Reduced Interference Sparse Time-Frequency Distributions for Compressed Observations

Branka Jokanović, *Student Member, IEEE,* and Moeness Amin, *Fellow, IEEE*

***Abstract*—Traditional quadratic time-frequency distributions are not designed to deal with randomly undersampled signals or data with missing samples. The compressed data measurements introduce noise-like artifacts in the ambiguity domain, compounding the problem of separating the signal auto-terms and cross-terms. In this paper, we propose a multi-task kernel design for suppressing both the artifacts and the cross-terms, while preserving the signal desirable auto-terms. The proposed approach results in highly concentrated time-frequency signature. We evaluate our approach using various polynomial phase signals and show its benefits, especially in the case of strong artifacts.**

***Index Terms*—Ambiguity domain, reduced interference distribution, sparsity, time-frequency kernel design.**

I. Introduction

MOST of the signals encountered in practice have time-varying spectra. Examples of nonstationary signals include microDoppler signals, ECG, EEG and speech. Traditionally, joint time-frequency (TF) signal representations are used for the analysis of these signals [1]-[6]. Due to the various applications in which non-stationary signals arise, it is difficult to define a single time-frequency signal representation (TFSR) which provides an ideal TF signature for all signal types.

The short-time Fourier transform (STFT) is the simplest linear TFSR [7], [8]. This transform is obtained by computing the Fourier transform over a sliding window in time. The square modulus of the STFT is the spectrogram. The major drawback of the spectrogram is that its efficiency depends on the employed window size and shape. Improvement in resolution can be achieved through the use of Wigner-Ville distribution (WVD), which is obtained as the Fourier transform of the signal instantaneous autocorrelation function (IAF). WVD is an ideal TFSR for linearly frequency modulated (FM) signals, i.e., chirps. However, due to the data bilinear products, WVD suffers from cross-terms when multi-component signals are analyzed. In order to suppress the cross-terms in WVD, a class of reduced interference distributions (RID) has been defined [9] which belongs to the general Cohen's class of quadratic time-frequency distributions (QTFDs) [10]. These distributions are better explained in the ambiguity domain where desirable signal auto-terms are generally located near the origin and along the two axes, in contrast with the cross-terms. Kernels with low-pass filter characteristics reduce cross-terms while preserving auto-terms. Accordingly, the two-dimensional (2D) Fourier transform of the kernelled ambiguity function yields a cross-terms mitigated TFSR.

The exponential growth of data demands new ways of collecting, representing and analyzing samples. In recent years, compressive sensing, dealing with few observations, has attracted a widespread interest [11]-[13]. Missing observations or random sampling in the field of radar can be due to range ambiguity, discarding noisy measurements, hardware simplification, sampling frequency limitations, logistical restrictions on data collection/storage, or co-existence of various wireless services with active or passive sensing platforms [14]-[16]. The possibility of acquiring less measurements than in the case of Nyquist is also beneficial in many other fields, such as seismology, biomedicine and astronomy [17]-[20]. QTFDs are traditionally defined to deal with data which is Nyquistly sampled or over sampled.

Consider the following linear model,

$$\mathbf{y} = \mathbf{\Psi}\mathbf{s}, \tag{1}$$

where $\mathbf{\Psi}$ maps the Nyquistly sampled data $\mathbf{s}$ into compressed measurements given by vector $\mathbf{y}$. Since (1) is an underdetermined system, additional information is necessary for finding a solution. Compressive sensing observes the case when the data is sparse in some transform domain.

Stationary signals and sparsity in the frequency domain have been extensively studied in the literature [21]-[23]. However, the majority of non-stationary data are not sparse in the frequency domain, i.e., they occupy large bandwidths. When observed in the joint time-frequency domain, these signals can be considered sparse. Fig. 1 depicts ideal TFSRs of a mono-component and multi-component signals. It is evident that most of the TF domain is vacant. Sparsity in the TF plane can present itself over the entire 2D plane or locally (for each time instant or frequency).

A. Related work

Several approaches dealing with sparse TFSRs have recently emerged [24]-[36]. Since there is a significant variety between some of these approaches, it is not simple to provide a unified framework. One way to divide these methods is according to their parametric/non-parametric perspective. Works reported in [29], [37] pertain to the parametric approaches. They operate on the time-domain data to estimate the signal parameters. The estimation can be performed either globally, incorporating all data, or locally, over short data segments. This parametric perspective is motivated by the need to perform classifications based on features related to the signal estimated parameters.

Authors are with the Center for Advanced Communications, Villanova University, Villanova, PA 19085, USA e-mail: (branka.jokanovic, moeness.amin@villanova.edu).

Manuscript received November xx, 2014; revised Month xx, 2015.

However, this approach requires known signal structures, e.g., sinusoidal FM or chirps. As such, it works well when there is a good match between the assumed and actual signal characteristics, but becomes sensitive to deviations in the assumed model. As an example of a non-parametric approach, we refer to the work presented in [24] where sparsity is observed over the TF plane. The compressed measurements are defined in the ambiguity domain through the application of an appropriate mask. This approach does not consider the case of random or missing samples in time. Due to incomplete data, the signal ambiguity function becomes contaminated by noise-like artifacts which render any masked ambiguity domain data unreliable. This, subsequently, yields undesirable TFSR, whether produced by sparse reconstruction or 2D FT. An alternative approach is to perform the TF reconstruction operating directly on the data and then, through the use of ambiguity domain filtering, reduce the presence of cross-terms. This process requires sequential operations which can be time consuming, considering the fact that we are dealing with 2D representations of data. Another drawback is that in reconstructing the TFSR, we are, in essence, recovering the cross-terms which will later be discarded through some kernel application.

For dictionary composed of Gabor atoms, conditions which guarantee successful reconstruction with high probability were derived in [34]. Approaches based on Gabor dictionary stem from the linear TFSRs, while in our work we focus on developing quadratic TFSR. The latter are more suitable for polynomial phase signals [38]. The authors in [36] develop a spectrotemporal pursuit and explicitly impose conditions for smoothness in time and sparsity in frequency. Using the link between the sparsity-promoting priors and the expectation-maximization method [39], an efficient algorithm for solving the estimation problem is proposed. However, as noted in [36], some a priori knowledge of the spectrum behavior is needed in order to choose the window size.

Most of the aforementioned methods use fixed dictionaries, and thus they are not fully data-driven. On the other hand, methods like Empirical mode decomposition (EMD) and Synchrosqueezed wavelet transforms use dictionaries which are learned in the decomposition process [40], [41]. Even though these data-driven approaches have been used in practice, there is still a lack of strong mathematical understanding about the entire process. Some of the issues, such as the "beating" phenomenon, are analyzed for both EMD and Synchrosqueezed transform in [42], [43]. In [35], the authors combine EMD and compressive sensing to obtain sparse TFSR. This work offers several advantages over traditional EMD, but some difficulties remain, as reported in [44]. Another EMD-based work that deals with gaps in the data is presented in [41]. Upon dividing signal into sub-signals without gaps, intrinsic mode functions (IMFs) are computed and then interpolated. The rationale given is that it is easier to interpolate modes of the signal, i.e., IMFs, than interpolating the signal itself. This approach becomes inefficient when dealing with highly interrupted signals underlined by random missing samples. In this case, there would be many sub-signals, some of which would contain only few samples making the EMD process less applicable.

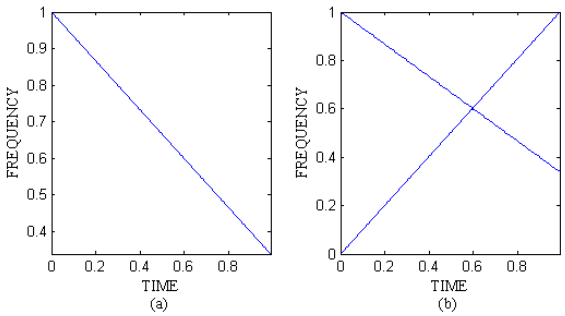

Fig. 1. Ideal TFSRs: (a) Chirp; (b) Two crossing chirps. Axes are normalized.

B. Main contribution

In this paper, we pursue the non-parametric perspective by introducing sparsity-aware TF kernels in Cohen's class. We analyze the effect of missing samples in the joint-variable domain and show that this effect manifests itself in terms of artifacts which follow certain pattern that does not favor the application of signal independent kernels. We then present a new class of data-dependent TF kernels that lead to reduced interference sparse distributions (RISD). These kernels aim at improving sparsity in the TF domain while reducing interference. The kernel design interplay auto-term preservation, cross-term suppression, and sparsity of the TFSR and decides on the low-pass filter characteristics as an optimum tradeoff between these objectives. The TF kernel design is, therefore, defined as an optimization problem. Based on the measurements in the time-lag domain, we search for the kernel in the ambiguity domain which yields sparse and cross-terms free TFSR. The Gini index [45]-[47] is used as the sparsity measure and controls the extent of the low pass kernel in the ambiguity domain. In the problem formulation, the kernel and ambiguity function appear as one term and, as such, the optimization is performed on the kerneled ambiguity function.

