Quasi-Periodic Non-negative Matrix Factorization for Phonocardiographic signals denoising

Nafissa DIA[1,2], Julie Fontecave-Jallon[1], Pierre-Yves Gumery[1], Bertrand Rivet[2]
[1]Univ. Grenoble Alpes, CNRS, CHU Grenoble Alpes, Grenoble INP*, TIMC-IMAG, F-38000 Grenoble, France
[2]Univ. Grenoble Alpes, CNRS, Grenoble INP*, GIPSA-lab, 38000 Grenoble, France
*Institute of Engineering Univ. Grenoble Alpes
[1]surname.name@univ-grenoble-alpes.fr [2]bertrand.rivet@gipsa-lab.grenoble-inp.fr

***Abstract*—Mechanical cardiac activity may be monitored with phonocardiographic (PCG) signals giving access to cardiac sounds. However, many noises interfere with cardiac information in raw signals and denoising such signals is necessary before interpretation. Non-negative Matrix Factorization (NMF) is of interest for time-frequency representations to separate noise and signal components. In this paper, to exploit the quasi-periodicity of the PCG, a quasi-periodic NMF (QP-NMF), based on multiplicative updates derived from a Majoration-Minimization algorithm, is proposed to decompose the PCG spectrograms. Numerical simulations show the good behavior of the proposed method to separate quasi-periodic components from the others. Finally, applied on real noisy PCG signals, QP-NMF shows its interest compared to an unsupervised NMF to denoise PCG signals.**

I. Introduction

Phonocardiogram signals, noted PCG, are cardiac signals recorded with acoustic sensors of microphonic type. They give access to cardiac sounds (sounds of the heart valves) which provide information about mechanical function of the heart. Two sounds are particularly audible, noted respectively S_1 and S_2 and corresponding to the closure of respectively the atrial-ventricular valves (beginning of the ventricular systole) and the aortic and pulmonary valves (onset of the ventricular diastole). As shown in Fig. 1, a PCG signal is therefore a succession of two bumps S_1 and S_2, following the R peak of a synchronous electrocardiogram (ECG) signal. Both ECG and PCG signals are quasi-periodic, due to the physiological phenomenon of variation in the time interval between heartbeats (heart rate variability). However raw PCG signals from cardiac microphones positioned on the chest are most often disturbed by many ambient interference (gastric or respiratory noises, cough ...) necessary to be removed before physiological interpretation.

PCG denoising has been investigated by many approaches of interest, such as adaptive filtering [1], Kalman filtering [2], wavelets [3], Empirical Modal Decomposition (EMD) [4] or Non-negative Matrix Factorization (NMF) [5], [6]. Among these methods, the NMF [7], [8] is particularly adapted to take into account the quasi-periodic property of physiological signals, since the algorithm applied to spectrograms is well-suited to identify events with particular spectrum and temporal regularity, such as quasi-periodicity [9]. Indeed, the NMF method decomposes a signal in k components, by approximating an observation matrix $X \in \mathbb{R}_+^{m\times n}$ of positive or zero coefficients by a product of two matrices with non-negative values $W \in \mathbb{R}_+^{m\times k}$ and $H \in \mathbb{R}_+^{k\times n}$: $X = V + N$ where $V = WH$ and N is the residual error from the approximation. In the specific case of physiological signal spectrogram observations, W is therefore considered as the matrix of frequency patterns and H the matrix of time activations. However, the quasi-periodic structure of PCG signals could be even more exploited.

In this paper, we propose to include to the cost functions some specific criteria of quasi-periodicity. This will allow to extract signal components out of noisy PCG signals, based on the important property of quasi-periodicity of PCG signals.

Fig. 1. Synchronous ECG and PCG signals. Time interval between heartbeats is not constant.

II. Quasi-periodic NMF

In this section, a brief recall of using NMF for source separation is presented (Section II-A) before the proposed extension to a quasi-periodic NMF (Section II-B) and the detailed algorithm in Section II-C.

