Simultaneous Representation and Separation for Multiple Interference Allied with Approximation Message Passing

Gui-sheng WANG‡1, Shu-fu DONG1, Hong-yi HANXIAO 1 and Bo YANG 1

1 Air Force Engineering University, Xi’an 710077, China
‡ Correspondence: wgsfuyun@163.com

Received: date; Accepted: date; Published: date

Abstract: — Broadband reliable communication is a competitive 5G technology for cognitive communication scenarios, but meanwhile introduces multiform interference to existing broadband transform domain communication system (TDCS) transmission. In order to facilitate the improvement of the anti-jamming performance for the coexistence of diverse interference and TDCS signals in wireless heterogeneous networks, it is important to separated and eliminate various interference to TDCS systems. In this paper, a novel sparse learning method-based cognitive transformation framework of interference separation is formulated for accurate interference recovery, which can be efficiently solved by iteratively learning the prior sparse probability distribution of the interference support. To further improve the separation accuracy and iterative convergence, the principal component analysis and Bayesian perspective in orthogonal base learning are exploited to singly recover the multiple interference and TDCS signals. Moreover, utilizing different sparsity states of spectrum analysis, the proposed novel interference separation algorithm is extended to simultaneous separation based on state evolving of approximation message passing, which iteratively learns the belief propagation posteriors and keeps shrunk by iterative shrinkage threshold. Simulation results demonstrate that the proposed methods are effective in separating and recovering the sparse diversities of interference to TDCS systems, and significantly outperform the state-of-the-art methods.

Keywords: transform domain communication system; interference separation; sparse representation; orthogonal base learning; Bayesian learning; approximation message passing

1. Introduction

With the 5G technology evolving gradually, communication frequency and bandwidth are constantly improving, which rapidly promote the reliable communication towards the high rate, low delay and large capacity [1-3]. According to the Nyquist sampling theorem, much high-speed information with instant requirements places conventional communication systems in a dilemma, which results in overburdened implementations and high costs for the hardware. Recently, compressed sensing (CS) [4,5] realizes the information compression processing with its sampling ability far lower than Nyquist rate, and it can accurately recover the original signal through many proven reconstruction algorithms, which is especially suitable for signal acquisition and processing of broadband communication.

As a typical broadband application of cognitive radio, transform domain communication system (TDCS) [6,7] develops the broadband multiple access and high spectrum efficiency with low interception and low detection performance. The TDCS has been demonstrated to facilitate the mitigation of interference from the perspectives both the transmitter and receiver rather than the receiver only, and provide the capability of avoiding spectral bands occupied by interference sources or jammers. Simultaneously, its framework enables the intelligent utilization of the scarce frequency spectrum available under complex and confrontational conditions through frequency pool nulling and
spectral broadening via transformation analysis to synthesize adaptive and interference-free waveforms for high spectrum efficiency [8]. The TDCS, given its advantages of wide-band and high-speed transmission, can support vehicular and unmanned aerial vehicle (UAV) networks for intelligent sensing, control, driving, and other vehicle-to-vehicle (V2V) applications, but it is susceptible to electromagnetic interference that impairs communication efficiency. Therefore, ensuring a low probability of interception (LPI) and susceptibility to interference is essential to maintain the security and reliability of communication in tactical edging vehicular networks. As the electromagnetic environment becomes increasingly complex with more and more severe confrontations in the battlefield, all related units are cumulatively faced with the threat of communication interruption, blocking or jamming. Therefore, it has a practical significance to develop some fit methods of anti-jamming performance with high spectrum efficiency for broadband TDCS.

Since most of all broadband communication signals can be compressed, that is, they can be well represented with sparse signals after selecting an appropriate basis for transformation [9], which provides a theoretical basis for sparse reconstruction and separation of signals. Recently, plenty of work has been done by exploiting sparse characteristic of the interference. According to diversities of sparsity in some dimensions or domains, they convert the targeted interference or communication signal into the corresponding transform domain, which enhances the performance of detection and separation for interference [10,11]. From a general view, suppression methods of interference mainly include three categories based on CS, selective measurement, sparse separation and reconstruction alignment. Among them, the selective measurement of interference [12] utilizes diverse representation characteristics of the measurement matrix so as to achieve the selectivity of signal receiving and the separability of compression sampling. A compressed filtering method was proposed based on the orthogonality of the compressed components between communication signals and encountered interference, which realized the effective suppression of jamming [13]. The principle of selective measurement is simple and efficient, and its pertinence appears quite distinct. However, the optimal interference subspace for measurement needs to be designed with specificity according to different interference characteristics, thus its scope of applications are limited relatively. On top of that, the applied methods more widely are sparse separation and reconstruction alignment. Moreover, the reconstruction alignment of interference [14,15] can reduce the negative disturbance on the communication signals primarily in time domain by employing the innovative alignment of interference after recovery with high precision. The improved interference estimation method in [16] realized the interference cancellation and suppression in time domain via estimating the sparsity awareness degree, and simulating results were proven that it is suitable for the jamming scenario of high power intensity, nevertheless in the actual, the elimination of interference is just passable under the precision impacts and the high-cost hardware improvement of antenna structure. Accordingly, the sparse separation method [17] generally utilizes the separability of sparse representation for interference and communication signals under the complete or over-complete dictionaries, which achieves the sparse separation between the electromagnetic interference and signals. A recovery and suppression method for noise interference was proposed based on the sparse separability of communication signals and interference, and it is found that the improved method not only equips with
simple calculation but also runs with low complexity. Above all, a sparse dictionary can be conducted for separation as long as the interference meets the requirements of sparsity, so it is more suitable for the joint processing of diversified interference.

However, communication systems are vulnerable to suffer manifold forms of interference readily, involving common noise interference and impulse interference of electronic devices; besides, there exist numerous malicious types of hostile interference encountered in the entire communication process, such as multi-tone interference, partial-band interference and comb spectrum interference, etc., which severely restrict the communication reliability and availability [18]. Although the developing sparse methods based on CS are competent for estimating a variety of sparse interference whereas they only for certain categories, such as the impulse interference or partial-band interference with continuous block sparse characteristics in time domain or frequency domain, ignoring the other discrete and variable sparsity types, including chirp interference and comb spectrum interference [19]. More importantly, due to the fact that different forms of interference present scattered sparse characteristics in the time, frequency or wavelet domain, it is fairly tough to achieve the optimal sparse separation if only utilizing traditional dictionary learning [20,21] and sparse separation methods [22,23], which will seriously affect the performance of interference separation and suppression. Especially for the coexistence of diverse interference, the assumption seemed incredible that one valid transformation separating or dictionary learning method would effectively dealing with the separation and elimination of various interference.

Different from the aforementioned existing schemes, the emerging and powerful morphological analysis theory and techniques, drawing tremendous research attention recently, can be a great inspiration to achieve a both efficient and reliable method of interference separation. In the research on morphological analysis, approximate message passing (AMP) has been exploited for solving split analysis [24,25]. Nevertheless, the conventional AMP method was not designed for interference separation. Moreover, the state-of-the-art research on sparse representation-based interference separation and recovery using iterative AMP is insufficient in literature. To fill this gap, a sparse learning method inspired cognitive transformation framework is formulated, and a novel algorithm called interference separation altering with relaxed AMP (ISAR-AMP) is proposed to iteratively learn the sparse distribution. The proposed method is capable of learning, recovering and separation multiform interference more efficiently and more accurately than state-of-the-art counterparts, which improve distinctly the effectiveness and robustness of dealing with diversified interference in broadband spectrum.

The primary contributions of this paper listed as follows.

1. The interference separation problem are analyzed based on the receiver of TDCS situated in different spectrum states, and it can be divided into the known sparsity and unknown sparsity states where there are the issues to resolve about single separation and simultaneous separation of sparse interference, respectively, which results in the diversity of separating interference.

2. An efficient learning allied with base orthogonality (ELABO) is proposed for the sensed signal with unknown sparsity, the recovery and separation for sparse interference can be achieved by orthogonal base learning for representation and multiple Bayesian learning for recovery, where an
An improved sparse method is represented for diversities based on principal component analysis and iterative weighted least squares.

3. A novel interference separation algorithm called ISAR-AMP based on state evolving is proposed for the sensed signal with known sparsity, which iteratively learns the belief propagation posteriors in the min-sum manner. To further reduce the number of messages caused by the extra burden of loopy graph connections per iteration, the remaining iterations are shrunk by the iterative shrinkage threshold, which achieves better separation accuracy and convergence rate.

The rest of this paper is organized as follows. Section II describes the statistical models of TDCS signals, the primary interference forms as well as the responding transformations. Section III presents the corresponding separation problems of single recovery and simultaneous elimination for diverse interference, respectively, according to the probable sparsity conditions on sensing spectrum. After that, an efficient learning method is proposed for the sensed signal with unknown sparsity in Section IV, and an improved simultaneous separation algorithm for the sensed signal with known sparsity is described in Section V. The simulation results are reported in Section VI with discussions in detail, followed by the conclusions in Section VII.

2. Related work

2.1 TDCS Signal

Different from direct sequence spread spectrum that mitigates interference at the receiver, transform domain communication system (TDCS) smartly synthesizes an adaptive wave form to avoid interference at the transmitter. The transmission sequence in TDCS has good auto-correlation and cross-correlation performance to achieve low LPI and orthogonality. Fig. 1 shows the processing steps of a general TDCS transmitter.

