

A Sparsity-Perspective to Quadratic Time-Frequency Distributions

Moeness G. Amin, Branka Jokanovic, Yimin D. Zhang, and Fauzia Ahmad

Center for Advanced Communications, Villanova University, Villanova, PA 19085, USA

E-mail: {moeness.amin, branka.jokanovic, yimin.zhang, fauzia.ahmad}@villanova.edu

Abstract

We examine nonstationary signals within the framework of compressive sensing and sparse reconstruction. Most of these signals, which arise in numerous applications, exhibit small relative occupancy in the time-frequency domain, casting them as sparse in a joint-variable representation. We present two general approaches to incorporate sparsity into time-frequency analysis, leading to what we refer to as sparsity-aware quadratic time-frequency distributions. Both approaches exploit time-frequency sparsity under full data and compressed observations. In the first approach, quadratic time-frequency distributions are derived based on optimal multi-task kernel design. In this case, sparseness in the time-frequency domain presents itself as a new design task, adding to the two fundamental tasks of auto-term preservation and cross-term suppression. In the second approach, sparse reconstruction is used, in lieu of the Fourier transform, to obtain quadratic time-frequency distributions from compressed measurements observed in the time domain or the joint-variable domain. It is shown that multiple measurement vector methods and block sparsity techniques play a fundamental role in improving signal local frequency representations. Examples of both approaches are provided. Analysis is supported by results based on simulated data, electromagnetic modeled data, and real Doppler and micro-Doppler data measurements of radar returns associated with human motions.

Keywords:

Sparse reconstruction, time-frequency representation, block sparsity, multiple measurement

vector, micro-Doppler.

1. Introduction

This review article deals with signal time-frequency (TF) signature reconstruction from complete and incomplete data. Incomplete data include missing observations or random sampling, and can be due to fading channels, discarding noisy measurements, hardware simplification, sampling frequency limitations, logistical restrictions on data collections and storage, or a result of co-existence between wireless communication systems and systems performing active or passive sensing [1, 2, 3]. The article groups recent developments and potential future advances in sparse nonstationary signal analysis into two fundamental approaches, both exploiting signal sparseness over the joint-variable TF domain.

There are numerous applications where nonstationary signals are present at the transmitter, receiver, or both. This covers astronomical, biological and man-made signals and spans ubiquitous active and passive sensing modalities, including sonar, radar, and ultrasound [1, 4, 5, 6]. Nonstationary signals, especially frequency modulated (FM) signals, are typically employed by smart jamming [7, 8], and also characterize speech and electromyographic recordings [9, 10]. In radar and sonar systems, the Doppler frequency is used to estimate the radial velocity of a target which can be a constant, linear, or nonlinear function of time. The rotation, vibration, and coning motion of a target or its parts may produce periodic Doppler modulations of the received signal, referred to as the micro-Doppler effect, which are best revealed in the time-frequency domain [11, 12, 13, 14, 15, 16, 17].

Unlike sinusoidal signals, FM signals are wideband, that is, the signals occupy the entire bandwidth under Nyquist sampling. In this respect, they are not globally sparse when represented in the frequency domain. However, owing to their power concentration in frequency at the different time instants, these signals are instantaneously narrowband. In this regard, the time-frequency signatures of a large class of nonstationary signals occupy

small regions in the TF domain. This property casts the signals as sparse in the joint-variable representations [18] and has recently invited sparse signal reconstruction and CS techniques [19, 20, 21, 22, 23] to play an important role in TF signal analysis and processing [24, 25, 26, 27, 28]. Opposite to stationary signals, where frequency sparsity can be assumed globally, local frequency reconstruction of a single- or a multi-component nonstationary signal is deemed to outperform the signal global reconstructions in which all data measurements are considered.

The objectives of recent research activities in sparsity-aware quadratic time–frequency distributions (SA-QTFDs) is to 1) utilize the signal joint-variable domain sparsity, 2) combat the adverse effect of missing samples on time–frequency signal representations, and 3) define effective sampling and data collection strategies for signals with time-varying spectral characteristics. Meeting these objectives will lead to fast data acquisition, improved signal and target detection and classification in radar, communications, sonar, and satellite navigation. Further, the research in SA-QTFDs benefits other important emerging applications in the area of big data where sparsity of the joint-variable domain provides a vehicle for reduction in data storage as well as efficiency in data recovery.

QTFDs seek to combine the instantaneous power and spectral energy density of the signal into one TF domain representation. However, missing samples and randomly under-sampled nonstationary signals give rise to artifacts that spread over both the TF domain and the ambiguity domain, which are related by the two-dimensional Fourier transform [29, 30]. These artifacts obscure the signal components and their instantaneous frequencies (IFs). Efforts and attempts to use traditional TF smoothing kernels for reducing missing samples artifacts proved ineffective and unsuccessful. This is because the ambiguity-domain low-pass filter characteristics, underlying signal-independent kernels which are applied to mitigate the signal cross-terms [31, 32, 33, 34, 35, 36], offer limited benefits against an exceedingly noisy ambiguity function (AF). On the other hand, missing samples can misguide the signal-

dependent adaptive kernels [37] into capturing the wrong areas of the AF. These difficulties and inabilities to apply QTFDs in their known nominal forms call for new approaches wherein sparsity can be either leveraged in TF kernel design or directly used in constrained optimization problems for TF signature reconstructions.

Compressed sensing (CS) has been studied extensively in many applications, including radar [38, 39, 40, 41, 42, 43]. In CS, a sparse representation of a signal is projected onto a much smaller measurement space. This leads, in general, to decreasing the data-acquisition requirements from the time, logistic, and hardware complexity perspectives. It is then possible to record a small number of linear measurements of a signal and, from those measurements, reconstruct the complete set of all samples that can be recorded conventionally. The required number of observations is slightly higher than the signal sparsity level, but far fewer than the signal ambient dimension. Although vastly applied in many applications, little consideration has been given to sparse reconstruction of nonstationary signals.

We consider the problem of SA-QTFDs using complete and compressed observations, following two general approaches, both are important and key to the understanding of the offerings of sparsity in enhancing nonstationary signal analysis. In the first approach, new QTFDs, within Cohen’s class, are introduced through a novel kernel design which has sparsity in the TF domain as one of its primary goals. These multi-task kernels combine low-pass filtering for reduced interference distributions (RIDs), with sparsity in the TF domain, yielding robustness to missing data. The result is a superior distribution over that obtained through conventional data-independent or data-dependent kernel design. Different sparsity measures applied in the TF domain can be used to solve for the optimum multi-task TF kernels.

In the second approach, one departs from Cohen’s class and replace the Fourier transform (FT), which connects time, time-lag, and ambiguity domains to the TF domain, by a corresponding linear dictionary and solve the respective sparse reconstruction and opti-

mization problem. This approach underlines most of the recent contributions of SA-QTFDs. The work in [18, 25, 44, 45] performs sparse reconstructions from windowed data in the time domain and, in this respect, it parallels the short-time Fourier transform (STFT) and the spectrogram. Sinusoidal and chirp atoms have both been used within each window to form the dictionary matrix with the latter outperforming the former due to its better approximation of the local frequency behavior of most FM signals. Aside from the time domain, sparse reconstructions are carried out in [24] from compressed observations in the ambiguity domain, and in [29] from the instantaneous autocorrelation function (IAF) domain, i.e., the time-lag domain. The difference between the ambiguity domain and the IAF domain is that missing samples in time lead to missing samples in the IAF, but not in the AF, which only becomes noisy. Additionally, reconstruction from the IAF domain allows the use of the dictionary matrix with a reduced size and enables exploitation of local sparsity over a short time period. A clear and fundamental role of a multiple measurement vector (MMV) model, block sparsity, and multi-task Bayesian compressive sensing (BCS) techniques in revealing the signal local power behavior was established in [46]. The MMV model can arise from using multiple data windows [47] reminiscent of the multiple window spectrogram [48, 49, 50, 51, 52, 53, 54, 55]. BCS enables, through the priors, the incorporation of the contiguity property of most TF signatures and thus enhances sparse optimization solutions. It should be noted that, in the above two approaches, we deal with complex data whereas a missing sample implies that both the sample's real and imaginary parts are unavailable. Further, missing samples are drawn from a uniform distribution. As such, data with contiguous missing samples are unlikely to occur neither are they enforced. Reference [56] addresses this case and presents solutions based on empirical mode decomposition.

A hybrid approach combining the aforementioned two approaches can also be used. In this case, sparsity-aware TF kernels are designed and applied to the AF similar to the first approach. A sparsity measure is then used to produce a signal power distribution in the

TF domain. In this respect, sparsity is used twice, in kernel design as well as in obtaining QTFDs from the ambiguity domain.

The above two approaches and their hybrid schemes constitute a nonparametric perspective to SA-QTFDs. There is also a parametric dictionary based approach which is directly applied to the time-domain data to estimate the signal parameters [27]. This parametric perspective is justified by the need, in many applications, to perform classification based on features related to the estimated signal parameters. However, this approach specifically deals with nonstationary signals with *a priori* known structures, such as chirps and sinusoidal FM signals. As such, it works well when there is a good match between the assumed and the actual signal characteristics, but remains sensitive to deviations from the assumed model.

This paper also considers interpolation as a method to deal with missing data and as an alternative to the above CS approaches. Upon obtaining the interpolated data, one can then proceed with the computation of TF distributions. Since missing data samples in time introduce missing samples in the time-lag domain, interpolation can be performed in either the time or the time-lag domain. It is shown in [57] that interpolation of the IAF outperforms data interpolation in time, as it acts on reducing cross-terms through its underlying low-pass filter characteristics. As such, it provides a better approximation of the TF signature when the complete data is considered. In this respect, IAF interpolation is better suited for QTFDs even though it exhibits more missing samples than those originally occurring in time. We compare QTFDs with and without data interpolation and contrast their performance with sparse signal reconstruction.

In addition to deterministic signals, QTFDs were used in the past to estimate the time-varying spectrum of nonstationary random processes [58, 59, 60]. When dealing with under-spread nonstationary random processes, compressive sensing techniques have been recently applied to estimate the minimum variance spectrum [61]. This work parallels that in [24]

for deterministic signals, but views the applied kernel and compressed observations in the ambiguity domain as an averaging process to reduce spectral variance. In this regard, it can be considered part of the second approach underlying SA-QTFDs. We focus in this paper on deterministic signals, and refer the reader to reference [61] for detailed information on examples and applications of sparse reconstruction of nonstationary random process with sparsity in the TF domain. SA-QTFDs may employ a variety of sparse reconstruction algorithms which may differ in accuracy, convergence, and complexity. These include greedy algorithms and techniques based on convex optimization. Many of those techniques guarantee reliable reconstruction only when the dictionary, linearly relating observations with the sparse signature, meets the so-called restricted isometry property (RIP) condition. This condition can limit the capability of these algorithms in achieving high-resolution sparse signal reconstruction. Sparse Bayesian learning methods [62], such as the multi-task CS (MT-CS) algorithm [63], form a different class of sparse signal reconstruction algorithms that are insensitive to the RIP restriction. These algorithms are originally designed to recover real-valued sparse solutions. The complex multi-task BCS (CMT-BCS) algorithm [64], which enforces group sparsity between real and imaginary components of complex sparse entries, is effective to handle problems that involve complex entries as in the case of a sparsity-aware time-frequency signal representation (SA-TFSR).

The organization of the paper is as follows. Section 2 presents a review of commonly applied quadratic TF analysis methods and sparse optimization techniques. The interpolation approach for estimating missing data samples is given in Section 3. Section 4 motivates the TF sparsity problem and presents the overall two-fold framework for SA-QTFDs. In Section 5, we expand Cohen's class of QTFDs to include sparsity promoting TF kernels and provide an example of the multi-task kernel design. In Section 6, we describe sparse reconstruction of the signal TF signature using compressed observations measured in the time, time-lag, and ambiguity domains. For the case of time-domain observations, both sinusoidal and chirp

dictionaries are used for local frequency reconstruction. Single and multiple data windows are considered, yielding single and multiple vector measurements. We identify BCS as the technique of preference due its robustness to mutual coherence of the dictionary matrix as well as its suitability in accounting for the signal continuous TF properties. In Section 7, we present the parametric dictionary approach and show its offerings and limitations. Section 8 is the conclusion.

2. Background

2.1. Time-Frequency Distributions

Cohen's class of QTFDs of a nonstationary signal $x(n)$, $n = 0, 1, 2, \dots, N-1$, given a kernel function $C(p, m)$, is defined as

$$\begin{aligned}
 D_{xx}(n, k) &= \sum_{p=-N/2}^{N/2-1} \sum_{m=-N/2}^{N/2-1} C(p, m) A_{xx}(p, m) e^{j(-mk-pn)2\pi/N}, \\
 A_{xx}(p, m) &= \sum_{n=0}^{N-1} x\left(n + \frac{m}{2}\right) x^*\left(n - \frac{m}{2}\right) e^{-j2\pi pn/N},
 \end{aligned} \tag{1}$$

where $j = \sqrt{-1}$ and $(\cdot)^*$ denotes complex conjugation. $A_{xx}(p, m)$ is AF of signal $x(n)$, with p and m , respectively, denoting the frequency shift (also referred to as the Doppler frequency) and the time lag. The kernel $C(p, m)$ acts as a filter and places different weightings on the AF. The majority of signals in practice have auto-terms located near the origin in the ambiguity domain, while the signal cross-terms are distant from the time-lag and frequency-shift axes. This property has motivated researchers to propose RID kernels with low-pass filter characteristics to suppress cross-terms and preserve auto-terms. RIDs can be signal-independent or signal-dependent. The former involves applying a fixed kernel to smooth the Wigner-Ville distribution (WVD), for which $C(p, m) = 1, \forall p, \forall m$. On the other hand, the most well-known signal-dependent kernel is the adaptive optimal kernel (AOK), which

is formulated as an optimization problem [37]:

$$\begin{aligned}
& \underset{C(r,q), \forall r, \forall q}{\text{maximize}} && \sum_{r=0}^{N-1} \sum_{q=0}^{N-1} r |A_{xx}(r, q)C(r, q)|^2 \\
& \text{subject to} && C(r, q) = e^{-\frac{r^2}{2\sigma^2(q)}}, \\
& && \sum_{q=0}^{N-1} \sigma^2(q) \leq \alpha.
\end{aligned} \tag{2}$$

The objective of the AOK is to maximize the energy in the ambiguity domain in order to preserve the auto-terms. Note that the AF and the kernel function in (2) are defined in the polar coordinates in terms of the radius r and the aspect angle q . The Gaussian constraint in $C(r, q)$ forces the kernel to have low-pass filter characteristics, which lends itself to cross-term reduction, if not suppression. The second constraint specifies the volume under the optimal kernel. The choice of α influences the tradeoff between cross-term suppression and auto-term preservation.

