^i is the hyperparameter for the signal element s_j^i . The zero-mean Gaussian *priori* is independent for each signal element. By applying Bayes' rule, the *a posteriori* probability of the original signal is given by

$$P(\mathbf{s}^i | \mathbf{y}^i, \boldsymbol{\alpha}^i, \beta) = \frac{P(\mathbf{y}^i | \mathbf{s}^i, \beta) P(\mathbf{s}^i | \boldsymbol{\alpha}^i)}{P(\mathbf{y}^i | \boldsymbol{\alpha}^i, \beta)} \quad (3)$$

$$\sim \mathcal{N}(\mathbf{s}^i | \boldsymbol{\mu}^i, \boldsymbol{\Sigma}^i),$$

where $\mathbf{A} = \text{diag}(\boldsymbol{\alpha}^i)$. The covariance and the mean of the signal are given by

$$\boldsymbol{\Sigma}^i = \left(\beta^{-2} (\Phi^i)^T \Phi^i + \mathbf{A} \right)^{-1}, \quad (4)$$

$$\boldsymbol{\mu}^i = \beta^{-2} \boldsymbol{\Sigma}^i (\Phi^i)^T \mathbf{y}^i.$$

Then, we obtain the estimation of the signal, $\hat{\mathbf{s}}^i$, which is given by

$$\hat{\mathbf{s}}^i = \left((\Phi^i)^T \Phi^i + \beta^2 \mathbf{A} \right)^{-1} (\Phi^i)^T \mathbf{y}^i. \quad (5)$$

In order to estimate the hyperparameters $\boldsymbol{\alpha}^i$ and \mathbf{A} , the maximum likelihood function based on observations is given by

$$\boldsymbol{\alpha}^i = \arg \max_{\boldsymbol{\alpha}^i} P(\mathbf{y}^i | \boldsymbol{\alpha}^i, \beta) \quad (6)$$

$$= \arg \max_{\boldsymbol{\alpha}^i} \int P(\mathbf{y}^i | \mathbf{s}^i, \beta) P(\mathbf{s}^i | \boldsymbol{\alpha}^i) d\mathbf{s}^i,$$

where, by integrating out s^i and maximizing the *posterior* with respect to α^i , the hyperparameter diagonal matrix \mathbf{A} is estimated. Then, the signal can be reconstructed using (5).

The matrix \mathbf{A} plays a key role in the signal reconstruction. The hyperparameter diagonal matrix \mathbf{A} can be used to transfer the mutual *a priori* information by sharing the same \mathbf{A} among all signals [23]. In such a way, if signals have many common nonzero elements, the signal reconstruction will benefit from such a similarity. However, when the similarity level is low, the transferred “wrong” information may impair the signal reconstruction [23].

Alternatively, we find a soft approach to integrating *a priori* information in a robust way. An exponential prior distribution is imposed on the hyperparameter α^i controlled by the parameter λ^i . The previously reconstructed signal elements will impact the λ^i and change the α^i distribution to yield *a priori* information. Then, the hyperparameter α^i conditioned on λ^i will join the current signal estimation using the maximum a posteriori (MAP) criterion, which is to fuse *a priori* information.

3.2. Yielding A Priori Information. The key idea of our TBCS algorithm is to *impose an exponential distribution on the hyperparameter α_j^i and exchange information among different BCS signal reconstruction procedures using the exponential distribution in a turbo iterative way.* In each iteration, the information from other BCS procedures will be incorporated into the exponential *a priori* and then used for the signal reconstruction of the current BCS signal reconstruction procedure being considered. Note that, in the standard BCS [21], a Gamma distribution with two parameters is used for α_j^i . The reason we adopt an exponential distribution here is that we need to handle only one parameter for the exponential distribution, which is much simpler than the Gamma distribution, while both distributions belong to the same family of distributions.

We assume that hyperparameter α_j^i satisfies the exponential prior distribution, which is given by

$$P(\alpha_j^i | \lambda_j^i) = \begin{cases} \lambda_j^i e^{-\lambda_j^i \alpha_j^i} & \text{if } \alpha_j^i \geq 0, \\ 0 & \text{if } \alpha_j^i < 0, \end{cases} \quad (7)$$

where λ_j^i ($\lambda_j^i > 0$) is the hyperparameter of the hyperparameter α_j^i . By assuming mutual independence, we have that

$$P(\alpha^i | \lambda^i) = \left(\prod_{j=1}^N \lambda_j^i \right) \exp \left(\sum_{j=1}^N -\lambda_j^i \alpha_j^i \right). \quad (8)$$

By choosing the above exponential prior, we can obtain the marginal probability distribution of the signal element depending on the parameter λ_j^i by integrating α_j^i out, which is given by

$$\begin{aligned} P(s_j^i | \lambda_j^i) &= \int P(s_j^i | \alpha_j^i) P(\alpha_j^i | \lambda_j^i) d\alpha_j^i \\ &= (2\pi)^{-(1/2)} \Gamma\left(\frac{3}{2}\right) \lambda_j^i \left(\lambda_j^i + \frac{(s_j^i)^2}{2} \right)^{-(3/2)}, \end{aligned} \quad (9)$$

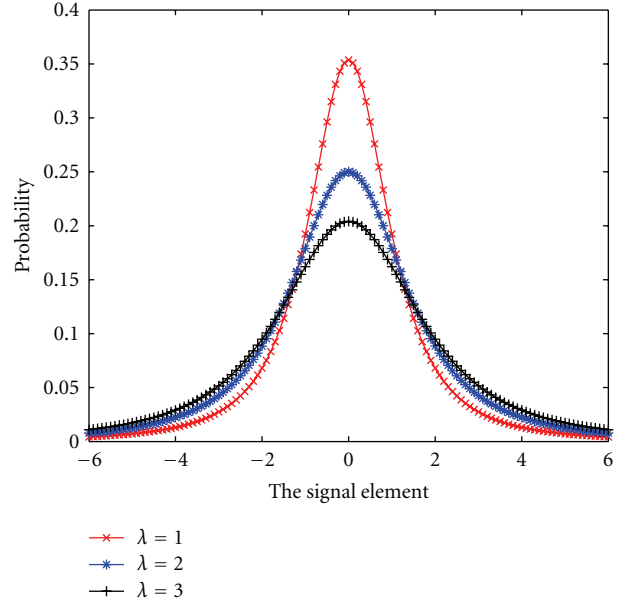


FIGURE 3: The distribution $P(s_j^i | \lambda_j^i)$.

where $\Gamma(\cdot)$ is the gamma function, defined as $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$. The detailed derivation is shown in Appendix A.

Figure 3 shows the signal element distribution conditioned on the hyperparameter λ_j^i . Obviously, the bigger the parameter λ_j^i is, the more likely the corresponding signal element can take a larger value. Intuitively, this looks very much like a Laplace prior which is sharply peaked at zero [20]. Here, λ_j^i is the key of introducing *a priori* information based on reconstructed signal elements.

Compared with the Gamma prior distribution imposed on the hyperparameter λ_j^i [21, 25], the exponential distribution has only one parameter while the Gamma distribution has two degrees of freedom. In many applications (e.g., communication networks), transferring one parameter is much easier and cheaper using the exponential distribution than handling two parameters. The exponential prior distribution does not degrade the performance, which can encourage the sparsity (see Appendix A). Also, using the exponential distribution is computationally tractable, which can produce *a priori* information for mutual information transfer.

Now the challenge is, given the j th reconstructed signal element s_j^b from the b th BCS procedure, how one yields *a priori* information to impact the hyperparameters in the i th BCS procedure for reconstructing the j th signal element s_j^i . When multiple BCS procedures are performed to reconstruct the original signals (no matter whether they are in time sequence or in parallel), the parameters of the exponential distribution, λ_j^i , can be used to convey and incorporate *a priori* information from other BCS procedures. To this end, we consider the conditional probability, $P(\alpha_j^i | s_j^b, \lambda_j^i)$, for α_j^i , given an observation element from another BCS procedure, s_j^b ($b \neq i$), and λ_j^i . Since the proposed algorithm

does not use a specific model for the correlation of signals at different BCS procedures, we propose the following simple assumption when incorporating the information from other BCS procedures into λ_j^i , for facilitating the TBCS algorithm.

Assumption. For different i and b , we assume that $\alpha_j^i = \alpha_j^b$, for all i, b .

Essentially, this assumption implies the same locations of nonzero elements for different BCS procedures. In other words, the hyperparameter α_j^i for the j th signal element is the same over different signal reconstruction procedures. Then, mutual information can be transferred through the shared hyperparameter α_j^i as proposed in [23]. However, the algorithm in [23] is a centralized MBCS algorithm, so the signal reconstructions for different tasks cannot be performed until all measurements are collected. Note that this technical assumption is only for deriving the algorithm for information exchange. It does not mean that the proposed algorithm only works for the situation in which all signals share the same locations of nonzero elements. Our proposed algorithm based on this assumption can provide a flexible and decentralized way to transfer mutual *a priori* information.

