Robust Time-Frequency Reconstruction by Learning Structured Sparsity
Lei Jiang, Haijian Zhang, and Lei Yu
Signal Processing Lab., School of Electronic Information, Wuhan University, China

Abstract—Time-frequency distributions (TFDs) play a vital role in providing descriptive analysis of non-stationary signals involved in realistic scenarios. It is well known that low time-frequency (TF) resolution and the emergency of cross-terms (CTs) are two main issues, which make it difficult to analyze and interpret practical signals using TFDs. In order to address these issues, we propose the U-Net aided iterative shrinkage-thresholding algorithm (U-ISTA) for reconstructing a near-ideal TFD by exploiting structured sparsity in signal TF domain. Specifically, the signal ambiguity function is firstly compressed, followed by unfolding the ISTA as a recurrent neural network. To consider continuously distributed characteristics of signals, a structure-aware thresholds are learned from enormous neighboring TF coefficients. The proposed U-ISTA model is trained by both non-overlapped and overlapped synthetic signals involving closely and far located non-stationary components. Experimental results demonstrate that the robust U-ISTA achieves superior performance compared with state-of-the-art algorithms, and gains a high TF resolution with CTs greatly eliminated even in low signal-to-noise ratio (SNR) environments.

Index Terms—time-frequency distribution, compressive sensing, iterative shrinkage-thresholding, deep learning

I. INTRODUCTION

DURING the past decades, various time-frequency distributions (TFDs), arising in the field of time-frequency (TF) signal analysis and processing [1—4], have been reported and developed to deal with a wide range of non-stationary signals whose magnitudes and frequencies change over time, such as radar signals [5—7], communication signals [8], audio signals [11—13], and biomedical signals [14, 15]. The vast majority of practical signals are non-stationary, thus seeking a near-ideal TF representation of non-stationary signals is of great significance to signal processing applications including filtering [16, 17], localization [18, 20], separation [21—23], classification [24—26], etc. In this paper, we emphasize on the robust reconstruction of TF domain mixture signal under different signal-to-noise ratio (SNR) conditions.

All the while, the short-time Fourier transform (STFT) is the most commonly used linear TFD whose resolution is restricted to fixed sliding window [19], and the wavelet transform (WT) [27] with adaptive window size was proposed to attain changeable TF resolution. Though linear TFDs have desirable property, e.g., low implementation complexity, nonlinear TFDs can achieve more concentrated TF energy distributions [28]. Typical nonlinear TFDs include polynomial TFDs [29, 30], Wigner-Ville distribution (WVD) [31, 32] and ambiguity function (AF) [25], however, they suffer from inner and outer cross-terms (CTs) caused by the nonlinear structure of TFDs, which might plague signal TF analysis and thus need to be reduced. From a post-processing point of view, the reassignment method assigns the average energy in certain domain to the gravity center of energy distributions, e.g., reassigned spectrogram (RSP) algorithm [33]. Nevertheless, it is still difficult to avoid CTs and TFD distortion especially in low SNR environments. The above difficulty highlights the requirement for improving TF reconstruction performance by breaking the tradeoff between high TF resolution and significant CTs [34, 35].

An increasing number of kernel function based TF analysis methods emerge to seek a compromise between suppressing CTs and preserving high TF resolution [36—41]. Since auto-terms (ATs) are distributed around the AF domain origin while most of CTs locate far away from the origin, classical Cohen’s class distributions [42] focus on designing various kernel functions in AF domain to remove CTs. On the one hand, some kernel functions could be signal independent, e.g., B-distribution (BD) [36] and extended modified B-distribution (EMBD) [37], which utilizes separable kernels to control the smoothing requirements between time and frequency axes. On the other hand, some kernel functions for special signals may be signal dependent as well, e.g., radially Gaussian kernel (RGK) [38] and multi-directional distribution (MDD) [39], which is specifically designed for piece-wise linear frequency-modulated signals. Besides, extracting useful information directly from typical TFDs such as WVD would be an alternative strategy to remove CTs [40, 41].

