A NOVEL THRESHOLDING TECHNIQUE FOR THE DENOISING OF MULTICOMPONENT SIGNALS

Duong-Hung Pham and Sylvain Meignen

Laboratoire Jean Kuntzmann, Université Grenoble Alpes, 38401 Saint-Martin-d'Hères, France

ABSTRACT

This paper addresses the issues of the denoising and retrieval of the components of multicomponent signals from their short-time Fourier transform (STFT). After having recalled the hard-thresholding technique, in the STFT context, we develop a new thresholding technique by exploiting some limitations of the former. Numerical experiments illustrating the benefits of the proposed method to retrieve the modes of noisy multicomponent signals conclude the paper.

Index Terms— Time-frequency, STFT, denoising, thresholding, AM/FM multicomponent signals

1. INTRODUCTION

Linear time-frequency (TF) analysis provides a powerful tool for characterizing and manipulating signals whose statistics vary in time, but also multicomponent signals (MCSs), defined as a superposition of amplitude-and frequency-modulated (AM-FM) modes. Various methods are reported in the literature for modeling and analyzing such signals, as for instance those based on Hilbert-transform [1–3], empirical-mode-decomposition (EMD) [4, 5], or on the short-time Fourier transform (STFT), the latter being the simplest and most popular [6]. Nevertheless, it suffers from a serious limitation that the information it provides is spread over the TF plane as a result of the *uncertainty principle*. Many efforts have been made to cope with this issue, as for instance *reassignment methods* [7] of which synchrosqueezing transforms (SST) has received considerable attention in recent years [8, 9]. Indeed, not only do these transforms enhance the TF representation but also allow for mode reconstruction. However, the reassignment operators are sensitive to noise, and local frequency integration carried out directly on the STFT rather than on the reassigned transform appeared to be a good alternative to SST for mode reconstruction [10] [11]. In a more general context than that of MCSs, a widely used noise removal technique, originally designed in the wavelet context, is called *hard-thresholding* (HT) [12], which basically consists of thresholding the wavelet coefficients below

a certain level before signal reconstruction. In the present paper, we build on HT, adapted to the STFT context, to enable a more efficient mode retrieval of the modes of noisy MCSs.

To do so, after having introduced some useful definitions in Section 2, we introduce hard-thresholding (HT) denoising technique and then how we adapt it to the context of MCSs, in the following section. Then, we introduce, in Section 4, our new denoising technique designed by handling some of HT limitations. Finally, numerical simulations, carried out on both simulated and real signals in Section 5, demonstrate the improvements brought by the proposed technique with respect to state-of-the-art signal denoising algorithms.

2. DEFINITIONS

In this section, we first recall some basic definitions which will be useful in the sequel.

2.1. Discrete-Time STFT and Reconstruction Formula

Let us consider a discrete signal $f \in l_2(\mathbb{Z})$ of length L and a filter $g \in l_2(\mathbb{Z})$ supported on $\{-M, \cdots, M\}$ such that $2M + 1 < N$, where N is the number of frequency bins, the discrete-time short-time Fourier transform (STFT) can be written by [13]:

$$V_f^g(m, \frac{k}{N}) = \sum_{n \in \mathbb{Z}} f[n]g[n-m]e^{-i2\pi \frac{k(n-m)}{N}}.$$

If $\|g\|_2 = 1$, one has the following reconstruction formula:

$$f[n] = \sum_{m=n-M}^{m=n+M} \sum_{k=0}^{N-1} V_f^g(m, \frac{k}{N})g[n-m]\frac{e^{i2\pi \frac{k(n-m)}{N}}}{N}, \quad (1)$$

which requires to know $\left(V_f^g(m, \frac{k}{N})\right)_k$ for $-M \leq m \leq L - 1 + M$. To recover f using only $\left(V_f^g(m, \frac{k}{N})\right)_k$ for $0 \leq m \leq L-1$, we assume f is L-periodic instead of finite, thus making STFT also L-periodic, and leading to:

$$f[n] = \sum_{m=n-M}^{n+M} \sum_{k=0}^{N-1} V_f^g(m \bmod L, \frac{k}{N})g[n-m]\frac{e^{i2\pi \frac{k(n-m)}{N}}}{N}, \quad (2)$$

where $m \bmod L$ means m *modulo* L.

