Compressive sensing meets time-frequency: An overview of recent advances in time-frequency processing of sparse signals

Ervin Sejdić[a], Irena Orović[b], Srdjan Stanković[b]

[a]*Department of Electrical and Computer Engineering, Swanson School of Enginering, University of Pittsburgh, Pittsburgh, PA, 15261, USA*
[b]*Faculty of Electrical Engineering, University of Montenegro, Podgorica, Montenegro*

Abstract

Compressive sensing is a framework for acquiring sparse signals at sub-Nyquist rates. Once compressively acquired, many signals need to be processed using advanced techniques such as time-frequency representations. Hence, we overview recent advances dealing with time-frequency processing of sparse signals acquired using compressive sensing approaches. The paper is geared towards signal processing practitioners and we emphasize practical aspects of these algorithms. First, we briefly review the idea of compressive sensing. Second, we review two major approaches for compressive sensing in the time-frequency domain. Thirdly, compressive sensing based time-frequency representations are reviewed followed by descriptions of compressive sensing approaches based on the polynomial Fourier transform and the short-time Fourier transform. Lastly, we provide brief conclusions along with several future directions for this field.

Keywords:
Compressive sensing, time-frequency analysis, time-frequency dictionary, nonstationary signals, sparse signals

1. Introduction

Time-frequency analysis provides a framework for a descriptive analysis of non-stationary signals whose models are not available or easily constructed [1], [2]. For such signals, time or frequency domain descriptions typically do not offer comprehensive details about changes in signal characteristics [3]. The main issue with the time domain representation is that it provides no details about the frequency content of those signals, and sometimes, even the time content can be difficult to interpret [4]. The frequency domain on the other hand provides no easily understood timing details about the occurrence of various frequency components [5]. In other words, timing details are buried within the phase spectrum of a signal, which is the most common reason for only analyzing the amplitude spectrum of a signal obtained via the Fourier transform. To combine timing and spectral into a joint representation, a time variable is typically introduced into a Fourier-based analysis to obtain two-dimensional, redundant representations of non-stationary signals [6]. Such representation provide a description of spectral signal changes as a function of time, that is, the description of time-varying energy concentration changes along the frequency axis. In an ideal case, these two-dimensional signal representations would combine instantenous frequency spectrum with global temporal behavior of a signal [7], [8], [9],[10], [11], [12], [13], [14], [15].

Time-frequency analysis is often employed in the analysis of complex non-stationary signals (e.g., physiological signals [16], [17], [18], [19], [20], [21], [22], [23], mechanical vibrations [24], [25], [26], [27], audio signals [28], [29], [30], radar signals [31], [32], [33], [34], [35], [36]). However, continuously monitoring such signals for an extended period of time can impose heavy burdens on data acquisition and processing systems, even when sampling these non-stationary signals at low sampling rates. To avoid these data acquisition and processing burdens, compressive sensing aims to compress signals during a data acquisition process, rather than afterwards [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52].

In this paper, we review recent advances that combine the ideas of time-frequency and compressive sensing analyses. Section 2 reviews the main ideas behind compressive sensing. In Section 3, we introduce the main approaches to obtain compressed samples in the time-frequency domain. Several different approaches are presented here including compressive sensing in the ambiguity domain, but also compressive sensing of non-stationary signals using time-frequency dictionaries. We also reviewed compressive sensing approaches that relied on the short-time Fourier transform and the polynomial Fourier transform. Compressive sensing based time-frequency representations are described in Section 4. Conclusions and future directions are provided in Section 5.

2. Compressive Sensing

In traditional signal processing, the Shannon-Nyquist sampling theorem mandates that any signal needs to be sampled at least twice the highest frequency present in the signal to be able to accurately recover information present in the signal. The traditional sampling approach can yield a large number of samples, and compressive strategies are often used immediately after sampling in order to reduce storage requirements or transmission complexities. While this has been a prevailing approach for many years, it is clearly a redundant approach as most of acquired samples are disregarded immediately after sampling. To avoid these redundant steps, compressive sensing has been proposed and it postulates a signal can be recovered using a fewer number of samples than required by the Shannon-Nyquist theorem [38], [39], [53], [40], [54] [55], [56], [57], [58], [59], [60], [61], [62], [63].

The main idea behind compressive sensing is to combine sensing and compression steps into a single step during a data acquisition process [39], [40], [42], [64], [65]. Compressive sensing approaches typically acquire signals at sub-Nyquist rates (e.g., one tenth of the Nyquist rate) and signals can be accurately recovered from these samples with *a certain probability* [39]. These approaches work very well for K-sparse signals, i.e., signals that can be represented by K bases in an N-dimensional space. In other words, compressive sensing approaches will acquire $M \ll N$ samples that will encode a K-sparse signal of dimension N by computing a measurement vector $\mathbf{y}$ of a signal vector $\mathbf{s}$ [66], [67], [68], [69]:

$$\mathbf{y} = \mathbf{\Phi s} \tag{1}$$

where $\mathbf{\Phi}$ represents an $M \times N$ sensing matrix [40]. The signal vector $\mathbf{s}$ can be recovered from sparse samples by utilizing a norm minimization approach:

$$\min ||\mathbf{s}||_0 \text{ subject to } \|\mathbf{y} - \mathbf{\Phi s}\|_2 < \xi \tag{2}$$

where ξ is measurement noise, $||s||_0$ represents the number of nonzero entries of $\mathbf{s}$ and $\|\bullet\|_2$ is the Euclidian norm. However, it should be mentioned that it is not guaranteed that eqns. (1) and (2) will provide an accurate representation of sparse signals. In some applications that are sensitive to small changes such as medical diagnostic applications, it is almost mandatory to achieve almost perfect recovery of these sparse signals, otherwise compressive sensing schemes are not useful at all in medical diagnostic applications. To reach these almost perfect reconstructions of sparse signals, compressive sensing can be performed in other domains (i.e., other than the time domain), which yields a new reformulation of the compressive sensing appraoch proposed in (1) as [64], [67]:

$$\mathbf{y} = \mathbf{\Phi}\mathbf{s} = \mathbf{\Phi}\mathbf{\Psi}\mathbf{x}. \tag{3}$$

Here, $\mathbf{x}$ is the vector of expansion coefficients representing the sparse representation of the signal $\mathbf{s}$ in the basis $\mathbf{\Psi}$. A very good example of this change is representing a single sinusoid in the frequency domain. This transformation would enable us to represent such a sinusoid with by a two-sparse vector. In this paper, this change of the domain is achieved by representing a signal in the time-frequency domain, rather than using its time-domain samples.

It should be understood that the compressive sensing approach proposed by eqn. (3) affects the sparsity in the transform domain, which then inherently affects the number of measurements needed to reconstruct a signal. This is assessed using the so-called coherence measure between the matrices $\mathbf{\Phi}$ and $\mathbf{\Psi}$ [70], [71], [72], [73]:

$$\mu(\mathbf{\Phi}, \mathbf{\Psi}) = \sqrt{N} \max |\langle \phi_k, \psi_j \rangle| \tag{4}$$

where N is the signal length, ϕ_k is the k^{th} row of $\mathbf{\Phi}$, and ψ_j is the j^{th} row of $\mathbf{\Psi}$. Smaller values of the coherence measure typically denote that a smaller number of random measurements is needed to accurately reconstruct a signal.

3. Time-frequency based compressive sensing

The time-frequency domain represent an ideal domain to sparsely represent nonstationary signals for several different reasons. First, it is very difficult to represent nonstationary signals sparsely either in time or frequency domains. For example, a frequency modulated signal is concentrated along its instantaneous frequency in the time-frequency domain, and most of other values are equal to zero. But, its frequency domain representation has many non-zero components, and its time domain representation typically has many (large) amplitude changes that can be difficult to compress. Therefore, such a frequency modulated signal or any other signal with complex non-stationary structures should be compressively sampled in the time-frequency domain, as their representations are often sparse in the time-frequency domain [74], [75], [76]. Second, recent advances in computational resources enabled fast manipulations of large matrices, which are required for compressive sensing of nonstationary signals in the time-frequency domain [77].

In this section, we will overview two major approaches for compressive sensing of nonstationary signals in the time-frequency domain. We will begin with compressively sampling a nonstationary signal in the ambiguity domain as proposed in [78] with understanding that this approach is only applicable for quadratic time-frequency representations. A more general approach is to utilize time-frequency dictionaries to obtain a sparse time-frequency representation of a nonstationary signal, which can be then used to compressively sensed such a signal.

3.1. Compressive sensing in the ambiguity domain

As mentioned in the previous paragraph, the ambiguity domain provides a suitable framework to compressively sampled non-stationary signals. To achieve representations in the ambiguity domain, we can start with the Wigner-Ville distribution, $WVD(t,f)$, and take the two-dimensional Fourier transform of it to obtain the ambiguity domain representation [1], [79]:

$$A_x(\nu,\tau) = \mathcal{F}_{2D}\{WVD(t,f)\} \tag{5}$$

where $\mathcal{F}_{2D}$ is the forward and inverse two-dimensional Fourier operator. The ambibuity domain offers an opportunity to supress or completely remove cross-terms, which plague the quadratic time-frequency representations, as cross-terms are typically displaced from the origin in the ambiguity domain, and the auto-terms are typically centered around the origin. Therefore, low-pass filtering by multiplying the ambiguity representation of the signal,$A_x(\nu,\tau)$, by a kernel function $k(\nu,\tau)$:

$$\mathcal{A}_x(\nu,\tau) = A_x(\nu,\tau)k(\nu,\tau). \tag{6}$$

However, it should be mentioned here that compressive sensing appraoches here are mostly used to obtain enhanced time-frequency signal energy localization in the time-frequency domain. Specifically, we compressively sample the ambiguity domain representation of the signal in order to obtain a very sparse time-frequency domain signal representation. This is achieved by solving the l_1-norm minimization problem to obtain a sparse time-frequency distribution $\widehat{\Upsilon}_x(t,f)$:

$$\widehat{\Upsilon}_x(t,f) = \arg\min_{\Upsilon_x(t,f)} \|\Upsilon_x(t,f)\|_1 \tag{7}$$

$$F_{2D}\{\Upsilon_x(t,f)\} - \mathcal{A}_x^M = 0\big|_{(\nu,\tau)\in\Omega} \tag{8}$$

where $\mathcal{A}_x^M$ denotes the set of samples from the ambiguity domain in the region defined by the mask $(\nu,\tau) \in \Omega$, $\Upsilon_x(t,f)$ denotes the time-frequency distribution, and $\|\cdot\|_1$ denotes the ℓ_1 norm. Noise distributions can be approximated as follows:

$$\widehat{\Upsilon}_x(t,f) = \arg\min_{\Upsilon_x(t,f)} \|\Upsilon_x(t,f)\|_1 \tag{9}$$

$$\|F_{2D}\{\Upsilon_x(t,f)\} - \mathcal{A}_x^M\|_2 \leq \epsilon\big|_{(\nu,\tau)\in\Omega} \tag{10}$$

One has to carefully select samples in the ambiguity domain via an appropriate ambiguity function masking, which is formed as a small, mostly rectangular, area around the origin in the ambiguity plane. This resembles an approach taken to achieve high-resolution time-frequency distributions [80], [81], as we design the mask to pass auto-terms, and reduce cross-terms.