The paper is organized as follows. In Section II, we provide a brief review of traditional QTFDs. The effect of compressed data on the ambiguity function is analyzed in Section III. The proposed kernel design is introduced in Section IV, while Section V contains simulations results. Conclusion is given in Section VI.

II. Traditional TF kernels

The IAF can be used to define the class of RIDs. For a complex-valued signal sampled with period T, i.e., $x(n) = x(nT)$, the IAF is formulated over time n and lag m as,

$$R_{xx}(n,m) = x(n+m)x^*(n-m). \tag{2}$$

The discrete Fourier transform (DFT) of the IAF over time yields the ambiguity function (AF) defined over Doppler frequency p and lag m,

$$A(p,m) = \sum_{n=-N/2}^{N/2-1} R_{xx}(n,m)e^{-j2\pi np/N}, \tag{3}$$

where $\jmath = \sqrt{-1}$ and N is the signal length. In the ambiguity domain, the desirable signal auto-terms are located at and around the origin, whereas cross-terms reside at distant positions. This property motivated the introduction of various TF kernels which possess low-pass filtering characteristics [10]. After applying a kernel $C(p,m)$, the resulting RID is obtained through the 2D Fourier transform,

$$RID(n,k) = \sum_p \sum_m A(p,m)C(p,m)e^{-\jmath 2\pi np/N} e^{-\jmath 2\pi mk/N}. \tag{4}$$

QTFDs should assume certain properties in order to be valid distributions of signal energy in the TF plane. In the case of RIDs, these properties are defined by the employed kernel. Some of desirable kernel properties are:

- Realness,
$$C(p,m) = C^*(-p,-m). \tag{5}$$
- Signal instantaneous power,
$$C(p,0) = 1. \tag{6}$$
- Power density spectrum,
$$C(0,m) = 1. \tag{7}$$

The last two properties are also known as marginal properties. Existing time-frequency kernels can be divided into two groups: signal-independent and signal-dependent. Most of the signal-independent kernels satisfy aforementioned properties. Fig. 2 depicts two commonly used signal-independent kernels. The major drawback of the signal-independent kernels is that they have fixed shapes in the ambiguity domain. This property renders them inflexible in accommodating a large class of signals and prevent them from achieving optimum tradeoff between auto-term preservation and cross-term suppression, even under the inclusion of an adjustable parameter σ. Additionally, and more importantly, these kernels are not designed to handle data with missing samples. Signal-dependent kernel design, on the other hand, is formulated as an optimization problem under two separate constraints [48]. One constraint forces the kernel to have low-pass filtering characteristics, whereas the other constraint specifies the volume under the optimal kernel. Such optimization can be performed either over the entire data record or windowed data. With no missing samples, the adaptive kernel generally outperforms the signal-independent kernels. However, under random data observations, existing signal-dependent kernels may be misguided when finding the optimal solution. Missing samples cause artifacts in the ambiguity domain [28] which can be mistakenly interpreted as signal regions of interest. Capturing these regions results in highly cluttered time-frequency distributions.

III. The Effect of Missing Samples on the Ambiguity Domain

In this section, we analyze the effect of missing samples on the ambiguity function. We show that preserving the realness and marginal properties does not lend itself to desirable TFSR when dealing with incomplete data.

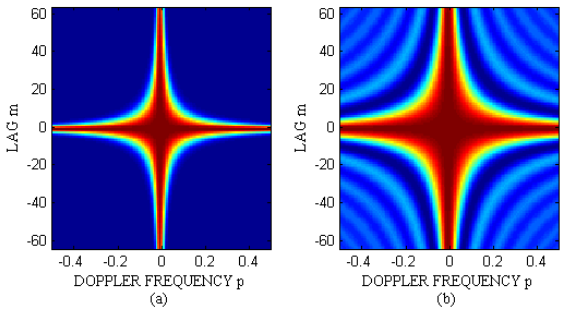

Fig. 2. Some of commonly used signal-independent time-frequency kernels: (a) Choi-Williams $C(p,m) = e^{\frac{-p^2m^2}{\sigma^2}}$; (b) Born-Jordan $C(p,m) = \frac{\sin(pm/\sigma)}{pm/\sigma}$.

Signal with missing samples can be represented as a modulated version of the original signal $s(n)$,

$$x(n) = s(n)\varphi_p(n), \tag{8}$$

where $\varphi_p(n)$ is the sampling pattern in time and can be represented as the sum of impulses at random positions n_p, i.e., $\varphi_p(n) = \Sigma_{n_p}\delta(n-n_p)$. These impulses determine the positions of available samples. The corresponding pattern of missing samples at positions n_m is given by,

$$\sum_{n_m} \delta(n-n_m) = 1 - \varphi_p(n). \tag{9}$$

These formulations show that missing samples could be viewed as a special type of impulsive noise. The ambiguity function of $x(n)$ becomes,

$$A_x(p,m) = \sum_{n=-N/2}^{N/2-1} R_{ss}(n,m)[1 - \sum_{n_m}\delta(n-n_m+m)] \\ [1 - \sum_{n_m}\delta(n-n_m-m)]e^{-\jmath 2\pi np/N}, \tag{10}$$

where $R_{ss}(n,m)$ denotes the IAF of the original signal $s(n)$. The ambiguity function of the compressed observations can be represented as the sum of the ambiguity function of the full data $A_s(p,m)$ and the artifacts caused by the missing samples $\Delta(p,m)$,

$$A_x(p,m) = A_s(p,m) + \Delta(p,m). \tag{11}$$

Define the ideal kernel, $C_x(p,m)$, for the missing sample case which satisfies

$$A_x(p,m)C_x(p,m) = A_s(p,m)C_s(p,m), \tag{12}$$

where $C_s(p,m)$ is a desirable RID kernel which, when applied to complete data $A_s(p,m)$, properly attenuates the cross-terms and captures the auto-terms. Accordingly from (11) and (12), we obtain

$$C_x(p,m) = \frac{A_s(p,m)C_s(p,m)}{A_s(p,m)+\Delta(p,m)} = \frac{C_s(p,m)}{1+\frac{\Delta(p,m)}{A_s(p,m)}}. \tag{13}$$

The above equation implies that the ideal kernel is signal dependent and it has complex values which, therefore, violate the realness property of traditionally used kernels.

The artifacts in (11) resemble noise in the sense that they are spread over the ambiguity domain. However, they follow a certain pattern and can be represented in the form of three terms,

$$\Delta(p,m) = V_1(p,m) + V_2(p,m) + V_3(p,m), \qquad (14)$$

where

$$V_1(p,m) = \sum_{n=-N/2}^{N/2-1} R_{ss}(n,m) \sum_{n_m} \delta(n-n_m+m) e^{-j2\pi np/N}; \qquad (15)$$

$$V_2(p,m) = \sum_{n=-N/2}^{N/2-1} R_{ss}(n,m) \sum_{n_m} \delta(n-n_m-m) e^{-j2\pi np/N}; \qquad (16)$$

$$V_3(p,m) = \sum_{n=-N/2}^{N/2-1} R_{ss}(n,m) \sum_{n_m} \delta(n-n_m+m) \sum_{n_m} \delta(n-n_m-m) e^{-j2\pi np/N}. \qquad (17)$$

By further developing each of the three artifact terms, we obtain the following expressions,

$$V_1(p,m) = -\sum_{n_m} s(n_m) s^*(n_m-2m) e^{-j2\pi p(n_m-m)/N}; \qquad (18)$$

$$V_2(p,m) = -\sum_{n_m} s(n_m+2m) s^*(n_m) e^{-j2\pi p(n_m+m)/N}; \qquad (19)$$

$$V_3(p,m) = \sum_{n_m} \delta(-2m) s(n_m) s^*(n_m-2m) e^{-j2\pi p(n_m-m)/N} + \sum_{n=-N/2}^{N/2-1} \sum_{n_m} [\delta(n-n_m+m) s(n+m) \sum_{n_l \neq n_m} \delta(n-n_l-m) s^*(n-m)] e^{-j2\pi np/N}. \qquad (20)$$

Besides the first two terms, which reside at $\forall m$, a component of the third term is always located at $m = 0$, i.e., along the Doppler frequency axis. This discourages the use of traditional RID kernels which, due to marginal properties, capture all values along the $m = 0$. Another conclusion which can be drawn from (18)-(20) is that the noise pattern in the ambiguity domain indeed depends on the values of the missing samples and their positions. Accordingly, it is expected that some parts of the ambiguity function will be more affected than others.