Classical time-frequency dictionaries have been combined with compressive sensing (CS) algorithms to establish sparse representations of non-stationary signals [43], e.g., matching pursuit (MP) [44] and orthogonal matching pursuit (OMP) [45]. The continuous development of CS technology brings new solutions to signal analysis and reconstruction by exploiting the sparsity of signal in transform domains [46—58]. Flandrin et al. [51] proposed the $\ell_1$-app algorithm relying on convex optimization to recover sparse TFDs of signals. The sparse representation of $\ell_1$-app is constructed with the 2-dimensional (2D) Fourier transform pair of quadratic AF and WVD. A small number of samples around the AF domain origin are treated as observation vector since they contain most energy of auto-terms, while their corresponding Fourier coefficients are used to built the measurement matrix. With approximate $\ell_1$ sparsity constraint, sparse TFDs with negligible CTs have been reconstructed [51]. However, when the signal’s TFD involves complex structure, e.g., crossing and closely-located components, the performance of $\ell_1$-app
might degrade. Moreover, fixed sampling geometry neglects the TF correlation structure of the analyzed signal and inner CTs resulting from the nonlinearity. Therefore, adaptive sampling geometry methods have been proposed to improve TF reconstruction performance [52]–[54]. Inspired by MDD [39], Moghadian et al. [55] proposed a signal-dependent sampling geometry algorithm which gains high TF resolution for signals composed of closely-located components. In addition to convex optimization algorithms, iterative algorithms have also been considered due to their low computational complexity and fast convergent rates. Volaric et al. [56] combined the fast intersection of the confidence intervals (FICI) rule with a two step iterative shrinkage thresholding algorithm (TwIST) [57], and proposed the FICI-TwIST algorithm. Since missing data samples are commonly encountered resulting from sensor failure and jammed measurements, Amin et al. [58] considered a real-life scenario and proposed the missing data iterative sparse reconstruction (MI-SR) algorithm based on instantaneous auto-correlation function (IAF) by taking account of practical burst missing data.

Recently, data-driven learning algorithms have been booming with the development of deep learning technology [59]–[63]. Benefiting from a large number of data, most of learning based algorithms pay attention to learning self-dependency between thresholds and signal amplitudes so that a one-to-one correspondence between each threshold and signal amplitude is obtained. It is well known that the structured sparsity model going beyond simple sparsity is feasible in many practical scenarios by exploiting the inter-dependency structure among signal coefficients [64]–[66]. Wu et al. [67] proposed a continuous structure Bayesian method to recover signals with missing data. Intuitively, considering the priori knowledge of structured sparsity, e.g., the continuity of signal TF signatures, will give rise to more robust recovery performance. Consequently, sophisticated TF reconstruction algorithms by learning structured sparsity prior of signal’s TFDs are worthy for further investigation.

In this paper, we propose a new deep learning based structured sparse TF reconstruction algorithm, which is hereafter referred to as the U-Net aided iterative shrinkage-thresholding algorithm (U-ISTA). Unlike other iterative algorithms with manual or self-dependent thresholds [59]–[61], [68], [69], the structure-aware thresholds in U-ISTA are learned via the U-Net classifier [70] to reveal underlying relationship among the signal sparsity coefficients in TF domain. The U-Net is recognized as an adaptive block involved into the unfolded ISTA, i.e., a one-to-many correspondence between one threshold and adjacent TF coefficients is built up so that the thresholds depend on both amplitude and structure information of signals. Experimental results demonstrate that the proposed end-to-end U-ISTA network greatly eliminates the CTs while obtaining a near-ideal TF representation even for spectrally-overlapped and closely-located signal components in low SNR environments.

The rest of this paper is organized as follows. Section II introduces the proposed U-ISTA. Experiments on various kinds of synthetic signals are presented in Section III. Finally, Section IV concludes this paper.

II. THE PROPOSED U-ISTA

A. Problem Statement in TF Domain

The quadratic WVD of a non-stationary signal \( x(t) \) is defined as

\[
WVD_x(t, f) = \mathcal{F}_{\tau \to t} \left\{ z \left( t + \frac{\tau}{2} \right) z^* \left( t - \frac{\tau}{2} \right) \right\},
\]

where \( z(t) \) is the analytic associate of \( x(t) \). For a signal \( x(t) \) containing a single linear frequency-modulated component with infinite duration, its WVD will exhibit ideal signal TF signature and give an exact estimation of signal instantaneous frequency (IF). Nevertheless, practical signals have finite duration, which inevitably results in energy leakage in signal TF domain. In contrast, when the signal \( x(t) \) contains multiple components, the inner and outer CTs emerge especially when nonlinear components are involved.

