Detecting and localizing interference in theTF plane

S. Meignen[◇], M. Colominas [⊥], N. Laurent [‡], T. Oberlin [†]

[◇] LJK, UGA, [⊥] University of Entre-Rios, Argentina,
[‡] Gipsa-Lab, UGA, [†] Supaero, Univ. Toulouse

the $9^t h$ of november 2023

Outline

- For $f \in L^1(\mathbb{R})$, STFT defined for a sliding window $g \in L^\infty(\mathbb{R})$ by:

$$V_f^g(t,\xi) = \int_{\mathbb{R}} f(\tau)g(\tau - t)e^{-2i\pi\xi(\tau-t)}d\tau.$$

The spectrogram is the squared modulus of the STFT.

- Multicomponent signals (MCSs) defined as a superimposition of AM-FM components or modes, used to model non stationary signals:

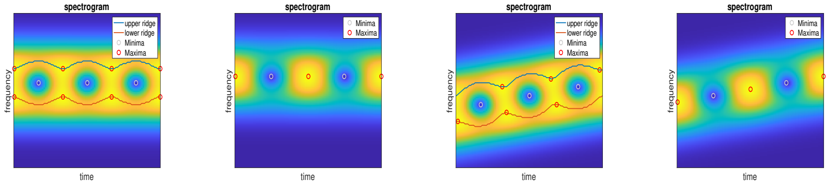
$$f(t) = \sum_{p=1}^{P} f_p(t) \quad with \;\; f_p(t) = A_p(t)e^{i2\pi\phi_p(t)}.$$

$A_p(t) > 0, \phi_p'(t) > 0$ and $\phi_{p+1}'(t) > \phi_p'(t)$ for all t.

Ideal TF (ITF) representation: $\mathrm{TI}_f(t,\omega) = \sum_{p=1}^{P} A_p(t)\delta\left(\omega - \phi_p'(t)\right).$

Pure tones separation from the spectrogram

- A mode is associated with a chain of local maxima of the spectrogram along the frequency axis (LMFs), called a ridge.
- In the presence of several modes, it is not always possible to associate with each mode a ridge, presence of time-frequency bubbles in case of strong interference.
- It is however possible to give a necessary and sufficient condition for two pure tones to create exactly two separate ridges.



- In the case of a two pure tones signal, and using the Gaussian window $g(t) = e^{-\pi \frac{t^2}{\sigma^2}}$, one has the following result:

Proposition

Let $f(t) = f_1(t) + f_2(t)$ with $f_1(t) = Ae^{i2\pi\xi_1 t}$ and $f_2(t) = e^{i2\pi\xi_2 t}$, where $\xi_1 < \xi_2$ and $A > 0$. The modes f_1 and f_2 are associated with two chains of LMFs if and only if:

$$\alpha := \sqrt{\frac{\pi}{2}}\sigma(\xi_2 - \xi_1) > 1 \text{ and}$$
$$|\log(A)| < -2\,\mathrm{arcosh}(\alpha) + 2\alpha\sqrt{\alpha^2 - 1}.$$

- This proposition can be extended to parallel linear chirps.

[1] S. Meignen, N. Laurent and T. Oberlin, " One or Two Ridges? An Exact Mode Separation Condition for the Gabor Transform", IEEE Signal Processing Letters, vol. 29, pp. 2507-2511, 2022.

Localizing interference in the TF plane

- The previous proposition is of little practical interest because it requires the knowledge of the signal parameters.
- It is also restrictive in the sense it gives an analytic solution only for pure harmonics or parallel linear chirps.
- We are going to explain how to localize interference corresponding to when it is not possible to assign a ridge to each mode, by means of a specific ridge detector.
- For that purpose, we need to give a brief overview of ridge detection.