1) Spectrum Identification

In the first step, the available spectrum is estimated to ensure an interference-free transmission utilizing the period-gram and auto-regressive methods, and the interference is detected and eliminated by adaptive filtering or maximum likelihood according to the sensing spectrum. Then, a spectrum mask \( M_T(w) \) is determined by the magnitude shaping and transformation analysis, \( M_T(w) = [m_1, m_2, \ldots, m_N] \), where \( N \) spectrum bands are separated from the entire spectrum band, \( T \) is the optimal transform domain for interference elimination, and \( m_i \in \{0, 1\} \).
2) Random Phase

The complex poly-phase vector is produced by a unique pseudo-random code selected randomly from an \( m \)-sequence with respect to the spectrum mask \( M_r(w) \) element-by-element, \( \mathbf{O}(\varphi) = [e^{i\varphi_1}, e^{i\varphi_2}, \ldots, e^{i\varphi_{2N/}}] \) and \( \varphi_i \in \{2^{-i} \pi, 2^{-i+1} \pi, \ldots, 2\pi\} \). Meanwhile, the multiple assessment is applied to each user by assigning pairs from a different \( m \)-sequence. The mapping rules achieve the noise-like property of \( \mathbf{B}(w) \), which is expressed as

\[
\mathbf{B}(\varphi) = M_r(w) \mathbf{O}(\varphi) = \sum_{i=1}^{N} m_i e^{i\varphi_i} \quad (1)
\]

3) Base Function

The desired energy is distributed equally or diversely in the spectrum nulling to scale the magnitude \([h_1, h_2, \ldots, h_N]\) according to limitations of the peak-to-average power ratio (PAPR), \( h_i \geq 0, i=1, 2, \ldots, N \). And then the appropriate inverse transform is performed for the energy-injected waveforms \( \mathbf{B}_w \) to generate the buffered base function \( b_n(t) \) in the time domain:

\[
b_n(t) = T^{-1} \left[ \mathbf{B}_w(w) \right] = T^{-1} \left[ \sum_{i=1}^{N} h_i m_i e^{i\varphi_i} \right] = \sum_{i=1}^{N} B_1 e^{i\varphi_i} e^{\frac{2\pi i}{N} n_i} \quad (2)
\]

4) Modulation

Generally, cyclic shift keying (CSK) is utilized for modulation in TDCS to adapt noise-like properties. For binary CSK (BCSK), the symbols are generated by circularly shifting the waveform over \( T_s/2 \) (half period). Then, the information bit stream \([r_1, r_2, \ldots, r_N]\) is synthesized by mapping the buffered base function, \( r_i \in \{0, 1\} \), and the transmitted waveform with the normalized energy factor \( \xi \) can be expressed as

\[
s_n(t) = \xi \sum_{i=1}^{N} B_1 e^{i\varphi_i} e^{\frac{2\pi i}{N} n_i} e^{-j2\pi w_i t}, l = 0, 1, \ldots, N-1 \quad (3)
\]

5) Transmission

The signal transmitted through the channel will be corrupted by additive Gaussian noise and external jamming; the channels are mainly dominated by the LOS component, and despite the possible limited multipath fade due to ground reflections, the propagation delay is mainly caused by path loss. Meanwhile, given the potentially relative mobility, Doppler frequency shifts may affect the channel and, consequently, the spectrum allocation. Therefore, the Clarke flat fade channel model proposed in [26] was employed to estimate the influence of Doppler frequency shifts, and the received TDCS signal can be given by

\[
s(t) = \kappa \xi \sum_{i=1}^{N} B_1 e^{i(r_i + \theta_i)} e^{\frac{2\pi i}{N} n_i} e^{-j2\pi w_i t} \quad (4)
\]

where \( \kappa \) is the fading factor, and \( \varphi \) is the Doppler shift.

In general, most signals processed in communication systems can be compressed by transforming them into coefficients in some respect. However, many of the transform coefficients of natural signals are close to zero [27]. Thus, the signals occupy only a small fraction of the entire transformed signal space and can be represented sparsely by a few elementary components out of a given collection. Consequently, sparsity can also be observed in other classes of natural signals [28].

When the length of the original signal is \( N \), there exist \( W=2^N \) possible forms of the transmitted signal...
in the entire signal space $S$. Only a small fraction of the space is occupied, which can be represented sparsely as

$$s(t) = \Psi^T \theta, \Psi^T = [S_1, S_2, \ldots, S_n]^T$$  \hspace{1cm} (5)

where $\Psi^T \in \mathbb{R}^{N \times L}$ is a sparse dictionary for signal space $S \in \mathbb{R}^{N \times L}$, and most coefficients $\theta^T \in \mathbb{R}^{L \times 1}$ are apparently equivalent to zero.

### 2.2 Sparse Interference elimination

Practically, communication systems are subject to a wide range of interference in complex environments, and the types of interference signals can be divided mainly into three categories, which include impulse interference, carrier interference, and direct noise interference [29]. These latter two categories can be further divided as well. For example, carrier interference occurs in many forms, such as sine, square and saw-tooth waves, and direct noise interference can be divided into different categories according to its modulation mode, such as amplitude modulation interference, frequency modulation (or phase modulation) interference, and hybrid modulation interference. In addition, the above categories of interference can be divided into generally based on their spectrum characteristics into targeted, multi-frequency and blocking interference categories [30]. For instance, the narrow-band interference is representative of targeted interference. Here, targeted interference typically represents interference with a high degree of similarity to the interfered signals. Multi-frequency interference interferes with multiple carriers, and is characterized as frequency-division, time-division, and comprehensive multi-frequency interference. Among these, multi-tone interference is applied broadly to tactical communications. Blocking interference has the characteristics of wide broadband coverage, such as in the case of chirp interference and comb-spectrum interference. Considering the composition and application of jammer in practical applications, the Gaussian noise substituted for the interference source generator is investigated in this paper.

With the rapid development of signal processing theory, it provides many practical responses to these dilemmas in transforming analysis where it is urgent to improve capabilities of secure and reliable communication for anti-jamming and anti-interception performance, such as principal component analysis, linear discriminant analysis and sparse transformation analysis as well as the corresponding derived methods [31-33]. They convert the detected signal into a new form for pursuing essential correlations, which will commendably adapt to the diversity of generalized signal forms and the selection of optimal transformation. Many methods have been proposed to suppress the interference, including filtering in the time domain and mitigating in the transform domain, which have achieved a good suppression effect on common interference, such as narrowband interference and impulse interference [34]. Sangho et al. utilized the spectrum sensing method for interference signals to opportunistically avoid the jamming spectrum; specifically, they designed a tactical data link with cognitive anti-jamming capability for improved bit-error-rate performance [35]. Yeh et al. verified the system performance of resisting impulse interference and continuous wave interference under the effect of multipath fading and shadow[36]. However, most of these conventional methods can only suppress a specific interference, and reconstruction of the interference cannot be achieved by varying the
parameters; consequently, these methods fail to achieve effective and robust performance in communication systems.

3. Problem Statement

Sparse representation builds the basis of emerging compressive sensing methods, which represents a sparse signal as a series of other forms and most of sparse coefficients close to zero. Thus, the sparse model can be illustrated by the various interference signals as

$$ j = \Psi J \theta J $$

(6)

where $\Psi J \in \mathbb{R}^{N \times N}$ is the complete dictionary for the particular interference $j \in \mathbb{R}^{N \times t}$, and $\theta J \in \mathbb{R}^{N \times t}$ is the corresponding sparse coefficients. According to the characteristic distribution of sparse dictionary and coefficients, we will achieve it for the separation of different formed interference.

The detection of synthesized signal forms at the receiver is not affected in the spectrum sensing stage of TDCS because its signals are highly sparse and noise-like. And then, the synthesized forms may be determined whether the interference exists or not according to the obtained transformation domain. The problem based on hypothesis testing can be presented as

$$ x = \begin{cases} s + j + n, & H_0 \\
 s + n, & H_1 \end{cases} $$

(7)

where $x$ is the perceived electromagnetic information, $s$ is the TDCS communication signal, $j$ (if exists) is electromagnetic interference, and $n$ is the environmental noise.

If the assumption $H_0$ is true, different interference forms, such as the narrowband interference and the multi-tone interference, should be considered. According to the dictionary for signal or interference obtained independently by orthogonal basis learning, the spectrum estimation with sparse representation in transformed domains may be further given by

$$ y' = A y = A (s + j + n) = A (\Psi s \theta_s + \Psi j \theta_j) + n' $$

(8)

where $y'$ denotes the measured spectrum, $A \in \mathbb{R}^{M \times N}$ denotes the measurement matrix, and $n'$ is a measured noise in the compressive sensing procedure.

When a TDCS system is put into the complex electromagnetic environment, the receiver will be situated in different spectrum states, and the received signal can be divided into the known sparsity and unknown sparsity states, which results in the diversity of separating for interference.

3.1 Signal Sensed with Unknown Sparsity

When the receiver loses connections with the transmitter resulting from the strong interference or some potential threats, the signal sparsity states will be changeable or unknown, and we will get into the dilemma without any information about sparse distribution of the signal. Therefore, we can only rely on the pre-collected sparsity library for the unknown signals by spectrum sensing at the receiver, and then achieve the recovery for interference separately.

According to the obtained dictionary for interference using multiple dictionary leaning, the spectrum estimation in multi-task compressed sensing with the noise-like TDCS signals is:

$$ y_k = \Phi s_k (s_k + j_k) + n_k = \Phi s_k \Psi \theta_s + (\Phi s_k + n_k) = A \theta_s + n \theta $$

(9)
where the compressed signal $y_k \in \mathbb{R}^{M \times 1}$, $\Phi_k \in \mathbb{R}^{M \times N}$ is the common measurement matrix, $\Psi_{jk} \in \mathbb{R}^{N \times L}$ is the interference dictionary, and the multiple coefficients $\theta_{jk} \in \mathbb{R}^{L \times 1}$, and $n_{jk}$ is a noise modeled as a Gaussian form with unknown variance $\sigma_k^2$.

In the analytical transforming process, the diverse characteristics of interference signals present on the different transform domains. For the general transformed framework, the optimal sparse transformation $\hat{T}(\cdot)$ will form the restricted conditions for the recovery of the interference signal $j$ based on the assumptions of the separate characteristics for interference and sparse signals, which is illustrated as

$$\min_{\hat{T}(\cdot)} \| \theta_j \|_1$$

s.t. $\| y - A(j + s) \|_2 \leq \varepsilon$

$$\| \theta_j \|_1 \leq K$$

where $\varepsilon$ denotes the recovery error, $K$ denotes the sparsity of interference signal, and $\hat{T}(\cdot)$ denotes the optimal sparse transformation. When the transformed framework $T(\cdot)$ is a strict framework and the measurement matrix $A$ meets the restricted isomorphic property (RIP) about the framework, the unknown interference $j$ can be reconstructed by the $l_1$ analysis.

### 3.2 Signal Sensed with Known Sparsity

In this case, both the transmitter and the receiver keep communicating with each other when the encountered interference varies in its types or parameters, which will result in degrading the system performance. To address this issue, we can achieve the simultaneous separation of interference and signal based on the obtained orthogonal bases, respectively, which will further improve the anti-jamming performance.

According to the dictionary obtained by multiple dictionary leaning, the estimation in multi-task compressed sensing with sparse representation may be given by

$$y_k = \Phi(k_s + j_k) + n_k = \Phi_{s} \Psi_{n} \theta_{n} + \Phi_{j} \Psi_{j} \theta_{j} + n_k$$

where $y_k \in \mathbb{R}^{M \times 1}$ is the compressed signal; $\Phi_{s} \in \mathbb{R}^{M \times N}$ is the common measurement matrix; $\Psi_{n} \in \mathbb{R}^{N \times L}$ and $\Psi_{j} \in \mathbb{R}^{N \times L}$ are multiple dictionaries for TDCS signals and interference, respectively; $\theta_{n} \in \mathbb{R}^{L \times 1}$ and $\theta_{j} \in \mathbb{R}^{L \times 1}$ are multiple coefficients. In detail, $n_k$ is modeled as a zero-mean Gaussian noise with unknown variance $\sigma_k^2$.

