An Online Subspace Denoising Algorithm for Maternal ECG Removal from Fetal ECG Signals

Marzieh Fatemi, Reza Sameni

An Online Subspace Denoising Algorithm for Maternal ECG Removal from Fetal ECG Signals

Marzieh Fatemi and Reza Sameni*

January 2015

***Abstract*— Noninvasive extraction of fetal electrocardiogram (fECG) from multichannel maternal abdomen recordings is an emerging technology used for fetal cardiac diagnosis. The strongest interference for the fECG is the maternal ECG (mECG), which is not totally removed through conventional methods including blind source separation (BSS). In this work, we address the problem of offline maternal cardiac signal removal and introduce an online subspace denoising procedure for mECG cancellation. The proposed method is a general online denoising framework, which can be used for the extraction of the signal subspace from noisy multichannel observations in low signal-to-noise ratios, using suitable prior information of the signal or noise. The method is fairly generic and may also be useful for the separation of other signals and noise even in the cases that BSS assumptions are not satisfied. The performance of the proposed technique is evaluated on both real and synthetic data and has shown significant outperformance as compared with the state of the art methods.**

***Index Terms*— Online subspace denoising; semi-blind source separation; maternal ECG cancellation; noninvasive fetal ECG extraction; online generalized eigenvalue decomposition.**

I. Introduction

The fetal electrocardiogram (fECG) provides vital information about the fetal cardiac status. Recent measurement and processing technologies have enabled the extraction of the fECG noninvasively, from an array of sensors placed on the maternal abdomen [1]. One of the most challenging issues in this context is to remove maternal cardiac (mECG) interferences, without affecting the fECG. The mECG can be up to two orders of magnitude stronger than the fECG [1].

Previous work in mECG removal include spatial filtering [2], adaptive filtering [3]–[5], template subtraction techniques [6], [7] and Kalman filtering [8]–[10]. Although adaptive and Kalman filters can be used for multichannel data denoising, previous research have mainly used them for single-channel ECG denoising, i.e., have not incorporated the inter-channel dependencies of the data. However, in some of the single-channel mECG removal methods the fECG is removed with the mECG during periods of mECG and fECG temporal overlap [8]. This problem can be avoided by using multiple channels. A well-known multichannel technique for extraction of fECG is *blind source separation* (BSS) using *independent component analysis* (ICA), which has been shown to be more accurate and robust as compared to other approaches [11]. However, a basic limitation in conventional ICA is that its performance degrades in presence of full-rank Gaussian noise [12], resulting in residual mECG within the fECG. It is therefore better to remove the mECG before applying ICA techniques [13].

In recent years, a deflation subspace decomposition procedure, which we call *denoising by deflation* (DEFL), was proposed for signal subspace separation from full-rank noisy multichannel observations [8], [13]–[15]. An interesting application of this framework is for mECG removal from maternal abdominal recordings [13]. The method has resulted in very good fECG separation, especially in low signal-to-noise ratios (SNR). Yet, a limiting factor of DEFL is its offline block-wise procedure of generalized eigenvalue decomposition (GEVD), as the core of this algorithm, which is a drawback for its applicability in real-time online fECG extraction. In this work, we introduce an online version of DEFL, called *online denoising by deflation* (ODEFL), for eliminating the mECG from maternal abdominal recordings using recent developments in online GEVD [16]. As with the offline version, the proposed method is fairly general and may be applied to various applications depending on the prior knowledge about signal and noise subspaces.

II. Data Model

Electrical signals recorded from the abdomen of a pregnant woman consist of mixtures of various signals including the mECG, fECG, baseline wanders and muscle contractions considered as noise. Biopotentials recorded at the body surface are low frequency signals compared with the high propagation velocity of the electrical signals and the sensor distances [17]. Therefore, the following linear instantaneous data model has been shown to be rather realistic for modeling multichannel maternal abdominal signals [13]:

$$\begin{aligned}\mathbf{x}(t) &= \mathbf{H}_m(t)\mathbf{s}_m(t) + \mathbf{H}_f(t)\mathbf{s}_f(t) + \mathbf{H}_\eta(t)\mathbf{v}(t) + \mathbf{n}(t) \\ &\triangleq \mathbf{x}_m(t) + \mathbf{x}_f(t) + \boldsymbol{\eta}(t) + \mathbf{n}(t)\end{aligned} \tag{1}$$

where $\mathbf{s}_m(t)$, $\mathbf{s}_f(t)$ and $\mathbf{v}(t)$ are, respectively, the maternal ECG source, fetal ECG source and structured noises (such as electrode movements and muscle contractions). $\mathbf{n}(t)$ is full-rank measurement noise and $\mathbf{H}_m(t)$, $\mathbf{H}_f(t)$ and $\mathbf{H}_\eta(t)$ are the transfer functions that model the propagation media onto the body surface [18]. In a realistic model, the cardium (of the mother and fetus) should be considered as a distributed signal source. Therefore, $\mathbf{s}_m(t)$ and $\mathbf{s}_f(t)$ are generally full-rank signals [8]; but the effective number of dimensions can be less depending on the sensor positioning and SNR.

*School of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran. Email: reza.sameni@gmail.com, rsameni@shirazu.ac.ir Web: www.sameni.info.

Manuscript Under Review.

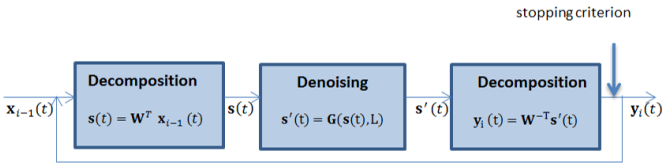


Fig. 1. General block-wise deflation scheme (i : the number of iteration, $\mathbf{x}_0(t) = \mathbf{x}(t)$)

The overall objective of noninvasive fECG extraction is to extract $\mathbf{x}_f(t)$ from this mixture. Among the different interferences and noises, the mECG is the dominant interference, which cannot be fully separated from the fECG through conventional ICA, due to its full-rank nature, high amplitude, and background noise. This results in residual components within the extracted fECG. The DEFL algorithm was proposed to overcome this issue [13]. Before introducing its online version, DEFL is explained in the following section.

III. Block-wise Denoising by Deflation

For multichannel noisy data $\mathbf{x}(t) \in \mathbb{R}^N$, DEFL consists of a sequence of *linear decomposition*, *denoising* and *linear recomposition* (Fig. 1). In the linear decomposition stage, the data is transformed to another subspace using a matrix $\mathbf{W} \in \mathbb{R}^{N\times N}$, obtained by applying GEVD on a pair of matrices containing second or higher order statistics of the data. A special case of GEVD, known as *periodic component analysis*, is explained further in the next section. Here, $\mathbf{W}$ is a matrix used as a *feature enhancer*, which transforms the data into a linear mixture ranked from most to least resemblance to the desired property. This enhancement is a result of maximizing the *Rayleigh quotient* in the GEVD procedure (cf. Section IV). Therefore, the SNR of the data is amplified within the first few channels of the transform domain, allowing better signal/noise separability. Next, the first L components (in the transform domain) are passed through a denoising stage, which separates their signal/noise contents. Finally, the residual signals and the $N-L$ unchanged channels are back-projected to the original subspace. These three stages construct a single iteration of the denoising algorithm, which can be summarized as follows:

$$\mathbf{y}(t) = \mathbf{W}^{-T}\mathbf{G}\big(\mathbf{W}^T\mathbf{x}(t), L\big) \tag{2}$$

where $\mathbf{x}(t)$ is the input and $\mathbf{y}(t)$ is the output of each iteration, $\mathbf{G}(\cdot,\cdot)$ is the denoising operator applied to the first L channels of the input, and $\mathbf{W}$ is as defined above. The procedure is repeated in multiple iterations over the output of the previous iteration, until all the undesired components within the data are eliminated. The number of iterations can be selected using a stopping criterion that is application dependent and measures the quality of the signal according to its characteristics. For instance, the *periodicity measure* (PM) criterion that shall be defined in Section VI-D can be used to indicate how much the maternal ECG is removed in each of the channels. An interesting property of this algorithm is that, unlike most PCA and ICA denoising schemes, the data dimensionality is preserved. Moreover, due to the denoising block between the linear projection stages, in overall it is a nonlinear filtering scheme, which can deal with full-rank and even non-additive terms. Apparently, the method is only applicable when prior information about the signal/noise subspaces is available. In previous studies, this algorithm has been used for various applications [13], [19], [20]. Despite its vast applications, the block-wise nature of the algorithm has limited its application to batch processing. In this work, an online extension of DEFL is presented using recent developments in online GEVD.

IV. Offline versus Online πCA

A. Offline πCA

Considering $\mathbf{x}(t) \in \mathbb{R}^N$ as multichannel ECG observations, *periodic component analysis* (πCA) can be used to transform $\mathbf{x}(t)$ into $s(t) = \mathbf{w}^T\mathbf{x}(t)$, such that $s(t)$ maximizes a measure of periodicity with the ECG heart rate period, while keeping the signal energy bounded [21].