As an illustrative example of this approach, let us consider a sinusoidally-modulated signal with additive white Gaussian noise. For illustrative purposes, the time-frequency representations in Figure 1 are of size 60×60 points. Here, we consider a very small rectangular mask of size 7×7 points in the ambiguity domain, which represents just over one percent of the total number of samples. Now, let us consider the time-frequency representations of the signal. Figure 1(a) depicts the signal in the ambiguity domain which is observed as a domain of observations. The Wigner distribution is considered as a standard time-frequency distribution that can be derived from the ambiguity function in Figure 1(a) and it is the most appropriate to the considered signal type. This standard form of the time-frequency distribution is used only for the comparison purpose and it is calculated assuming the full set of samples from the ambiguity domain is available, Figure 1(b). Lastly, the compressive sensing based sparse time-frequency representation is presented in Figure 1(c). Unlike the standard time-frequency representation that requires full set of samples, the sparse representation is calculated from a very limited number of available samples. Despite this fact, we may observe that compared to the standard

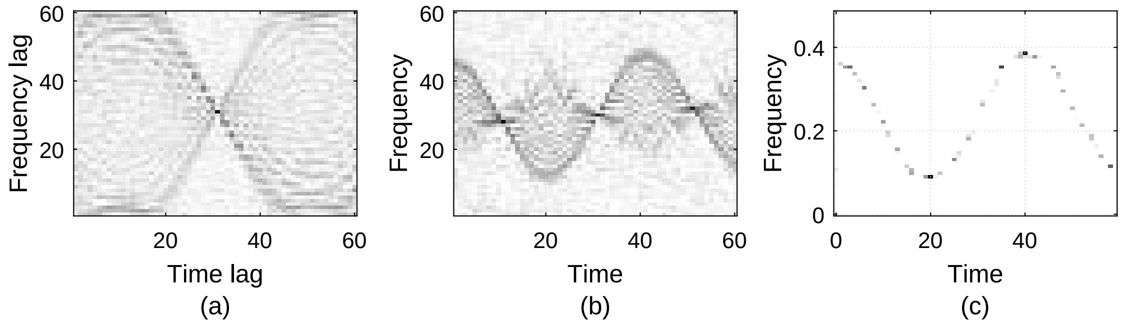


Figure 1: Representations of a sinusoidally-modulated signal in the time-frequency domain using: (a) the Wigner distributions, (b) the ambiguity function, (c) the resulting sparse time-frequency representation.

distribution (Figure 1(b)) the sparse representation achieved using compressive sensing approach offered a number of distinct advantages. First, it reduced the noise influence on the time-frequency representation, that is, it almost removed it completely from the distribution. Second, a number of non-zero terms in the sparse representation is about 50, which represents slightly over one percent of the total number of points in the time-frequency domain. Hence, by compressively sensing the ambiguity representation of the signal, not only did we manage to compress the signal representation, but we also de-noised it.

An inherent issue with this approach is one requires the representation of a signal in the ambiguity domain. Hence, any hardware implementation of this approach is quite complex. Furthermore, this approach is applicable only to quadratic signal representations, which limits our further signal manipulations (e.g., obtaining its time domain samples).

3.2. Compressive sensing based on time-frequency dictionary and matching pursuit

A more widely adopted approach is to utilize time-frequency dictionaries and obtain sparse signal representations using eqns. (2) and (3). Needless to say, this approach has very high computational costs which hindered its applications for many years. However, matching pursuit [82] or its variations (e.g., [83]) are very useful to avoid computational burdens associated with traditional compressive sensing approaches. Here, it is important to mention that a wide selection of time-frequency dictionaries exists, such as Gabor frames, curvelet frames, wavelet frames, overcomplete Fourier dictionaries or any combinations of these dictionaries [70], [84], [85], [86], [87], [88], [89], [90], [91].

Compressive sensing based implementation of a typical matching pursuit approach starts with an initial approximation of the signal, $\widehat{x}^{(0)}(m) = 0$, and the residual, $R^{(0)}(m) = x(m)$. Here m represent the M time indices that are uniformly or non-uniformly distributed, that is, M time indices compressively acquired. At each subsequent stage, the matching pursuit algorithm identifies a dictionary atom with the strongest contribution to the residual and adds it to the current approximation:

$$\widehat{x}^{(k)}(m) = \widehat{x}^{(k-1)}(m) + \alpha_k \phi_k(m) \tag{11}$$

$$R^{(k)}(m) = x(m) - \widehat{x}^{(k)}(m) \tag{12}$$

where $\alpha_k = \langle R^{(k-1)}(m), \phi_k(m) \rangle / \left\| \phi_k(m) \right\|^2$. The process continues till the norm of the residual $R^{(k)}(m)$ does not exceed required margin of error $\varepsilon > 0$: $||R^{(k)}(m)|| \leq \varepsilon$ [82], or a number of bases, $\mathfrak{n}_{\mathfrak{B}}$, needed for signal approximation should satisfy $\mathfrak{n}_{\mathfrak{B}} \leq \mathcal{K}$. Lastly, an approximation of a compressively sampled signal is obtained using L bases as

$$x(n) = \sum_{l=1}^{L} \langle x(m), \phi_l(m) \rangle \phi_l(n) + R^{(L)}(n) \tag{13}$$

where ϕ_l are L bases from the dictionary with the strongest contributions. L bases used in the signal approximation are obtained regardless of the implemented stopping criterion.

The approach based on time-frequency dictionaries is suitable for any post-processing of compressively sampled signals. For example, we can easily obtain any time-frequency representation of a signal using this L-bases based approximation:

$$\mathcal{TF}\{x(n)\} = \sum_{l=1}^{L} \langle x(m), \phi_l(m) \rangle \mathcal{TF}\{\phi_l(n)\} \tag{14}$$

where $\mathcal{TF}\{\}$ is a time-frequency operator (e.g., the S-transform or short-time Fourier transform) [1], [92].

As mentioned in the previous paragraphs, M samples can be acquired in uniform or nonuniform manners and the exact time values are needed to acquire proper values of the time-frequency dictionary. Nevertheless, many real-life conditions may prevent us from acquiring such exact times, and there is a need to estimate the sampling time instances. This is achiavable via annihilating filters contributions [38], [93], [94], which rely on determining the roots of an autoregressive filter in order to estimate the sampling instances.

3.2.1. A case study of a time-frequency dictionary for compressive sensing: Modulated discrete prolate spheroidal sequences

Discrete prolate spheroidal sequences were proposed by Slepian in 1978 [95]. For N samples and a normalized half-bandwidth value, W, a discrete prolate spheroidal sequence, $v_k(n, N, W)$, is defined as the real solution of[95]:

$$\sum_{m=0}^{N-1} \frac{\sin[2\pi W(n-m)]}{\pi(n-m)} v_k(m, N, W) = \lambda_k(N, W) v_k(n, N, W) \quad k = 0, 1, ..., N-1 \tag{15}$$

with $0 < W < 0.5$, and $\lambda_k(N, W)$ being non-zero eigenvalues of (15). The amplitude of these eigenvalues can be also approximated for fixed k and large N as

$$1 - \lambda_k(N, W) \sim \frac{\sqrt{\pi}}{k!} 2^{\frac{14k+9}{4}} \alpha^{\frac{2k+1}{4}} [2-\alpha]^{-(k+0.5)} N^{k+0.5} e^{-\gamma N} \tag{16}$$

where $\alpha = 1 - \cos(2\pi W)$ and $\gamma = \log\left[1 + \frac{2\sqrt{\alpha}}{\sqrt{2}-\sqrt{\alpha}}\right]$. It can be shown that the first $2NW$ eigenvalues are very close to 1 while the rest rapidly decays to zero [95]. These eigenvalues are also the eigenvalues of an $N \times N$ matrix $C(m, n)$ defined as [95]:

$$C(m, n) = \frac{\sin[2\pi W(n-m)]}{\pi(n-m)} \quad m, n = 0, 1, ..., N-1. \tag{17}$$

By time-limiting a discrete prolate spheroidal sequence, $v_k(n, N, W)$, we can obtain an eigenvector of $C(m, n)$. The discrete prolate spheroidal sequences are doubly orthogonal, that is, they are orthogonal on the infinite set $\{-\infty, ..., \infty\}$ and orthonormal on the finite set $\{0, 1, ..., N-1\}$.

In recent years, discrete prolate spheroidal sequences were used to obtain sparse signal representations especially in the cases when these sequences and an analyzed signal are in the same frequency band [96], [97]. Nevertheless, when the sequences and the signal are not aligned in the frequency domain, a larger number of discrete prolate spheroidal sequences is needed to obtain an accurate approximation and resulting approximations are often not sparse. To avoid this issue with discrete prolate spheroidal sequences, modulated discrete prolate spheroidal sequences were proposed in [96], which are defined as:

$$M_k(N, W, \omega_m; n) = \exp(j\omega_m n) v_k(N, W; n) \tag{18}$$

where $\omega_m = 2\pi f_m$ is a modulating frequency. Modulated discrete prolate spheroidal sequences are also doubly orthogonal, have most of properties of original discrete prolate spheroidal sequences and are bandlimited to the frequency band $[-W + \omega_m : W + \omega_m]$ [96].

Choosing a proper modulation frequency ω_m requires some a priori knowledge, or a guess. The simplest case is when a signal is confined to a known band $[\omega_1; \omega_2]$, then the modulating frequency, ω_m, and the bandwidth of the modulated discrete prolate spheroidal sequences are given by

$$\omega_m = \frac{\omega_1 + \omega_2}{2} \tag{19}$$

$$W = \left| \frac{\omega_2 - \omega_1}{2} \right| \tag{20}$$

as long as both satisfy:

$$|\omega_m| + W < \frac{1}{2}. \tag{21}$$

Nevertheless, the exact frequency band is not known in many practical applications. In general, we will only have details about a relatively wide frequency band. To account for many different possibilities, we proposed to construct a time-frequency dictionary containing bases which reflect various bandwidth scenarios [96]. This approach was used in a number of recent compressive sensing contributions [98], [92], [99], [100].

A sample case, depicted in Figure 2(a), involves a signal consisting of four basis functions from a 25-band dictionary based on modulated discrete prolate spheroidal functions with the normalized half-bandwidth equal to $W = 0.495$ and $N = 256$. For both uniform and non-uniform sampling, only 42 samples were needed to accurately recover the signal (less than 17% of the total number of samples) and its spectrograms based on regular and irregular sample times as shown in 2(c) and (d). A greater percentage of samples was required for this case in comparison to the first case as the second case is recovered almost exactly.

3.3. Compressive sensing in the time-frequency domain using short-time Fourier transform sparsity

Many of the signals appearing in real applications have a sparse representation in the Fourier transform domain, but also in the short-time Fourier transform domain [101], [102]. However, when the signal is affected by the noise, the number of non-zero components significantly increases thus ruining the sparsity, as shown in Figure 3(a) (its sorted

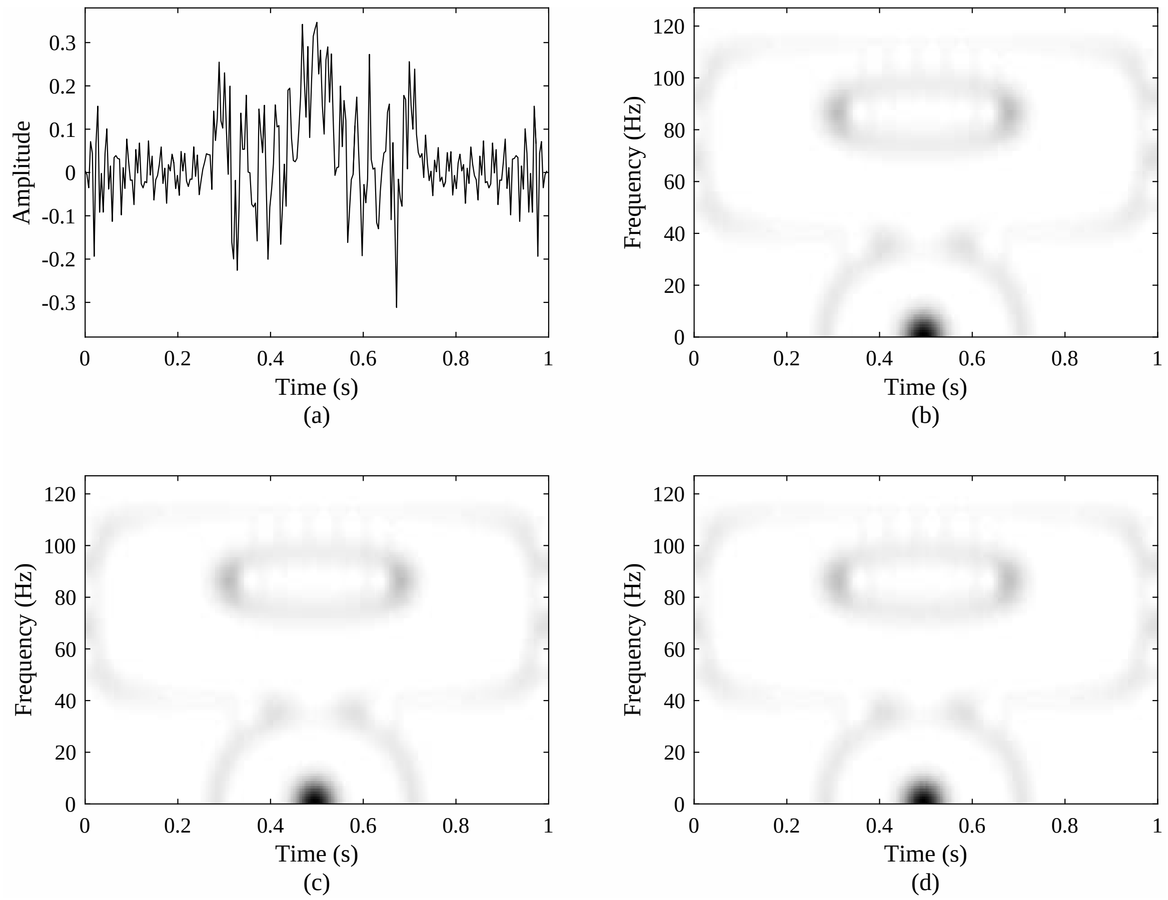


Figure 2: The time domain representation of the consider signal is shown in (a). Spectrograms of: (b) the original signal; (c) the signal based on equal distance samples; (d) the signal based on irregular samples.