2.2. Multicomponent Signals

In what follows, we are going to analyze so-called *multicomponent signals* of the form,

$$f[n] = \sum_{l=1}^{K} f_l[n], \text{ with } f_l[n] = A_l[n]e^{i2\pi\phi_l[n]}, \ 0 \le n \le L-1$$

for some K, where $A_l[n] > 0$, $\phi_l'[n] > 0$ for any n and l. In the following, $A_l[n]$ is often called instantaneous amplitude (IA) of mode l and $\phi_l'[n]$ its instantaneous frequency (IF). Also, we assume the f_ls, called modes or components, are separated with resolution Δ, i.e. for all $l \in \{1, \cdots, K-1\}$:

$$\phi_{l+1}'[n] - \phi_l'[n] > 2\Delta \text{ for } \forall n.$$

One of the goals of TF analysis is to recover f_l from f.

3. SIGNAL DENOISING FROM STFT

3.1. Hard-Thresholding Denoising Technique

One of the most commonly used techniques for signal denoising is hard-thresholding (HT). It consists of computing and then thresholding the STFT of a noisy signal before reconstructing it [12, 13]. For instance, when f is a complex signal contaminated by a complex white Gaussian noise Φ with variance σ^2, namely $\tilde{f}[n] = f[n] + \Phi[n]$, $V_\Phi^g(m, \frac{k}{N})$ is also Gaussian with zero mean and satisfies:

$$\text{Var}\left(\Re\{V_\Phi^g(m, \frac{k}{N})\}\right) = \text{Var}\left(\Im\{V_\Phi^g(m, \frac{k}{N})\}\right) = \sigma^2\|g\|_2^2,$$

where $\Re\{Z\}$ (*resp.* $\Im\{Z\}$) is the real (*resp.* imaginary) part of complex number Z. Assuming the variance of the noise σ^2 is known, HT for STFT means to threshold coefficients as:

$$\overline{V}_{\tilde{f}}^g(m, \frac{k}{N}) = \begin{cases} V_{\tilde{f}}^g(m, \frac{k}{N}), & \text{if } |V_{\tilde{f}}^g(m, \frac{k}{N})| \ge 3\sqrt{2}\sigma\|g\|_{2,m} \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

where $\|g\|_{2,m}$ is the l_2 norm of the possibly truncated filter g (near the signal boundaries) used in the computation of $V_{\tilde{f}}^g(m, \frac{k}{N})_k$. Note that the normalization factor $\sqrt{2}$ is used because one thresholds the modulus of STFT and not its real or imaginary parts. Then, the reconstruction of the signal f is performed by replacing $V_f^g(m, \frac{k}{N})$ by $\overline{V}_{\tilde{f}}^g(m, \frac{k}{N})$ in (2).

3.2. Hard-Thresholding in the Context of MCSs

In the context of MCSs, since one is interested not only in denoising the signal but also in recovering the modes, one first proceeds with ridge extraction and then performs the denoising as explained hereafter.

3.2.1. Ridge Extraction from Noisy STFT

To estimate the ridge $(m, \phi_l'(m))$ and assuming knowledge of the number K of modes, we use the same algorithm as in [7] or [8], originally proposed in [14]. The approximation $(\psi_l)_{1 \le l \le K}$ of $(\phi_l')_{1 \le l \le K}$ is found as a local minimum of the functional:

$$\sum_{l=1}^{K} \sum_{m=0}^{L-1} -|V_{\tilde{f}}^g(m, \psi_l(m))|^2 + \lambda\psi_l'(m)^2 dt + \beta\psi_l''(m)^2,$$

where λ and β are chosen regularization parameters. It was shown in [15] through a numerical study, that to use regularization terms resulted in less accurate ridge estimation, so they will be set to zero in our simulations.

3.2.2. STFT Thresholding

Once an estimate of the ridges are found, we select the coefficient above the threshold provided by HT in the vicinity of the detected ridges, namely we define, for mode l:

$$\eta_{l,m}^{[1]} = \underset{k}{\arg\max}\left\{\frac{k}{N} < \psi_l(m), |V_{\tilde{f}}^g(m, \frac{k}{N})| < 3\sqrt{2}\sigma\|g\|_{2,m}\right\}$$

$$\eta_{l,m}^{[2]} = \underset{k}{\arg\min}\left\{\frac{k}{N} > \psi_l(m), |V_{\tilde{f}}^g(m, \frac{k}{N})| < 3\sqrt{2}\sigma\|g\|_{2,m}\right\},$$

and then set: $J_{l,m} := [N\psi_l(m) - \eta_{l,m}^{[1]}, N\psi_l(m) + \eta_{l,m}^{[2]}]$.

