A Novel Pseudo-Bayesian Approach for Robust Multi-Ridge Detection and Mode Retrieval

Quentin Legros[1,2]
[1]*LTCI*
Télécom Paris
Palaiseau, FRANCE
quentin.legros@telecom-paris.fr

Dominique Fourer[2]
[2]*Laboratoire IBISC*
Université d'Évry-Val-d'Essonne
Évry, FRANCE
dominique.fourer@univ-evry.fr

Abstract—This paper introduces a novel approach for extracting the elementary components present in an observed non-stationary mixture signal. Our technique based on a pseudo-Bayesian approach operates in the time-frequency plane and sequentially estimates the ridge of each component that is required for mode extraction. We compare our results with those obtained with the state-of-the-art Brevdo method which has shown its efficiency for disentangling multicomponent noisy signals. Our results reveal an improvement of the reconstruction performance when compared to the state of the art.

Index Terms—ridge extraction, time-frequency, random walk, variational methods, robust divergence.

I. Introduction

Disentangling natural signals in audio, biomedicine, seismic or radar applications is a challenging problem arising in a large variety of real-world scenarios [1]. Such signals can be modeled as a mixture of superimposed amplitude- and frequency-modulated (AM/FM) elementary components (the modes), which convey relevant information about the signal content. Following this goal, time-frequency (resp. time-scale) analysis [2] offers an elegant framework where MultiComponent Signals (MCS) are projected into a 2-dimensional plane allowing to efficiently observe the instantaneous frequency trajectory (the ridge) of each mode. These ridges are full of interest for signal analysis since they enable applications such as signal enhancement, denoising, information extraction, and source separation [3], [4]. Several techniques previously proposed in the literature for ridge detection and mode retrieval follow the idea that ridges correspond to interpolated curves of detected local maxima along the frequency axis [5], [6]. It is shown that the ridges approximate the instantaneous frequency (IF) of components for which the estimation quality depends on the noise level and the presence of interfering components [7], [8]. Such approach was investigated for example in audio applications based on the tracking of partials [9] which aim at predicting the trajectory of harmonic signal components made of a fundamental frequency (F_0) and its integer multiples. In [10], Brevdo *et al.* proposed a mode extraction algorithm based

on a total variation optimization approach combined with the synchrosqueezing transform [11] allowing for mode retrieval using a sharpened and reversible time-frequency representation. On the other hand, Flandrin introduced the Delaunay triangulation method [12] which uses the zeros of the spectrogram for extracting signal components in the time-frequency plane. Despite promising results, all these approaches suffer from several limitations for modes reconstruction, in particular in low Signal-to-Noise Ratio (SNR) configurations due to a lack of robustness to noise.

In this work, we propose a novel approach for extracting the ridge associated with the IF of a signal from its time-frequency representation (TFR). Similarly to [13], we consider a simple observation model and we introduce a robust approach in order to circumvent the limitations occurring through its inaccuracy. More precisely, a variety of robust divergence [14]–[16] is used instead of the classical Kullback Leibler divergence (KLD) to account for model mismatch in the presence of noise or of several frequency components. The estimated IF is assigned to a temporal prior model accounting for the expected small changes between successive time instants, providing a regularized inference process in a pseudo-Bayesian (PB) framework. Finally, we present our algorithm where a sequential estimation strategy is adopted to extract several ridges independently. This paper is organized as follows. In Section II, we present the fundamentals of time-frequency analysis using the Short-Time Fourier Transform (STFT) for ridge estimation and mode retrieval before introducing our observation model. Section III presents the different similarity measures used in the estimation process and our new proposed estimation strategy is described in Section IV. The performance of the proposed method is comparatively assessed in Section V with numerical results. Conclusions and future work are finally reported in Section VI.