A. NMF for source separation

Let us assume that the observed matrix X is the linear mixture of two sources $V^{(1)}$ and $V^{(2)}$ so that

$$X = V^{(1)} + V^{(2)} + N, \tag{1}$$

with $V^{(i)} = W^{(i)}H^{(i)}$ ($W^{(i)} \in \mathbb{R}_+^{m\times k_i}$, $H^{(i)} \in \mathbb{R}_+^{k_i\times n}$ for $i \in \{1,2\}$). A classical way to tackle the estimation of $V^{(1)}$ and $V^{(2)}$ is to apply a two steps algorithm. The first step performs a NMF of $X = V + N$ where $V = WH$ with $W = [W^{(1)}, W^{(2)}]$ and $H = [{H^{(1)}}^T, {H^{(2)}}^T]^T$ ($\cdot^T$ is the

transposition operator). In this paper, the euclidian distance is used

$$\hat{W}, \hat{H} = \arg \min_{W \in \mathbb{R}_+, H \in \mathbb{R}_+} C_D(X|W,H), \quad (2)$$

with

$$C_D(X|W,H) = \frac{1}{2}\|X - WH\|_F^2, \quad (3)$$

and the optimization is based on a majoration-minimization (MM) algorithm using multiplicative updates [10] to decouple the variables. Then in a second step, a clustering of the estimated patterns of $\hat{W}$ and $\hat{H}$ into $\{\hat{W}^{(1)}, \hat{H}^{(1)}\}$ and $\{\hat{W}^{(2)}, \hat{H}^{(2)}\}$ is done to identify the two sources.

However, this unsupervised NMF is often inaccurate in clustering the estimated patterns into specific components related to a single source. For instance, the two sources can partially share some spectral patterns so that the related activation patterns remain mixtures of the two sources. To tackle this problem some prior knowledge on the spectral patterns and/or on the activation ones can be added to the NMF criterion (2) to enforce the separation during the NMF decomposition. For instance, one can use a spectral and/or temporal reference (e.g., [11]–[13]).

B. Criterion for quasi-periodic NMF

In our application and in addition to the data fitting term $C_D(X|W,H)$, some penalization terms are used to constrain as much as possible $V^{(1)}$ and $V^{(2)}$ to be related to the PCG signal and to the other sources of interference, respectively.

The main prior knowledge is on the activation patterns since the heart produces the beats regularly in time (Fig. 1). However, the PCG signal is only a quasi-periodic signal: all the beats have slightly different durations and each beat pattern is sightly different from the others. To take into account the changes of rhythm, a time-wrapping matrix $T^{(1)}$ is introduced: $V^{(1)} = W^{(1)}H^{(1)}T^{(1)}$ so that the components of $H^{(1)}$ have now a regular rhythm instead of the original quasi-regular one. For the considered application, this matrix $T^{(1)}$ is estimated from the ECG signal by detecting the R peaks [14]. In practice, the observed data are pre-time-wrapped by the inverse time-wrapping $\left(T^{(1)}\right)^{-1}$ leading to $X\left(T^{(1)}\right)^{-1} = W^{(1)}H^{(1)} + W^{(2)}H^{(2)}\left(T^{(1)}\right)^{-1} + N\left(T^{(1)}\right)^{-1}$. For the sake of simplicity, in the remaining of Section II, $X\left(T^{(1)}\right)^{-1}$, $H^{(2)}\left(T^{(1)}\right)^{-1}$ and $N\left(T^{(1)}\right)^{-1}$ are simply denoted by X, $H^{(2)}$ and N, respectively.

To constrain $V^{(1)}$ to be quasi-periodic, the first term is thus a penalization on $H^{(1)}$ to impose that its k_1 components are P-periodic:

$$C_\pi^P\left(W^{(1)}, H^{(1)}\right) = \frac{1}{2}\sum_{k=1}^{k_1}\sum_{l=P+1}^{n} \beta_k^{(1)}\left[H_{k,l-P}^{(1)} - H_{k,l}^{(1)}\right]^2, \quad (4)$$

where $\beta_k^{(1)} = \sum_{f=1}^m \left(W_{f,k}^{(1)}\right)^2$. Indeed, it enforces the activation pattern of a beat to be similar to the previous beat. As in [15], to avoid trivial solution, $H^{(1)}$ decreasing toward 0, the $\beta_k^{(1)}$ are introduced.