Ridge Point Definition

- Let us define

$$V_f^g(t,\eta) = \quad |V_f^g(t,\eta)| e^{2i\pi\Psi(t,\eta)}, \qquad (1)$$

- First characterization of a ridge point $\partial_t \Psi(t,\eta) = \eta$.
- Alternatively, a ridge point can be viewed as a LMF, namely a TF point such that $\partial_\eta |V_f^g(t,\eta)|^2 = 0$,
- When $g(t) = e^{-\pi \frac{t^2}{\sigma^2}}$, then $g'(t) = -\frac{2\pi}{\sigma^2} t g(t)$, and thus:

$$\begin{aligned}
\partial_t \Psi(t,\eta) &= \frac{1}{2\pi} \Im\left\{ \frac{\partial_t V_f^g(t,\eta)}{V_f^g(t,\eta)} \right\} \\
&= \eta - \frac{1}{2\pi} \Im\left\{ \frac{V_f^{g'}(t,\eta)}{V_f^g(t,\eta)} \right\} = \eta + \Im\left\{ \frac{1}{\sigma^2} \frac{V_f^{tg}}{V_f^g} \right\} \qquad (2) \\
&= \eta - \Im\left\{ \frac{\partial_\eta V_f^g}{2i\pi\sigma^2 V_f^g} \right\} = \eta + \frac{1}{4\pi\sigma^2} \frac{\partial_\eta |V_f^g|^2}{|V_f^g|^2},
\end{aligned}$$

$\Im\{X\}$ denotes the imaginary part of the complex number X.

Classical ridge detection

- Every mode of a MCS occupies a "ribbon" around its IF $\phi'_p(t)$, classically estimated by means of ridge detection (RD).
- RD carried out by applying for p from 1 to P the following peeling algorithm (discrete TF formalism):

$$\max_{c_p} \sum_{n=0}^{N-1} |V_{f,p}^g[n, c_p[n]]|^2, \quad \text{s.t. } |\Delta c_p[n]| \leq B_f,$$

where c_p is the p^{th} estimated ridge, $\Delta z[n] = z[n+1] - z[n]$, while B_f upper bound for the frequency modulation of the modes.
- $V_{f,p}^g$ is recursively defined as follows:

$$V_{f,p}^g[n,q] = \begin{cases} 0, \text{ if } q \in [c_p[n] - B_f, c_p[n] + B_f] \\ V_{f,p-1}^g[n,q], \text{ otherwise,} \end{cases}$$

where $V_{f,1}^g = V_f^g$.
- Lack of adaptivity since B_f kept *constant* regardless of the frequency modulation of the modes.
- The ridges thus detected are not made of discrete LMFs.

Adaptive ridge detection

- Ridges can alternatively be detected using a local chirp-rate estimate.
- It is based on complex reassignment operators

$$\tilde{\omega}_f(t,\xi) = \frac{\partial_t V_f^g(t,\xi)}{2i\pi V_f^g(t,\xi)} \text{ and } \tilde{t}_f(t,\xi) = t - \frac{\partial_\xi V_f^g(t,\xi)}{2i\pi V_f^g(t,\xi)},$$

and also on the following complex frequency modulation operator:

$$\tilde{q}_f(t,\xi) = \frac{\partial_t \tilde{\omega}_f(t,\xi)}{\partial_t \tilde{t}_f(t,\xi)} = \frac{\partial_t \left(\frac{\partial_t V_f^g(t,\xi)}{V_f^g(t,\xi)} \right)}{2i\pi - \partial_t \left(\frac{\partial_\xi V_f^g(t,\xi)}{V_f^g(t,\xi)} \right)}.$$

- $\hat{q}_f(t,\xi) = \Re\left\{\tilde{q}_f(t,\xi)\right\} = \phi''(t)$, when f is a Gaussian modulated linear chirp ($f(t) = A(t)e^{2i\pi\phi(t)}$, with $\phi(t) = at + \frac{c}{2}t^2$ and $\log(A)$ also polynomial with degree at most 2).