2.2. Effect of Missing Samples on QTFDs

Missing data and random measurements cause [29]: a) missing entries in the IAF and generation of noise-like artifacts in the WVD and AF; and b) reduction in the magnitude of the auto-term entries due to lost energies. Fig. 1 compares the WVD, AF, and the IAF of a signal consisting of two parallel polynomial phase components [29]. In the WVD, the artifacts due to 60% missing data make TF signature recognition difficult.

Exploitation of TF kernels may improve the QTFD. However, it is shown in [30] that, in order for a kernel under incomplete data to produce the same results as that applied to complete data, it must lose its nominal characteristics of being real, symmetric and satisfying the marginal properties. This shows that nominal data-independent kernels will not meet their intended goal of reduced artifacts, which typically increase with a higher number of missing samples.

As shown in Fig. 2(a), the AOK is superior to a data-independent kernel and results in significant artifact reduction while recovering the auto-terms. However, because of the high artifacts along the lag axis in the ambiguity domain arising from the impulsive-like effect of missing samples (see Fig. 2(b)), the AOK becomes misguided and captures the wrong terms, leading to a cluttered TF domain in the form of vertical strips. Fig. 2(c) shows the Choi-Williams distribution (CWD) [35] as an example of using a data-independent distribution kernel. Similar results were obtained using more recent kernels, including the B-distribution [65]. While both of these kernels effectively reduce the noise-like artifacts, the missing energy of the auto-terms is not well recovered, and the vertical strips remain pronounced. For comparison, the spectrogram is shown in Fig. 2(d) demonstrating poor resolution, but not resulting in a clutter TFD such as in Figs. 2(a) and 2(b).

2.3. Compressive Sensing and Sparse Reconstruction

For a brief review of the main ideas underlying CS, consider a signal vector $\mathbf{x} \in \mathbb{C}^N$, which is K -sparse. That is, the vector \mathbf{x} has at most K nonzero elements where $K < N$. The set of positions of the nonzero elements of \mathbf{x} is referred to as the support of \mathbf{x} . Only $M \ll N$ linear non-adaptive measurements of the signal \mathbf{x} are recorded, i.e.,

$$\mathbf{y} = \mathbf{\Phi}\mathbf{x} + \boldsymbol{\epsilon}, \quad (3)$$

where $\mathbf{y} \in \mathbb{C}^M$ is the measurement vector, $\mathbf{\Phi} \in \mathbb{C}^{M \times N}$ is a known dictionary matrix, which is typically designed with some degree of randomness, and $\boldsymbol{\epsilon} \in \mathbb{C}^M$ is an unknown complex Gaussian noise vector.

Given the single measurement vector (SMV) model in (3), the sparse signal vector \mathbf{x} can be recovered uniquely with a high probability from the measurement vector \mathbf{y} provided that the matrix $\mathbf{\Phi}$ has some desirable attributes and the dimension M of the measurement vector \mathbf{y} is at least of the order of $K \log(N/K)$ [66]. Several sparse reconstruction algorithms

Algorithm 1 - OMP

Inputs: measurement vector \mathbf{y} , matrix \mathbf{A} , initial residual $\mathbf{r}_0=\mathbf{y}$, $\Phi_0=\emptyset$ Output: signal vector \mathbf{x}

At each iteration i until the stopping criterion is met do:1. Find the j -th column of matrix \mathbf{A} which has the highest correlation with current residual,

$$j = \arg \max | \langle \mathbf{r}_{i-1}, \mathbf{A}_j \rangle |.$$

2. Augment the matrix of chosen atoms,

$$\Phi_i = [\Phi_{i-1}; \mathbf{A}_j].$$

3. Solve the least square problem,

$$\mathbf{x}_i = \arg \min \|\mathbf{y} - \Phi_i \mathbf{x}_i\|_2.$$

4. Calculate the new residual,

$$\mathbf{r}_i = \mathbf{y} - \Phi_i \mathbf{x}_i.$$

are available in the literature, ranging from those based on ℓ_1 -norm convex optimization to iterative greedy algorithms. The ℓ_1 -norm related reconstruction scheme solves the convex optimization problem (or its variants) as follows:

$$\begin{aligned} & \underset{\mathbf{x}}{\text{minimize}} && \|\mathbf{x}\|_1 \\ & \text{subject to} && \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 \leq \varepsilon_0, \end{aligned} \tag{4}$$

where $\|\cdot\|_1$ denotes the ℓ_1 -norm, $\|\cdot\|_2$ represents the Euclidean (ℓ_2)-norm, and the parameter ε_0 is a regularizer. There are many solvers of the ℓ_1 -regularized formulation in (4) or its variants, such as the least absolute shrinkage and selection operator (LASSO) [23]. On the other hand, greedy algorithms iteratively build the sparse solution by identifying the support set, either one or multiple terms at a time. The orthogonal matching pursuit (OMP) algorithm is a popular greedy algorithm [22] which can be summarized in Algorithm 1.

In some applications, a sequence of measurement vectors, $\mathbf{y}_l, l = 1, \dots, L$ with $L > 1$, is acquired from slowly varying sources or similar scenes. The underlying ensemble of signals $\{\mathbf{x}_l\}_{l=1, \dots, L}$ share a common sparse support, resulting in what is commonly referred to

as the multiple measurement vector (MMV) model. The mapping from the measurements $\{\mathbf{y}_l\}_{l=1,\dots,L}$ to the signals $\{\mathbf{x}_l\}_{l=1,\dots,L}$ can be cast as a block/group sparse reconstruction problem, which can be solved with a number of algorithms, such as the block OMP and the group LASSO algorithms [67, 68]. With a common dictionary, the multiple measurements and signals are arranged into matrices and the signal ensemble recovery problem is formulated as a matrix norm minimization problem, which can be solved using, for example, the regularized M-FOCUSS algorithm [69].

Recently, Bayesian based CS algorithms have received much attention because they generally achieve better reconstruction performance than those obtained by the aforementioned conventional sparse reconstruction approaches [70, 71, 72]. The BCS [62] and its extended multi-task CS (MT-CS) [63] algorithms constitute an important family of algorithms to recover sparse signals under the SMV and MMV models, respectively. The BCS approaches have originally been designed for real-valued sparse solutions. Complex-valued signals and measurements are typically handled by decomposing them into independent real and imaginary components. Effective BCS algorithms for recovering complex-valued signals under both SMV and MMV models has been recently developed [64, 73].

3. Linear Interpolation for TF Signature Estimation

Before embarking on the discussion of SA-QTFDs, we examine QTFDs obtained from interpolated data. Interpolation attempts to uncompress the observations and, in so doing, it reduces errors when performing TF analysis. Interpolation is very common in signal processing and, in general, can be formulated as,

$$\mathbf{y} = \mathbf{H}\mathbf{x}, \tag{5}$$

where matrix $\mathbf{H} \in \mathbb{C}^{N \times N}$ is typically sparse and represents a linear operator that maps vector $\mathbf{x} \in \mathbb{C}^N$ which has missing samples into vector $\mathbf{y} \in \mathbb{C}^N$ which, for the considered interpolation problem, contains the coefficients of the interpolation kernel. The above equation does not require iteration as in the case for sparse reconstruction. It represents a time-invariant filter operation for the case of uniform down-sampling, where the number of interpolated samples B between any two consecutive observations remains the same. In the case of random under-sampling, B becomes time-dependent, amounting to time-varying filtering operations.

In applying interpolators to nonstationary signals, and prior to computing TF distributions, one has the option to either interpolate the raw data in the time domain or to interpolate the bilinear data products, referred to as IAF, in the time-lag domain. Without loss of generality, we use a linear interpolator as a representative of classical interpolators. When directly operating on the randomly sampled data, the interpolated data in the frequency domain become

$$Y(k) = \sum_{p=1}^P X(k; p)H(k; p), \quad k = 0, \dots, N-1, \quad (6)$$

where P is the number of interpolation segments, $Y(k)$ is the DFT of \mathbf{y} at frequency index k , $X(k; p)$ is the DFT of \mathbf{x} over the p th interpolation segment, and $H(k; p)$ is the interpolator frequency response over the same segment.

Equation (6) implies a time-varying low-pass filtering operation. The exclusion of high frequency components due to filtering would depend on B . In essence, fewer missing samples cause a less truncated frequency spectrum. Consequently, when dealing with under-sampling of critically sampled data, part of the high signal bandwidth will be lost, leading to signal distortion.

Performing interpolation in the time-lag domain corresponds to applying low-pass filter-

ing in the ambiguity domain for each lag m ,

$$A_{yy}(k, m) = \sum_{p=1}^{P(m)} A_{xx}(k, m; p)H(k, m; p), \quad k, m = 0, \dots, N-1, \quad (7)$$

where $P(m)$ is the number of interpolation segments used at lag m , $A_{xx}(k, m; p)$ is the AF over the p th interpolation segment, and $H(k, m; p)$ is the interpolator frequency response over the same segment.

In this case, removal of high frequency components occurs in the ambiguity domain. Since the signal auto-terms are of low-pass characteristics with high power concentration at or near the origin, the filtering operation in (7) is considered less harmful to the TF signature than that of the time-domain interpolation in (6).

In order to demonstrate the specific role of linear interpolators in different domains, we examine their performance for polynomial phase signals. In all cases, the WVD is computed using the interpolated data. For comparison, we examine both the WVD and the CWD applied to incomplete data. The results demonstrate the advantage of using the linear interpolator in the time-lag domain. In all plots, the frequency axis is normalized.

Consider a sinusoidal FM signal

$$x(n) = e^{j32 \sin(4\pi n/N) + jn\pi/2},$$

where $N = 256, n = 0, \dots, N-1$, and 40% of noiseless data is randomly missing. Fig. 3(a) shows the WVD with high level of artifact which is the consequence of missing samples in the time domain. The effects of applying the linear interpolator in different domains are shown in Figs. 3(b) and 3(c). As depicted in Fig. 3(c), the WVD, obtained from the interpolated IAF over time, reveals the desired TF signature, even though some artifact is still present. Also, we show results using other interpolation techniques, namely splines

and nearest neighbor method (Fig. 4). We can notice that these methods exhibit similar behavior as linear interpolation.

In the next example, we observe a multi-component signal,

$$x(n) = e^{j4.8\pi(2n/N-1)^3+j64n/N} + e^{j4.8\pi(2n/N-1)^3-j192n/N},$$

where $N = 256, n = 0, \dots, N - 1$. The signal is randomly under-sampled and 50% of the samples are missing. The results in Fig. 5 illustrate, as in the previous cases, the specific filtering role of the interpolator in the ambiguity domain. However, the cross-terms in Fig. 5(b) are not completely removed by the linear interpolator, indicating the dependence of the filter cutoff frequency on the number of missing samples. A higher number of missing samples will amount to decreasing the cutoff frequency of the corresponding filter, but at the cost of estimating more unknown points. Fig. 5(c) shows the result corresponding to 35% of the observed data. While fewer cross-terms are produced due to a narrower interpolator filter, the auto-terms are adversely affected due to the same reason.

This analysis sheds new light on the role of interpolators as a cross-term removal method. Another way to interpolate in the time-lag domain is to estimate the missing samples for a given time over the lag variable, which is shown in [57] to be inferior to that performed in the same domain over the time variable. Out of the aforementioned three ways to interpolate the data, the one with the filtering effect in the ambiguity domain can be considered most attractive for Doppler signals. The simple implementation of classical interpolators motivates their use in the TF signature estimation problem, but they remain inferior to SA-QTFDs, as demonstrated in Section 5 (see Fig. 11).

4. Concept and Overall Approaches for SA-QTFDs

Owing to their instantaneous narrowband characteristics and high power concentrations over the joint TF variables, the signatures of a large class of nonstationary signals occupy small regions in the TF domain. In other words, most of the TF domain is vacant. This property casts the signals as sparse in the joint-variable representations and has recently invited sparse signal reconstruction and CS techniques to play an important role in TF signal analysis and processing [24, 25, 26, 27, 28, 29, 30, 44, 45, 46, 47]. We focus on sparsity-aware nonparametric TFSRs based on quadratic distributions. We reconsider the offerings of QTFDs from a sparsity perspective and accommodate a broad class of real-world signals beyond frequency hopping and signals representable by time-varying autoregressive models [74].

Conceptually, consider an ideal TF representation of a sinusoidal FM signal and a linear FM (LFM) signal as shown in Fig. 6. The IF signal characterization demonstrates: a) Sparsity in the two-dimensional (2-D) TF domain by the virtue of perfect power localization, making the signal appear as a wavy line in a plane populated by zero values; b) Frequency sparsity, where the intersection of a vertical time slice and the wavy line produces only two nonzero values over frequency which, in turn, establishes a $K=2$ sparsity property; and c) Time sparsity along a frequency slice. With non-ideal practical TFSR, the time and frequency slices are, respectively, replaced by short windows and narrowband filters, as depicted in Fig. 6(b). In this case, sparsity is reduced when compared to the ideal case, due to the inclusion of consecutive points. The TFSR in Fig. 6(b) is provided by the spectrogram, which is the magnitude squared of the STFT.

Micro-Doppler signals are generally sparse and exhibit significant changes in their structures over time. Radar micro-Doppler is the phenomenon of the observed micro-motions on the top of the bulk main Doppler component of a target's motion. It has been the subject of research over a number of years focusing on the additional information that can be extracted

from this signal [75]. Such information can then be exploited in a variety of applications for security, law enforcement, urban warfare, and search and rescue, where the detection, tracking, and classification of many human targets moving in a cluttered environment is of paramount importance. These signals are also associated with radar returns from vibrating, oscillating, and rotating targets [43].

As illustrated in Fig. 7, sparsity-aware quadratic TF analysis follows two different approaches. The first approach is to maintain the FT operations connecting the four domains of time, IAF, AF, and TF. These transformations can be generalized through Cohen’s class, described in (1), to include kernel operations that are responsible for removing cross-terms. In this case, the kernel, which thus far has been designed with two purposes of cross-term suppressions and auto-term preservations, is now required to be multi-task, i.e., also promoting a sparse QTFD. In the second approach, we replace the FT by a linear dictionary and solve a sparse reconstruction problem to recover the QTFD from complete or compressed data measured in any of the three other domains, as depicted in Fig. 8. This approach typically deals with SMV and can be extended to exploit MMVs and make use of multi-task BCS. A hybrid approach that combines the above two approaches is also considered.

5. Sparse Kernel Design – A Multi-task Design

Designing a proper kernel has been an important task in TF analysis since the introduction of RIDs. Even though various kernel designs have been proposed in the literature, their primary purpose has been invariably the capture of the true signal components. Traditional kernels of signal-independent forms did not observe the case of missing samples. Robustness to missing data can, however, be achieved by imposing sparsity constraints in the kernel design. In so doing, kernels become multi-task in the sense that, in addition to auto-term preservation and cross-term suppression, they are tasked to provide sparse TFSRs.