Two priors are used to enforce the split of the cardiac activity in $V^{(1)}$ and the other interferences in $V^{(2)}$: $H^{(1)}$ and $H^{(2)}$ should be different to each other and $H^{(2)}$ should not be P-periodic. This first constraint is ensured by

$$C_{\neq}\left(\left\{W^{(i)}, H^{(i)}\right\}_i\right) = \sum_{k=1}^{k_1}\sum_{j=1}^{k_2}\lambda_k^{(1)}\lambda_j^{(2)}\sum_{l=1}^{n} H_{k,l}^{(1)}H_{j,l}^{(2)}, \quad (5)$$

where $\lambda_k^{(i)} = \sum_{f=1}^m W_{f,k}^{(i)}$. Indeed, the last summation in (5) is nothing else but the scalar product between the k-th component of $H^{(1)}$ and the j-th one of $H^{(2)}$. Again, the $\lambda_k^{(i)}$ are used to avoid trivial solutions $H^{(i)} = 0$. The non-periodicity criterion is defined by

$$C_{n\pi}^P\left(W^{(2)}, H^{(2)}\right) = \frac{1}{2}\sum_{k=1}^{k_2}\beta_k^{(2)}\sum_{l=P+1}^{n} H_{k,l-P}^{(2)}H_{k,l}^{(2)}. \quad (6)$$

Indeed, the last summation is the value of the auto-correlation function of the k-th component of $H^{(2)}$ for a lag equal to P.

Nevertheless, this straight implementation of the priors is not numerically suitable even with the $\beta_k^{(i)}$ and $\lambda_k^{(i)}$ terms. Indeed, the constraint (5) can be minimized by degenerated solutions where $H_{k,l}^{(1)} = 0$ while $H_{j,l}^{(2)} \neq 0$ (or the reverse) without splitting the cardiac activation patterns in $H^{(1)}$ and the interferences in $H^{(2)}$. To avoid such a degenerated solution, a temporal smoothness constraint on the components of $H^{(i)}$ is introduced: $H_{k,l}^{(i)}$ should be similar to the previous sample $H_{k,l-1}^{(i)}$:

$$C_s\left(W^{(i)}, H^{(i)}\right) = \frac{1}{2}\sum_{k=1}^{k_i}\sum_{l=2}^{n}\beta_k^{(i)}\left[H_{k,l-1}^{(i)} - H_{k,l}^{(i)}\right]^2. \quad (7)$$

It is worth noting that $C_s\left(W^{(i)}, H^{(i)}\right) = C_\pi^1\left(W^{(i)}, H^{(i)}\right)$.

Finally, the quasi-periodic NMF criterion is defined by

$$\begin{aligned} C_{QP}(X|W^{(1)}, H^{(1)}, W^{(2)}, H^{(2)}) &= C_D(X|W,H) \\ &+ \gamma_\pi C_\pi^P\left(W^{(1)}, H^{(1)}\right) + \gamma_s^{(1)} C_s\left(W^{(1)}, H^{(1)}\right) \\ &+ \gamma_{\neq} C_{\neq}\left(\left\{W^{(i)}, H^{(i)}\right\}_i\right) \\ &+ \gamma_{n\pi} C_{n\pi}^P\left(W^{(2)}, H^{(2)}\right) + \gamma_s^{(2)} C_s\left(W^{(2)}, H^{(2)}\right), \end{aligned} \quad (8)$$

where γ are ponderation coefficients between the different criteria.

C. Algorithm: auxiliary functions

To minimize (8), an alternating MM algorithm is used to derive multiplicative updates that ensure the non-negativity of the estimated $W^{(i)}$ (resp. $H^{(i)}$) given $H^{(i)}$ (resp. $W^{(i)}$). To this end, auxiliary functions $L(\cdot|\tilde{\cdot})$ (where $\tilde{\cdot}$ stands for the current estimation) of each term in (8) are expressed in Appendix. Vanishing the derivative of the overall auxiliary function of (8) leads to the update equations:

$$\forall i \in \{1,2\}, \quad W^{(i)} \leftarrow W^{(i)} \odot \frac{X\left(H^{(i)}\right)^T}{\tilde{X}\left(H^{(i)}\right)^T + \Psi^{(i)}} \quad (9)$$

and[1]

$$\forall i \in \{1,2\}, \quad H^{(i)} \leftarrow H^{(i)} \odot \frac{\left(W^{(i)}\right)^T X + \Phi^{(i)}}{\left(W^{(i)}\right)^T \tilde{X} + \Xi^{(i)}}, \qquad (10)$$

where $\tilde{X} = W^{(1)}H^{(1)} + W^{(2)}H^{(2)}$ and $\odot$ and the division are the element wise multiplication and division, respectively.