II. Time-frequency Analysis Rationale

A. Short-Time Fourier Transform

The STFT of a signal x, using a differentiable analysis window h, can be defined at each time instant t and each

angular frequency ω, as:

$$F_x^h(t,\omega) = \int_{\mathbb{R}} x(u)h(t-u)^* \, \mathbf{e}^{-j\omega u} \, \mathrm{d}u, \tag{1}$$

with $j^2 = -1$ and z^* the complex conjugate of z. The spectrogram provides a TFR computed by $|F_x^h(t,\omega)|^2$. Plancherel theorem allows to rewrite Eq. (1) in the frequency domain for a better understanding of the signal content, as:

$$F_x^h(t,\omega) = \int_{\mathbb{R}} F_x(\Omega + \omega)F_h(\Omega)^* \, \mathbf{e}^{j\Omega t} \, \frac{\mathrm{d}\Omega}{2\pi}, \tag{2}$$

with $F_x(\omega) = \int_{\mathbb{R}} x(t) \, \mathbf{e}^{-j\omega t} dt$ (resp. F_h) the Fourier transform of signal x (resp. h). When $h(t_0) \neq 0$, the STFT admits a synthesis formula allowing to recover the signal x with a delay $t_0 \geq 0$, as:

$$x(t-t_0) = \frac{1}{h(t_0)} \int_{-\infty}^{+\infty} F_x^h(t,\omega) \, \mathbf{e}^{j\omega(t-t_0)} \, \frac{\mathrm{d}\omega}{2\pi}, \tag{3}$$

where the integration region can be reduced to the frequency support of the signal x. In the remainder, we assume a Gaussian analysis window $h(t) = \frac{1}{\sqrt{2\pi}T} \, \mathbf{e}^{-\frac{t^2}{2T^2}}$, with a Fourier transform which can be expressed as $F_h(\omega) = \mathbf{e}^{-\frac{\omega^2 T^2}{2}}$.

B. Ridge Estimation and Component Reconstruction

We assume that the signal x is a mixture made of K superimposed AM-FM components expressed as:

$$x(t) = \sum_{k=1}^{K} x_k(t) \quad, \text{ with } x_k(t) = a_k(t) \, \mathbf{e}^{j\phi_k(t)}, \tag{4}$$

where $a_k(t)$ and $\phi_k(t)$ are respectively the time-varying amplitude and frequency of the k-th component. Each component (or mode) can be characterized by a ridge located at its IF that is defined as the derivative of the phase w.r.t. time $\frac{d\phi_k}{dt}(t)$. The STFT of each component k can be approximated [17] by:

$$F_k^h(t,\omega) \approx x_k(t)F_h(\omega - \frac{d\phi_k}{dt}(t)), \tag{5}$$

and its IF (*i.e.* ridge) can be estimated using the local maxima of $|F_x^h(t,\omega)|$. When the components are separable ($|\frac{d\phi_i}{dt}(t) - \frac{d\phi_k}{dt}(t)| \geq \epsilon, \forall i \neq k, \epsilon > 0$), they can be disentangled and reconstructed separately using Eq. (3) with an integration region limited to the vicinity of the ridge such as [17]:

$$x_k(t) \approx \frac{1}{h(0)} \int_{|\omega - \frac{d\phi_k}{dt}(t)| < \epsilon} F_x^h(t,\omega) \, \mathbf{e}^{j\omega t} \, \frac{\mathrm{d}\omega}{2\pi} \tag{6}$$

where ϵ is an arbitrary small threshold.

C. Observation model

Given a MultiComponent Signal (MCS) (*cf.* Eq. (4)), we aim at extracting separately the superimposed modes x_k which fully describe x. However, only a noisy version y of x is known in practice such that $y = x + \epsilon$, with ϵ standing for an additive noise (assumed white Gaussian). Thus, we consider as observation matrix $\boldsymbol{S}$ the $M \times N$ spectrogram