The objective function of πCA is as follows:

$$\mathbf{w}^* = \underset{\mathbf{w}}{\operatorname{argmax}} \frac{\mathbf{E}_t\{s(t)s(t+\tau_t)\}}{\mathbf{E}_t\{s(t)^2\}} = \underset{\mathbf{w}}{\operatorname{argmax}} \frac{\mathbf{w}^T\mathbf{C}_\tau\mathbf{w}}{\mathbf{w}^T\mathbf{C}\mathbf{w}} \tag{3}$$

where $\mathbf{C} \triangleq \mathbf{E}_t\left\{\mathbf{x}(t)\mathbf{x}(t)^T\right\}$, $\mathbf{C}_\tau \triangleq \mathbf{E}_t\left\{\mathbf{x}(t)\mathbf{x}(t+\tau_t)^T\right\}$, with the expectations taken over all the samples t, and τ_t is a variable period calculated using the reference (here the maternal) ECG R-wave peaks. Equation (3), is in the form of the *Rayleigh quotient*, which is maximized by solving the following GEVD problem for $\mathbf{W}$:

$$\mathbf{W}^H\mathbf{C}_\tau\mathbf{W} = \mathbf{\Lambda}, \qquad \mathbf{W}^H\mathbf{C}\mathbf{W} = \mathbf{I} \tag{4}$$

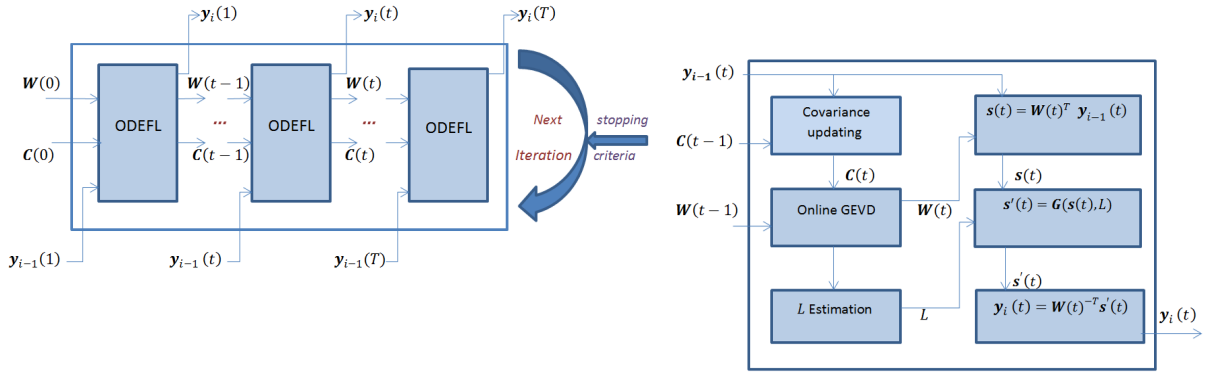
where $\mathbf{W} = [\mathbf{w}_1, \cdots, \mathbf{w}_N]$ is a matrix of generalized eigenvectors and $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \cdots, \lambda_N)$ is a diagonal matrix containing the generalized eigenvalues on its diagonal. It can be shown that $\mathbf{w}^* = \mathbf{w}_1$, i.e., the eigenvector corresponding to the largest generalized eigenvalue λ_1 maximizes (3). Moreover, if $\mathbf{C}$ and $\mathbf{C}_\tau$ are symmetric matrices, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$ are real and the components of $\mathbf{s}(t) = \mathbf{W}^T\mathbf{x}(t)$ are ranked according to their resemblance with the desired ECG [21]. πCA is a batch offline algorithm. In order to make it online (as required in ODEFL), the covariance matrices should be updated online.

B. Online Estimation of Covariance Matrices for πCA

For online applications, the signal statistics contained in $\mathbf{C}$ and $\mathbf{C}_\tau$, can vary in time. In order to re-estimate them as the signal evolves, the temporal averaging in the definitions of $\mathbf{C}$ and $\mathbf{C}_\tau$ can be replaced with a weighted sum as follows [22]:

$$\begin{aligned}\mathbf{C}(t) &= \sum_{i=0}^{t} \beta^{t-i}\mathbf{x}_s(i)\mathbf{x}_s(i)^T \\ \mathbf{C}_\tau(t) &= \sum_{i=0}^{t} \gamma^{t-i}\mathbf{x}_s(i)\mathbf{x}_s(i+\tau_i)^T\end{aligned} \tag{5}$$

where we have assumed that the signal indexes start from $t = 0$ and $\beta \in [0,1]$ and $\gamma \in [0,1]$ are forgetting factors. This is an *infinite impulse response*-like formulation, in which



(a) Block diagram for the whole data, $\mathbf{y}_0(t) = \mathbf{x}(t)$

(b) Block diagram for each sample

Fig. 2. General ODEFL scheme

all samples in the range $0 \leq i \leq t$ contribute in estimating the covariance matrices; but with smaller weights to the older samples. Alternatively, one might prefer a *finite impulse response* form, in which the samples do not have any effect beyond a finite window length.

The weighted sum in (5) can be replaced with the following recursion formulas, in favor of computational and memory efficiency:

$$\begin{aligned} \mathbf{C}(t) &= \beta \mathbf{C}(t-1) + \mathbf{x}(t)\mathbf{x}(t)^T \\ \mathbf{C}_\tau(t) &= \gamma \mathbf{C}_\tau(t-1) + \mathbf{x}(t)\mathbf{x}(t+\tau_t)^T \end{aligned} \tag{6}$$

The forgetting factors adapt the tracking capability of the algorithm in stationary or non-stationary environments. For stationary data, selecting $\beta = \gamma = 1$ incorporates all the samples with identical weights. For non-stationary data, the value is chosen less than 1, which is similar to a growing sliding window with the effective window length of $1/(1-\beta)$ for $t \gg 1$ [22].

In order to guarantee the symmetry of $\mathbf{C}$ and $\mathbf{C}_\tau$ (for having real generalized eigenvalues extracted by GEVD), we can make the following modification after each update:

$$\mathbf{C}(t) \leftarrow \frac{\mathbf{C}(t) + \mathbf{C}(t)^T}{2}, \quad \mathbf{C}_\tau(t) \leftarrow \frac{\mathbf{C}_\tau(t) + \mathbf{C}_\tau(t)^T}{2} \tag{7}$$

C. Online πCA using online GEVD

In recent research several algorithms have been proposed for online GEVD [16], [23]. These algorithms are commonly based on online extensions of (3) or (4). The method that we propose in this study is based on *incremental common spatial pattern* (ICSP) [16]. In this method, the first generalized eigenvector corresponding to the largest generalized eigenvalue is estimated first using the following recursion formula:

$$\mathbf{w}_1(t) = \frac{\mathbf{w}_1^T(t-1)\mathbf{C}_\tau(t)\mathbf{w}_1(t-1)}{\mathbf{w}_1^T(t-1)\mathbf{C}(t)\mathbf{w}_1(t-1)} \mathbf{C}_\tau^{-1}(t)\mathbf{C}(t)\mathbf{w}_1(t-1) \tag{8}$$

where $\mathbf{w}_1(t)$ is the principal eigenvector at time index t. The matrix $\mathbf{C}_\tau^{-1}(t)$ can be calculated recursively using the matrix inversion lemma:

$$\begin{aligned} \mathbf{C}_\tau^{-1}(t) = \; & \gamma \mathbf{C}_\tau^{-1}(t-1) \\ & - \frac{\gamma \mathbf{C}_\tau^{-1}(t-1)\mathbf{x}(t)\mathbf{x}^T(t+\tau_t)\mathbf{C}_\tau^{-1}(t-1)}{\gamma^{-1} + \mathbf{x}^T(t)\mathbf{C}_\tau^{-1}(t-1)\mathbf{x}(t+\tau_t)} \end{aligned} \tag{9}$$

The other minor eigenvectors are computed in a sequential (deflation) manner. For instance for the second generalized eigenvector, the new matrices $\hat{\mathbf{C}}_\tau = \mathbf{C}_\tau$ and $\hat{\mathbf{C}} = \left[\mathbf{I} - \frac{\mathbf{C}\mathbf{w}_1\mathbf{w}_1^T}{\mathbf{w}_1^T\mathbf{C}\mathbf{w}_1}\right]\mathbf{C}$ are used in (8).

V. Online Subspace Decomposition Algorithm for mECG Removal

Having developed the online version of πCA, DEFL can now be made online. The overall procedure is similar to the offline DEFL, except that the πCA projection, denoising and back-projection stages are all updated sample by sample, instead of block-wise. The general scheme of the ODEFL algorithm is shown in Fig. 2, which can be summarized as follows: In this algorithm, $\mathbf{x}(t)$ is the input multi-channel data, $\mathbf{y}_i(t)$ $(1 \leq i \leq K)$ is the output of each iteration, K is the number of iterations, T is the number of samples, and $\mathbf{G}_i(\cdot, L)$ is the denoising function for removing undesired parts[1], applied to the first L channels in iteration i.

We should note that since the online πCA algorithm requires the R-peak locations for calculating $\mathbf{C}_\tau(t)$, the update of this matrix has a minimum delay of one ECG beat, which can be fixed to the longest expected mECG beat gap. Therefore, the ODEFL output has a fixed delay with its input. This fixed lag, which is of order of a second, is acceptable for almost any online application of ODEFL.

VI. Benchmark Algorithms and Synthetic Dataset

In this section, the benchmark ANC, Kalman filter and the synthetic ECG generation procedure are presented.

[1]Note that the projection (and back-projection) algorithms and the denoising scheme can generally be customized for each iterations. This generalization has not been considered in this work.