Based on the assumption, we obtain

$$\begin{aligned} P(\alpha_j^i | s_j^b, \lambda_j^i) &= \frac{P(s_j^b, \alpha_j^i | \lambda_j^i)}{P(s_j^b, \lambda_j^i)} \\ &= \frac{P(s_j^b | \alpha_j^i) P(\alpha_j^i | \lambda_j^i)}{\int P(s_j^b | \alpha_j^i) P(\alpha_j^i | \lambda_j^i) d\alpha_j^i} \\ &= \frac{(\lambda_j^i + (s_j^b)^2/2)^{3/2} \exp(-(\lambda_j^i + (s_j^b)^2/2)\alpha_j^i)}{\Gamma(3/2)} \\ &= \frac{(\tilde{\lambda}_j^i)^{3/2} \exp(-\tilde{\lambda}_j^i \alpha_j^i)}{\Gamma(3/2)}, \end{aligned} \quad (10)$$

where $\Gamma(\cdot)$ is the gamma function, defined as $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$. The detailed derivation is given in Appendix A.

Obviously, the posterior $(\alpha_j^i | s_j^b, \lambda_j^i)$ also belongs to the exponential distribution [27]. Compared with the original *prior* distribution in (7), given the j th reconstructed signal element s_j^b from the b th BCS procedure, the hyperparameter λ_j^i in the i th BCS procedure controlling *a priori* distribution is actually updated to $\tilde{\lambda}_j^i$, which is given by

$$\tilde{\lambda}_j^i = \lambda_j^i + \frac{(s_j^b)^2}{2}. \quad (11)$$

If the information from n BCS procedures b_1, \dots, b_n is introduced, the parameter $\tilde{\lambda}_j^i$ is then updated to

$$P(\alpha_j^i | s_j^{b_1}, s_j^{b_2}, \dots, s_j^{b_n}, \lambda_j^i) = \frac{(\tilde{\lambda}_j^i)^{(2n+1)/2} \exp(-\tilde{\lambda}_j^i \alpha_j^i)}{\Gamma((2n+1)/2)}, \quad (12)$$

where

$$\tilde{\lambda}_j^i = \lambda_j^i + \frac{\sum_{i=1}^n (s_j^{b_i})^2}{2}. \quad (13)$$

The derivation details are given in Appendix A.

Equations (11) and (13) show how the single or multiple signal elements s_j^b , $j = 1, 2, \dots, N$, $n = 1, 2, \dots$, from other BCS procedures impact the hyperparameter of the signal element s_j^i , $j = 1, 2, \dots, N$ at the same location in the i th BCS signal reconstruction. Note that the b th BCS signal reconstruction may be previously performed or is ongoing with respect to the i th BCS procedure. This provides significant flexibility to apply our TBCS in different situations.

3.3. Incorporating A Priori Information into BCS. Now, we study how to incorporate the *a priori* information obtained in the previous subsection into the signal reconstruction procedure. In order to incorporate *a priori* information, provided by the external information, we maximize the log posterior based on (6), which is given by

$$\begin{aligned} L(\alpha^i) &= \log P(\mathbf{y}^i | \alpha^i, \beta) P(\alpha^i | \{\mathbf{s}^b\}, \lambda^i) \\ &= \log P(\mathbf{y}^i | \alpha^i, \beta) + \log P(\alpha^i | \{\mathbf{s}^b\}, \lambda^i). \end{aligned} \quad (14)$$

Therefore, the estimation of α^i not only depends on the local measurements, which are in the first term $\log P(\mathbf{y}^i | \alpha^i, \beta)$, but also relies on the external signal elements $\{\mathbf{s}^b\}$ through the parameter λ^i , which are in the second term $\log P(\alpha^i | \{\mathbf{s}^b\}, \lambda^i)$.

An expectation maximization (EM) method can be utilized for the signal estimation. Recall that the signal vector \mathbf{s}^i is Gaussian distributed conditioned on α^i , while α^i also conditionally depends on the parameters λ^i . Equation (3) shows that the conditional distribution of \mathbf{s}^i satisfies $\mathcal{N}(\mu, \Sigma)$. Then, applying a similar argument to that in [21], we consider \mathbf{s}^i as hidden data and then maximize the following posterior expectation, which is given by

$$E_{\mathbf{s}^i | \mathbf{y}^i, \alpha^i} [\log P(\mathbf{s}^i | \alpha^i, \beta) P(\alpha^i | \lambda^i)]. \quad (15)$$

By differentiating (15) with respect to α^i and setting the differentiation to zero, we obtain an update, which is given by

$$\alpha_j^i = \frac{3}{(s_j^i)^2 + \Sigma_{jj}^i + 2\lambda_j^i}, \quad (16)$$

where Σ_{jj}^i is the j th diagonal element in the matrix Σ^i . The detail of the derivation is given in Appendix B. Basically, the hyperparameters α^i are interactively estimated and most of them will tend to infinity, which means that most corresponding signal elements are zero. Only the nonzero signal elements are estimated.

Considering the computation of the matrix inverse (with complexity $O(n^3)$) associated with the process, the EM algorithm has a large computational cost. Even though a

Cholesky decomposition can be applied to alleviate the calculation [28, 29], the EM method still incurs a significant computational cost. We will provide an incremental method for the optimization to reduce the computational cost.

4. Incremental Optimization

In this section, we utilize an incremental optimization to incorporate transferred *a priori* information and optimize the posterior function. Due to the inherent sparsity of the signal, the incremental method finds the limited nonzero elements by separating and testing a single index one by one, which alleviates the computational cost compared with the EM algorithm. Note that the key principle is similar to that of the fast relevance vector machine algorithm in [21]. However, the incorporation of the hyperparameter λ^i brings significant difficulty for deriving the algorithm. For convenience, we abbreviate α^i as α and \mathbf{y}^i as \mathbf{y} because we are focusing on the current signal estimation.

In order to introduce *a priori* knowledge, the target log posterior function can be written as

$$\begin{aligned}\alpha &= \arg \max_{\alpha} L(\alpha) \\ &= \arg \max_{\alpha} (\log P(\mathbf{y} | \alpha, \beta^2) P(\alpha | \mathbf{x})) \\ &= \arg \max_{\alpha} (\log P(\mathbf{y} | \alpha, \beta^2) + \log P(\alpha | \{\mathbf{s}^b\}, \lambda)) \\ &= \arg \max_{\alpha} (L_1(\alpha) + L_2(\alpha)),\end{aligned}\quad (17)$$

where $L_1(\alpha)$ is the term of signal estimation from local observation and $L_2(\alpha)$ introduces *a priori* information from other external BCS procedures.

In contrast to the complex EM optimization, the incremental algorithm starts by searching for a nonzero signal element and iteratively adds it to the candidate index set for the signal reconstruction, an algorithm which is similar to the greedy pursuit algorithm. Hence, we isolate one index, assuming the j th element, which is given by

$$\begin{aligned}L(\alpha) &= L(\alpha_{-j}) + l(\alpha_j) \\ &= L_1(\alpha_{-j}) + l_1(\alpha_j) + L_2(\alpha_{-j}) + l_2(\alpha_j),\end{aligned}\quad (18)$$

where $l_1(\alpha_j)$ is the separated term associated with the j th element from the posterior function $L(\alpha_j)$. The remaining term is $L_1(\alpha_{-j})$, resulting from removing the j th index.

Initially, all the hyperparameters λ_j , $j = \{1, 2, \dots, N\}$, are set to zero. When the transferred signal elements are zero, that is, $s_j^b = 0$, $j = \{1, 2, \dots, N\}$, the updated hyperparameters will also be zeros, that is, $\tilde{\lambda}_j^i = 0$, $j = \{1, 2, \dots, N\}$, according to (11) and (13). This implies no *prior* information and the term $L_2(\alpha) = 0$ based on (7), which is equivalent to the original BCS algorithm [5, 25].

Suppose that the external information from other BCS procedures is incorporated, that is, $s_j^b \neq 0$, $\tilde{\lambda}_j^i \neq 0$, and $L_2(\alpha) \neq 0$. We target maximizing the separated term by

considering the remaining term $L(\alpha_{-j})$ as fixed. Then, the posterior function separating a single index is given by

$$\begin{aligned}l(\alpha_j) &= l_1(\alpha_j) + l_2(\alpha_j) \\ &= \frac{1}{2} \left(\log \frac{\alpha_j}{\alpha_j + g_j} + \frac{h_j^2}{\alpha_j + g_j} \right) + \log \lambda_j - \lambda_j \alpha_j,\end{aligned}\quad (19)$$

where

$$\begin{aligned}g_j &= \phi_j^T E_{-j}^{-1} \phi_j, \\ h_j &= \phi_j^T E_{-j}^{-1} \mathbf{y}, \\ E_{-j} &= \beta^2 I + \sum_{k \neq j} \alpha_k^{-1} \phi_k \phi_k^{-1},\end{aligned}\quad (20)$$

where ϕ_j is the j th column vector of the matrix Φ . The detailed derivation is provided in Appendix C. Then, we seek for a maximum of the posterior function, which is given by

$$\alpha_j^* = \arg \max_{\alpha_j} l(\alpha_j) = \arg \max_{\alpha_j} (l_1(\alpha_j) + l_2(\alpha_j)). \quad (21)$$

When there is no external information incorporated, the optimal hyperparameter is given by [25]

$$\alpha_j^* = \arg \max_{\alpha_j} (l_1(\alpha_j)), \quad (22)$$

where

$$\alpha_j^* = \begin{cases} \frac{h_j^2}{g_j^2 - h_j}, & \text{if } g_j^2 > h_j, \\ \infty, & \text{otherwise.} \end{cases} \quad (23)$$

When external information is incorporated, to maximize the target function (19), we compute the first-order derivative of $l(\alpha_j)$, which is given by

$$\begin{aligned}l'(\alpha_j) &= \frac{g_j}{2\alpha_j(\alpha_j + g_j)} - \frac{h_j^2}{2(\alpha_j + g_j)^2} - \lambda_j \\ &= \frac{f(\alpha_j, g_j, h_j, \lambda_j)}{\alpha_j(\alpha_j + g_j)^2},\end{aligned}\quad (24)$$

where $f(\alpha_j, g_j, h_j, \lambda_j)$ is a cubic function with respect to α_j . By setting (24) to zero, we get the optimum α_j^* .