Generally, the unknown vector $T(s)$ and $T(j)$ can be represented sparsely in the multiple transform analysis $T$, due to sparsity constraints alone. To achieve transforming a sparse object with less residuals, two sparse signals are transformed diversely $T_1(\cdot), T_2(\cdot)$ into the desired forms when the measured signal $y$ and measurement matrix $A$ are given as well as the sparsity is in the scope of the cardinality of support sets, which will achieve it for the allied sparse separation of interference $j$ and signal $s$ [37]. The problem can be described as
\[
\min_{T_1(s), T_2(j)} \|T_1(s)\|_0 + \|T_2(j)\|_0,
\]
\[
s.t. \quad \|T_1(s)\|_1 \leq K_1,
\]
\[
\|T_2(j)\|_1 \leq K_2,
\]
\[
\|y - A(s + j)\|_p \leq \epsilon.
\]

Furthermore, based on the property of non-convexity for the \( l_0 \) norm, the aforementioned problem can be resolved by the \( l_1 \)-split-analysis [38] when the transformed frameworks \( T_1(\cdot), T_2(\cdot) \) are un-correlation with each other, which can be derived from

\[
\min_{T_1(s), T_2(j)} \|T_1(s)\|_1 + \|T_2(j)\|_1,
\]
\[
s.t. \quad \|T_1(s)\|_1 \leq K_1,
\]
\[
\|T_2(j)\|_1 \leq K_2,
\]
\[
\|y - A(s + j)\|_p \leq \epsilon.
\]

Specially, the accurate recovery of unknown signals \( s \) and \( j \) can be achieved with great probability when \( A \) is a Gaussian random matrix, and the theoretical error between the optimal solution and real solution will be forced into

\[
\|\hat{s} - s\| + \|\hat{j} - j\| \leq C_0 \epsilon + C_1 \sqrt{\|\hat{T}_1(s)_{K_1} - \hat{T}_1(s)\|_1 + \|\hat{T}_2(j)_{K_1} - \hat{T}_2(j)\|_1} / (K_1 + K_2)
\]

where \( C_0 \) and \( C_1 \) are constants according to the signal recovery.

In the conventional perspective of \( l_1 \)-split-analysis, morphological component analysis (MCA) is deemed to be an effective approach for decomposing combined components [39]. For the indeterminate principle in compressive sensing, a sparse signal cannot be represented by more than two sparse base frameworks simultaneously, which means the number of nonzero elements in non-sparse coefficients is far larger than that in sparse coefficients. Therefore, the mixed components \( j \) may be sparse representation with difference by designing the appropriate base framework, which is widely adopted in \( l_1 \)-split-analysis. And the separation method of base framework achieves separation singly utilizing the optimal sparse representation in the obtained over-complete dictionary, which has its advantage on flexibility and generality of sparse signals. For more complex models, however, the aforementioned method has the extra operating burden on transformation and analysis, which leads to the separation errors much more than optimum and is not suitable for the recovery of complex forms in sparse transforming separation. More practical solvers recently, the TFOCS [40] is provided with much computing complexity even in examples of simple linear combinations, and applied more complicatedly to our interference separation problem allied with multi-transforms. In these cases, the extra computing complexity \( O(L^3) \) will get exponential increase with the number of signals \( L \). Hence, it requires to be efficiently solved by a novel algorithm for the separation problem of interference multi-transforms in the more complex electromagnetic situations.

4. Efficient Learning Allied with Base Orthogonality
The optimization problem in Eq. (10) has been solved based on a number of structured libraries investigated in previous studies, such as time-frequency dictionaries, wavelet packets and cosine packets. Most of these libraries can be considered as corresponding to orthogonal bases in some cases. Herein, we denote these libraries as $S$, and an adaptive method [41] is utilized for selecting the optimal basis from $S$ that delivers near-optimal sparsity representations on the order of $N \log(N)$ in terms of time. This method is formulated as follows.

$$\min \Xi\{T_s(x); T_s \in S\}$$  \hfill (15)

Here, the term $\Xi\{T_s(x)\} = \sum_i e[T_s(x_i)]$ represents the level of entropy associated with information theory, and $e[T_s(x_i)]$ is an entropy function with scalar representation, which is defined as

$$e[T_s(x_i)] = \min \left\{\|F_i(x_i) - \Psi^* \Theta^\dagger\|^2 + \alpha \|\Theta^\dagger\|^2, \Lambda_i^{-2}\right\}$$  \hfill (16)

where the parameters $\Lambda_i = \lambda^2 (1 + t_i)^2$. Here, $\lambda > 0$ is the threshold, and $t_i = \sqrt{2 \log(M_i)}$, where $M_i$ represents the total number of distinct vectors occurring among all bases in the library. For example, $M_i = n \log(n)$ for the structured library based on wavelet packets.

According to the definition of information entropy [42], we can obtain the optimal transformation analysis $T_s$ as

$$T_s^*(x) = \arg \min \Xi\{T_s(x); T_s \in S\}$$  \hfill (17)

Here there exists the probability of occurrence exceeding $(1 - \epsilon/M_i)$ when Eq. (17) satisfies the condition $\|T_s^*(x) - T_s(x)\| \leq \frac{\lambda}{t_i - \lambda} \Lambda_i \cdot \min \Xi\{T_s(x) - T_s(x)\}$.

4.1 Orthogonal base learning for representation

The above analyses indicate that one of the subproblems for various forms of the perceived TDCS signal $s$ or interference $j$ (unified as $x \in \mathbb{C}^{N \times 1}$) between (10) and (13) would be adaptively solved by obtaining the optimal orthogonal dictionary $\Psi_s$, $\Psi_j$ (unified as $\Psi^* \in \mathbb{C}^{N \times N}$) respectively through a transformation learning approach, while ensuring that the $\Psi^*$ obtained at each iteration is an orthogonal matrix. An iterative learning approach for the multiple sparse dictionary is presented in this section to ensure orthogonalization of the obtained $\Psi^*$ at each iteration.

Assuming that a random orthogonal dictionary $\Psi, \in \mathbb{C}^{N \times N}$ and a signal dataset $x \in \mathbb{C}^{N \times 1}$ (whether $s(t)$ and $j(t)$ is separate or together) are given initially at the $t^{th}$ iteration, then the nested optimization problem is found as an optimum $\Theta^* \in \mathbb{C}^{N \times 1}$ of

$$\Theta_t = \arg \min_{\Theta \in \mathbb{C}^{N \times 1}} \|x - \Psi \Theta\|^2$$  \hfill (18)

For a temporary fixed $\Psi^*$, the optimal coefficient vector $\Theta^*$ in (18) may be obtained by Moore-Penrose inverse as

$$\Theta^* = I_k(x, \Psi^*) \Psi^{\dagger} x$$  \hfill (19)
where $I_k (x, \Psi)$ e $\mathbb{C}^{N \times N}$ represents a sparse diagonal matrix with $K$ unitary entries, and otherwise zero entries. In each training step, the $K$ largest entries $|\theta|_a$ are selected from $\Psi^H x$ and the corresponding positions $I_{ka}$ in $I_k (x, \Psi)$ are fixed to ones.

Supposing that the obtained coefficient $\Theta$ is a short-term solution, we adopt SVD decomposition for the iterative representation of $\Psi \in \mathbb{C}^{N \times N}$, which is given by

$$
\begin{align*}
U \cdot Z \cdot V^H &= x \Theta^H \\
\Psi' &= U \cdot I \cdot V^H
\end{align*}
$$

(20)

where $U \in \mathbb{C}^{N \times N}$, $I \in \mathbb{C}^{N \times N}$, $Z \in \mathbb{C}^{N \times N}$, and $V \in \mathbb{C}^{N \times N}$.

It is observed that the iterative forms are caught divergence away from the optimum in (19) and (20), whereas the stochastic descent approach motivates the process convergence. Based on the principal component analysis (PCA) [33] for the sparse recovery process, the optimization problem with weighted $l_2$ analysis can be derived from (18)

$$
\begin{align*}
\min_{\theta} & \|x - \Psi' \theta\|_F \|x - \Psi' \theta\|_F^{-1} \\
\text{s.t.} & \|\Theta\|_0 \leq K \\
& \|x - \Psi' \theta\|_F \neq 0
\end{align*}
$$

(21)

However, the listed model in (21) is out of the unique condition but simple for solutions, which is depicted as:

$$
\begin{align*}
\min_{\theta} & \|\Psi' \theta - \Psi' \theta\|_F \\
\text{s.t.} & \|\Theta\|_0 \leq K \\
& \|x - \Psi' \theta\|_F = 0
\end{align*}
$$

(22)

The above-mentioned depiction promotes the allied forms of sparse representation and estimation error whereas deserving to be settled separately. Then, the generalized cost function $J(\cdot)$ (also denoted as the overall error) may be obtained by the extra weight from (21) and (22) as follows:

$$
J(x, \Psi', \omega) = \|w (x - \Psi' \theta)\|_F + \|w (\Psi' \theta - \Psi' \theta)\|_F + \|w\|_F
$$

(23)

where $w \in \mathbb{C}^{N \times N}$ denotes a diagonal matrix for the weighted error.

For minimizing the overall estimation error and simplifying the forms during the iterative procedure, the cost function is utilized by the iterative weighted least squares to avoid getting stuck into the local optima, which may be obtained by

$$
J(x, \Psi', W, a) = \left(x^H W x - x^H \Psi' I_k (x, \Psi') W \Psi' x\right) + a^H W a + W^{-1}
$$

(24)

where $W = \omega W^H \in \mathbb{C}^{N \times N}$, and $a \in \mathbb{C}^{N \times 1}$ denotes the iterative forms of estimation error, yielding the measurement at the ($t+1$)th iteration
\[ a_{i+1} = \min_{a_i} \left\{ a_i, \| \Psi_t^{i+1} \theta - \Psi_t^i \theta \|_F \right\} \] (25)

Furthermore, the iteration stopping criteria are determined by

\[ \Psi_{i+1} = \arg \min_{\Psi_i} \left\{ J(x, \Psi_i^i; W, a_{i+1}) \right\} \] (26)
\[ W_{i+1} = \arg \min_{W_i} \left\{ J(x, \Psi_i^i; W, a_{i+1}) \right\} \] (27)

It is noted that the respective gradients of generalized cost function are given by

\[ \frac{\partial J}{\partial \Psi_i^i} = -x^i \Psi_i^i I_k (x, \Psi_i^i) W_i x \] (28)
\[ \frac{\partial J}{\partial W_i} = x^i x^i - x^i \Psi_i^i I_k (x, \Psi_i^i) \Psi_i^i W_i x + a^i a - W_i^{-2} \] (29)

Therefore, the minimum of (29) can be indicated as explicitly

\[ W_{i+1} = \frac{1}{\sqrt{\| x - \Psi_i^i \theta \|_F + \| a \|}} \] (30)

Then, the stochastic gradient descent method is employed to minimize the cost function in (23), and the global optimum \( \Psi_{t+1} \) at the \( (t+1) \)th iteration may be formulated as

\[ \Psi_{t+1} = OR[\Psi_t + \beta_{t+1} x, x^i W_{t+1}, \Psi_t] \] (31)

where \( x \) is the randomly selected from \( x \) with respect to descending distribution of the gradient function, \( \beta_{t+1} > 0 \) is the learning step, and the Gram–Schmidt orthogonalization operator \( OR[\cdot] \) is performed on the updated base.