Missing samples produce artifacts in the ambiguity domain. The artifacts pattern de-

depends on the values of the missing samples as well as their positions. Thus, some parts of the ambiguity function will be more affected than others. This problem motivated the introduction of optimum kernels which can maintain their desirable performance despite the heavily cluttered ambiguity domain. We invoke the sparsity condition of the QTFD which carries a different optimality than that embedded in the commonly used AOK. The latter can itself be used as a platform for the new kernel formulation.

The optimization problem in a discrete form is formulated as follows:

$$\begin{aligned}
& \underset{\sigma}{\text{minimize}} && \| \mathbf{A} - \mathbf{A} \odot \mathbf{C} \|_2 + \eta \| \mathbf{W}(\mathbf{A} \odot \mathbf{C})\mathbf{W}^H \|_1 \\
& \text{subject to} && C(r, q) = e^{-\frac{r^2}{2\sigma^2(q)}}, \\
& && \sum_{q=0}^{N-1} \sigma^2(q) \leq \alpha,
\end{aligned} \tag{8}$$

where η represents the regularization parameter and \mathbf{W} represents the discrete Fourier transform (DFT) matrix. \mathbf{A} corresponds to the short-time AF, i.e., the AF of the windowed signal centered at a time instant. In addition, \odot denotes element-wise multiplication and $(\cdot)^H$ denotes Hermitian conjugate. As such, we search for $\sigma(q)$ in a radial Gaussian kernel, which would fulfill the three desired tasks of auto-term preservation, cross-term suppression, and TF sparseness. In the proposed objective function, two terms are minimized. The first term is an error between the original AF and the kernelled AF. Minimization of this term amounts to preserving as much energy as possible in the ambiguity domain. The second term imposes sparsity in the TF domain and its role is to maintain good signal localization properties under incomplete data. After computing $\sigma(q)$ and the corresponding kernel function, we perform the 2-D DFT to obtain the corresponding QTFD. Since we are searching for the TF SR of one time slice, we use only one column of the obtained matrix corresponding to the centered window instant. In order to find the value of $\sigma(q)$, the gradient method can be used. This method requires computing the derivative of the ℓ_1 -norm penalty term. This term is not

differentiable, motivating the use of the sub-gradient method where σ is updated at each iteration [76].

Below, we demonstrate the effectiveness of the proposed kernel for the signal considered in Fig. 2, where 50% of the noise-free data is randomly missing. We compare the proposed kernel QTFD with the AOK distribution, the CWD, and the spectrogram. The results are depicted in Fig. 9, and they clearly demonstrate the effectiveness of the proposed approach. As shown in Figs. 9(b) and 9(c), the CWD and the spectrogram fail to capture the signal TF signature under missing samples. On the other hand, the AOK based distribution in Fig. 9(a) signifies the two signal components, but is highly cluttered. The result of the proposed sparse adaptive kernel design, as shown in Fig. 9(d), successfully suppresses such artifacts.

The previous approach seeks the sparsest QTFD which could be obtained based on the AF of the incomplete data. Thus, part of the clutter will always be captured since the AF itself is distorted. This challenge can be addressed by incorporating the estimation of the AF in the process of kernel design. The following design solves for kerneled AF $\mathbf{A}_{\mathbf{C}}$ and kernel \mathbf{C} ,

$$\begin{aligned}
& \underset{\mathbf{A}_{\mathbf{C}}, \mathbf{C}(\sigma)}{\text{maximize}} && GI[\text{vec}(\text{FT}_2(\mathbf{A}_{\mathbf{C}}))] \\
& \text{subject to} && \|\mathbf{r}_{\mathbf{xx}}^m - \mathbf{Q}(\mathbf{C})\mathbf{a}_{\mathbf{C}}^m\|_2 \leq \varepsilon, \quad m = 1, 2, \dots, N, \\
& && \mathbf{C}(\sigma) \text{ is a low-pass kernel,}
\end{aligned} \tag{9}$$

where $\text{vec}(\cdot)$ denotes the vectorization operation, FT_2 denotes the 2-D DFT, GI denotes the Gini index, and σ is a parameter which determines the kernel extent. $\mathbf{a}_{\mathbf{C}}^m$ is the m th row of $\mathbf{A}_{\mathbf{C}}$, arranged in a column, representing the AF of the m th lag. The TF sparsity is measured using the Gini index, which is defined for sorted real-valued signal vector \mathbf{x} , $x(1) \leq \dots \leq x(N)$ as,

$$GI(\mathbf{x}) = 1 - 2 \sum_{n=1}^N \frac{x(n)}{\|\mathbf{x}\|_1} \frac{N - n + 0.5}{N}. \tag{10}$$

Signals with a small number of nonzero values have a high GI [30]. It should be noted that, just like in the previous case when dealing with matrices, in order to compute the Gini index, a vectorized form of the QTFD is used. The first constraint in (9) is responsible for data fitting, as it minimizes the ℓ_2 -norm of residuals. This constraint relates the measurements in the time-lag domain $\mathbf{r}_{\mathbf{xx}}^m$ and kerneled AF $\mathbf{a}_{\mathbf{C}}^m$ for each lag m . Matrix \mathbf{Q} is a partial DFT matrix and is determined by the extent of kernel \mathbf{C} . By solving the ℓ_2 -norm optimization problem, we obtain the kerneled AF for each m , $\mathbf{a}_{\mathbf{C}}^m$. Setting the kernel to have low-pass filtering characteristics discards the high-frequency columns in the partial DFT matrix. The second constraint in (9) is necessary for ensuring that the auto-terms are captured without the inclusion of cross-terms. Any low-pass filter shape can be applied. Since the optimum \mathbf{C} and $\mathbf{A}_{\mathbf{C}}$ are unknown, the optimization in (9) can be computationally demanding when compared to some of the well-known one-variable optimizations, e.g., basis pursuit. An efficient algorithm was proposed in [30] for solving (9).

In order to evaluate the performance of the kernel design in (9), we consider a noiseless signal consisting of two crossing chirps:

$$x(n) = e^{j64\pi(n/N)^2} + e^{-j64\pi(n/N)^2},$$

where $N = 128$ and $n = 0, \dots, N - 1$. The above signal is sampled according to the Nyquist theorem, followed by random removal of 50% of the samples. Both the CWD and the AOK based distributions of incomplete data (Figs. 10(a) and 10(b)) contain clutter introduced by the missing samples. Also, we can notice vertical lines in the cluttered CWD. These lines are impulses which, due to the marginal properties of the CWD kernel, are captured along the Doppler frequency axis in the ambiguity domain, as discussed in [30]. In Fig. 10(c), the QTFD based on the proposed kernel design is shown. Even though 50% of samples are missing, the TF signature clearly depicts the two components.

In the second example, we observe the multicomponent signal observed in Section 3 (Fig. 5). The phase of two signal components is a third degree polynomial. As in the previous example, the signal is under-sampled, 65% of samples are randomly missing. The proposed kernel design provides the TF signature without gaps or vertical lines which exist in both the the AOK based distribution and the CWD (Fig. 11).

6. Non-parametric Sparse reconstruction of Signal TF Signatures

6.1. Ambiguity Domain Sparse Reconstruction

Reconstructions of QTFDs from few ambiguity domain observations is based on selecting suitable AF samples close to the origin and then searching for the sparsest QTFD, in ℓ_1 -norm, which best approximates the AF observations in the least square error sense. The argument made in favor of this technique, which is part of the second approach in Fig.7, is that masking the AF around the origin would have two opposite effects on the QTFD: it reduces cross-terms due to the mask low-pass filter characteristics, but it also degrades the TF resolution and IF localization. This degradation is combated by using sparse reconstruction in place of the crude Fourier inversion. The problem can be formulated as follows,

$$\begin{aligned} \mathbf{D}_x &= \arg \min_{\mathbf{D}} \|\text{vec}(\mathbf{D})\|_1 \\ \text{subject to } &\|\text{vec}(\text{FT}_2\{\mathbf{D}\}) - \text{vec}(\mathbf{A}_x)\|_2^2 \leq \varepsilon, \end{aligned} \quad (11)$$

where \mathbf{D}_x is the estimated TFSR matrix, and \mathbf{A}_x is the AF of the observed signal. The work in [24] on this subject also considers the case where there is a strict equality constraint rather than the ℓ_2 -norm. It is noteworthy that the results of this approach heavily depend on the choice of the ambiguity domain mask. The authors in [24] use the uncertainty principle to select the size of the mask in the order of the signal length. In [77], this mask is decided by applying an adaptive optimum kernel, which is described in (2).

The QTFD obtained by sparse reconstruction of a masked AF works well with complete data (Fig.12(a)). In the case of random under-sampled data, the AF samples within the mask become cluttered, rendering this technique ineffective. This is shown in Fig. 12(b) for a signal composed of a sinusoidal FM and a chirp when 50% of samples are missing.

6.2. Local Sparse Reconstruction from a Single Window

One of the most straightforward sparse reconstructions of local signal frequency characteristics is achieved by applying a sliding time window, reminiscent of the STFT. Using a partial Fourier basis, one can proceed to apply greedy algorithms or convex optimization techniques to find the sparsest local frequency contents that describe the observations within the window. Since we deal with sliding windows and pursue reconstruction at every time sample, it is prudent to use computationally efficient algorithms for spectral recovery. This justifies the use of greedy algorithms which are easier to implement than linear programming methods. In this case, the application of the OMP should go beyond the first iteration for a single component signal. This is attributed to the fact that the OMP strives to find the column (atom) of the dictionary matrix which is most correlated with the windowed observations. With a Fourier dictionary, this correlation corresponds to performing the FT of the observations and then searching for the peak location. Accordingly, estimating the IF over a sliding window provides the same results for both the single-iteration OMP and the spectrogram. Of course, due to the orthogonal projection underlying the OMP second and higher iterations, the results of these two methods differ when dealing with multi-component signals. The work in [25] proposed to overcome the above similarity by performing several OMP iterations followed by IF estimation based on averaging, frequency of occurrence, and maximum value of the iteration results. The follow-up work in [18] considered the same problem, and discussed the paradigm of sparsity, number of observations, and window length. It is shown that, for LFM signals, the sparsity of the local frequencies is proportional to the window length, thus adding another condition on window selections, besides

the conventional temporal and spectral resolution requirements. It was maintained that stable signal reconstruction within a sliding window depends on the underlying probability distribution function guiding the random sampling intervals. It is important to mention that, in this approach as well as in classical window-based TF methods, continuity in time is implicitly assumed using overlapping sliding windows. However, as reported in [78], highly structured TF representations can benefit from explicitly imposing time smoothness. In [78], this enforcement was achieved using Bayesian framework.

In order to demonstrate the benefits of local sparse reconstruction, we consider computational electromagnetic modeling data corresponding to a human walking straight toward a pulse-Doppler radar (which means 0° azimuth). Vertical polarization is considered. The radar operates around a carrier frequency of 1 GHz with a bandwidth of 80 MHz. Also, we consider a walking cycle that takes 2 seconds. The finite-difference time-domain (FDTD) method is used as the electromagnetic problem solver. We randomly remove 45% of the data. Fig. 13 shows that the global signal reconstruction fails once again to depict the torso and limbs movements. The local signal reconstruction over short windows resembles the spectrogram with all uniformly sampled data available. This resemblance will lend itself to proper motion classification under missing or random observations. Applications of the same method to real human gait data was considered in [43].

It has been shown that the sinusoidal dictionary causes competing requirements on the window size. That is, including more observations for stable recovery leads to an increased bandwidth and, in turn, less sparsity. The chirp dictionary can break, or at least loosen, this adverse window length-sparsity interlocking. It is shown in [44, 45] that local frequency reconstruction using chirp atoms better represents the approximate piece-wise chirp behavior of most Doppler TF signatures. This enables the utilization of longer windows for accurate signal reconstruction. Because a constructed chirp spans the window length, we cannot designate the reconstruction results to a mid-window reference sample, as in the case

of sinusoidal basis. Accordingly, with highly overlapping sliding windows, one may avoid conflicting results by resorting to TF averaging so as to obtain a unique IF estimate [45].

This example demonstrates the performance of local reconstruction with chirp dictionary and sinusoid dictionary when applied to a three-component signal which contains two different signal types. We sample the data at the Nyquist rate, and then randomly discard some of the samples. The sampling frequency is $F_s = 256\text{Hz}$, the total signal length is $N = 256$, and only 50% of the data is used to estimate the IFs. A 70-point data window is used when computing the TF representation. The three-component signal is expressed as

$$x(n) = e^{j[(0.1F_s)\cos(2\pi n) + 2\pi(0.25F_s)n]} + e^{j[(0.1F_s)\cos(2\pi n) + 2\pi(0.35F_s)n]} \\ + e^{j[2\pi(0.4F_s)n - (0.3F_s)\frac{n^2}{2N}]} + \epsilon(n),$$

where $n = 0, 1, 2, \dots, N - 1$, with SNR = 30 dB. The result in Fig. 14 demonstrates that the local frequency reconstruction using sinusoid atoms fails to resolve the signal, whereas the chirp dictionary yields a desirable TF representation. Sparse reconstruction of local signal frequencies using a chirp dictionary is shown in [44, 45] to outperform the chirp transform, especially under compressed observations.

6.3. Multiple Window Reconstruction of Time-Frequency Distributions

Just as the single-window based local frequency reconstruction, the multiple-window reconstruction technique follows the second approach in Fig. 6 and is considered part of the sparse construction approach for nonstationary signals. It is inspired by the fact that multiple-window spectrograms outperform their single-window counterparts. One commonly used set of windows is the Hermite functions [79]. Another possibility is to use windows which are the dominant eigenvectors of a desired RID kernel [29]. The CWD could be approximated by a weighted sum of spectrograms where the weights are the respective eigenvalues of the corresponding kernel eigendecomposition. As an example, Fig. 15 shows

the eigenvalues and the first six dominant eigenvectors of the CWD kernel [47]. It is clear that the eigenvalues decay rapidly, and those after the 6th term are negligible. We use $L = 6$ terms in the sequel for a high-fidelity TFSR reconstruction. Note that $L = 6$ eigen-decompositions are data-independent and thus can be computed only once prior to QTFD reconstructions.

To develop the sparse reconstruction approach which parallels the multi-window spectrograms, we consider the linear model with the corresponding multiple-window STFT

$$\mathbf{s}_l(t) = \tilde{\mathbf{F}}\mathbf{x}_l(t), \quad l = 1, \dots, L, \quad (12)$$

where $\mathbf{s}_l(t)$ is the windowed data observations where the missing samples are removed, $\tilde{\mathbf{F}}$ is the inverse DFT matrix, and $\mathbf{x}_l(t)$ is the l th QTFD to be reconstructed. Eq. (12) represents an SMV when solved using sparse reconstruction techniques for each l separately. It can also define an MMV model, building on the approximately common support of $\mathbf{x}_l(t)$ in the TF domain for different l . In this case, each window defines a different data set, and jointly sparse signal recovery techniques can then be applied. These techniques include distributed greedy algorithms, such as simultaneous OMP and simultaneous subspace pursuit [80, 81, 82]. The common support characteristic and the corresponding sparse representation can also be cast as block sparsity or group sparsity, which is effectively solved using techniques that account for such property. Group sparsity solvers include block-sparsity based CS [68, 72], and multi-task BCS [63, 64, 73].