In what follows, we analyze the artifacts in the ambiguity domain using an approach based on the spectral density. We observe the spectral density of the IAF for fixed τ, i.e., $\Phi_{xx}(\omega)$. The bandwidth of the full data unaltered IAF is W over which the spectral density is assumed constant. Using the result given in the Appendix, we obtain

$$\Phi_{xx}(\omega) = W\frac{1-\beta^2}{T}\beta^2 + \frac{2\pi\beta^4}{T}\sum rect(\frac{1}{W}(\omega - \frac{2\pi n}{T})). \qquad (21)$$

T is the sampling period, whereas β and β^2 are probabilities that a sample is present in the time and time-lag domain, respectively. The first term in the above equation corresponds to the noise introduced by the missing samples, whereas the second term corresponds to the original unaltered IAF spectrum. As expected, more missing samples (smaller β) introduce not only higher noise level, but also higher distortion of signal auto-terms. Another important observation is that the noise level also depends on the support of the original IAF. We can define the signal-to-noise ratio ρ_A as,

$$\rho_A = \frac{2\pi\beta^2}{W(1-\beta^2)}, \qquad (22)$$

and use it to compare the effect of missing samples on the time and time-lag domains. For the former, if we assume that the signal occupies a frequency band of width Z, then the signal-to-noise ratio becomes $\rho_F = \frac{2\pi\beta}{Z(1-\beta)}$. Comparing these two ratios,

$$\rho = \frac{\rho_A}{\rho_F} = \frac{Z}{W}\frac{\beta}{1+\beta}. \qquad (23)$$

This relationship highlights the advantages of performing reconstruction in the ambiguity domain rather than in the frequency domain. Namely, most of the non-stationary signals are wide-band and occupy the entire frequency domain, while in the ambiguity domain, desired auto-terms may represent a fraction of the available bandwidth for given τ. Thus, even though there are more missing samples in the time-lag domain, $(1-\beta^2)$ compared to $(1-\beta)$, the relative signal small support in the ambiguity domain contributes to small noise level, making it easier to detect signal components in the ambiguity domain compared to the frequency domain. Fig. 3 shows how the value of ρ changes over different β and Z/W. It can be noticed that larger ratios of supports Z/W leads to more suitable detection in the ambiguity domain.

Our analysis in this section revealed the behaviour of ambiguity function in the presence of missing samples and showed that traditional RID kernels are unsuitable for randomly undersampled signals. This motivates the need for a new approach for kernel design.

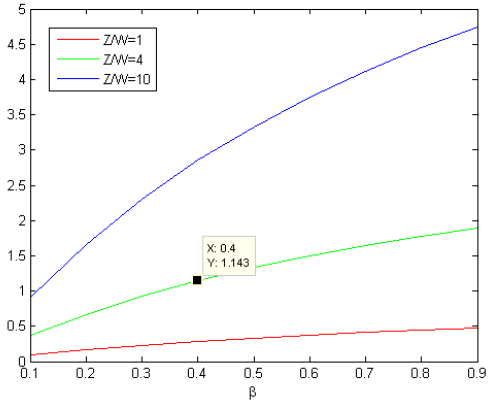

Fig. 3. Value of ρ over different percentage of present samples in time β. Point depicts the case when the support in the ambiguity domain is 4 times smaller than in the frequency domain, signal to noise ratio is higher in the ambiguity domain for $\beta \geq 0.4$.

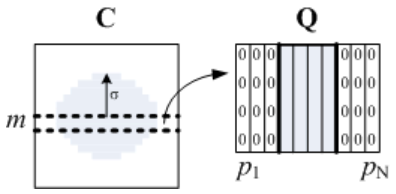

Fig. 4. Illustration of matrix $\mathbf{Q}$ for a simple circular shaped kernel $\mathbf{C}$ observed at a certain lag m.

IV. Time-frequency kernel design in the case of compressed data

In this section, we formulate the proposed RISD kernel design. We define the design problem as,

$$\begin{aligned} \underset{\mathbf{A_C},\mathbf{C}(\sigma)}{\text{maximize}} \quad & GI[((\mathbf{A_C})_{2DFT})_{vec}] \\ \text{subject to} \quad & \parallel \mathbf{r_{xx}}^m - \mathbf{Q}(\mathbf{C})\mathbf{a_C}^m \parallel_2 \leq \epsilon, m = 1, 2, ...N \\ & \mathbf{C} \text{ is a low-pass kernel} \end{aligned} \tag{24}$$

In (24), we seek the kernel $\mathbf{C}$ and the kerneled ambiguity function $\mathbf{A_C}$, dependent on parameter σ, which maximize the Gini index (GI) of TFSR. The latter is obtained as the 2D Fourier transform of $\mathbf{A_C}$. Matrix $\mathbf{Q}$, which is a partial discrete Fourier transform (DFT) matrix, relates the measurements in the time-lag domain to the kerneled ambiguity function.

In this respect, the proposed designed kernels identifies a new subclass of the general Cohen's class of TFDs. The Gini index is known as a sparsity measure [45], [46], which for sorted real-valued signal vector $\mathbf{x}, x(1) \leq ... \leq x(N)$ is defined as,

$$GI(\mathbf{x}) = 1 - 2\sum_{n=1}^{N} \frac{x(n)}{\|x\|_1} \frac{N - n + 0.5}{N}. \tag{25}$$

Signals with minimum number of non-zero values have maximum Gini index. It should be noted that in order to compute the Gini index, vectorized form of TFSR is used.

The first constraint in (24) is responsible for data fitting as it minimizes the ℓ_2 norm of residuals. Residuals in the constraint are expressed using the measurements in the time-lag domain for each lag m, $\mathbf{r_{xx}}^m$. The kernel and its extent determine the columns of matrix $\mathbf{Q}$. By solving this ℓ_2 problem, we obtain the kerneled ambiguity function for each m, $\mathbf{a_C}^m$. Setting the kernel to have low-pass filtering characteristics discards high frequency columns in the partial DFT matrix. This is illustrated in Fig. 4 which shows how the matrix $\mathbf{Q}$ is formed. The second constraint in (24) is necessary for ensuring that the auto-terms are captured without the inclusion of cross-terms. Without the general low pass filter characteristics of the kernel, cross-terms will be equally welcomed, as sparsity in the TF domain does not discriminate between auto and cross-terms. We employ a kernel function with circular shape of radius σ, even though any low-pass filter shape can be applied. For the sake of simplicity, the employed kernel has binary values, i.e., 1s and 0s. However, non-binary kernels can also be used such as the radial Gaussian kernel or the sinc kernel [10].

Since we do not know what the optimum $\mathbf{C}$ or $\mathbf{A_C}$ is, the optimization in (24) can be a computationally demanding process when compared to some of the well-known one-variable optimization processes, e.g., basis pursuit. However, by using knowledge about the sparsity and concentration behavior in the TF plane, we propose an efficient algorithm for solving (24). We first analyze the dependence of RID on parameter σ. We write the ambiguity function in terms of auto-terms and cross-terms. For a signal composed of L components, $s(n) = \sum_{l=1}^{L} s_l(n)$, the ambiguity function can be written as,

$$\begin{aligned} A(p,m) = & \sum_{l=1}^{L} \sum_{n=-N/2}^{N/2-1} s_l(n+m)s_l^*(n-m)e^{-\jmath 2\pi pn/N} \\ & + \sum_{l=1}^{L} \sum_{k=1,l\neq k}^{L} \sum_{n=-N/2}^{N/2-1} s_l(n+m)s_k^*(n-m)e^{-\jmath 2\pi pn/N}. \end{aligned} \tag{26}$$

The first term corresponds to auto-terms, whereas the second term corresponds to cross-terms. The 2D Fourier transform of the signal auto-terms provides the Wigner distribution without the cross-terms, i.e., $\mathbf{WD_{AT}}$, while the TF signature of cross-terms can be denoted as $\mathbf{WD_{CT}}$. The optimal σ_{opt} can be defined as the one which filters out the auto-terms while suppressing the cross-terms. We can observe three cases:

1) $\sigma \ll \sigma_{opt}$
   In this case, the kernel does not include cross-terms. However, part of the auto-terms may also be excluded. The resulting TF distribution is,
   $$RID(n,k) = \sum_u \sum_r WD_{AT}(u,r)\widetilde{C}(n-u,k-r), \tag{27}$$
   where $\widetilde{C}(n-u,k-r)$ represents the kernel in the TF domain. The narrowband property of $C(p,m)$ causes $\widetilde{C}(n,k)$ to be broad, leading to poor localization of the $RID(n,k)$. As such, the TFSR in (27) is neither concentrated nor sparse.
2) $\sigma_{opt} - \epsilon < \sigma < \sigma_{opt} + \epsilon$
   We consider this to be the optimal range of σ for capturing the auto-terms and suppressing the cross-terms. That is,
   $$RID(n,k) = WD_{AT}(n,k). \tag{28}$$
   The corresponding RID is well-concentrated and sparse.
3) $\sigma \gg \sigma_{opt}$
   Increasing the value of σ will improve the localization property of kernel in the TF domain, but will also include cross-terms. The higher value of σ implies more cross-terms in the TF plane. Thus, although RID is well-concentrated, it has more components and, therefore, is less sparse than in the previous case. In this case, the RID can be represented as,
   $$\begin{aligned} RID(n,k) = & WD_{AT}(n,k) + \\ & \sum_u \sum_r WD_{CT}(u,r)\widetilde{C}(n-u,k-r). \end{aligned} \tag{29}$$

A. Algorithm

Based on the above trade-off cases, we propose the following algorithm for solving (24).

procedure KERNEL DESIGN($\mathbf{A_C}, \mathbf{C}$)
  initialize $\sigma = \sigma_0$
  form a circular shaped kernel $\mathbf{C}$ of radius σ
  for each lag m **do**
    find the number K_m and the position set Ω_m of non-zero kernel values
    Based on the set Ω_m form matrix $\mathbf{Q}(\mathbf{C})$
    solve $\| \mathbf{r_{xx}}^m - \mathbf{Q}(\mathbf{C})\mathbf{a_C}^m \|_2$ using K_m
  end for
  if $GI[((\mathbf{A_C})_{2DFT})_{vec}]^{(i+1)} > GI[((\mathbf{A_C})_{2DFT})_{vec}]^{(i)}$ **then**
    increase σ and go to step 3
    otherwise stop and return 2DFT of $\mathbf{A_C}^i$
  end if
end procedure

In the initial step, we set a value for kernel parameter σ. We can set this value as small as possible. This parameter defines the frequency range over which the auto-terms will be reconstructed in the first iteration, $i = 1$. After forming the kernel, we solve for the ambiguity function for each lag in the time-lag domain. For a lag m, we determine the set Ω_m consisting of Doppler frequencies which are non-zero in the kernel $\mathbf{C}$. Based on this set, matrix $\mathbf{Q}(\mathbf{C})$ can be formulated and used to solve the under-determined problem which minimizes the residuals. It should be noted that $\mathbf{r_{xx}}^m$ contains the data measurements from the time-lag domain, including the cross-terms. Since the number of non-zero Doppler frequencies K_m is set by the kernel, and these components appear within a sub-band, greedy algorithms like Block Orthogonal matching pursuit (BOMP) [49] can be used, and ϵ in (24) does not have to be specified. After obtaining the low-pass filtered ambiguity function $\mathbf{A_C}$, the 2D Fourier transform is computed and TFSR is obtained. The Gini index of the result is measured which determines the stopping criterion of the algorithm. When sparsity of the TFSR in $i+1$ iteration is smaller than that in the ith iteration, procedure stops and the result from previous iteration is used as the final TFSR.

B. Choosing the sparsity measure

In what follows, we elaborate on the choice for the sparsity measure. Several measures of sparsity have appeared in the literature [47]. Among those measures, some deal with concentration property. Concentration measures have been successfully applied for determining the quality of TFSR [50]-[52]. Since concentration coincides with sparsity, multiple options are available for measuring the concentration/sparsity in the TF plane. However, each measure has different attributes. In [47], the authors proposed a set of intuitive attributes which each sparsity measure should have. They examined several sparsity measures, and proved that the Gini index and pq-mean satisfy all attributes, making them better choice compared to commonly used ℓ_1 norm. The pq-mean measure of real-valued vector $\mathbf{x}$ is defined as,

$$pq(\mathbf{x}) = -\frac{(\frac{1}{N}\sum_{k=1}^{N} x^p(k))^{1/p}}{(\frac{1}{N}\sum_{k=1}^{N} x^q(k))^{1/q}}. \quad (30)$$

When $p = 4$ and $q = 2$, and the forth power is computed, one of well-known TFSR concentration measures is obtained [51]. In statistics, this measure is better known as kurtosis. Even though this measure satisfies few desirable properties for measuring sparsity, it has one major shortcoming. For the case when there are multiple components, the above measure will not favor the distribution where each component is highly localized. Rather, it favors the case where one component is much more concentrated than others. In other words, TFSR with cross-terms is preferred over the TFSR which contains only auto-terms. Mixed norm represents an alternative measure, and for matrix $\mathbf{X}$ is defined as [53],

$$\ell_{21}(\mathbf{X}) = (\sum_k (\sum_j |X(k,j)|^2)^{1/2}). \quad (31)$$

This norm favors the group sparsity which is an intuitive manner for measuring the concentration of TFSR since TF signature for majority of signals is clustered around the instantaneous frequency. However, this measure is based on the use of ℓ_1 norm which is shown to be inferior when compared to the Gini index.

It is important to note that none of these sparsity measures are specifically designed to distinguish between the auto and the cross-terms. In our recent work [54], we compared several sparsity measures in order to determine which measure is the most efficient in the TF domain. Besides the previously described measures, the ℓ_0 norm and Renyi entropy were also included in the comparison. Results showed that, among the employed measures, the Gini index provided the TFSR with the most suppressed cross-terms for both noiseless and noisy data. Another significant result in [54] is that the most efficient solution for the noiseless full data was obtained with the ℓ_0 norm. This supports the argument that the sparsity, i.e., the number of occupied frequency bins, is significantly altered in the presence of cross-terms. One important property of the Gini index, which could explain its success in the underlying problem, is that the index is a normalized measure with values within the range [0,1] for any given signal vector. This property makes the Gini index a natural tool for measuring sparsity of different representations of the same signal. This conforms well with our case where changing the kernel parameters provides a different representation of the same signal.

Figs. 5 and 7 show TFSRs when three different values of σ are used. Figs. 5, 6 correspond to the signal composed of two 3-degree polynomial phase components, whereas Figs. 7, 8 show results for two crossing chirps. For each σ, the sparsity measure is computed. Based on the results in Figs. 6 and 8, we can notice that, in the two examples, all measures successfully cast first value of σ as the least sparse. The results of measures for the second and the third case are close in values, and they demonstrate the difficulty which arises when distinguishing

the auto and cross-terms. Among the employed measures, the Gini index successfully determined the case when auto-terms are captured and cross-terms are not included.

V. Simulations

In order to evaluate the performance of the proposed RISD kernel design, various signal types are observed. Namely, we consider randomly under-sampled polynomial phase signals. It is shown that the TFSR based on the multi-task kernel design provides improved TF signature representation when compared to the traditional QTFDs. Additionally, we include results of other sparse reconstruction approaches for comparison. In all plots, the frequency axis is normalized with respect to the sampling frequency.

Example 1: In this example, we consider a signal consisting of two crossing chirps:

$$x(n) = e^{j16\pi n^2} + e^{-j16\pi n^2}.$$

The above signal is sampled according to Nyquist theorem followed by random removal of 50% of the samples. Fig. 9 shows the Choi-Williams (CW) distribution [55] and Adaptive optimal kernel (AOK) distribution [48], both in the case of full and incomplete data. Compared with the distributions using full data (Fig. 9 (a,b)), missing samples introduce noise which clutters both time-frequency distributions, as evident in Fig. 9 (c,d). Also, we can notice vertical lines in the cluttered CW distribution. These lines are impulses which, due to the marginal properties of the CW kernel, are captured along the $m = 0$ axis in the ambiguity domain, as discussed in Section III. In Fig. 10(a), the TFSR based on the proposed kernel design is shown. Even though 50% of samples are missing, the TF signature clearly depicts the two components.

For comparison, the results based on other reconstruction techniques are also shown (Fig. 10 (b,c)). The TF plot in Fig. 10(b) is obtained using the approach in [24]. Namely, first we compute ambiguity function based on the IAF and then, using a rectangular mask, we obtain the compressed measurements which are used for the reconstruction of the TFSR. We choose an appropriate mask for this signal in order to remove cross-terms. The signal auto-terms are distorted due to the induced noise effect of the missing samples, rendering the reconstruction process unreliable. In Fig. 10(b), the TFSR depicts two distorted chirps. Another way of providing the TF signature is through local reconstruction in which we partition the data into overlapping segments and carry signal reconstruction over each segment separately [32]. The OMP algorithm is used for the reconstruction [56]. The result of this approach, shown in Fig. 10(c), depicts two clear components, but the resolution is poor.