3.2.3. Mode Reconstruction

With this in mind the lth mode is reconstructed through:

$$f_l[n] = \sum_{m=n-M}^{n+M} \sum_{k \in J_{l,m}} V_f^g(m \bmod L, \frac{k}{N})g[n-m]\frac{e^{i2\pi\frac{k(n-m)}{N}}}{N}, \quad (4)$$

The HT process is illustrated on mono component signal in Figure 1: the STFT of a noisy linear chirp computed with a Gaussian window with variance $\sigma_s = 0.15$ is displayed in Figure 1 (a), along with the modulus of the noise-free, noisy and hard-thresholded STFTs at $m = L/2$, in Figure 1 (b). The threshold used is also displayed on this latter figure.

This technique suffers from some serious drawbacks, which can hardly be seen if one only studies the modulus of STFT. Indeed, on a ridge associated with one mode the real and imaginary parts of the STFT are oscillating, while its modulus is almost constant. This has the consequence that the effect of hard-thresholding is very different on the real part and imaginary parts of the STFT: an illustration of this is given in Figure 1 (c) and (d) where we display them still at time $m = L/2$ (the reverse situation could be obtained by selecting a different time). In fact, when the imaginary and real parts of the STFT on the ridge have very different amplitudes, the one with the lowest amplitude may be of the order or below the noise level, even in the vicinity of the ridge (see Figure 1 (d)). In these instances, the noise may shift the local extremum of either the real or imaginary parts of the STFT,
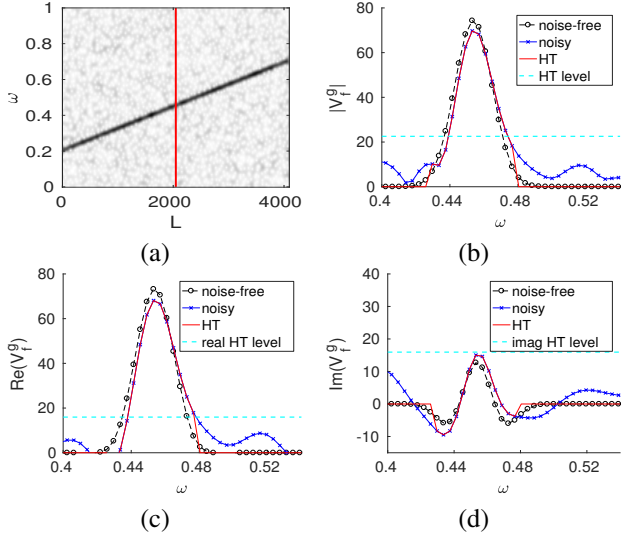
Fig. 1. Illustration of HT for signal denoising: (a): modulus of the STFT of a noisy linear chirp (SNR = 0 dB); (b): modulus of STFTs (noise-free, noisy and hard-thresholded) at $m = L/2$ along with the noise threshold; (c): the real part of STFT displayed in (b); (d): same as (c) but for the imaginary part.

depending on their relative amplitudes, and may create symmetry distortion. These phenomena are however not taken into account by hard-thresholding techniques. Furthermore, the STFT of f is regular with respect to the frequency, as a result of the properties of the convolution, and the hard-thresholded transform is always irregular along the frequency axis. After having explained how we compute in practice the noise threshold in the next subsection, we will explain why the real and the imaginary parts of STFT should attain a local extremum along the frequency axis at the instantaneous frequency corresponding to a mode, while being symmetric with respect to that point. This will be crucial in the description of the new denoising algorithm that follows.

3.3. Threshold Estimation

In real-life applications, HT level $\gamma := \sigma \|g\|_2$ needs to be estimated. To handle this, we use, in this paper, the robust estimator proposed in [12]:

$$\hat{\gamma}_p = \frac{\underset{0 \leq k \leq N-1}{\text{median}} \left| \Re \left(V_{\tilde{f}}^g (m \bmod L, \frac{k}{N}) \right)_m \right|}{0.6745},$$

where median represents the median of the coefficients.

4. A NEW METHOD FOR THE DENOISING OF THE STFT OF MCSS

In this section, we introduce a new denoising technique, called *Shifted-Symmetrized-Regularized Hard-Thresholding* (SSR-HT), which aims at compensating for the three above

mentioned flaws of HT: potential shift of the thresholded transform, lack of symmetry and regularity. To explain why the transform need to be symmetric with respect to the instantaneous frequency of a mode along the frequency axis, we use the local linear chirp approximation already exploited in [6] and then detail the novel algorithm for STFT denoising.