$S_y^h = |F_y^h|^2$ with $F_y^h[n,m] \approx F_y^h(nT_s, 2\pi \frac{m}{MT_s})$ the discrete-time STFT computed at time index $n \in \{0,1,...,N-1\}$ and frequency bin $m \in \{0,1,...,M-1\}$, T_s being the sampling period and M (resp. N) being the maximal number of frequency bins (resp. time indices). $L = \frac{T}{T_s}$ is the time spread of the Gaussian analysis window h used in this study. We state $\boldsymbol{s}_n = [\boldsymbol{S}]_{n,:} = [s_{n,0}, \ldots, s_{n,M-1}]^\top$ is of size $M \times 1$ where $s_{n,m}$ corresponds to the spectrogram value $S_y^h[n,m]$. Assuming for now that y is fully described by only one component, and neglecting any other additional sources apart from the signal, the observation model can be expressed for any $z \in [0, M-1]$ as

$$p(z|\bar{m}_n) = g(z - \bar{m}_n), \tag{7}$$

where $g(m) = \frac{2\sqrt{\pi}L}{M} \, \mathbf{e}^{-\left(\frac{2\pi mL}{M}\right)^2}$ is the normalized and discretized squared modulus of the Fourier transform of h, and $\bar{m}_n$ is the ridge position at time n. Given $\bar{m}_n$, the elements in $\boldsymbol{s}_n$ are assumed independent and the joint likelihood can be formulated as

$$p(\boldsymbol{s}_n|\bar{m}_n) = \prod_{m=0}^{M-1} p(s_{n,m}|\bar{m}_n). \tag{8}$$

III. ALTERNATIVE DIVERGENCES

Whilst the proposed model is extremely simple, it becomes limited when either more than one ridge is present in the TFR or when external noise sources corrupt the signal y. For that reason, inferring estimates from the observations through maximum likelihood estimation (MLE) for instance can provide poor results. However, this choice is motivated by the low cost of the estimation process. Nonetheless, not only the similarity between the observation model and the data distribution have a positive impact on the estimation performance, but also the cost function to be optimized [14], [16], [18]. For instance, performing MLE is equivalent to minimize the KLD between the empirical data distribution $\hat{p}_n(z)$, and the postulated model in (7) such that

$$D_{\mathrm{KL}}(\hat{p}_n(z)||p(z|\bar{m}_n)) = \int \hat{p}_n(z) \log\left(\frac{\hat{p}_n(z)}{p(z|\bar{m}_n)}\right) \mathrm{d}z. \tag{9}$$

Note that it is possible to reach acceptable performance in the case of model mismatch by modifying the cost function to be minimized. This is equivalent to replace the similarity measure associated with the optimization problem. In particular, using alternative divergences instead of the classical KLD can improve the estimation estimation performance in cases involving discrepancies between the data and the observation model [14], [16], [19]. Two divergences are considered in this paper, namely the β-divergence (β-D): $\quad D_\beta(\hat{p}_n(z)||p(z|\bar{m}_n)) =$

$$\int -\frac{1+\beta}{\beta}\hat{p}_n(z)p(z|\bar{m}_n)^\beta + p(z|\bar{m}_n)^{1+\beta}$$
$$+ \frac{1}{\beta}\hat{p}_n(z)^{1+\beta}\mathrm{d}z, \quad \beta > 0, \tag{10}$$

and the Rényi divergence (RD): $D_\alpha(\hat{p}_n(z)||p(z|\bar{m}_n)) =$

$$\frac{1}{\alpha-1}\log\left(\int \hat{p}_n(z)^\alpha p(z|\bar{m}_n)^{1-\alpha}dz\right), \quad \alpha > 0, \alpha \neq 1. \tag{11}$$

A special case to be mentioned is when minimizing the β-D (resp. RD) with $\beta \to 0$ (resp. $\alpha \to 1$), the resulting estimator reduces to MLE. While minimizing the β-D can produce robust estimators depending on the choice of β [19], the RD allows to control, through α, the mass covering of the observation [19]. Indeed, the logarithm of the quotient in the KLD penalizes the region where $\hat{p}_n(z)$ is under-estimated by the model [19]. The parameter α in Eq. (11) controls the weight given to each observation during the minimization process. Selecting $\alpha \in]0, 1[$ remains to concentrate on regions where $\frac{\hat{p}_n(z)}{p(z|\bar{m}_n)}$ is large, while considering $\alpha > 1$ will favor the regions where the true distribution is underestimated, highlighting broader approximated distribution.

