Instantaneous Frequency and Amplitude Estimation in Multi-Component Signals Using EM-based Algorithm

Quentin Legros

November 10, 2023

Collaborative work

- Dominique Fourer
- Sylvain Meignen
- Marcelo A. Colominas

Introduction

- Focus on multi-component signals (MCS).

$$x(n) = \sum_{k=1}^{K} x_k(n), \quad \text{with } x_k(n) = \alpha_k(n) e^{j2\pi\phi_k(n)}$$

MCS

- Mixture of K superimposed components.
- $a_k(n)$ and $\phi_k(n)$ the time-varying amplitude and phase of component k.

Objectives

- IF and IA estimation.
- Robust to noise and modes overlapping.
- Acceptable computational time.

Plan

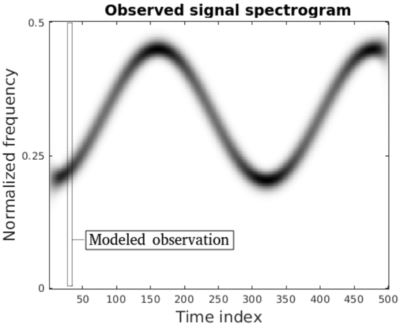
1. Observation model
2. Prior models
3. Estimation strategy
4. Results
5. Conclusion

Plan

Observation model

Motivation

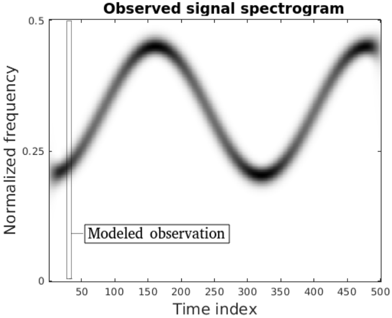
- Spectrogram : squared modulus of the STFT.
- Model vertical spectrogram slices.
- 1D signal observed for a fixed time instant $n \in [0, N-1]$.
- Bayesian framework.



Observation model

Motivation

- Spectrogram : squared modulus of the STFT.
- Model vertical spectrogram slices.
- 1D signal observed for a fixed time instant $n \in [0, N-1]$.



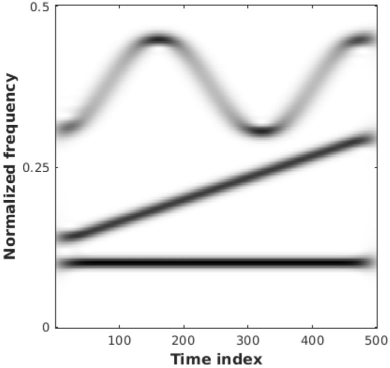
- $g(m) = \frac{2\sqrt{\pi}L}{M} e^{-\left(\frac{2\pi mL}{M}\right)^2}$.
- m : frequency in $[0, M-1]$.
- L : time spread of the analysis window

- Known Gaussian analysis window.
- Expected signal shape at given time indexes.
- Sinusoidal components only.

Observation model

Motivation

- Spectrogram : squared modulus of the STFT.
- Model vertical spectrogram slices.
- 1D signal observed for a fixed time instant $n \in [0, N-1]$.



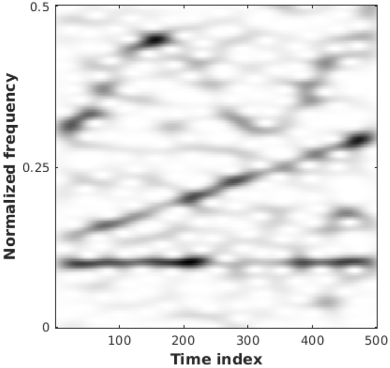
Requirements

- Each ridges has to be modelled.
- Solution : a distribution for each ridge.
- Mixture model.
- Different positions and amplitudes.
- Instantaneous frequency (IF).
- Instantaneous amplitude (IA).

Observation model

Motivation

- Spectrogram : squared modulus of the STFT.
- Model vertical spectrogram slices.
- 1D signal observed for a fixed time instant $n \in [0, N-1]$.



Requirements

- Noise has to be considered in the model.
- Additive noise modeling.

Limits

- Components interference.
- Multiplicative noise.