The sparse representation in the proposed U-ISTA is built based on the 2D Fourier relationship between WVD and AF [51]. The signal’s AF is interpreted as a joint TF correlation function, and can be calculated by 2D Fourier transform of the signal’s WVD

\[
AF_z(\nu, \tau) = \mathcal{F}_{\nu \to \tau, f \to \nu, \tau} \{ WVD_z(t, f) \}.
\]

The discrete version of analytic signal \( z(t) \) is defined as \( z(n) \), \( n = 1, 2, ..., N \), where \( N \) denotes the number of time samples. For brevity, we assume the number of frequency samples is equivalent to that of time samples. Thus, the discrete matrix form in (2) can be written as

\[
A = DWD^T,
\]

where \( A \in \mathbb{R}^{N \times N} \), \( W \in \mathbb{R}^{N \times N} \), \( D \in \mathbb{R}^{N \times N} \) is the discrete Fourier transform matrix, and \( T \) denotes the transpose operator. Next, the Kronecker product is used to obtain a convenient representation for the matrix form in (3)

\[
a = \Psi \omega,
\]

where \( a \in \mathbb{R}^{N^2 \times 1} \) denotes the vectorization of the matrix \( A \) formed by stacking the columns of \( A \) into a single column vector, \( \Psi = D \otimes D \in \mathbb{R}^{N^2 \times N^2} \), where \( \otimes \) represents the Kronecker product. Similarly, \( \omega \in \mathbb{R}^{N \times 1} \) denotes the vectorization of the matrix \( W \) formed by stacking the columns of \( W \) into a single column vector.

Since the ATs corresponding to the clean TF signatures mainly distribute around the origin in AF domain, whereas the CTs locate far away from the AF domain origin. This offers an opportunity to remove CTs by appropriately choosing samples in AF domain, which is realized through multiplying a low pass filter \( \Phi(\nu, \tau) \) in AF domain

\[
AF’_z(\nu, \tau) = AF_z(\nu, \tau) \odot \Phi(\nu, \tau),
\]

where \( \odot \) denotes element-wise multiplication, thus a new observation vector \( a' \) can be obtained, and \( \Phi(\nu, \tau) \) is actually a mask function with the sample set \( \Omega \) corresponding to the sampling rectangular area centered around the AF domain origin

\[
\Phi(\nu, \tau) = \begin{cases} 
1, & (\nu, \tau) \in \Omega \\
0, & \text{otherwise}
\end{cases}
\]
which has totally $M = D_o \times D_r$ samples, where $D_o$ and $D_r$ denote the two dimensions of the sampling rectangle.

According to the above analysis and the sparse sampling in [5], the expression in [4] is reformulated as a sparse representation

$$a' = \Psi' \omega,$$

where $a' \in \mathbb{R}^{M \times 1}$ is the CS observation vector, and $\Psi' \in \mathbb{R}^{M \times N^2}$ is the CS measurement matrix consisting of certain Fourier transform vectors which are relevant to $\Phi(\nu, \tau)$. Note that when the size of $\Omega$ is much less than that of $\omega$, i.e., $M \ll N^2$, the expression in [7] can be viewed as a classical CS sparse reconstruction problem by solving the following LASSO problem [71]

$$\hat{\omega} = \arg \min_\omega \frac{1}{2}||a' - \Psi' \omega||^2_2 + \lambda||\omega||_1,$$

where $\lambda > 0$ is a regularization parameter. The objective of the proposed U-ISTA is to learn the CT-free $\omega$ in [8], which is expected to be close to ideal TF representation.

### B. U-ISTA Network Architecture

An efficient method for solving the problem in [8] is ISTA, of which each iteration includes matrix-vector multiplication involving $\Psi'$ and $\Psi'^T$, followed by a shrinkage-thresholding step [72]

$$\omega^{(k+1)} = h_\theta \left( \omega^{(k)} - \frac{1}{L} \Psi'^T \left( \Psi' \omega^{(k)} - a' \right) \right),$$

where the soft shrinkage function is denoted as $h_\theta(\omega) = \text{sign}(\omega)(|\omega| - \theta)$, where $\theta = \lambda/L$ is a fixed threshold, and $L$ is the largest eigenvalue of $\Psi'^T \Psi'$. The above shows that all the thresholds in standard ISTA merely depend on regularization parameter $\lambda$ and the largest eigenvalue $L$. Extended researches related to ISTA pay increasing attention to proper selection of thresholds [68], [69]. In addition, learning all parameters is considered in Learned ISTA (LISTA) [59], thus theoretical fast convergence rate as well as lower error rates are gained [60], [61].

However, most of existing supervised algorithms pay less attention to the inter-dependencies between signal coefficients, e.g., the TF continuity of analyzed signal. Inspired by the fact that exploring TF structured sparsity will lead to better signal reconstruction, we propose a new U-Net aided ISTA network architecture, as illustrated in Fig. 1. At first, we unfold the ISTA as a resemble recurrent neural network (RNN) like the LISTA [39], for which it is more convenient to introduce structured sparse prior. Specifically, we make a choice of fixing $\Psi'$ in every layer other than learning it through data since Fourier transform matrix has an explicit physical significance. Furthermore, with a proper selection of threshold, i.e., making small thresholding values in nonzero entries while setting large values in elsewhere, higher convergence rate and lower possible error will be achieved [60].