values are shown in Figure 3(b)). By applying the L-estimation over columns of the short-time Fourier transform matrix, we may discard most of the unwanted coefficients from the short-time Fourier transform domain [103], [32]. However, many useful coefficients are also discarded in this process and we need to recover them using the compressive sensing approach. In the matrix form the short-time Fourier transform vector calculated at the time instant n using a rectangular window of size M can be defined as follow:

$$\begin{bmatrix} STFT(n,0) \\ STFT(n,1) \\ \vdots \\ STFT(n,M-1) \end{bmatrix} = \begin{bmatrix} \Psi(0,0) & \dots & \Psi(0,M-1) \\ \Psi(1,0) & \dots & \Psi(1,M-1) \\ \vdots & \dots & \vdots \\ \Psi(M-1,0) & \dots & \Psi(M-1,M-1) \end{bmatrix} \begin{bmatrix} x(n) \\ x(n+1) \\ \vdots \\ x(n+M-1) \end{bmatrix} \tag{22}$$

or in a more compact form:

$$\mathbf{STFT_M}(n) = \mathbf{\Psi}_M \mathbf{x}_M(n) \tag{23}$$

where $\mathbf{\Psi}_M$ is the discrete Fourier transform matrix of size $M \times M$:

$$\Psi(m,k) = \exp\left(-\frac{2\pi km}{M}\right) \; m = 0,1,...,M-1, k = 0,1,...,M-1 \tag{24}$$

The index M denotes the size of corresponding vectors. For the sake of simplicity, let us assume the non-overlapping windows, meaning that the short-time Fourier transform is calculated for time instants n taken with the step $M : \{0, M, 2M, ..., N-M\}$. The short-time Fourier transform calculation results in a set of short-time Fourier transform vectors: $\mathbf{STFT_M}(0), \mathbf{STFT_M}(M), ..., \mathbf{STFT_M}(N-M)$. Therefore, the STFT for all considered time instants $n \in \{0, M, 2M, ..., N-M\}$ is defined as follows [32]:

$$\begin{bmatrix} \mathbf{STFT_M}(0) \\ \mathbf{STFT_M}(M) \\ \vdots \\ \mathbf{STFT_M}(N-M) \end{bmatrix} = \begin{bmatrix} \mathbf{\Psi}_M & \dots & 0_M \\ 0_M & \dots & 0_M \\ \vdots & \dots & \vdots \\ 0_M & \dots & \mathbf{\Psi}_M \end{bmatrix} \begin{bmatrix} \mathbf{x}_M(0) \\ \mathbf{x}_M(M) \\ \vdots \\ \mathbf{x}_M(N-M) \end{bmatrix} \tag{25}$$

or equivalently: $STFT = \Theta\mathbf{x}$. The signal vector $\mathbf{x}$ consists of N samples and can be expressed using a sparse vector $\mathbf{X}$ of DFT coefficients:

$$\mathbf{x} = \mathbf{\Psi}_N^{-1}\mathbf{X} \tag{26}$$

where $\mathbf{\Psi}_N^{-1}$ is the N-point inverse Fourier transform. Hence, the relationship between the short-time Fourier transform and discrete Fourier transform vectors can be written as follows:

$$\mathbf{STFT} = \left(\mathbf{\Theta}\mathbf{\Psi}_N^{-1}\right)\mathbf{X} = \mathbf{AX}. \tag{27}$$

Now, let assume that the short-time Fourier transform is subject to L-estimation based filtering. After sorting the values of **STFT** and discarding a certain percent of the highest and the lowest components, we are left with missing data in the short-time Fourier transform domain. On the positions of discarded components, the zero values remain (Figure 3(c)). Hence, a compressive sensing problem can be observed in the short-time Fourier transform domain:

$$\mathbf{y} = \mathbf{A}_{cs}\mathbf{X}. \tag{28}$$

where

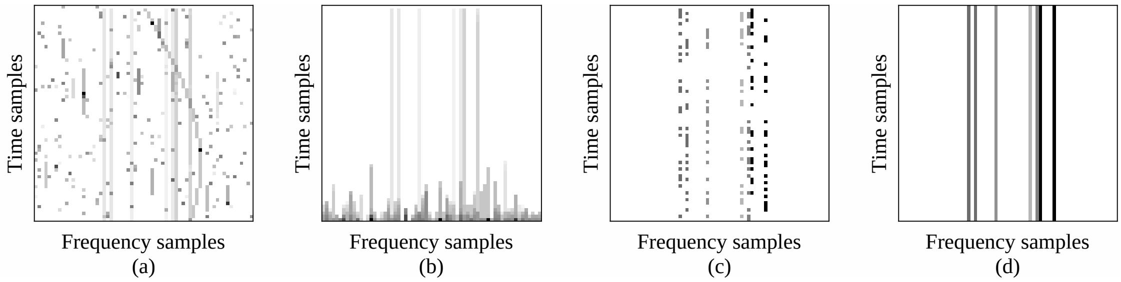


Figure 3: Compressive sensing and short-time Fourier transform: (a) noisy short-time Fourier transform; (b) sorted values of the noisy short-time Fourier transform; (c) available samples in the short-time Fourier transform after L-estimation; (d) reconstructed stationary components.

- $\mathbf{y}$ is a vector of N_a available values from **STFT** (i.e., nonzero values);
- $\mathbf{X}$ is a sparse discrete Fourier transform vector;
- $\mathbf{A}_{cs}$ is obtained from $\left(\mathbf{\Theta}\mathbf{\Psi}_N^{-1}\right)$ after removing the rows on the positions of missing samples in **STFT**.

In order to reconstruct **STFT**, the minimization problem is observed in the form:

$$\min \|\mathbf{X}\|_1 \text{ s.t. } \mathbf{y} = \mathbf{A}_{cs}\mathbf{X}. \tag{29}$$

The reconstructed stationary components in the short-time Fourier transform domain are shown in Figure 3(d).

The amplitudes of the reconstructed components in the DFT domain corresponds to the original signal components: [2, 2, 1.5, 1, 1,7, 3.5, 3.5]. The proposed method is compared with the results produced by an ideal case of notch filter (its inverse form), where we need to assume that all signal frequencies are known. The ideal notch filter response will pick all values along the considered frequencies, meaning that will pick also the noise producing wrong amplitudes of certain signal components as follows. The recovered amplitudes of components are as follows: [2, 2, 1.5, 1.14, 2.17, 3.6, 3.76] as shown in Figure 4.

Finally, let us consider a real-world radar signal. The radar signal consist of five rigid body components (stationary components) and three corner reflectors rotating at 60 RPM (nonstationary components). The stationary and non-stationary components intersects in both time and frequency dimensions. The short-time Fourier transformation

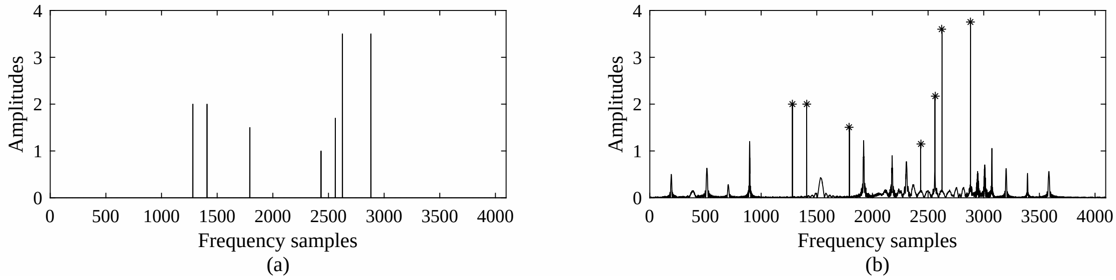


Figure 4: Comparing time-frequency based compressive sensing approaches and a notch filter: (a) reconstructed frequency components obtained using the proposed approach; (b) noisy components that would be selected by an ideal notch filter are marked by '*'.

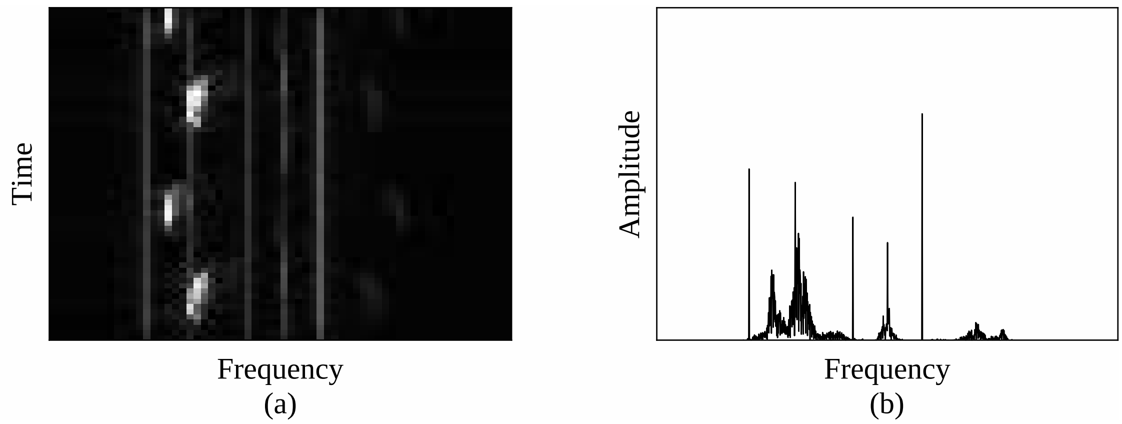


Figure 5: A sample radar signal: (a) the short-time Fourier transform of the observed signal; (b) the corresponding Fourier transform.

of the observed signal is shown in Figure 5(a), while the Fourier transform of the signal is shown in Figure 5(b).

The goal is to separate the micro-Doppler and rigid body components. As previously described, the short-time Fourier transform is calculated using non-overlapping windows (Figure 6(a)). The values in the short-time Fourier transform matrix are then sorted and 50% of lowest values are discarded. Namely, the micro-Doppler components appears to be smaller valued than the rigid body components in the sorted time-frequency signal representation because of their shorter duration. The remaining short-time Fourier transform values are shown in Figure 6(b) and represent the compressive sensing measurements that are subject to the compressive sensing reconstruction procedure. The reconstructed stationary rigid body components are sown in Figure 6(c), while the remaining micro-Doppler components are shown in Figure 6(d).

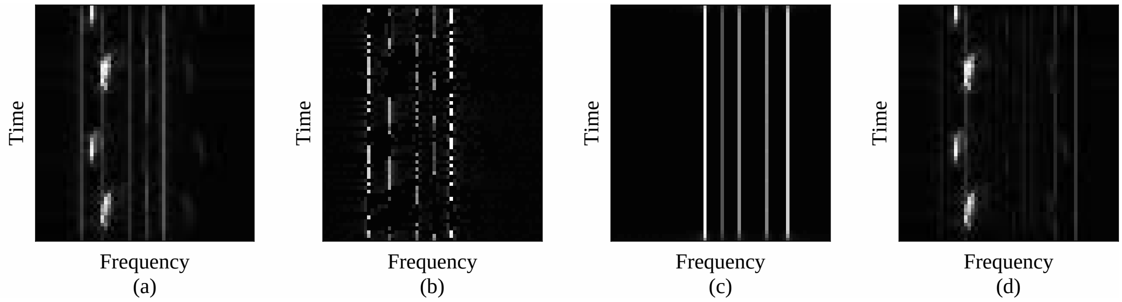


Figure 6: Compressive sensing and the short-time Fourier transform: (a) the short-time Fourier transform of the signal; (b) available samples in the short-time Fourier transform after L-estimation; (c) extracted rigid body components; (d) extracted micro-Doppler components.