Simulation results are provided in Figs. 16 and 17 to demonstrate the effectiveness of the proposed multiple-window sparse reconstruction of TF signatures of nonstationary signals [47]. Both SVM- and MMV-based sparse reconstruction techniques are applied and compared. In the former, the TFSR is separately reconstructed using the SMV for each window l in (12), and the results are weighted with the respective eigenvalues and then

averaged. We consider the case where only 25% random time-domain samples of the data are available. Fig. 16 shows the observed signal waveform as well as its spectrogram and the CWD. Neither distribution shows clear TF signatures. The QTFD obtained from the MMV model through the CMT-BCS algorithm yields much cleaner results, as evident in Fig. 17(a), whereas the weighted sum of the individually reconstructed TFSR results, as depicted in Fig. 17(b), shows inferior TF signatures and a high level of undesired artifacts. For comparison, the sparse reconstruction using a single rectangular window is depicted in Fig. 17(c), which does not depict a clear signature of the chirp component either. It is worth noting that single and multiple window signature reconstructions do not suffer from vertical line type of clutter as those techniques using smoothing TF kernels.

In the next example, we observe a signal composed of two components with third degree polynomial phase. Results in Fig. 18(a) demonstrate the effectiveness of the MMV approach using common support. Single window reconstruction (Fig. 18(c)) suffers from gaps in the TF signature which occurred due to reconstruction errors. On the other hand, separate reconstruction of multiple windows in Fig. 18(b) provides blurred signature around the true IF law. In sparse reconstruction, we cannot fully quantify the performance gain of using multiple windows. However, we put forward the common approximate support property underpinning multiple windows as the reason behind the improved TF representation. For demonstration, we used the eigenvectors corresponding to the first, second, and fourth dominant eigenvalues, and the results are included in Fig. 18(d). As evident from this figure, the reconstructed TFSR exhibits a phenomenon resembling the ripple effect due to the support mismatch. Further investigation is required in order to gain insights into the improvements achieved by using MMV approach.

6.4. Structure-Aware Bayesian Compressive Sensing for Time-Frequency Distribution Reconstruction

When performing sparse reconstruction of the TFSR, we observe that the signal IFs show continuous signatures in the 2-D TF domain. As such, the reconstruction performance can be improved by utilizing the specific TFSR contiguity structure [46].

To encourage contiguity in the joint TF domain, we utilize data in the neighboring time instants. The coefficient of a pixel in the TF domain is expressed as the product of two variables, i.e., $x(t, f) = w(t, f)z(t, f)$, where $z(t, f)$ is a Bernoulli distributed binary number that determines whether the pixel takes a nonzero value, whereas $w(t, f)$ determines the value of the pixel when it is determined to be nonzero. The Bernoulli distribution parameters of $z(t, f)$ of a pixel at (t, f) are determined by the values of its neighboring pixels at time instants $t - 1$ and $t + 1$. More specifically, the continuous IF property allows us to place different priors to the distribution of $z(t, f)$ into several different patterns, prompting a nonzero entry when it falls between a single diagonal pair of nonzero neighboring pixels. For example, in [46], three-level patterns are used, i.e., “strong rejection”, “weak rejection”, and “strong acceptance”, as shown in Fig. 19. For this purpose, two variables, i.e., the number of neighboring pixels and the number of diagonal pairs with nonzero entries, are defined. Conceptually, strong acceptance is set when the neighboring pixels exhibit a single diagonal pattern (Fig. 19(c)). On the other hand, the prior is set to strongly reject the pixel under test when all the neighboring pixels are zero, or a high number neighboring diagonal pairs are present (Fig. 19(a)), and weakly reject the pixel when the neighboring pixels have a single nonzero entry (Fig. 19(b)).

For illustration, consider two parallel polynomial phase signals with 128 samples [46]. The component with a lower frequency spans samples 33 to 128, whereas the other one spans samples 1 to 96. Fig. 20(a) shows the TFSR results reconstructed from the IAF, kernalled with the AOK, using the OMP, where the sparsity is specified as 2 throughout

all time samples. Fig. 20(b) shows the results using the BCS without incorporating the continuous IF structure. It is observed that in both cases the TFSRs are not smoothly connected and there are spurious entries toward the two ends where the true sparsity is one. The results in Fig. 20(c) are obtained from structure-aware BCS, which clearly show smooth TFSR signatures with much less spurious entries.

6.5. Sparse Reconstruction from Instantaneous Autocorrelation Function

Early work on sparse reconstruction of the QTFDs focused on the ambiguity domain, oftenkerneled, as the domain of observations. It exploited the 2-D FFT relationship between the ambiguity domain and the TF domain [24]. Alternatively and more efficiently, QTFDs can be obtained through sparse reconstruction of data in the IAF domain, as pointed out in Section 5. In this case, the two domains of the IAF and the TF are related by a 1-D FFT and the sparse reconstruction is performed by utilizing the Fourier relationship between the autocorrelation lags and frequency spectrum for each time instant. This yields a number of advantages in terms of simplicity, flexibility, and performance [29]:

(1) The size of the dictionary matrix is significantly reduced and so is the overall solution complexity and the execution time. For an N -point FFT, the dictionary matrix corresponding to the 2-D FFT relationship requires a dimension of $N^2 \times N^2$, whereas that corresponding to the 1-D case is only $N \times N$. The latter also permits parallel computation of the QTFD for all time instants.

(2) It enables exploitation of local sparsity, instead of the global sparsity, for improved performance. Because the IAF and TF domains share the same time dimension, the QTFD is obtained individually for each time instant. As such, we can specify sparsity in the frequency domain entries for each time instant. This is particularly important for FM signals as the local sparsity specification avoids the cluttered QTFD entries around TF regions with strong QTFD presence. In Fig. 21, we plot the results corresponding to the same signals as in Fig. 20, but the results are reconstructed by OMP and BCS exploiting global sparsity rather

Table 1 - Absolute distance results for different approaches

MMV approach	Separate reconstruction	Single window reconstruction	Kernel reconstruction	Linear interpolation
0.3297	0.3568	0.3516	0.3343	0.3396

than local sparsity. For the OMP method, the total sparsity in Fig. 21(a) is the same as that in Fig. 20(a). It is evident that nonzero entries are reconstructed along the impulsive artifacts, whereas few nonzero entries are obtained toward the edges because of low power concentration as a result of the zero-padding in the TF analysis.

(3) IAF-based TFSR reconstruction also permits the exploitation of signal TFSR structures, such as the TF signature continuity with adjacent TF domain pixels, as discussed in Section 6.4. Such treatment, particularly the consideration of signal structures in a local time region, is difficult when reconstructing QTFD from the ambiguity domain.

6.6. Comparison of Different Approaches

In this section, we compare some of the previously described approaches applied to the noiseless signal composed of a chirp and a component with a third degree polynomial phase:

$$x(n) = e^{j51.2\pi(n/N-1)^3} + e^{j64\pi(n/N-1)^2},$$

where $N = 256$ and $n = 0, \dots, N - 1$. We assume that 50% of samples are randomly missing, and the results obtained from different TF reconstruction methods are shown in Fig. 22. All the results are compared with respect to the true TF signature, i.e., the IF law which is a derivative of the components phases. The absolute difference is used as a metric, and the results are given in Table 1.

We notice that the MMV approach and the kernel approach give the smallest errors. The worst result is obtained using the multiple windows with separate reconstructions. The single window approach shows gaps in the reconstructed TF signature, while linear interpolation contains more clutter when compared to the other techniques.

7. Parametric SA-TFSR

The parametric approach for sparse-aware TF analysis focuses on separation of multiple sinusoidal FM micro-Doppler components and the estimation of their parameters, such as the Doppler repetition period, the Doppler amplitude, and the initial phase. In this case, the sparse reconstruction may benefit from the fact that all components, belonging to radar reflection from the same rotating, vibrating, or oscillating targets share the same Doppler repetition period. This benefit may be realized by the pruned orthogonal matching pursuit (POMP) [27]. The radar echo from a coning target is given by

$$y(t) = \sum_{\kappa=1}^K a_{\kappa} e^{j \frac{4\pi}{\lambda} d_{\kappa} \sin(\omega t + \theta_{\kappa})}, \quad t = t_1, t_2, \dots, t_N, \quad (13)$$

where ω refers to the Doppler repetition frequency, defining the angular velocity of the target, a_{κ} is the complex reflectivity of the κ th scatterer, λ is the wavelength, d_{κ} is the Doppler amplitude and is dependent on the spatial position of the κ th scatterer, θ_{κ} is the initial phase, and K is the number of dominant scatterers. The model of sparse representation is expressed as

$$\mathbf{Y} = \Phi(\boldsymbol{\mu})\mathbf{X}, \quad (14)$$

where $\boldsymbol{\mu}$ is the unknown key parameter vector that represents the common characteristics of all nonzero entries in the sparse vector \mathbf{X} which contains information about the signal parameters. The desired solution of parametric sparse representation can be expressed as

$$\{\hat{\boldsymbol{\mu}}, \hat{\mathbf{X}}\} = \arg \min_{\boldsymbol{\mu}, \mathbf{X}} \|\mathbf{Y} - \Phi(\boldsymbol{\mu})\mathbf{X}\|_2^2 \leq \varepsilon_1, \quad (15)$$

where ε_1 is the noise threshold. The parametric sparse TF analysis approach adopts the analytic form of basis functions in (14) and, as such, accurately fits the received signal characteristics. This type of model matching represents a key advantage over the nonparametric

counterpart of SA-QTFDs. One approach to solve for the sparse \mathbf{X} is by discretizing the three parameter values $(\omega, d_\kappa, \theta_\kappa)$ in (13) and form a large size dictionary, which will likely be computationally prohibitive and, due to limited observations, may lead to unstable solution. A more efficient algorithm is based on separating the common frequency parameter from individual parameters in the signal decomposition process. Since the recovery error is more sensitive to the wrong value of ω than to the wrong values of d_κ, θ_κ , the authors in [27] attempt to reduce the computational costs associated with the choice of ω . In the first iteration all candidate values of ω are included. However, in subsequent iterations the set of candidates is reduced by excluding half of the ω values which produce the highest residual errors. In so doing, the problem size can be significantly reduced [27]. This algorithm was used to generate the following results.

Consider two sinusoidal FM signals with respective parameter sets of $(\omega, d_1, \theta_1)=[4\pi, 0.015, 1]$ and $(\omega, d_2, \theta_2) = [4\pi, 0.003, 3]$. Using (15), the three parameters of each sinusoid are properly estimated per Fig. 23(a). Fig. 23(b) depicts the TF signature of the two signals using the estimated parameters. However, as previously mentioned, a deviation from the model would cause this approach to fail and to become inferior to the nonparametric approach [44, 45].

8. Conclusions

The emergence of the areas of compressive sensing and sparse reconstruction has brought a new perspective to QTFD of nonstationary signals. In essence, the high power concentration in the TF domain casts these signals as sparse and, as such, invites sparse optimization techniques to play a fundamental role in TF signature estimations. This new paradigm has ignited the development of SA-QTFDs and motivated new approaches for designing TF kernels for revealing the local frequency behavior of multi-component FM signals. This paper categorized current research activities involving Fourier basis in SA-QTFD into those

approaches that aim at identifying new members of Cohen’s class of QTFDs and others which solve sparse optimization problems using Fourier dictionary. A hybrid approach is a one that uses the Fourier transform as well as a Fourier dictionary. In all three cases, the domain of complete or compressed observations may vary, but the common destination is the TF domain. The paper showed significant improvements offered by SA-QTFD over the traditional QTFD approaches which do not incorporate, directly or indirectly, sparsity in their formulations. It was shown that tradeoffs between temporal and spectral resolutions as well as TF signal contiguity are better served using block-sparsity and Bayesian CS techniques. There remain many stimulating problems in SA-QTFD which encourage further developments in this area covering several applications that involve nonstationary signals.

9. References

- [1] C. J. Law, G. Jones, D. C. Backer, W. C. Barott, G. C. Bower, C. Gutierrez-Kraybill, P. K. G. Williams, D. Werthimer, Millisecond imaging of radio transients with the pocket correlator, *The Astrophysical Journal* 742 (1) (2011) 1–13.
- [2] S. Sodagari, A. Khawar, T. C. Clancy, R. McGwier, A projection based approach for radar and telecommunication systems coexistence, in: *Proc. IEEE Globecom, Anaheim, CA, 2012*, pp. 5010–5014.
- [3] A. M. Zoubir, V. Koivunen, Y. Chakhchoukh, M. Muma, Robust estimation in signal processing, *IEEE Signal Process. Mag.* 29 (4) (2012) 61–80.
- [4] W. C. Knight, R. G. Pridham, S. M. Kay, Digital signal processing for sonar, *Proc. IEEE* 69 (11) (1981) 1451–1506.
- [5] T. Misaridis, J. A. Jensen, Use of modulated excitation signals in medical ultrasound. Part I: Basic concepts and expected benefits, *IEEE Trans. Ultrason., Ferroelectr., Freq. Control* 52 (2) (2005) 177–191.
- [6] C. Clemente, A. Balleri, K. Woodbridge, J. J. Soraghan, Developments in target micro-Doppler signatures analysis: radar imaging, ultrasound and through-the-wall radar, *EURASIP J. on Advances in Signal Process.* 2013 (1) (2013) 1–18.
- [7] M. G. Amin, L. Zhao, A. R. Lindsey, Subspace array processing for the suppression of FM jamming in GPS receivers, *IEEE Trans. Aerosp. Electron. Syst* 40 (1) (2004) 80–92.