First approach

- The idea of modulation-based ridge detector (MB-RD) is to extract the ridges one after the other using the modulation operator $\widehat{q}_f$.
- The latter is discretized on a TF grid, on which n is the time index, ranging from 0 to $N-1$ and k is the frequency index, ranging from 0 to $M-1$. $\widehat{q}_f$ associated with an $N \times M$ matrix $\widehat{\mathbf{q}}_f$.
- The signal f lasts for T seconds, the time index n corresponds to the time $\frac{n}{N}T$, and the frequency index k to the frequency $\frac{k}{M}\frac{N}{T}$.
- To build the first ridge, denoted by the vector φ of frequency indices of length N, one picks a time index n, and computes

$$\varphi[n] = \underset{0 \leq k \leq M-1}{\operatorname{argmax}} |\mathbf{V}_f^g[n,k]|, \tag{3}$$

where $\mathbf{V}_f^g$ is the matrix corresponding to the discretization of V_f^g on the TF grid.

- Then, the next point on the ridge is computed as:

$$\varphi[n+1] := \underset{k,\ |k-\varphi[n]-\frac{MT^2}{N^2}\widehat{\mathbf{q}}_f[n,\varphi[n]]|\leq C}{\operatorname{argmax}} |\mathbf{V}_f^g[n+1,k]|. \tag{4}$$

- Performing a first order Taylor expansion of ϕ_p', assuming f_k is a linear chirp, one obtains:

$$\phi_p'\left(\frac{n+1}{N}T\right) = \phi_p'\left(\frac{n}{N}T\right) + \frac{T}{N}\phi_p''\left(\frac{n}{N}T\right).$$

Now, if one assumes $\phi_p'(\frac{n}{N}T) \approx \frac{\varphi[n]}{M}\frac{N}{T}$ for some k, one has:

$$\begin{aligned}\frac{\varphi[n+1]}{M}\frac{N}{T} \approx \frac{\varphi[n]}{M}\frac{N}{T} + \frac{T}{N}\phi_p''\left(\frac{n}{N}T\right) \approx \frac{\varphi[n]}{M}\frac{N}{T} + \frac{T}{N}\widehat{\mathbf{q}}_f[n,\varphi[n]] \\ \Leftrightarrow \varphi[n+1] \approx \varphi[n] + \frac{MT^2}{N^2}\widehat{\mathbf{q}}_f[n,\varphi[n]],\end{aligned} \tag{5}$$

which justifies the range for k in Eq. (4).

- The parameter C is then used to cope with potential errors in the approximation given by Eq. (5).

[2] M. Colominas, S. Meignen and D-H. Pham, "Fully Adaptive Ridge Detection Based on STFT Phase Information", IEEE Signal Processing Letters, vol. 27, no. 1, pp. 620-624, 2020.

- The ridge obtained that way not necessarily made of discrete LMFs.
- $\widehat{\mathbf{q}}_f$ not necessarily accurate (local linear chirp approximation).
- RD is continued even if the detected points are irrelevant, and this often happens in noisy situations.
- Finally, the detected ridge depends on the initialization time index n.

- To deal with all these issues, the concept of *relevant ridge portions* (RRPs) was introduced in [3].
- To define a generic RRP, denoted by φ, one first selects a time index n and a frequency index k, such that $[n, k]$ is an LMF and sets $\varphi[n] := k$. Then, one defines:

$$F[k] := k + \frac{MT^2}{N^2}\widehat{\mathbf{q}}_f[n, k], \tag{6}$$

meaning that $\varphi[n+1] \approx F[\varphi[n]]$, when the modulation operator accurately estimates the chirp-rate on the ridge.
- Conversely, assuming $\varphi[n+1]$ is known, one also has the relation $\varphi[n] \approx B[\varphi[n+1]]$, with

$$B[k] := k - \frac{MT^2}{N^2}\widehat{\mathbf{q}}_f[n+1, k]. \tag{7}$$

[3] N. Laurent and S. Meignen," A Novel Ridge Detector for Non Stationary Multicomponent Signals: Development and Application to Robust Mode Retrieval", IEEE TSP, vol. 69, pp. 3325-3336, 2021.

