Separation of Physiological Signals Using Minimum Norm Projection Operators

James D. Wilson and Jens Haueisen, Member, IEEE

Abstract—Objective: This paper presents the development of a fast and robust method which can be applied to multichannel physiologic signals for the purpose of either removing a selected interfering signal or separating signals that arise from temporally correlated and spatially distributed signals such as maternal or fetal cardiac waveform recordings. Methods: Projection operators based upon both the weighted and un-weighted minimum norm equations are presented. The weighted formulation uses models based on signal covariance and the un-weighted formulation requires that a statistical model be built using time-locked averaging. Results: We present examples that demonstrate the utility of our projection operators when applied to maternal and fetal magneto-cardiograms. In addition, we demonstrate the ability to separate fetal breathing signals from both maternal and fetal cardiac signals. Conclusion: The method is effective, robust, fast, and does not require significant input from a user. Significance: Although we demonstrate the utility of our projection operators applied to biomagnetic signals, the method can easily be adapted to other applications where the goal is to either separate or suppress selected signal components.

I. INTRODUCTION

ACTIVE current sources in the body produce measurable magnetic fields and electric potentials. Such recordings may be the result of several spatially distinct sources arising from the various organs such as the heart, brain, and muscle. It is common that a number of spatially distinct sources associated with an organ are temporally correlated. The adult human heart is the most obvious example in which the P wave and QRS complex arise from tissue depolarization in spatially distinct regions while the T wave arises from repolarization, yet all are temporally correlated. It is often desirable to select one signal to study while discarding others, or it may be desirable to suppress a single large interfering signal to reveal weaker signals that are of interest. In practice, a range of methods may be applied, sometimes in sequence, to achieve the desired result [1]–[20]. Our aim here is to separate the maternal magneto cardiogram (mMCG), the fetal magneto cardiogram (fMCG), and the fetal breathing signal. Our methods are based upon the minimum norm approach that is commonly used to estimate the strength of spatially distributed current dipole [21]–[23].

We first summarize the minimum norm method as presented in [22]. Then, we introduce our new approach, which uses the minimum norm idea and applies it to time series at the sensor level. The foundational assumption is that the magnetic fields produced by a large number of current dipoles can be represented by a set of lead fields associated with a set of surrogate current sources as presented in the following matrix formulation:

\[ X = LQ. \] (1)

After including the system noise, the observed data is given by

\[ Y = X + n. \] (2)

\( Y \) is the observed data and each row of \( Y \) is the time series from one sensor, \( X \) is the modeled signals arising from current dipoles within the body, and \( n \) describes the system noise. \( X, Y, \) and \( n \) are size \( N_c \) (number of channels) by \( N_t \) (number of time samples). \( Q \) has dimension \( N_q \) (number of sources) by \( N_t \) and \( L \) has dimension \( N_q \) by \( N_c \). The matrix \( Q \) describes the time series of the biologically active sources that give rise to the observed signals and each column of the matrix \( L \) is the lead field associated with each current source. We only consider the case where the set of equations defined by (1) and (2) is underdetermined.

We take the covariance to be estimated by the expectation operator. Then,

\[ C_X = E \{ XX^T \} = \sum_{q=1}^{N_q} \{ E \{ QQ^T \} \} L^T = LC_QL^T, \] (3)

and assuming independence between noise and currents gives

\[ C_Y = LC_QL^T + C_n. \] (4)

Continuing with the formulation as shown in [22], the weighted minimum norm estimate for the current sources is

\[ \hat{C} = \begin{bmatrix} \hat{Q} \\ \hat{n} \end{bmatrix} = \begin{bmatrix} C_QL^T \\ C_n \end{bmatrix} \begin{bmatrix} LC_QL^T + C_n \end{bmatrix}^{-1} Y. \] (5)

If no estimate of the source covariance is available, it is common practice to use zero order Tikhonov regularization [22],...
[23] so that (5) reduces to

\[ \hat{Q} = L^T(LL^T + C_l)^{-1}Y. \]  

(6)

We now outline the basic concept of our new approach. As an alternative to the model of (1), one can think of using time-averaged observations at the sensor level, or templates that arise from the activity of a group of physiologically linked current sources like those found in the heart to describe the signals themselves. We refer to the averaged data as the statistical model produced by a group of coordinated sources. We will derive our projection operators to work on the sensor level data only. Neither the number of sources, nor their spatial separation, nor the lead fields are considered. We propose replacing (1) with a model-based system of equations such that

\[ X = MS \]  

(7)

where \( M \) is a model that is derived from the time locked averages of various signals of interest and \( S \) is the set of explanatory signals associated with each component of the model. \( M \) is analogous to \( L \) in that it is an embodiment of the spatial information in the signals, and \( S \) is analogous to \( Q \) in that \( S \) is the temporal embodiment of the signals. We note that both (1) and (7) are assumed approximations to a more complex reality. Starting with (7), we then apply the minimum norm solutions from (5) or (6) to derive our projection operators.