3.4. Compressive sensing of signals based on the polynomial Fourier transform

The sparse representation of polynomial phase signals can be achieved by applying the polynomial Fourier transform [104]. The polynomial Fourier transform of signals $s(n)$ can be defined as follows:

$$X(k_1, ..., k_L) = \sum_{n=0}^{N-1} s(n) \exp\left\{-j\frac{2\pi}{N}\left(\frac{n^2k_2}{2} + ... + \frac{n^Lk_L}{L!}\right) - j\frac{2\pi}{N}nk_1\right\} \tag{30}$$

where the polynomial coefficients are assumed to be bounded integers. If $s(n)$ is a mono-component polynomial phase signal of the form:

$$s(n) = A\exp\left\{j\frac{2\pi}{N}\left(n\varphi_1\frac{n^2\varphi_2}{2} + ... + \frac{n^L\varphi_L}{L!}\right)\right\} \tag{31}$$

and if a set of polynomial Fourier transform coefficients $(k_2, k_3, ..., k_L)$ match the signal phase parameters $(\varphi_2, \varphi_3, ..., \varphi_L)$:

$$k_2 = \varphi_2, k_3 = \varphi_3, ..., k_L = \varphi_L \tag{32}$$

then we will obtain the sinusoid in the polynomial Fourier transform domain at the position $k_1 = \varphi_1$. Otherwise, the polynomial Fourier transform of $s(n)$ is not sparse. In that sense, the polynomial Fourier transform transform can be observed as the discrete Fourier transform of $s(n)$ demodulated by the exponential term $d(n)$:

$$X(k_1, ..., k_L) = \sum_{n=0}^{N-1} s(n)d(n) \exp\left\{-j\frac{2\pi}{N}nk_1\right\} \tag{33}$$

where

$$d(n) = \exp\left\{-j\frac{2\pi}{N}\left(\frac{n^2k_2}{2} + ... + \frac{n^Lk_L}{L!}\right)\right\} \tag{34}$$

The situation becomes more complex when $s(n)$ is a K-component polynomial phase signal:

$$s_K(n) = \sum_{i=1}^{K} A_i \exp\left\{ j\frac{2\pi}{N}\left(na_{1i} + \frac{n^2 a_{2i}}{2} + ... + \frac{n^L a_{Li}}{L!} \right) \right\}. \tag{35}$$

The coefficients of demodulation term $d(n)$ should be then chosen to correspond to one of the components:

$$k_2 = \varphi_{i2}, k_3 = \varphi_{i3}, ..., k_L = \varphi_{iL}. \tag{36}$$

As a result, the ith signal component is demodulated and becomes a sinusoid in the polynomial Fourier transform domain [105]. The polynomial Fourier transform representation is not strictly sparse as in the case of a single polynomial phase signal, but the ith component will be dominant in the polynomial Fourier transform spectrum. In that sense, we might say that if $k_2 = \varphi_{i2}, k_3 = \varphi_{i3}, ..., k_L = \varphi_{iL}$ is satisfied, then the polynomial Fourier transform is compressible with the dominant ith component. Note that the sparsity (compressibility) in the polynomial Fourier transform domain is observed with respect to the single demodulated component. Thus, we need to change the values of polynomial Fourier transform coefficients k_2, k_3, k_L within a certain range $[k_{min}, k_{max}]$ until we obtain a dominant component in the polynomial Fourier transform domain, which means that we revealed one of the K signal components: $k_2 = \varphi_{i2}, k_3 = \varphi_{i3}, ..., k_L = \varphi_{iL}$, $i \in [1, K]$.

In the compressive sensing context, the signal vector $\mathbf{s}$ is randomly undersampled having only $N_a \ll N$ available samples. It means that the demodulation vector $\mathbf{d}$ should be also calculated only for N_a available instants. Now, the measurement vector $\mathbf{y}$ can be defined as follows [105]:

$$\mathbf{y} = \mathbf{s}(n_a)\mathbf{d}(n_a) = \mathbf{x}(n_a) \tag{37}$$

where n_a denotes available sample positions.

The vector form of the polynomial Fourier transform definition given by eqn. (33) is given by:

$$\mathbf{X} = \mathbf{\Psi}_N \mathbf{x} \tag{38}$$

where $\mathbf{\Psi}_N^N$ is a $N \times N$ discrete Fourier transform matrix. From eqns. (37) and (38), we may write:

$$\mathbf{y} = \mathbf{\Psi_{N_a}^{-1}X} \tag{39}$$

Algorithm 1 Calculate compressive sensing based polynomial Fourier transform

Require: $n_a > 0$

for $k_j = k_{min} : \text{step} : k_{max},\ j = 2, ..., L$ **do**

$\mathbf{y} = \mathbf{s}(n_a)\mathbf{d}(n_a)$

$\mathbf{X} = \mathbf{\Psi}_{N_a}\mathbf{y}$

if $k_2 = \varphi_{i2}, k_3 = \varphi_{i3}, ..., k_L = \varphi_{iL} \Leftrightarrow$ there is a dominant sinusoid $\mathbf{X_i}$ **then**

$k_{1i} = \arg\max(\mathbf{X})$

Save $\mathbf{k_i} = (k_1, k_2, ..., k_L) = (\varphi_{1i}, \varphi_{2i}, ..., \varphi_{Li})$

end if

end for

where $\mathbf{\Psi_{N_a}^{-1}}$ is the partial random inverse Fourier matrix of size $N_a \times N$ obtained by omitting rows from inverse discrete Fourier transform matrix $\mathbf{\Psi_{N_a}^{-1}}$ that correspond to the unavailable samples. If the demodulation term is chosen such that $k_2 = \varphi_{i2}, k_3 = \varphi_{i3}, ..., k_L = \varphi_{iL}$ then $\mathbf{X}$ can be observed as a demodulated version of the ith signal component $\mathbf{X}_i$, having the dominant ith component in the spectrum with the support k_{1i}. Other components in spectrum are much lower than $\mathbf{X}_i$ and could be observed as noise. The minimization problem can be written in the form:

$$\min \|\mathbf{X}_i\|_1 \ \text{ subject to } \ \left\|\mathbf{y}\mathbf{\Psi_{N_a}^{-1}}\mathbf{X}\right\|_2 < \xi. \tag{40}$$

The procedure can be described in the form of pseudo code as shown in Algorithm 1. Hence, as a result of this phase we have identified the sets of signal phase parameters: $\mathbf{k_i} = (k_1, k_2, ..., k_L) = (\varphi_{1i}, \varphi_{2i}, ..., \varphi_{Li})$.

Next, we need to recover the exact components amplitudes. Denote the set of available signal samples positions by $\mathbf{n}_a = (n_1, n_2, , n_{N_a})$. In order to calculate the exact amplitudes $A_1, A_2, ..., A_K$ of K signal components, we observe the set of equations in the form:

$$\begin{bmatrix} s(n_1) \\ s(n_2) \\ \vdots \\ s(n_{N_a}) \end{bmatrix} = \begin{bmatrix} \Phi(1,1) & \dots & \Phi(1,K) \\ \Phi(2,1) & \dots & \Phi(2,K) \\ \vdots & \dots & \vdots \\ \Phi(N_a,1) & \dots & \Phi(N_a,K) \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_K \end{bmatrix} \tag{41}$$

where

$$\Phi(j,i) = \exp\left\{ j\frac{2\pi}{N}\left(n_j k_{1i} + \dots n_j^L k_{L_i}\right)\right\} \ j = 1, ..., N_a; i = 1, ..., K. \tag{42}$$

In other words we have another system of equations given by:

$$\mathbf{s}(n_a) = \boldsymbol{\Phi}\alpha \tag{43}$$

where $\alpha = [A_1, ..., A_K]^T$ contains the desired K signal amplitudes. The rows of Φ correspond to positions of measurements $n_1, n_2, ..., n_{N_a}$, and columns correspond to the K components with phase parameters $(k_1, k_2, ..., k_L) = (\varphi_{1i}, \varphi_{2i}, ..., \varphi_{Li})$, for $i = 1, ..., K$. The solution of the observed problem can be obtained in the least square sense as:

$$\alpha = (\boldsymbol{\Phi}^{*}\boldsymbol{\Phi})^{-1}\, \boldsymbol{\Phi}^{*}\sim(n_a) \tag{44}$$

Let us consider a polynomial phase signal in the form that consists of three chirp components:

$$\begin{aligned} s(t) &= s_1(t) + s_2(t) + s_3(t) \\ &= \exp\left(-j\pi\varphi_{21}t^2 + j\pi\varphi_{11}t\right) + \exp\left(-j\pi\varphi_{22}t^2 + j\pi\varphi_{12}t\right) + \exp\left(-j\pi\varphi_{23}t^2 + j\pi\varphi_{13}t\right) \\ &= \exp\left(-j\pi 8Tt^2 + j\pi 16Tt\right) + \exp\left(-j\pi 32Tt^2 + j\pi 16Tt\right) + \exp\left(-j\pi 8Tt^2 + j\pi 16Tt\right) \end{aligned}$$

where the signal parameters are given as: $t = [-1/2, 1/2)$ with step $\Delta t = 1/1024$, $T = 32$, and the total signal length is 1024 samples. Observe that there are three chirp components with the rates: $\varphi_{21} - 8T$, $\varphi_{22} = 32T$, $\varphi_{23} = 8T$. The demodulation term is given in the form $d(t) = \exp(j2\pi k_2 Tt^2)$. We need to search for parameter k_2. Thus, we change the values of parameter k_2 within a predefined range to match components one after the other. The discrete Fourier transform spectrum of full length signal before applying the demodulation term (as a part of the polynomial Fourier transform) is shown in Figure 7(a). When $k_2 = -8T$, the first component is detected and it becomes dominant in the spectrum (Figure 7(b)). The same situation appears when $k_2 = 32T$ (Figure 7(c)) and $k_2 = 8T$ (Figure 7(d)).

We may observe that the polynomial Fourier transform with an appropriate demodulation term $d(t)$ can be considered as compressible for the dominant component matched by the demodulation term. As such, it is amenable for compressive sensing reconstruction.