By setting (24) to zero, we get the optimum solution for the posterior likelihood function $l(\alpha_j)$, which is given by

$$\alpha_j = \begin{cases} \alpha_j^*, & \text{if } g_j^2 > h_j, \\ \infty, & \text{otherwise.} \end{cases} \quad (25)$$

The details are given in Appendix D.

Therefore, in each iteration only one signal element is isolated and the corresponding parameters are evaluated. After several iterations, most of the nonzero signal elements are selected into the candidate index set. Due to the sparsity of the signal, after a limited number of iterations, only a few signal elements are selected and calculated, which greatly increases the computational efficiency.

5. Case Study: Space-Time Turbo Bayesian Compressed Sensing for UWB Systems

The TBCS algorithm can be applied in various applications. A typical application is the UWB communication/positioning system. Our proposed TBCS algorithm will be applied to the UWB system to fully exploit the redundancies in both space and time, which is called *Space-Time Turbo BCS (STTBCS)*. In this section, we first introduce the UWB signal model. Then, the structure to transfer spatial and temporal *a priori* information in the CS-based UWB system is explained in detail. Finally, we summarize the STTBCS algorithm.

5.1. UWB System Model. In a typical UWB communication/positioning system, suppose that there is only one transmitter, which transmits UWB pulses on the order of nano- or sub-nanosecond. As shown in Figure 1, several receivers, or base stations, are responsible for receiving the UWB echo signals. The time is divided into frames. The received signal at the i th base station and the k th frame in the continuous time domain is given by

$$\mathbf{s}^{ik}(t) = \sum_{l=1}^L a_l p'(t - t_l), \quad (26)$$

where L is the number of resolvable propagation paths, a_l is the attenuation coefficient of the l th path, and t_l is the time delay of the l th propagation path. We denote by $p(t)$ the transmitted Gaussian pulse and by $p'(t)$ the corresponding received pulse which is close to the original pulse waveform but has more or less distortions resulting from the frequency-dependent propagation channels. At the same frame or time slot, there is only one transmitter but multiple receivers which are closely placed in the same environment. Therefore, the received echo UWB signals at different receivers are similar at the same time, thus incurring spatial redundancy. In other words, the received signals share many common nonzero element locations. Typically, around 30–70% of nonzero element indices are the same in one frame according to our experimental observation [30]. In particular, no matter what kind of signal modulation is used for UWB communication, such as pulse amplitude modulation (PAM), on-off keying (OOK), or pulse position modulation (PPM), the UWB echo signals among receivers are always similar, and thus the spatial redundancy always exists. In this case, the spatial redundancy can be exploited for good performance using the proposed space TBCS algorithm.

In one base station, the consecutively received signals can also be similar. Suppose that, in UWB positioning systems, the pulse repetition frequency is fixed. When the transmitter moves, the signal received at the i th base station and the $(k + 1)$ th frame can be written as

$$\mathbf{s}^{i(k+1)}(t) = \sum_{l=1}^{L'} a'_l p'(t - \tau - t_l). \quad (27)$$

Compared with (26), τ stands for the time delay which comes from the position change of the transmitter. In high precision

positioning/tracking systems, this τ is always relatively small, which makes the consecutive received signals similar. Due to the similar propagation channels, the numbers L and L' , as well as a_l and a'_l , are similar in consecutive frames. This leads to the temporal redundancy. In our experiments, about 10–60% of the nonzero element locations in two consecutive frames are the same [30]. Then, this temporal redundancy can be exploited for good performance by using the Time TBCS algorithms. Actually, there exist both spatial and temporal redundancies in the UWB communication/positioning system. Therefore we can utilize the STBCS algorithm for good performance.

To archive a high precision of positioning and a high speed communication rate, we have to acquire ultrahigh resolution UWB pulses, which demands ultrahigh sampling rate ADCs. For instance, it requires picosecond level time information and 10 G sample/s or even higher sampling rate ADCs to achieve millimeter (mm) positioning accuracy for UWB positioning systems [28], which is prohibitively difficult. UWB echo signals are inherently sparse in the time domain. This property can be utilized to alleviate the problem of an ultrahigh sampling rate. Then the high-resolution UWB pulses can be indirectly obtained and reconstructed from measurements acquired using lower sampling rate ADCs.

The system model of the CS-based UWB receiver can use the same model as that in Figure 2. The received UWB signal at the i th base station is first “compressed” using an analog projection matrix [6]. The hardware projection matrix consists of a bank of Distributed Amplifiers (DAs). Each DA functions like a wideband FIR filter with different configurable coefficients [6]. The output of the hardware projection matrix can be obtained and digitized by the following ADCs to yield measurements. For mathematical convenience, the noise generated from the hardware and ADCs is modeled as Gaussian noise added to the measurements. When several sets of measurements are collected at different base stations, a joint UWB signal reconstruction can be performed. This process is modeled in (1).

5.2. STTBCS: Structure and Algorithm. We apply the proposed TBCS to UWB systems to develop the STTBCS algorithm. Figure 4 illustrates the structure of our STTBCS algorithm and explains how mutual information is exchanged. For simplicity, only two base stations (BS1 and BS2) and two consecutive frames of UWB signals (the k th and $(k + 1)$ th) in each base station are illustrated. For each BCS procedure, Figure 4 also depicts the dependence among measurements, noise, signal elements, and hyperparameters.

In the STTBCS, multiple BCS procedures in multiple time slots are performed. Between BS1 and BS2, the signal reconstruction for $\mathbf{s}^{1(k+1)}$ and $\mathbf{s}^{2(k+1)}$ is carried out simultaneously while the information in \mathbf{s}^{1k} and \mathbf{s}^{2k} , the previous frame, is also used.

Algorithm 1 shows the details of the STTBCS algorithm. We start with the initialization of the noise, hyperparameters α , and the candidate index set Ω (an index set containing all possibly nonzero element indices). Then, the information

- (1) The hyperparameter α is set to $\alpha = [\infty, \dots, \infty]$.
 The candidate index set $\Omega = \emptyset$.
 The noise is initialized to a certain value without any *prior* information, or utilize the previous estimated value.
 The parameter of the hyperparameter $\lambda : \lambda = [0, \dots, 0]$;
- (2) Update λ using (11) and (13) from the previous reconstructed nonzero signal elements.
 This introduces temporal *a priori* information.
- (3) **repeat**
- (4) Check and receive the ongoing reconstructed signal elements from other simultaneous BCS reconstruction procedures to update the parameter λ ; this is to fuse spatial *a priori* information.
- (5) Choose a random j th index; Calculate the corresponding parameter g_j and h_j as shown in (C.4) and (C.5).
- (6) **if** $(g_j)^2 > h_j$ and $\tilde{\lambda}_j \neq 0$ **then**
- (7) Add a candidate index: $\Omega = \Omega \cup j$;
- (8) Update α_j by solving (24).
- (9) **else**
- (10) **if** $(g_j)^2 > h_j$ and $\tilde{\lambda}_j = 0$ **then**
- (11) Add a candidate index: $\Omega = \Omega \cup j$
- (12) Update α_i using (23).
- (13) **else if** $(g_j)^2 < h_j$ **then**
- (14) Delete the candidate index: $\Omega = \Omega \setminus \{j\}$ if the index is in the candidate set.
- (15) **end if**
- (16) **end if**
- (17) Compute the signal coefficients \mathbf{s}_Ω in the candidate set using (5).
- (18) Send out the ongoing reconstructed signal elements \mathbf{s}_Ω to other BCS procedures as spatial *a priori* information.
- (19) **until** converged
- (20) Re-estimate the noise level using (28) and send out the noise level for the next usage.
- (21) Send out the reconstructed nonzero signal elements for the next time utilization as temporal *a priori* information.
- (22) Return the reconstructed signal.

ALGORITHM 1: Space-time turbo bayesian compressed sensing algorithm.

from previous reconstructed signals and from other base stations is utilized to update the hyperparameter λ . The terms g_j and h_j are also computed. The term $g_j^2 > h_j$ is then used to add the j th element from the candidate index set. A convergence criterion is used to test whether the differences between successive values for any α_j , $j = \{1, 2, \dots, N\}$ are sufficiently small compared to a certain threshold. When the iterations are completed, the noise level β will be reestimated from setting $\partial L / \partial \beta = 0$ using the same method in [21], which is given by

$$(\beta^2)^{\text{new}} = \frac{\|y - \Sigma \mathbf{s}\|^2}{N - \sum_{i=1}^M (1 - \alpha_i \Sigma_{ii})}, \quad (28)$$

where Σ_{ii} is the diagonal element in the matrix Σ . The details of the above STTBCS algorithm are summarized

in Algorithm 1. Note that only the nonzero signal element which is shown from the local measurements can introduce *a priori* information and thus update the hyperparameter $\tilde{\lambda}_j$. In other words, only if it satisfies $g_j^2 > h_j$ can the parameter $\tilde{\lambda}_j$ be updated. This avoids the adverse effects from wrong *a priori* information to add a zero signal element into the candidate index set.