IV. ESTIMATION STRATEGY

We propose to extract the ridge by sequentially estimating its position $\bar{m}_n$ for each time instant n. Thus, a pseudo-Bayesian framework is adopted to infer estimates from the observations, which first requires the computation of the pseudo-posterior distribution. The alternative divergences presented in Section III are used to derive pseudo-likelihood terms, and we assume the $\bar{m}_n$ are assigned a prior distribution $p(\bar{m}_n)$ whose choice is discussed in Section IV-B.

A. Pseudo-posterior distribution

When the KLD is considered as the similarity measure, the posterior distribution can be obtained in practice by maximizing the evidence lower-bound (ELBO) [20]. The resulting distribution writes

$$p(\bar{m}_n|\boldsymbol{s}_n) \propto e^{-MCE_{KL}(\bar{m}_n)}p(\bar{m}_n), \tag{12}$$

where $CE_{KL}(\bar{m}_n) = -\frac{1}{M}\sum_m log\left(p(s_{n,m}|m_n)\right)$ is the KLD cross entropy (CE) term between the postulated model and the empirical data distribution. In a similar fashion to [15], [16], alternative pseudo-posterior distributions are built using the CE derived from Eq. (10) and Eq. (11). The CE derived from the β-D writes, for a fixed time n, $CE_\beta(\bar{m}_n) =$

$$-\frac{1+\beta}{\beta}\sum_m p(s_{n,m}|\bar{m}_n)^\beta + \int p(z|\bar{m}_n)^{1+\beta}dz, \tag{13}$$

and that of the Rényi, $CE_\alpha(\bar{m}_n) =$

$$\frac{1}{\alpha-1}\log\left(\sum_m s_{n,m}{}^\alpha p(s_{n,m}|\bar{m}_n)^{1-\alpha}\right). \tag{14}$$

B. Online approach

We consider now a set of N sequential temporal signals $\boldsymbol{s}_n$. Once the $p(\bar{m}_n|\boldsymbol{s}_n)$ has been computed, it becomes possible to infer an estimate $\hat{m}_n$ of the current ridge position at time n. We perform this estimation through minimum mean squared error (MMSE), and propagate this information along the

temporal direction using a prior model $p(\bar{m}_{n+1})$ to regularize the estimation at time $n+1$. Since weak informative prior models are preferable over wrong ones, the following Gaussian Random Walk (GRW) prior model is selected to alleviate issues related to the propagation of spurious estimations

$$p(\bar{m}_{n+1}) \propto \mathcal{N}(\hat{m}_n, \sigma_{rw}^2)\hat{\mathcal{N}}_n, \tag{15}$$

where $\mathcal{N}(\hat{m}_n, \sigma_{rw}^2)$ is a Gaussian random walk modelling the expected evolving of the IF, and $\hat{\mathcal{N}}_n$ is a Gaussian approximation of $p(\bar{m}_n|\boldsymbol{s}_n)$ obtained by minimizing the KLD, remaining to match the first and second moments. Following the assumed density filtering principle, this model aims to propagate the available information at each time using the previous estimate. The variance σ_{rw}^2 determines how far $\hat{m}_{n+1}$ is allowed to be from $\hat{m}_n$ using the three-sigma rule of thumb, forcing the ridge to present limited variations. Since the initialization coupled with the slow evolving constraint imposed by the GRW limit the performance for the first iterations, a backward correction step is achieved after the estimation of the whole ridge.

C. Sequential multi-ridge estimation

Here, we extend the estimation process to MCS by accounting for the presence of multiple ridges by applying the proposed strategy sequentially. Indeed, an iterative process based on the number of components K (assumed to be known) is proposed. Note that K can be estimated if unknown (using [21] for instance). Since there is no reason that repeating the same process allows for the extraction of distinct modes, the TFR is updated after each ridge detection by removing the energy located at the corresponding IF. More precisely, a temporary spectrogram $\boldsymbol{R}_k$, containing only the estimated ridge, is constructed and subtracted to the data $\boldsymbol{S}_k$ to generate an updated spectrogram $\boldsymbol{S}_{k+1}$. $\boldsymbol{R}$ can be constructed straightforwardly using a sequence $g(\hat{m}_n), n \in [0, M-1]$. The proposed method is then repeated on the novel TFR until K modes have been extracted. The overall procedure is detailed in Algorithm. 1.