Observation model

Mixture model

- Mixture of Gaussian distributions plus a uniform term.
- K Gaussian distributions to model the K signal components.
- A single uniform term to model arbitrary distributed noise.

$$p(s_{n,m}|\boldsymbol{w}_n, \boldsymbol{m}_n) = \sum_{k=1}^{K} w_{n,k} g(m - m_{n,k}) + \frac{\left(1 - \sum_{k=1}^{K} w_{n,k}\right)}{M}$$

- Spectrogram columns as $\boldsymbol{s}_n = [s_{n,0}, \ldots, s_{n,M-1}]^\top$.
- Ridge positions $\boldsymbol{m}_n = [m_{n,1}, \ldots, m_{n,K}]^\top$ associated with $\phi'_n = \{\frac{d\phi_k}{dn}(n)\}_{k=1}^{K}$
- Mixture weight $\boldsymbol{w}_n = [w_{n,1}, \ldots, w_{n,K}]^\top$.

Observation model

Mixture weight

- $w_{n,k}$ is the probability to observe the kth component in $\boldsymbol{s}_n$.
- $1 - \sum_{k=1}^{K} w_{n,k}$ is the probability to observe noise in $\boldsymbol{s}_n$.

$$w_{n,k} = \frac{a_{n,k}}{\sum_{k=1}^{K} a_{n,k} + M b_n}$$

- b_n : average noise amplitude at time n.
- $\boldsymbol{w}_n = [w_{n,1}, \ldots, w_{n,K}]^\top$ belongs to $[0,1]^K$ and $\sum_k w_{n,k} \leq 1$.
- Joint likelihood function :

$$p(\boldsymbol{S}|\boldsymbol{W}, \boldsymbol{M}) = \prod_n p(\mathrm{s}_n|\boldsymbol{w}_n, \boldsymbol{m}_n)$$

- with $\boldsymbol{S} = \{\boldsymbol{s}_n\}_{n=0}^{N-1}$, $\boldsymbol{W} = \{\boldsymbol{w}_n\}_{n=0}^{N-1}$ and $\boldsymbol{M} = \{\boldsymbol{m}_n\}_{n=0}^{N-1}$.

Plan

Prior model : Total variation

- Need of prior distributions to complete the Bayesian model.
- A weak uniform prior model is assigned to the mixture weights $\boldsymbol{W}$.
- Models the lack of prior information.
- Need a prior model for $\boldsymbol{M}$.

Prior model : Total variation

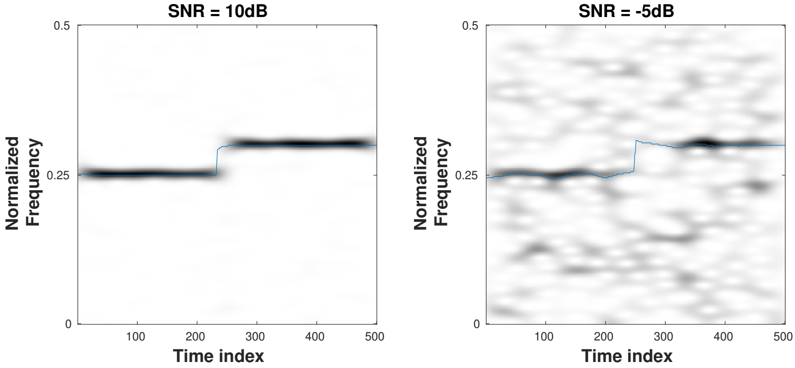
- Ridges can be split or destroyed in the presence of noise.
- $\rightarrow$ Enforcing spatial smoothness between successive IF estimates.
- $\rightarrow$ Constraining the derivative of the estimates
- Markov random field (MRF) Total variation (TV) prior model on $\boldsymbol{M}$ to preserve sharp edges[1].

$$p(\boldsymbol{M}|\epsilon) \propto \exp\left[-\epsilon \sum_{k=1}^{K} \left\|\Delta^{1} \boldsymbol{m}_{.,k}\right\|_{1}\right]$$

- with Δ^1. denoting the first order finite difference.
- $\boldsymbol{m}_{.,k}$ the k-th row of $\boldsymbol{M}$.
- ϵ an arbitrary fixed user-defined hyper-parameter.

[3] L. I. Rudin and S. Osher and E. Fatemi, Nonlinear total variation based noise removal algorithms, 1992.