6. Simulation Results

Numerical simulations are conducted to evaluate the performance of the proposed TBCS algorithm, compared with the MBCS [23] and original BCS algorithms [5]. We use spike signals and experimental UWB echo signals [26] for the performance test. The quality of the reconstructed signal

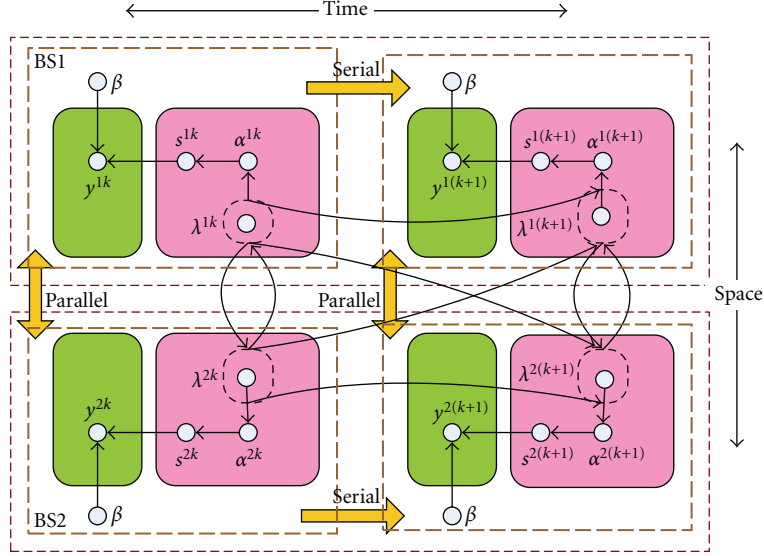


FIGURE 4: Block diagram of space-time turbo Bayesian compressed sensing.

is measured in terms of the reconstruction percentage, which is defined as

$$1 - \frac{\|\mathbf{s} - \tilde{\mathbf{s}}\|_2}{\|\mathbf{s}\|_2}, \quad (29)$$

where \mathbf{s} is the true signal and $\tilde{\mathbf{s}}$ is the reconstructed signal.

Our TBCS algorithm performance is largely determined by how the introduced signal is similar to the objective signal. In other words, we consider how many common nonzero element locations are shared between the objective signal and the introduced signals. Then we define the similarity as

$$P_s = \frac{K_{\text{com}}}{K_{\text{obj}}}, \quad (30)$$

where K_{obj} is the number of nonzero signal elements in the objective unrecovered signal, K_{com} is the number of the common nonzero element locations among the transferred reconstructed signals and objective signal, and P_s represents the similarity level as a percentage. Note that, without loss of generality, we only consider the relative number of common nonzero element locations to measure the similarity, ignoring any amplitude correlation. Hence, when $P_s = 100\%$, it does not mean that the signals are the same but means that they have the same nonzero element locations; the amplitudes may not be the same.

Our TBCS algorithm performance is compared with MBCS and BCS using different types of signals, different similarity levels, noise powers, and measurement numbers.

6.1. Spike Signal. We first generate four scenarios of spike signals with the same length $N = 512$, which have the same number of 20 nonzero signal elements with random locations and Gaussian distributed (mean = 0, variance = 1) amplitudes. One spike signal is selected as the objective signal, as shown in Figure 5. With respect to the objective signal, the other three signals have a similarity of 25%, 50%, and 75%, which will be introduced as *a priori* information.

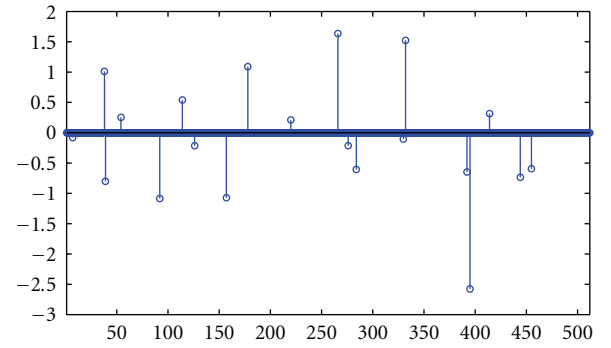


FIGURE 5: Spike signal with 20 nonzero elements in random locations.

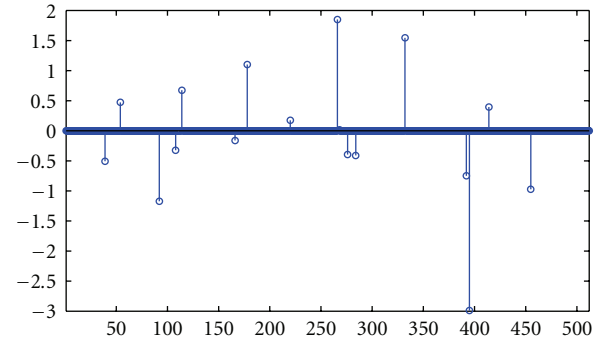


FIGURE 6: Reconstructed spike signal using MBCS with 75% similarity.

The objective signal is then reconstructed using the original BCS, MBCS, and TBCS algorithms, respectively, with the same number of measurements ($M = 62$) and the same noise variance 0.15 ($\text{SNR} \approx 6$ dB). We also investigate the performance gain (in terms of reconstruction percentage) at each iteration.

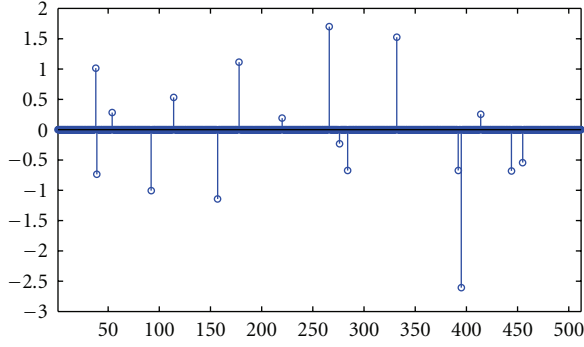


FIGURE 7: Reconstructed spike signal using TBCS with 75% similarity.

Figures 6 and 7 show the reconstructed spike signal using MBCS and TBCS, respectively, by introducing the spike signal with a similarity of 75%. The reconstruction percentage using TBCS is 92.7% while it is 57.5% using MBCS. The comparison of the two figures shows that TBCS can recover most of the original signal while MBCS fails to reconstruct the signal with so few measurements ($M = 62$) in spite of using a high-similarity signal as *a priori* information.

Figures 8, 9, and 10 show, when transferred signals have a similarity of 25%, 50%, and 75%, respectively, how much signal reconstruction percentage can be achieved at each iteration using the BCS, MBCS, and TBCS algorithms. The simulations are run 100 times, over which the results are averaged. It is clear that our proposed TBCS is much better than the BCS at each iteration. Particularly, when the similarity is 25%, MBCS is worse than BCS while our TBCS achieves higher performance at each iteration than BCS. For instance, at iteration 25 in Figure 8, TBCS can achieve a reconstruction percentage of 61.7%, while BCS can reach 42.2% and MBCS only recovers 35.6%. It shows that, at a low similarity, our TBCS can still achieve good performance at every iteration, compared with MBCS and BCS. Moreover, with a high similarity, the performance gap between TBCS and MBCS is enlarged at each step. For example, at iteration 21 with a similarity of 25% in Figure 8, TBCS can achieve a reconstruction percentage of 59.7%, while MBCS can reach 28.2%. Hence, the performance gap is 31.5%. When the similarity is 75% in Figure 10, the performance gap is increased to 50.9% because TBCS can reach 80.5%, while MBCS achieves 29.6% at the 21st iteration.

6.2. UWB Signal. The tested scenarios are the experimental UWB echo pulses from various UWB propagation channels in practical indoor residential, office, and clean, line-of-sight (LOS) and non-line-of-sight (NLOS) environments, which are drawn from experimental IEEE 802.15.4a UWB propagation models [26]. In a typical UWB communication/positioning system where receivers are distributed in the same environment, the received UWB echo signals are more or less similar. We test performance of original BCS, TBCS, and MBCS algorithms with different similarity levels.

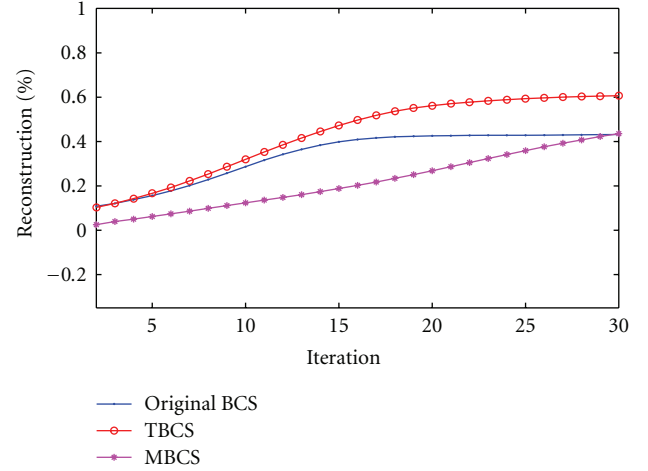


FIGURE 8: Performance gain in each iteration with 25% similarity.