To keep descent trend of the cost function, the one-dimensional search method is utilized for the criteria and the learning step in (31) should satisfy

\[ \beta_{t+1} = \arg \min_{\beta > 0} J(\Psi_t + \beta x, x^i W_{t+1}, \Psi_t) \] (32)

For further simplifying the generalized computation of the learning step, it can be exploited commonly as the exponential form with respect to \( \beta_t \), we can iteratively update \( \beta_{t+1} \) by

\[ \beta_{t+1} = \beta_t \left( \frac{t}{t+1} \right)^r \beta_0, t = 1, 2, \ldots, t_{\text{max}} \in \mathbb{Z}^+, r > 0 \] (33)

where \( \beta_0 \) is the initial learning step; \( r \) is an adjustable parameter, \( r < 1 \); and \( t_{\text{max}} \) is the iterations.

Furthermore, after normalizing the basis vector, the selected form \( x_{rs} \) is updated by iterations

\[ x_{rs} = x_{rs} - \Psi_{t+1} \Psi_{t+1}^i x_{rs} \] (34)

After completing the iterative learning process, the obtained orthogonal basis is not required normalization until the sparsity achieves the predetermined settings.

Noticeably, after applying the Gram-Schmidt process for orthogonalization and normalization of the bases, relationships between \( \| a_{i+1} \| \leq \| a_i \|_F \) and \( \| \Psi_{t+1} \| \geq \| \Psi_t \|_F \) are restricted with an arbitrary \( x \) at successive iterations. This provides the corresponding restriction:

\[ J(x, \Psi_{t+1}; W, a) \leq J(x, \Psi_t; W, a) \] (35)
This means the convergent properties of the orthogonal base learning, and it will converge to the global minimum at the end.

4.2 Multiple Bayesian Learning for Recovery

Because one measurement matrix implicitly measures all the coefficients when multiple compressive sensing is conducted for the interference, each compressive measurement is performed with random linear combination of the sparse basis in \( \Psi_{j_k} \), and each random projection corresponds to one measurement matrix [43]. Since the measured interference \( j_k \) is compressive in the basis \( \Psi_{j_k} \), one measurement may approximate \( \theta_{j_k} \) (that is \( j_k \)) accurately by solving an \( l_1 \)-regularized formulation from (10), thus achieving the single recovery for sensed interference with the obtained orthogonal base \( \Psi_{j_k} \):

\[
\theta_{j_k}^* = \arg \min_{\theta_{j_k}} \| y_k - \Phi_{s_k} \theta_{j_k} \|_2 + \lambda_k \| \theta_{j_k} \|_1
\]

Despite many techniques available to perform the inversion for each compressive measurement separately, one measurement may conduct multiple sets of measurements, and many of the measurements are statistically related when repeated measurements are taken in similar scenes or for the interrelated tasks. Therefore, the inversion problem in (36) may be expressed from a Bayesian perspective with a prior belief.

In the context of a regression analysis from (36) and a temporarily-fixed interference \( s_k \) with independence, we assume

\[
y'_k = y_k - \Phi_k \Psi_{s_k} \theta_{s_k} = A_k \theta_{j_k} + n_k
\]

where \( A_k = \Phi_k \Psi_{j_k} \). The likelihood function for \( \theta_{j_k} \) based on the observed \( y'_k \) \( \epsilon \mathbb{R}^{M	imes 1} \) is expressed as

\[
\begin{align*}
\{ p(y'_k | \theta_{j_k}) &= N(y'_k | A_k \theta_{j_k}, \sigma^2) \\
p(y'_k) &= N(A_k \theta_{j_k}, \sigma^2)
\end{align*}
\]

Due to the assumption of independence of the \( y'_k \), the likelihood of \( y'_k \) for the coefficients \( \theta_{j_k} \) and parameters \( \sigma^2 \) of the observed \( y'_k \) can be written as

\[
p(y'_k | \theta_{j_k}, \sigma^2) = \frac{p(y'_k | \theta_{j_k}, \sigma^2)}{p(y'_k | \sigma^2)} = \left( 2\pi \sigma^2 \right)^{-N/2} \exp \left\{ -\frac{\| y'_k - A_k \theta_{j_k} \|^2}{2\sigma^2} \right\}
\]

The coefficients \( \theta_{j} \) are assumed to be drawn from the zero-mean Gaussian prior distributions with the hyperparameter vector \( \alpha^*=\{\alpha_1, \alpha_2, \ldots, \alpha_L\} \) independently shared by all related tasks, and we may obtain

\[
p(\theta_j | \alpha) = \prod_{k=1}^{K} N(\theta_{j_k} | 0, \alpha^{-1})
\]

To promote sparsity of the coefficients \( \theta_{j_k} \), Gamma priors are defined by the hyper-priors over \( \alpha \), which are formulated as

\[
p(\alpha) = \prod_{l=1}^{L} \Gamma(\alpha_l | a, b) = \prod_{l=1}^{L} \Gamma(\alpha_l)^{-1} b^a \alpha_l^{-a-1} e^{-b \alpha_l}
\]
where parameters $a$ and $b$ are fixed to small positive values.

Given the measurements from the statistically-related samples by Bayesian rule, the posterior density function over the hypermeters $\alpha$, coefficients $\theta_J$, and parameters $\sigma^2=\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_K^2\}$ of the observed $y'k$ is

$$
p(\theta_n, a, \sigma_i^2 | y'_i) = \frac{p(y'_i | \theta_n, a, \sigma_i^2) p(\theta_n, a, \sigma_i^2)}{p(y'_i)}
$$

(42)

Since the posterior $p(\theta_n, a, \sigma_i^2)$ cannot be computed directly, we utilize the probability distribution of the observed $y'_i$:

$$
p(y'_i) = \int p(y'_i | \theta_n, a, \sigma_i^2) p(\theta_n, a, \sigma_i^2) d\theta_n d\sigma_i^2
$$

(43)

Then the posterior of $p(\theta_n, a, \sigma_i^2 | y'_i)$ is decomposed as:

$$
p(\theta_n, a, \sigma_i^2 | y'_i) = p(\theta_n | y'_i, a, \sigma_i^2) p(a, \sigma_i^2 | y'_i)
$$

(44)

We can obtain the posterior distribution over the coefficients $\theta_n$ based on the conditional probability density functions:

$$
p(y'_i | a, \sigma_i^2) = \int p(y'_i | \theta_n, \sigma_i^2) p(\theta_n | a) d\theta_n
$$

(45)

Furthermore, the posterior density function for the coefficients can be derived from (45) and thus expressed as

$$
p(\theta_n | y'_i, a, \sigma_i^2) = \frac{p(y'_i | \theta_n, \sigma_i^2) p(\theta_n | a)}{p(y'_i | a, \sigma_i^2)}
$$

(46)

where $\hat{\theta}_n \in \mathbb{R}^{L \times 1}$ and $\Sigma_i \in \mathbb{R}^{L \times L}$, $\hat{\theta}_n = \mu_i = \Sigma_i A_i^T D_i^{-1} y' $, $\Sigma_i = (A_i^T A_i + \hat{Q})^{-1}$. Here, $Q = \text{diag}(a) \in \mathbb{R}^{L \times L}$, $D = \text{diag}(\sigma^2) \in \mathbb{R}^{K \times K}$. Therefore, we obtain the final expressions of the recovered $j_k$ from (40)

$$
j_k^* = \Psi_n \hat{\theta}_n = \Psi_n \Sigma_i A_i^T D_i^{-1} y'
$$

(47)

Above all, according to the updating parameters acquired, we can obtain the closed form solutions $j_k$ by iteratively alternating $\mu_k$ and $\Sigma_k$ concurrently until the convergence reaches a certain degree.

5. Interference Separation Altering with Relaxed AMP

Above we showed that, one measurement matrix implicitly measures all the coefficients when multiple compressive sensing is conducted for the signals, each compressive measurement is performed with random linear combination of the sparse basis, and each random projection corresponds to one measurement matrix, resulting in the problem of allied sparse separation in (11) is substantially more expensive than ordinary separation rules for the redundant computation even the orthogonal base is obtained efficiently. Therefore, multitask compressive sensing approach is a better option than other
methods for recovery and separation of various forms of interference, which could be collected from different locations or transformations and extend the application of the methods to diverse interference forms \[44\].

After the base learning, the first step of the separation is to approximate the ideal components following the message passing rule in a factor graphical model where there exist a variable node set \( V \) (\( N \) variable nodes) and a factor node set \( F \) (\( M \) factor nodes) in each separation task, which is illustrated in Fig. 2.

**Figure 2.** The message passing rule in a factor graphical model obtained for multi-task interference separation problem.

Universally, given the independence between the interference \( j \) and signal \( s \), the posterior joint distribution \( p(s, j; y) \) and marginal likelihood function \( p(y | s, j) \) are derived from the Bayesian rule:

\[
P(s, j; y) = p(y | s, j) p(s, j) = p(y | s, j) p(s) p(j) \\
\propto p(T_1(s)) p(T_2(j)) \prod_{r=1}^{R} p(y_r | s_r, j_r) \tag{48}
\]

where \( R \) denotes the number of measured tasks, \( p(T_1(s)) \) and \( p(T_2(j)) \) are two priors in optimal transformation \( T_1 \) and \( T_2 \), respectively. Given the representation in the factor graph \[45\], the transformation of a particular distribution \( p(\cdot) \) into a distribution domain \( T \) is defined as

\[
P(T_1(s)) = \arg\min_{q \in \mathcal{F}_1} d(s \| q) = \arg\min_{q \in \mathcal{F}_1} \int s(x) \log \frac{s(x)}{q(x)} dx \tag{49}
\]

\[
P(T_2(j)) = \arg\min_{q \in \mathcal{F}_2} d(j \| q) = \arg\min_{q \in \mathcal{F}_2} \int j(x) \log \frac{j(x)}{q(x)} dx \tag{50}
\]

where \( d(\cdot) \) denotes the Kullback-Leibler divergence.