Using these notations, $\varphi[n+1]$ is then defined as satisfying $[n, \varphi[n]] \sim [n+1, \varphi[n+1]]$, where the relation $\sim$ corresponds to the following definition :

Definition

Let $\mathbf{m}$ be a vector with values in $[\![0, M-1]\!]$, and $[n, \mathbf{m}[n]]$ and $[n+1, \mathbf{m}[n+1]]$ two LMFs, then:

$$[n, \mathbf{m}[n]] \sim [n+1, \mathbf{m}[n+1]]$$
$$\Leftrightarrow \begin{cases} \mathbf{m}[n+1] & := \underset{k,\ [n+1,k]\ \text{LMF}}{\arg\min} \left|k - F[\mathbf{m}[n]]\right| \\ \mathbf{m}[n] & := \underset{k,\ [n,k]\ \text{LMF}}{\arg\min} \left|B[\mathbf{m}[n+1]] - k\right|, \end{cases}$$

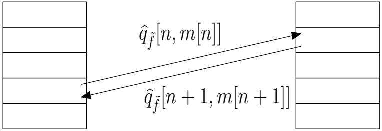
with the functions F and B defined in Eq. (6) and (7), respectively.

- $[n+1, \varphi[n+1]]$ (resp. $[n, \varphi[n]]$) is the closest LMF to $[n, \varphi[n]]$ (resp. $[n+1, \varphi[n+1]]$) at time index $n+1$ (rest. n) in the direction given by $\widehat{\mathbf{q}}_f[n, \varphi[n]]$ (resp. $-\widehat{\mathbf{q}}_f[n+1, \varphi[n+1]]$).
- $[n, \varphi[n]] \sim [n+1, \varphi[n+1]]$ means that $\widehat{\mathbf{q}}_f$ computed at these LMFs correspond to a stable orientation.
- The relation $\sim$ is then used to define φ iterating the procedure forward and backward, from time index n. Note that, when the relation $\sim$ cannot be satisfied at a time index, the detection procedure stops, which is why one uses the term "ridge portion", hence the notation RRP.

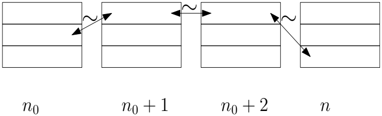
Illustrations of the procedure

- The *constraints* between LMFs can be viewed as:

$[n, m[n]] \sim [n+1, m[n+1]]$



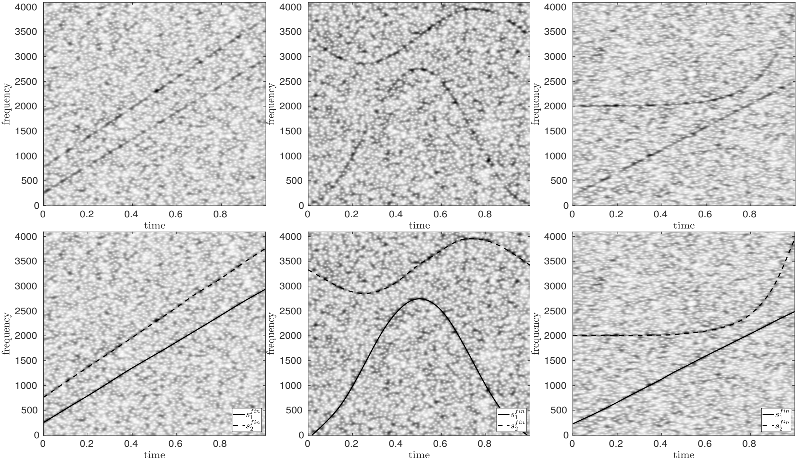
$[n_0, m[n_0]] \leftrightarrow [n, m[n]]$



- A relevant ridge portion(RRP) $\mathcal{R}_i$ is a finite set of LMFs sharing relation $\leftrightarrow$ (when the relation $\leftrightarrow$ no longer satisfied the ridge is interrupted).