ALGORITHM 1

1: Input: TFR $\boldsymbol{S}_0$, GRW mean $\hat{m}_0$ and variance σ_0, K, g.
2: **for** Component $k = 1$ to K **do**
3:     **for** Time bin $n = 0$ to $N-1$ **do**
4:         Compute the prior model $p(\bar{m}_n)$ by matching moments.
5:         Compute the pseudo-posterior $p(m_n|\boldsymbol{s}_n)$ in Eq. (12).
6:         Estimate $\hat{m}_n$ by MMSE.
7:     **end for**
8:     Repeat steps 4 to 6 iterating from $n = N-1$ to 0
9:     Compute $\boldsymbol{R}_k$.
10:     Update the TFR $\boldsymbol{S}_{k+1} = \max\left[(\boldsymbol{S}_k - \boldsymbol{R}_k), 0\right]$.
11: **end for**

D. Divergence parameters and motivation

Although the postulated observation model in Eq. (7) is incomplete when multiple ridges are observed, the proposed strategy circumvents this limitation as long as they do not overlap. While estimating the divergence hyperparameters is out of the scope of this paper, we remain interested by the

property of the resulting estimators. The lack of accuracy of the model in the presence of outliers motivates the use of the β-D, reducing model mismatch limitations since a broader distributions will be used to infer estimates. Note that the proposed strategy allows for the presence of any additive noise as long as its distribution does not concentrate around specific regions and if the divergence hyperparameters are chosen accordingly. Conversely, the RD provides control over the mass covering of the observations, and is motivated by the presence of close ridges. In this situation, the remaining energy corresponding to the tails of distinct ridges will accumulate and skew them. Selecting $\alpha < 1$ allows to produce thinner distributions, as the divergence will seek the mode of the data distribution. This property would be profitable in the case of close ridge implying for instance the presence of remaining energy after the ridge removal.

V. NUMERICAL RESULTS

In this section, we discuss the ridge detection and mode reconstruction performance of our proposal[1] compared to the Brevdo method [10]. We consider a non-stationary signal made of 3 components which correspond to a sinusoid, a linear chirp and a sinusoidal FM chirp depicted in Fig. 1.
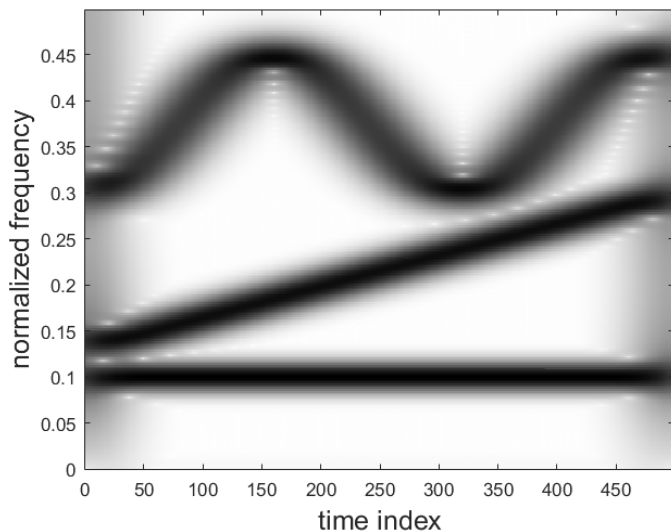

Fig. 1. Spectrogram of the analyzed multicomponent signal.