Example 2: In the second example, we observe the multi-component signal,

$$x(n) = e^{j2.4\pi(n-1)^3+j16n} + e^{j2.4\pi(n-1)^3-j48n}.$$

The phase of both components is a third degree polynomial. As in the previous example, the signal is undersampled, 70% of samples are randomly missing. Fig. 11 shows the effect of missing samples on the TFSRs. In this example, the B-distribution kernel [57] is used as an representative of signal-independent kernels which do not satisfy marginal properties. Even though vertical lines are absent, some parts of the signal TF signature become missing, as evident in Fig. 11(d). Reconstruction results are depicted in Fig. 12. The reconstruction using measurements in the ambiguity domain (Fig. 12(b)) fails to capture the signal TF signature when a large portion of data is missing. Local reconstruction, shown in Fig. 12(c), contains gaps in the TF signature. These gaps correspond to time instants which have large number of missing samples within a window, as evident in Fig. 12(d). Additionally, even over the regions of TF plane where we can recognize two components, the TFSR, like spectrogram, suffers from poor resolution. The proposed kernel design provides the TF signature without gaps or vertical lines and with better resolution than local reconstruction (Fig. 12(a)).

Example 3: We observe a similar signal as in the previous example, but with three gaps, i.e., the missing samples occur in groups. The gaps are represented by red asterisk in Fig. 13 (d). The starting and ending points of the gaps are as follows: (36, 40), (70, 80) and (106, 110). Reconstruction based on the proposed approach is given in Fig. 13 (a) while that associated with the chirp dictionary [37], [58] is depicted in Fig. 13 (b). In Fig. 13 (c), we show the reconstruction based on the interpolation of the IMFs. First, the signal is subdivided into four contiguous portions. IMFs, which are computed for each of these four portions, are interpolated in order to mitigate the effect of the gaps. The EMD gap-filling concept was presented in [41] where a sophisticated technique to interpolate gaps in the IMFs is used. However, this approach requires manual setting of the parameters and it is based on the appearance of signal gaps. Here, we use simple interpolation based on the spline functions [59]. The signal is reconstructed as the sum of the interpolated IMFs and the spectrogram is computed (Fig. 13 (c)). We can notice that both the proposed approach and IMF based approach have significant clutter around the second gap, which is the largest among the three gaps. However, the clutter is more suppressed in the proposed approach. The method based on the chirp dictionary successfully fills in the gaps, but also shows clear deviations which occur when the employed dictionary does not appropriately represent the local signal behavior.

Example 4: In this example, computational EM modeling data corresponding to a human walking straight toward a pulse-Doppler radar (which means 0 azimuth) is observed. Vertical polarization is considered. The radar operates around 1 GHz, with 80 MHz bandwidth. It is assumed that a walking cycle takes 2 seconds. The EM solver, Finite Difference Time Domain (FDTD) is used. Analysis of human gait signals plays an important role in various military applications and TFSRs are commonly used to represent these signals [60]. Fig. 14 depicts AOK distribution computed using full and compressed data. Fig. 14(b) clearly shows the negative impact of missing samples on the TFSR. Reconstruction results are shown in Fig. 15. The proposed kernel design (Fig. 15(a)) successfully depicts TF signature which is not the case with local reconstruction that suffers from poor resolution (Fig.

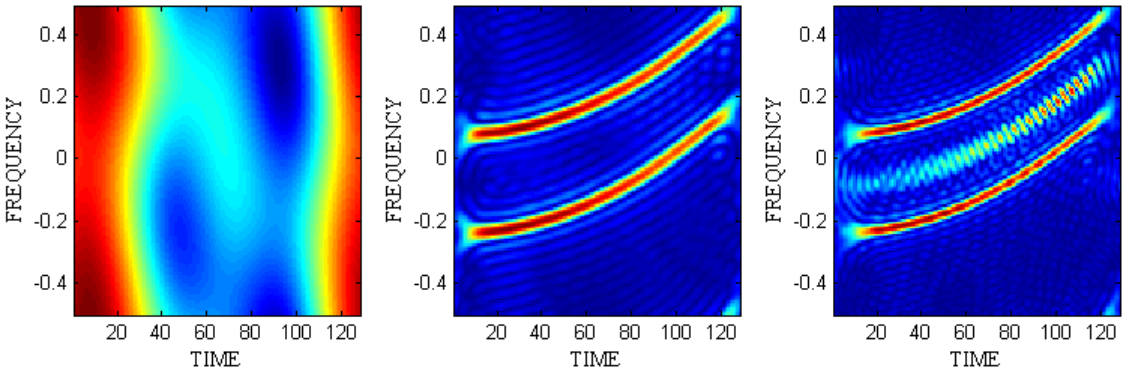

Fig. 5. TFSRs of a signal composed of two 3-degree polynomial phase. The corresponding kernel parameter is (a) $\sigma = 0.01N$, (b) $\sigma = 0.1N$, (c) $\sigma = 0.2N$. N is signal length, $N = 128$.

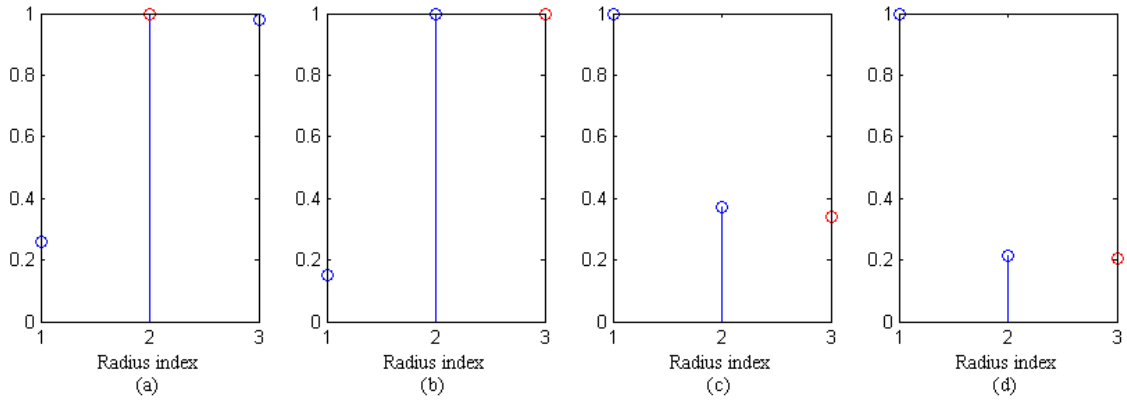

Fig. 6. Different sparsity measures applied to the TFSRs in Fig. 5. Radius indices correspond to the cases Fig. 5 (a), (b) and (c), respectively. (a) Gini index, (b) Kurtosis, (c) ℓ_{21} norm, (d) ℓ_1 norm. Red circle denotes the point where sparsest representation occurs according to each measure.

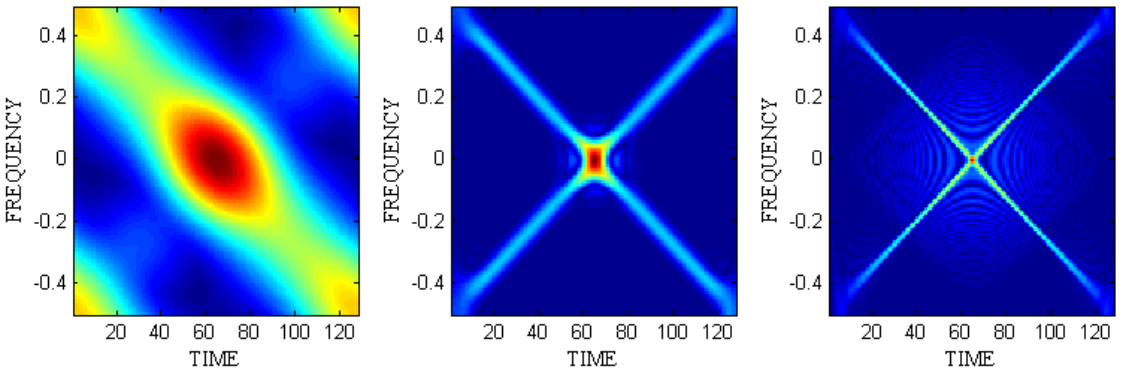

Fig. 7. TFSRs of two crossing chirps. The corresponding kernel parameter is (a) $\sigma = 0.01N$, (b) $\sigma = 0.1N$, (c) $\sigma = 0.2N$. N is signal length, $N = 128$.

15(b)).