Results



Observation

- Spectrograms of a signal made of two portions of noisy sinusoidal waves
- Enables the estimates to make sharp transitions when the ridges are split

Prior model : Laplacian

Motivation

- Other choice: constraining the mean curvature of estimated ridges.
- Bound IFs second derivatives.
- MRF Laplacian prior model on $\boldsymbol{M}$ to ensure smooth estimation[2].

$$p(\boldsymbol{M}|\lambda) \propto \exp\left[-\frac{\lambda}{2}\sum_{k=1}^{K}||\Delta^2 \boldsymbol{m}_{.,k}||_2^2\right]$$

- with Δ^2 denoting the second order finite difference.
- L^2-norm penalization
- λ an arbitrary fixed user-defined hyper-parameter.

[4] X. Wang, Laplacian Operator-Based Edge Detectors, 2007.

Plan

Estimation strategy

Posterior distribution

- Assumed independence between $\boldsymbol{W}$ and $\boldsymbol{M}$.
- Clarity: hyper-parameters are omitted in the sequel.
- Bayes rule to compute the joint posterior distribution.

$$p(\boldsymbol{W}, \boldsymbol{M} | \boldsymbol{S}) \propto p(\boldsymbol{S} | \boldsymbol{W}, \boldsymbol{M}) p(\boldsymbol{M}) p(\boldsymbol{W})$$

Limitations

- Challenging joint estimation of $(\boldsymbol{W}, \boldsymbol{M})$.
- Multimodal likelihood w.r.t. $\boldsymbol{M}$.
- Presence of multiple ridges.

Estimation strategy

EM algorithm

- Challenging joint estimation of $(\boldsymbol{W}, \boldsymbol{M})$.
- Due to the unobserved variable $\boldsymbol{M}$.
- EM algorithms[3] are particularly adapted to address this problem.
- The shape of the observation model is well suited for such methods.
- $\rightarrow$ Marginalizing over the hidden parameter $\boldsymbol{M}$.

Marginal maximum a posteriori estimation

$$\hat{\boldsymbol{W}}_{\text{MMAP}} = \underset{\boldsymbol{W}}{\operatorname{argmax}} \sum_{\boldsymbol{M}} p(\boldsymbol{W}, \boldsymbol{M} | \boldsymbol{S}) = \underset{\boldsymbol{W}}{\operatorname{argmax}} \; p(\boldsymbol{W} | \boldsymbol{S})$$

[5] G. McLachlan and T. Krishnan, The EM algorithm and extensions, 2007.

Estimation strategy

EM algorithm

- Unobserved variables → compute iteratively MMAP estimates.
- At each iteration, two main steps are performed.
- At each iteration of the algorithm, two main steps are performed.
- Given $\boldsymbol{W}^{(i)}$ the current estimation of $\boldsymbol{W}$ at iteration i.

EM-steps

$$Q(\boldsymbol{W}|\boldsymbol{W}^{(i)}) = E_{\boldsymbol{M}|\boldsymbol{W}^{(i)},\boldsymbol{S}}\left[\log(p(\boldsymbol{W},\boldsymbol{M}|\boldsymbol{S}))\right]$$

$$\boldsymbol{W}^{(i+1)} = \underset{\boldsymbol{W}}{\mathrm{argmax}}\ Q(\boldsymbol{W}|\boldsymbol{W}^{(i)}).$$

Estimation strategy

Stochastic approach

- Intractable expectation due to the MRF prior models.
- Solution : Stochastic EM algorithm.
- Multimodal conditional distribution $p(\boldsymbol{M}|\boldsymbol{W}^{(i)}, \boldsymbol{S})$ and overlapping components would require long simulation steps.
- Instead we adopt a strategy similar to mean field-like approximations[6]

[6] G. Celeux and F. Forbes and N. Peyrard, EM procedures using mean field-like approximations for Markov model-based image segmentation, 2003.