$$\Psi^{(1)} = \gamma_{\neq} 1_m 1_m^T W^{(2)} H^{(2,1)} + \gamma_s^{(1)} W^{(1)} \mathrm{diag}\Big(\Delta_1 H^{(1)}\Big) + \gamma_\pi W^{(1)} \mathrm{diag}\Big(\Delta_P H^{(1)}\Big),$$

with 1_m a column vector of m ones, $H^{(2,1)} = H^{(2)}\left(H^{(1)}\right)^T$, $\Delta_P H^{(1)}$ is a vector whose k-th entry is $\sum_{l=P+1}^{n}\left(H^{(1)}_{k,l} - H^{(1)}_{k,l-P}\right)^2$ and $\mathrm{diag}(\cdot)$ is a diagonal matrix whose diagonal elements are its arguments.

$$\Psi^{(2)} = \gamma_{\neq} 1_m 1_m^T W^{(1)} H^{(1,2)} + \gamma_s^{(2)} W^{(2)} \mathrm{diag}\Big(\Delta_1 H^{(2)}\Big) + \gamma_{n\pi} W^{(2)} \mathrm{diag}\Big(\Pi_P H^{(2)}\Big),$$

where $H^{(1,2)} = H^{(1)}\left(H^{(2)}\right)^T$ and $\Pi_P H^{(2)}$ is a vector whose k-th entry is $\sum_{l=P+1}^{n} H^{(2)}_{k,l} H^{(2)}_{k,l-P}$.

$$\Phi^{(1)} = \gamma_\pi \mathrm{diag}\big(\boldsymbol{\beta}^{(1)}\big)\Big[2H^{(1)} + H^{(1)}_{-P} + H^{(1)}_{P}\Big] + \gamma_s^{(1)} \mathrm{diag}\big(\boldsymbol{\beta}^{(1)}\big)\Big[2H^{(1)} + H^{(1)}_{-1} + H^{(1)}_{1}\Big],$$

where $H^{(1)}_{-P}$ (resp. $H^{(1)}_{P}$) is the matrix $H^{(1)}$ whose columns are shifted to the right (resp. left) by P columns and $\boldsymbol{\beta}^{(i)} = [\beta_1^{(i)}, \cdots, \beta_{k_i}^{(i)}]^T$.

$$\Xi^{(1)} = \gamma_{\neq}\, \boldsymbol{\lambda}^{(1)}\left(\boldsymbol{\lambda}^{(2)}\right)^T H^{(2)} + 4\gamma_\pi \mathrm{diag}\big(\boldsymbol{\beta}^{(1)}\big) H^{(1)} + 4\gamma_s^{(1)} \mathrm{diag}\big(\boldsymbol{\beta}^{(1)}\big) H^{(1)},$$

where $\boldsymbol{\lambda}^{(i)} = [\lambda_1^{(i)}, \cdots, \lambda_{k_i}^{(i)}]^T$.

$$\Phi^{(2)} = \gamma_s^{(2)} \mathrm{diag}\big(\boldsymbol{\beta}^{(2)}\big)\Big[2H^{(2)} + H^{(2)}_{-1} + H^{(2)}_{1}\Big],$$

and

$$\Xi^{(2)} = \gamma_{\neq} \boldsymbol{\lambda}^{(2)}\Big(\boldsymbol{\lambda}^{(1)}\Big)^T H^{(1)} + \gamma_{n\pi} \mathrm{diag}\big(\boldsymbol{\beta}^{(2)}\big)\Big[H^{(2)}_{-P} + H^{(2)}_{P}\Big] + 4\gamma_s^{(2)} \mathrm{diag}\big(\boldsymbol{\beta}^{(2)}\big) H^{(2)}.$$

The overall algorithm, called quasi-periodic NMF and denoted QP-NMF, estimates $\left\{W^{(i)}, H^{(i)}\right\}_i$ by alternating the updates of $W^{(i)}$ (9) and $H^{(i)}$ (10).

III. Results

The proposed QP-NMF methodology has been evaluated on simulated and real signals.

[1]Note that equation (10) does not hold for the first P and last P columns of the matrices due to the side effects. The complete equations are not given due to the lack of space but they can be easily obtained.