4.1. Local Linear Chirp Approximation

Let us consider the second-order approximation close to the lth ridge, of the STFT defined as [6]:

$$V_f^g(m, \frac{k}{N}) \approx f_l[m] \widehat{gc_t}(\frac{k}{N} - \phi_l'(m)), \qquad (5)$$

where $f_l[m] = A_l[m] e^{i2\pi \phi_l(m)}$ and $c_t(\tau) = e^{i\pi \phi_l''(m)\tau^2}$. When g is the Gaussian window $\sigma_s^{-\frac{1}{2}} e^{-\pi \frac{x^2}{\sigma_s^2}}$, this approximation leads to [6]:

$$V_f^g(m, \frac{k}{N}) \approx \left| V_{f_l}^g(m, \frac{k}{N}) \right| e^{i2\pi \Psi_l(m, \frac{k}{N})}, \qquad (6)$$

where $\Psi_l(m, \frac{k}{N}) = -\frac{\phi_l''(m)}{2r(m)^2} \left(\frac{k}{N} - \phi_l'(m) \right)^2 - \frac{\theta(m)}{4\pi} + \phi_l(m)$ with $\theta(m) = \arctan(-\phi''(m)\sigma_s^2)$ and $r(m) = \sigma_s^{-2} \left(1 + \phi''(m)^2 \sigma_s^4 \right)^{\frac{1}{2}}$, and where

$$\left| V_{f_l}^g(m, \frac{k}{N}) \right| = A_l(m)\sigma_s^{-\frac{1}{2}} r(m)^{-\frac{1}{2}} e^{-\frac{\pi(\omega - \phi_l'(m))^2}{\sigma_s^2 r(m)^2}}.$$

We thus remark that in the noise-free case, and for a fixed time m, $\left| V_f^g(m, \frac{k}{N}) \right|$, $\Re \left(V_f^g(m, \frac{k}{N}) \right)$ and $\Im \left(V_f^g(m, \frac{k}{N}) \right)$ are even functions in the vicinity of the lth ridge, and all reach an extremum at $\frac{k}{N} = \phi_l'(t)$. However, these properties disappear for noisy signal, see Figure 1. In the designing of the denoising algorithm that follows, we are going to exploit these properties of the STFT.

4.2. Novel Algorithm for MCS Denoising

The key idea of SSR-HT is to improve the representation of the real and imaginary parts of the hard-thresholded noisy STFT, at each time instant. It is important to first explain that we deal with the modes one at a time, that is we improve the STFT of one mode and then proceed with its reconstruction. More precisely, to deal with the shifts induced by noise on the STFT of mode l (see Figure 1 (d)), we propose, at each time m, to reallocate the closest extremum to $\psi_l(m)$ of both the imaginary and real parts of STFT to $\psi_l(m)$. Having shifted the real and imaginary parts, we symmetrize them by replacing their values along the frequency axis and on each side of $\psi_l(m)$, more precisely in the interval $\widehat{J}_{l,m} = [\psi_l(m) - \frac{\eta}{N}, \psi_l(m) + \frac{\eta}{N}]$ with $\eta = \max(\eta_{l,m}^{[1]}, \eta_{l,m}^{[2]})$, by the average of its value and of that at the symmetric point with respect to $\psi_l(m)$. This enables to remove the symmetry distortion induced by noise on the STFT. Finally, to deal with

the lack of regularity generated by the hard-thresholding process, we use piecewise cubic Hermite interpolation to smooth the resulting real and imaginary parts of STFT. From now on, we call this process SSR-HT for "Shifted-Symmetrized-Regularized Hard-Thresholding". After having computed the denoised STFT $\widehat{V}^g_{\tilde{f}}$ via SSR-HT, the reconstructed mode l is obtained by replacing $V^g_{\tilde{f}}$ by $\widehat{V}^g_{\tilde{f}}$ and $J_{l,m}$ by $\widehat{J}_{l,m}$ in (4).

The details of the novel denoising algorithm are given in Algorithm 1, for the extraction of the lth mode. As an illustration, the result of SSR-HT on the real and imaginary parts of STFT displayed in Figure 1 (c) and (d) is shown in Figure 2. In this figure, one notices a clear improvement of the denoised STFT brought by SSR-HT compared with that resulting from the application of HT. We are going to quantify this in the following section.


