Fetal Magnetocardiographic Signals Extracted by 'Signal Subspace' Blind Source Separation

Carlo Salustri, Giulia Barbati, and Camillo Porcaro

Abstract—In this paper, we apply independent component analysis to fetal magnetocardiographic data. In particular, we propose an extension of the "cumulant-based iterative inversion" algorithm in order to achieve a two-step "signal subspace" subdivision, which allows the user to control the number of components to be estimated by analyzing the eigenvalues distribution in an interactive way. Our results show that this method is a powerful means not only for the extraction of the cardiac signals from the background noise but also for a sharp separation of the baby’s heart from the mother’s.

Index Terms—Blind source separation, independent component analysis, magnetocardiography, signal subspace.

I. INTRODUCTION

The electric and magnetic signals generated by the fetal heart are always embedded in a background noise originating from electronic interference and from a multitude of electronic and biological sources (respiration, gastric activity, muscle contractions), which hide a major portion of the cardiac wave complex, often making waveform analysis quite difficult if not impossible. Moreover, due to the variability of the fetus’ position with respect to the sensors, the fetal heart wave is sometimes barely distinguishable from the mother’s. For all these reasons, the extraction of the fetal heart signal from the background noise and its complete separation from the mother’s signal are prerequisites for an assessment of the fetus’ well being. Recently, some authors have explored the feasibility of applying independent component analysis (ICA) to data recorded by electrocardiographic electrodes applied on the mother’s abdomen [1], [2]. In such configuration, the application of ICA may face limitations due to problems inherently related to tissue conductivity, electrode efficiency, and other factors that may cause the signal mixture at the sensors to be “noninstantaneous.” For this reason, we describe here an ICA performed on the magnetic fields recorded by an array of magnetic sensors distributed over the abdomen of pregnant women. We expected the quality of the analysis to be enhanced by the fact that the magnetic fields generated by the mentioned biological sources are virtually not distorted by the presence of human tissues in the frequency range of interest [3]. In particular, we propose in this paper, an extension of the “cumulant based iterative inversion” algorithm [4] to include a “signal subspace” subdivision, which allows the user to control the number of components to be estimated by analyzing the eigenvalues distribution in an “interactive” way. We will refer to our extended algorithm as cumulant-based iterative inversion in signal subspace (CISS) throughout the paper.

II. METHOD

Recordings were assumed to be representable as \( x(t) = As(t) + n(t) \), i.e., a linear mixture of statistically independent sources plus additive Gaussian noise. The first step of our procedure consisted in “whitening” the data. Generally, whitened data \( \tilde{x} \) are obtained by singular value decomposition of the original data’s covariance matrix \( R_{xx} = E[xx^T] \), i.e.,

\[
\text{svd}(R_{xx}) \rightarrow \Lambda^{-\frac{1}{2}}V^T x = Q x = \tilde{x}
\]

where the \((M \times M)\) matrix \( V \) contains the eigenvectors associated with the eigenvalues of

\[
\Lambda = \text{diag}\{\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_M\}
\]

in descending order. The whitening matrix \( Q \) is such that \( R_{\tilde{x}\tilde{x}} = E[\tilde{x}\tilde{x}^T] = I_M \), where the subscript \( M \) indicates the order of \( I \).

Instead of following this approach, we preferred to perform a dimensionality reduction by seeking a separation between signal and noise subspaces; this procedure has been shown to be applicable to cases in which the noise covariance matrix can be modeled as \( R_{nn} = \sigma_n^2 I_M \) and the noise variance is relatively small (i.e., the signal-to-noise ratio (SNR) is high above some threshold) [5]. We judged that our data matched these conditions since both maternal and fetal heartbeats appeared generally much stronger than the noise in our MEG recordings. One assumes that the first \( K \) of the \( M \) eigenvalues in \( \Lambda \) form a \( K \)-dimensional signal subspace, whereas the remaining \((M-K)\) define a \((M-K)\)-dimensional noise subspace. If one takes the mean value of the \((M-K)\) “minor” eigenvalues as representing the noise variance \( \sigma_n^2 \), then the \((K \times M)\) “quasi-whitening” matrix can be computed as

\[
\tilde{Q} = \hat{A}_S^{-\frac{1}{2}}V_S^T = (A_S - \sigma_n^2 I_K)^{-\frac{1}{2}}V_S^T
\]

and the new \( K \)-dimensional data vector can then be expressed as \( \tilde{x} = \tilde{Q}x \). The matrix \( \tilde{Q} \) is such that \( R_{\tilde{x}\tilde{x}} = E[\tilde{x}\tilde{x}^T] \equiv I_K \), where again the subscript \( K \) indicates the order of \( I \). Note that now \( R_{\tilde{x}\tilde{x}} \) does not equal exactly the identity matrix, but this does not disturb our procedure since the ICA algorithm we have subsequently used does not require an orthogonal mixing matrix [4].