Algorithm 1 Online denoising by deflation (ODEFL)

$\mathbf{y}_0(t) \leftarrow \mathbf{x}(t)$ ▷ Initialize with the input data
for $i = 1 \rightarrow K$ **do** ▷ The number of parallel stages of ODEFL
  $\mathbf{z}(t) \leftarrow \mathbf{y}_{i-1}(t)$ ▷ In each stage, initialize with the previous iteration output
  $\mathbf{C}(0) \leftarrow \mathbf{I}$; $\mathbf{C}_\tau(0) \leftarrow \mathbf{I}$
  for $t = 1 \rightarrow T$ **do** ▷ For all samples of the data
    ▷ Updating the covariance matrices
    $\mathbf{C}(t) \leftarrow \beta\mathbf{C}(t-1) + \mathbf{z}(t)\mathbf{z}(t)^T$
    $\mathbf{C}_\tau(t) \leftarrow \gamma\mathbf{C}_\tau(t-1) + \mathbf{z}(t)\mathbf{z}(t+\tau_t)^T$
    $\mathbf{C}(t) \leftarrow (\mathbf{C}(t) + \mathbf{C}(t)^T)/2$
    $\mathbf{C}_\tau(t) \leftarrow (\mathbf{C}_\tau(t) + \mathbf{C}_\tau(t)^T)/2$
    $\mathbf{C}_\tau^{-1}(t) \leftarrow \gamma\mathbf{C}_\tau^{-1}(t-1) - \dfrac{\gamma\mathbf{C}_\tau^{-1}(t-1)\mathbf{z}(t)\mathbf{z}^T(t)\mathbf{C}_\tau^{-1}(t-1)}{\gamma^{-1} + \mathbf{z}^T(t)\mathbf{C}_\tau^{-1}(t-1)\mathbf{z}(t)}$
    ▷ Online GEVD $(\mathbf{C}(t), \mathbf{C}_\tau(t))^T$
    $\hat{\mathbf{C}} \leftarrow \mathbf{C}(t)$
    for $j = 1 \rightarrow N$ **do** ▷ Iterate over the separating matrix channels
      $\mathbf{w}_j(t) \leftarrow \dfrac{\mathbf{w}_j^T(t-1)\mathbf{C}_\tau(t)\mathbf{w}_j(t-1)}{\mathbf{w}_j^T(t-1)\hat{\mathbf{C}}(t)\mathbf{w}_j(t-1)}\mathbf{C}_\tau^{-1}(t)\hat{\mathbf{C}}(t)\mathbf{w}_j(t-1)$
      $\mathbf{w}_j(t) \leftarrow \mathbf{w}_j(t)/\|\mathbf{w}_j(t)\|$
      $\hat{\mathbf{C}} \leftarrow \left[\mathbf{I} - \dfrac{\hat{\mathbf{C}}\mathbf{w}_j\mathbf{w}_j^T}{\mathbf{w}_j^T\hat{\mathbf{C}}\mathbf{w}_j}\right]\hat{\mathbf{C}}$
    end for
    $\mathbf{s}(t) \leftarrow \mathbf{W}^T(t)\mathbf{z}(t)$
    Estimate L, the number of mECG dimensions
    $\mathbf{s}'(t) \leftarrow \mathbf{G}_i(\mathbf{s}(t), L)$
    $\mathbf{y}_i(t) \leftarrow \mathbf{W}^{-T}(t)\mathbf{s}'(t)$
  end for
end for

A. Kalman Filter

The *Kalman filter* (KF) and its nonlinear version *extended Kalman filter* (EKF) are methods for estimating hidden states of a system, having its dynamics and a set of observations. This filter was adapted by Sameni *et al.*, for estimating ECG signals from noisy measurements [8]–[10]. In summary, using a polar extension of McSharry-Clifford's ECG model [24], the following *state space* and *observation* models were used as the ECG dynamic model:

$$\begin{cases} \theta_{k+1} = (\theta_k + \omega\delta) \ mod \ 2\pi \\ z_{k+1} = z_k - \sum_i \delta \frac{\alpha_i\omega}{b_i^2}\Delta\theta_i \exp[-\frac{\Delta\theta_i^2}{2b_i^2}] + \eta \end{cases} \tag{10}$$

$$\begin{cases} s_k = z_k + v_k \\ \phi_k = \theta_k + u_k \end{cases} \tag{11}$$

where $\Delta\theta_i = (\theta_k - \theta_i) \ mod \ 2\pi$, δ is the sampling period, η is an additive noise, and the summation is taken over finite number of Gaussian signals used for modeling P, Q, R, S and T waves with amplitude, center and width parameters α_i, θ_i and b_i, respectively. The variable z_k, the amplitude of the noiseless ECG at time instant k, and θ (known as the *cardiac phase*), are assumed as state variables for this model. The parameters $\theta_i, \omega, \alpha_i, b_i$ and η are i.i.d Gaussian random variables considered as process noise vectors. In the observation equations, s_k and ϕ_k are amplitude and phase of the noisy observation ECG and v_k and u_k are observation noise vectors of the ECG and its phase (cf. [9], [10] for further details). Using an EKF the ECG signal z_k can be estimated from the background noise v_k. For our application of interest, z_k is the maternal ECG, which should be estimated and removed from the maternal abdominal sensors.

B. Adaptive Noise Cancellation

Adaptive noise cancellation (ANC) is a well-known method for online signal denoising developed by Widrow *et al.* [3]. Standard ANC consists of a *primary input* that is the corrupted signal and a *reference input* containing the noise that is somehow correlated with the primary noise. The weights of the filter adaptively change over time to retrieve an estimate of the noise (the weight update algorithm depends on cost function). Finally, by subtracting the filter output (noise estimate) from the primary input, the primary signal is estimated and the corrupted signal is denoised. For mECG cancellation, the reference input is obtained by a mECG channel recorded directly from the maternal chest. The primary input is obtained by maternal abdomen recordings containing both maternal and fetal ECG. For multichannel recordings, the ANC is applied to each channel separately. As discussed in [10], the drawback of conventional ANC for ECG denoising is that the reference ECG should be morphologically similar to contaminating ECG. However, since the ECG morphology highly depends on

the lead position, the mECG contaminated over the maternal abdominal leads do not necessarily resemble the chest lead ECG morphology. As a result, the performance of this method widely defers from one channel to another, which leads into a weak overall performance over the whole channels, as compared to other methods. Nonetheless, the method remains as a well-know benchmark for mECG cancellation. More rigorously, considering $n(t)$ as the mECG, $s(t)$ as the non-mECG (fECG plus background signals), $d(t) = s(t) + n(t)$ as the noisy observations and $x(t)$ as the reference mECG, and $\mathbf{w} = [w_0, \cdots w_{p-1}]$ as the weight coefficient of length p, and using a *least mean squares* (LMS) algorithm, the output of an ANC is obtained as follows: where T is the number

Algorithm 2 Adaptive noise cancelation (ANC) algorithm

1: **for** $t = 1 \rightarrow T$ **do**
2: $\quad \mathbf{x}(t) = [x(t), x(t-1), \cdots x(t-p+1)]^T$
3: $\quad \hat{n}(t) = \mathbf{w}^T \mathbf{x}(t)$
4: $\quad \hat{s}(t) = d(t) - \hat{n}(t)$
5: $\quad \mathbf{w}(t+1) = \mathbf{w}(t) + 2\mu \hat{s}(t)\mathbf{x}(t)$
6: **end for**

of data samples, $\hat{n}(t)$ and $\hat{s}(t)$ are estimates of primary noise and primary signal, respectively. The parameter μ is a step size that controls the filter stability and convergence rate and should be in the range $[0, \lambda_{max}]$, where λ_{max} is the greatest eigenvalue of the covariance matrix $\mathbf{R} = \mathbf{E}\left\{\mathbf{x}(t)\mathbf{x}(t)^T\right\}$ [25, Ch. 9]. Recently, other extensions of the ANC have also been introduced for fECG extraction. One of the extensions that is used in this study for comparison is a multistage adaptive filter [5]. The modified ANC consists of two sequential adaptive filters, which enables the application of different kinds of adaptive algorithms such as LMS, *recursive least squares* (RLS) and *normalized least mean square* (NLMS) in one filter. Another aspect of this method is that the primary and reference inputs are applied to the algorithm after a sequence of operations such as squaring and/or rescaling to increase reliability of the algorithm to situations in which the maternal ECG in the primary input is not quite similar to the reference input.

C. ICA Based BSS and Denoising

ICA based BSS was first used in [11] for fECG extraction from maternal abdominal sensors. This method exploits the statistical independence and spatial diversity of the sources (here the maternal and fetal heart signals plus noises) for separating fECG from other signals. In classical ICA schemes, it is assumed that the signals $\mathbf{x}(t) \in \mathbb{R}^N$, observed from N body sensors, are linear mixtures of N independent sources $\mathbf{s}(t) \in \mathbb{R}^N$:

$$\mathbf{x}(t) = \mathbf{A}(t)\mathbf{s}(t) \tag{12}$$

in which the mixing matrix $\mathbf{A}(t) \in \mathbb{R}^{N \times N}$ models the propagation media and $\mathbf{s}(t)$ contains the source signals. ICA methods are then used to find the separating matrix $\mathbf{B}(t)$ such that $\hat{\mathbf{s}}(t) = \mathbf{B}(t)\mathbf{x}(t)$ can be considered as an estimate of the sources and $\hat{\mathbf{A}}(t) = \hat{\mathbf{B}}^{-1}(t)$ as an estimate of mixing matrix. Among the different ICA algorithms, the *joint diagonalization of eigenmatrices* (JADE) is used in this work [26].