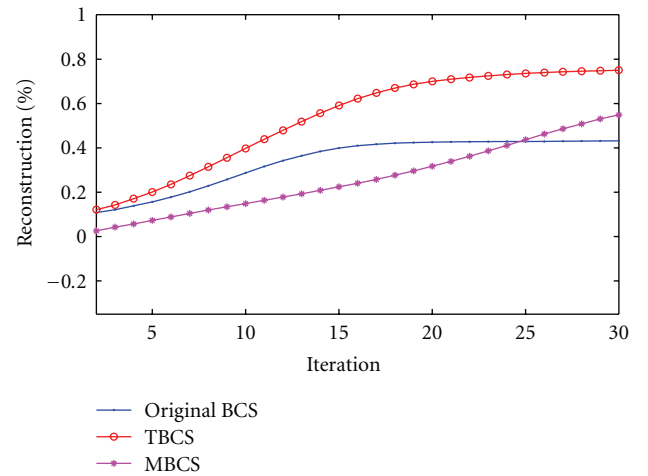


FIGURE 9: Performance gain in each iteration with 50% similarity.

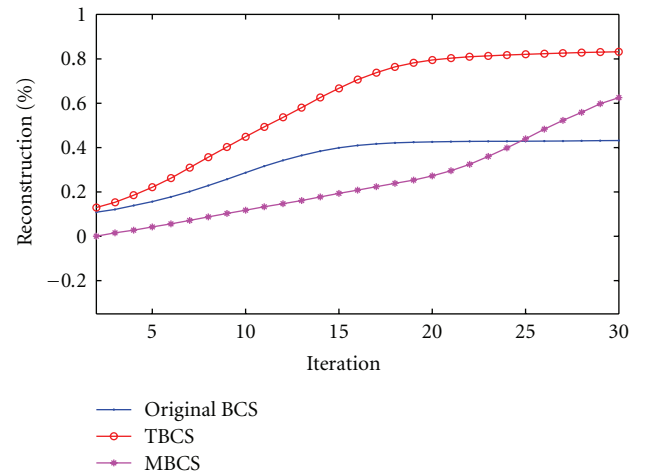


FIGURE 10: Performance gain in each iteration with 75% similarity.

Figure 11 shows the reconstructed UWB echo signals using the original BCS and our TBCS algorithms. The test UWB echo signals S0 (not shown in Figure 11), S1, S2, S3,

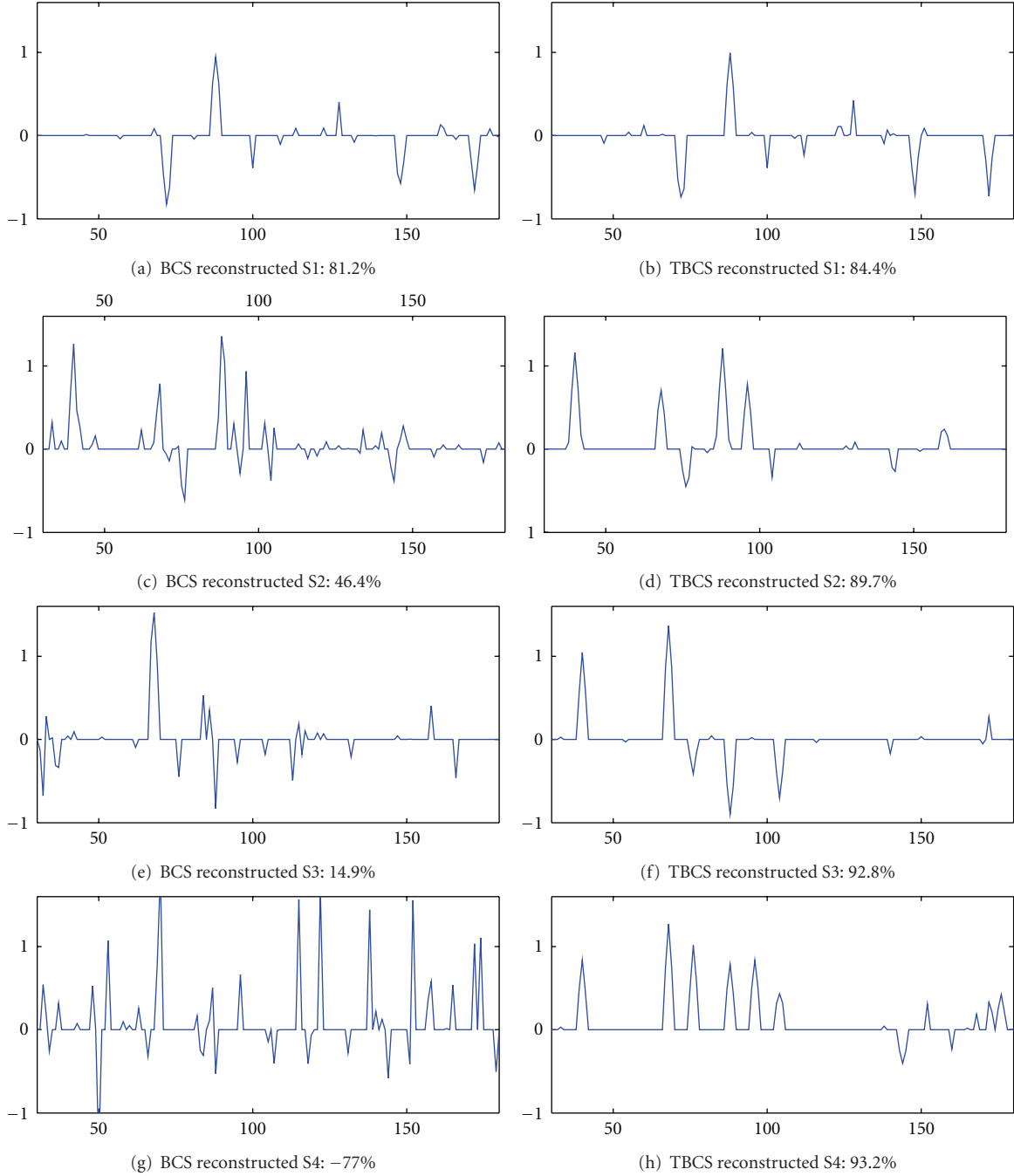


FIGURE 11: The performance of original BCS and TBCS. The UWB echo signals S1, S2, S3, and S4 with length $N = 512$ are reconstructed using the BCS and TBCS algorithms but only a section (length = 150) is shown. In the TBCS algorithm, the reconstructed signal S0 (not shown) is transferred to other signal reconstruction as *a priori* information. The number of measurements, SNR, similarity, and reconstruction percentage are (a) and (b) measurements $M = 60$; SNR = 9.2 dB; with respect to S0, the similarity in S1 is 11.5%; the reconstruction percentages of S1 using BCS and TBCS algorithms are 81.2% and 84.4%, respectively. (c) and (d) $M = 60$, SNR = 17.7 dB; with respect to S0, similarity in S2 is 31.3%; the reconstruction percentages are 46.4% and 89.7%. (e) and (f) $M = 50$, SNR = 12.4 dB; 61.0% similarity; the reconstruction percentages are 14.9% and 92.8%. (g) and (h) $M = 70$, SNR = 15.1 dB; 98.1% similarity; the reconstruction percentages are -77.0% and 93.2%.

and S4 are drawn from the IEEE802.15 UWB propagation model [26], in which the reconstructed S0 as *a priori* information is transferred to the other four signal scenarios. With respect to S0, the similarity levels in S1, S2, S3, and S4 are 11.5%, 31.3%, 61.0%, and 98.1%, respectively. For

each signal, both algorithms utilize the same number of measurements with the same SNR level for reconstruction. For clarity, only a portion of the UWB signal scenario is expanded to illustrate the waveform details of the reconstructed pulses. It is clearly observed from Figure 11

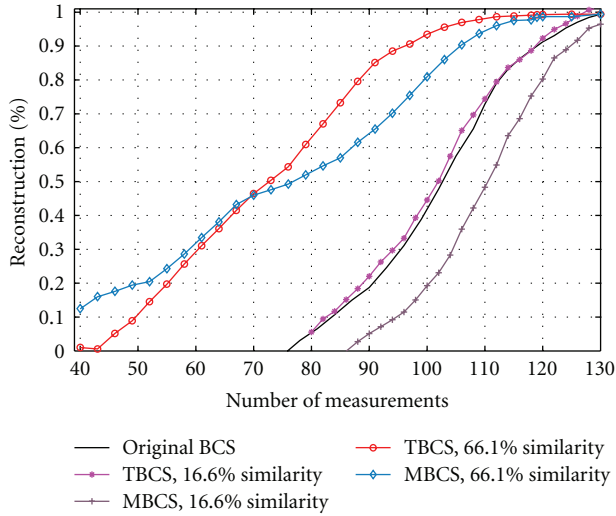


FIGURE 12: Performance comparison at different similarity levels without noise.

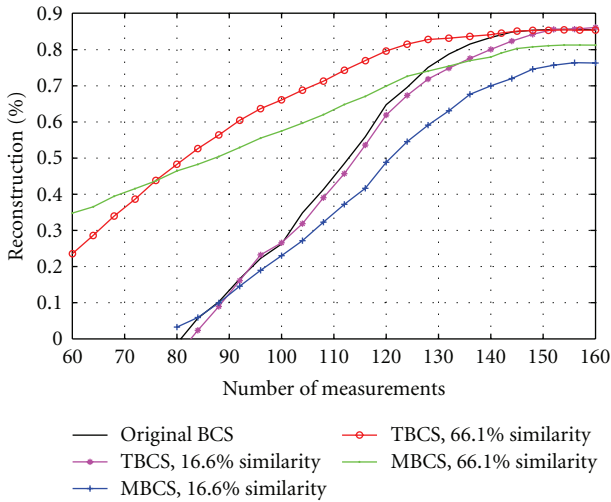


FIGURE 13: Performance comparison at different similarity levels in the presence of noise.

that our TBCS is much better than the original BCS for different similarity levels. The reconstruction percentages using TBCS are much higher than those using original BCS by introducing *a priori* information with the same number of measurements. Moreover, the performance gap is increasing with the growth of the similarity level. For instance, with a similarity of 11.5% for reconstructing the signal S1 in Figures 11(a) and 11(b), the difference of reconstruction percentages using BCS and TBCS is only 3.2% (84.4–81.2%). When the similarity level is 98.1% for reconstructing the signal S4 in Figures 11(g) and 11(h), the difference is increased to 170.2% (93.2–(–77)). Therefore, with a higher similarity level, higher performance gain can be achieved.