The key point of this procedure is of course the determination of \( K \). Several ways have been proposed in the literature: Cichocki and Amari suggest finding empirically the gap in the eigenvalues distribution of the data covariance matrix, in order to separate signal and noise subspaces [5]. They also consider the alternative use of one of two information theory criteria, namely Akaike’s information criterion (AIC) or the minimum description length criterion (MDL) [6]. Beside the fact that both these procedures present relevant computational problems, we decided not to use them because they have been proven to provide estimates of the number of sources that are rather rough and very sensitive to SNR variations and especially to the number of data samples available [7]. Another problem with both AIC and MDL is that they were derived assuming that the data vectors \( x(t) \) have a strict Gaussian distribution: usually, the Gaussianity assumption does not hold exactly in the BSS context, even though data vectors are mixtures of the sources and therefore often have quasi-Gaussian distributions. Recently, “a rank detection criterion” has been proposed, that evaluates the above-mentioned gap by means of numerical analysis procedures, to provide a “maximally stable” decomposition of a data covariance matrix into signal and noise subspaces [7]. The authors show that such criterion is insensitive to variations in the SNR and the number of data samples, overcoming the major drawbacks of AIC and MDL; moreover, they show that their criterion may indicate whether a classification into signal and noise subspaces is stable and well-conditioned or unstable and ill-conditioned. However,

Manuscript received August 5, 2003; revised November 20, 2004. *Corresponding author.

G. Barbati and C. Porcaro are with the Center of Medical Statistics and Information Technology, AfaR, San Giovanni Calibita—Fatebenefratelli Hospital, Isola Tiberina, 00186 Rome, Italy (e-mail: giulia.barbati@afar.it; camillo.porcaro@afar.it).

Digital Object Identifier 10.1109/TBME.2005.846708
they applied this criterion in microwave radio applications; the fetal heartbeat varies much with respect to the maternal one and we verified that this criterion was too restrictive for our data, never producing signal subspace dimensions of more than four components. It must be stressed that when looking for a gap separating the signal and the noise subspaces, one comes across another gap inside the signal subspace, due to the energy difference between maternal and fetal signals (this happened in the vast majority of our cases). Consequently, we thought that, in applying the “rank detection criterion,” we would have run the risk of either seeing the fetal component mixed together with the noise one, or mixed with one of the mother’s, which is precisely the opposite of what we wanted to do.

Because of all these considerations, we used the following “interactive” scheme for the determination of $K$, which takes into account the particular characteristics of the data available and includes a “control cycle” on subsequent independent components (IC).

**Step a)** After the $M$ eigenvalues are delivered in decreasing order $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_M$, where $M$ corresponds to the number of sensors, each eigenvalue $\lambda_j$ was divided by the order of magnitude $10^i$ of the greatest eigenvalue $\lambda_1$ (for example, if $\lambda_1 = 98.765.4321$, then every $\lambda_j$ was divided by $10^i$). In this way we obtained a new set of eigenvalues $\tilde{\lambda}_j = \lambda_j / 10^i$.

**Step b)** Each $\tilde{\lambda}_j$ was rounded up to the $i$th decimal digit and we checked whether and which two eigenvalues were identical.

- **Step b.1)** If $\tilde{\lambda}_j = \tilde{\lambda}_{j+1}$, then the signal subspace dimension was $K = j - 1$ and the round up parameter remained $i$. Quasi-whitened data were $K$-dimensional, and we proceeded to the ICA phase.
- **Step b.1.1)** If the ICA algorithm converged, then we judged the “quality of separation” by visual inspection (quality was defined satisfactory when maternal and fetal heartbeats resulted positively represented by different components).
- **Step b.1.1.1)** If the quality was satisfactory, the process ended.
- **Step b.1.1.2)** If the quality was not satisfactory, then we set $i = i - 1$ and went back to Step b.
- **Step b.2)** If the ICA algorithm did not converge, then we set $i = i - 1$ and went back to Step b.
- **Step b.2.1)** If there were no identical eigenvalues, we set $i = i - 1$ and went back to Step b.