Estimation strategy

Approximating distribution

- The idea is to replace $p(\boldsymbol{M}|\boldsymbol{W}^{(i)}, \boldsymbol{S})$ to make the expectation tractable.

$$Q(\boldsymbol{W}|\boldsymbol{W}^{(i)}) = \sum_{n,k} \sum_{m=0}^{M-1} p(m_{n,k} = m|\boldsymbol{s}_n, w_n^{(i)}) \log(p(\boldsymbol{s}_n, \hat{m}_{n,k} = m|\boldsymbol{w}_n))$$

- First, we compute an approximation $\tilde{p}(\boldsymbol{M}|\bar{\boldsymbol{M}})$ of $p(\boldsymbol{M})$.
- Auxiliary random sample : $\bar{\boldsymbol{M}}$
- Using a hot-started 2-step Gibbs sampler.
- Bayes rule $\tilde{p}(\boldsymbol{M}|\boldsymbol{W}^{(i)}, \boldsymbol{S}) = \frac{p(\boldsymbol{S}|\boldsymbol{W}^{(i)}, \boldsymbol{M})\tilde{p}(\boldsymbol{M}|\bar{\boldsymbol{M}})}{p(\boldsymbol{S})}$.

Estimation strategy

Sequential MMAP

- Current estimation of $\boldsymbol{M}$ from $\tilde{p}(\boldsymbol{M}|\boldsymbol{W}^{(i)}, \boldsymbol{S})$ through sequential MMAP estimation.

Simplified M-step

- Current estimate $\boldsymbol{M}$, denoted by $\boldsymbol{M}^{(i+1)}$, is used at each iteration to simplify the M-step.

$$\boldsymbol{W}^{(i+1)} = \underset{\boldsymbol{W}}{\operatorname{argmax}} \; Q(\boldsymbol{W}|\boldsymbol{W}^{(i)}) \approx \underset{\boldsymbol{W}}{\operatorname{argmax}} \; \log \left(p(\boldsymbol{S}|\boldsymbol{W}, \boldsymbol{M}^{(i+1)}) \right) + C,$$

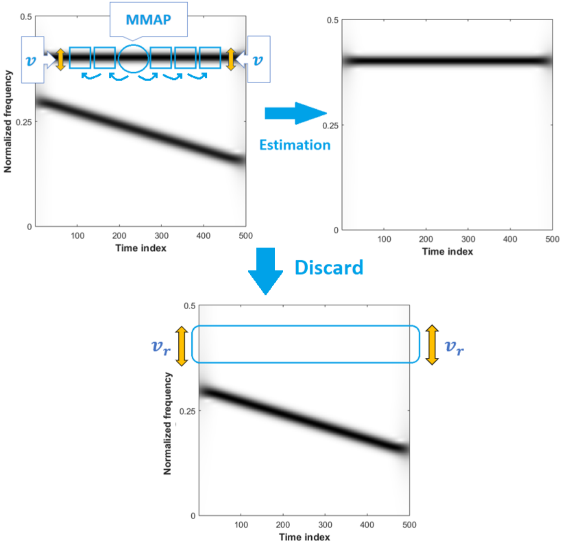
- where $C \sim \log(p(\boldsymbol{M})p(\boldsymbol{W}))$ is a constant.
- The final estimate is denoted as $\hat{\boldsymbol{W}}$

Estimation strategy

Sequential MMAP estimation

- At each EM iteration, $\boldsymbol{M}$ is estimated from $\tilde{p}(\boldsymbol{M}|\boldsymbol{W}^{(i)}, \boldsymbol{S})$.
- But it is multimodal.
- Sequential approach selecting most likely spaced estimates.
  - First Estimate through MAP estimation from $\tilde{p}(\boldsymbol{M}|\boldsymbol{W}^{(i)}, \boldsymbol{S})$.
  - Propagate the estimation in a neighborhood of size v.
  - Once a ridge is estimated : discard
  - Set $\tilde{p}(\boldsymbol{M}|\boldsymbol{W}^{(i)}, \boldsymbol{S})$ in a neighborhood of size v_r around the ridge to zeros.
  - Iterate K times.