A. Robustness to Noise Analysis

We first consider the case of a linear chirp only (the second component presented in Fig. 1) merged with a white Gaussian noise with a SNR varying from -20 to 20 dB. We assess our approach using respectively the KLD, the β-D and the RD, compared with the Brevdo method [10]. The STFT was computed using $M = 500$ and $L = 50$ and we initialized our prior model using $m_0 = M/2$, $\sigma_0^2 = (M/2)^2/12$ and $\sigma_{rw}^2 = 2$. Hence, after estimation of the ridges, the signal is reconstructed (from Eq. (3)) by considering a vicinity of

[1]Matlab code freely available at: https://fourer.fr/eusipco21/

8 frequency bins around the estimated IF. The estimation performance of the ridge positions is assessed using the relative mean squared error RMSE $= \frac{1}{NM^2}\|\bar{m}_n - \hat{m}_n\|_2^2$, which is displayed for the competing methods in Fig 2.

It can first be remarked that the proposed approach performs better than the Brevdo one for almost all divergences and parameters choice in all scenarios. The exceptions are at high SNR with the β-D using $\beta \sim 1$, and at very low SNR (< -10dB). An interesting behaviour of the reconstruction RMSE can be noted when using the β-D. Indeed, there is an inversion of the tendency between the low and high SNR regimes. In the absence of important noise (SNR > -5dB), $\beta \sim 0$ provide the best estimation while the best results for low SNR are obtained using the robust approach ($\beta \sim 1$). Interestingly, the KLD is the most satisfying criterion until -10dB, after what the β-D provides the best reconstruction. The seeking property of the Rényi divergence allows efficient estimation at high SNR since thinner posterior distribution are computed in those cases. It also allows to circumvent the particular nature of the noise spreading in the TFR at low SNR [12]. We also assess the quality of the mode retrieval, which are comparatively evaluated in terms of Reconstruction Quality Factor (RQF): $10 \log_{10}\left(\frac{\|x\|^2}{\|x-\hat{x}\|^2}\right)$ where x (resp. $\hat{x}$) stands for the reference (resp. estimated) signal. The RQF related to the experiments performed in Fig. 2 are displayed in Fig. 3. As expected, the better ridge estimation performance of the proposed method allows for more efficient mode retrieval than Brevdo. However, the remaining noise present in the TFR bins used for reconstruction makes those performance less evident.
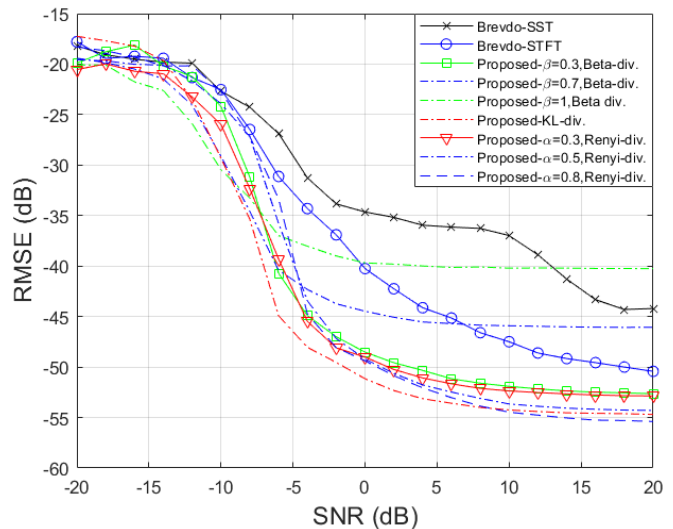

Fig. 2. RMSE (in dB) of the ridge positions (averaged over 100 realizations) obtained with the different competing methods.