The performance of the original BCS, MBCS, and TBCS at different similarity levels is then compared. We select three UWB echo signals S5, S6, and S7 with the same dimension $N = 512$. The additive noise variance is only 0.01, implying a very high SNR. The reconstructed signals

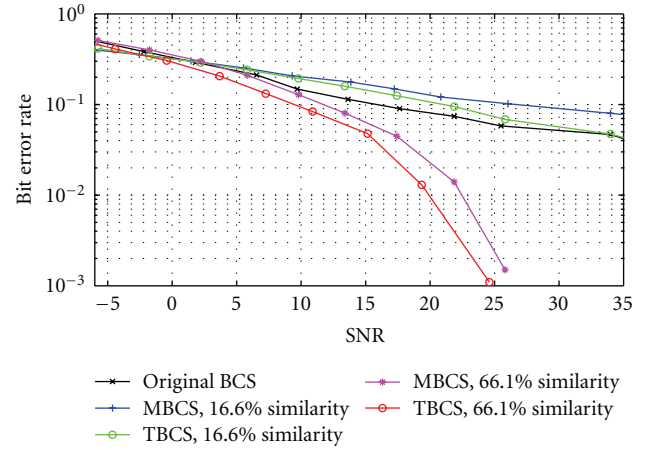


FIGURE 14: BER performance using different algorithms.

S6 and S7 as *a priori* information are transferred to the signal reconstruction for S5. With respect to S6 and S7, the similarities in S5 are 16.3% and 64.4%, respectively. The signal S5 is recovered with different numbers of measurements using the original BCS, TBCS, and MBCS algorithms. Figure 12 shows the reconstruction percentages versus the number of measurements for the signal S5. Obviously, at a low similarity level, the MBCS performance is substantially worse than the original BCS whereas our TBCS achieves a performance equaling that of the original BCS performance. For a high similarity level, both MBCS and TBCS are much better than the original BCS due to the benefits of high similarity transferred from the signal S7. This demonstrates that our TBCS achieves a good balance between local observations and *a priori* information, leading to a more robust performance than the MBCS.

In the presence of more noise interference, our TBCS still outperforms MBCS and BCS, as shown in Figure 13. We use the same signals S5, S6, and S7 but the noise variance is increased to 0.4. We observe that our TBCS exhibits good performance, as shown in Figure 12. Particularly in the presence of noise, when the number of measurements is large enough ($M > 150$). At a low similarity level, the MBCS can achieve a maximum reconstruction percentage of 74.5% while our TBCS algorithm is able to accomplish a maximum reconstruction percentage of 86.9%. At a high similarity level, MBCS can reach a maximum of 80.1% while our TBCS algorithm is still able to accomplish a maximum of 86.9%. Therefore, by introducing *a priori* information, the proposed TBCS algorithm can significantly reduce the number of measurements and improve the capability of combating noise.

Figure 14 shows the Bit Error Rate (BER) for an example UWB communication system using different algorithms. We utilize Binary Phase Shift Keying (BPSK) modulation to transfer the data since biphasic modulation is one of the easiest methods to implement. The performance of the TBCS, MBCS, and the original BCS algorithms is compared for the UWB communication system. The BER is tested using different noise levels with the same number of measurements ($M = 112$). With so few measurements, using the BCS algorithm leads to a high BER at different SNR. It is

also observed that, at a low similarity level, the TBCS performance is much better than the MBCS algorithm. At a high similarity level, the BER performance using the TBCS and MBCS algorithms are much better than that using the original BCS algorithm, while TBCS is the best. Therefore, by applying our TBCS algorithm in the UWB communication system, it can reduce the BER, provide more tolerance of the noise, and thus achieve the best performance when compared with the MBCS and BCS algorithms.

7. Conclusion

This paper has proposed an efficient approach to exploit and integrate the spatial and temporal *a priori* information existing in sparse signals, for example, UWB pulses. The turbo BCS algorithm has been designed to fully exploit *a priori* information from both space and time. Numerical simulation results have shown that the proposed TBCS outperforms the MBCS and traditional BCS, in terms of the robustness to noise and reduction of the required amount of samples.

Appendices

A. Proof of (9) and (10)

We first show the derivation of (9), which is given by

$$\begin{aligned} P(s_j^i | \lambda_j^i) &= \int P(s_j^i | \alpha_j^i) P(\alpha_j^i | \lambda_j^i) d\alpha_j^i \\ &= \int \left(\frac{\alpha_j^i}{2\pi} \right)^{-(1/2)} \exp\left(-\frac{(s_j^i)^2 \alpha_j^i}{2}\right) \lambda_j^i \exp(-\lambda_j^i \alpha_j^i) d\alpha_j^i \\ &= \frac{\lambda_j^i}{(2\pi)^{1/2}} \int (\alpha_j^i)^{-(1/2)} \exp\left(-\left(\lambda_j^i + \frac{(s_j^i)^2}{2}\right) \alpha_j^i\right) d\alpha_j^i. \end{aligned}$$

Let

$$\begin{aligned} t &= \left(\lambda_j^i + \frac{(s_j^i)^2}{2} \right) \alpha_j^i \\ &= \frac{\lambda_j^i}{(2\pi)^{1/2}} \int \left(\frac{t}{\lambda_j^i + ((s_j^i)^2/2)} \right)^{1/2} \exp(-t), \\ &\quad d\left(\frac{t}{\lambda_j^i + ((s_j^i)^2/2)} \right) \\ &= \frac{\lambda_j^i}{(2\pi)^{1/2}} \left(\lambda_j^i + \frac{(s_j^i)^2}{2} \right)^{-(3/2)} \int t^{1/2} \exp(-t) dt \\ &= (2\pi)^{-(1/2)} \Gamma\left(\frac{3}{2}\right) \lambda_j^i \left(\lambda_j^i + \frac{(s_j^i)^2}{2} \right)^{-(3/2)}, \end{aligned} \tag{A.1}$$

where $\Gamma(\cdot)$ is the gamma function, defined as $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$. We have $\Gamma(3/2) = \int_0^\infty t^{1/2} e^{-t} dt$. Because both distributions belong to the exponential distribution family, the marginal distribution is still in the same family. It is also observed that the marginal distribution $P(s_j^i | \lambda_j^i)$ is sharply peaked at zero, which encourages the sparsity. Therefore, the chosen exponential a prior distribution in the hierarchical Bayesian framework can be recognized and encourage the sparsity of the reconstructed signal.

Based on the assumption $\alpha_j^b = \alpha_j^i$, we have the same derivation:

$$\begin{aligned} P(s_j^b | \lambda_j^i) &= \int P(s_j^b | \alpha_j^i) P(\alpha_j^i | \lambda_j^i) d\alpha_j^i \\ &= \int \left(\frac{\alpha_j^i}{2\pi} \right)^{-(1/2)} \exp\left(-\frac{(s_j^b)^2 \alpha_j^i}{2}\right) \lambda_j^i \exp(-\lambda_j^i \alpha_j^i) d\alpha_j^i \\ &= (2\pi)^{-(1/2)} \Gamma\left(\frac{3}{2}\right) \lambda_j^i \left(\lambda_j^i + \frac{(s_j^b)^2}{2} \right)^{-(3/2)}. \end{aligned} \tag{A.2}$$

In order to obtain (10), we utilize the above equations. Then the derivation of the posterior is given by

$$\begin{aligned} P(\alpha_j^i | s_j^b, \lambda_j^i) &= \frac{P(s_j^b | \alpha_j^i) P(\alpha_j^i | \lambda_j^i)}{P(s_j^b, \lambda_j^i)} \\ &= \frac{P(s_j^b | \alpha_j^i) P(\alpha_j^i | \lambda_j^i)}{\int P(s_j^b | \alpha_j^i) P(\alpha_j^i | \lambda_j^i) d\alpha_j^i} \\ &= \frac{(\alpha_j^i)^{-1} (2\pi)^{-(1/2)} \lambda_j^i \exp\left(-\left(\frac{(s_j^b)^2 \alpha_j^i}{2} - \lambda_j^i \alpha_j^i\right)\right)}{(2\pi)^{-(1/2)} \Gamma(3/2) \left(\lambda_j^i + \frac{(s_j^b)^2}{2}\right)^{-(3/2)} \lambda_j^i} \\ &= \frac{\left(\lambda_j^i + \frac{(s_j^b)^2}{2}\right)^{3/2} \exp\left(-\left(\lambda_j^i + \frac{(s_j^b)^2}{2}\right) \alpha_j^i\right)}{\Gamma(3/2)} \\ &= \frac{(\tilde{\lambda}_j^i)^{3/2} \exp(-\tilde{\lambda}_j^i \alpha_j^i)}{\Gamma(3/2)}. \end{aligned} \tag{A.3}$$