Illustrations

It is possible to gather together RRPs in a noisy environment and reconstruct the ridges using a variational approach



Limitations of RRPs for strong interference detection

- A ridge point cannot belong to several RRPs, and one slightly changes the definition of the ridge portions to allow for their merging.
- The new type of ridge portions are now denoted by ERRPs (for *extended relevant ridge portions*) and are slightly different from RRPs.
- To build a first ERRP, which we denote by φ_1, we consider the set

$$\mathcal{D}_1 = \{[n,k]\ LMF,\ \ n \in [\![0, N-1]\!],\ \ k \in [\![0, M-1]\!]\}\,, \tag{8}$$

and then an initial point:

$$[n_1, \varphi_1[n_1]] = \underset{[n,k]\in\mathcal{D}_1}{\operatorname{argmax}} |\boldsymbol{V}_f^g[n,k]|. \tag{9}$$

- Starting with $[n_1, \varphi_1[n_1]]$, the associated ERRP corresponds to the RRP passing through that point, plus the two ending points at which the relation $\sim$ is no longer satisfied.
- The ending points can belong to several chains of LMFs, contrary to the other points on the ERRP.

- Having defined φ_1, its associated TF domain reads:

$$\mathcal{E}(\varphi_1) := \{[n, \varphi_1[n]],\ \varphi_1[n] \text{ defined}\},$$

- One introduces $\mathcal{D}_2 := \mathcal{D}_1 \setminus \mathcal{E}(\varphi_1)$, then defines

$$[n_2, \varphi_2[n_2]] := \operatorname*{argmax}_{[n,k]\in\mathcal{D}_2} |\boldsymbol{V}_f^g[n,k]|,$$

  finally, φ_2 is detected following the same procedure as for φ_1. Associated TF domain $\mathcal{E}(\varphi_2) := \{[n, \varphi_2[n]],\ \varphi_2[n] \text{ defined}\}$.
- The following ERRPs (for $p \geq 3$), are computed iteratively:

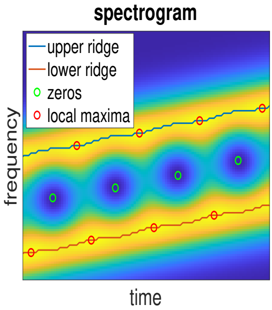
$$\mathcal{D}_p := \mathcal{D}_{p-1} \setminus \mathcal{E}(\varphi_{p-1}),$$

  and then, starting with

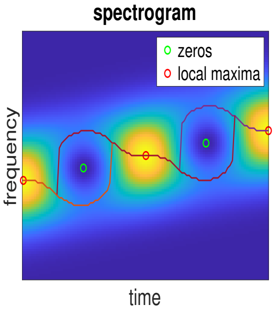
$$[n_p, \varphi_p[n_p]] := \operatorname*{argmax}_{[n,k]\in\mathcal{D}_p} |\boldsymbol{V}_f^g[n,k]|,$$

  the p^{th} ERRP is detected in the same way as for φ_1, and $\mathcal{E}(\varphi_p)$ correspond to its TF domain.
- Such a procedure (ERRP-RD) is carried out until the detected ERRP has a length below some predefined threshold.

Illustration



(a)



(b)

Noisy Case

- One considers $\tilde{f} = f + \varepsilon$, with ε a complex Gaussian white noise.
- Detection of ERRPs initialized using only LMFs associated with the signal part of the spectrogram with a high probability.
- Assuming the variance of the noise is σ_ε^2, $\boldsymbol{V}_\varepsilon^g[n,k]$ is Gaussian with zero mean and satisfies:

$$\mathrm{Var}\left(\Re\{\boldsymbol{V}_\varepsilon^g[n,k]\}\right) = \mathrm{Var}\left(\Im\{\boldsymbol{V}_\varepsilon^g[n,k]\}\right) = \sigma_\varepsilon^2 \|g\|_2^2.$$