Estimation strategy



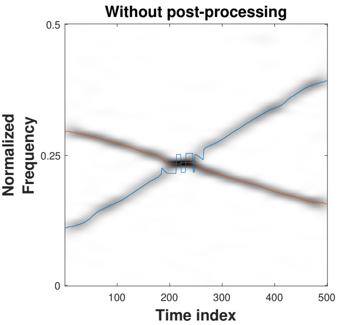
Estimation strategy

- Our method depends on two parameters
  - The frequency radius v centered around the previous estimate.
  - The radius v_r of discarded frequencies.
- Need assumptions to automatize ridge estimation.
- v_r is set to three standard deviations (three-sigma rule of thumb) of the data distribution.
- More difficult for v : balance between two conflicting aspects:
  - Successive IF estimates have to be close to each other (small v)
  - v needs to be greater than v_r to jump over discarded values.
- We set $v = 4v_r$, which enables jumps over removed regions of overlap.

Estimation strategy

Limitation

- Discard : creates gap on overlapping regions.
- Need to refine the ridges.



Post-processing

- Detection of overlapping regions (close IF).
- Polynomial interpolation of remaining pieces.

Estimation strategy

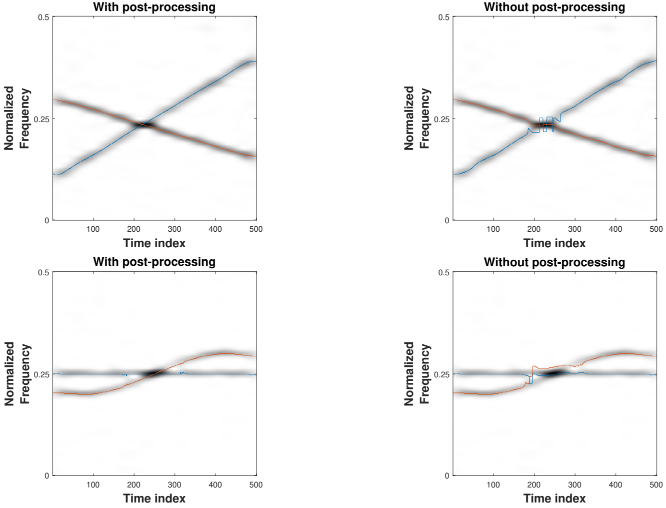


Figure: Spectrograms of a MCS merged with additive white Gaussian noise (SNR = 10 dB). Left: with post-processing. Right : without post-processing.

Estimation strategy

Amplitude estimation

- Amplitude estimation after the estimation of $\boldsymbol{W}$ and $\boldsymbol{M}$.
- The relation between $w_{n,k}$, the signal amplitude $\alpha_k(n)$ and the analysis window θ can be expressed as

$$w_{n,k} = \frac{\alpha_k^2(n)\|F_\theta\|_2^2}{\sum_{k=1}^{K} \alpha_k^2(n)\|F_\theta\|_2^2 + Mb_n},$$

- with $F_\theta = \frac{M}{2L\sqrt{\pi}}$ the Fourier transform of the analysis window θ.
- The amplitude can not be directly estimated.

Solution

- We propose to approximate the denominator by setting $\bar{\boldsymbol{s}}_n = \sum_{m=0}^{M-1} s_{n,m}$ as

$$E_{\boldsymbol{s}_n|\boldsymbol{a}_n,\boldsymbol{m}_n,b_n}[\bar{\boldsymbol{s}}_n] = \sum_{k=1}^{K} \alpha_k^2(n)\|F_\theta\|_2^2 + Mb_n$$

Estimation strategy

Amplitude estimation

- We propose to approximate the denominator by setting $\bar{\boldsymbol{s}}_n = \sum_{m=0}^{M-1} s_{n,m}$ as

$$E_{\boldsymbol{s}_n|\boldsymbol{a}_n,\boldsymbol{m}_n,b_n}\left[\bar{\boldsymbol{s}}_n\right] = \sum_{k=1}^{K} \alpha_k^2(n)\|F_\theta\|_2^2 + Mb_n$$

- Allows to approximate the denominator by the expected values of $\bar{\boldsymbol{s}}_n$.
- It directly follows that $\alpha_k^2(n) = \frac{w_{n,k}E[\bar{\boldsymbol{s}}_n]}{\|F_\theta\|_2^2}$.
- This assumes $\bar{\boldsymbol{s}}_n$ to be a good approximation of $E[\bar{\boldsymbol{s}}_n]$.