In fECG applications, due to multidimensional nature of the sources, source signals are categorized into sets of multichannel components including mECG, fECG and noise subspaces as described in *multidimensional ICA* (MICA) [26] and *blind source subspace separation* (BSSS) [27] methods. Suppose that $\hat{\mathbf{s}}_f(t) = [\hat{s}_{f_1}(t), \cdots, \hat{s}_{f_M}]$ represents M-dimensional fetal components and the reminding components of $\hat{\mathbf{s}}(t)$ include mECG and noises. Accordingly, the corresponding columns of the mixing matrix are stored in $\hat{\mathbf{A}}_f(t) = [\hat{\mathbf{a}}_{f1}, \cdots, \hat{\mathbf{a}}_{fM}]$. As a result, the contribution of the fetal signals in the observation signals can be obtained by:

$$\hat{\mathbf{x}}_f(t) = \hat{\mathbf{A}}_f(t)\hat{\mathbf{s}}_f(t) \tag{13}$$

in which $\hat{\mathbf{x}}_f(t)$ is the extracted fECG signal in the original domain. A known drawback of conventional ICA is that they cannot preserve the order, sign and amplitude of the sources [28]. Therefore, for automatic applications, reliable source type detection and block-wise sign/amplitude correction is required to identify and correct the fECG sources among the other extracted components. In practice, due to the rather structured morphology of the ECG, the significant amplitude of the mECG compared to the fECG and accessible of prior information about the mECG, it is easier to identify mECG rather than fECG. Therefore, in this work, we detect mECG signals using some channel assessment criteria.

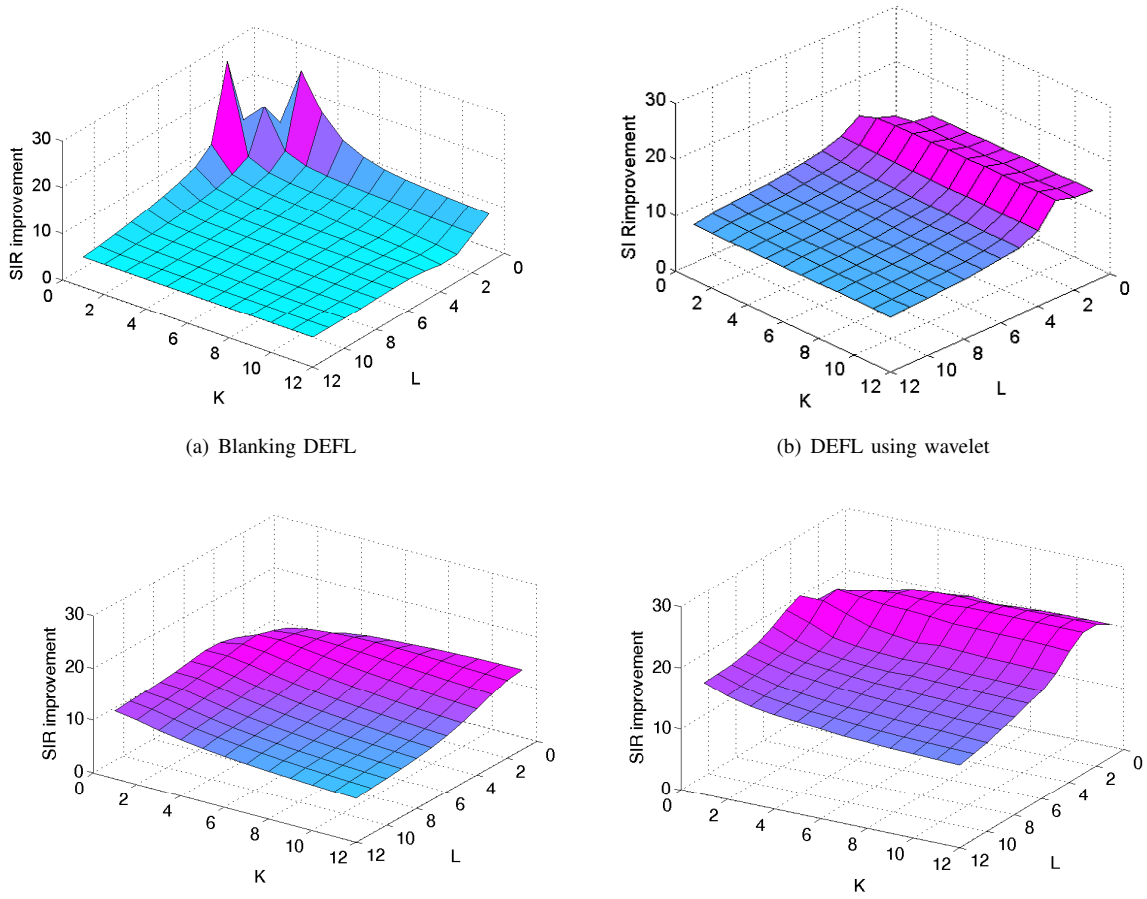
D. Synthetic ECG Generation

For quantitative analysis, maternal and fetal ECG mixtures are generated using a realistic model adopted from the *open-source electrophysiological toolbox* (OSET) [18], [29]:

$$\begin{aligned}\mathbf{x}(t) &= \mathbf{H}_m(t)\mathbf{s}_m(t) + \mathbf{H}_f(t)\mathbf{s}_f(t) + \mathbf{H}_\eta(t)\mathbf{v}(t) + \mathbf{n}(t) \\ &\triangleq \mathbf{x}_m(t) + \mathbf{x}_f(t) + \boldsymbol{\eta}(t) + \mathbf{n}(t)\end{aligned} \tag{14}$$

This model is based on the single dipole model of the heart, which assumes three geometrically orthogonal lead pairs, known as the Frank lead electrodes (or the *vectorcardiogram* (VCG)), and a linear propagation media for the body volume conductor to map the three dimensions to body-surface potentials, using a Dower-like transformation [30]. Although the single dipole model is only an approximation of the true cardiac activity [31], the model was found to be accurate enough for the hereby presented study.

Based on this model, we generated a three dimensional $\mathbf{s}_m(t)$ and $\mathbf{s}_f(t)$ representing the ECG signal of maternal and fetus heart, using a $3D$ VCG. The ECG sources are then mapped to 12 body surface channels using the $\mathbf{H}_m(t)$ and $\mathbf{H}_f(t)$ matrices, which model the propagation media. As a result, both maternal and fetal ECG are distributed in all the body surface ECG channels, but with only three underlying dimensions. A realistic full rank noise with desired SNR is also added to the signal using the idea proposed in [18]. Accordingly, 10000 samples of twelve lead synthetic maternal/fetal ECG mixtures were generated with a sampling rate of 500Hz. Following (1), consider $\mathbf{x}(t)$ as the noisy input



(a) Blanking DEFL

(b) DEFL using wavelet

(c) DEFL using Template subtraction

(d) DEFL using Kalman

Fig. 3. Sensitivity of the SIR improvement versus K and L parameters in four denoising strategies

observations, $\mathbf{x}_f(t)$ as the fECG signal, $\mathbf{x}_m(t)$ as interference and $\boldsymbol{\eta}(t) + \mathbf{n}(t)$ as noise for the fECG. The overall *signal to interference plus noise ratio* (SINR), defined in [13], can be used to quantify the data quality *before denoising*:

$$\text{SINR} \triangleq \frac{E_{t,i}\left\{x_{f_i}(t)^2\right\}}{E_{t,i}\left\{\left[x_{m_i}(t) + \eta_i(t) + n_i(t)\right]^2\right\}} \tag{15}$$

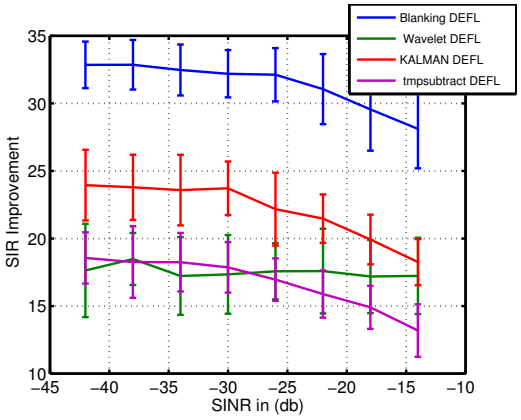
where $x_{f_i}(t)$, $x_{m_i}(t)$, $\eta_i(t)$, and $n_i(t)$ are the i-th element of the vectors, $\mathbf{x}_f(t)$, $\mathbf{x}_m(t)$, $\boldsymbol{\eta}(t)$, and $\mathbf{n}(t)$, respectively. For synthetic data, the SINR can be set to arbitrary ratios by scaling the mixing matrices $\mathbf{H}_m(t)$, $\mathbf{H}_f(t)$, $\mathbf{H}_\eta(t)$ and the noise variances in (1) by appropriate factors (cf. [13] for further details). After applying the denoising procedure, in order to evaluate the effectiveness of mECG cancellation algorithm, the *signal-to-interference ratio* (SIR), *similarity measure* (SM) and *periodicity measure* (PM) [13] are used:

$$\begin{aligned}
\text{SIR} &\triangleq 10\log\left(\frac{E\left\{\mathbf{x}_s(t)^2\right\}}{E\left\{\hat{\mathbf{x}}_m(t)^2\right\}}\right)\\
\text{PM} &\triangleq \left|\frac{\text{tr}\left(E\left\{\mathbf{y}(t)\mathbf{y}(t+\tau_t)^T\right\}\right)}{\text{tr}\left(E\left\{\mathbf{y}(t)\mathbf{y}(t)^T\right\}\right)}\right| \times 100\\
\text{SM} &\triangleq \frac{\left|E\left\{\mathbf{y}(t)\mathbf{x}_s(t)^T\right\}\right|}{\sqrt{E\left\{\mathbf{y}(t)^2\right\}E\left\{\mathbf{x}_s(t)^2\right\}}}
\end{aligned} \tag{16}$$