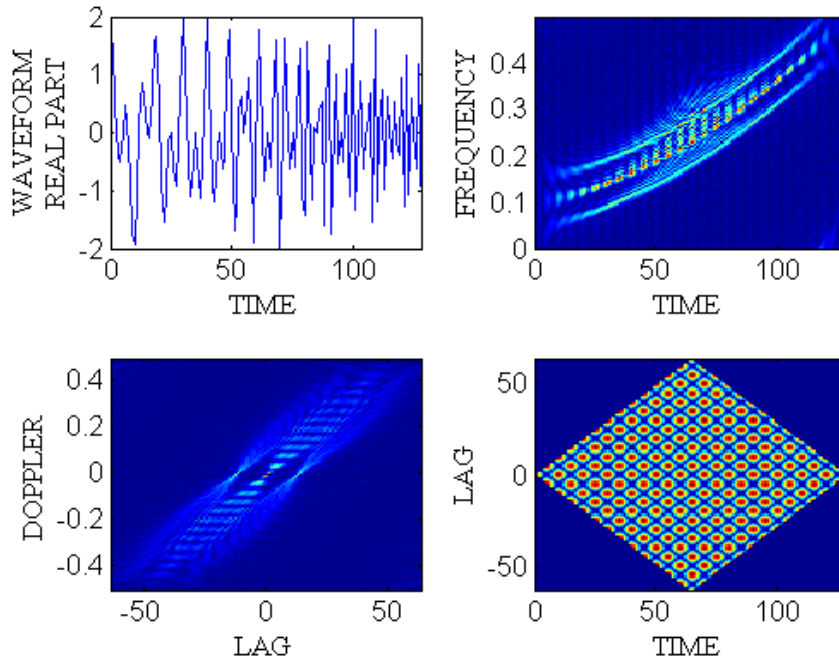
- [8] S. R. Lach, M. G. Amin, A. R. Lindsey, Broadband interference excision for software-radio spread-spectrum communications using time-frequency distribution synthesis, *IEEE J. Sel. Areas Commun.* 17 (4) (1999) 704–714.
- [9] M. G. Hall, A. V. Oppenheim, A. S. Willsky, Time-varying parametric modeling of speech, *Signal Process.* 5 (3) (1983) 267–285.
- [10] E. Shwedyk, R. Balasubramanian, R. Scott, A nonstationary model for the electromyogram, *IEEE Trans. Biomed. Eng.* (5) (1977) 417–424.
- [11] V. C. Chen, Analysis of radar micro-Doppler with time-frequency transform, in: *Proc. 10th IEEE Workshop Stat. Signal and Array Process.*, Pocono Manor, PA, 2000, pp. 463–466.
- [12] V. C. Chen, H. Ling, W. Miceli, Editorial: Time-frequency analysis for synthetic aperture radar and feature extraction, *IEE Proc. Radar, Sonar and Navigation* 150 (4) (2003) 201–202.
- [13] P. Setlur, M. Amin, F. Ahmad, Analysis of micro-Doppler signals using linear FM basis decomposition, in: *Proc. SPIE*, Vol. 6210, Orlando, FL, 2006.
- [14] Y. Kim, H. Ling, Human activity classification based on micro-Doppler signatures using a support vector machine, *IEEE Trans. Geosci. Remote Sens.* 47 (5) (2009) 1328–1337.
- [15] D. Tahmoush, J. Silvius, Radar micro-Doppler for long range front-view gait recognition, in: *IEEE Int. Conf. Biometrics: Theory, Applicat., and Systems (BTAS'09)*, Washington, D.C., 2009, pp. 1–6.
- [16] G. Smith, K. Woodbridge, C. Baker, H. Griffiths, Multistatic micro-Doppler radar signatures of personnel targets, *IET Signal Process.* 4 (3) (2010) 224–233.
- [17] L. Stankovic, I. Djurovic, T. Thayaparan, Separation of target rigid body and micro-Doppler effects in ISAR imaging, *IEEE Trans. Aerosp. Electron. Syst.* 42 (4) (2006) 1496–1506.
- [18] M. Amin, B. Jokanovic, T. Dogaru, Reconstruction of locally frequency sparse nonstationary signals from random samples, in: *Proc. European Signal Process. Conf. (EUSIPCO)*, Lisbon, Portugal, 2014, pp. 1771–1775.
- [19] D. L. Donoho, Compressed sensing, *IEEE Trans. Inf. Theory* 52 (4) (2006) 1289–1306.
- [20] E. J. Candès, J. Romberg, T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, *IEEE Trans. Inf. Theory* 52 (2) (2006) 489–509.
- [21] E. J. Candès, M. B. Wakin, An introduction to compressive sampling, *IEEE Signal Process. Mag.* 25 (2) (2008) 21–30.
- [22] J. A. Tropp, A. C. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit, *IEEE Trans. Inf. Theory* 53 (12) (2007) 4655–4666.

- [23] R. Tibshirani, Regression shrinkage and selection via the lasso, *J. Royal Stat. Society* (1996) 267–288.
- [24] P. Flandrin, P. Borgnat, Time-frequency energy distributions meet compressed sensing, *IEEE Trans. Signal Process.* 58 (6) (2010) 2974–2982.
- [25] B. Jokanović, M. Amin, S. Stanković, Instantaneous frequency and time-frequency signature estimation using compressive sensing, in: *Proc. SPIE*, Vol. 8714, Baltimore, MD, 2013.
- [26] L. Stankovic, I. Orovic, S. Stankovic, M. Amin, Compressive sensing based separation of non-stationary and stationary signals overlapping in time-frequency, *IEEE Trans. Signal Process.* 61 (18) (2013) 4562–4572.
- [27] G. Li, P. K. Varshney, Micro-Doppler parameter estimation via parametric sparse representation and pruned orthogonal matching pursuit, *IEEE J. Sel. Topics Appl. Earth Observ.* 7 (12) (2014) 4937–4948.
- [28] R. Sevimli, M. Tofghi, E. Cetin, Range-Doppler radar target detection using denoising within the compressive sensing framework, in: *Proc. European Signal Process. Conf. (EUSIPCO)*, Lisbon, Portugal, 2014.
- [29] Y. D. Zhang, M. G. Amin, B. Himed, Reduced interference time-frequency representations and sparse reconstruction of undersampled data, in: *Proc. European Signal Process. Conf. (EUSIPCO)*, Marrakech, Morocco, 2013, pp. 1–5.
- [30] B. Jokanovic, M. Amin, Reduced interference sparse time-frequency distributions for compressed observations, *IEEE Trans. Signal Process.*, under review.
- [31] P. Flandrin, *Time-frequency/time-scale analysis*, Academic Press, New York, NY, 1998.
- [32] B. Boashash (Ed.), *Time-Frequency Signal Analysis and Processing. A Comprehensive Reference*, 1st Edition, Elsevier, Oxford, U.K., 2003.
- [33] L. Cohen, Time-frequency distributions—a review, *Proc. IEEE* 77 (7) (1989) 941–981.
- [34] P. Flandrin, M. Amin, S. McLaughlin, B. Torresani, Time-frequency analysis and applications [from the guest editors], *IEEE Signal Process. Mag.* 30 (6) (2013) 19–150.
- [35] H. I. Choi, W. J. Williams, Improved time-frequency representation of multicomponent signals using exponential kernels, *IEEE Trans. Acoust. Speech Signal Process.* 37 (6) (1989) 862–871.
- [36] W. J. Williams, J. Jeong, Reduced interference time-frequency distributions, in: B. Boashash (Ed.), *Time-Frequency Signal Analysis: Methods and Applications*, Longman Cheshire, Melbourne, Australia, 1992, pp. 74–97.
- [37] D. L. Jones, R. G. Baraniuk, An adaptive optimal-kernel time-frequency representation, *IEEE Trans. Signal Process.* 43 (10) (1995) 2361–2371.

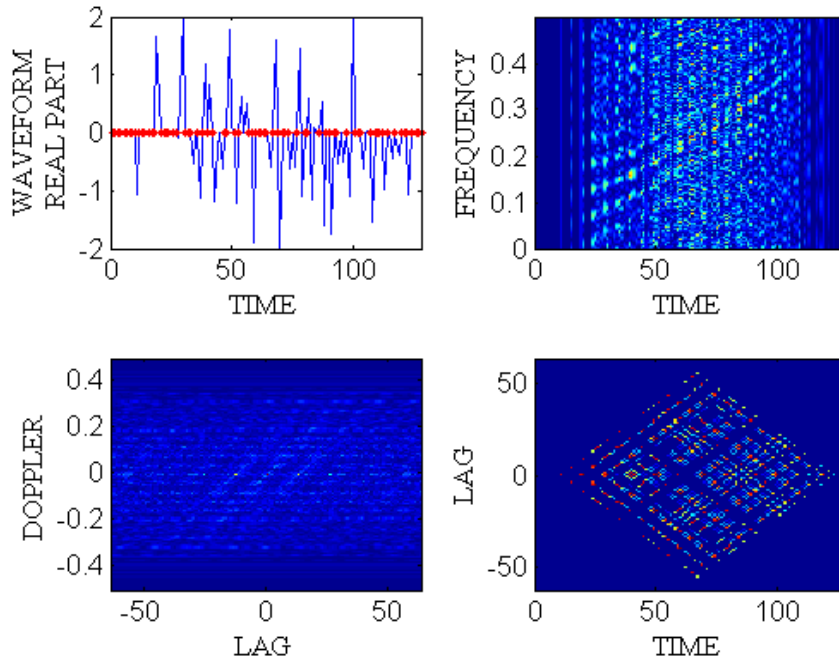
- [38] R. Baraniuk, P. Steeghs, Compressive radar imaging, in: Proc. IEEE Radar Conf., Waltham, MA, 2007, pp. 128–133.
- [39] M. A. Herman, T. Strohmer, High-resolution radar via compressed sensing, *IEEE Trans. Signal Process.* 57 (6) (2009) 2275–2284.
- [40] A. C. Gurbuz, J. H. McClellan, W. R. Scott, A compressive sensing data acquisition and imaging method for stepped frequency GPRs, *IEEE Trans. Signal Process.* 57 (7) (2009) 2640–2650.
- [41] L. C. Potter, E. Ertin, J. T. Parker, M. Cetin, Sparsity and compressed sensing in radar imaging, *IEEE Proc.* 98 (6) (2010) 1006–1020.
- [42] J. H. Ender, On compressive sensing applied to radar, *Signal Process.* 90 (5) (2010) 1402–1414.
- [43] M. G. Amin (Ed.), *Compressive Sensing for Urban Radar*, CRC Press, Boca Raton, FL, 2014.
- [44] Y. Nguyen, M. Amin, M. Ghogho, D. McLernon, Time-frequency signature sparse reconstruction using chirp dictionary, in: Proc. SPIE, Baltimore, MD, Apr. 2015.
- [45] M. Amin, Y. Nguyen, M. Ghogho, Local Doppler frequency sparse reconstructions using chirp atoms, in: Proc. IEEE Radar Conf. (Radarcon), May, 2015.
- [46] Q. Wu, Y. D. Zhang, M. G. Amin, Continuous structure based Bayesian compressive sensing for sparse reconstruction of time-frequency distributions, in: Proc. Int. Conf. Digital Signal Process. (DSP), Hong Kong, China, 2014, pp. 831–836.
- [47] M. G. Amin, Y. D. Zhang, B. Jukanovic, Time-frequency signature reconstruction from random observations using multiple measurement vectors, in: Proc. IEEE ICASSP, Florence, Italy, 2014.
- [48] M. Bayram, R. G. Baraniuk, Multiple window time-frequency and time-scale analysis, in: Proc. SPIE, Denver, CO, 1996, pp. 174–185.
- [49] G. Fraser, B. Boashash, Multiple window spectrogram and time-frequency distributions, in: Proc. IEEE ICASSP, Vol. 4, Adelaide, Australia, 1994, pp. IV–293.
- [50] F. Cakrak, P. J. Loughlin, Multiple window non-linear time-varying spectral analysis, in: Proc. IEEE ICASSP, Vol. 4, Seattle, WA, 1998, pp. 2409–2412.
- [51] M. Hansson, Multiple window decomposition of time-frequency kernels using a penalty function for suppressed sidelobes, in: Proc. IEEE ICASSP, Toulouse, France, May 2006, pp. 2883–2886.
- [52] W. J. Williams, S. Aviyente, Optimum window time-frequency distribution decompositions, in: Proc. Asilomar Conf., Vol. 1, Pacific Grove, CA, 1998, pp. 817–821.
- [53] M. G. Amin, Spectral decomposition of time-frequency distribution kernels, *IEEE Trans. Signal Process.* 42 (5) (1994) 1156–1165.

- [54] J. W. Pitton, Time-frequency spectrum estimation: An adaptive multitaper method, in: Proc. IEEE-SP Int. Symp. Time-Frequency and Time-Scale Analysis, IEEE, Pittsburgh, PA, 1998, pp. 665–668.
- [55] I. Orović, N. Žarić, S. Stanković, M. Amin, A multiwindow time-frequency approach based on the concepts of robust estimate theory, Proc. IEEE ICASSP (2011) 3584–3587.
- [56] A. Moghtaderi, P. Borgnat, P. Flandrin, Gap-filling by the empirical mode decomposition, in: Proc. ICASSP, Kyoto, Japan, 2012, pp. 3821–3824.
- [57] B. Jokanovic, M. Amin, T. Dogaru, Time-frequency signal representations using interpolations in joint-variable domains, IEEE Geosci. Remote Sens. Lett. 12 (1) (2015) 204–208.
- [58] W. Martin, P. Flandrin, Wigner-Ville spectral analysis of nonstationary processes, IEEE Trans. Acoust., Speech, Signal Process. 33 (6) (1985) 1461–1470.
- [59] S. Hearon, M. G. Amin, Minimum variance time-frequency distribution kernels, IEEE Trans. Signal Process. 43 (5) (1995) 1258–1262.
- [60] A. M. Sayeed, D. L. Jones, Optimal kernel design for nonstationary spectrum estimation, IEEE Trans. Signal Process. 43 (2) (1995) 478–491.
- [61] A. Jung, G. Taubock, F. Hlawatsch, Compressive spectral estimation for nonstationary random processes, IEEE Trans. Inform. Theory 59 (5) (2013) 3117–3138.
- [62] S. Ji, Y. Xue, L. Carin, Bayesian compressive sensing, IEEE Trans. Signal Process. 56 (6) (2008) 2346–2356.
- [63] S. Ji, D. Dunson, L. Carin, Multitask compressive sensing, IEEE Trans. Signal Process. 57 (1) (2009) 92–106.
- [64] Q. Wu, Y. D. Zhang, M. G. Amin, B. Himed, Complex multitask Bayesian compressive sensing, in: Proc. IEEE ICASSP, Florence, Italy, 2014.
- [65] B. Barkat, B. Boashash, A high-resolution quadratic time-frequency distribution for multicomponent signals analysis, IEEE Trans. Signal Process. 49 (10) (2001) 2232–2239.
- [66] R. Baraniuk, M. Davenport, R. DeVore, M. Wakin, A simple proof of the restricted isometry property for random matrices, Constructive Approximation 28 (3) (2008) 253–263.
- [67] L. Jacob, G. Obozinski, J.-P. Vert, Group Lasso with overlap and graph Lasso, in: Proc. Int. Conf. on Machine Learning, Montreal, QC, Canada, pp. 433–440.
- [68] M. Yuan, Y. Lin, Model selection and estimation in regression with grouped variables, Roy. Statist. Soc. B 68 (1) (2006) 49–67.
- [69] S. F. Cotter, B. D. Rao, K. Engan, K. Kreutz-Delgado, Sparse solutions to linear inverse problems with