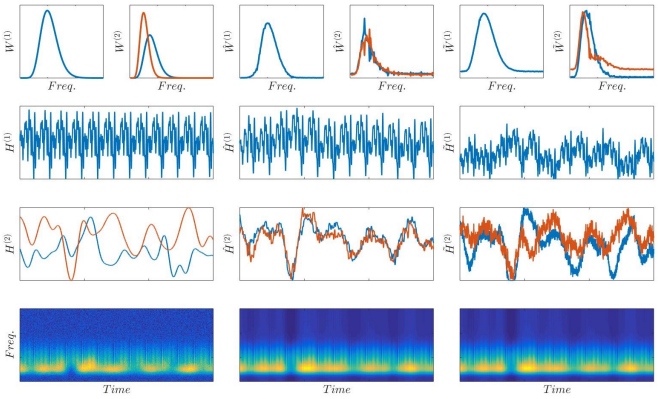

Fig. 2. Comparison of QP-NMF and NMF. Left column: simulated patterns. Middle column: estimated patterns by QP-NMF. Right column: estimated patterns by NMF. The last row plots X for the left column and $\hat{V}$ for the middle and right columns.

A. Numerical simulations

In these numerical experiments, a non-negative matrix $X \in \mathbb{R}_+^{n\times m}$ ($n = 5000$, $m = 256$) has been generated according to (1). A single quasi-periodic component ($k_1 = 1$) is generated and the interference sources have two components ($k_2 = 2$). The activation patterns $H^{(i)}$ have been randomly generated as Gaussian processes so that they are non-negative and $H^{(1)}$ has the desired quasi-periodic property. The spectral patterns $W^{(i)}$ have been randomly generated from gamma distributions. The additive noise N is uniformly distributed and white so that the average signal-to-noise ratio is 6dB. The ponderation coefficients are chosen for QP-NMF as $\gamma_\pi = 1$, $\gamma_s^{(1)} = 10$, $\gamma_{\neq} = 10^{-4}$, $\gamma_s^{(2)} = 10$ and $\gamma_{n\pi} = 10^{-3}$. The simulations are drawn randomly 50 times by regenerating all the parameters related to the patterns and with random initialization. For each simulation, the quality of the estimation is assessed by the error on $\hat{V}^{(i)} = \hat{W}^{(i)}\hat{H}^{(i)}$ defined as

$$\epsilon\big(\hat{V}^{(i)}\big) = \big\|\hat{V}^{(i)} - V^{(i)}\big\|_F, \qquad (11)$$

where $\hat{W}^{(i)}$ and $\hat{H}^{(i)}$ are the estimates. The figure 2 illustrates the simulated patterns and the estimated ones. As one can see, the proposed QP-NMF succeeded to estimate $V^{(1)}$ while the unsupervised NMF (applied with $k = 3$ components) failed: $H^{(1)}$ is definitively better estimated by QP-NMF than by NMF. This is confirmed by the average performance (Fig. 3). Indeed, while the error on $\hat{V}$ is almost the same for the two methods ($\epsilon(\hat{V}) = 40$), meaning that the sum $V^{(1)} + V^{(2)}$ is estimated with the same accuracy, the QP-NMF provides a better estimation of $\hat{V}^{(1)}$ than by the NMF method ($\epsilon(\hat{V}^{(1)}) = 11$ vs. 27 in median values). The QP-NMF also leads to slightly worse estimation for $V^{(2)}$ than the NMF method ($\epsilon(\hat{V}^{(2)}) = 42$ vs. 33 in median values).

B. Real PCG signals

Behaviour of the proposed QP-NMF is also analyzed for real PCG signals with additive real interference signals.

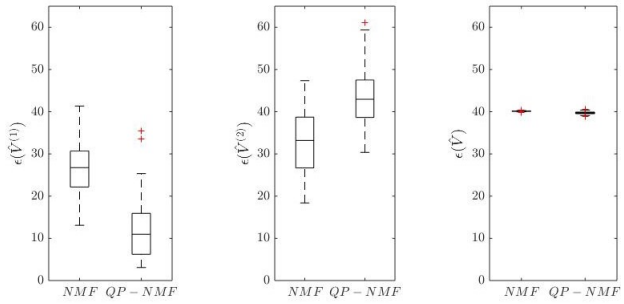

Fig. 3. Performance of QP-NMF algorithm on simulated spectrogram observations, by comparison to standard NMF. The distance between estimations and reference simulations are computed for the quasi-periodic component, the free components and the whole set of components.