Next, we consider a small number of randomly selected available samples, i.e., 128 out of 1024 are available (12.5% of the total signal length). The missing samples within $s(t)$ are considered as zero values, and then the demodulation term is applied iteratively for

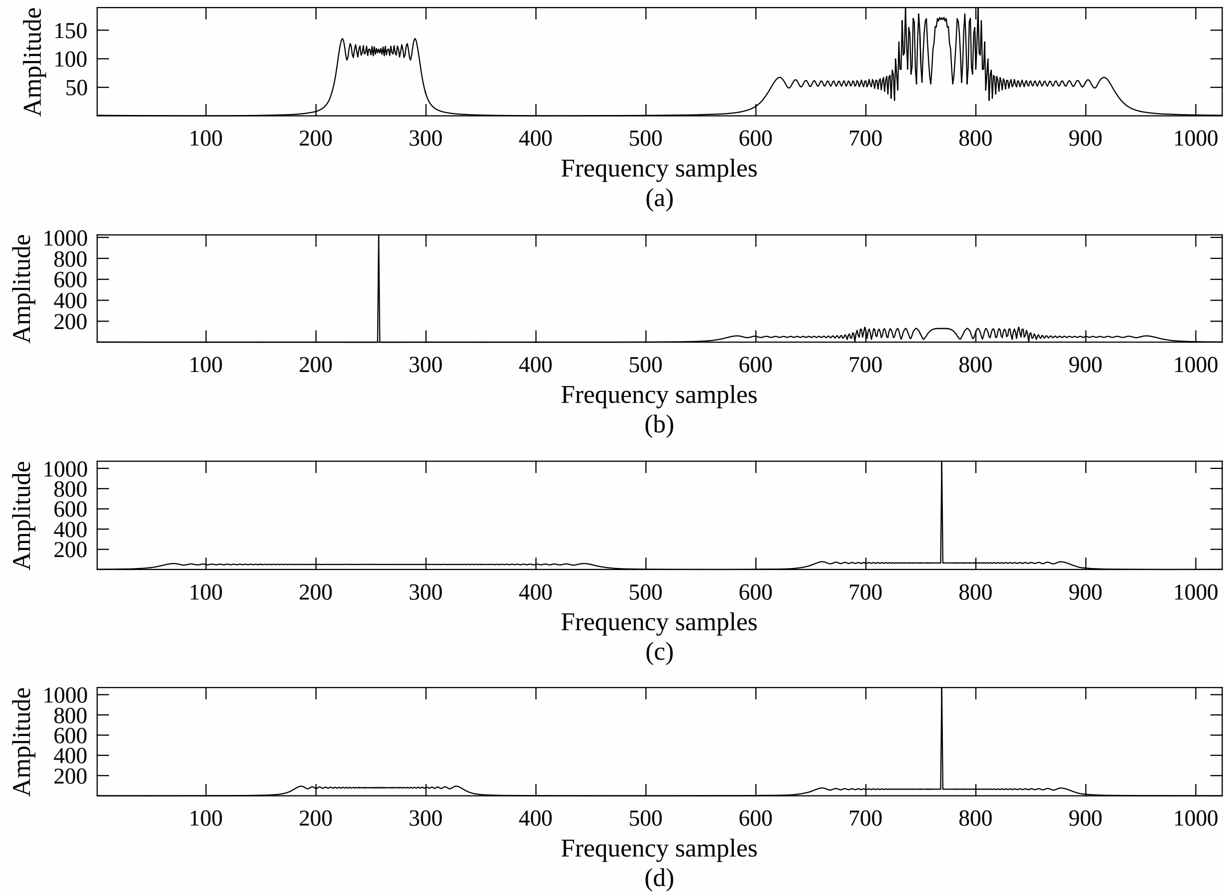


Figure 7: The effects of demodulation on the discrete Fourier transform spectrum: (a) the discrete Fourier transform of $s(t)$ before demodulation; (b) demodulation with $k_2 = -8T$; (c) demodulation with $k_2 = 32T$; (d) demodulation with $k_2 = 8T$.

a range of values k2. The results for the PFT when k_2 matches $\varphi_{21} = -8T$, $\varphi_{22} = 32T$ and $\varphi_{23} = 8T$ are given in Figures 8(a)-(c). Note that noise appears as a consequence of missing samples that are set to zero value in order to calculate the initial polynomial Fourier transform. However, the demodulated components are prominent in the spectrum (Figure 8). For the illustration, Figure 9 depicts the initial polynomial Fourier transform when k_2 does not match any of φ_{21}, φ_{22} and φ_{23}. The spectrum is noisy with no dominant components revealed.

The compressive sensing reconstruction method needs to determine the support of components revealed after the appropriate demodulation as shown in Figure 8. In the same time, it should ignore the cases with inappropriate k_2. The simplest solution can be achieved using a threshold derived for the single iteration compressive sensing reconstruction algorithm, proposed in [106]. When the threshold (horizontal red line in Fig.

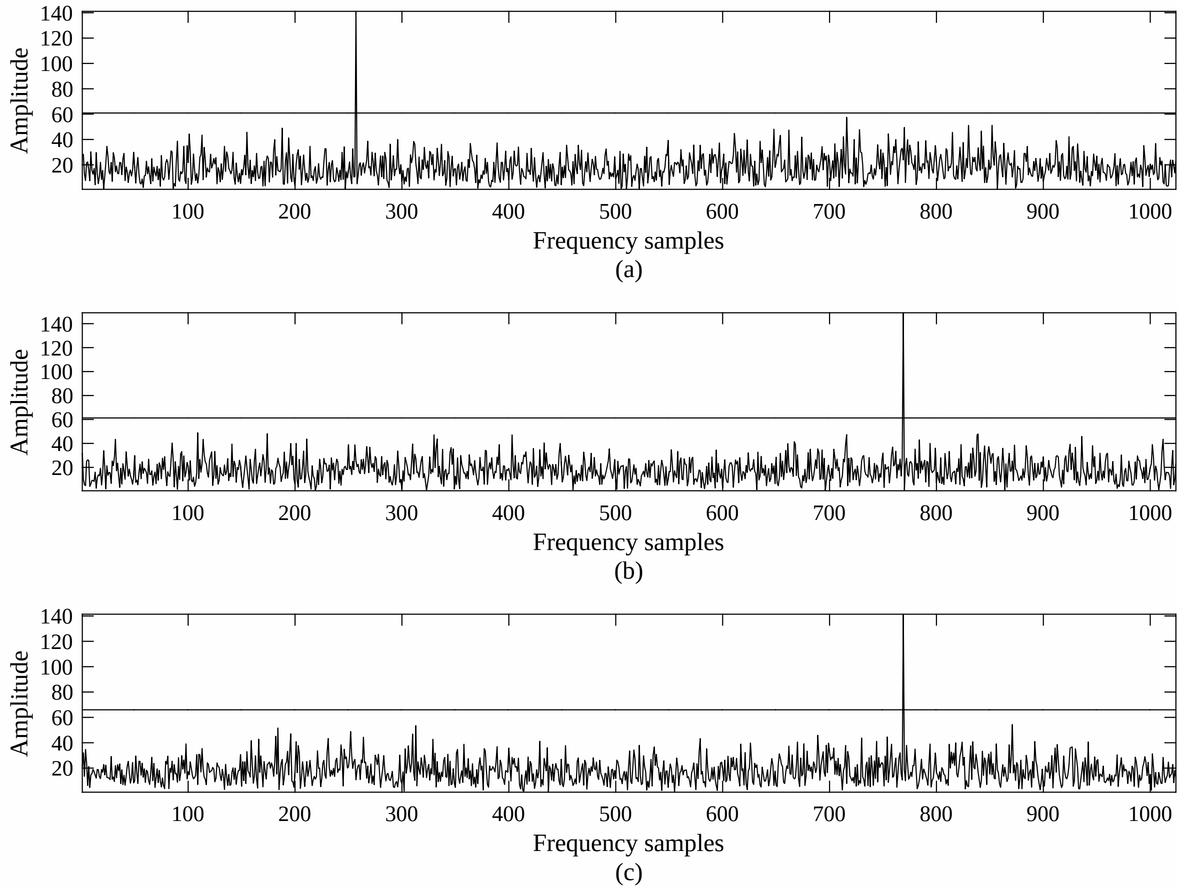


Figure 8: The polynomial Fourier transform for the compressive sensing case: (a) demodulation with $\varphi_{21} = -8T$; (b) demodulation with $\varphi_{22} = 32T$; (c) demodulation with $\varphi_{23} = 8T$.

3) is applied on the spectrum after demodulation with appropriate k_2 (k_2 matches either $\varphi_{21} = -8T$, $\varphi_{22} = 32T$ and $\varphi_{23} = 8T$), a support of demodulated component is returned as a result: $\varphi_{11} = -8T$, $\varphi_{12} = -16T$ and $\varphi_{13} = 16T$. Otherwise, when the threshold is applied to the spectrum after demodulation with inappropriate (wrong k2), the method returns no support (Figure 9).

4. Compressive sensing based time-frequency representations

Most of time-frequency distributions can be observed as the Fourier transforms of the local autocorrelation functions. In order to produce highly localized energy distributions, the autocorrelation functions must locally approximate a sinusoidal signal at each time sample. The overall instantaneous frequency characteristics are obtained based on the individual sinusoidal frequencies from different shifted signal windows [107], [108],

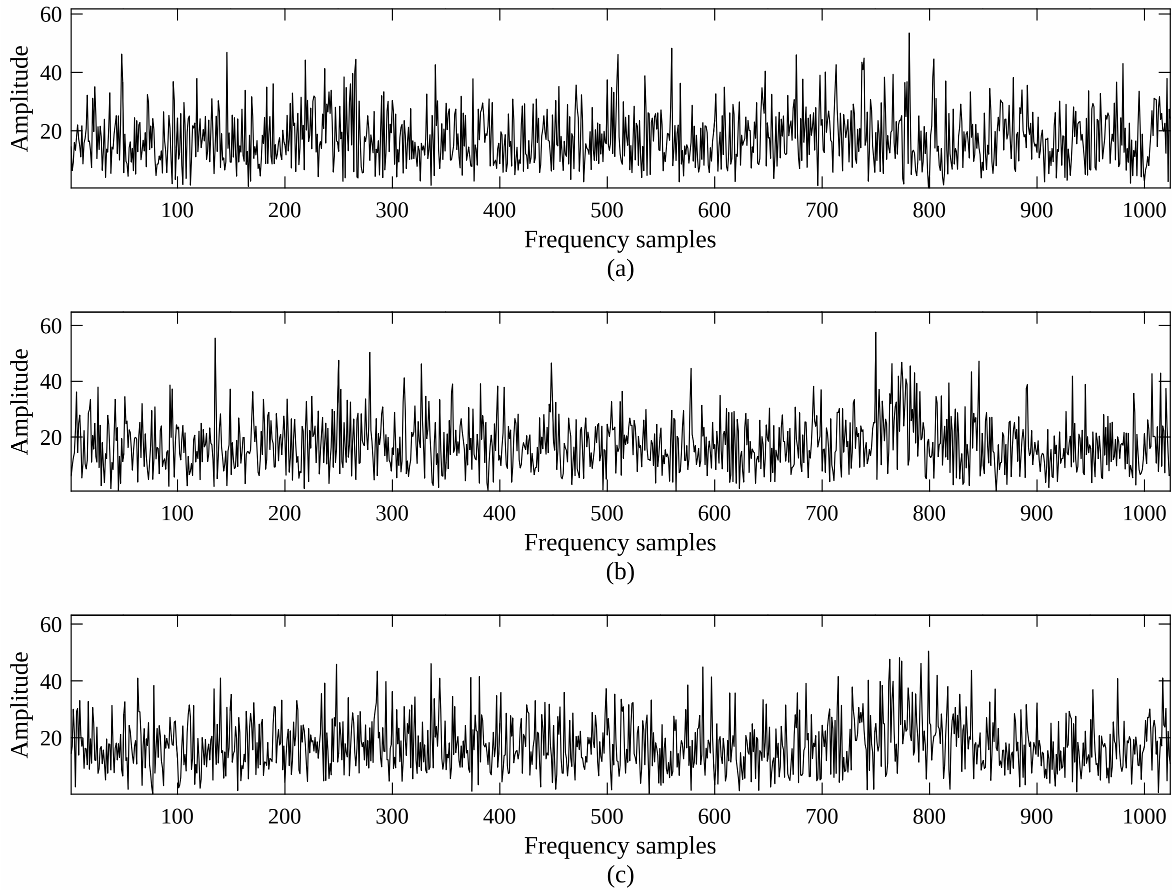


Figure 9: Compressive sensing of the polynomial Fourier transform with an incorrect demodulation term: (a) demodulation with $k_2 = 18T$; (b) demodulation with $k_2 = 12T$; (c) demodulation with $k_2 = 16T$

[109], [110]. In the context of compressive sensing, we observe the case when a signal is represented by a small set of random samples. Consequently, only a small percent of total autocorrelation function samples are considered as available for time-frequency distrubition calculations and instantenous frequency estimations. The standard form of time-frequency distributions calculated from coarsely under-sampled autocorrelation functions would be seriously degraded by noise, with amplitudes of components being much below their true values. As a solution, we can consider compressive sensing based time-frequency representations obtained by applying the compressive sensing reconstruction algorithm to autocorrelation functions, in lieu of the Fourier transform, to achieve an ideal time-frequency signal representations [111], [112].