To bypass the original global separation problem in (13) with marginalizing functions, we utilize an alternative quadratic approximation and Taylor series expansions for solving the separation problem. Correspondingly, the proposed algorithm is divided into two phases- approximation and shrinkage, which is illustrated in detail as follow.
5.1 Approximation

In the min-sum manner, the belief propagation posteriors of individual signals \( s \) are estimated by the local message-passing rule, which are given by

\[
\alpha_{j \rightarrow i}(s_j) = \min_{s_i} p(y_{ik} | x_i) \prod_j \alpha_{j \rightarrow i}(s_j) \prod_j \beta_{j \rightarrow i}(J_j)
\]

(51)

\[
\alpha_{j \rightarrow i}(s_j) = p(T_i(s_i)) \prod_i \alpha_{i \rightarrow j}(s_i)
\]

(52)

Similarly, the marginal message functions are measured for the individual interference \( j \)

\[
v_{k \rightarrow j}(J_k) = \min_{x_k} p(y_{nk} | x_k) \prod_j \eta_{j \rightarrow k}(x_k)
\]

(53)

\[
v_{k \rightarrow j}(J_k) = p(T_k(s_k)) \prod_j v_{j \rightarrow k}(J_j)
\]

(54)

where \( u(*) \), \( v(*) \), \( \alpha(*) \), \( \beta(*) \), \( \gamma(*) \), \( \xi(*) \) denote the passing message between the factor node and the variable node.

Then, in view of the opposite direction for passing message at the \( t \)th iteration, the update rules of the expectation propagation [46] at the \((t+1)\)th iteration are expressed as

\[
\alpha_{i \rightarrow j}^{t+1}(s_i) = \frac{p(T_i(s_i)) \prod_i \alpha_{i \rightarrow j}^{t}(s_i)}{u_{i \rightarrow j}^{t}(s_i)}
\]

(55)

\[
\beta_{j \rightarrow i}^{t+1}(J_j) = \frac{p(T_j(J_j)) \prod_j \beta_{j \rightarrow i}^{t}(J_j)}{v_{j \rightarrow i}^{t}(J_j)}
\]

(56)

In this case, each updated passing message may be approximated as a complex Gaussian distribution, and the closed marginal posteriors are denoted as

\[
\alpha_{j \rightarrow i}^{t+1}(s_j) = p(T_i(s_i)) \prod_i \alpha_{i \rightarrow j}^{t+1}(s_i)
\]

(57)

\[
\beta_{j \rightarrow i}^{t+1}(J_j) = p(T_j(J_j)) \prod_j \beta_{j \rightarrow i}^{t+1}(J_j)
\]

(58)

Under the conditions of fixed point solutions for the cost function in (11) with LASSO forms [47], the approximated signal \( s \) and interference \( j \) are estimated by MAP:

\[
s'_i = \arg \min_{s_i} \int \alpha_j(s_j) d(s_j)
\]

(59)

\[
j'_i = \arg \min_{J_i} \int \beta_j(J_j) d(J_j)
\]

(60)

5.2 Shrinkage

To further reduce the number of messages caused by the extra burden of loopy graph connections per iteration, the remaining iterations are shrunk and the iterative shrinkage threshold (IST) [48] is applied to the optimal solutions:
where the iterative form $A_{i\tau}$ is complied with an independent identically distribution of $A_{i\tau}\cdot \tau$ denotes a moderate threshold, which illustrates the overall degree of recovery error with the parameter $\varphi>0$. Here, we define as

$$
\log \varphi_{i}^{t+1} = \log \varphi_{i}^{t} + \tau_{i}^{t} \log \left( \frac{t}{t+1} \right)
$$

Then, we neglect the high order term in the common AMP after the state evolving is altered the estimation, and the residual error $e_{r}$ with Onsager term [49] is approximated through the quadratic Taylor extension, which is obtained by

$$
eq_{r}^{t+1} \approx y_{t} - A_{r} x_{r}^{t} + b_{t} e_{r}^{t}
$$

where $b_{t}$ denotes a reactive factor.

Furthermore, a universal framework of sparse separation for interference is needed to fit all types and present the block-sparse (clustered row-sparse), element-sparse, and finite-difference characteristics simultaneously whereas the relatively constant signal is deemed as infinite difference sparsity in a transitory period, such as partial-band interference around the center tone interfering will become a continuously clustered row-sparse (block sparse) interference, multi-tone interference with marked sparsity can be considered as an element-sparse sporadically interference, and comb-spectrum interference is decorated finite-difference with signal, which forms the diverse ISTs for sparse separation.

1) Separation of finite-difference and block sparsity

For the finite-difference signal $s$ based on the exponential prior, a total-variation soft threshold [50] is applied to estimate the iterative optimum, which is evaluated by

$$
\hat{s}_{i}^{t+1} = \eta_{i} \left( s_{i}^{t}; \tau_{i}^{t} \right) = \arg \min_{x_{i}^{t+1}} \| x_{i}^{t+1} - x_{i}^{t} \|_{TV} + b_{t} \| x_{i}^{t+1} - x_{i}^{t} \|_{2}
$$

where the total-variation norm $\| \cdot \|_{TV}$ is defined as

$$
\| x \|_{TV} = \sum_{m,n} \sqrt{\| x_{m+1,n} - x_{m,n} \|^{2} + \| x_{m,n+1} - x_{m,n} \|^{2}}
$$

And for the block sparse interference $j$ based on the block Gaussian prior, a block soft threshold [51] is applied to pursuit the group sparsity, and the estimation is taken from

$$
\hat{j}_{i}^{t+1} = \eta_{j} \left( j_{i}^{t}; \tau_{i}^{t} \right) = \arg \min_{x_{i}^{t+1}} \| x_{i}^{t+1} \|_{F} + b_{t} \| x_{i}^{t+1} - x_{i}^{t} \|_{F} = \frac{x_{i}^{t+1}}{\| x_{i}^{t+1} \|_{F}} \cdot \max \left( \| x_{i}^{t+1} \|_{F}, \tau_{i}^{t} \right)
$$

where $B$ denotes the number of clustered elements in row (also named block size).

2) Separation of finite-difference and element sparsity

Similarly, the iterative estimation for the finite-difference signal $s$ is evaluated by a total-variation
norm:

\[ \tilde{s}_{i}^{t+1} = \eta_{r}(s'_{j} ; r'_{t}) = \arg \min_{s_{j}} \|s'_{j} + b_{l} \tilde{s}_{j} - x_{j}\|_{p} \]  

(67)

And for the element sparse interference \( j \) based on the Laplace prior, a direct soft threshold [52] is applied to estimate the discrete sparsity, which is given by

\[ \tilde{j}_{j}^{t+1} = \eta_{l}(j_{j} ; r'_{t}) = \arg \min_{j_{j}} \|j_{j} + b_{l} \tilde{j}_{j} - x_{j}\|_{p} \approx \text{sgn}(\frac{x_{j}}{\|x_{j}\|}) \max \{\|x_{j}\|_{p} , r_{t}^{*}\} \]  

(68)

3) Separation of finite-difference and finite-difference

Similarly, the iterative estimation for the finite-difference signal \( s \) is evaluated by a total-variation norm:

\[ \tilde{s}_{i}^{t+1} = \eta_{r}(s'_{j} ; r'_{t}) = \arg \min_{s_{j}} \|s'_{j} + b_{l} \tilde{s}_{j} - x_{j}\|_{p} \]  

(69)

And for the finite-difference interference \( j \) is given by

\[ \tilde{j}_{j}^{t+1} = \eta_{l}(j_{j} ; r'_{t}) = \arg \min_{j_{j}} \|j_{j} + b_{l} \tilde{j}_{j} - x_{j}\|_{p} \]  

(70)

6. Simulations and Results

In this section, we present exemplary comparisons to illustrate the effectiveness and complexity of the proposed ELABO and ISARAMP methods, which includes two different phases of the sparsity separation. The first phase is the sparsity measurement and transformation learning for ELABO algorithm, where the recovery results were compared with orthogonal sparse coding (OSC), geodesic flow OSC (GF-OSC) and their derivative transformation OSC (TS-OSC) for sparse representation [27,28]. Furthermore, the second is sparsity analysis and interference separation for ISARAMP algorithm, where the numerical examples were compared with mixture via AMP (MIX-AMP), improved MIX-AMP (IMIX-AMP), and state evolving AMP (SE-AMP) for sparse separation [41,49].

To generate a synthetic dataset, various signals and several types of interference are considered for diversity measurement. For signals, the information rate was 1 kbit/s, the sampling frequency was 1 kHz, the length of information code was 1,000, the modulation pattern was BCSK, the length of m-sequence was 1,024, and the base function was generated by hard threshold with mean amplitude. Moreover, the initial signal-to-noise ratio was 3 dB, and the transmitter suffered slight narrowband interference with interference-to-noise ratio -3 dB. The typical six-tap multipath channel model was specified by cost207TUx6 for typical urban areas with the Rician fading. As for the sparse diversities of interference parameters, the three components of element-sparse multi-tone (Mt) interference with interval 5 kHz were represented as discrete interference in the simulations. The bandwidth of block-sparse partial-band (Pb) interference was 50 kHz, and its frequency overlap remained 5%. Moreover, the initial carrier frequency of finite-difference chirp (Cp) interference was 75 kHz and its adjustable parameter was 130 whereas the initial carrier frequency of block-sparse comb-spectrum (Cs) interference was 25 kHz and its adjustable parameter was 160. More commonly, the response of element-sparse impulse (Ip) interference was \( (4\sigma(t), 2\sigma(t), 1\sigma(t)) \) and its power was -8 dBW under environment noise. The initial carrier frequency of finite-difference amplitude modulation for noise (An) interference was 80 kHz and its power was -5 dBW under environment noise. The Bernoulli
matrix was chosen as the measurement in the numerical experiments. And the initial signal-to-noise-ratio is 3 dB whereas the initial interference-to-noise-ratio is 8 dB where the environment noise is assumed as a Gaussian white noise channel. The detailed settings of interference datasets considered in this subsection are listed in Table 1.

**Table 1.** Characteristics of interference datasets in simulations.

| Interference types | Initial parameters | Intervals | Variables |
|--------------------|-------------------|-----------|-----------|
| Mt                 | Frequency $f_0=30,55,85$ kHz | 5         | $f_0+5t$ |
| Pb                 | Frequency $f_0=80$ kHz, bandwidth $B_w=50$ | 5         | $f_0+5t$ |
| Cp                 | Frequency $f_0=75$ kHz, adjustment $k=130$ | 5         | $k+5t$   |
| Cs                 | Frequency $f_0=25$ kHz, adjustment $k=160$ | 5         | $f_0+5t$ |
| Ip                 | Response $H=[4,2,1]$, power $P=-8$ dBW | 0.1       | $P=0.1t$ |
| An                 | Frequency $f_0=80$ kHz, power $P=-5$ dBW | 0.1       | $P=0.1t$ |

![Fig. 1](image)

Fig. 1 presents the sparse characteristics of the different interference signals listed above by STFT, and the results indicate that the interference signals include obvious sparse characteristics, and sparsity measurement may be effectively used to reduce the complexity of signal analysis even increase the processing speed.
Figure 1. Sparse Characteristics and power distribution of the aforementioned interference by applying STFT: (a) Mt interference; (b) Pb interference; (c) Cp interference; (d) Cs interference; (e) Ip interference; (f) An interference.