Amplitude estimation

- The final amplitude estimate thus reads:

$$\hat{\alpha}_k(n) = \sqrt{\frac{\hat{w}_{n,k}\bar{\boldsymbol{s}}_n}{\|F_\theta\|_2^2}}.$$

Plan

1. Observation model
2. Prior models
3. Estimation strategy
4. Results
5. Conclusion

Results

Experiments

- IF estimation performance.
- Comparison with different approaches [7,8,9].
- Proposed approach using either TV or Laplacian prior models.
- White Gaussian noise with various Signal-to-noise ratio (SNR).
- Simple cases first.

[7] Q. Legros and D. Fourer, A novel pseudo-Bayesian approach for robust multi-ridge detection and mode retrieval, 2021.

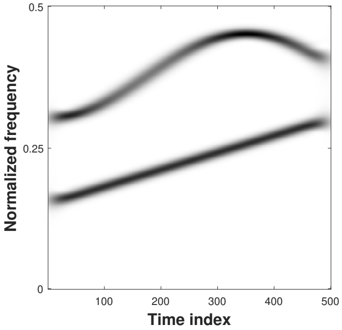
[8] E. Brevdo and N. S. Fuckar and G. Thakur and H-T. Wu, The Synchrosqueezing algorithm: a robust analysis tool for signals with time-varying spectrum, 2011.

[9] N. Laurent and S. Meignen, A Novel Ridge Detector for Nonstationary Multicomponent Signals: Development and Application to Robust Mode Retrieval, 2021.

Results

Experiments

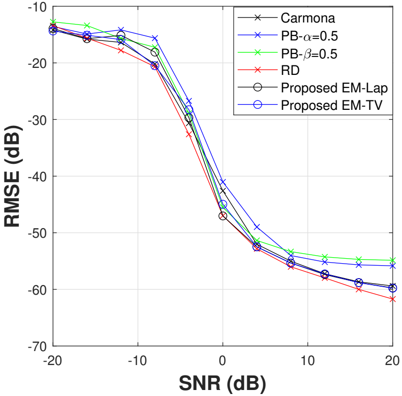
- Relative mean squared error: $\text{RMSE} = \frac{1}{NM^2} \sum_{k=1}^{K} \sum_{n=0}^{N-1} (\bar{m}_{n,k} - \hat{m}_{n,k})^2$.
- $\bar{m}_{n,k}$ (resp. $\hat{m}_{n,k}$) is the actual (resp. estimated) normalized IF of the k-th component at the n-th time instant.



Reconstruction

- MCS with two components
- Prior hyperparameters: $\epsilon = 10^{-2}$ and $\lambda = 10^{-1}$.

Results



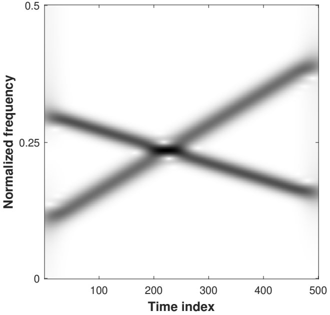
Observations

- Similar behavior.
- RD slightly better at high SNR.

Results

Experiments

- Similar experiment but with overlapping components.
- The RD method is no longer considered (not adapted).
- Use instead 3DRD method[10]

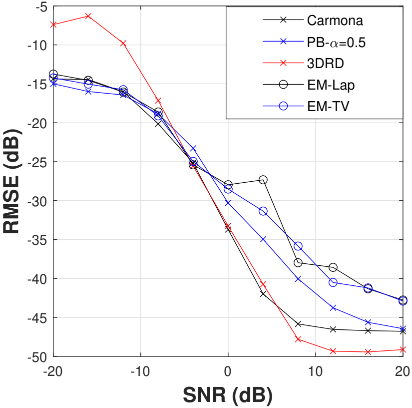


Reconstruction

- 3DRD : 3D separation and estimation.
- Carmona method jump parameters $\rightarrow$ set to v and v_r.

[10] X. Zhu and H. Yang and Z. Zhang and J. Gao and N. Liu, Frequency-chirprate reassignment, 2020.

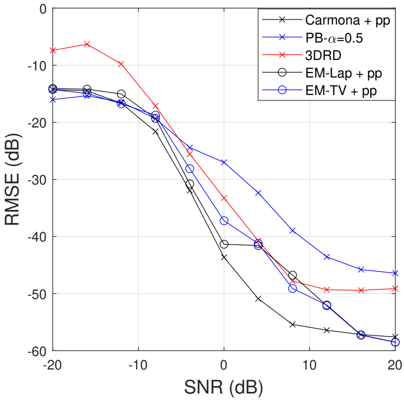
Results



Observations

- Without post-processing.
- Poor performances.
- Oscillations in overlapping region.