- multiple measurement vectors, *IEEE Trans. Signal Process.* 53 (7) (2005) 2477–2488.
- [70] M. E. Tipping, Sparse Bayesian learning and the relevance vector machine, *J. machine learning research* 1 (2001) 211–244.
- [71] L. Poli, G. Oliveri, F. Viani, A. Massa, MT-BCS-based microwave imaging approach through minimum-norm current expansion, *IEEE Trans. Antennas Propag.* 61 (9) (2013) 4722–4732.
- [72] Q. Wu, Y. D. Zhang, F. Ahmad, M. G. Amin, Compressive sensing based high-resolution polarimetric through-the-wall radar imaging exploiting target characteristics, *IEEE Antennas Wireless Propag. Lett.* 14 (2015) 1043–1047.
- [73] Q. Wu, Y. Zhang, M. Amin, B. Himed, Multi-task Bayesian compressive sensing exploiting intra-task dependency, *IEEE Signal Process. Lett.* 22 (4) (2015) 430–434.
- [74] D. Angelosante, G. Giannakis, N. Sidiropoulos, Sparse parametric models for robust nonstationary signal analysis: Leveraging the power of sparse regression, *IEEE Signal Process. Mag.* 30 (6) (2013) 64–73.
- [75] L. A. Cirillo, A. M. Zoubir, M. G. Amin, Parameter estimation for locally linear FM signals using a time-frequency Hough transform, *IEEE Trans. Signal Process.* 56 (9) (2008) 4162–4175.
- [76] M. Schmidt, G. Fung, R. Rosales, Fast optimization methods for l1 regularization: A comparative study and two new approaches, in: *Proc. European Conf. Machine Learning, Warsaw, Poland, 2007*, pp. 286–297.
- [77] P. Addesso, M. Longo, S. Marano, V. Matta, I. Pinto, M. Principe, Sparsifying time-frequency distributions for gravitational wave data analysis, in: *Proc. Int. Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing, Pisa, Italy, 2015*, pp. 160–164.
- [78] D. Ba, B. Babadi, P. L. Purdon, E. N. Brown, Robust spectrotemporal decomposition by iteratively reweighted least squares, *Proc. of the National Academy of Sciences* 111 (50) (2014) 5336–5345.
- [79] B. Jokanovic, M. G. Amin, Y. D. Zhang, Reconstruction of multi-window time-frequency representation based on Hermite functions, in: *Proc. SPIE, Vol. 9109, Baltimore, MD, 2014*.
- [80] J. A. Tropp, A. C. Gilbert, M. J. Strauss, Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit, *Signal Processing* 86 (3) (2006) 572–588.
- [81] D. Baron, M. Duarte, S. Sarvotham, M. B. Wakin, R. G. Baraniuk, Distributed compressed sensing, preprint, 2009. Online available <http://arxiv.org/abs/0901.3403>.
- [82] J.-M. Feng, C.-H. Lee, Generalized subspace pursuit for signal recovery from multiple-measurement vectors, in: *IEEE Wireless Commun. and Networking Conf., 2013*, pp. 2874–2878.



(a)



(b)

Figure 1: Real-part signals, WVD (time vs. frequency), AF (lag vs. Doppler), and IAF (time vs. lag) of the two-component signals. (a) with full data; (b) with 60% of missing data.

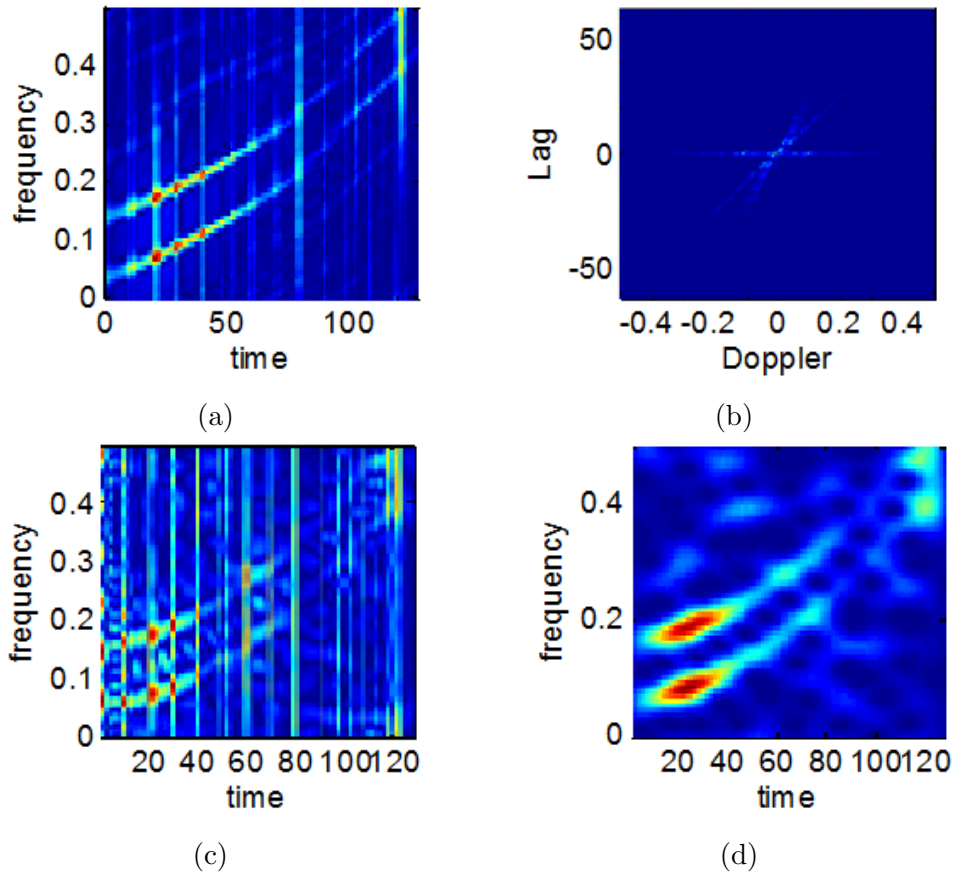


Figure 2: QTFD of missing data. (a) QTFD with an adaptive optimal kernel; (b) AF after the AOK is multiplied; (c) the CWD; and (d) the spectrogram.

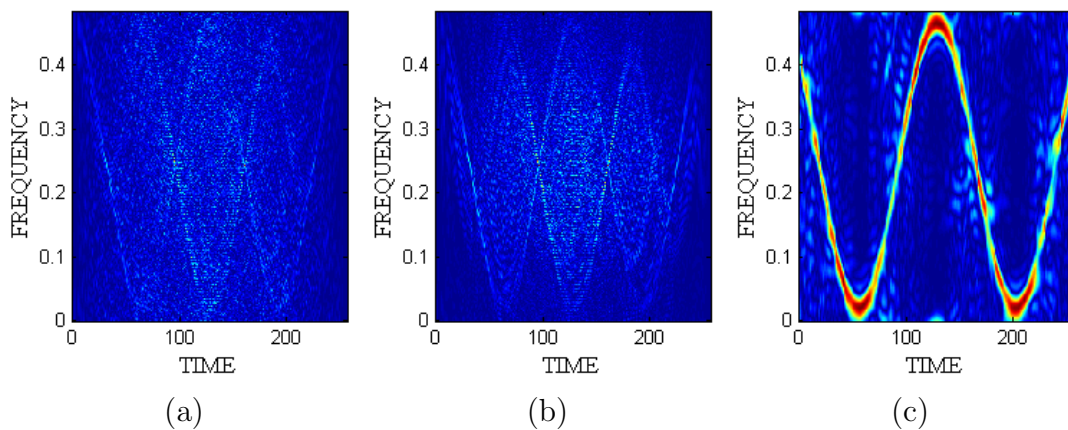


Figure 3: TFSRs of an incomplete sinusoidal FM signal: (a) the WVD obtained by using incomplete data; (b) the WVD obtained after using linear interpolation in the time domain; (c) the WVD obtained after using linear interpolation along the time axis in the IAF.

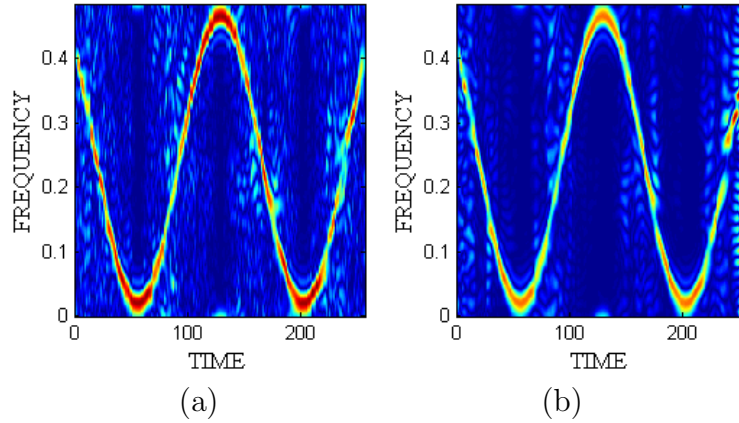


Figure 4: Other interpolation methods for an incomplete sinusoidal FM signal: (a) Nearest neighbor; (b) Splines.

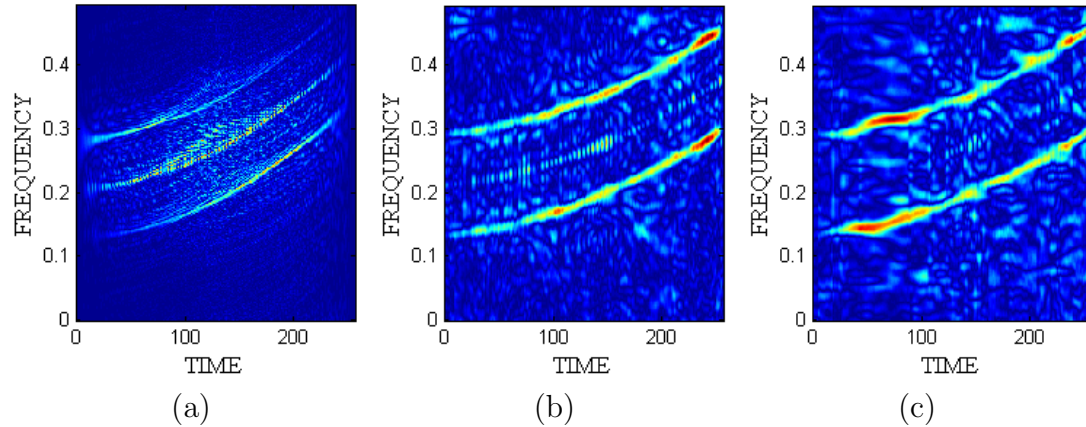


Figure 5: TFSRs of a randomly under-sampled multi-component signal: (a) the WVD obtained after using linear interpolation in the time domain; (b) the WVD obtained after using linear interpolation along the time axis in the IAF; (c) the WVD obtained after using a linear interpolation along the time axis in the IAF with 35% of data samples present.

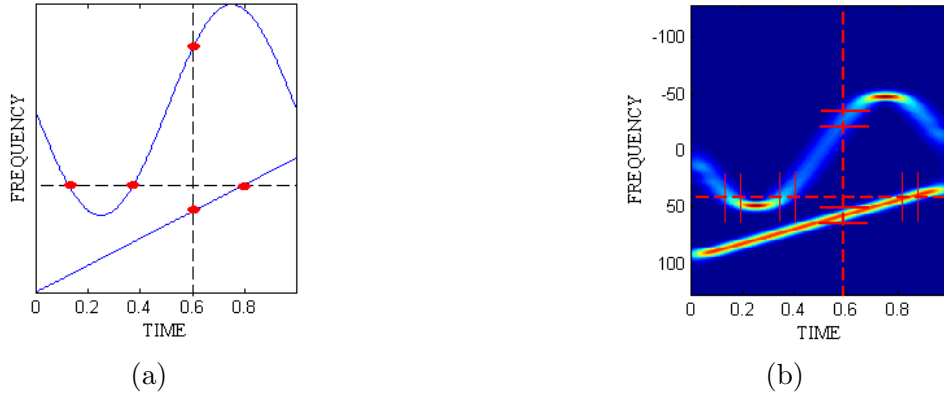


Figure 6: TFSR for a given multi-component signal. (a) ideal distribution; (b) spectrogram.

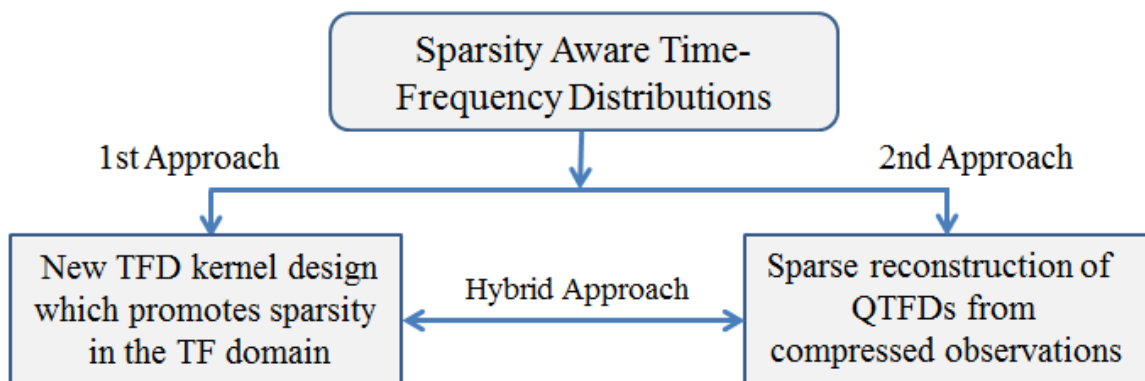


Figure 7: Dual-based approach for SA-TFSRs.

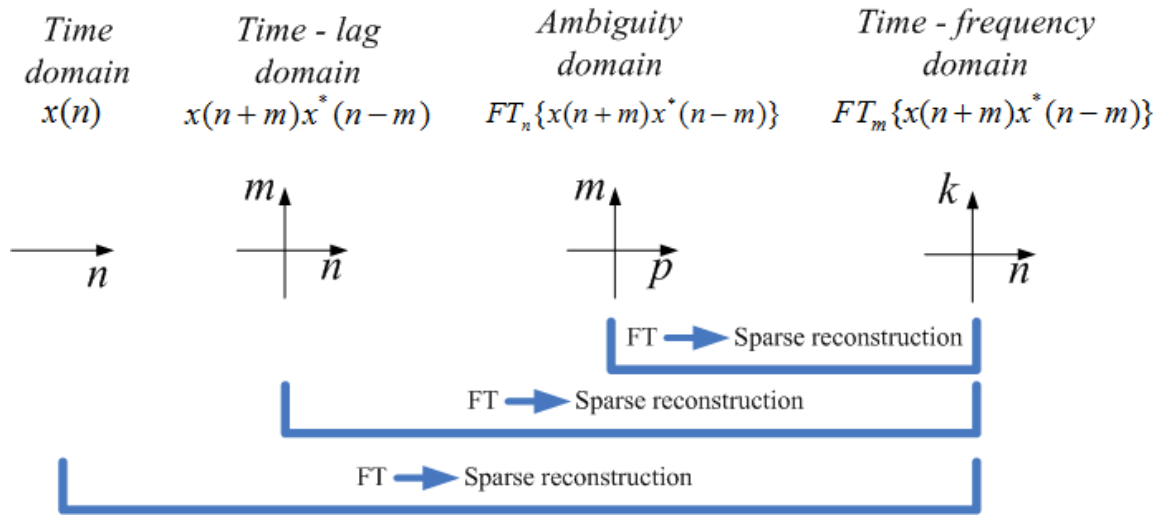


Figure 8: TFSR obtained from different domains through sparse reconstruction.