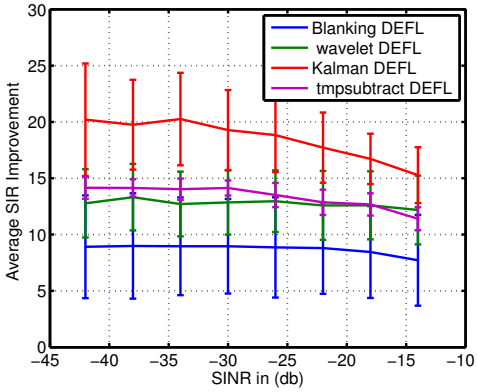
where $\text{tr}(\cdot)$ represents matrix trace and $\mathbf{x}_s(t) \triangleq \mathbf{x}_f(t) + \boldsymbol{\eta}(t) + \mathbf{n}(t)$, is the summation of all non-mECG components, which we call the *mECG complement* and is produced during synthetic signal generation. $\mathbf{y}(t)$ denotes the denoised signal and $\hat{\mathbf{x}}_m(t)$ denotes the mECG (noise) residue in the output. Since the aim of this method is to remove mECG, $\mathbf{y}(t)$ should be equal to $\mathbf{x}_s(t)$, in an ideal mECG canceller. Therefore, the residual mECG artifact within the signal is:

$$\hat{\mathbf{x}}_m(t) = \mathbf{y}(t) - \mathbf{x}_s(t) \tag{17}$$

In the later presented results, *SIR improvement* is defined as the output SIR minus the input SIR in dB. Therefore, SIR improvement is a measure of mECG cancellation in dB. The PM, measures the amount of periodicity of denoised data



(a) SIR improvement (for best parameter selection)



(b) SIR improvement (Average over all parameters in range)

Fig. 4. SIR improvement versus SINR in four denoising strategies

according to the period of the reference mECG. By definition, $0 \leq \text{PM} \leq 100$ (PM $= 0$ for fully aperiodic signals and PM $= 100$ for a fully periodic signal). This measure indicates the amount of mECG components that still exists in the output. It should be noted that the reduction of PM is only a necessary (but not sufficient) measure for the algorithm success, since the PM might decrease due to an increase of noise or at a cost of losing the fECG. Therefore, a compliment measure is required, which assures the fidelity of the remaining components. For this, we measure the correlation coefficient (SM) between the denoised data and the original signal components, $\mathbf{x}_s(t)$, by definition $0 \leq \text{SM} \leq 1$. A SM value close to 1 indicates that the algorithm has preserved the non-mECG components (including the fECG) in its output.

VII. Parameter Estimation

All algorithms used for comparison have parameters that require optimization, which is described in this section.

A. Extended Kalman Filter Parameters

For estimating the parameters of the Gaussian kernels used in the extended Kalman filter, the ensemble average of the mECG are extracted as a single beat average template. Next, the parameters are estimated by applying a nonlinear least squares error algorithm to fit the ECG template, using the open-source packages of OSET [29]. The other parameters and covariance matrices are initialized following the methods developed in [10].

B. ANC Parameters

The standard and the more recent multi-stage ANC are implemented using a 5-tap FIR filter (20ms window length at a 250Hz sampling frequency) with a step size equal to $\mu = 1e^{-6}$. Both parameters were found as the optimal values, by searching over a grid of possible values in varying SINR. The maternal ECG reference, required for ANC, is selected directly from $\mathbf{x}_m(t)$ in equation (14) during the generation of synthetic data. Since $\mathbf{x}_m(t)$ is a pure mECG without other noise and interferences, each of its channels can play the role of the chest lead ECG required as reference.

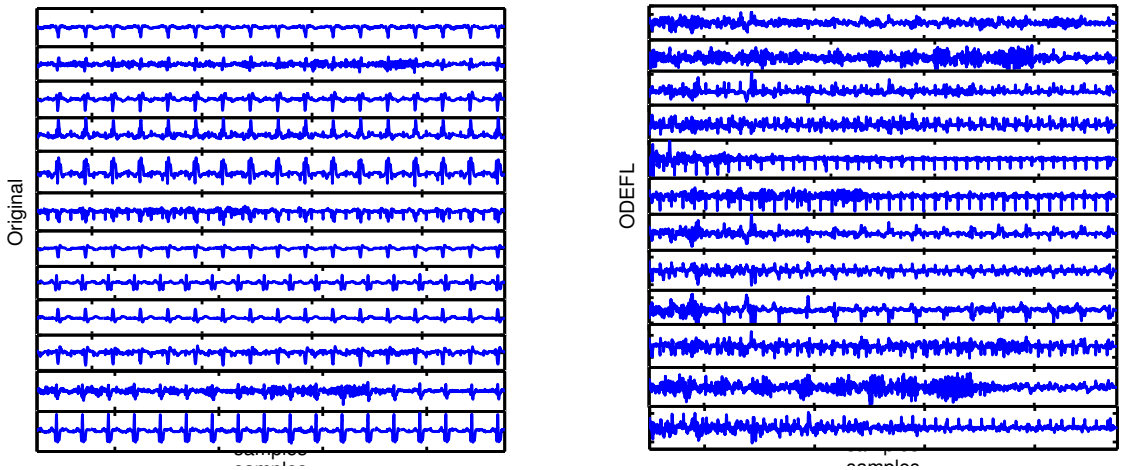
C. Wavelet Parameter Estimation

Following the results in [10], the *Coiflets3* mother wavelet with 6 levels of decomposition, using the *Stein's unbiased risk estimate* (SURE) shrinkage rule, single level rescaling and a soft thresholding strategy were used as the optimal denoising setup of the wavelet-based ECG denoiser (cf. [10] for a detailed discussion).

D. DEFL and ODEFL Parameters

The optimum number of iterations, K, the number of channels to be denoised in each iteration, L, and the strategy used for denoising are critical (and application-dependent) issues that highly influence the performance of DEFL and ODEFL. The parameter K, provides the capability of eliminating full-rank and possibly nonlinearly superposed noise, which is beyond the capabilities of conventional ICA techniques. The parameter L, may be considered as the *effective number of dimensions* of the signal and noise subspaces. The denoising function $\mathbf{G}(\cdot,\cdot)$, used for signal and noise subspace separation, also influences the overall performance of both DEFL and ODEFL. In practice, all of these parameters should be tuned according to the application. In this work, a Monte Carlo simulation was carried out to investigate the sensitivity of DEFL and ODEFL algorithms, with respect to the denoising function and the values of L and K. The performance was investigated using 700 simulated data, generated according to the scheme in Section VI-D, in different input SINRs, in the range of -35dB to -5dB in 5dB steps. Fig. 3 shows the average SIR improvements versus K and L using four denoising strategies $\mathbf{G}(\cdot,\cdot)$. In the first strategy, which we call *blanking DEFL*, the first L channels of $\mathbf{s}(t)$ are simply set to zero (similar to hard-thresholding in the wavelet denoising). In the second strategy, wavelet denoising was used as the denoiser using the optimal parameters explained in Section VII-C. In the third strategy, the single-channel extended Kalman filtering scheme proposed in [9], [10] is used as the denoiser. In the fourth strategy, the single-channel template subtraction technique proposed in [7] is used as the denoiser.

The results of optimizing the parameters of all methods are shown in Fig. 4 from different viewpoints. In Fig. 4(a), the



(a) Sample ECG data (b) After mECG removal

Fig. 5. A sample of synthetic ECG with $-20db$ SINR and $20sec$ length (a) before and (b) after mECG removal. The peaks remaining after mECG removal are the fECG plus background noise.

SIR improvement versus different SINRs is calculated for the best values of K and L parameters. In Fig. 4(b), the average SIR improvement over the average of the whole values of K and L in the range of studied parameters is calculated versus different SINRs.