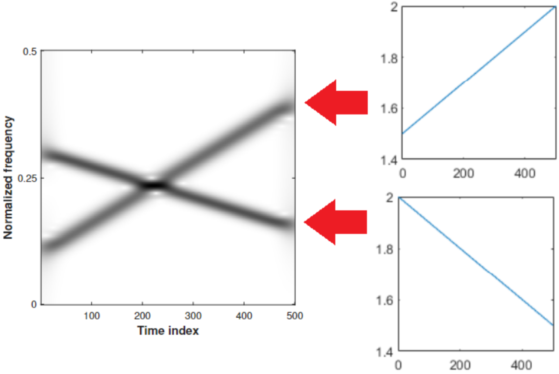
Results



Observations

- With post-processing.
- EM and Carmona → best performances.

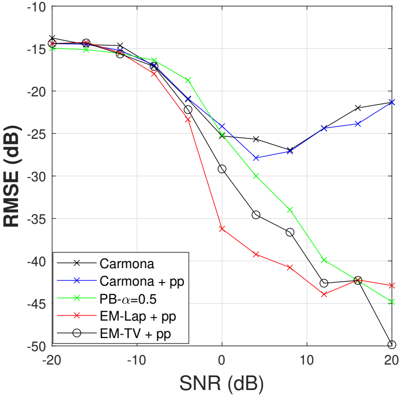
Results



Observations

- Amplitude modulated components.
- Decreasing (resp. increasing) amplitude assigned to decreasing (resp. increasing) chirp.

Results



Experiments

- EM is the most efficient method.
- Carmona limited (does not insure smoothness of component amplitude)
- With and without post-processing.

Results

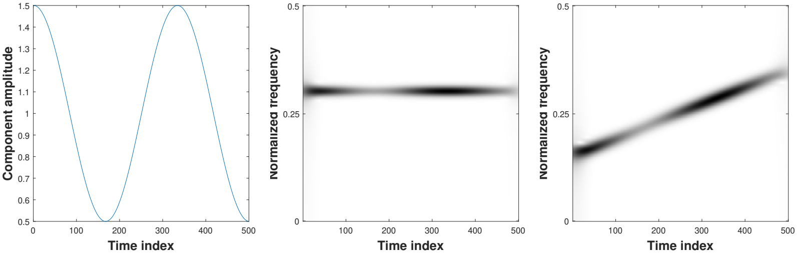
Amplitude

- Relative mean absolute error $\text{RMAE} = \frac{1}{NK} \sum_{k=0}^{K-1} \sum_{n=0}^{N-1} |\bar{\alpha}_k(n) - \hat{\alpha}_k(n)|$.
- $\bar{\alpha}_k(n)$ (resp. $\hat{\alpha}_k(n)$) is the actual (resp. estimated) amplitude.
- Comparison: PB[11] and deterministic approach (Local)[12].
- Oracle equivalent for each method (IF known).
- Laplacian prior only.

[11] Q. Legros, D. Fourer, Pseudo-Bayesian Approach for Robust Mode Detection and Extraction Based on the STFT, 2022.

[12] D. Fourer and F. Auger and G. Peeters, Local AM/FM parameters estimation: application to sinusoidal modeling and blind audio source separation, 2018.

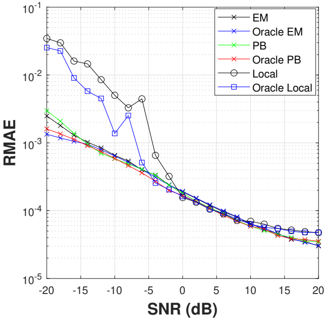
Results

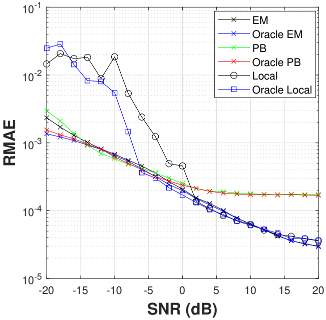


Observations

- Pure tone and linear chirp signals.
- Sinusoidal amplitude.
- slight frequency modulation.