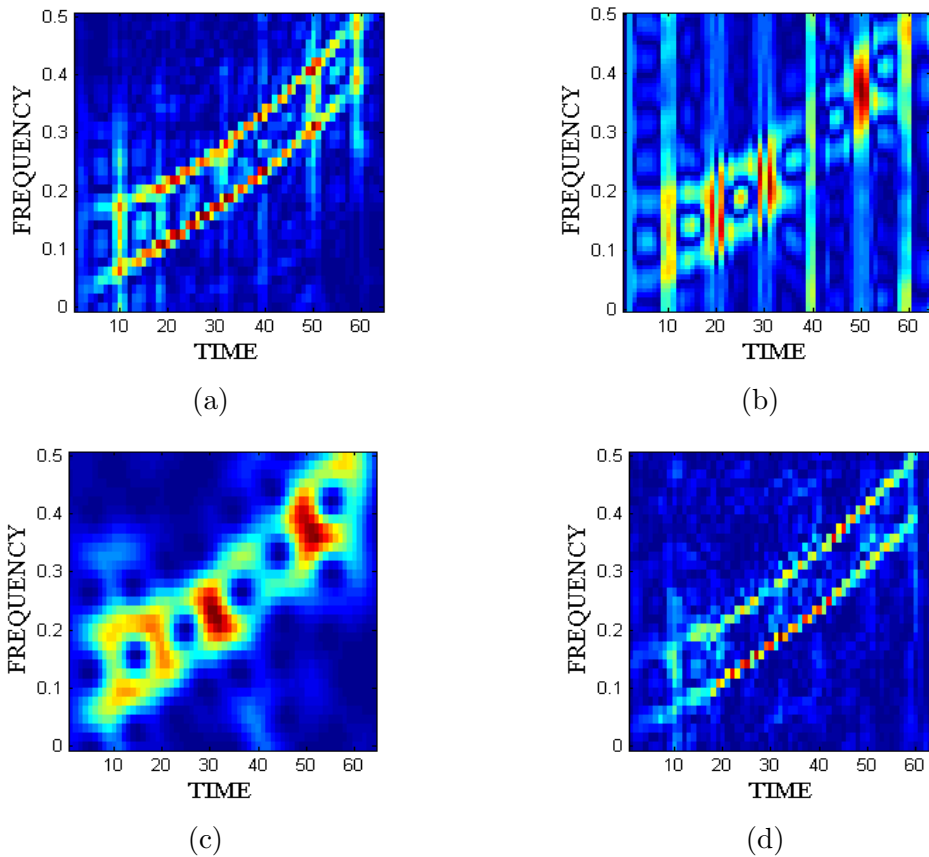


Figure 9: RIDs of a signal consisting of two polynomial phase signals when 50% of the data is missing. (a) adaptive optimal kernel; (b) the CWD; (c) the spectrogram; (d) our approach.

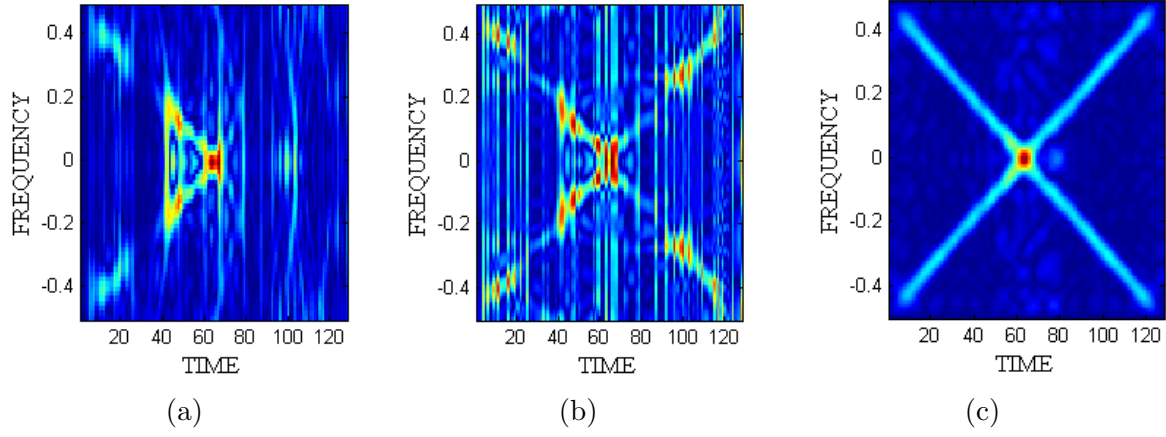


Figure 10: TFSRs of a signal consisting of two chirps: (a) AOK based distribution of compressed data; (b) the CWD of compressed data; (c) the sparsity aware kernel approach.

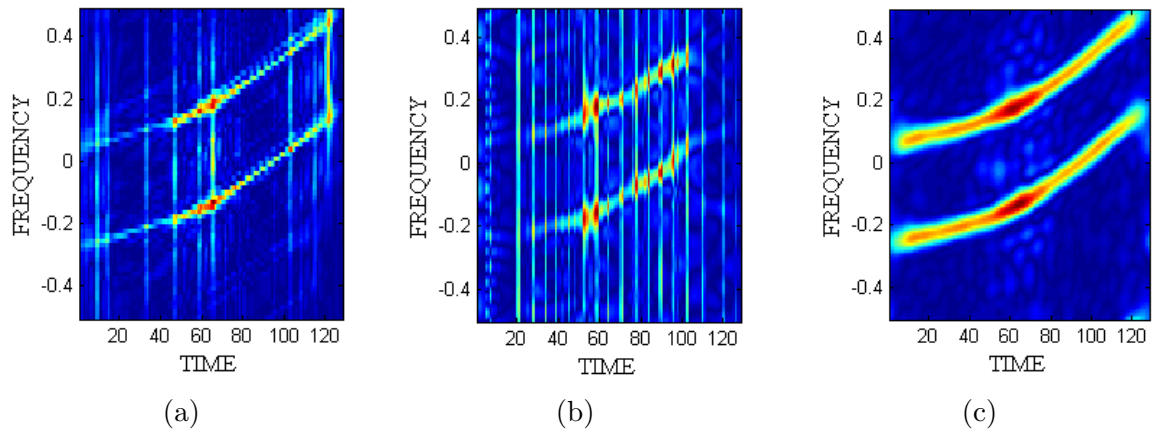


Figure 11: TFSRs of the signal considered in Fig. 5: (a) AOK based distribution of compressed data; (b) the CWD of compressed data; (c) the sparsity aware kernel approach.

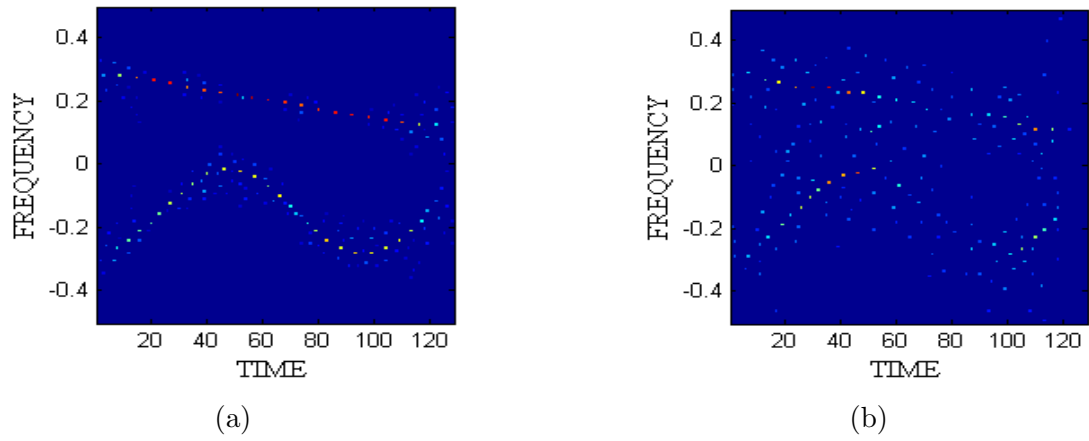


Figure 12: Sparse reconstruction using the measurements in the ambiguity domain: (a) reconstructed TFSR when using complete data; (b) reconstructed TFSR when 50% of samples are missing.

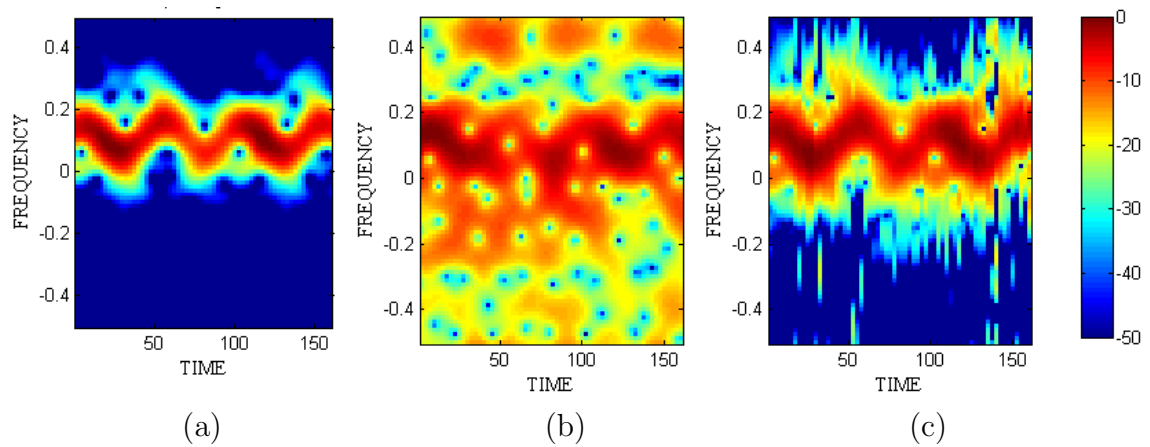
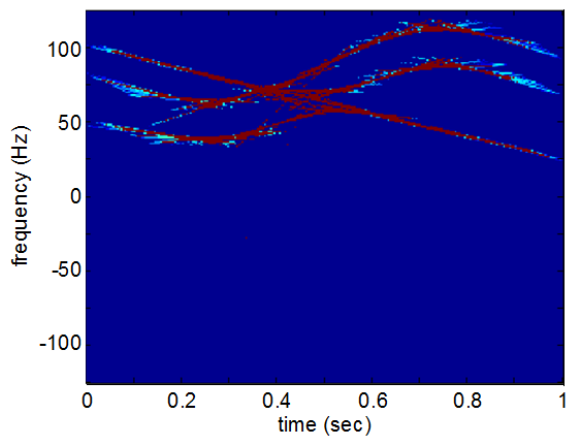
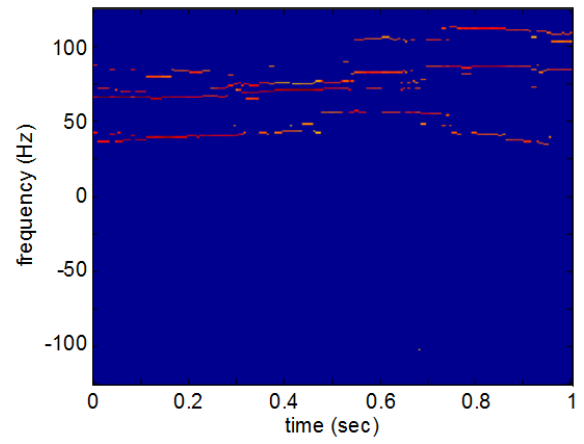


Figure 13: (a) Spectrogram obtained from data with no missing samples; (b) global reconstruction; (c) local reconstruction of the human gait signal when 45% of the data is missing.

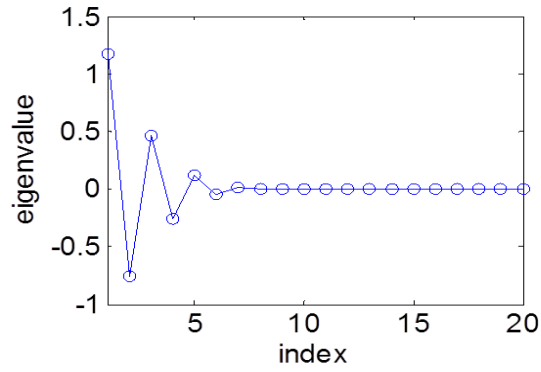


(a)

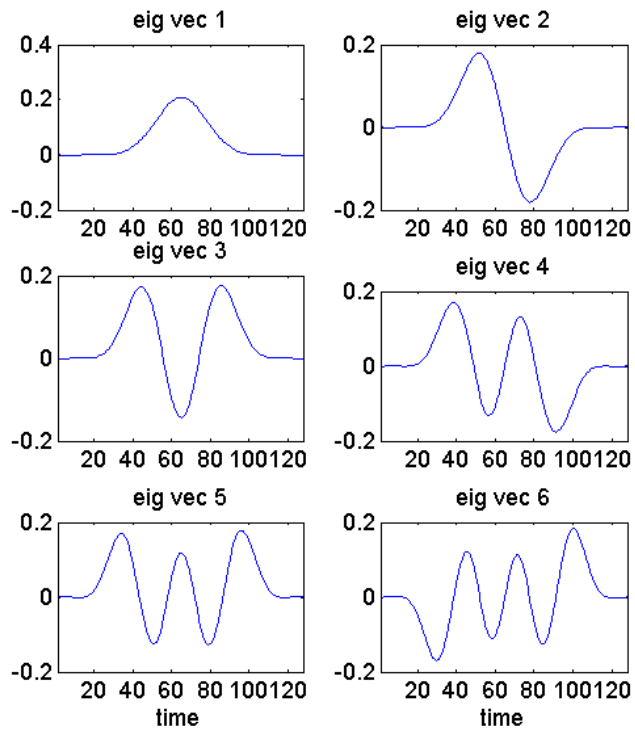


(b)

Figure 14: Local reconstruction of a three-component signal when 50% of the data is missing using (a) chirp atoms; (b) sinusoid atoms.