B. Multi-Ridge Detection and Mode Retrieval

We now consider the whole MCS displayed in Fig. 1 for comparative assessment purpose. In this regard, we present the resulting RQFs obtained with the competing methods in Table I for a SNR = 10 dB and $L = 20$.
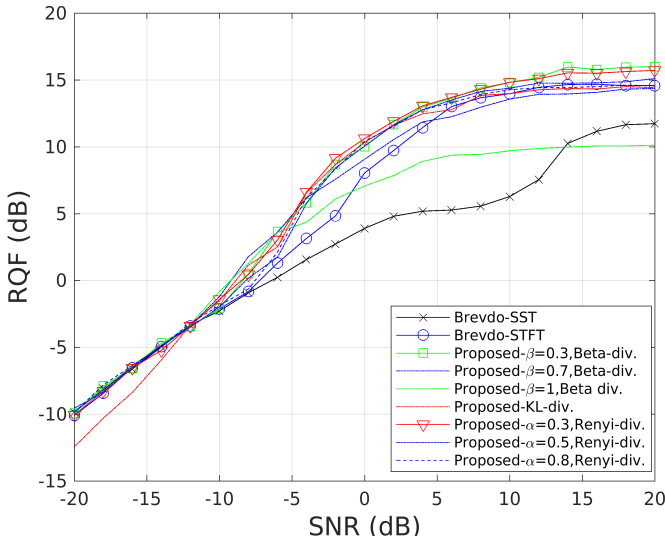
Fig. 3. RQF of competing methods (averaged over 100 realizations).

TABLE I
RQF OF EACH COMPONENTS (AVERAGED OVER 100 REALIZATIONS) FOR
THE DIFFERENT COMPETING APPROACHES FOR A SNR = 10 dB.

	Sinusoid	Linear chirp	Sin. FM chirp	Average
Brevdo-STFT	16.10	**15.46**	2.86	11.47
Brevdo-Synchrosqueezing	16.43	15.34	5.24	12.34
PB-β-D, $\beta = 0.5$	**16.71**	15.22	9.13	13.69
PB-β-D, $\beta = 0.8$	16.45	14.92	5.49	12.29
PB-KLD	2.46	2.65	1.18	2.10
PB-RD, $\alpha = 0.5$	16.59	15.24	**9.57**	**13.80**
PB-RD, $\alpha = 0.8$	15.44	15.22	7.84	12.83

The results show that our method obtains the best averaged RQF using RD ($\alpha = 0.5$). We note that in the presence of moderate noise level (SNR=10dB), performing MLE does not efficiently reconstruct the signal. Two differences to the previous experiment in Fig. 2 have to be highlighted. First, the SNR does not involve the same amount of noise since three components are present in the considered signal. Then, the presence of multiple ridges in the TFR involves important model mismatch, since the incomplete model in Eq. (7) does not only neglect noise disturbances but assumes for the presence of a single component. Even though the reconstruction performances of the proposed algorithm are almost equivalent to those of the Brevdo approach for both the sinusoid and the linear chirp it performs much better for the recovery of the sinusoidal FM chirp with a maximum RQF of 9.57 dB.

VI. CONCLUSION

In this paper, we have introduced a novel pseudo-Bayesian procedure for estimating the ridge position in order to retrieve the elementary components from a noisy MCS. The robust approach can provide better results than the Brevdo method while the simple observation model allows similar computational cost. An alternative procedure to the classical Bayesian approach has been adopted to account for the lack of accuracy of the observation model. More precisely, the β-D improves the robustness of the estimation process and allows for model

mismatch, while the RD provides adapted tools to account for the challenge of close ridges. The comparison with other approaches has demonstrated a significant improvement of the estimation performance of the ridge position. Moreover, the proposed algorithm seems well adapted for the extraction of frequency modulated signals. Future work will further investigate the alpha-beta divergence [19], merging the property of both β-D and RD. Nonetheless, the hyperparameters and mask width choices need to be considered. The method could also be improved using the synchrosqueezing method.