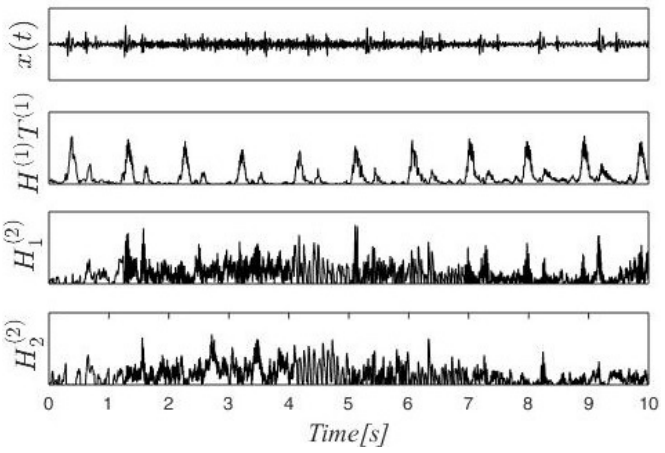

Fig. 4. Time activations from QP-NMF applied on one noisy PCG signal. $x(t)$ is the noisy PCG signal, $H^{(1)}T^{(1)}$ is the quasi-periodic component of the signal (corresponding to the "clean" PCG part) and $H_i^{(2)}$ is the i-th interference component ($i = 1, 2$).

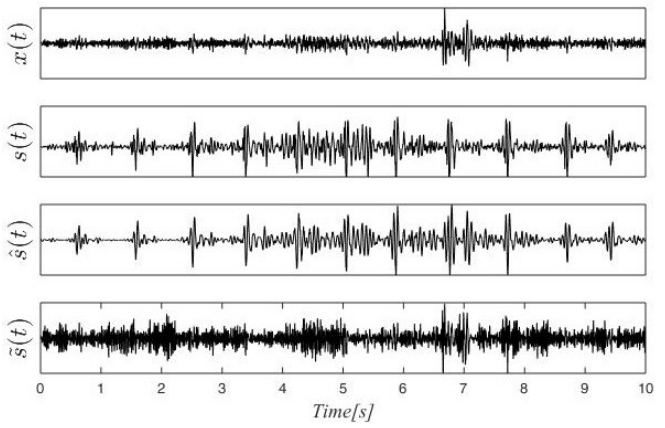

Fig. 5. Real PCG signals denoising. $x(t)$ is the noisy PCG signal, $s(t)$ the original clean PCG, $\hat{s}(t)$ (respectively $\tilde{s}(t)$) the denoised PCG estimation thanks to Wiener filtering based on the components separated by QP-NMF (respectively by standard NMF). Both NMF are applied with $k = 3$ components and consider $k_1 = 1$ signal component and $k_2 = 2$ interference components.

The database previously proposed to Signal Separation Evaluation Campaign in 2016 (SiSEC 2016) [16] is used for evaluation. It consists of sixteen samples with a duration varying from 10 to 70 seconds. Each sample is composed of a clean PCG $s(t)$, an interference signal $n(t)$, the artificially noisy PCG $x(t) = s(t) + n(t)$ and the synchronous ECG $ecg(t)$. As for the simulation evaluation, we consider $k_1 = 1$ quasi-periodic component and $k_2 = 2$ interference components. The algorithm is applied with the following values for penalization parameters: $\gamma_\pi = .7$, $\gamma_s^{(1)} = 10$, $\gamma_{\neq} = 5e-3$, $\gamma_{n\pi} = 4e-1$, $\gamma_s^{(2)} = 10$.

Fig. 4 illustrates the time activations from QP-NMF for one noisy PCG signal. The quasi-periodic part of the PCG signal is well represented by $H^{(1)}$ component and does not seem to be present in $H^{(2)}$ components, which mainly consist of the interferences. So as to judge the quality of separation by QP-NMF between the quasi-periodic components and the interference components, estimation of denoised PCG signals can be realized and compared to the original "clean" PCG signals of the database. QP-NMF allows to define signal and noise components, from which spectral densities can be computed to define a Wiener filter. This filter is then applied to the noisy PCG to estimate a denoised PCG reconstruction, noted $\hat{s}(t)$. A similar approach is applied on components out of a standard NMF. From the 3 components estimated, we visually choose the signal component, as the one which is the most periodic, according to the ECG reference. We consider as noise components the 2 remaining ones. The estimation of denoised PCG from standard NMF is obtained according to the related Wiener filter and is noted $\tilde{s}(t)$. Fig. 5 illustrates the reconstruction for both QP-NMF and standard NMF for one real noisy PCG signal. We observe that the impulse noises around 7s does not appear anymore in the denoised PCG $\hat{s}(t)$ with QP-NMF, which is not the case for $\tilde{s}(t)$ obtained after standard NMF. This example shows that it is of interest to constrain the activation profiles $H^{(1)}$ to be quasi-periodic.