(a)



(b)

Figure 15: Eigenvalues and eigenvectors of the rotated kernel. (a) The first 20 dominant eigenvalues of the rotated autocorrelation-domain kernel, sorted by the magnitude; (b) The eigenvectors of the first six dominant terms.

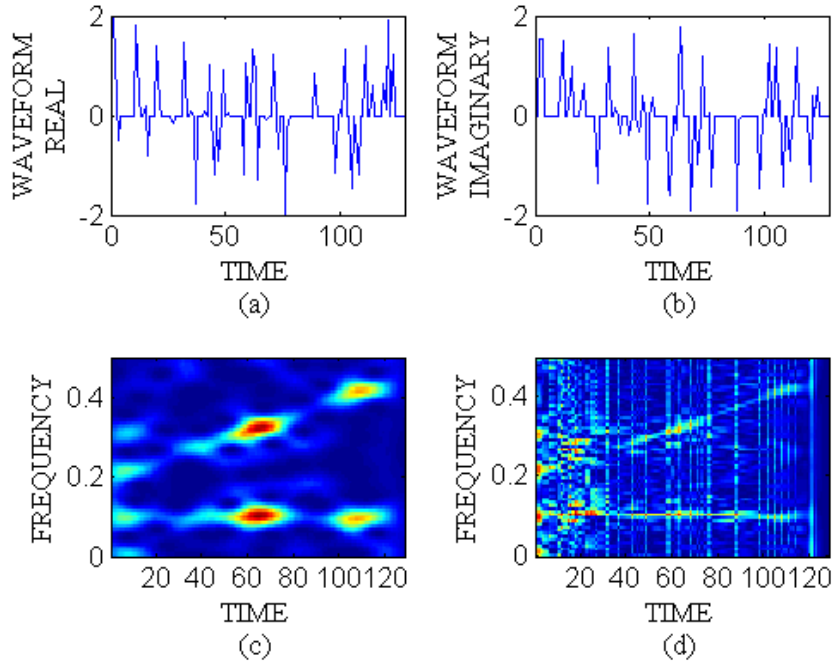


Figure 16: Signal waveform of 25% observation data with its spectrogram and CWD. (a) real-part of waveform; (b) imaginary-part of waveform; (c) spectrogram; and (d) CWD.

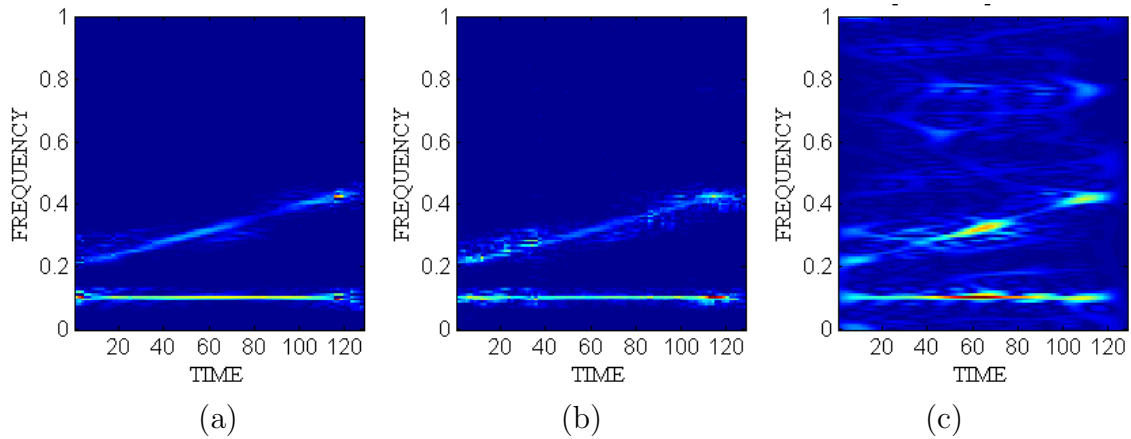


Figure 17: TFSR reconstructed using 25% of observed data. (a) sparse reconstruction from MMV approach; (b) sum of the six individually reconstructed TFSRs; and (c) sparse reconstruction using a single rectangular window.

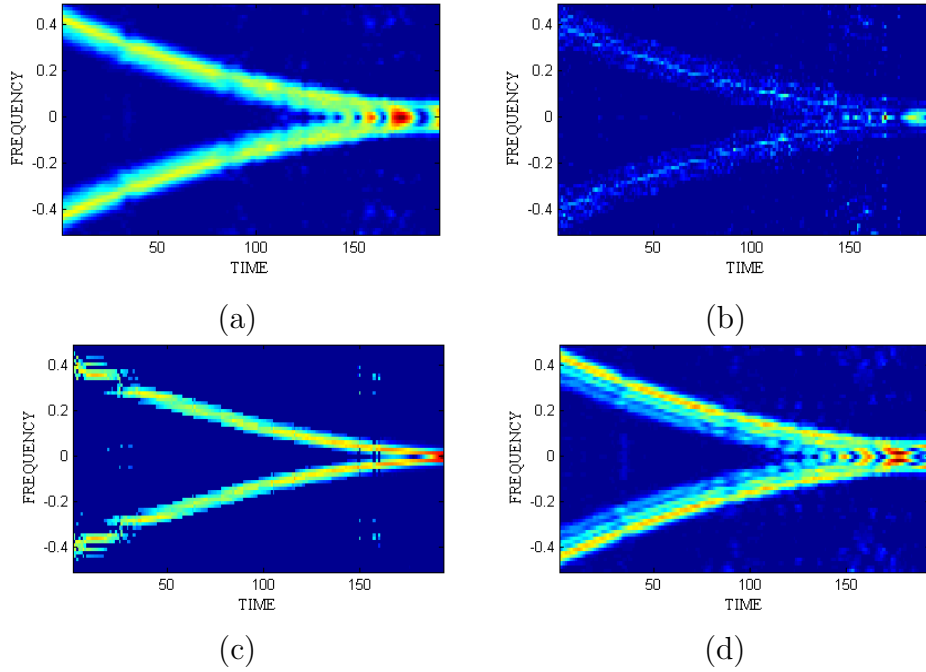


Figure 18: TFSR reconstructed using 50% of observed data. (a) sparse reconstruction from MMV approach; (b) sum of the four individually reconstructed TFSRs; and (c) sparse reconstruction using a single gaussian window. (d) sparse reconstruction using MMV approach when first, second and fourth eigenvectors are used as windows.

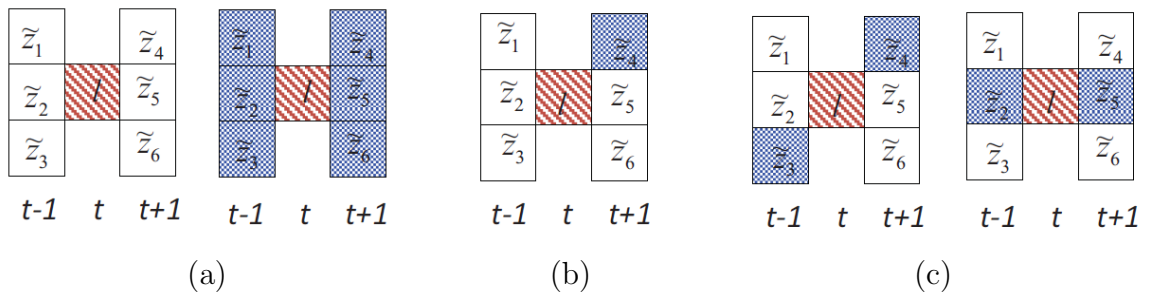


Figure 19: Examples of TFSR patterns. White squares denote entries with zero entries, blue shaded squares denote the entries with nonzero amplitudes, and red shaded squares denote the entry under test. (a) strong rejection; (b) weak rejection; (c) strong acceptance.

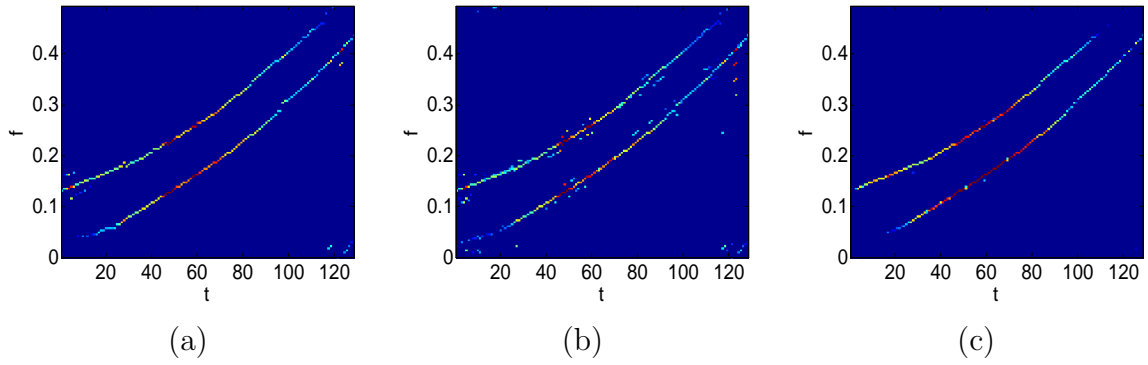


Figure 20: TF SRs reconstructed using (a) OMP; (b) BCS; and (c) structure-aware BCS.

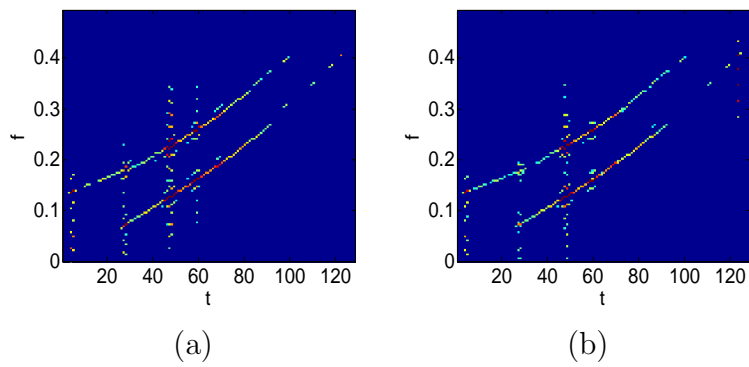


Figure 21: TF SRs reconstructed exploiting global sparsity by (a) OMP; (b) BCS.

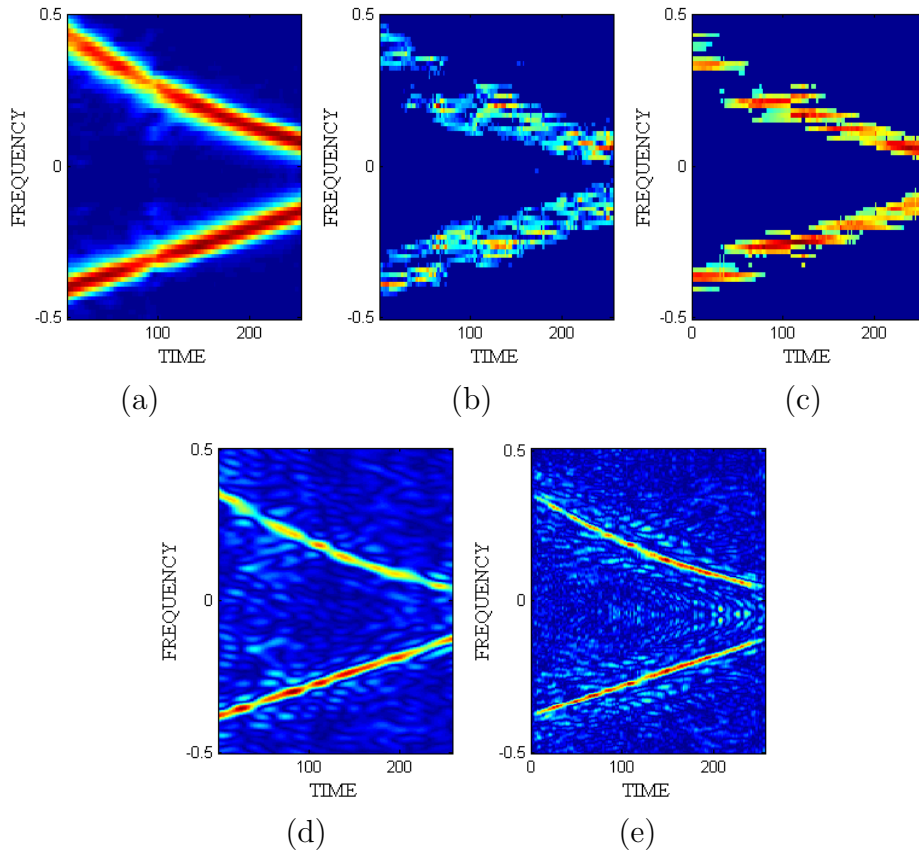


Figure 22: Different techniques applied to the same data (a) sparse reconstruction using MMV approach; (b) sum of the three individually reconstructed TF SRs; (c) sparse reconstruction using a single window; (d) kernel approach; and (e) linear interpolation.

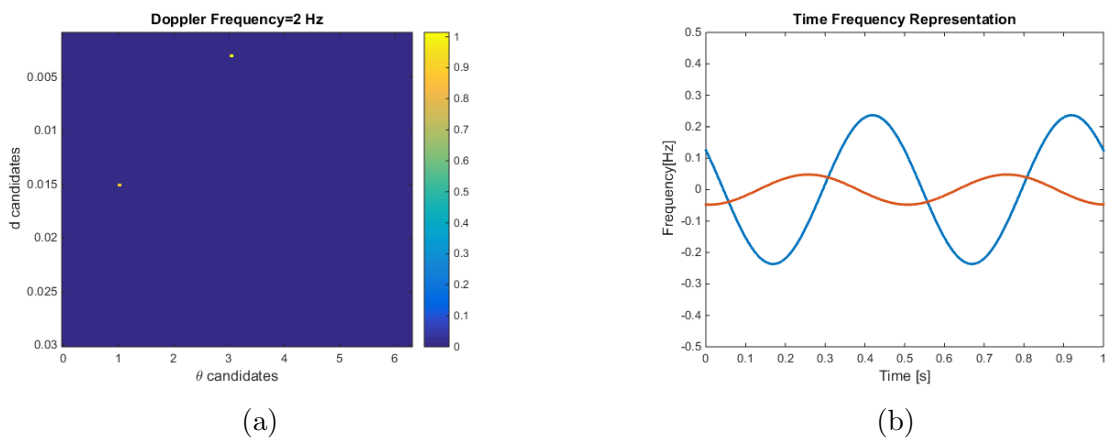


Figure 23: (a) Estimated parameters; (b) TF signature of the two signals using the estimated parameters.