6.1 Sparsity Measurement and Transformation Learning

We investigated the self-adaption and superiority of our proposed ELOBA algorithm for different sparsity of interference, and comparisons of transformation learning were divided into two cases - the diversity of algorithms and base orthogonality. The representative types of interference with distinct sparsity were selected, which includes the element sparse Mt interference and Ip interference, the block sparse Cs interference and the finite-difference An interference. More obviously, the obtained bases of aforementioned four types of interference were converted into gray-scale images, and the middle parts (6×6 diamonds) of the base images were present in Fig. 2, which display the sensitivity of interference for particular frequency and spatial locations.

Figure 2. Gray-scale representatives of interference base images obtained by the ELABO: (a) Mt interference; (b) Cs interference; (c) Ip interference; (d) An interference.

More specifically, Fig. 3 presents the relative estimation errors for the representative interference in transformation analysis, which is obtained by the noiseless and noise measurement, respectively. The relative estimation error is defined as \( \frac{\| f - \hat{f} \|_F}{\| f \|_F} \), and data statistics were analyzed in Table 2.

These results demonstrate the recovery diversities of ELABO applied to multiple forms of interference and the relative errors were obtained ultimately in a convergent way. Obviously, their recovery performances were affected by the measured noise and degraded a lot. Interestingly, for ELABO algorithm, better estimation effect can be achieved for the block-sparse Cs interference with the noiseless measurement whereas the element-sparse Ip interference does better under the impact of measured noise. That is, block-sparse signals are more accessible for the intrinsic correlations adjacenty, and the element-sparse Ip interference is less sensitive for the noise in time domain than the
similar Mt interference, which indicate the moderate adjustment for parameters under noiseless and noise measurement.

**Figure 3.** Relative estimation errors for diverse interference recovery obtained by ELABO: (a) Relative error under the noiseless measurement; (b) Relative error under the noise measurement.

**Table 2.** Data statistics of estimation analyzed under noiseless and noise measurement.

| Statistics  | Mt                | Cs                | Ip                | An                |
|------------|-------------------|-------------------|-------------------|-------------------|
| Maximum    | 0.03604/0.2032    | 0.0296/0.06938    | 0.04823/0.04473   | 0.05683/0.0903    |
| Minimum    | 0.01492/0.07578   | 0.01066/0.04379   | 0.02185/0.02766   | 0.03757/0.06404   |

Furthermore, we compared the estimation performance of proposed ELABO with existing algorithms OSC, GF-OSC and TS-OSC for interference sparse representation in Fig. 4, which includes the block-sparse Pb interference and the infinite-difference Cp interference. We note that the relative estimation errors obtained by ELABO and TS-OSC decreased and achieved convergence more rapidly while the descent trend of ELABO is more robust for successive iterations than TS-OSC. Moreover, initial mean convergent speed (at 10th iteration) of ELABO for sparse interference is higher than OSC,
GF-OSC and TS-OSC about 30.9%, 16.1% and 5.47% respectively whereas the final estimation of GF-OSC achieves similar improvement to ELABO. Furthermore, the complexity comparisons of algorithms are examined for the infinite-difference Cp interference during 50 iterations in Fig. 5 and the running time of ELABO is consumed less than OSC, GF-OSC and TS-OSC about 45.4%, 45.7% and 11.1%, respectively, which present the superiority of performance for interference representation.

**Figure 4.** Relative estimation errors with respect to iteration epochs when applying the OSC, GF-OSC, TS-OSC and ELABO algorithms to sparse interference: (a) the block-sparse Pb interference; (b) the infinite-difference Cp interference.
Figure 5. Complexity comparisons of OSC, GF-OSC, TS-OSC and ELABO algorithms applied for the infinite-difference Cp interference during 50 iterations.

6.2 Sparsity Analysis and Interference Separation

We studied the performance of our proposed ISAR-AMP algorithm for different distribution sparsity of interference, and comparisons of interference separation were divided into three cases according to the intrinsic sparsity which includes the element sparse interference, the block sparse interference and the finite-difference interference. In addition, we keep TDCS signals constant without the impact of interference at receiver and they only vary with the surrounding spectrum at transmitter. The projection matrix is constructed by a Gaussian distribution when the signal-to-noise ratio is 5 dB, and then the matrix is normalized the rows to a unit value. For recovery indicator, we select the relative overall estimation error Oval_err to evaluate the threshold $\tau_r$ of interference separation in (61), and the relative separation distance error Dist_err during the iterations is defined by $\text{Dist}_\text{err} = \frac{||s - tj||_F + ||j - st||_F}{||s + j||_F}$. Furthermore, the relative separation error for interference and signal $j_\text{err}$, $s_\text{err}$ are normalized respectively by $j_\text{err} = \frac{||j_{t+1} - j_t||_F}{||j_t||_F}$, $s_\text{err} = \frac{||s_{t+1} - s_t||_F}{||s_t||_F}$.

The estimation and separation results for block sparse interference and finite-difference signal in Fig. 6-7 show that the performance of ISAR-AMP algorithm was able to match the separation requirements. These results went beyond previous reports, showing that both Pb and Cs interference nearly achieved the bound in separation examples with similarity. Whenever the independent block sparse Pb interference or correlative block sparse Cs interference determined, the distance error $\text{Dist}_\text{err}$ and the signal error $s_\text{err}$ verged to the ideal state gradually whereas the overall error $\text{Oval}_\text{err}$ and the interference error $j_\text{err}$ descended with fluctuation during iterations. This is due to the fact that the size or the number of blocks cannot correspond to the real distribution initially and it deserves consuming more time or iterations for larger block size.
Figure 6. Relative estimation and separation errors when applying ISAR-AMP algorithm to the block sparse Pb interference: (a) relative overall estimation error; (b) relative separation distance error; (c) relative separation error for interference; (d) relative separation error for signal.

Figure 7. Relative estimation and separation errors when applying ISAR-AMP algorithm to the block sparse Cs interference: (a) relative overall estimation error; (b) relative separation distance error; (c) relative separation error for interference; (d) relative separation error for signal.
We present another separation case for element-sparse interference and finite-difference signal in Fig. 8-9. The results show that, whenever the Mt interference or the Ip interference is selected, the separation examples achieved the bound similarly. Moreover, the distance error $\text{Dist}_{\text{err}}$ and the interference error $j_{\text{err}}$ decreased gradually and rapidly converged while the overall error $\text{Oval}_{\text{err}}$ and the signal error $s_{\text{err}}$ are not readily discernible for the ideal state. The tendency difference of iterative results is coming from the fact that the spectrum with random distribution for Mt interference requires more learning steps than the clustered distribution for Ip interference, which indicates that the detailed frequency allocations will promote the Mt interference separation.

**Figure 8.** Relative estimation and separation errors when applying ISAR-AMP algorithm to the element sparse Mt interference: (a) relative overall estimation error; (b) relative separation distance error; (c) relative separation error for interference; (d) relative separation error for signal.
Figure 9. Relative estimation and separation errors when applying ISAR-AMP algorithm to the element-sparse Ip interference: (a) relative overall estimation error; (b) relative separation distance error; (c) relative separation error for interference; (d) relative separation error for signal.

More generally applicable results of finite-difference interference and finite-difference signal are presented in Fig. 10-11, which presents the relative estimation and separation errors obtained with respect to iteration epochs when applying ISAR-AMP algorithm. We note that the proposed method achieved the finite-difference interference separation eventually and the distance error Dist_err converged to the ideal state. Furthermore, whenever the Cp interference or the An interference is confirmed, the interference error \( j_{\text{err}} \) and the signal error \( s_{\text{err}} \) are impacted by random searching steps during the learning period without obtaining the priors in advance, which results in the overall error Oval_err with great fluctuation.
**Figure 10.** Relative estimation and separation errors when applying ISAR-AMP algorithm to the finite-difference sparse Cp interference: (a) relative overall estimation error; (b) relative separation distance error; (c) relative separation error for interference; (d) relative separation error for signal.
6.3 Separation Performance and Running Time

Since our proposed ISAR-AMP algorithm is designed for the interference separation with different distribution sparsity, the experiments were also including three sparse separation cases above, and the performance of estimation was compared with MIX-AMP, IMIX-AMP and SE-AMP algorithms. Here, the selected datasets are same as section 6.2 and the projection matrix is normalized from the i.i.d.-Gaussian distribution when the signal-to-noise ratio is 5 dB. For estimation indicator, we select the normalized mean squared error for signal and interference \( \text{NMSE}_s, \text{NMSE}_j \) are respectively given by:

\[
\text{NMSE}_s = \frac{\sum || \hat{s}_r - s_r ||_F / || s_r ||_F}{R}, \\
\text{NMSE}_j = \frac{\sum || \hat{j}_r - j_r ||_F / || j_r ||_F}{R}.
\]

Fig. 12 presents the estimation and separation results for block sparse interference and finite-difference signal when applying the above algorithms. We note that the proposed ISAR-AMP algorithm has the best performance by exploiting the spatial difference of the block sparse \( P_b \) interference to simultaneously improve the separation accuracy, and even the final overall error of ISAR-AMP algorithm is obtained the improvement about 68.30% compared to SE-AMP. It is also shown that the proposed algorithm achieves a further convergence gain of about 66.67% over the ISAR-AMP algorithm and gets much rapider than MIX-AMP or IMIX-AMP algorithm by imposing shrinkage on the passing message. Moreover, it can be observed that the proposed approach outperforms the conventional sparse separation algorithms of MIX-AMP, IMIX-AMP and SE-AMP by approximately 28.52%, 31.06% and 8.78% in consuming time for 100 iterations, respectively. It is also verified from Tab. 3 that the high accuracy is obtained by the proposed sparse separation method for \( P_b \) interference whereas the NMSE of the proposed algorithm are asymptotically approaching the accuracy of SE-AMP.
Table 3. Diverse estimation errors and running time for the block sparse Pb interference applied with ISAR-AMP, MIX-AMP, IMIX-AMP and SE-AMP algorithms.

| Performance | MIX     | IMIX    | SE      | ISAR    |
|-------------|---------|---------|---------|---------|
| Accuracy    | NMSE_s /e^{-3} | 1.2943  | 1.2942  | 1.3126  | 1.2944  |
|             | NMSE_j /e^{-3} | 2.4999  | 1.4036  | 1.4042  | 1.4051  |
| Complexity  | Time /s   | 1.4584754 | 1.5123366 | 1.1429943 | 1.0425304 |

More generally, the estimation and separation results for element sparse interference and finite-difference signal are illustrated in Fig. 13, which are applied with ISAR-AMP, MIX-AMP, IMIX-AMP and SE-AMP algorithms. We note that the final overall error of the proposed ISAR-AMP algorithm is obtained by 1.735e^{-8} by estimating the discrete difference of the element sparse Mt interference, which improved the separation accuracy greatly than the SE-AMP. In addition, due to similar reduction for the number of messages during the iterations, it is also shown that the proposed algorithm achieves the approximated convergence speed with the ISAR-AMP algorithm whereas they still achieve much higher gains than MIX-AMP or IMIX-AMP algorithm. Moreover, it can be observed that the complexity of proposed approach demonstrates its superiority over the existing MIX-AMP, IMIX-AMP and SE-AMP by approximately 23.62%, 29.38% and 3.50% for 100 iterations, respectively. It is also observed from Tab. 4 that the proposed sparse separation method for Mt interference shares similar performance with SE-AMP about the high accuracy.