Resorting to the power of advanced deep learning technology, the structure-aware thresholds could be learned benefiting from sufficient training data. For the sake of exploiting the continuity structure behind sparsity, we need to expand effective receptive field so as to attain long-range spatial dependencies among coefficients of the signal’s TFD. To achieve this, the typical U-Net classifier, which combines shallow with semantic features to assign a label to each pixel, is employed to explore TF correlation structure of the signal, e.g., structure-aware thresholding values are prone to be larger in ATs while smaller values are learned in CTs. Additionally, we employ only three layers in encoder and decoder path to balance between complexity and high-resolution, in the meanwhile, channels are considerably reduced.

In the framework of U-ISTA, the solution to the problem in [8] is written as

$$\begin{align*}
\omega^{(k+1)} &= h_\theta(\omega^{(k)} + Q^{(k)} a') \\
P^{(k+1)} &= I - t^{(k+1)} \Psi'^T \Psi' \\
Q^{(k+1)} &= t^{(k+1)} \Psi'^T
\end{align*}$$

where $k = 0, 1, 2, \ldots, K - 1$, resulting in a $K$-layer network, $t^{(k)}$ is a scalar which is the key to the convergence rate of network [61], and $I$ denotes identity matrix. Hence, the structure-aware threshold is defined as

$$\theta^{(k)} = U \{ \omega^{(k)} \}, \quad k = 0, 1, 2, \ldots, K - 1$$

where $U \{ \cdot \}$ represents the mapping from the $k$th iteration signal TF reconstruction to the structure-aware thresholds at the $k$th U-Net block. Generally, in order to achieve a near-ideal TF representation, the proposed U-ISTA adaptively controls thresholds according to the analyzed signal structure by learning local dependencies among signal amplitudes, which further makes a compensation for simple sparsity.

### III. Numerical Experiments

We train the proposed U-ISTA network on synthetic non-stationary signal data for 2500 epochs using NVIDIA GeForce
GTX 1080 GPUs with 0.001 learning rate. To be specific, the number of iterations in \((11)\) is set to \(K = 5\), and the training data contain various signal mixtures consisting of randomly matched linear frequency-modulated (LFM) and nonlinear sinusoidal frequency-modulated (SFN) components, with SNR levels ranging from 5 dB to 25 dB. Furthermore, the smooth \(\ell_1\) loss function is employed to obtain a robust network. For all reconstruction results, amplitudes are coded logarithmically with a dynamic range of 20 dB. The U-ISTA takes \(29 \times 29\) sampling rectangular area while the \(\ell_1\)-app algorithm merely samples \(13 \times 13\) points.\(^1\) The proposed U-ISTA is compared against some commonly used state-of-the-art algorithms, including BD \([36]\), WVD \([2]\), \(\ell_1\)-app \([51]\), and RSP \([33]\) \(^2\). In our simulation, the averaged normalized mean square error (NMSE) is used as a performance assessment criterion

\[
\text{NMSE} = 10 \log_{10} \left\{ \frac{\| \omega - \hat{\omega} \|_2^2}{\| \omega \|_2^2} \right\}. \tag{12}
\]

In the following, we consider five different cases with both linear LFM component as well as nonlinear SFM component, and each component has \(N = 128\) time samples. The visualization of the five non-stationary mixtures in 2-D TF domain is provided in the first columns of Fig. 2 and Fig. 3, which examine various situations including complex conditions, e.g., closely-located and spectrally-overlapped components. To verify the effectiveness of different algorithms on varying signal amplitudes, an amplitude modulated (AM) function is imposed on each component. As shown in Fig. 2 and Fig. 3, the reconstructed TFDs by different recovery algorithms are presented in favorable noise environment (SNR = 45 dB) and adverse noise environment (SNR = 5 dB), respectively.

**Case 1: signal composed of two far-located LFs**

Considering a mixture signal which contains two far-located LFM components:

\[
x_1(t) = a(t) \cos \left\{ 2\pi \left( 0.0002(t^2 - t_0^2) + 0.441(t - t_0) \right) \right\}
\]

\[
x_2(t) = a(t) \cos \left\{ 2\pi \left( 0.0004(t^2 - t_0^2) + 133(t - t_0) \right) \right\}
\]

where \(a(t) = \exp \left\{ -(0.0016t - 1)^2 \pi \right\} \) is the AM function, and \(t_0\) equals to 64 in all five cases.