According to Figs. 3 and 4(a), by setting appropriate values for L and K, blanking DEFL has better performance as compared to wavelet, template subtraction and Kalman denoising strategies, which is due to the fact that when the signal space dimensions are obtained, the algorithm completely removes all the noise space dimensions preserving the signal unchanged. In practice, the appropriate value of K can be estimated using some stopping criterion such as the PM criterion. The suitable value of L can also be calculated using related methods for estimating the signal/noise dimensionality [32], [33]. For nonstationary data, K and L can also be updated in time[2]. According to Fig. 4, although blanking DEFL performs best for the suitable parameters, it is sensitive to the proper choice of K and L and its performance highly degrades for inappropriate parameters. On the other hand, the wavelet, template subtraction and Kalman denoising strategies are more robust to the choice of parameters, since increasing K and L beyond their optimal values does not significantly degrade the SIR improvements. As a result, using denoising methods such as wavelet, template subtraction or Kalman filter in DEFL, instead of banking DEFL are more appropriate in practice. From Fig. 4 it is also seen that the Kalman filter outperforms the template subtraction and wavelet denoiser in SIR improvement and robustness to its parameters. This result was anticipated, as the Kalman filter is a model-based approach, which benefits from prior knowledge of the signal. Besides, as compared to the template subtraction method, the Kalman filter performance relies on both the model and the observation, which makes it effectively adaptive to different SNR scenarios. The necessity of a signal model is at the same time a limitation of this method in practice as compared to the non model-based methods. Yet, for simplicity, in what follows the first denoising strategy (blanking the first L components) with $K = 1$ and $L = 3$ are used for evaluation of both DEFL and ODEFL algorithms. The other parameters of ODEFL are the forgetting factors β and γ. These factors should be chosen according to the degree of data (non-)stationary within the range $[0, 1]$. In the studied database, the ECG signal and noise were both stationary. Hence, we chose $\beta = \gamma = 1$, i.e., the algorithm does not forget the old samples.

[2]According to our empirical results, for ECG signals, the update should be done over long temporal windows (tens of seconds) rather than short windows, otherwise the performance degrades.

E. ICA Denoising Parameters

The free parameter in ICA denoising is the number of mECG components (effective mECG dimension) that should be removed after the source separation stage. For synthetic data, according to our prior knowledge about generating three dimensional $\mathbf{s}_m(t)$ for the maternal heart, we choose $L = 3$. For real data, it is also chosen as $L = 3$ that is experimentally found as the best parameter for this dataset and eliminates the most dominant components of the mECG. In general, the number of mECG channels can be adaptively obtained during the denoising process by morphological similarity, or by using the notion of *effective number of dimensions* [34]. In this work, the mECG identification for both real and synthetic data is accomplished by computing the correlation coefficient defined in (16) between the maternal reference signal (chest lead ECG) and the different source channels extracted by ICA. The top L channels with the highest correlations are selected as the mECG components. These channels are set to zero and the remaining channels are back-projected to the original subspace. This strategy is rather similar to a single stage of the DEFL algorithm.

VIII. Results and Evaluation

The proposed algorithm has been evaluated on both real and synthetic data and compared with the block-wise DEFL

[13], single-channel ECG Kalman Filter [8]–[10], standard ANC [3], the modified ANC [5], template subtraction [7], ICA denoising [11] and the wavelet denoiser applied to each channel separately. The detailed results are discussed in this section.

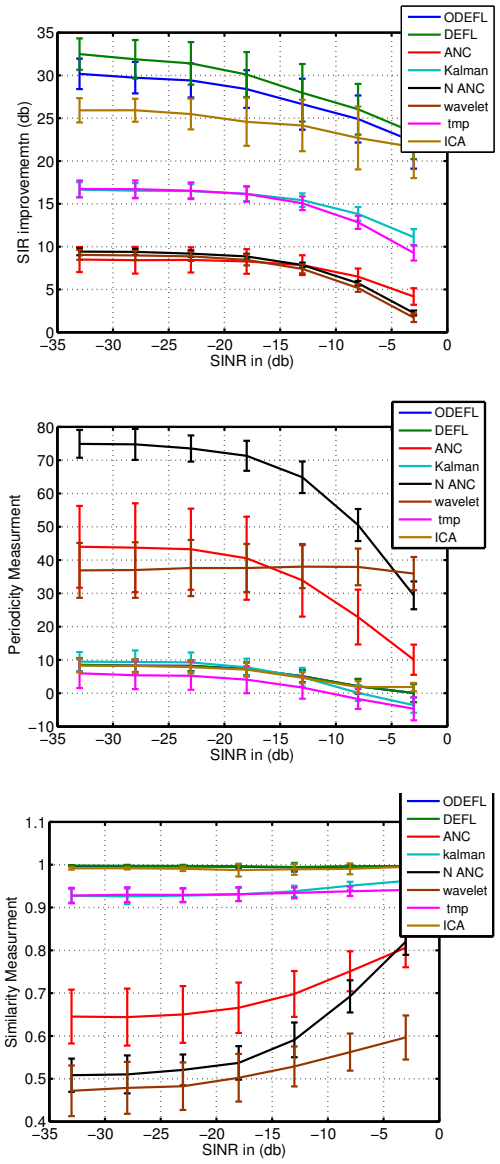


Fig. 6. (Top) SIR improvement, (Middle) PM and (Bottom) SM versus input SINRs. The PM and SM of DEFL and ODEFL have overlapped.

A. Simulated Data

The simulated data mode was discussed in Section VI-D. For visual inspection, a typical 20s length synthetic ECG with SINR of -20dB, along with the corresponding denoised output is shown in Fig. 5. One can see that the maternal ECG is distributed in all the simulated channels. The denoised output indicates that the maternal ECG is removed in almost all channels without affecting the fetal ECG.

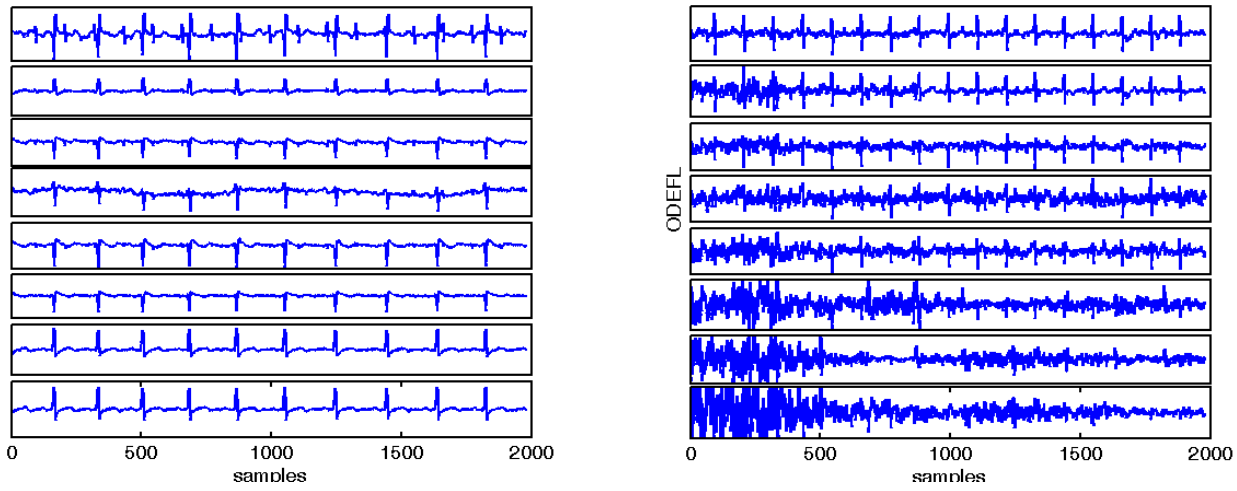
For a quantitative evaluation, the proposed algorithm was compared with other benchmark methods using 1000 different ensembles of simulated data and noise, in different input SINRs. The average and standard deviation of SIR improvements, PM, and SM are shown in Fig. 6.

Accordingly, DEFL outperforms all methods and is only slightly better than the ODEFL. The outperformance of DEFL as compared to ODEFL is reasonable due to the offline and exact calculation of the covariance matrices used in DEFL. However, the difference is negligible as compared to the advantages of ODEFL for online and nonstationary applications. As shown, DEFL and ODEFL, which are based on prior knowledge of the ECG periodicity have outperformed ICA. This is due to the fact that DEFL and ODEFL can deal with situation that ICA assumptions are not satisfied. In fact, ICA algorithms despite their vast and effective applications have some intrinsic ambiguities according to their simplified assumptions. Typically, it is assumed that the number of independent sources is fixed and equal to the number of sensors. The signal mixture is considered instantaneous and time-invariant. However, these assumptions are not always satisfied in practice. As a result, the performance of ICA degrades in presence of full-rank Gaussian noise, correlated and/or distributed sources [15], resulting in residual mECG within the fECG. Moreover, the ranking property of DEFL and ODEFL, in comparison to permutation ambiguity of the ICA, helps the reliable and automatic detection of fECG/mECG signal in long recordings [15], while for ICA it is necessary to have robust source type identification methods that recognize the mECG components among others. Therefore, ICA denoising is very sensitive to the correct identification of mECG in the projected domain. Overall, DEFL, ODEFL and ICA denoising outperform all other methods in both low and high SNR scenarios. This can be due to the fact that the ANC, wavelet, template subtraction and Kalman filtering schemes are all single-channel, while DEFL, ODEFL and ICA benefit from the spatial information within multiple channels to obtain higher SNR.

Among the single-channel methods, the performance of Kalman filter and template subtraction is similar in low SNR scenarios and outperforms other single-channel methods, while in high SNR Kalman filter has superior performance. The reason is that the Kalman filter relies on both the observation and dynamic model, depending on the signal quality. When the data is too noisy, the Kalman filter tracks the prior dynamic model rather than relying on the observation. Therefore, in low SNR, Kalman filter performance is identical to template subtraction. However, in high SNR the Kalman filter benefits from the information within the observations, making it better than template subtraction. The low performance of ANC, as mentioned before, can be explained due to the fact that the reference signal used in ANC (here the chest lead mECG) does not necessarily resemble the morphology of the mECG superposed over the abdominal leads, which significantly downgrades ANC's performance.