REFERENCES

[1] D. Fourer, J. Harmouche, J. Schmitt, T. Oberlin, S. Meignen, F. Auger, and P. Flandrin, "The ASTRES toolbox for mode extraction of non-stationary multicomponent signals," in *Proc. EUSIPCO'17*, Aug. 2017.
[2] P. Flandrin, *Time-Frequency/Time-Scale analysis.* Acad. Press, 1998.
[3] S. Meignen, T. Oberlin, and S. McLaughlin, "A new algorithm for multicomponent signals analysis based on synchrosqueezing: With an application to signal sampling and denoising," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 5787–5798, 2012.
[4] D. Fourer, F. Auger, and G. Peeters, "Local AM/FM parameters estimation: application to sinusoidal modeling and blind audio source separation," *IEEE Signal Processing Lett.*, vol. 25, Oct. 2018.
[5] R. A. Carmona, W. L. Hwang, and B. Torrésani, "Characterization of signals by the ridges of their wavelet transforms," *IEEE Trans. Signal Process.*, vol. 45, no. 10, pp. 2586–2590, 1997.
[6] ——, "Multiridge detection and time-frequency reconstruction," *IEEE Trans. Signal Process.*, vol. 47, no. 2, pp. 480–492, 1999.
[7] L. Stankovic, M. Dakovic, and V. Ivanovic, "Performance of spectrogram as IF estimator," *Electronics Lett.*, vol. 37, no. 12, 2001.
[8] D. Fourer, F. Auger, K. Czarnecki, S. Meignen, and P. Flandrin, "Chirp rate and instantaneous frequency estimation: Application to recursive vertical synchrosqueezing," *IEEE Signal Processing Lett.*, vol. 24, no. Issue 11, pp. 1724–1728, Nov. 2017.
[9] M. Lagrange, S. Marchand, and J.-B. Rault, "Using linear prediction to enhance the tracking of partials," in *Proc. IEEE ICASSP*, vol. 4, 2004.
[10] E. Brevdo, N. S. Fuckar, G. Thakur, and H.-T. Wu, "The synchrosqueezing algorithm: a robust analysis tool for signals with time-varying spectrum," *Arxiv preprint*, 2011.
[11] I. Daubechies and S. Maes, "A nonlinear squeezing of the continuous wavelet transform," *Wavelets in Medecine and Bio.*, pp. 527–546, 1996.
[12] P. Flandrin, "Time-frequency filtering based on spectrogram zeros," *IEEE Signal Processing Lett.*, vol. 22, no. 11, Nov. 2015.
[13] Q. Legros, J. Tachella, R. Tobin, A. Mccarthy, S. Meignen, G. S. Buller, Y. Altmann, S. Mclaughlin, and M. E. Davies, "Robust 3d reconstruction of dynamic scenes from single-photon lidar using beta-divergences," *IEEE Transactions on Image Processing*, vol. 30, pp. 1716–1727, 2021.
[14] A. Basu, I. R. Harris, N. L. Hjort, and M. Jones, "Robust and efficient estimation by minimising a density power divergence," *Biometrika*, vol. 85, no. 3, pp. 549–559, 1998.
[15] Y. Li and R. E. Turner, "Rényi divergence variational inference," *arXiv preprint arXiv:1602.02311*, 2016.
[16] F. Futami, I. Sato, and M. Sugiyama, "Variational inference based on robust divergences," in *Proc. Intern. Conf. on Artificial Intel. and Stat., (AISTATS)*, A. J. Storkey and F. Pérez-Cruz, Eds., vol. 84. Playa Blanca, Lanzarote, Canary Islands, Spain: PMLR, 2018, pp. 813–822.
[17] T. Oberlin, S. Meignen, and V. Perrier, "Second-order synchrosqueezing transform or invertible reassignment? towards ideal time-frequency representations," *IEEE Trans. Signal Process.*, vol. 63, no. 5, 2015.
[18] Q. Legros, S. McLaughlin, Y. Altmann, S. Meignen, and M. E. Davies, "Robust depth imaging in adverse scenarios using single-photon lidar and beta-divergences," in *Proc. Sensor Signal Processing for Defence Conference (SSPD)*. IEEE, 2020, pp. 1–5.
[19] J.-B. Regli and R. Silva, "Alpha-beta divergence for variational inference," *arXiv preprint arXiv:1805.01045*, 2018.
[20] X. Yang, "Understanding the variational lower bound," Univ. Maryland, Tech. Rep., 2017.
[21] V. Sucic, N. Saulig, and B. Boashash, "Estimating the number of components of a multicomponent nonstationary signal using the short-term time-frequency Rényi entropy," *EURASIP Journal on Advances in Signal Processing*, vol. 2011, no. 1, pp. 1–11, 2011.