When SNR = 45 dB, it is seen from the first row of Fig. 2 that the CTs are significantly suppressed by \(\ell_1\)-app, RSP and U-ISTA, while BD and WVD suffer from varying degrees of CT interference. It is worth mentioning that the proposed U-ISTA achieves near-ideal TF concentration which is close to ideal TFD. The reconstruction results when SNR = 5 dB are given in the first row of Fig. 3.\(^3\) note that although the performance of the U-ISTA degrades in contrast to the high SNR case, the proposed U-ISTA still obtains the highest TF resolution compared with other recovery algorithms, which are prone to be seriously disturbed by strong noise.

\(^1\)The performance of \(\ell_1\)-app degrades as the size of sampling area increases, thus the number of samples in \(\ell_1\)-app is smaller than that of U-ISTA to reduce the probability of sampling CTs in AF domain.

\(^2\)We would like to thank the authors in these works for sharing source codes, which are available online.

**Case 2: signal composed of two closely-located LFs**

Considering a mixture signal which contains two closely-located LFM components:

\[
x_1(t) = a(t) \cos \left\{ 2\pi \left( -0.0003(t^2 - t_0^2) + 0.3164(t - t_0) \right) \right\}
\]

\[
x_2(t) = a(t) \cos \left\{ 2\pi \left( 0.00001(t^2 - t_0^2) + 211(t - t_0) \right) \right\}
\]

When signal’s components are closely located to each other, the CTs tend to become serious, thus posing a challenge to the existing literature. The reconstruction results in high and low SNR levels are shown in the second rows of Fig. 2 and Fig. 3 respectively. It is seen that the results of U-ISTA give high TF resolution without CTs even though partly deformed when SNR = 5 dB. On the contrary, severe distortion appears in \(\ell_1\)-app while other algorithms suffer from considerable CTs and noise influence, which make them hard to analyze and interpret the mixture signal.

**Case 3: signal composed of two overlapped LFs**

Considering a mixture signal which contains two overlapped LFM components:

\[
x_1(t) = a(t) \cos \left\{ 2\pi \left( -0.0006(t^2 - t_0^2) + 0.29(t - t_0) \right) \right\}
\]

\[
x_2(t) = a(t) \cos \left\{ 2\pi \left( 0.0011(t^2 - t_0^2) + 0.109(t - t_0) \right) \right\}
\]

Recovery results are shown in the third rows of Fig. 2 and Fig. 3. It is observed that the \(\ell_1\)-app fails to represent the signal since ATs and CTs are mixed up around the origin of AF domain. When SNR = 5 dB, we note that the CTs are greatly reduced for the RSP algorithm, however, the deformation of its signal TFD occurs at the intersection. Benefitting from a large number of training data, the U-ISTA has more stable performance and succeeds to recover the signal TFD with high TF resolution. It should be emphasized that the proposed U-ISTA can accurately restore the information at the intersection, which is a difficult problem to be solved.

**Case 4: signal composed of an LFM and an SFM**

Considering a mixture signal which contains an LFM component and an SFM component:

\[
x_1(t) = a(t) \cos \left\{ 2\pi \left( 0.0002(t^2 - t_0^2) + 0.297(t - t_0) \right) \right\}
\]

\[
x_2(t) = a(t) \cos \left\{ 0.27\pi(t - t_0) \right. \}
\]

\[
+ 9.98 \sin \left( 0.0156\pi(t - t_0) - \phi_0 \right) - 9.98 \sin \phi_0 \}
\]

where \(\phi_0 = -1.403\).

In this case, we can obtain the similar observation and conclusion to those in **Case 1**. The only difference is that the nonlinearity of the SFM component will cause inner CTs, which can be perceived in Fig. 2. As seen from the fourth rows of Fig. 2 and Fig. 3, the \(\ell_1\)-app, RSP and U-ISTA can still gain free CTs, while BD and WVD undergo both inner and outer CTs. Furthermore, the performance of \(\ell_1\)-app and RSP algorithms deteriorates as the SNR level decreases. In contrast, the U-ISTA preserves a superb TF concentration irrespective of signal types, which verifies the effectiveness and robustness of the proposed U-ISTA.
Fig. 2: TF reconstructions of the five cases at SNR = 45 dB and NMSE is annotated above each image (From top to bottom: Case 1 to Case 5. From left to right: Ideal TFD, BD [36], WVD [2], \( \ell_1 \)-app [51], RSP [33], and the proposed U-ISTA).

**Case 5: signal composed of overlapped LFM and SFM**

Considering a mixture signal which contains overlapped LFM component and SFM component:

\[
\begin{align*}
    x_1(t) &= a(t) \cos \left\{ 2\pi \left( -0.0004(t^2 - t_0^2) + 0.117(t - t_0) \right) \right\} \\
    x_2(t) &= a(t) \cos \left\{ 0.42\pi (t - t_0) \\
    & \quad +15.9 \sin \left( 0.0156\pi (t - t_0) - \phi_0 \right) - 0.0156 \sin \phi_0 \right\}
\end{align*}
\]

where \( \phi_0 = -1.832 \).