We can start with a definition of time-frequency distributions defined as the Fourier transform of a higher order local autocorrelation function:

$$TFD(t,f) = \int_{-\infty}^{+\infty} R(t,\tau)e^{-j2\pi\tau} d\tau. \quad (45)$$

The general form of the local autocorrelation function can be defined as follows:

$$R(t,\tau) = \prod_{i=1}^{P/2} x^{b_i}(t + a_i\tau)x^{*b_i}(t - a_i\tau) \quad (46)$$

where P is an even number representing the order of a distribution, while the coefficients a_i and b_i depend on a particular time-frequency distribution. Without loss of generality, the rectangular window function is assumed. For instance,

1. The Wigner distribution is obtained when $P = 2$ and $a_1 = 1/2$, and $b_1 = 1$.
2. The L-Wigner distribution is obtained when $P = 2$, $a_1 = 1/(2L)$, and $b1 = L$ as it yields $R(t,\tau) = x^L\left(t + \frac{\tau}{2L}\right)x^{*L}\left(t - \frac{\tau}{2L}\right)$.
3. For $P = 4$ and $a_1 = 0.675$, $b_1 = 2$, $a_2 = -0.85$, $b2 = 1$, the auto-correlation function in given by $R(t,\tau) = x^2\left(t + 0.675\tau\right)x^{*2}\left(t - 0.675\tau\right)x\left(t - 0.85\tau\right)x^*\left(t + 0.85\tau\right)$, which leads to the polynomial distribution.

Now, assume that only a small number of M random samples from $R(t,\tau)$ are available in each windowed part where $M \ll N$ holds (N is the total number of samples within the window). For the sake of simplicity we may write the autocorrelation function in the form:

$$R(t,\tau) = x_1x_2\ldots x_p = \prod_{i=1}^{P} x_i \quad (47)$$

and

$$R(t,\tau) = \begin{cases} R(t,\tau_{n_m}) & m = 1,...,M \\ 0 & \text{otherwise} \end{cases} \quad (48)$$

where the discrete signal terms are denoted by vectors x_i, while τ_{n_m} for $m = 1,...,M$ are random positions of available samples in one lag-window. Hence, we have:

$$\|R(t,\tau)\|_{l_0} = \left\|\prod_{i=1}^{P} x_i\right\|_{l_0} = M. \quad (49)$$

The standard TFD calculated on the basis of $R(t,\tau)$ with missing samples, would be affected by the noise due to the missing samples. Namely, the missing samples needs to be considered as zero values in $R(t,\tau)$ which will produce noise in the time-frequency domain. Hence, we consider the possibility to apply the concept of CS reconstruction in order to provide a noise-free time-frequency representation. If we observe the vector of autocorrelation samples for a single time instant tj denoted as $R(t,\tau)$, then the optimization problem can be defined as follows:

$$\min \|\mathbf{X}(t_j,\tau)\|_{l_1} \text{ subject to } R(t_j,\tau) = \mathbf{A}\mathbf{X}(t_j,f) \tag{50}$$

where for the observed t_j, $\mathbf{X}(t_j,f)$ represents a sparse vector belonging to an ideal time-frequency representation at the time instant t_j. The matrix $\mathbf{A}$ is the Fourier transform based compressive sensing matrix. The minimization problem can be solved using some of the known compressive sensing reconstruction algorithms, such as the orthogonal matching pursuit.

Let us consider an illustrative example in the form:

$$x_1(t) = \exp(j160\pi t^3 - j190\pi t) \tag{51}$$

The polynomial distribution of the fourth order is considered for a time-frequency representation of the observed signal. The amount of available samples $\mathbf{x}_1$ is 35% of the total number of window samples ($N = 128$ samples). The standard polynomial distribution calculated with zero values on the positions of missing samples in auto-correlation function is shown in Figure 10(a). In order to provide the compressive sensing based time-frequency representation, the orthogonal matching pursuit is applied to each windowed signal part in order to recover sparse spectrum corresponding to each particular time instant. It means that in the case of the considered monocomponent signal, a single frequency component is obtained for each time instant, resulting in an ideal time-frequency representation as shown in Figure 10(b).

Similarly, let us consider a signal with a fast varying instantenous frequency given in the form:

$$x_2(t) = \exp(j6\sin(2.4\pi t) + j3\cos(1.5\pi t) - j20\pi t^2) \tag{52}$$

As a suitable time-frequency distribution, the complex-time distribution is usually considered for fast varying instantaneous frequency estimations, since it provides significant

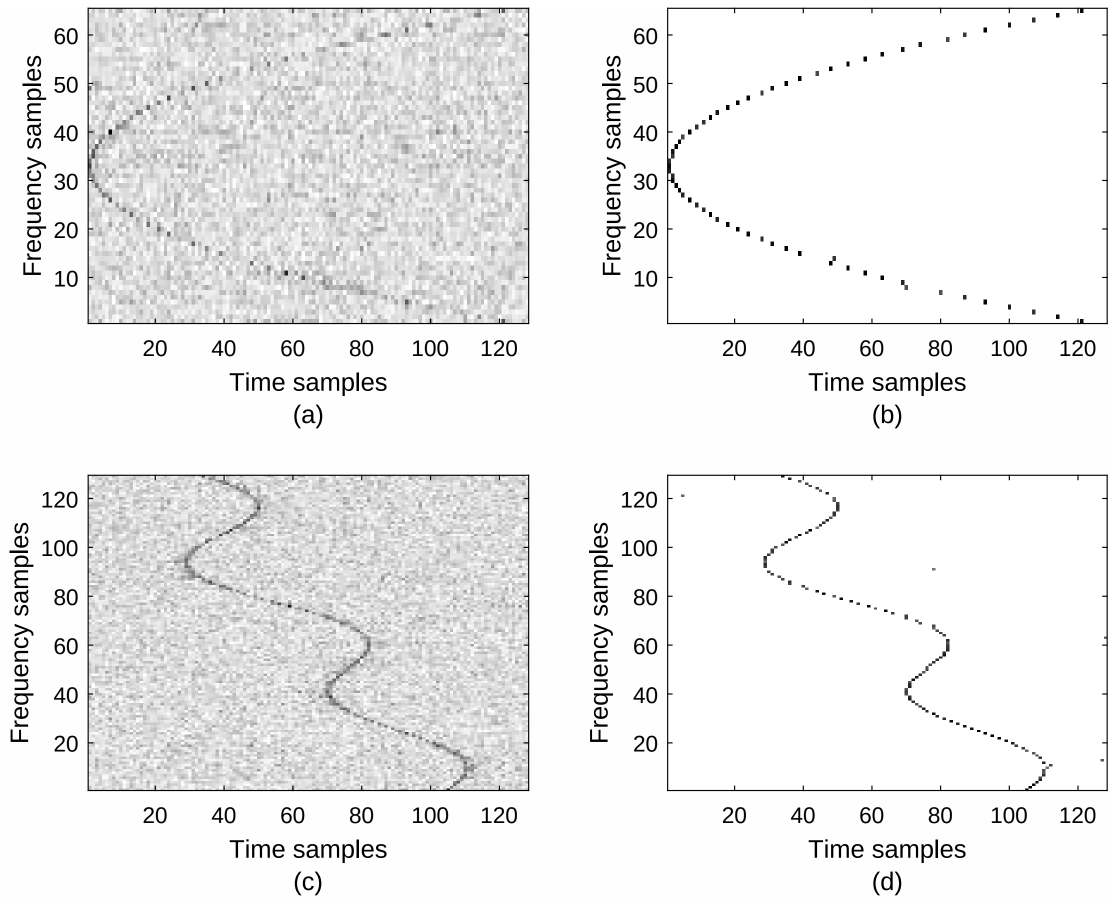
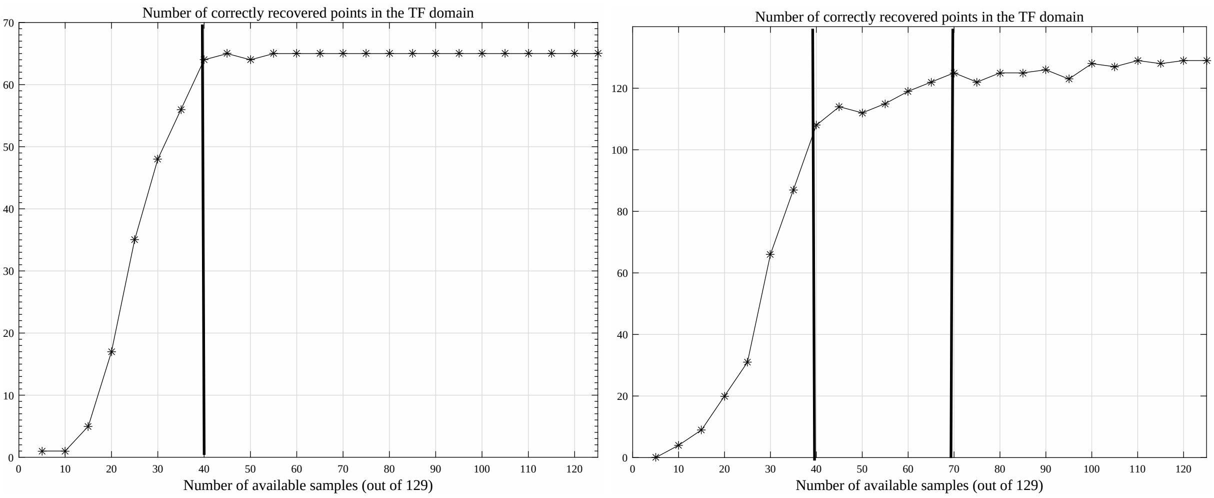


Figure 10: Comparison between the traditional time-frequency distributions and their compressive sensing based equivalents: (a) the traditional polynomial distribution $x_1(t)$, (b) a compressive sensing based polynomial distribution of $x_1(t)$ (c) the traditional complex-time distribution of $x_2(t)$, (d) a compressive sensing variant of the complex-time distribution of $x_2(t)$.

concentration improvements with respect to the quadratic but also polynomial distributions [113], [114]. A commonly used complex-lag distribution is defined for the autocorrelation function in the form:

$$R(t,\tau) = x\left(t+\frac{\tau}{4}\right)x^{-1}\left(t-\frac{\tau}{4}\right)x^{-j}\left(t+j\frac{\tau}{4}\right)x^{j}\left(t-j\frac{\tau}{4}\right) \tag{53}$$

Let us assume that for $\mathbf{x}_2$, there is approximately 40% of available samples. The standard form of the complex-time distributions and its improved compressive sensing version are provided in Figures 10(c) and 10(d). The compressive sensing based complex-time distribution provides almost an ideal representation for the instantaneous frequency estimation. Compared to the standard form, the compressive sensing based form is noiseless and highly compressible providing a set of enhanced peaks along the instantaneous fre-



(a) (b)

Figure 11: Estimated numbers of points to accurately estimate instantaneous frequency with (a) the polynomial distribution; (b) the complex-time distribution.

quency, while the other values in the tim-frequency plane are zeros.

Next, we estimated the required number of available samples that can provide an accurate instantaneous frequency representation after the reconstruction in the time-frequency domain. The results are shown in Figures 11a and 11b for the polynomial distribution and the complex-time distribution, respectively. The assumed signal length is 129 samples. In the case of the polynomial distribution, a high level of precision (100% of points are exactly reconstructed) is achieved with 30% of available samples (40 samples out of 129). For the complex-time distribution, we can observe that the acceptable precision is achieved even with 30% of available samples (app. 40 samples out of 129), while in the case when the amount of available samples is 55% or more (app. 70 samples or more), the accurate estimation is achieved.

While these illustrative examples depict that the major advantage of these approaches is the fact that we can obtain very accurate representations of non-stationary signals even a small number of samples, it needs to be pointed out that these compressive sensing based time-frequency approaches are computationally more expensive than traditional time-frequency approaches. The increased computational complexity is due to the optimization procedure implemented to recover missing samples.

5. Conclusions and future directions

In this review paper, we summarized recent advances regarding compressive sensing and time-frequency analysis. All these recent contributions demonstrate that compressive sensing provided a framework for sparse time-frequency processing of non-stationary signals. Based on the current contributions, we anticipate the following future directions:

- There is a great need to develop hardware solutions for all these signal processing schemes consider in this paper. Hardware developments are severely lagging the algorithmic development, which currently leaves many questions unanswered when it comes to practical applicability of these algorithms.

- Compressive sensing based time-frequency representations address a major issue associated with traditional time-frequency representations, that is, the ability to obtain a time-frequency representation of a signal using only a small number of random samples. However, the major disadvantage of these approaches is that they are much more computationally expansive than traditional time-frequency approaches. Hence, future research directions include the development of computationally inexpensive compressive sensing based time-frequency representations, that have the computational cost of the same order as traditional time-frequency methods.