Algorithm 1 SSR-HT Algorithm

1: **for** all time m **do**
2:     **Input:** $\eta^{[1]}_{l,m}, \eta^{[2]}_{l,m}, V^g_{\tilde{f}}(t_m, \omega), \psi_l(m)$
3:     $\eta = \max(\eta^{[1]}_{l,m}, \eta^{[2]}_{l,m})$,
4:     $\widehat{J}_{l,m} = [\psi_l(m) - \frac{\eta}{N}, \psi_l(m) + \frac{\eta}{N}]$
5:     **Shift operation:**
6:     **if** $\widehat{J}_{l,m} \in [0,1]$ **then**
7:         $S_R(m) = \psi_l(m) - \underset{\omega \in \widehat{J}_{l,m}}{\mathrm{argmax}} \left| \Re(V^g_{\tilde{f}}(m,\omega)) \right|$
8:         $S_I(m) = \psi_l(m) - \underset{\omega \in \widehat{J}_{l,m}}{\mathrm{argmax}} \left| \Im(V^g_{\tilde{f}}(m,\omega)) \right|$
9:
10:         Shift $\Re(V^g_{\tilde{f}}(m,.))$ by $S_R(m)$
11:         Shift $\Im(V^g_{\tilde{f}}(m,.))$ by $S_I(m)$
12:     **Symmetrization:** for $\nu \in [0,\eta]$ compute:
13:     $\dfrac{\left(V^g_{\tilde{f}}(m, \psi_l(m) + \frac{\nu}{N}) + V^g_{\tilde{f}}(m, \psi_l(m) - \frac{\nu}{N}) \right)}{2} \to X$
14:     $X \to V^g_{\tilde{f}}(m, \psi_l(m) \pm \frac{\nu}{N})$
15:     **Smoothing step:**
16:     smooth real and imaginary parts with piecewise cubic
17:     Hermite interpolation.
18:   **Output:** $\widehat{V}^g_{\tilde{f}}(m,\omega)$
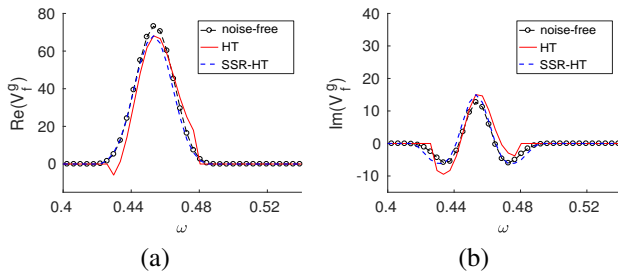
Fig. 2. Illustration of SSR-HT on the real and imaginary parts of the STFT displayed in Fig.1 (c) and (d).

5. NUMERICAL EXPERIMENTS

This section investigates the performance of the new denoising technique by comparing it, on both a simulated MCS and a set of real signals, with hard-thresholding (HT) and Block-Thresholding [16], denoted by BT hereafter. The Matlab code implementing the method and the scripts generating all the figures of this paper can be found `github.com/phamduonghung/ICASSP2018`.

Let us first consider a simulated MCS composed of two cosine chirps containing high frequency modulations. We first display the modulus of the STFT of the signal in Figure 3 (a). This signal is then contaminated by some additive white Gaussian noises such that input Signal-to-Noise Ratio (SNR) range is from -5 to 30 dB. Denoising results, depicted in Figure 3 (b), display the output SNR associated with the reconstruction of the two modes with respect to input SNR, and show that SSR-HT clearly outperforms the other denoising techniques: the gain is 2dB on average compared with HT.
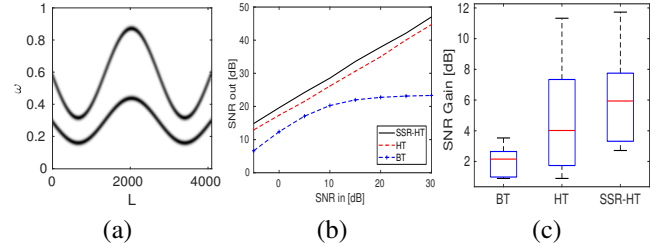
Fig. 3. Denoising performance of three techniques: a): STFT of the simulated MCS; (b): on the signal displayed in (a); (c) on real noisy PCG signals (using boxplot [17]).