So the parameter λ_j^i is updated to $\tilde{\lambda}_j^i$, which is given by

$$\tilde{\lambda}_j^i = \lambda_j^i + \frac{(s_j^b)^2}{2}. \tag{A.4}$$

For transferred multiplied reconstructed signal elements $s_j^{b_1}, s_j^{b_2}, \dots, s_j^{b_n}$, the posterior function also belongs to the exponential distribution family. As shown in (12), the parameter λ_j^i is updated to

$$\begin{aligned}
 & P(\alpha_j^i | s_j^{b_1}, s_j^{b_2}, \dots, s_j^{b_n}, \lambda_j^i) \\
 &= \frac{P(s_j^{b_1} | \alpha_j^i) P(s_j^{b_2} | \alpha_j^i) \cdots P(s_j^{b_n} | \alpha_j^i) P(\alpha_j^i | \lambda_j^i)}{P(s_j^{b_1}, s_j^{b_2}, \dots, s_j^{b_n}, \lambda_j^i)} \\
 &= \frac{P(s_j^{b_1} | \alpha_j^i) P(s_j^{b_2} | \alpha_j^i) \cdots P(s_j^{b_n} | \alpha_j^i) P(\alpha_j^i | \lambda_j^i)}{\int P(s_j^{b_1}, \alpha_j^i | \lambda_j^i) P(s_j^{b_2}, \alpha_j^i | \lambda_j^i) \cdots P(s_j^{b_n}, \alpha_j^i | \lambda_j^i) d\alpha_j^i} \\
 &= \frac{P(s_j^{b_1} | \alpha_j^i) P(s_j^{b_2} | \alpha_j^i) \cdots P(s_j^{b_n} | \alpha_j^i) P(\alpha_j^i | \lambda_j^i)}{\int P(s_j^{b_1} | \alpha_j^i, \lambda_j^i) P(s_j^{b_2} | \alpha_j^i, \lambda_j^i) \cdots P(s_j^{b_n} | \alpha_j^i, \lambda_j^i) P(\alpha_j^i | \lambda_j^i) d\alpha_j^i} \\
 &= \frac{\left(\lambda_j^i + \sum_{i=1}^n (s_j^{b_i})^2 / 2 \right)^{(2n+1)/2} \exp\left(-\left(\lambda_j^i + \sum_{i=1}^n (s_j^{b_i})^2 / 2 \right) \alpha_j^i\right)}{\Gamma((2n+1)/2)} \\
 &= \frac{(\tilde{\lambda}_j^i)^{(2n+1)/2} \exp(-\tilde{\lambda}_j^i \alpha_j^i)}{\Gamma((2n+1)/2)}. \tag{A.5}
 \end{aligned}$$

The distributions $P(s_j^{b_1} | \alpha_j^i), P(s_j^{b_2} | \alpha_j^i), \dots$, and $P(s_j^{b_n} | \alpha_j^i)$ are conditionally independent from each other. In this case, the parameter is updated to

$$\tilde{\lambda}_j^i = \lambda_j^i + \frac{\sum_{i=1}^n (s_j^{b_i})^2}{2}, \tag{A.6}$$

where n represents the total number of *a priori* information $s_j^{b_1}, s_j^{b_2}, \dots, s_j^{b_n}$.

Therefore, the above derivations show how the single or multiple signal elements $s_j^{b_i}, j = 1, 2, \dots, N, n = 1, 2, \dots$, from the other BCS procedures update the hyperparameters in the i th BCS signal reconstruction procedure.

B. Derivation of (15)

One strategy to maximize (17) is to exploit an EM method, treating the s^i as hidden data, and maximize the following expectation:

$$E_{s^i | y^i, \alpha^i} [\log P(s^i | \alpha^i) P(\alpha^i | \lambda^i)]. \tag{B.1}$$

The operator $E_{s^i | y^i, \alpha^i}$ denotes an expectation of the posterior $P(s^i | y^i, \alpha^i, \lambda^i, \beta)$ with respect to the distribution

over the s^i given the data and hidden variables. Through differentiation with respect to α^i we get

$$\begin{aligned}
 & \frac{\partial}{\partial \alpha_j^i} E_{s^i | y^i, \alpha^i} [\log P(s^i | \alpha^i, \beta) P(\alpha^i | \lambda^i)] \\
 &= E_{s^i | y^i, \alpha^i} \left[\frac{\partial}{\partial \alpha_j^i} (\log P(s^i | \alpha^i, \beta) + \log P(\alpha^i | \lambda^i)) \right] \\
 &= -\frac{1}{2} E_{s^i | y^i, \alpha^i} \left[-2(s_j^i)^2 - 4\lambda_j^i + \frac{6}{\alpha_j^i} \right] \\
 &= E_{s^i | y^i, \alpha^i} \left((s_j^i)^2 \right) + 2\lambda_j^i - \frac{3}{\alpha_j^i}. \tag{B.2}
 \end{aligned}$$

According to (3), we have

$$E_{s^i | y^i, \alpha^i} \left((s_j^i)^2 \right) = \Sigma_{jj}^i + (\mu_j^i)^2. \tag{B.3}$$

We set (B.2) to 0, which yields an update for α_j^i :

$$\alpha_j^i = \frac{3}{(s_j^i)^2 + \Sigma_{jj}^i + 2\lambda_j^i}. \tag{B.4}$$

C. Derivation of (19)

For the $L_1(\alpha)$, as shown in [21], we have that

$$\begin{aligned}
 L_1(\alpha) &= -\frac{1}{2} (N \log 2\pi + \log |E| + y^T E^{-1} y) \\
 &= L_1(\alpha_{-j}) + l_1(\alpha_j), \tag{C.1}
 \end{aligned}$$

where,

$$\begin{aligned}
 E &= \sigma^2 I + \Phi A^{-1} \Phi^T \\
 &= \sigma^2 I + \sum_{k \neq j} \alpha_k^{-1} \phi_k \phi_k^{-1} + \alpha_j^{-1} \phi_j \phi_j \\
 &= E_{-j} + \alpha_j^{-1} \phi_j \phi_j, \tag{C.2}
 \end{aligned}$$

$$\begin{aligned}
 L_1(\alpha_{-j}) &= -\frac{1}{2} (N \log 2\pi + \log |E_{-j}| + y^T E_{-j}^{-1} y), \\
 l_1(\alpha_j) &= \frac{1}{2} \left(\log \alpha_j - \log (\alpha_j + g_j) + \frac{h_j^2}{\alpha_j + g_j} \right). \tag{C.3}
 \end{aligned}$$

The quantities g_j, h_j , and E_{-j} are given by

$$g_j = \phi_j^T E_{-j}^{-1} \phi_j, \tag{C.4}$$

$$h_j = \phi_j^T E_{-j}^{-1} y, \tag{C.5}$$

$$E_{-j} = \beta^2 I + \sum_{k \neq j} \alpha_k^{-1} \phi_k \phi_k^{-1}, \tag{C.6}$$

where ϕ_j is the j th column vector of the matrix Φ .

In order to find the critical point, the differentiation of $l_1(\alpha_j)$ is given by

$$\frac{\partial l_1(\alpha_j)}{\partial \alpha_j} = \frac{\alpha_j^{-1} g_j^2 - (h_j^2) + g_j}{2(\alpha_j + g_j)^2} = 0. \quad (\text{C.7})$$

It is easy to maximize $l_1(\alpha_j)$ with respect to α_j by taking the first and second derivatives. Then the maximum point α_j is given by

$$\alpha_j^* = \begin{cases} \frac{h_j^2}{g_j^2 - h_j}, & \text{if } g_j^2 > h_j^2, \\ \infty, & \text{otherwise.} \end{cases} \quad (\text{C.8})$$

The second derivative is

$$\frac{\partial^2 l_1(\alpha_j)}{\partial \alpha_j^2} = \frac{-2\alpha_j^2(\alpha_j + g_j)(\alpha_j^{-1} g_j^2 - h_j^2 + g_j) - (\alpha_j + g_j)^2 g_j^2}{2\alpha_j^2(\alpha_j + g_j)^4}. \quad (\text{C.9})$$

Taking the critical point α_j^* into the second derivative expression, we have known that

$$\frac{\partial^2 l_1(\alpha_j = \alpha_j^*)}{\partial \alpha_j^2} = \frac{-g_j^2}{2\alpha_j^2(\alpha_j + g_j)^2}. \quad (\text{C.10})$$

Obviously, it is always negative, and therefore function $l_1(\alpha_j)$ achieves the maximum at α_j^* , which is unique.

D. Derivations about (24) and (25)

The first derivative of the $l_2(\alpha_j)$ is $l_2'(\alpha_j) = -\lambda_j$. All together the first differentiation of the posterior $l(\alpha_j)$ is given by

$$\begin{aligned} l'(\alpha_j) &= l_1'(\alpha_j) + l_2'(\alpha_j) \\ &= \left(\frac{g_j}{2\alpha_j(\alpha_j + g_j)} - \frac{h_j^2}{2(\alpha_j + g_j)^2} \right) - \lambda_j \\ &= \frac{1}{2} \left[\frac{1}{\alpha_j} - \frac{1}{\alpha_j + g_j} - \frac{h_j^2}{(\alpha_j + g_j)^2} - 2\lambda_j \right]. \end{aligned} \quad (\text{D.1})$$

By setting the (D.1) to zero, we can find the optimum α_j^* for (25).