VI. Conclusion

In this paper, we introduced reduced interference sparse TF distributions with the aim of providing highly localizable TF signature for incomplete or compressed data. The TF kernel design was defined as the optimization problem over two variables. In order to effectively solve this problem, we proposed an algorithm which takes into account the specific nature of the kernel spread and the auto- and cross-term distributions in the ambiguity domain. The algorithm iterates between the TF domain and ambiguity domain and measures sparsity using the Gini's index. We examined the properties of desired kernels under missing samples and demonstrated

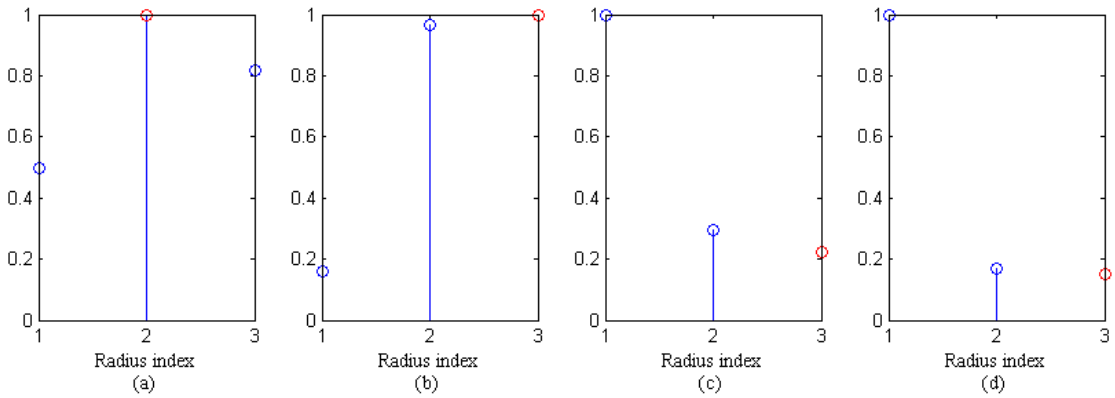

Fig. 8. Different sparsity measures applied to the TFSRs in Fig. 7. Radius indices correspond to the cases Fig. 7 (a), (b) and (c), respectively. (a) Gini index, (b) Kurtosis, (c) ℓ_{21} norm, (d) ℓ_1 norm. Red circle denotes the point where sparsest representation occurs according to each measure.

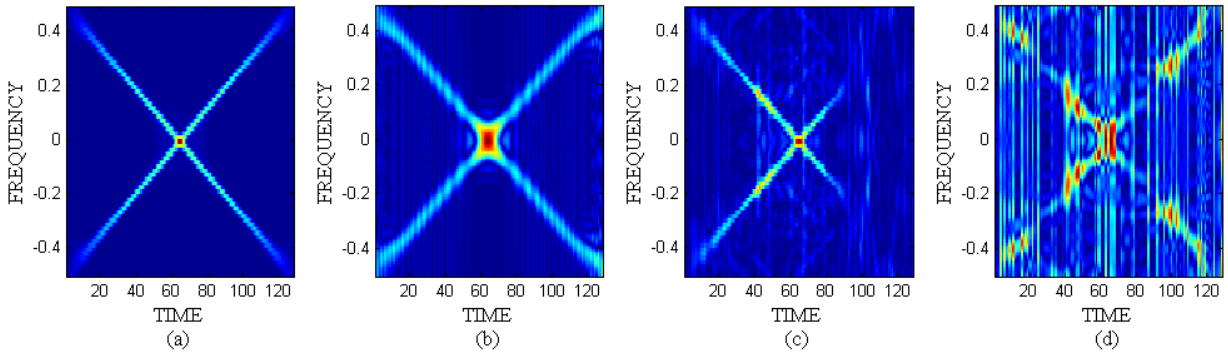

Fig. 9. TFSRs of signal consisting of two chirps: (a) AOK distribution of full data, (b) CW distribution of full data, (c) AOK distribution of compressed data, (d) CW distribution of compressed data.

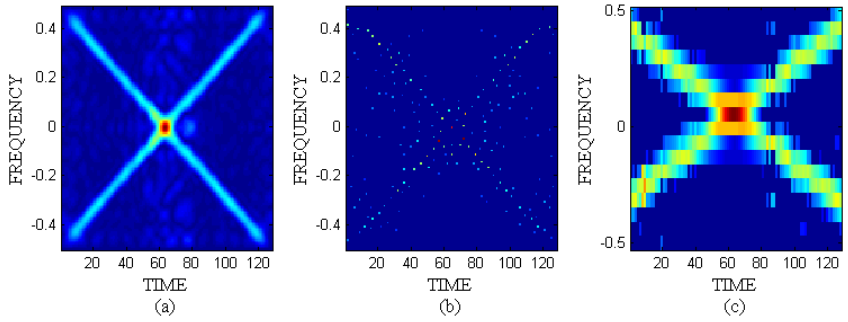

Fig. 10. TFSRs obtained using different reconstruction approaches: (a) Proposed kernel design, (b) Reconstruction of the TFSR using ambiguity domain, (c) Local reconstruction.

that they cannot be satisfied by the traditional TF kernels in Cohen's class. Simulation results illustrated the benefits of using the proposed approach, even in the case when a large part of data is missing.

APPENDIX

SPECTRAL DENSITY OF THE IAF

Define the signal with missing samples as $x(t) = s(t)\varphi_p(t)$, where both the complete signal $s(t)$ and the sampling pattern $\varphi_p(t)$ are continuous. Even though this assumption is not feasible in practice, it enables relatively simple derivations for impulse processes. We assume that the two processes $s(t)$ and $\varphi_p(t)$ are independent which, considering the randomness of missing samples, is a valid assumption. The IAF of randomly sampled signal is,

$$R_{xx}(t,\tau) = s(t+\tau)\varphi_p(t+\tau)s^*(t-\tau)\varphi_p(t-\tau). \quad (32)$$

If we observe IAF for fixed τ, then we can define spectral density of $R_{xx,\tau}(t)$ and examine the behaviour of IAF in the ambiguity domain,

$$\Phi_{xx}(\omega) = \Phi_{ss}(\omega) * \Phi_{\varphi}(\omega), \quad (33)$$

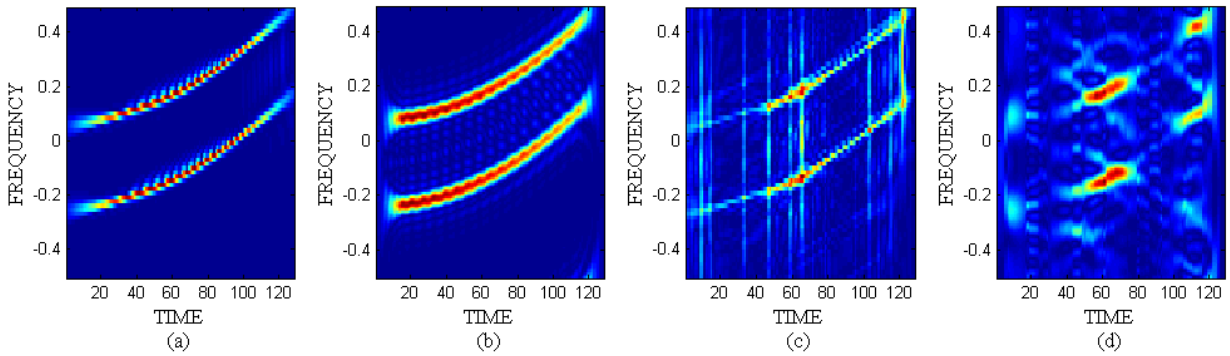

Fig. 11. TFSRs of signal consisting of two 3-degree polynomial components: (a) AOK distribution of full data, (b) B-distribution of full data, (c) AOK distribution of compressed data, (d) B-distribution of compressed data.

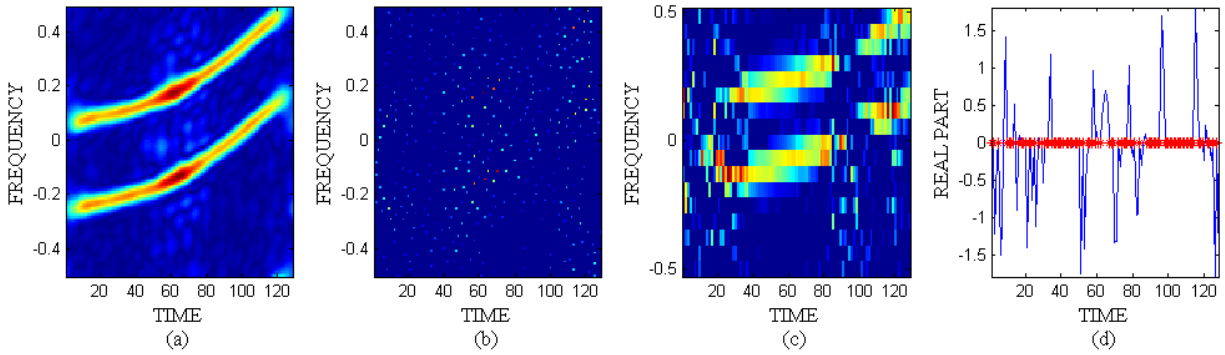

Fig. 12. TFSRs obtained using different reconstruction approaches: (a) Proposed kernel design, (b) Reconstruction of the TFSR using ambiguity domain, (c) Local reconstruction (d) Real part of signal, missing samples are denoted by red asterisk.