IV. Conclusion

In this paper, we have proposed a new supervised NMF method to constrain explicitly some components to be quasi-periodic. This approach may be of interest to denoise quasi-periodic signals, such as PCG, since it allows the separation of signal and noise components directly during the components estimation compared to an unsupervised NMF that sorts the components after their estimation. The proposed method has been evaluated on simulated signals, highlighting that quasi-periodic part of the observations are better estimated with supervised QP-NMF than with unsupervised standard NMF. Moreover, observations on noisy real PCG recordings have been carried out and preliminary results are promising since the QP-NMF seems to remove some specific interference occurring in daily situations.

In future works, an automatic selection of the penalization parameters will be studied as well as the extension of this QP-NMF to model several quasi-periodic components with different periods.

APPENDIX

For the generic periodic penalization C_π^P (also used for the smoothness C_s with $P=1$)

$$L_\pi^P\Big(W^{(1)}\Big|\tilde{W}^{(1)}\Big) = C_\pi^P\Big(W^{(1)}, H^{(1)}\Big)$$

since C_π^P is a convex function with respect to $W^{(1)}$ and

$$L_\pi^P\Big(H^{(1)}\Big|\tilde{H}^{(1)}\Big) = \frac{1}{4}\sum_{k=1}^{k_1}\beta_k^{(1)}\sum_{l=P+1}^{n}\Big[\Big(2H_{k,l}^{(1)} - \tilde{H}_{k,l}^{(1)} - \tilde{H}_{k,n-P}^{(1)}\Big)^2 + \Big(-2H_{k,l-P}^{(1)} + \tilde{H}_{k,l-P}^{(1)} + \tilde{H}_{k,n}^{(1)}\Big)^2\Big].$$

For the non-periodic penalization $C_{n\pi}^P$

$$L_{n\pi}^P\Big(W^{(2)}\Big|\tilde{W}^{(2)}\Big) = C_{n\pi}^P\Big(W^{(2)}, H^{(2)}\Big),$$

since $C_{n\pi}^P$ is a convex function with respect to $W^{(2)}$ and

$$L_{n\pi}^P\Big(H^{(2)}\Big|\tilde{H}^{(2)}\Big) = \sum_{k=1}^{k_2}\beta_k^{(2)} \times \sum_{l=P+1}^{n}\left[\Big(H_{k,l}^{(2)}\Big)^2\frac{\tilde{H}_{k,n-P}^{(2)}}{2\tilde{H}_{k,l}^{(2)}} + \Big(H_{k,l-P}^{(2)}\Big)^2\frac{\tilde{H}_{k,n}^{(2)}}{2\tilde{H}_{k,l-P}^{(2)}}\right].$$

For the antagonism penalization $C_{\neq}$

$$L_{\neq}\Big(W^{(1)}, W^{(2)}\Big|\tilde{W}^{(1)}, \tilde{W}^{(2)}\Big) = \sum_{k=1}^{k_1}\sum_{j=1}^{k_2}\sum_{f=1}^{m}\sum_{l=1}^{m}H_{k,j}^{(1,2)} \times \left[\Big(W_{f,k}^{(1)}\Big)^2\frac{\tilde{W}_{l,j}^{(2)}}{2\tilde{W}_{f,k}^{(1)}} + \Big(W_{l,j}^{(2)}\Big)^2\frac{\tilde{W}_{f,k}^{(1)}}{2\tilde{W}_{l,j}^{(2)}}\right],$$

where $H_{k,j}^{(1,2)} = \sum_{l=1}^{n}H_{k,l}^{(1)}H_{j,l}^{(2)}$ and

$$L_{\neq}\Big(H^{(1)}, H^{(2)}\Big|\tilde{H}^{(1)}, \tilde{H}^{(2)}\Big) = \sum_{k=1}^{k_1}\sum_{j=1}^{k_2}\lambda_k^{(1)}\lambda_j^{(2)} \times \sum_{l=1}^{n}\left[\Big(H_{k,l}^{(1)}\Big)^2\frac{\tilde{H}_{j,l}^{(2)}}{2\tilde{H}_{k,l}^{(1)}} + \Big(H_{j,l}^{(2)}\Big)^2\frac{\tilde{H}_{k,l}^{(1)}}{2\tilde{H}_{j,l}^{(2)}}\right].$$

ACKNOWLEDGMENT

This work has been partially supported by the LabEx PERSYVAL-Lab (ANR-11-LABX-0025-01) funded by the French program Investissements d'avenir. This work is also partially supported by the French National Research Agency, as part of the SurFAO project (ANR-17-CE19-0012).