- Adaptations of these new algorithms in many different areas is still an open questions. While there are applications that have highly redundant information and can tolerate errors (e.g., communication systems) for which compressive sensing provides excellent results, there are many more that require compressive sensing to provide perfect reconstruction every time (e.g., most of medical diagnostic applications). Hence, the time-frequency based compressive sensing approaches have the largest value in these applications requiring very high accuracies, as typical compressive sensing approaches based on random basis dictionaries are not suitable.

In conclusion, this paper provides a concise summary of the work in compressive sensing for sparse time-frequency precessing. While the framework provides very powerful tools to process sparse non-stationary signals, we strongly believe it is still in its early stages, and it is expected that further research and applications of the existing schemes will grow in the near future. In our companion paper [115], we describe MATLAB functions used to generate figures presented in this manuscript.

Acknowledgments

The work of E. Sejdić was partially funded by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number R01HD074819. The content is solely the responsibility of E. Sejdić and does not necessarily represent the official views of the National Institutes of Health.

References

[1] E. Sejdić, I. Djurović, and J. Jiang, "Time-frequency feature representation using energy concentration: An overview of recent advances," *Digital Signal Processing*, vol. 19, no. 1, pp. 153–183, Jan. 2009.

[2] A. N. Akansu and R. A. Haddad, *Multiresolution Signal Decomposition: Transforms, Subbands, and Wavelets.* San Diego: Academic Press, 2001.

[3] M. Akay, Ed., *Time-Frequency and Wavelets in Biomedical Signal Processing.* Piscataway, NJ: IEEE Press, 1998.

[4] R. E. Challis and R. I. Kitney, "Biomedical signal processing (in four parts) - part 1: Time-domain methods," *Medical and Biological Engineering and Computing*, vol. 28, no. 6, pp. 509–524, Nov. 1990.

[5] I. Orović, M. Orlandić, S. Stanković, and Z. Uskoković, "A virtual instrument for time-frequency analysis of signals with highly nonstationary instantaneous frequency," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 3, pp. 791–803, March 2011.

[6] B. Boashash, Ed., *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*, 2nd ed. Amsterdam: Elsevier, 2016.

[7] K. Gröchenig, *Foundations of Time-Frequency Analysis.* Boston: Birkhäuser, 2001.

[8] S. Stanković, I. Orović, and E. Sejdić, *Multimedia Signals and Systems: Basic and Advanced Algorithms for Signal Processing*, 2nd ed. New York, NY: Springer US, 2016.

[9] A. Cohen and J. Kovačević, "Wavelets: The mathematical background," *Proceedings of the IEEE*, vol. 84, no. 4, pp. 514–522, Apr. 1996.

[10] S. Stanković, I. Orović, and A. Krylov, "Video frames reconstruction based on time-frequency analysis and Hermite projection method," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, no. 1, p. 970105, 2010.

[11] F. Hlawatsch and G. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Magazine*, vol. 9, no. 2, pp. 21–67, Apr. 1992.

[12] X. Li, G. Bi, S. Stankovic, and A. M. Zoubir, "Local polynomial fourier transform: A review on recent developments and applications," *Signal Processing*, vol. 91, no. 6, pp. 1370–1393, 2011.

[13] S. Stanković and I. Orović, "Time-frequency rate distributions with complex-lag argument," *Electronics Letters*, vol. 46, no. 13, pp. 950–952, June 2010.

[14] P. J. Kootsookos and B. C. L. B. Boashash, "A unified approach to the STFT, TFDs, and instantaneous frequency," *IEEE Transactions on Signal Processing*, vol. 40, no. 8, pp. 1971–1982, Aug. 1992.

[15] S. Stanković, I. Orović, and V. Sucic, "Averaged multiple l-spectrogram for analysis of noisy nonstationary signals," *Signal Processing*, vol. 92, no. 12, pp. 3068–3074, 2012.

[16] S. Blanco, R. Q. Quiroga, O. A. Rosso, and S. Kochen, "Time-frequency analysis of electroencephalogram series," *Physical Review E*, vol. 51, no. 3, pp. 2624–2631, Mar. 1995.

[17] C. Li, C. Zheng, and C. Tai, "Detection of ECG characteristic points using wavelet

transforms," *IEEE Transactions on Biomedical Engineering*, vol. 42, no. 1, pp. 21–28, Jan. 1995.

[18] S. Blanco, R. Q. Quiroga, O. A. Rosso, and S. Kochen, "Time-frequency analysis of electroencephalogram series II. Gabor and wavelet transforms," *Physical Review E*, vol. 54, no. 6, pp. 6661–6672, Dec. 1996.

[19] P. M. Bentley, P. M. Grant, and J. T. E. McDonnell, "Time-frequency and time-scale techniques for the classification of native and bioprosthetic heart valve sounds," *IEEE Transactions on Biomedical Engineering*, vol. 45, no. 1, pp. 125–128, Jan. 1998.

[20] M. Stridh, L. Sörnmo, C. J. Meurling, and S. B. Olsson, "Sequential characterization of atrial tachyarrhythmias based on ECG time-frequency analysis," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 1, pp. 100–114, Jan. 2004.

[21] I. Orović, S. Stanković, T. Chau, C. M. Steele, and E. Sejdić, "Time-frequency analysis and Hermite projection method applied to swallowing accelerometry signals," *EURASIP Journal on Advances in Signal Processing*, pp. 323 125–1–7., 2010.

[22] E. Sejdić, K. A. Lowry, J. Bellanca, M. S. Redfern, and J. S. Brach, "A comprehensive assessment of gait accelerometry signals in time, frequency and time-frequency domains," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 3, pp. 603–612, May 2014.

[23] H. Bleton, S. Perera, and E. Sejdić, "Cognitive tasks and cerebral blood flow through anterior cerebral arteries: a study via functional transcranial doppler ultrasound recordings," *BMC Medical Imaging*, vol. 16, no. 1, p. 22, 2016.

[24] A. G. Rehorn, E. Sejdić, and J. Jiang, "Fault diagnosis in machine tools using selective regional correlation," *Mechanical Systems and Signal Processing*, vol. 20, no. 5, pp. 1221–1238, Jul. 2006.

[25] H. Oehlmann, D. Brie, M. Tomczak, and A. Richard, "A method for analysing gearbox faults using time-frequency representations," *Mechanical Systems and Signal Processing*, vol. 11, no. 4, pp. 529–545, Nov. 1997.

[26] W. J. Staszewski, K. Worden, and G. R. Tomlinson, “Time-frequency analysis in gearbox fault detection using the Wigner-Ville distribution and pattern recognition,” *Mechanical Systems and Signal Processing*, vol. 11, no. 5, pp. 673–692, Sep. 1997.

[27] J. Lin and L. Qu, “Feature extraction based on Morlet wavelet and its application for mechanical fault diagnosis,” *Journal of Sound and Vibration*, vol. 234, no. 1, pp. 135–148, Jun. 2000.

[28] G. Tzanetakis and P. Cook, “Musical genre classification of audio signals,” *IEEE Transactions on Speech and Audio Processing*, vol. 10, no. 5, pp. 293–302, Jul. 2002.

[29] K. Umapathy, S. Krishnan, and S. Jimaa, “Multigroup classification of audio signals using time-frequency parameters,” *IEEE Transactions on Multimedia*, vol. 7, no. 2, pp. 308–315, Apr. 2005.

[30] B. Boashash and P. O’Shea, “A methodology for detection and classification of some underwater acoustic signals using time-frequency analysis techniques,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 11, pp. 1829–1841, Nov. 1990.

[31] LJ. Stanković, T. Thayaparan, and M. Daković, “Signal decomposition by using the S-method with application to the analysis of HF radar signals in sea-clutter,” *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4332–4342, Nov. 2006.

[32] L. Stanković, S. Stanković, T. Thayaparan, M. Daković, and I. Orović, “Separation and reconstruction of the rigid body and micro-Doppler signal in ISAR Part I - theory,” *IET Radar, Sonar Navigation*, vol. 9, no. 9, pp. 1147–1154, 2015.

[33] ——, “Separation and reconstruction of the rigid body and micro-Doppler signal in ISAR Part ii - statistical analysis,” *IET Radar, Sonar Navigation*, vol. 9, no. 9, pp. 1155–1161, 2015.

[34] L. Stanković, I. Orović, S. Stanković, and M. Amin, “Compressive sensing based separation of nonstationary and stationary signals overlapping in time-frequency,” *IEEE Transactions on Signal Processing*, vol. 61, no. 18, pp. 4562–4572, Sept 2013.

[35] Y. Sun and J. Li, "Time-frequency analysis for plastic landmine detection via forward-looking ground penetrating radar," *IEE Proceedings - Radar, Sonar and Navigation*, vol. 150, no. 4, pp. 253–261, Aug. 2003.

[36] Y. Zhang, M. Amin, and G. Frazer, "High-resolution time-frequency distributions for manoeuvring target detection in over-the-horizon radars," *IEE Proceedings - Radar, Sonar and Navigation*, vol. 150, no. 4, pp. 299–304, Aug. 2003.

[37] L. Brechet, M.-F. Lucas, C. Doncarli, and D. Farina, "Compression of biomedical signals with mother wavelet optimization and best-basis wavelet packet selection," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 12, pp. 2186–2192, Dec. 2007.

[38] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, Jun. 2002.

[39] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[40] W. Dai and O. Milenković, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Transaction on Information Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.

[41] I. Orović, A. Draganić, and S. Stanković, "Sparse time-frequency representation for signals with fast varying instantaneous frequency," *IET Radar, Sonar Navigation*, vol. 9, no. 9, pp. 1260–1267, 2015.

[42] K.-K. Poh and P. Marziliano, "Compressive sampling of EEG signals with finite rate of innovation," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, pp. 183 105–1–12, 2010.

[43] A. M. Abdulghani, A. J. Casson, and E. Rodriguez-Villegas, "Compressive sensing scalp EEG signals: implementations and practical performance," *Medical & Biological Engineering & Computing*, vol. 50, no. 11, pp. 1137–1145, 2012.

[44] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, “Compressed sensing mri,” *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 72–82, March 2008.

[45] J. H. Ender, “On compressive sensing applied to radar,” *Signal Processing*, vol. 90, no. 5, pp. 1402–1414, 2010.

[46] B. Han, F. Wu, and D. Wu, “Image representation by compressive sensing for visual sensor networks,” *Journal of Visual Communication and Image Representation*, vol. 21, no. 4, pp. 325–333, 2010.

[47] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, “Compressed channel sensing: A new approach to estimating sparse multipath channels,” *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1058–1076, June 2010.

[48] Y. Rivenson, A. Stern, and B. Javidi, “Overview of compressive sensing techniques applied in holography,” *Applied Optics*, vol. 52, no. 1, pp. A423–A432, Jan 2013.

[49] L. Stanković, S. Stanković, and M. Amin, “Missing samples analysis in signals for applications to l-estimation and compressive sensing,” *Signal Processing*, vol. 94, pp. 401–408, 2014.

[50] D. Craven, B. McGinley, L. Kilmartin, M. Glavin, and E. Jones, “Compressed sensing for bioelectric signals: A review,” *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 2, pp. 529–540, March 2015.

[51] S. Stanković, I. Orović, and M. Amin, “L-statistics based modification of reconstruction algorithms for compressive sensing in the presence of impulse noise,” *Signal Processing*, vol. 93, no. 11, pp. 2927–2931, 2013.

[52] I. Orovic, V. Papic, C. Ioana, X. Li, and S. Stankovic, “Compressive sensing in signal processing: Algorithms and transform domain formulations,” *Mathematical Problems in Engineering*, 2016, article ID 7616393, 16 page.

[53] J. Bobin, J. L. Starck, and R. Ottensamer, “Compressed sensing in astronomy,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 5, pp. 718–726, Oct 2008.

[54] Y. Wiaux, L. Jacques, G. Puy, A. M. M. Scaife, and P. Vandergheynst, "Compressed sensing imaging techniques for radio interferometry," *Monthly Notices of the Royal Astronomical Society*, vol. 395, no. 3, p. 1733, 2009.