- $\frac{|\boldsymbol{V}_\varepsilon^g|^2}{\sigma_\varepsilon^2 \|g\|_2^2}$ is χ_2 distributed with two degrees of freedom, the probability that $|\boldsymbol{V}_\varepsilon^g[n,k]| \geq \beta \sigma_\varepsilon \|g\|_2$ is lesser than 1% if $\beta = 3$.
- To estimate $\gamma = \sigma_\varepsilon \|g\|_2$, we use the robust estimator proposed in:

$$\hat{\gamma} := \frac{\mathrm{median}\left|\Re\left\{\boldsymbol{V}_{\tilde{f}}^g[n,k]\right\}_{n,k}\right|}{0.6745}.$$

- Based on this analysis, one defines:

$$\mathcal{S}(\beta) := \left\{[n,k], |\boldsymbol{V}_{\tilde{f}}^g[n,k]| \geq \beta\hat{\gamma}\right\}, \tag{10}$$

One considers that only the LMFs in $\mathcal{S}(3)$ to detect ERRPs.

- To detect the first ERRP, one uses the same algorithm as previously, but starting this time with

$$[n_1, \varphi_1[n_1]] := \underset{[n,k]\in\tilde{\mathcal{D}}_1}{\operatorname{argmax}} |\boldsymbol{V}_{\tilde{f}}^{g}[n,k]|,$$

with $\tilde{\mathcal{D}}_1 = \mathcal{D}_1 \cap \mathcal{S}(3)$, and then performing ERRP detection using points in $\tilde{\mathcal{D}}_1$. The set $\mathcal{E}(\varphi_1)$ being defined as in the noiseless case, we put:

$$\tilde{\mathcal{D}}_2 = \tilde{\mathcal{D}}_1 \setminus \mathcal{E}(\varphi_1).$$

- Then, the detection of φ_2 follows the same procedure as in the noiseless case, replacing $\mathcal{D}_2$ by $\tilde{\mathcal{D}}_2$ to find the initial points, and bearing in mind that the points on the ERRP have to be in $\tilde{\mathcal{D}}_1$.
- Finally, the detection of the next ERRPs ($p \geq 3$) involves points in $\tilde{\mathcal{D}}_1$ and is based on the same framework as in the noiseless case, replacing $\mathcal{D}_p$ by $\tilde{\mathcal{D}}_p := \tilde{\mathcal{D}}_{p-1} \setminus \mathcal{E}(\varphi_{p-1})$.

Time-Frequency Bubbles

- A particular type of strong interference in the TF plane corresponds to the notion of Time-Frequency Bubbles (TFBs).
- They occur when the signal is locally associated with a "circular" set of LMFs in the TF plane (interference between two modes).
- Such structures can also be present in noise, as a result of the interence between two *logons*.
- To clarify this notion of TFBs in our discrete TF setting, we propose the following definition using ERRPs:

Definition

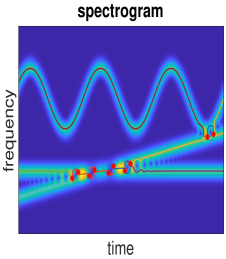
Two ERRPs create a TFB when they have two points in common, and when, in the region delimited by these ERRPs between these two points, there is a single zero of the spectrogram.

Algorithm to detect TFBs

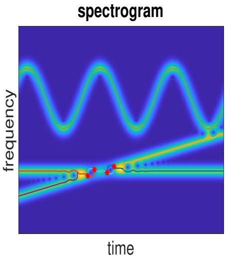
1. Find the ERRPs that have two points in common.
2. Compute the number of zeros of the spectrogram inside the TF domain delimited by these two points and the associated ERRPs.
3. If this number equals one, this pair of points is associated with a TFB.

We coin such a pair of points *TFB points*, as they localize a TFB in the TF plane.

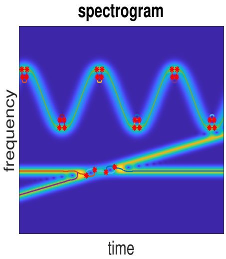
IIllustration



(a):small σ



(b):intermediate σ



(c): large σ

The detection of TFB points is helpful to find an appropriate σ.

Singular points detection in spectrograms of polyharmonic signals

- Another application of ERRP-RD is to localize *singular points* in polyharmonic signals, such as voice signals, that a priori do not contain any mode crossings.
- By singular points, we recall that we mean TF locations where two ERRPs merge (without being necessarily a TFB point).
- ERRP detection enables to localize such points, the number of which varies with respect to σ.