Results

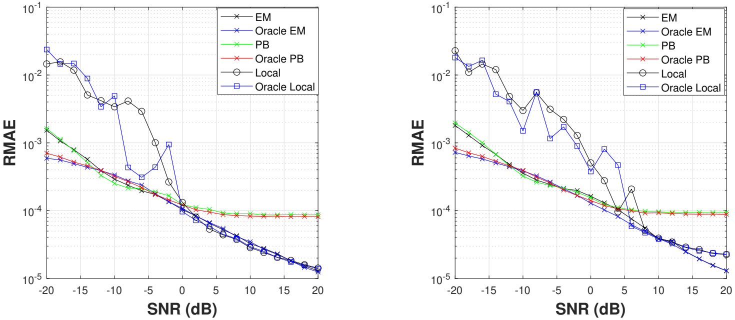




Observations

- Similar results at high SNR.
- Local approach limited for negative SNR.
- EM and PB similar to their Oracles → low impact of IF performances.

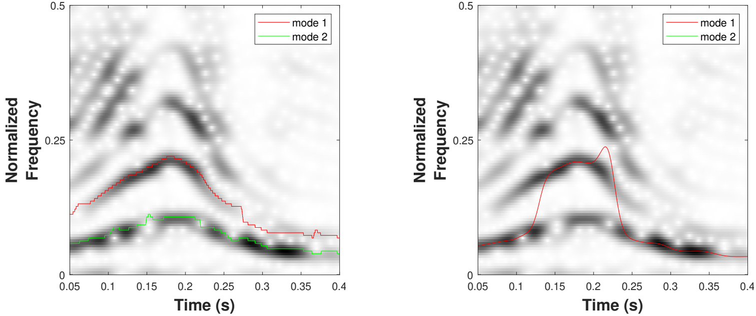
Results



Observations

- Applied now on previous MCS.
- One linear amplitude (from 2 to 0.5) and one Sinusoidal amplitude.
- Best performance obtained with EM → PB limited at high SNR.

Results

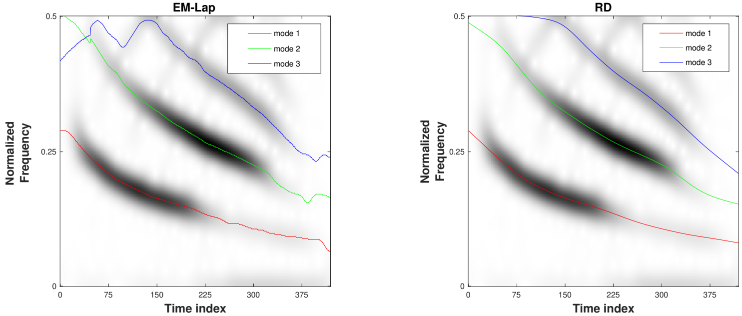


Observation

- Speech signal of Japanese native speakers[13]
- Comparison with RD algorithm.
- Many gaps, discontinuities and spurious local maxima.

[13] D. Fourer and T. Shochi and J-L. Rouas and A. Rilliard, Perception and manipulation of Japanese attitudes, 2016.

Results



Observation

- Bat signal
- EM exhibit oscillations - RD provides smooth estimates centered around the IF.
- RD : troubles to estimate mode 3 perfectly.

Plan

Conclusions and perspectives

Conclusions

- A new model for IF and IA estimation of MCS modes in the presence of noise.
- Adapted to component overlap and low SNR scenarios.
- Stochastic variant of the EM algorithm (low computational cost).

Conclusions and perspectives

Future works

- Consider the modulation rate[14].
- Generalizing the estimation process (amplitude, hyperparameters).
- Extension to generalized EM algorithm[15].

[14] M. Colominas and S. Meignen and D. H. Pham, Time-Frequency Filtering Based on Model Fitting in the Time-Frequency Plane, 2019.
[15] G. McLachlan and T. Krishnan, The EM algorithm and extensions, 2007.

Thanks for your attention !

Contact me
quentin.legros@univ-orleans.fr

Paper
https://arxiv.org/pdf/2203.16334.pdf

Code
https://github.com/QuentinLEGROS/TSP2023