Figure 13. Comparisons of relative overall estimation error for the element sparse Mt interference applied with ISAR-AMP, MIX-AMP, IMIX-AMP and SE-AMP.

Table 4. Diverse estimation errors and running time for the element sparse Mt interference applied with ISAR-AMP, MIX-AMP, IMIX-AMP and SE-AMP algorithms.

| Performance | MIX     | IMIX    | SE      | ISAR    |
|-------------|---------|---------|---------|---------|
For separating the finite-difference sparse interference and finite-difference sparse signal, these results in Fig. 14 demonstrate that the proposed ISAR-AMP algorithm yields a better performance about 30.29% than SE-AMP by exploiting the infinite-difference. Furthermore, despite it only provides the inconspicuous improvement in accuracy and running time, the traditional MIX-AMP, IMIX-AMP and SE-AMP algorithms cannot guarantee convergence beyond some certain settings of finite-difference sparse interference. This is due to the fact that the variable prior of finite-difference sparse interference deserves more learning steps and known parameters to further improve the performance.

| Performance | MIX | IMIX | SE | ISAR |
|-------------|-----|------|----|------|
| Accuracy    |     |      |    |      |
| NMSE_s /e^{-3} | 1.3643 | 1.3629 | 1.3627 | 1.3624 |
| NMSE_j /e^{-3} | 1.0140 | 1.0138 | 1.0137 | 1.0701 |
| Complexity  |     |      |    |      |
| Time /s     | 0.3779117 | 0.7114728 | 0.4591012 | 0.5963075 |

**Figure 14.** Comparisons of relative overall estimation error for the finite-difference sparse Cp interference applied with ISAR-AMP, MIX-AMP, IMIX-AMP and SE-AMP.

**Table 5.** Diverse estimation errors and running time for the finite-difference sparse Cp interference applied with ISAR-AMP, MIX-AMP, IMIX-AMP and SE-AMP algorithms.

6.4 Measuring Applications and Impact Factors
To further exploit the impact factors on separation performance of ISAR-AMP based on the divided three sparse cases including block sparse Pb interference, element sparseMt interference and finite-difference sparse Cp interference, we select the interference spectrum, interference power and noise power from the frequency and spatial domains. In details, the variable spectrum dataset of measured interference is determined by the settings in Table 1, where the frequency range of Pb interference varies from 80 kHz to 330 kHz, the frequency range of Mt interference varies from [30, 55, 85] kHz to [280, 305, 335] kHz, and the adjustable rate range of Mt interference varies from 130 to 380. For the interference and noise power, all components amplitudes of interference are ranging from 0 to 5 with the interval 0.1 when the noise power is constant; besides, all interference-noise-ratios (INRs) of interference change from -5 dB to 0 dB with the interval 0.1 dB. Then, we select the mean squared error for signal and interference as estimation indicators $\text{NMSE}_s$ and $\text{NMSE}_j$, which are defined by $\text{MSE}_s = \frac{1}{\|s\|_p} \|s - \hat{s}\|_p$ and $\text{MSE}_j = \frac{1}{\|j\|_p} \|j - \hat{j}\|_p$, respectively. The experiment results for three sparse interference separation cases during 100 iterations are conducted in Fig. 16-18.

The MSEs of block sparse Pb interference separation in Fig. 15 show that the performance of ISAR-AMP is impacted by frequency, amplitude and INRs diversely in some degree. With the increase of the interference frequency in Fig. 15 (a), the separation performance present periodical regularity due to the bandwidth of Pb interference keeps constant. In contrast, the separation performance is get improvement with the increase of interference amplitude whereas the increase of the MSEs is much slower with a positive correlation between INRs. Hence, it can be concluded that the proposed IASR-AMP algorithm will greatly reduce the separation errors of measurement data required for accurate recovery when encountered high power or low INRs of Pb interference, achieving higher anti-jamming efficiency than conventional counterparts.
Figure 15. MSEs of block sparse Pb interference separation applied with ISAR-AMP algorithm under different impact factors where bar graphs denote the MSE_s and MSE_j whereas the line graph denotes the mean value of MSEs: (a) relations between MSEs and frequency; (b) relations between MSEs and amplitude; (c) relations between MSEs and INRs.

The separation performance of the proposed method for element sparse Mt interference versus the frequency, amplitude and INRs is depicted in Fig. 16. We noted that the separation performance cannot present obvious regularity with the increase of the interference frequency whereas the proposed algorithm reaches less separation errors when the interference amplitude and INRs grow more, respectively. It is thus validated that the proposed method can accurately separate the Mt interference with much higher power and high INRs under the conventional spectrum estimation. Since the measured Mt interference only occupies 3 locations with 25-30 kHz intervals in the frequency spectrum, it can be inferred that the proposed method is capable of effectively separating and recovering at least 30 kHz out-band Mt interference.
Figure 16. MSEs of element sparse Mt interference separation applied with ISAR-AMP algorithm under different impact factors where bar graphs denote the MSE_s and MSE_j whereas the line graph
denotes the mean value of MSEs: (a) relations between MSEs and frequency; (b) relations between MSEs and amplitude; (c) relations between MSEs and INRs.

To further exploit more complex distribution of sparsity interference, we compared the MSE performance of the proposed method for block sparse Cs interference under different impact factors, which is illustrated in Fig. 17. Apart from the influence of interference amplitude, the separation performance of the ISAR-AMP method achieves slight improvement or degradation with the increase of frequency or INRs. Furthermore, it can be observed that the proposed algorithm for dealing with the high power components of Cs interference significantly outperforms the state-of-the-art counterparts and even presents positive reciprocal relation with interference amplitude. The reason is that the component accumulations of Cs interference have enough discriminating difference between various signals even though lacking unknown parameters. This implies that the Cs can be more effectively separated and recovered in the proposed message passing framework of sparse approximation using the iterative learning.
Figure 17. MSEs of infinite-difference sparse Cs interference separation applied with ISAR-AMP algorithm under different impact factors where bar graphs denote the MSE_s and MSE_j whereas the line graph denotes the mean value of MSEs: (a) relations between MSEs and frequency; (b) relations between MSEs and amplitude; (c) relations between MSEs and INRs.

7. Conclusions

In this paper, a novel sparse learning method-based cognitive transformation framework of interference separation is formulated for the coexistence of diverse interference and TDCS systems. The combinatorial optimization problem of multiform interference separation is efficiently and accurately solved by the proposed sparse learning algorithms of ELABO and ISAR-AMP, which iteratively learns the prior sparse probability distribution, i.e., the sparse pattern, of the interference support by minimizing the loss function of state evolution. By imposing spectrum analysis on the sparsity states, the intricate interference separation problem is generalized from single separation and simultaneous separation. Furthermore, the principal component analysis and Bayesian perspective in orthogonal base learning ELABO is exploited to simultaneously recover the multiple interference and TDCS signals, which achieves better separation accuracy and iterative convergence. Moreover, the proposed novel interference separation algorithm based on state evolving iteratively learns the belief propagation posteriors in the min-sum manner, and the number of passing messages is shrunk by iterative shrinkage threshold. It is verified by theoretical analysis and numerical simulation results that the proposed algorithms outperform state-of-the-art counterparts in separation accuracy and computational complexity, which is especially suitable for interference separation and elimination of broadband communication.

Author Contributions: Conceptualization, W.G. and W.Y.; methodology, W.G. and H.G.; software, D.S.; validation, W.G. and S.Q.; formal analysis, W.G.; investigation, W.G.; resources, W.G.; data curation, Z.Y.; writing—original draft preparation, W.G.; writing—review and editing, W.Y.; visualization, H.G.; supervision, D.S.; project administration, S.Q.; funding acquisition, H.G.
**Funding:** This research was funded in part by the National Natural Science Foundation of China, grant number 61701521, and in part by the University Cooperation Project of Key Laboratory for Aerospace Information Technology, grant number KX162600022.

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

**References**

[1] L. Chettri and R. Bera, "A Comprehensive Survey on Internet of Things (IoT) Toward 5G Wireless Systems," in IEEE Internet of Things Journal, vol. 7, no. 1, pp. 16-32, Jan. 2020, doi: 10.1109/JIOT.2019.2948888.

[2] C. Fager, T. Eriksson, F. Barradas, K. Hausmair, T. Cunha and J. C. Pedro, "Linearity and Efficiency in 5G Transmitters: New Techniques for Analyzing Efficiency, Linearity, and Linearization in a 5G Active Antenna Transmitter Context," in IEEE Microwave Magazine, vol. 20, no. 5, pp. 35-49, May 2019, doi: 10.1109/MMM.2019.2898020.

[3] Ding, Zg., Xu, M., Chen, Y. et al. Embracing non-orthogonal multiple access in future wireless networks. Frontiers Inf Technol Electronic Eng 19, 322–339 (2018). https://doi.org/10.1631/FITEE.1800051

[4] J. Liu, X. Jiang, X. Tian, M. Mallick, K. Huang and C. Ma, "Hybrid Particle Filter Based Dynamic Compressed Sensing for Signal-Level Multitarget Tracking," in IEEE Access, vol. 8, pp. 17134-17148, 2020, doi: 10.1109/ACCESS.2020.2967550.

[5] H. Zayyani, R. Sari and M. Korki, "A Distributed 1-bit Compressed Sensing Algorithm for Nonlinear Sensors With a Cramer–Rao Bound," in IEEE Communications Letters, vol. 21, no. 12, pp. 2626-2629, Dec. 2017, doi: 10.1109/LCOMM.2017.2748943.

[6] J. Zhao, Q. Liu, X. Wang and S. Mao, "Scheduled Sequential Compressed Spectrum Sensing for Wideband Cognitive Radios," in IEEE Transactions on Mobile Computing, vol. 17, no. 4, pp. 913-926, 1 April 2018, doi: 10.1109/TMC.2017.2744621.

[7] Lin R, Bi G, Liu X, et al., 2014. On the modulation and signalling design for a transform domain communication system. IET Communications, 8(99):2909-2916.

[8] Hu S, Guan Y, Bi G, et al., 2012. Cluster-based transform domain communication systems for high spectrum efficiency. IET Communications, 6:2734-2739.