This case is similar to Case 3 except that the performance of \( \ell_1 \)-app and RSP at the intersection is slightly improved, as shown in the last rows of Fig. 2 and Fig. 3. Note that the \( \ell_1 \)-app recovers the signal TFD with some remaining CTs and artifacts, which may be derived from the unapt selection of sampling area. Besides, the proposed U-ISTA cannot completely remove the CTs, and the reason behind this may be the oversized sampling area, i.e., \( D_\nu \times D_\tau \) is too large so that excessive CTs are contained. It implies that selecting a more proper sampling area is worth investigating in future. When SNR = 5 dB, the last row of Fig. 3 illustrates that only U-ISTA preserves a high resolution TFD with CTs greatly eliminated.

The quantitative evaluation results on different recovery algorithms for the above five cases in various SNR levels are presented in Fig. 4 (a) to (e), where reconstruction accuracy at each SNR level is assessed via NMSE over 100 Monte Carlo runs. In addition, the averaged NMSEs over the five cases are shown in Fig. 4 (f). We observe that the NMSE
results are consistent with those in Fig. 2 and Fig. 3. By learning structured sparsity of training data with varying SNR levels, the U-ISTA has achieved a large performance gain compared to other algorithms in the entire SNR range from 5 dB to 45 dB. When the SNR drops below 5 dB, the strong noise will burden the learning blocks, thus resulting in performance deterioration. It is worth stressing that although our training data merely include synthetic signals with SNR levels ranging from 5 dB to 25 dB, the proposed U-ISTA still gains distinguished results outside this range, e.g., SNRs varying from 30 dB to 45 dB in Fig. 4 which verifies its good generalization ability. In addition to NMSE, quantitative evaluation in terms of $\ell_1$-distance to model and Rényi entropy can also be considered to compare different algorithms [51].

**IV. CONCLUSION**

In this paper, we attempt to reconstruct near-ideal TFD by proposing the data-driven U-ISTA model, which combines the unfolded ISTA with the U-Net block to learn structure-aware thresholds so that the TF correlation structure of signals is explored. The effectiveness and robustness of the proposed U-ISTA are examined via numerical experiments in terms of NMSE, demonstrating much more significant performance improvement on synthetic non-stationary data compared to state-of-the-art algorithms, including classical TFDs and CS based ones. Particularly, the proposed U-ISTA successfully
(a) **Case 1**: two far-located LFM components.

(b) **Case 2**: two closely-located LFM components.

(c) **Case 3**: two overlapped LFM components.

(d) **Case 4**: non-overlapped LFM and SFM components.

(e) **Case 5**: overlapped LFM and SFM components.

(f) Averaged NMSEs from **Case 1 to Case 5**.

Fig. 4: NMSE performance versus various SNR levels using different TF reconstruction algorithms: BD [36], WVD [2], $\ell_1$-app [51], RSP [33], and the proposed U-ISTA.
reconstructs more signal details, e.g., the intersection region of spectrally-overlapped components, while maintaining superior TF concentration and continuous characteristics of signals even for closely-located signal components. Although exposing the potential benefit by sampling in AF domain, it is believed that proper selection of AF samples would result in better reconstruction results, thus a learning strategy for optimal sampling is worth investigating for future research. In addition, we also believe that our U-ISTA model can be boosted by feeding a wide variety of synthetic data as well as real-life measured data.

REFERENCES

[1] L. Stanković, M. Daković, and T. Thayaparan, “Time-Frequency Signal Analysis With Applications,” Artech House, 2013.

[2] B. Boashash, “Time-Frequency Signal Analysis and Processing (Second Edition),” Academic Press, 2016.

[3] P. Flandrin, “Explorations in Time-Frequency Analysis,” Cambridge University Press, 2018.

[4] E. Sejdić, I. Djurić, and J. Jiang, “Timefrequency feature representation using energy concentration: An overview of recent advances,” Digital Signal Processing, vol. 19, no. 1, pp. 153 – 183, 2009.

[5] H. Zhang, D. Li, Y. Zhang, and T. Truong, “Simultaneous Radio Frequency and Wideband Interference Suppression in SAR Signals via Sparsity Exploitation in TimeFrequency Domain,” IEEE Trans. Aero. Electron. Syst., vol. 56, no. 10, pp. 5780–5793, Oct. 2018.

[6] V. C. Chen and Shei Qian, “Joint Time-Frequency Transform for Radar Range-Doppler Imaging,” IEEE Trans. Aero. Electron. Syst., vol. 34, no. 2, pp. 486–499, Apr. 1998.