The g_j and h_j^2 are not negative based on (C.4) and (C.5). We have $\alpha_j \geq 0$ and $\lambda_j > 0$ according to the exponential distribution as shown in (7), and $l'(\alpha_j) \rightarrow -2\lambda_j < 0$ as $\alpha_j \rightarrow +\infty$. Then, it has $l'(\alpha_j) > 0$ when $\alpha_j \rightarrow 0$. Therefore, for the function $l'(\alpha_j) = 0$, it has at least one positive root for $\alpha_j > 0$.

We rearrange (D.1) to

$$\begin{aligned} l'(\alpha_j) &= \frac{1}{2} \left[\frac{1}{\alpha_j} - \frac{1}{\alpha_j + g_j} - \frac{h_j^2}{(\alpha_j + g_j)^2} - 2\lambda_j \right] \\ &= \frac{f(\alpha_j, g_j, h_j, \lambda_j)}{\alpha_j(\alpha_j + g_j)^2}. \end{aligned} \quad (\text{D.2})$$

Setting (D.2) to zero is to let the numerator be zero, that is, $f(\alpha_j, g_j, h_j, \lambda_j) = 0$. To find the solution, we normalize the equation to reduce one parameter for convenience. Then we need to solve

$$\frac{f(\alpha_j, g_j, h_j, \lambda_j)}{-2\lambda_j} = \alpha_j^3 + B_0\alpha_j^2 + B_1\alpha_j + B_2 = 0. \quad (\text{D.3})$$

The corresponding coefficients are given by [31]

$$\begin{aligned} B_0 &= 2g_j, \\ B_1 &= \frac{g_j - 2\lambda_j g_j^2 - h_j^2}{-2\lambda_j}, \\ B_2 &= \frac{g_j^2}{-2\lambda_j}. \end{aligned} \quad (\text{D.4})$$

To solve the cubic function, we define intermediate components as

$$U = 2B_0^3 - 9B_0B_1 + 27B_2, \quad (\text{D.5})$$

$$V = (2B_0^3 - 9B_0B_1 + 27B_2)^2 - 4(B_0^2 - 3B_1)^3.$$

Then the solutions of the cubic function are given by

$$\begin{aligned} x_1 &= -\frac{1}{3} \left(B_0 + \sqrt[3]{\frac{U + \sqrt{V}}{2}} + \sqrt[3]{\frac{U - \sqrt{V}}{2}} \right), \\ x_2 &= -\frac{1}{3} \left(B_0 + \omega_1 \sqrt[3]{\frac{U + \sqrt{V}}{2}} + \omega_2 \sqrt[3]{\frac{U - \sqrt{V}}{2}} \right), \\ x_3 &= -\frac{1}{3} \left(B_0 + \omega_2 \sqrt[3]{\frac{U + \sqrt{V}}{2}} + \omega_1 \sqrt[3]{\frac{U - \sqrt{V}}{2}} \right), \end{aligned} \quad (\text{D.6})$$

where,

$$\begin{aligned} \omega_1 &= -\frac{1}{2} + \frac{\sqrt{3}}{2}i \\ \omega_2 &= -\frac{1}{2} - \frac{\sqrt{3}}{2}i. \end{aligned} \quad (\text{D.7})$$

Therefore, all those three roots x_1 , x_2 , and x_3 are critical points of the optimization function shown in (19). We choose the positive root which maximizes the optimization function in (19) as the optimum solution α_j^* for (25).

References

- [1] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [2] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.
- [3] H. Zayyani, M. Babaie-Zadeh, and C. Jutten, "Bayesian pursuit algorithm for sparse representation," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '09)*, pp. 1549–1552, Taipei, Taiwan, April 2009.

- [4] H. Zayyani, M. Babaie-Zadeh, and C. Jutten, "An iterative Bayesian algorithm for sparse component analysis in presence of noise," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4378–4390, 2009.
- [5] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2346–2356, 2008.
- [6] D. Yang, H. Li, G. D. Peterson, and A. Fathy, "Compressed sensing based UWB receiver: hardware compressing and FPGA reconstruction," in *Proceedings of the 43rd Annual Conference on Information Sciences and Systems (CISS '09)*, pp. 198–201, Baltimore, Md, USA, March 2009.
- [7] Z. Wang, G. R. Arce, B. M. Sadler, J. L. Paredes, S. Hoyos, and Z. Yu, "Compressed UWB signal detection with narrowband interference mitigation," in *Proceedings of the IEEE International Conference on Ultra-Wideband (ICUWB '08)*, pp. 157–160, Singapore, September 2008.
- [8] D. Yang, H. Li, and G. D. Peterson, "Feedback orthogonal pruning pursuit for pulse acquisition in UWB communications," in *Proceedings of the 20th IEEE Personal, Indoor and Mobile Radio Communications Symposium (PIMRC '09)*, Tokyo, Japan, September 2009.
- [9] P. Zhang, Z. Hu, R. C. Qiu, and B. M. Sadler, "Compressive sensing based ultra-wideband communication system," in *Proceedings of the IEEE International Conference on Communications (ICC '09)*, Dresden, Germany, June 2009.
- [10] D. Baron, M. B. Wakin, M. F. Duarte, S. Sarvotham, and R. G. Baraniuk, "Distributed compressed sensing," available online, <http://arxiv.org/abs/0901.3403>.
- [11] W. Wang, M. Garofalakis, and K. Ramchandran, "Distributed sparse random projections for refinable approximation," in *Proceedings of the 6th International Symposium on Information Processing in Sensor Networks (IPSN '07)*, pp. 331–339, Cambridge, Mass, USA, April 2007.
- [12] D. Leviatan and V. N. Temlyakov, "Simultaneous approximation by greedy algorithms," IMI Report 2003, University of South Carolina at Columbia, 2003.
- [13] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation. Part I: greedy pursuit," *Signal Processing*, vol. 86, no. 3, pp. 572–588, 2006.
- [14] N. Vaswani and W. Lu, "Modified-CS: modifying compressive sensing for problems with partially known support," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '09)*, pp. 488–492, Seoul, Korea, July 2009.
- [15] E. Berg and M. P. Friedlander, "Joint-sparse recovery from multiple measurements," available online, <http://www.arxiv.org/abs/0904.2051v1>.
- [16] F. Zhang and H. D. Pfister, "On the iterative decoding of highrate LDPC codes with applications in compressed sensing," available online, <http://arxiv.org/abs/0903.2232>.
- [17] X. Tan and J. Li, "Computationally efficient sparse Bayesian learning via belief propagation," *IEEE Transactions on Signal Processing*, vol. 58, no. 4, pp. 2010–2021, 2010.
- [18] D. Baron, S. Sarvotham, and R. G. Baraniuk, "Bayesian compressive sensing via belief propagation," *IEEE Transactions on Signal Processing*, vol. 58, no. 1, pp. 269–280, 2010.
- [19] P. Schniter, L. C. Potter, and J. Ziniel, "Fast Bayesian matching pursuit," in *Proceedings of the Information Theory and Applications Workshop (ITA '08)*, pp. 326–332, La Jolla, Calif, USA, January-February 2008.
- [20] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Bayesian compressive sensing using laplace priors," *IEEE Transactions on Image Processing*, vol. 19, no. 1, pp. 53–63, 2010.
- [21] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *Journal of Machine Learning Research*, vol. 1, no. 3, pp. 211–244, 2001.
- [22] D. Yang, H. Li, G. D. Peterson, and A. E. Fathy, "UWB signal acquisition in positioning systems: Bayesian compressed sensing with redundancy," in *Proceedings of the 43rd Annual Conference on Information Sciences and Systems (CISS '09)*, pp. 192–197, Baltimore, Md, USA, March 2009.
- [23] S. Ji, D. Dunson, and L. Carin, "Multitask compressive sensing," *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 92–106, 2009.
- [24] Y. Qi, D. Liu, D. Dunson, and L. Carin, "Bayesian multi-task compressive sensing with dirichlet process priors," in *International Conference on Machine Learning*, Helsinki, Finland, July 2008.
- [25] M. E. Tipping and A. C. Faul, "Fast marginal likelihood maximisation for sparse Bayesian models," in *Proceedings of the International Conference on Artificial Intelligence and Statistics*, Key West, Fla, USA, June 2003.
- [26] A. F. Molisch et al., "IEEE 802.15.4a channel model—final report," IEEE 2004, <http://www.ieee802.org/15/pub/04/15-04-0662-02-004a-channel-model-final-report-r1.pdf>.
- [27] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, CRC Press, New York, NY, USA, 2nd edition, 2003.
- [28] D. Yang, G. D. Peterson, H. Li, and J. Sun, "An FPGA implementation for solving least square problem," in *Proceedings of the IEEE Symposium on Field Programmable Custom Computing Machines (FCCM '09)*, pp. 303–306, Napa, Calif, USA, April 2009.
- [29] D. Yang, G. Peterson, and H. Li, "High performance reconfigurable computing for Cholesky decomposition," in *Proceedings of the Symposium on Application Accelerators in High Performance Computing (UIUC '09)*, Urbana, Ill, USA, July 2009.
- [30] D. Yang, A. Fathy, H. Li, G. D. Peterson, and M. Mahfouze, "Millimeter accuracy UWB positioning system using sequential sub-sampler and time difference estimation algorithm," in *Proceedings of the IEEE Radio and Wireless Symposium (RWS '10)*, New Orleans, Fla, USA, January 2010.
- [31] W. S. Anglin and J. Lambek, *The Heritage of Thales*, Springer, New York, NY, USA, 1993.