This procedure differs from other criteria mainly in the fact that it contains an empirical control cycle (the user moves through the described steps on the basis of his/her judgment of quality and performance); attempts to develop automatic mechanisms for the determination of $K$ are currently under way. One could object at this point that defining the quasi-whitening matrix by (3), which leads to express the new $K$-dimensional data vector as $\tilde{\mathbf{x}} = \mathbf{Q}_{k+} \mathbf{x}$, may actually amplify the noise rather than suppress it, when the mixing matrix is ill-conditioned, i.e. some of the signal subspace eigenvalues $\lambda_k, \lambda_{k+1}, \ldots$ are very small; this is indeed a difficult problem since the mixing matrix is by definition unknown, and one cannot really know when it is ill-conditioned. In practical applications, one deals with the eigenvalues distribution of the data covariance matrix, not with the one of the mixing matrix. In some cases, a “regularization” can be applied to overcome this difficulty [5] but, in the cases we have examined, the estimated noise was systematically lower than the smallest eigenvalue of the estimated signal subspace and we did not feel the need for a “regularization” approach. Consequently, in our case, we have verified in all our data that the difference between $\lambda_k$ and $\lambda_{k+1}$ was big enough to be reasonably sure about a separation of the cardiac signal from the noise. Most of all, we took advantage of the fact that a model was available to us (i.e., we knew we were looking for signals of cardiac nature), which allowed us to check the output of the quasi-whitening process. However, in cases where no model is available, the proposed procedure might be too risky.

### III. RESULTS

Magnetocardiographic data were recorded from twelve pregnant women participating in our hospital’s pregnancy assistance program, using a 25-sensor array placed over the mother’s abdomen. Details and features of the sensor system are described elsewhere [8]. Data were acquired continuously for 4 min at 250 Hz sampling rate with a 0.48–64 Hz bandpass prefiltering.

In all 12 recordings, the ICA algorithm extracted independent components that could be clearly associated to cardiac sources. We used spectral coherence between extracted components and recorded traces to safely discriminate maternal and fetal components. Although signals generated by the mothers’ hearts generally dominated the recorded MCG traces, amplitudes showed a variety of fetus/mother ratios from case to case, most likely due to the different fetal positions with respect to the sensors array. Fig. 1 shows the results of our procedure in two representative cases: case A, in which the fetal cardiac amplitude was...
comparable to the maternal one, and case B, in which the fetal signal was barely visible in the recorded traces. For both cases, a time section of the highest ranking extracted ICs representing the mother’s heart and the IC representing the fetus’ heart are superimposed to the same time period of the MCG signal.

In case A, performing the quasi-whitening phase with the initial 25-dimensional data matrix produced no identical eigenvalues. By recursively applying the above-described dimensionality reduction procedure, we selected nine independent components, which accounted for 99.71% of the total variance. As Fig. 1 shows, in case A, the fetus’ relatively high-amplitude QRS complex allowed an easy eye identification of the independent components representing both the mother’s and the fetus’ cardiac sources. Notice how the rather deep S deflection of the maternal cardiac signal is represented by a separate independent component: it must be remembered that, because of the stationarity assumption, more than one independent component may describe the different stages of the ventricular process. In case B, instead, 11 eigenvalues, accounting for 99.96% of total variance, were selected by the quasi-whitening phase.

As an example of comparison between our $K$ determination procedure and the “rank detection criterion,” we show in Fig. 2 the traces representing the independent components extracted by the two procedures in case A. Fig. 2(a) shows that out of the nine ICs delivered by our procedure, three clearly represented the maternal (IC3 and IC7) and the fetal (IC6) heartbeats; Fig. 2(b) shows instead that maternal and fetal heartbeats resulted mixed together in IC3 (and also in IC4) when $K$ was determined by the rank detection criterion.

IV. CONCLUSION

In this paper, we have assumed the fetus’ and the mother’s hearts as statistically independent sources of magnetic fields: under this assumption, we have proposed the use of a two-step BSS procedure which, by defining two signal subspaces (signal and noise) and identifying a number of independent components, extracted the cardiac signals from the background noise and effectively separated the fetus’ from the mother’s traces. We have shown the effectiveness of this method even in cases where the fetal signal is almost invisible to the naked eye in the MCG trace.

REFERENCES