REFERENCES

[1] M. Tinati, A. Bouzerdoum, and J. Mazumdar, "Modified adaptive line enhancement filter and its application to heart sound noise cancellation," in *Proc. Int. Symp. on Signal Processing and Its Applications (ISSPA)*, Gold Coast Australia, 1996, pp. 815–818.

[2] S. Charleston and M. Azimi-Sadjadi, "Reduced order Kalman filtering for the enhancement of respiratory sounds," *IEEE Transactions on Biomedical Engineering*, vol. 43, no. 4, pp. 421–424, apr 1996.

[3] S. R. Messer, J. Agzarian, and D. Abbott, "Optimal wavelet denoising for phonocardiograms," *Microelectronics Journal*, vol. 32, no. 12, pp. 931–941, dec 2001.

[4] O. Beya, E. Fauvet, and O. Laligant, "EDA, approche non linéaire de débruitage des signaux cardiaques," in *Proc. CORESA*, Le Creusot, France, nov 2013.

[5] N. Dia, J. Fontecave-jallon, B. Rivet, and P.-y. Guméry, "Application De La Factorisation Non-Négative Des Matrices (NMF) Pour Le Débruitage Des Signaux Phonocardiographiques (PCG)," in *Proc. GRETSI*, Juan-Les-Pins, France, 2017.

[6] N. Dia, J. Fontecave-jallon, P.-y. Gumery, and B. Rivet, "Denoising Phonocardiogram signals with Non-negative Matrix Factorization informed by synchronous Electrocardiogram," in *Proc. European Signal Processing Conference (EUSIPCO)*, 2018, pp. 1–5, (Submitted to).

[7] D. D. Lee and H. S. Seung, "Algorithms for Non-negative Matrix Factorization," in *Proc. Neural Information Processing Systems (NIPS)*, Vancouver, Canada, 2001.

[8] F. Sedighin, M. Babaie-Zadeh, B. Rivet, and C. Jutten, "Multimodal Soft Nonnegative Matrix Co-Factorization for Convolutive Source Separation," *IEEE Transactions on Signal Processing*, vol. 65, no. 12, pp. 3179–3190, jun 2017.

[9] M. Niegowski and M. Zivanovic, "ECG-EMG separation by using enhanced non-negative matrix factorization," in *Proc. Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Chicago, USA, aug 2014, pp. 4212–4215.

[10] C. Févotte and J. Idier, "Algorithms for Nonnegative Matrix Factorization with the β-Divergence," *Neural Computation*, vol. 23, no. 9, pp. 2421–2456, sep 2011.

[11] D. Kitamura, H. Saruwatari, K. Shikano, K. Kondo, and Y. Takahashi, "Music signal separation by supervised nonnegative matrix factorization with basis deformation," in *Proc. Int. Conf. Digital Signal Processing (DSP)*, 2013, pp. 1–6.

[12] N. Seichepine, S. Essid, C. Fevotte, and O. Cappe, "Soft Nonnegative Matrix Co-Factorization," *IEEE Transactions on Signal Processing*, vol. 62, no. 22, pp. 5940–5949, nov 2014.

[13] N. Souviraa-Labastie, A. Olivero, E. Vincent, and F. Bimbot, "Multi-Channel Audio Source Separation Using Multiple Deformed References," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 11, pp. 1775–1787, nov 2015.

[14] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Transactions on Biomedical Engineering*, vol. BME-32, no. 3, pp. 230–236, March 1985.

[15] N. Seichepine, S. Essid, C. Fevotte, and O. Cappe, "Soft nonnegative matrix co-factorizationwith application to multimodal speaker diarization," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, may 2013, pp. 3537–3541.

[16] 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 *Proc. Int. Conf. Latent Variable Analysis and Signal Separation (LVA-ICA)*. Springer, Cham, feb 2017, pp. 323–332.