On the relation between singularities of the modulation operator and TFBs

Proposition

Let $f(t) = f_1(t) + f_2(t)$ with $f_1(t) = Ae^{i2\pi\xi_1 t}$ and $f_2(t) = e^{i2\pi\xi_2 t}$, where $\xi_1 < \xi_2$, the singularities of $\widehat{q}_f$ are the TF points (t,η) satisfying:

$$V_f^g(t,\eta)V_f^{t^2g}(t,\eta) - (V_f^{tg}(t,\eta))^2 = 0. \tag{11}$$

Now, if we further assume that $\alpha := \sqrt{\frac{\pi}{2}}\sigma(\xi_2 - \xi_1) \leq 1$, the singularities of $\widehat{q}_f$ are located at $(t_{k,1},\eta^*)$ or $(t_{k,2},\eta^*)$, with

$$\eta^* = \frac{\xi_1 + \xi_2}{2} + \frac{\log(A)}{2\pi\sigma^2(\xi_2 - \xi_1)}$$

$$t_{k,1} = \frac{k - \frac{\arccos(-1+2\alpha^2)}{2\pi}}{\xi_2 - \xi_1} \; and \; t_{k,2} = \frac{k + \frac{\arccos(-1+2\alpha^2)}{2\pi}}{\xi_2 - \xi_1}$$

with $k \in \mathbb{Z}$.

It is interesting to analyze where these singularities are located with respect to the zeros and the local maxima of the spectrogram. As, the spectrogram reads:

$$\begin{aligned}|V_f^g(t,\eta)|^2 =& A^2\hat{g}^2(\eta-\xi_1)+\hat{g}^2(\eta-\xi_2)\\ &+2A\hat{g}(\eta-\xi_1)\hat{g}(\eta-\xi_2)\cos(2\pi(\xi_2-\xi_1)t),\end{aligned}$$

we can derive the following:

Proposition

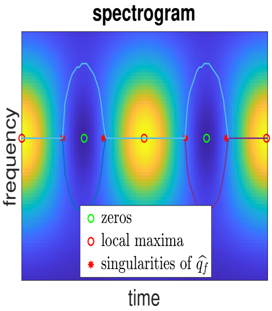
Considering the signal f as in the previous proposition, the zeros of its spectrogram are located at $(\tilde{t}_k,\eta^*)$, with $\tilde{t}_k=\frac{k+1/2}{\xi_2-\xi_1}$, $k\in\mathbb{Z}$, meaning they are aligned with the singularities of $\widehat{q_f}$.

Investigating the position of the singularities of $\widehat{q}_f$ with respect to that of local maxima, we show the following:

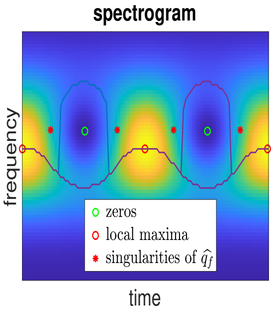
Proposition

Considering the signal f of the previous proposition, the zeros and the local maxima of its spectrogram are aligned only if $A = 1$.

This means that the singularities of $\widehat{q}_f$ are never aligned with local maxima except if $A = 1$.

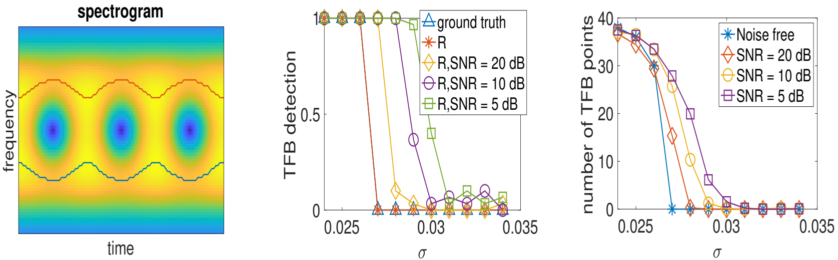

spectrogram
frequency
time
zeros
local maxima
singularities of $\widehat{q}_f$

(a)


spectrogram
frequency
time
zeros
local maxima
singularities of $\widehat{q}_f$

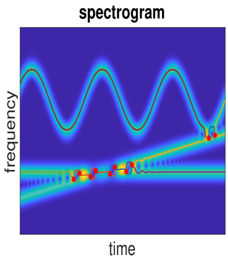
(b)