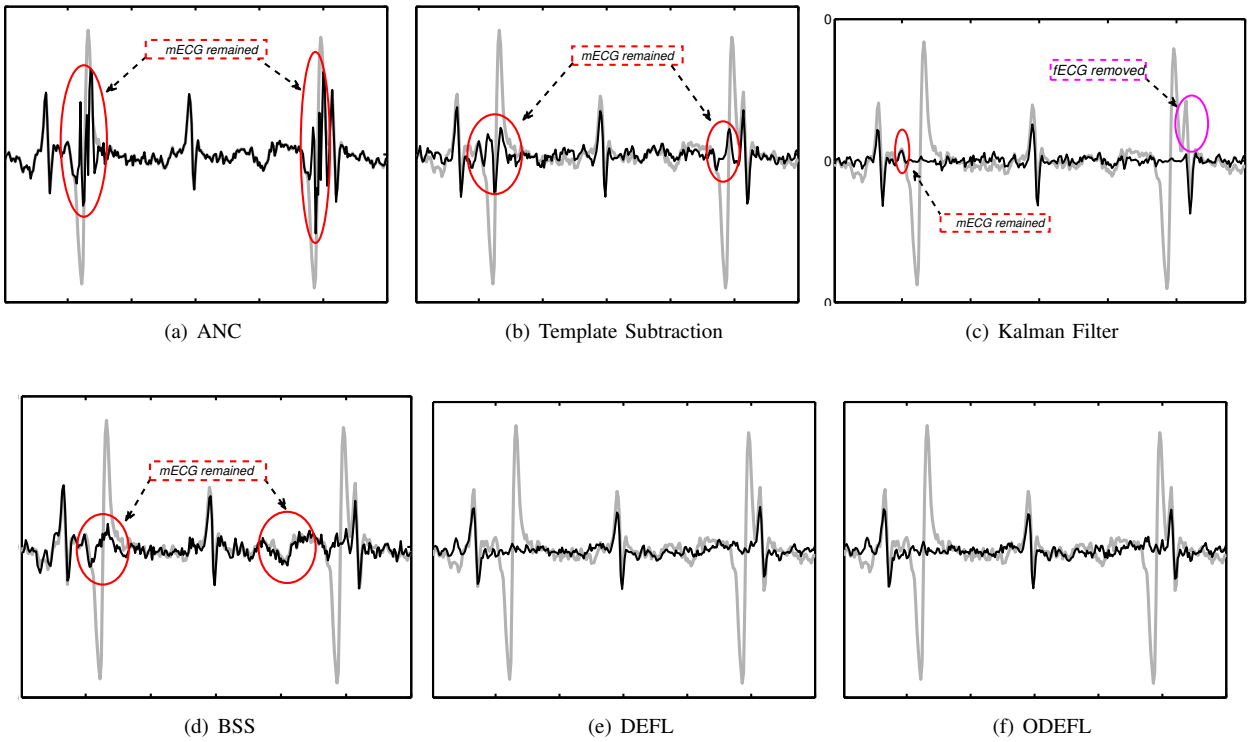
B. Real Data

The proposed algorithm was also applied to the widely used DaISy fECG dataset, shown in Fig. 7 [35]. This dataset consists of five abdominal and three thoracic channels, recorded



(a) DaISy dataset (b) After mECG removal

Fig. 7. DaISy dataset (a) before and (b) after mECG removal. The peaks remaining after mECG removal are the fECG plus background noise. Due to the sequential structure of ODEFL, the algorithm converges slower in the last channels.



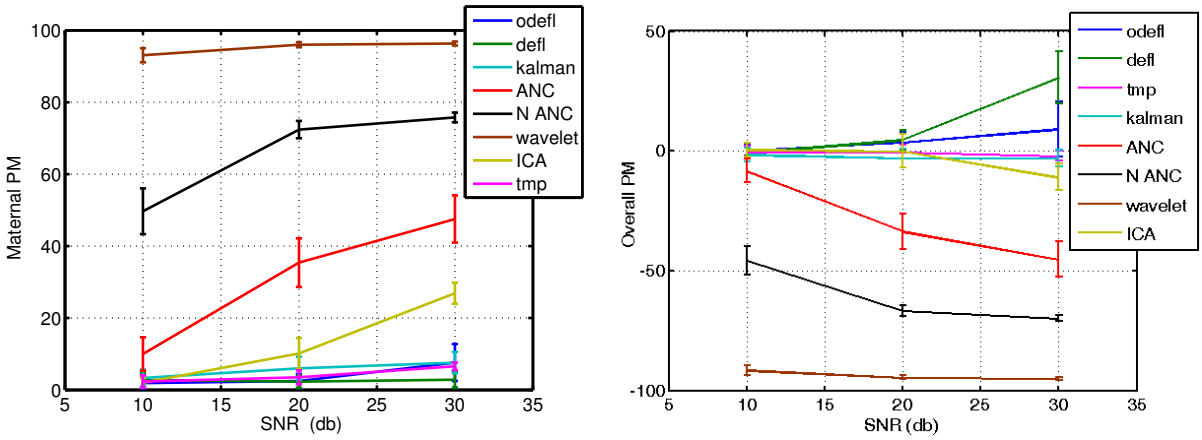
(a) ANC (b) Template Subtraction (c) Kalman Filter

(d) BSS (e) DEFL (f) ODEFL

Fig. 8. A typical data segment before (gray plots) and after (black plots) mECG cancellation. It is observed that the mECG is completely removed in DEFL and ODEFL methods with minimal effect on the fECG.

from the abdomen and chest of a pregnant woman, with a sampling rate of 250Hz. Due to the rather stationary nature of the dataset we selected $\beta = \gamma = 1$. The results of ODEFL on this dataset are shown in Fig. 7. It is seen that after about 40 samples (160ms), the algorithm has converged and the mECG is almost completely removed in the first channel; but it takes up to 500 samples (2s) for all channels to converge. This is due to the sequential nature of the proposed ODEFL algorithm. Fig. 8 shows a closer view of the results over two ECG beats. It is seen that DEFL and ODEFL outperformed the ANC, template subtraction, Kalman filter and ICA denoising. While DEFL and ODEFL have effectively removed the mECG, other methods have left some residual mECG or removed parts of the fECG. For numerical evaluation of the proposed method on real data, we altered the real DaISy abdominal signals as follows [13]:

$$\mathbf{x}(t) = \mathbf{G}[\mathbf{x}_0(t) + \mathbf{\Lambda}\mathbf{v}(t)] \tag{18}$$



(a) Maternal PM

(b) Overal PM

Fig. 9. Output Maternal PM and Overal PM versus SNR of added noise

where $\mathbf{x}_0(t)$ is the original real data in Fig. 7, $\mathbf{v}(t)$ is Gaussian white noise, $\mathbf{\Lambda} = diag(\lambda_1, \cdots, \lambda_N)$ is a diagonal matrix which controls the SNR per channel, $\mathbf{G} \in R^{N\times N}$ is an arbitrary random matrix and $\mathbf{x}(t)$ is the new noisy signal. $\mathbf{x}(t)$ is generated in three different input SNR of 30dB, 20dB and 10dB by changing the entries of $\mathbf{\Lambda}$. The proposed methods are then applied to $\mathbf{x}(t)$ by selecting $L = 3$ and $K = 2$. The algorithm is repeated over 1000 trials using different instances of $\mathbf{v}(t)$ and $\mathbf{G}$ in each trial. The PM was defined in (16) as a measure of algorithm performance in mECG cancellation. But we noted that it should be studied together with fECG preserving indeces, to assert the overall algorithm performance. For this we define the *overall periodicity measure* (OPM):

$$\text{OPM} = \text{fPM} - \text{mPM} \tag{19}$$

where mPM and fPM are maternal and fetal PM, respectively. Accordingly, $-100 \leq$ OPM ≤ 100, where higher values of OPM are indication of algorithm success in simultaneously removing the mECG and preserving the fECG. The average and standard deviation of the mPM and OPM are shown in Fig. 9 for the proposed and benchmark methods. We can see that the results on real data follow the same trend and order as the synthetic data results. The only exception is the ICA denoiser, which has inferior results for real data. This might be due to the fact that for real noisy data, mECG identification and estimation of L is difficult, resulting in degrading performance.

IX. Conclusion

In this paper, an online version of an iterative subspace denoising procedure was presented for removing maternal ECG from biosignals recorded from the abdomen of a pregnant woman. The proposed method is rather generic and may be applied to other blind and semi-blind source separation applications, in which the signal and noise mixtures are full-rank. It was shown that the proposed method outperforms the state of the art single channel denoising techniques, while it marginally performs as good as its offline version. Interestingly, it was shown that DEFL and ODEFL algorithms which is obtained by GEVD of only two second-order matrices, outperformed the classical ICA that uses more than two matrices containing the higher-order statistics of the observations. The outperformance can be related to the fact that DEFL and ODEFL can deal with situation in which the underlying assumptions of ICA are not satisfied. Moreover, DEFL and ODEFL benefit from the ranking property of the GEVD for mECG detection, while ICA suffers from permutation and sign ambiguities, which require the utilization of a robust source type identification method for mECG recognition in conjunction with the basic ICA. As a result, the proposed method is less complicated and more reliable for long datasets, as compared with batch ICA techniques.

The performance of ODEFL was investigated with different set of parameters using different denoising strategies including simple blanking, wavelet denoising, template subtraction and Kalman denoising.