[7] H. Zhang, G. Bi, W. Yang, S. G. Razul, and C. M. S. See, “IF estimation of FM signals based on time-frequency image,” IEEE Transactions on Aerospace and Electronic Systems, vol. 51, no. 1, pp. 326–343, 2015.

[8] T. Keller and L. Hanzo, “Adaptive Multicarrier Modulation: a Convenient Framework for Time-Frequency Processing in Wireless Communications,” Proceedings of the IEEE, vol. 88, no. 5, pp. 611–640, 2000.

[9] C. Wang, J. Ding, C. Chang, and L. Ouyang, “Learning Based Noise Identification Techniques Using Time-Frequency Analysis and the U-Net,” in 2019 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), China, Taipei, 2019, pp. 1–2.

[10] M. G. Amin, D. Borio, Y. D. Zhang, and L. Galleani, “Time-Frequency Analysis for GNSSs: From interference mitigation to system monitoring,” IEEE Signal Processing Magazine, vol. 34, no. 5, pp. 85–9, 2017.

[11] Y. Liu and N. Mesgarani, “Conv-TasNet: Surpassing Ideal Time-Scale Representations for Suppressing Cross Terms,” IEEE Transac-

[12] J. R. Fonollosa and C. T. Nikias, “Wigner Higher Order Moment Spectra: Definition, Properties, Computation and Application to Transient Signal Analysis,” IEEE Trans. Signal Process., vol. 41, no. 1, p. 245, Jan. 1993.

[13] F. Auger and P. Flandrin, “Improving the Readability of Time-Frequency and Time-Scale Representations by the Reassignment Method,” IEEE Trans. Signal Process., vol. 33, no. 5, pp. 1086–1089, May 1995.

[14] L. Zuo, M. Li, Z. Liu, and L. Ma, “A High-Resolution Time-Frequency Representation and the Cross-Term Suppression,” IEEE Transac-

[15] L. Zuo, M. Li, and X. Xia, “New Smoothened Time-Frequency Rate Representations for Suppressing Cross Terms,” IEEE Transactions on Signal Processing, vol. 65, no. 3, pp. 733–747, 2017.

[16] 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, Oct. 2001.

[17] B. Boashash, G. Azemi, and J. M. O’Toole, “Time-Frequency Processing of Nonstationary Signals: Advanced TFD Design to Aid Diagnosis with Highlights From Medical Applications,” IEEE Signal Process. Mag., vol. 30, no. 6, pp. 108–119, Nov. 2013.

[18] R. G. Baraniuk and D. L. Jones, “Signal-dependent Time-Frequency Analysis Using a Radially Gaussian Kernel,” Signal Process., vol. 32, no. 3, pp. 263–284, Jun. 1993.

[19] B. Boashash and S. Ouelha, “An Improved Design of High-Resolution Quadratic Time-Frequency Distributions for the Analysis of Nonstationary Using Component Spectral, Directional Compact Kernels,” IEEE Trans. Signal Process., vol. 65, no. 10, pp. 2701–2713, May 2017.

[20] N. Saulig, J. Lerga, Z. Barasckai, and M. Daković, “Adaptive Thresholding in Extracting Useful Information From Noisy Time-Frequency Distributions,” in International Symposium on Image and Signal Processing and Analysis (ISPA), Dubrovnik, Croatia, Sep. 2019, pp. 329–334.

[21] N. Saulig, J. Lerga, Z. Barasckai, and M. Daković, “Adaptive Thresholding in Extracting Useful Information Content From Noisy Signals Based on Structural Affinity of Clustered TFDs Coefficients,” IEEE Trans. Signal Process., vol. 67, no. 12, pp. 3154–3167, Jun. 2019.
[42] L. Cohen, “Time-frequency distributions—a review,” Proceedings of the IEEE, vol. 77, no. 7, pp. 941–981, 1989.

[43] A. Gholami, “Sparse TimeFrequency Decomposition and Some Applications,” IEEE Transactions on Geoscience and Remote Sensing, vol. 51, no. 6, pp. 3598–3604, 2013.

[44] S. G. Mallat and Z. Zhang, “Matching Pursuits With Time-Frequency Dictionaries,” IEEE Trans. Signal Process., vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[45] Y. C. Pati, R. Rezaifar, and P. S. Krishnaprasad, “Orthogonal Matching Pursuit: Recursive Function Approximation With Applications to Wavelet Decomposition,” in Proceedings of 27th Asilomar Conference on Signals, Systems and Computers, vol. 1, Nov 1993, pp. 40–44.