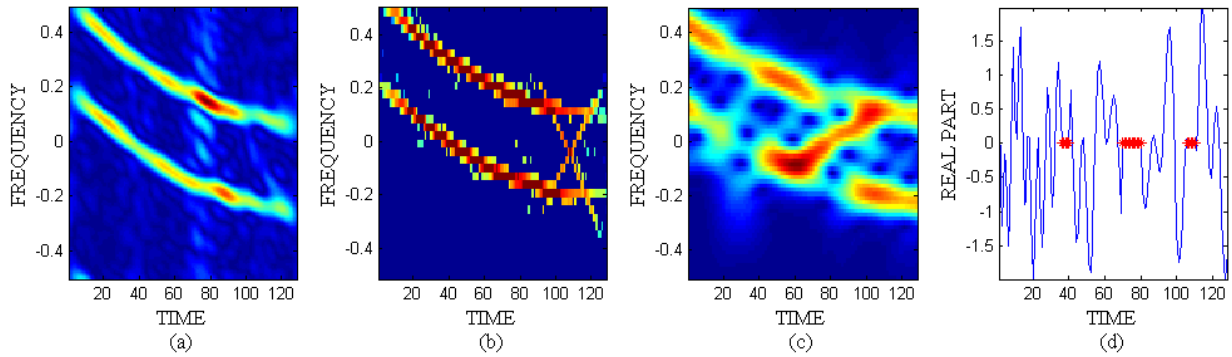

Fig. 13. TFSRs obtained using different reconstruction approaches: (a) Proposed kernel design, (b) Reconstruction of the TFSR using chirp dictionary, (c) Reconstruction based on interpolating the IMFs, (d) Real part of signal, missing samples are denoted by red asterisk.

where $*$ denotes convolution operation, $\Phi_{ss}(\omega)$ is the spectral density of the function $s(t+\tau)s^*(t-\tau)$, while $\Phi_{\varphi}(\omega)$ denotes spectral density of the corresponding sampling pattern function $\varphi_p(t+\tau)\varphi_p(t-\tau)$. Using a similar approach as in [61], [62], which operates on the compressed observations in the time domain and derives the spectral density in the frequency domain, the spectral density of $R_{xx,\tau}(t)$ becomes,

$$\Phi_{xx}(\omega) = \Phi_{ss}(\omega) * \frac{1-\beta^2}{T}\beta^2 + \Phi_{ss}(\omega) * \frac{2\pi\beta^4}{T}\sum \delta(\omega - \frac{2\pi n}{T}), \quad (34)$$

where T is the sampling period, β^2 is the probability that a sample is non-missing in the time-lag domain and β is the probability that a sample is present in the time domain.

REFERENCES

[1] 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.

[2] P. Setlur, M. Amin, and T. Thayaparan, "Micro-Doppler signal estimation for vibrating and rotating targets," in *Proc. 8th Int. Symp. Signal Process. and Appl.*, vol. 2, Sydney, Australia, Aug. 2005, pp. 639–642.

[3] M. Stridh, L. Sornmo, C. J. Meurling, and S. B. Olsson, "Sequential characterization of atrial tachyarrhythmias based on ECG time-frequency analysis," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 1, pp. 100–114, 2004.

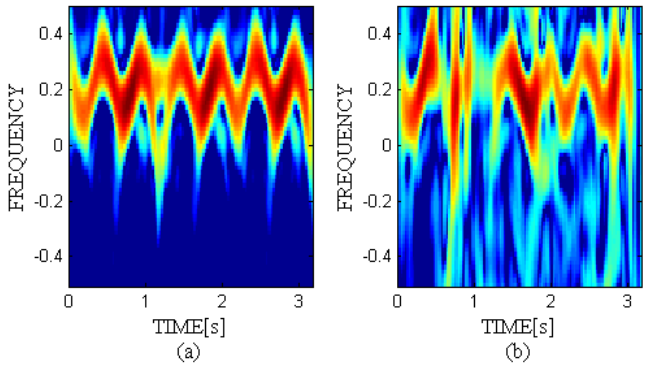

Fig. 14. TFSRs of human gait signal: (a) AOK distribution of full data, (b) AOK distribution of compressed data.

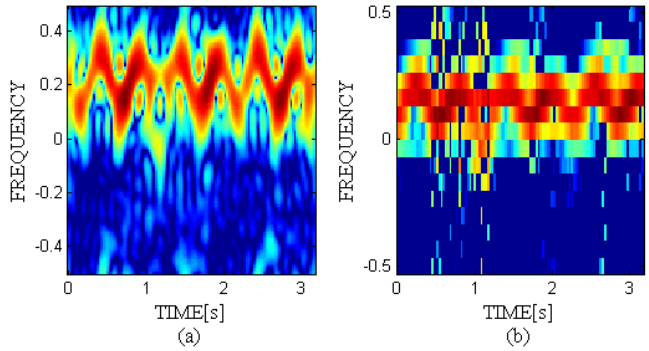

Fig. 15. TFSRs obtained using different reconstruction approaches: (a) Proposed kernel design, (b) Local reconstruction.

[4] I. Christov, G. Gómez-Herrero, V. Krasteva, I. Jekova, A. Gotchev, and K. Egiazarian, "Comparative study of morphological and time-frequency ECG descriptors for heartbeat classification," *Medical engineering & physics*, vol. 28, no. 9, pp. 876–887, 2006.

[5] F. Miwakeichi, E. Martınez-Montes, P. A. Valdés-Sosa, N. Nishiyama, H. Mizuhara, and Y. Yamaguchi, "Decomposing EEG data into space–time–frequency components using parallel factor analysis," *NeuroImage*, vol. 22, no. 3, pp. 1035–1045, 2004.

[6] O. Yilmaz and S. Rickard, "Blind separation of speech mixtures via time-frequency masking," *IEEE Trans. Signal Process.*, vol. 52, no. 7, pp. 1830–1847, 2004.

[7] L. Cohen, *Time-Frequency Analysis: Theory and Applications*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1995.

[8] B. Boashash, Ed., *Time-Frequency Signal Analysis and Processing. A Comprehensive Reference*, 1st ed. Oxford, U.K.: Elsevier, 2003.

[9] J. Jeong and W. J. Williams, "Kernel design for reduced interference distributions," *IEEE Trans. Signal Process.*, vol. 40, no. 2, pp. 402–412, 1992.

[10] L. Cohen, "Time-frequency distributions-a review," *Proc. IEEE*, vol. 77, no. 7, pp. 941–981, 1989.

[11] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[12] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Signal Process. Mag.*, vol. 52, no. 2, pp. 489–509, 2006.

[13] R. G. Baraniuk, "Compressive sensing," *IEEE Signal Processing Mag.*, vol. 24, no. 4, 2007.

[14] M. G. Amin, Ed., *Compressive Sensing for Urban Radar*. Boca Raton, FL: CRC Press, 2014.

[15] C. Law, G. Jones, D. Backer, W. Barott, G. Bower, C. Gutierrez-Kraybill, P. Williams, and D. Werthimer, "Millisecond imaging of radio transients with the pocket correlator," *The Astrophysical Journal*, vol. 742, no. 1, p. 12, 2011.

[16] J. M. Peha, "Wireless communications and coexistence for smart environments," *IEEE Pers. Commun.*, vol. 7, no. 5, pp. 66–68, 2000.

[17] J. D. Scargle, "Studies in astronomical time series analysis. II-Statistical aspects of spectral analysis of unevenly spaced data," *The Astrophysical Journal*, vol. 263, pp. 835–853, 1982.

[18] A. Duijndam, M. Schonewille, and C. Hindriks, "Irregular and sparse sampling in exploration seismology," in *Nonuniform Sampling*. Springer, 2001, pp. 479–518.

[19] F. J. Herrmann, M. P. Friedlander, and O. Yilmaz, "Fighting the curse of dimensionality: compressive sensing in exploration seismology," *IEEE Signal Processing Mag.*, vol. 29, no. 3, pp. 88–100, 2012.

[20] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic resonance in medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[21] M. F. Duarte and R. G. Baraniuk, "Spectral compressive sensing," *Applied and Computational Harmonic Analysis*, vol. 35, no. 1, pp. 111–129, 2013.

[22] H. Rauhut, "Stability results for random sampling of sparse trigonometric polynomials," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5661–5670, 2008.