According to these results and the experiments carried out in [10], we conclude that for single channel data the Kalman filter outperforms other ECG denoising schemes in different SINR scenarios, while the DEFL and ODEFL algorithms are better for multichannel data as they use the inter-channel correlations without requiring the mixing matrix of the data. Therefore, in future studies, the combination of the Kalman denoiser and ODEFL may result in superior results. Introducing an online method for automatic calculation of the algorithm parameters L, K, β and γ is also an interesting extension to the current work, which was partially studied in [34].

The performance of ODEFL is influenced by several aspects including the method used for online GEVD. In future studies, other online GEVD algorithms can be compared with the incremental common spatial pattern used in this work. Moreover, theoretical aspects of online GEVD and the convergence of DEFL and ODEFL should also be considered. A symmetric extension of the method for avoiding the problems of sequential source separation and error propagation is also interesting for practical applications.

REFERENCES

[1] R. Sameni and G. D. Clifford, "A Review of Fetal ECG Signal Processing; Issues and Promising Directions," *The Open Pacing, Electrophysiology & Therapy Journal (TOPETJ)*, vol. 3, pp. 4–20, November 2010.

[2] P. Bergveld and W. J. H. Meijer, "A New Technique for the Suppression of the MECG," *Biomedical Engineering, IEEE Transactions on*, vol. BME-28, no. 4, pp. 348–354, April 1981.

[3] B. Widrow, J. Glover, J. McCool, J. Kaunitz, C. Williams, H. Hearn, J. Zeidler, E. Dong, and R. Goodlin, "Adaptive Noise Cancelling: Principles and Applications," *Proc. IEEE*, vol. 63, no. 12, pp. 1692–1716, 1975.

[4] P. Strobach, K. Abraham-Fuchs, and W. Harer, "Event-synchronous cancellation of the heart interference in biomedical signals," *Biomedical Engineering, IEEE Transactions on*, vol. 41, no. 4, pp. 343 –350, april 1994.

[5] R. Swarnalath and D. V. Prasad, "A Novel Technique for Extraction of FECG using Multi Stage Adaptive Filtering ," *Journal of Applied Sciences*, vol. 10, no. 4, pp. 319–324, 2010.

[6] M. Ungureanu and W. Wolf, "Basic aspects concerning the event-synchronous interference canceller," *Biomedical Engineering, IEEE Transactions on*, vol. 53, no. 11, pp. 2240 –2247, nov. 2006.

[7] S. M. M. Martens, C. Rabotti, M. Mischi, and R. J. Sluijter, "A robust fetal ECG detection method for abdominal recordings," *Physiol Meas*, vol. 28, no. 4, pp. 373–388, Apr 2007. [Online]. Available: http://dx.doi.org/10.1088/0967-3334/28/4/004

[8] R. Sameni, "Extraction of Fetal Cardiac Signals from an Array of Maternal Abdominal Recordings," Ph.D. dissertation, Sharif University of Technology – Institut National Polytechnique de Grenoble, July 2008. [Online]. Available: http://www.sameni.info/Publications/Thesis/PhDThesis.pdf

[9] R. Sameni, M. B. Shamsollahi, and C. Jutten, "Model-based Bayesian filtering of cardiac contaminants from biomedical recordings," *Physiological Measurement*, vol. 29, no. 5, pp. 595–613, May 2008.

[10] R. Sameni, M. B. Shamsollahi, C. Jutten, and G. D. Clifford, "A Nonlinear Bayesian Filtering Framework for ECG Denoising," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 12, pp. 2172–2185, December 2007.

[11] V. Zarzoso and A. Nandi, "Noninvasive fetal electrocardiogram extraction: blind separation versus adaptive noise cancellation," *Biomedical Engineering, IEEE Transactions on*, vol. 48, no. 1, pp. 12–18, January 2001.

[12] D. Graupe, Y. Zhong, and M. H. Graupe, "Extracting fetal from maternal ecg for early diagnosis: theoretical problems and solutions - baf and ica," in *Proceedings of the fifth IASTED International Conference: biomedical engineering*, ser. BIEN '07. Anaheim, CA, USA: ACTA Press, 2007, pp. 352–356. [Online]. Available: http://dl.acm.org/citation.cfm?id=1295494.1295558

[13] R. Sameni, C. Jutten, and M. B. Shamsollahi, "A Deflation Procedure for Subspace Decomposition," *IEEE Transactions on Signal Processing*, vol. 58, no. 4, pp. 2363–2374, April 2010.

[14] R. Sameni, C. Jutten, M. Shamsollahi, and G. Clifford, "Extraction of Fetal Cardiac Signals," U.S. Patent US 2010/0 137 727 A1, June 3, 2010.

[15] M. Fatemi, M. Niknazar, and R. Sameni, "A robust framework for noninvasive extraction of fetal electrocardiogram signals," *Computing in Cardiology Conference (CinC), 2013*, pp. 201–204, Sept 2013.

[16] Q. Zhao, L. Zhang, A. Cichocki, and J. Li, "Incremental Common Spatial Pattern algorithm for BCI," in *Neural Networks, 2008. IJCNN 2008. (IEEE World Congress on Computational Intelligence). IEEE International Joint Conference on*, june 2008, pp. 2656 –2659.

[17] D. B. Geselowitz, "On the Theory of the Electrocardiogram," *Proc. IEEE*, vol. 77, pp. 857–876, Jun. 1989.

[18] R. Sameni, G. D. Clifford, C. Jutten, and M. B. Shamsollahi, "Multichannel ECG and Noise Modeling: Application to Maternal and Fetal ECG Signals," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, pp. Article ID 43 407, 14 pages, 2007, ISSN 1687-6172, doi:10.1155/2007/43407. [Online]. Available: http://www.hindawi.com/GetArticle.aspx?doi=10.1155/2007/43407

[19] L. Amini, R. Sameni, C. Jutten, G. Hossein-Zadeh, and H. Soltanian-Zadeh, "MR Artifact Reduction in the Simultaneous Acquisition of EEG and fMRI of Epileptic Patients," in *EUSIPCO2008 - 16th European Signal Processing Conf.*, Lausanne, Switzerland, August 25-29 2008.

[20] C. Gouy-Pailler, R. Sameni, M. Congedo, and C. Jutten, "Iterative Subspace Decomposition for Ocular Artifact Removal from EEG Recordings," in *Proc. of the 8th Intl. Conf. on Independent Component (ICA 2009)*, Paraty, Brazil, 2009, pp. 419–426.

[21] R. Sameni, C. Jutten, and M. B. Shamsollahi, "Multichannel Electrocardiogram Decomposition using Periodic Component Analysis," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 8, pp. 1935–1940, Aug 2008.

[22] B. Yang, "Projection approximation subspace tracking," *Signal Processing, IEEE Transactions on*, vol. 43, no. 1, pp. 95 –107, jan 1995.

[23] J. Yang, H. Xi, F. Yang, and Y. Zhao, "Rls-based adaptive algorithms for generalized eigen-decomposition," *Signal Processing, IEEE Transactions on*, vol. 54, no. 4, pp. 1177 – 1188, april 2006.

[24] P. E. McSharry, G. D. Clifford, L. Tarassenko, and L. A. Smith, "A Dynamic Model for Generating Synthetic Electrocardiogram Signals," *IEEE Trans. Biomed. Eng.*, vol. 50, pp. 289–294, mar 2003.

[25] S. Haykin, *Adaptive filter theory*, ser. Prentice-Hall information and system sciences series. Prentice Hall, 1996.

[26] J.-F. Cardoso, "Multidimensional independent component analysis," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '98)*, vol. 4, May 1998, pp. 1941–1944.

[27] L. de Lathauwer, B. de Moor, and J. Vandewalle, "Fetal electrocardiogram extraction by blind source subspace separation," *Biomedical Engineering, IEEE Transactions on*, vol. 47, no. 5, pp. 567 –572, may 2000.

[28] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. Wiley-Interscience, 2001.

[29] R. Sameni, *The Open-Source Electrophysiological Toolbox (OSET), version 2.1*, 2010. [Online]. Available: http://www.oset.ir

[30] L. Edenbrandt and O. Pahlm, "Vectorcardiogram synthesized from a 12-lead ECG: Superiority of the inverse Dower matrix," *J. Electrocardiol.*, vol. 21, p. 361, 1988.

[31] R. Sameni, C. Jutten, and M. B. Shamsollahi, "What ICA Provides for ECG Processing: Application to Noninvasive Fetal ECG Extraction," in *Proc. of the International Symposium on Signal Processing and Information Technology (ISSPIT'06)*, Vancouver, Canada, August 2006, pp. 656–661.

[32] R. Nadakuditi and A. Edelman, "Sample eigenvalue based detection of high-dimensional signals in white noise using relatively few samples," *Signal Processing, IEEE Transactions on*, vol. 56, no. 7, pp. 2625 –2638, july 2008.

[33] J. A. Lee and M. Verleysen, *Nonlinear Dimensionality Reduction*. Springer Science, 2007.

[34] R. Sameni and C. Gouy-Pailler, "An Iterative Subspace Denoising Algorithm for Removing Electroencephalogram Ocular Artifacts," *Journal of Neuroscience Methods*, vol. 225, no. 3, pp. 97–105, March 2014.

[35] B. De Moor, *Database for the Identification of Systems (DaISy)*. [Online]. Available: http://homes.esat.kuleuven.be/~smc/daisy/