[46] L. Stanković, E. Sejdić, S. Stanković, M. Daković, and I. Orovčić, “A Tutorial on Sparse Signal Reconstruction and Its Applications in Signal Processing,” Circuits Sys. Signal Process., vol. 38, no. 12601263, 2019.

[47] E. Sejdić, I. Orovčić, and S. Stanković, “Compressive sensing meets time-frequency: An overview of recent advances in timefrequency processing of sparse signals,” Digital Signal Processing, vol. 77, pp. 22 – 35, 2018.

[48] L. Yu, H. Sun, J. Barbot, and G. Zheng, “Bayesian compressive sensing for cluster structured sparse signals,” Signal Processing, vol. 92, no. 1, pp. 259 – 269, 2012.

[49] Q. Wu, J. Liu, and M. Amin, “Group Adaptive Matching Pursuit with Intra-group Correlation Learning for Sparse Signal Recovery,” Signal Processing, p. 107560, 2020.

[50] L. Stanković, M. Bajović, I. Stanković, J. Lerga, and M. Daković, “RANSAC-Based Signal Denoising Using Compressive Sensing,” arXiv:2003.12289, 2020.

[51] P. Flandrin and P. Borgnat, “Time-Frequency Energy Distributions Meet Compressed Sensing,” IEEE Trans. Signal Process., vol. 58, no. 6, pp. 2974–2982, Jun. 2010.

[52] I. Volaric, V. Sucic, and S. Stankovic, “A Data Driven Compressive Sensing Approach for Time-Frequency Signal Enhancement,” Signal Process., vol. 141, pp. 229–239, Dec. 2017.

[53] P. Addesso, M. Longo, S. Marano, V. Matta, I. M. Pinto, and M. Principe, “Sparsifying Time-Frequency Distributions for Gravitational Wave Data Analysis,” in International Workshop on Compressed Sensing and its Applications to Radar, Sonar and Remote Sensing (CoSeRa), Pisa, Jun. 2015, pp. 154–158.

[54] N. Whitelionis and H. Ling, “Radar Signature Analysis Using a Joint Time-Frequency Distribution Based on Compressed Sensing,” IEEE Transactions on Antennas and Propagation, vol. 62, no. 2, pp. 755–763, 2014.

[55] S. S. Moghadian and S. Gazor, “Sparsely Localized Time-Frequency Energy Distributions for Multi-Component LFM Signals,” IEEE Signal Process. Lett., vol. 27, pp. 6–10, 2020.

[56] I. Volaric, V. Sucic, and G. Bokelmann, “Sparse Time-Frequency Distribution Calculation with an Adaptive Thresholding Algorithm,” in International Symposium on Image and Signal Processing and Analysis (ISIPA), Dubrovnik, Croatia, Sep. 2019, pp. 341–346.

[57] J. M. Bioucas-Dias and M. A. T. Figueiredo, “A New TwIST: Two-Step Iterative Shrinkage/Thresholding Algorithms for Image Restoration,” IEEE Trans. Image Process., vol. 16, no. 12, pp. 3092–3004, Dec. 2007.

[58] V. S. Amin, Y. D. Zhang, and B. Himed, “Sparsity-based Time-Frequency Representation of FM Signals With Burst Missing Samples,” Signal Process., vol. 135, pp. 25–43, Feb. 2019.

[59] K. Gregor and Y. LeCun, “Learning Fast Approximations of Sparse Coding,” in International Conference on Machine Learning (ICML), 2010, pp. 399–406.

[60] X. Chen, J. Liu, Z. Wang, and W. Yin, “Theoretical Linear Convergence of Unfolded ISTA and Its Practical Weights and Thresholds,” in Advances in Neural Information Processing Systems, 2018, pp. 9061–9071.

[61] J. Liu, X. Chen, Z. Wang, and W. Yin, “ALISTA: Analytic Weights Are as Good as Learned Weights in LISTA,” in International Conference on Learning Representations (ICLR), 2019.

[62] H. Alaskar, “Deep Learning-Based Model Architecture for Time-Frequency Images Analysis,” International Journal of Advanced Computer Science and Applications, vol. 9, no. 12, 2018.

[63] Z. Sun, L. Zhou, and W. Wang, “Learning Time-Frequency Analysis in Wireless Sensor Networks,” IEEE Internet of Things Journal, vol. 5, no. 5, pp. 3388–3396, 2018.

[64] Y. C. Eldar and M. Mishali, “Robust Recovery of Signals From a Structured Union of Subspaces,” IEEE Trans. Inf. Theory, vol. 55, no. 11, pp. 5302–5316, Nov. 2009.

[65] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, “Model-based Compressive Sensing,” IEEE Trans. Inf. Theory, vol. 56, no. 4, pp. 1982–2001, Apr. 2010.