Finally, we study the performance of the proposed technique on noisy phonocardiogram (PCG) signals associated with the database used in the sixth Signal Separation Evaluation Campaign (SiSEC 2016) [18]. This database contains 16 noisy and noise-free signals, and we thus measure the quality of reconstruction in terms of SNR gain (i.e. output SNR-input SNR). The denoising results correspond to the boxplots in Figure 3 (c): while the SNR gain with SSR-HT is nearly 6, it equals 4 and 2 with HT and BT respectively, which confirms the potential interest of our new technique for real signals denoising.

6. CONCLUSION

We introduced in this paper a novel technique to denoise the short-time Fourier transform of multicomponent signals, inspired by the flaws of hard-thresholding. We have then used the denoised transform to perform mode reconstruction by local integration taking into account local frequency modulation and noise level. Numerical experiments demonstrated the effectiveness of the approach on multicomponent signals with cosine phases and a database of real PCG signals.

REFERENCES

[1] M. Feldman, "Time-varying vibration decomposition and analysis based on the hilbert transform," *Journal of Sound and Vibration*, vol. 295, pp. 518–530, 2006.

[2] F. Gianfelici, G. Biagetti, P. Crippa, and C. Turchetti, "Multicomponent am–fm representations: an asymptotically exact approach," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 3, pp. 823–837, 2007.

[3] G. Biagetti, P. Crippa, A. Curzi, S. Orcioni, and C. Turchetti, "Analysis of the emg signal during cyclic movements using multicomponent am–fm decomposition," *IEEE journal of biomedical and health informatics*, vol. 19, no. 5, pp. 1672–1681, 2015.

[4] M. Feldman, "Analytical basics of the emd: Two harmonics decomposition," *Mechanical Systems and Signal Processing*, vol. 23, no. 7, pp. 2059–2071, 2009.

[5] R. Sharma, L. Vignolo, G. Schlotthauer, M. A. Colominas, H. L. Rufiner, and S. Prasanna, "Empirical mode decomposition for adaptive am-fm analysis of speech: a review," *Speech Communication*, vol. 88, pp. 39–64, 2017.

[6] T. Oberlin, S. Meignen, and V. Perrier, "The fourier-based synchrosqueezing transform," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 315–319.

[7] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Transactions on Signal Processing*, vol. 43, no. 5, pp. 1068–1089, 1995.

[8] 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.

[9] D.-H. Pham and S. Meignen, "High-order synchrosqueezing transform for multicomponent signals analysis - with an application to gravitational-wave signal," *IEEE Transactions on Signal Processing*, vol. 65, no. 12, pp. 3168–3178, June 2017.

[10] S. Meignen, T. Oberlin, and S. McLaughlin, "Multi-component signal denoising with synchrosqueezing," in *Statistical Signal Processing Workshop (SSP), 2012 IEEE*, 2012, pp. 660–663.

[11] T. Oberlin, S. Meignen, and S. McLaughlin, "A novel time-frequency technique for multicomponent signal denoising," in *21st European Signal Processing Conference*, May 2013, pp. 5358–5362.

[12] D. Donoho and I. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, 1994.

[13] S. Mallat, *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*, 3rd ed. Academic Press, 2008.

[14] R. Carmona, W. Hwang, and B. Torresani, "Characterization of signals by the ridges of their wavelet transforms," *IEEE Transactions on Signal Processing*, vol. 45, no. 10, pp. 2586–2590, Oct 1997.

[15] S. Meignen, D.-H. Pham, and S. McLaughlin, "On demodulation, ridge detection and synchrosqueezing for multicomponent signals," *IEEE Transactions on Signal Processing*, vol. 65, no. 8, pp. 2093–2103, 2017.

[16] G. Yu, S. Mallat, and E. Bacry, "Audio denoising by time-frequency block thresholding," *IEEE Transactions on Signal Processing*, vol. 56, no. 5, pp. 1830–1839, May 2008.

[17] R. McGill, J. W. Tukey, and W. A. Larsen, "Variations of box plots," *The American Statistician*, vol. 32, no. 1, pp. 12–16, 1978.

[18] A. Liutkus, F.-R. Stöter, Z. Rafii, D. Kitamura, B. Rivet, N. Ito, N. Ono, and J. Fontecave, "The 2016 signal separation evaluation campaign," in *International Conference on Latent Variable Analysis and Signal Separation*. Springer, 2017, pp. 323–332.