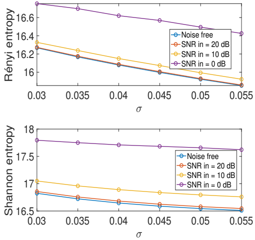
Validation of TFB detector

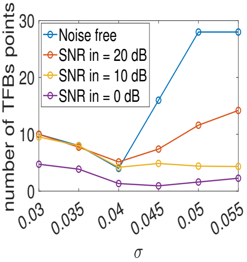


(a) (b) (c)

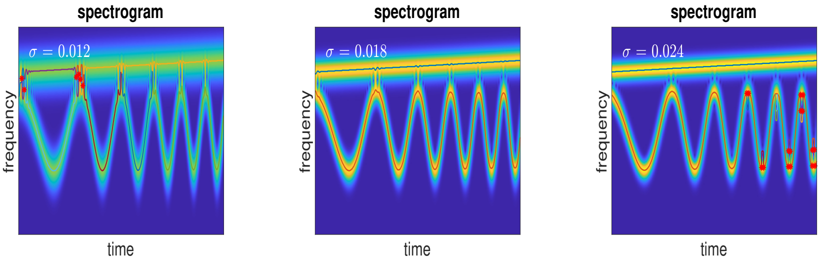
Finding the appropriate window length

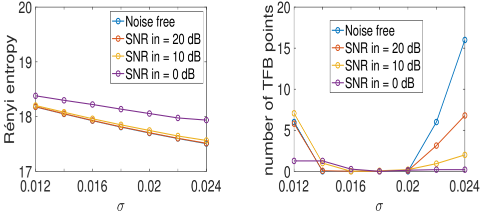




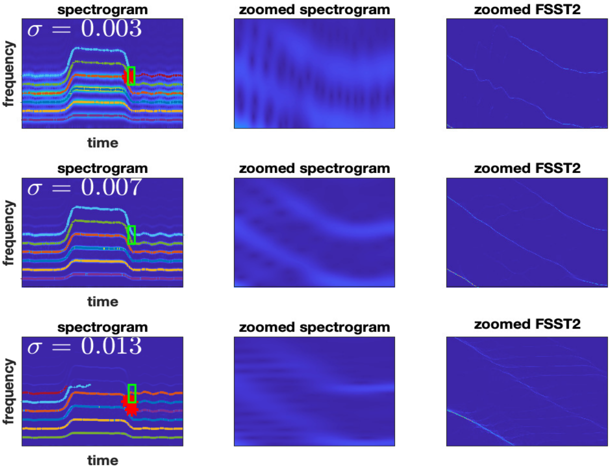


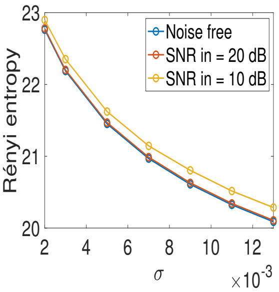
Another Illustration

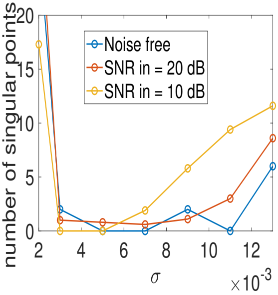




Detection of singular points on a voice signal







[4] S. Meignen and M. Colominas, "A New Ridge Detector Localizing Strong Interference in Multicomponent Signals in the Time-Frequency Plane", IEEE TSP, vol. 71, pp. 3413-3425, 2023.

Thanks for your attention!