[9] Foucart S, Rauhut H, 2013. A Mathematical Introduction to Compressive Sensing. Springer, New York, America, p.10-15.

[10] LIU S, XIAO L, HUANG F, et al. Impulsive noise recovery and elimination: A sparse machine learning based approach[J]. IEEE Transactions on Vehicular Technology, 2019, PP(99): 1-13.

[11] ZHANG Y, JIA X, YIN C, et al. NBI mitigation in DSSS communications via block sparse Bayesian learning[J]. Signal Processing, 2019, 158: 129-140.

[12] ZHANG Y, JIA X. Adaptive interference suppression for DSSS communications based on compressive sensing. Int J Commun Syst, 2018, 31(11): e3699.

[13] Zhang Y, Jia X, Kou B, 2015. Adaptive Multi-tone Jamming Suppression for DSSS Communications Based on Compressive Sensing. 8th International Congress on Image and Signal Processing,10:1323-1327.

[14] LIU S, FANG Y, DING W, et al. Two-dimensional structured-compressed-sensing-based NBI cancelation exploiting spatial and temporal correlations in MIMO Systems[J]. IEEE Transactions on Vehicular Technology, 2016, 65(11): 9020-9028.
[15] LIU S, FANG Y, SONG J, et al. Block sparse bayesian learning based NB-IoT interference elimination in LTE-Advanced systems[J]. IEEE Transactions on Communications, 2017, PP(99): 1-16.

[16] Liu S, Yang F, Ding W, et al., 2016. Double Kill: Compressive-Sensing-Based Narrow-Band Interference and Impulsive Noise Mitigation for Vehicular Communications. IEEE Transactions on Vehicular Technology, 65(7):5099-5111.

[17] Tengstrand S, Eliardsson P, Axell E, 2016. Mitigation of Multiple Impulse Noise Sources Through Selective Attenuation, IEEE Conference on Military Communications, 4:1-6.

[18] Axell E, Eliardsson P, Tengstrand S, et al., 2017. Power Control in Interference Channels With Class A Impulse Noise. IEEE Transactions Wireless Communications Letters, 6(1):1-6.

[19] Wang G, Wang Y, Huang G, et al., 2020. Classification methods with signal approximation for unknown interference. IEEE Access, https://doi.org/ 10.1109/ACCESS.2020. 2976128.

[20] Z. Shakeri, W. U. Bajwa and A. D. Sarwate, "Minimax Lower Bounds on Dictionary Learning for Tensor Data," in IEEE Transactions on Information Theory, vol. 64, no. 4, pp. 2706-2726, April 2018, doi: 10.1109/TIT.2018.2799931.

[21] A. Seghouane, A. Iqbal and K. Abed-Merai, "A Sequential Block-Structured Dictionary Learning Algorithm for Block Sparse Representations," in IEEE Transactions on Computational Imaging, vol. 5, no. 2, pp. 228-239, June 2019, doi: 10.1109/TCI.2018.2884809.

[22] J. Yang, Y. Guo, Z. Yang and S. Xie, "Under-Determined Convolutinal Blind Source Separation Combining Density-Based Clustering and Sparse Reconstruction in Time-Frequency Domain," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 66, no. 8, pp. 3015-3027, Aug. 2019, doi: 10.1109/TCSL.2019.2908394.

[23] T. T. Chan and Y. Yang, "Informed Group-Sparse Representation for Singing Voice Separation," in IEEE Signal Processing Letters, vol. 24, no. 2, pp. 156-160, Feb. 2017, doi: 10.1109/LSP.2017.2647810.

[24] N. Ruozzi and S. Tatikonda, "Message-Passing Algorithms: Reparameterizations and Splittings," in IEEE Transactions on Information Theory, vol. 59, no. 9, pp. 5860-5881, Sept. 2013, doi: 10.1109/TIT.2013.2259576.

[25] C. K. Thomas and D. Slock, "Low Complexity Static and Dynamic Sparse Bayesian Learning Combining BP, VB and EP Message Passing," 2019 53rd Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 2019, pp. 685-689, doi: 10.1109/IEEECONF44664.2019.9048860.

[26] Haykin,S. and M.Moher, Modern Wireless Communications, Prentice Hall, New York, 2005.

[27] Schiitze H, Barth E, Martinez T, 2016. Learning efficient data representations with Orthogonal Sparse Coding, IEEE Transactions on Computational Imaging, 2(3):1-77-189.

[28] Schiitze H, Barth E, Martinez T, 2015. Learning Orthogonal Sparse Representations by Using Geodesic Flow Optimization. International Conference on Neutral Networks,7:1-8.

[29] Axell E, Eliardsson P, Tengstrand S, et al., 2017. Power Control in Interference Channels With Class A Impulse Noise. IEEE Transactions Wireless Communications Letters, 6(1):1-6.

[30] G. Wang, Q. Ren and Y. Su, "The interference classification and recognition based on SF-SVM algorithm," 2017 IEEE 9th International Conference on Communication Software and Networks (ICCSN), Guangzhou, 2017, pp. 835-841, doi: 10.1109/ICCSN.2017.8230229.

[31] Yang Z, Tao R, Wang Y, et al., 2014. A Novel Multi-carrier Order Division Multi-access Communication System Based on TDCS with Fractional Fourier Transform Scheme. Wireless Personal Communications, 79:301-1320.

[32] Fraley C, Raftery AE, 2002. Model-based clustering, discriminant analysis, and density estimation, J. Amer. Statist. Assoc., 97(458): 611–631.
[33] A. Seghouane, N. Shokouhi and I. Koch, "Sparse Principal Component Analysis With Preserved Sparsity Pattern," in IEEE Transactions on Image Processing, vol. 28, no. 7, pp. 3274-3285, July 2019, doi: 10.1109/TIP.2019.2895464.

[34] G. Wang, Q. Sun, T. Ren and Y. Zhang, "Interference Estimation and Recovery in Transform Domain Communication System," 2020 International Conference on Computer Communication and Network Security (CCNS), Xi'an, China, 2020, pp. 196-200, doi: 10.1109/CCNS50731.2020.00050.

[35] Sangho Choe, Euny Ko, and Janghyuk Lee. Tactical Data Link with Cognitive Anti-Jamming Capability and its Simulator , 2015IEEE MILCOM,607-611.

[36] S. Y. Yeh, J. Y. Chu and T. Y. Hsu, "Poster: Anti-jamming automatic gain control and packet detection for vehicular receiver," in Proceedings of IEEE Vehicular Networking Conference, Kyoto, Japan, 2015, pp.155-156. doi:10.1109/VNC.2015. 7385564

[37] Eldar Y, Kutyniok G, et al., 2012. Compressed Sensing: Theory and Applications, Cambridge University Press, London, UK, p.380-389.

[38] X. Xu, J. Li, S. Li and A. Plaza, "Generalized Morphological Component Analysis for Hyperspectral Unmixing," in IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 4, pp. 2817-2832, April 2020, doi: 10.1109/TGRS.2019.2956562.

[39] A. Y. Yang, Z. Zhou, A. G. Balasubramanian, S. S. Sastry and Y. Ma, "Fast l1-Minimization Algorithms for Robust Face Recognition," in IEEE Transactions on Image Processing, vol. 22, no. 8, pp. 3234-3246, Aug. 2013, doi: 10.1109/TIP.2013.2262292.

[40] R. R. Coifman and M. V. Wickerhauser, "Entropy-based algorithms for best-basis selection", IEEE Trans. Info. Theory 38 (1992) p. 713-718.

[41] J. Kang, H. Jung and K. Kim, "Fast Signal Separation of 2-D Sparse Mixture via Approximate Message-Passing," in IEEE Signal Processing Letters, vol. 22, no. 11, pp. 2024-2028, Nov. 2015, doi: 10.1109/LSP.2015.2454003.

[42] L. B. Kish, G. A. Niklasson, C. G. Granqvist, D. K. Ferry and J. M. Smulko, "Facts and myths about zero-point thermal noise, and information entropy versus thermal entropy," 2017 International Conference on Noise and Fluctuations (ICNF), Vilnius, 2017, pp. 1-2, doi: 10.1109/ICNF.2017.7985927.

[43] Qisong Wu, Yimin D. Zhang, Moeness G. Amin, and Braham Himed. Multi-Task Bayesian Compressive Sensing Exploiting Intra-Task Dependency, IEEE Signal Processing Letters,

[44] David P. Wipf and Bhaskar D. Rao, An Empirical Bayesian Strategy for Solving the Simultaneous Sparse Approximation Problem, IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 55, NO. 7, JULY 2007,3704-3716.

[45] X. Meng, S. Wu, M. R. Andersen, J. Zhu and Z. Ni, "Efficient recovery of structured sparse signals using approximate message passing with structured spike and slab prior," in China Communications, vol. 15, no. 6, pp. 1-17, June 2018, doi: 10.1109/CC.2018.8398220.

[46] X. Meng, S. Wu, L. Kuang and J. Lu, "An Expectation Propagation Perspective on Approximate Message Passing," in IEEE Signal Processing Letters, vol. 22, no. 8, pp. 1194-1197, Aug. 2015, doi: 10.1109/LSP.2015.2391287.

[47] N. Radmanesh, I. S. Burnett and B. D. Rao, "A Lasso-LS Optimization with a Frequency Variable Dictionary in a Multizone Sound System," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 24, no. 3, pp. 583-593, March 2016, doi: 10.1109/TASLP.2016.2518801.
[48] V. W. Zibetti, E. S. Helou and D. R. Pipa, "Accelerating Overrelaxed and Monotone Fast Iterative Shrinkage-Thresholding Algorithms With Line Search for Sparse Reconstructions," in IEEE Transactions on Image Processing, vol. 26, no. 7, pp. 3569-3578, July 2017, doi: 10.1109/TIP.2017.2699483.

[49] K. Takeuchi, "Convolutional Approximate Message-Passing," in IEEE Signal Processing Letters, vol. 27, pp. 416-420, 2020, doi: 10.1109/LSP.2020.2976155.

[50] Y. Zhang, N. He, X. Zhen and X. Sun, "Image Denoising Based on the Wavelet Semi-soft Threshold and Total Variation," 2017 International Conference on Vision, Image and Signal Processing (ICVISP), Osaka, 2017, pp. 55-62, doi: 10.1109/ICVISP.2017.16.

[51] P. V. V. Kishore, A. S. C. S. Sastry, A. Kartheek and S. H. Mahatha, "Block based thresholding in wavelet domain for denoising ultrasound medical images," 2015 International Conference on Signal Processing and Communication Engineering Systems, Guntur, 2015, pp. 265-269, doi: 10.1109/SPACES.2015.7058262.

[52] D. L. Donoho, I. Johnstone and A. Montanari, "Accurate prediction of phase transitions in compressed sensing via a connection to minimax denoising", IEEE Trans. Inf. Theory, vol. 59, no. 6, pp. 3396-3433, Jun. 2013.