[23] S. Bourguignon, H. Carfantan, and J. Idier, "A sparsity-based method for the estimation of spectral lines from irregularly sampled data," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 575–585, 2007.

[24] P. Flandrin and P. Borgnat, "Time-frequency energy distributions meet compressed sensing," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 2974–2982, 2010.

[25] J. F. Gemmeke, H. Van Hamme, B. Cranen, and L. Boves, "Compressive sensing for missing data imputation in noise robust speech recognition," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 272–287, 2010.

[26] L. Stankovic, I. Orovic, S. Stankovic, and M. Amin, "Compressive sensing based separation of nonstationary and stationary signals overlapping in time-frequency," *IEEE Trans. Signal Process.*, vol. 61, no. 18, pp. 4562–4572, Sep. 2013.

[27] L. Chaparro, E. Sejdic, A. Can, O. Alkishriwo, S. Senay, and A. Akan, "Asynchronous representation and processing of nonstationary signals: A time-frequency framework," *IEEE Signal Process. Mag.*, vol. 30, no. 6, pp. 42–52, 2013.

[28] Y. D. Zhang, M. G. Amin, and B. Himed, "Reduced interference time-frequency representations and sparse reconstruction of undersampled data," in *Proc. Signal Process. Conf. (EUSIPCO)*, Marrakech, Morocco, Sep. 2013, pp. 1–5.

[29] G. Li and P. Varshney, "Micro-Doppler parameter estimation via parametric sparse representation and pruned Orthogonal matching pursuit," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. PP, no. 99, pp. 1–12, 2014.

[30] B. Jokanovic, M. G. Amin, Y. D. Zhang, and F. Ahmad, "Time-frequency kernel design for sparse joint-variable signal representations," in *Proc. Signal Process. Conf. (EUSIPCO)*, Lisbon, Portugal, Sep. 2014, pp. 1–5.

[31] B. Jokanović, M. G. Amin, and Y. D. Zhang, "Reducing noise in the time-frequency representation using sparsity promoting kernel design," in *Proc. SPIE*, Baltimore, MD, May 2014.

[32] M. G. Amin, B. Jokanovic, and T. Dogaru, "Reconstruction of locally frequency sparse nonstationary signals from random samples," in *Proc. Signal Process. Conf. (EUSIPCO)*, Lisbon, Portugal, Sep. 2014, pp. 1–5.

[33] L. Stankovic, S. Stankovic, and M. Amin, "Missing samples analysis in signals for applications to L-estimation and compressive sensing," *Signal Process.*, vol. 94, pp. 401–408, Jan. 2014.

[34] 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.

[35] T. Y. Hou and Z. Shi, "Adaptive data analysis via sparse time-frequency representation," *Advances in Adaptive Data Analysis*, vol. 3, no. 1–2, pp. 1–28, 2011.

[36] D. Ba, B. Babadi, P. L. Purdon, and E. N. Brown, "Robust spectrotemporal decomposition by iteratively reweighted least squares," *Proc. of the National Academy of Sciences*, vol. 111, no. 50, pp. 5336–5345, 2014.

[37] Y. T. Nguyen, M. G. Amin, M. Ghogho, and D. McLernon, "Time-frequency signature sparse reconstruction using chirp dictionary," in *Proc. SPIE*, Baltimore, MD, April 2015.

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

[39] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the royal statistical society. Series B (methodological)*, pp. 1–38, 1977.

[40] I. Daubechies, J. Lu, and H.-T. Wu, "Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool," *Applied and computational harmonic analysis*, vol. 30, no. 2, pp. 243–261, 2011.
[41] A. Moghtaderi, P. Borgnat, and P. Flandrin, "Gap-filling by the empirical mode decomposition," in *Proc. ICASSP*. IEEE, 2012, pp. 3821–3824.
[42] H.-T. Wu, P. Flandrin, and I. Daubechies, "One or two frequencies? the synchrosqueezing answers," *Advances in Adaptive Data Analysis*, vol. 3, no. 01n02, pp. 29–39, 2011.
[43] G. Rilling and P. Flandrin, "One or two frequencies? *t*he empirical mode decomposition answers," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 85–95, 2008.
[44] T. Y. Hou and Z. Shi, "Data-driven time–frequency analysis," *Applied and Computational Harmonic Analysis*, vol. 35, no. 2, pp. 284–308, 2013.
[45] H. Dalton, "The measurement of the inequality of incomes," *The Economic Journal*, pp. 348–361, 1920.
[46] J. L. Gastwirth, "The estimation of the Lorenz curve and Gini index," *The Review of Economics and Statistics*, pp. 306–316, 1972.
[47] N. Hurley and S. Rickard, "Comparing measures of sparsity," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4723–4741, 2009.
[48] D. L. Jones and R. G. Baraniuk, "An adaptive optimal-kernel time-frequency representation," *IEEE Trans. Signal Process.*, vol. 43, no. 10, pp. 2361–2371, 1995.
[49] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3042–3054, 2010.
[50] 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, 2009.
[51] D. L. Jones and T. W. Parks, "A high resolution data-adaptive time-frequency representation," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 38, no. 12, pp. 2127–2135, 1990.
[52] L. Stanković, "A measure of some time-frequency distributions concentration," *Signal Processing*, vol. 81, no. 3, pp. 621–631, 2001.
[53] A. Gramfort, M. Kowalski, and M. Hämäläinen, "Mixed-norm estimates for the M/EEG inverse problem using accelerated gradient methods," *Physics in medicine and biology*, vol. 57, no. 7, p. 1937, 2012.
[54] B. Jokanovic and M. G. Amin, "Sparsity and concentration measures for optimum quadratic time-frequency distributions of doppler signals," in *IEEE Radar Conf.*, May 2015.
[55] H. I. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 37, no. 6, pp. 862–871, Jun. 1989.
[56] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
[57] B. Barkat and B. Boashash, "A high-resolution quadratic time-frequency distribution for multicomponent signals analysis," *IEEE Trans. Signal Process.*, vol. 49, no. 10, pp. 2232–2239, 2001.
[58] L. Applebaum, S. D. Howard, S. Searle, and R. Calderbank, "Chirp sensing codes: *d*eterministic compressed sensing measurements for fast recovery," *Applied and Computational Harmonic Analysis*, vol. 26, no. 2, pp. 283–290, 2009.
[59] J. A. Parker, R. V. Kenyon, and D. Troxel, "Comparison of interpolation methods for image resampling," *IEEE Trans. Med. Imag.*, vol. MI-2, no. 1, pp. 31–39, Mar. 1983.
[60] T. Dogaru, C. Le, and G. Kirose, "Time-frequency analysis of a moving human doppler signature," DTIC Document, Tech. Rep., 2009.
[61] F. J. Beutler and O. A. Leneman, "The spectral analysis of impulse processes," *Information and Control*, vol. 12, no. 3, pp. 236–258, 1968.
[62] C. Luo and J. H. McClellan, "Discrete random sampling theory," in *Proc. ICASSP*. IEEE, 2013, pp. 5430–5434.

Branka Jokanovic received the B.Sc. and M.Sc. degrees from the Faculty of Electrical Engineering, University of Montenegro, in 2010 and 2012, respectively. Currently, she is a Research Assistant at the Villanova University where she is pursuing the Ph.D. degree. Her research interests include time-frequency analysis, compressive sensing and their application to radars.

Moeness Amin is the Director of the Center for Advanced Communications, Villanova University, Pennsylvania, USA. He is a Fellow of the Institute of Electrical and Electronics Engineers; Fellow of the International Society of Optical Engineering; Fellow of the Institute of Engineering and Technology; and Fellow of the European Association for Signal Processing. Dr. Amin is a Recipient of the 2014 IEEE Signal Processing Society Technical Achievement Award; Recipient of the 2009 Individual Technical Achievement Award from the European Association for Signal Processing; Recipient of the IEEE Warren D White Award for Excellence in Radar Engineering; Recipient of the IEEE Third Millennium Medal; Recipient of the 2010 NATO Scientific Achievement Award; Recipient of the 2010 Chief of Naval Research Challenge Award; Recipient of Villanova University Outstanding Faculty Research Award, 1997; and the Recipient of the IEEE Philadelphia Section Award, 1997. He was a Distinguished Lecturer of the IEEE Signal Processing Society, 2003-2004, and is currently the Chair of the Electrical Cluster of the Franklin Institute Committee on Science and the Arts. Dr. Amin has over 700 journal and conference publications in signal processing theory and applications. He co-authored 18 book chapters and is the Editor of the two books "Through the Wall Radar Imaging" and "Compressive Sensing for Urban Radar", published by CRC Press in 2011 and 2014, respectively.