[55] C. R. Berger, S. Zhou, J. C. Preisig, and P. Willett, "Sparse channel estimation for multicarrier underwater acoustic communication: From subspace methods to compressed sensing," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1708–1721, March 2010.

[56] Y. Chi, L. L. Scharf, A. Pezeshki, and A. R. Calderbank, "Sensitivity to basis mismatch in compressed sensing," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2182–2195, May 2011.

[57] M. F. Duarte and Y. C. Eldar, "Structured compressed sensing: From theory to applications," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4053–4085, Sept 2011.

[58] S. Qaisar, R. M. Bilal, W. Iqbal, M. Naureen, and S. Lee, "Compressive sensing: From theory to applications, a survey," *Journal of Communications and Networks*, vol. 15, no. 5, pp. 443–456, Oct 2013.

[59] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7465–7490, Nov 2013.

[60] R. Leary, Z. Saghi, P. A. Midgley, and D. J. Holland, "Compressed sensing electron tomography," *Ultramicroscopy*, vol. 131, pp. 70–91, 2013.

[61] A. Veeraraghavan, D. Reddy, and R. Raskar, "Coded strobing photography: Compressive sensing of high speed periodic videos," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 4, pp. 671–686, April 2011.

[62] S. Ganguli and H. Sompolinsky, "Compressed sensing, sparsity, and dimensionality in neuronal information processing and data analysis," *Annual Review of Neuroscience*, vol. 35, pp. 485–508, 2012.

[63] A. Massa, P. Rocca, and G. Oliveri, “Compressive sensing in electromagnetics - a review,” *IEEE Antennas and Propagation Magazine*, vol. 57, no. 1, pp. 224–238, Feb 2015.

[64] H. Rauhut, K. Schnass, and P. Vandergheynst, “Compressed sensing and redundant dictionaries,” *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 2210 –2219, May 2008.

[65] O. Teke, A. C. Gurbuz, and O. Arikan, “A robust compressive sensing based technique for reconstruction of sparse radar scenes,” *Digital Signal Processing*, vol. 27, pp. 23–32, 2014.

[66] R. G. Baraniuk, “Compressive sensing,” *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–121, Jul. 2007.

[67] E. J. Candes and M. B. Wakin, “An introduction to compressive sampling,” *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[68] C. R. Berger, S. Zhou, J. C. Preisig, and P. Willett, “Sparse channel estimation for multicarrier underwater acoustic communication: From subspace methods to compressed sensing,” *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1708–1721, March 2010.

[69] T. Strohmer, “Measure what should be measured: Progress and challenges in compressive sensing,” *IEEE Signal Processing Letters*, vol. 19, no. 12, pp. 887–893, Dec 2012.

[70] E. J. Candes, Y. C. Eldar, D. Needell, and P. Randall, “Compressed sensing with coherent and redundant dictionaries,” *Applied and Computational Harmonic Analysis*, vol. 31, no. 1, pp. 59–73, 2011.

[71] M. Elad, “Optimized projections for compressed sensing,” *IEEE Transactions on Signal Processing*, vol. 55, no. 12, pp. 5695–5702, Dec 2007.

[72] L. C. Potter, E. Ertin, J. T. Parker, and M. Cetin, “Sparsity and compressed

sensing in radar imaging," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1006–1020, June 2010.

[73] B. Adcock, C. Hansen, C. Poon, and B. Roman, "Breaking the coherence barrier: A new theory for compressed sensing," *Forum of Mathematics, Sigma*, vol. 5, 2017.

[74] G. E. Pfander and H. Rauhut, "Sparsity in time-frequency representations," *Journal of Fourier Analysis and Applications*, vol. 16, no. 2, pp. 233–260, 2010.

[75] Y. Wang, J. Xiang, Q. Mo, and S. He, "Compressed sparse timefrequency feature representation via compressive sensing and its applications in fault diagnosis," *Measurement*, vol. 68, pp. 70–81, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0263224115001049

[76] M. G. Amin, B. Jokanovic, Y. D. Zhang, and F. Ahmad, "A sparsity-perspective to quadratic timefrequency distributions," *Digital Signal Processing*, vol. 46, pp. 175–190, 2015.

[77] G. Orchard, J. Zhang, Y. Suo, M. Dao, D. T. Nguyen, S. Chin, C. Posch, T. D. Tran, and R. Etienne-Cummings, "Real time compressive sensing video reconstruction in hardware," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 2, no. 3, pp. 604–615, Sept 2012.

[78] P. Flandrin and P. Borgnat, "Time-frequency energy distributions meet compressed sensing," *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 2974–2982, Jun. 2010.

[79] T. A. C. M. Claasen and W. F. G. Mecklenbrauker, "The Wigner distribution - a tool for time frequency signal analysis: Part I Continuous time signals," *Philips Journal of Research*, vol. 35, no. 3, pp. 217–250, 1980.

[80] B. Boashash, G. Azemi, and J. M. O'Toole, "Time-frequency processing of nonstationary signals: Advanced TFD design to aid diagnosis with highlights from medical applications," *IEEE Signal Processing Magazine*, vol. 30, no. 6, pp. 108–119, Nov. 2013.

[81] B. Jokanovic and M. Amin, “Reduced interference sparse time-frequency distributions for compressed observations,” *IEEE Transactions on Signal Processing*, vol. 63, no. 24, pp. 6698–6709, Dec 2015.

[82] S. G. Mallat and Z. Zhang, “Matching pursuits with time-frequency dictionaries,” *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[83] D. Needell and J. A. Tropp, “Cosamp: Iterative signal recovery from incomplete and inaccurate samples,” *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, May 2009.

[84] L. Hu, Z. Shi, J. Zhou, and Q. Fu, “Compressed sensing of complex sinusoids: An approach based on dictionary refinement,” *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3809–3822, July 2012.

[85] Y. Zhang, B. Dong, and Z. Lu, “l_0 minimization of wavelet frame based image restoration,” *Mathematics of Computation*, vol. 82, pp. 995–1015, 2013.

[86] T. Y. HOU and Z. SHI, “Adaptive data analysis via sparse time-frequency representation,” *Advances in Adaptive Data Analysis*, vol. 03, no. 01n02, pp. 1–28, Apr. 2011.

[87] D. Angelosante, G. B. Giannakis, and N. D. Sidiropoulos, “Sparse parametric models for robust nonstationary signal analysis: Leveraging the power of sparse regression,” *IEEE Signal Processing Magazine*, vol. 30, no. 6, pp. 64–73, Nov. 2013.

[88] G. Taubock, F. Hlawatsch, D. Eiwen, and H. Rauhut, “Compressive estimation of doubly selective channels in multicarrier systems: Leakage effects and sparsity-enhancing processing,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 255–271, April 2010.

[89] T. Y. Hou and Z. Shi, “Data-driven timefrequency analysis,” *Applied and Computational Harmonic Analysis*, vol. 35, no. 2, pp. 284–308, 2013.

[90] A. C. Gurbuz, V. Cevher, and J. H. Mcclellan, "Bearing estimation via spatial sparsity using compressive sensing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 2, pp. 1358–1369, APRIL 2012.

[91] A. Perelli, L. D. Marchi, L. Flamigni, A. Marzani, and G. Masetti, "Best basis compressive sensing of guided waves in structural health monitoring," *Digital Signal Processing*, vol. 42, pp. 35–42, 2015.

[92] E. Sejdić and L. F. Chaparro, "Time-frequency representations based on compressive samples," in *Proc. of 21st European Signal Processing Conference (EUSIPCO'13)*, Marrakech, Morocco, Sep. 9–13, 2013, pp. 1 569 742 057–1–4.

[93] S. Senay, L. F. Chaparro, and L. Durak, "Reconstruction of nonuniformly sampled time-limited signals using prolate spheroidal wave functions," *Signal Processing*, vol. 89, no. 12, pp. 2585–2595, Dec. 2009.

[94] T. Blu, P.-L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, "Sparse sampling of signal innovations," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 31–40, Mar. 2008.

[95] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty - V: The discrete case," *The Bell System Technical Journal*, vol. 57, no. 5, pp. 1371–1430, May/Jun. 1978.

[96] E. Sejdić, M. Luccini, S. Primak, K. Baddour, and T. Willink, "Channel estimation using DPSS based frames," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2008)*, Las Vegas, Nevada, USA, Mar./Apr. 31–4, 2008, pp. 2849–2852.

[97] T. Zemen and C. F. Mecklenbräuker, "Time-variant channel estimation using discrete prolate spheroidal sequences," *IEEE Transactions on Signal Processing*, vol. 53, no. 9, pp. 3597–3607, Sep. 2005.

[98] E. Sejdić, A. Can, L. F. Chaparro, C. M. Steele, and T. Chau, "Compressive sampling of swallowing accelerometry signals using time-frequency dictionaries based on

modulated discrete prolate spheroidal sequences," *EURASIP Journal on Advances in Signal Processing*, vol. 2012, pp. 101–1–14, May 2012.

[99] J. Oh, S. Senay, and L. F. Chaparro, "Signal reconstruction from nonuniformly spaced samples using evolutionary Slepian transform-based POCS," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, pp. 367 317–1–12, 2010.

[100] E. Sejdić and L. F. Chaparro, "Recovering heart sounds from sparse samples," in *Proc. of 38th Annual Northeast Bioengineering Conference (NEBEC 2012)*, Philadelphia, PA, USA, Mar. 16–18, 2012, pp. 107–108.

[101] V. C. Chen, F. Li, S. S. Ho, and H. Wechsler, "Micro-Doppler effect in radar: phenomenon, model, and simulation study," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 1, pp. 2–21, Jan 2006.

[102] M. Martorella, "Novel approach for isar image cross-range scaling," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, no. 1, pp. 281–294, January 2008.

[103] I. Djurović and L. Stanković, "Realization of robust filters in the frequency domain," *IEEE Signal Processing Letters*, vol. 9, no. 10, pp. 333–335, Oct 2002.

[104] V. Katkovnik, "A new form of the Fourier transform for time-varying frequency estimation," *Signal Processing*, vol. 47, no. 2, pp. 187–200, Nov. 1995.

[105] S. Stanković, I. Orović, and L. Stanković, "Polynomial Fourier domain as a domain of signal sparsity," *Signal Processing*, vol. 130, pp. 243–253, 2017.

[106] ——, "An automated signal reconstruction method based on analysis of compressive sensed signals in noisy environment," *Signal Processing*, vol. 104, pp. 43–50, 2014.

[107] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal - part 1: Fundamentals," *Proceedings of the IEEE*, vol. 80, no. 4, pp. 520–538, Apr. 1992.

[108] ——, “Estimating and interpreting the instantaneous frequency of a signal - part 2: Algorithms and applications,” *Proceedings of the IEEE*, vol. 80, no. 4, pp. 540–568, Apr. 1992.

[109] B. Boashash, Ed., *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference.* Amsterdam: Elsevier, 2003.

[110] J. Lerga and V. Sucic, “Nonlinear IF estimation based on the pseudo WVD adapted using the improved sliding pairwise ICI rule,” *IEEE Signal Processing Letters*, vol. 16, no. 11, pp. 953–956, Nov 2009.

[111] I. Orović, S. Stanković, and T. Thayaparan, “Time-frequency-based instantaneous frequency estimation of sparse signals from incomplete set of samples,” *IET Signal Processing*, vol. 8, no. 3, pp. 239–245, May 2014.

[112] I. Orović and S. Stanković, “Improved higher order robust distributions based on compressive sensing reconstruction,” *IET Signal Processing*, vol. 8, no. 7, pp. 738–748, September 2014.

[113] S. Stanković and LJ. Stanković, “Introducing time-frequency distribution with a “complex-time” argument,” *Electronics Letters*, vol. 32, no. 14, pp. 1265–1267, Jul. 1996.

[114] I. Orović and S. Stanković, “A class of highly concentrated time-frequency distributions based on the ambiguity domain representation and complex-lag moment,” *EURASIP Journal on Advances in Signal Processing*, vol. 2009, no. 1, p. 935314, 2009.

[115] E. Sejdić, I. Orović, and S. Stanković, “A software companion for compressively sensed time-frequency processing of sparse nonstatinary signals,” *SoftwareX*, 2017, under consideration.