We develop projection operators that either pass a selected signal, for example the fetal heart, or suppress a selected signal, for example the maternal heart. While cardiac signals are excellent signals for this method, the method is by no means limited to cardiac signals.

II. THEORY

A. Model Approach

Equations (8) through (14) set up the basis for our method. We also neglect noise during the development of the projection operators. Let the noiseless signal \( X \) be made up of the summation signals arising from currents \( Q \) distributed throughout the body. Further, there will be grouped current sources, \( Q_1 \) (with associated \( L_1 \)), which are both spatially and temporally associated. They are best thought of as a collection of sources with various orientations distributed throughout a limited volume, but all working in a semisynchronous pattern. Then, \( X \) may be written as the sum of the signals from all signal groups so that

\[ X = X_1 + X_2 + X_3 + X_4 + \ldots \]  

(8)

Using (7), we assume that a good estimate of \( X \) is \( \hat{X} \) such that

\[ \hat{X} = \hat{M}\hat{S} \]  

(9)

where \( \hat{S} \) is computed using the unbiased minimum norm estimate from (6). In our example application, typically \( \hat{M} \) is size \([150 \times 10^4]\), \( \hat{S} \) is size \([150 \times 700]\), and \( \hat{S} \) is size \([700 \times 10^4]\). The meaning of \( \hat{S} \) will be made clear with the introduction of (12) below.

The matrix \( M \) is the statistical model of the time-varying observations and is obtained by averaging the original data based on an easily found repetitive event such as the maximum point of the maternal or fetal R wave (explained in more detail later). The matrix \( M \) is written as

\[ M = [m_1 \ldots m_{j-1} m_j \ldots m_{k-1} m_k \ldots m_{n}] \]  

(10)

where each \( m \) is a column vector \([N_c \times 1]\) that is found by computing the time-locked average of the signal \( X \) at the same point in a repeating waveform. For example, the column vectors \([m_1 \ldots m_{j-1}]\) might describe the averaged maternal cardiac cycle in all channels spanning a time from just before the maternal P wave until just after the maternal T wave. Likewise, \([m_j \ldots m_{k-1}]\) could describe the fetal cardiac cycle and \([m_k \ldots m_n]\) might be used to describe fetal breathing. We call each subgroup of \( M \) a statistical model for that physiological signal and \( M \) can be thought of as an array of concatenated templates. There are \( N_m \) columns in \( M \), and \( N_m \) is typically much larger than the number of channels in the system so that (7) is an underdetermined set of equations. In our data, \( MM^T \) is not full rank and requires regularization to invert. A three-component model as described in (10) can be seen in Fig. 14. Given \( M \) and the assumed form of (7), the explanatory matrix \( \hat{S} \) is computed using the minimum norm solution defined by (6)

\[ \hat{S} = (M^T)(MM^T + \lambda I)^{-1}Y. \]  

(11)

To get some sense of the meaning of \( S \), we ignore noise and, then, use (1) to give

\[ \hat{S} = (M^T)(MM^T + \lambda I)^{-1}LQ. \]  

(12)

The explanatory matrix, \( S \), is a scaled, linear mixture of the source activity \( Q \). In general, any one source will contribute to the signal in multiple rows of \( S \), and at any instant, there are probably multiple sources active. However, a group of spatial–temporal current sources, such as the cardiac sources, will primarily give rise to activity in a corresponding group of explanatory signals. Partitioning the matrix \( M \) into three sub-matrices as suggested by (10) gives the maternal cardiac cycle \( M_M = [m_1 \ldots m_{j-1}] \), the fetal cardiac cycle \( M_F = [m_j \ldots m_{k-1}] \), and fetal breathing signal \( M_B = [m_k \ldots m_n] \). Given \( M \) that is partitioned accordingly, there are three corresponding sub-matrices of \( \hat{S} \); \( \hat{S}_M \) is the first “j” rows of \( \hat{S} \), \( \hat{S}_F \) is a row “j” through “k − 1,” and \( \hat{S}_B \) is row “k” through “m.” Then,

\[ \hat{X} = M_M \hat{S}_M + M_F \hat{S}_F + M_B \hat{S}_B. \]  

(13)

Equation (13) is a key to group-level signal separation. As seen in (13), the sub-matrices of \( \hat{S} \) align with the grouped columns of \( M \) that create them through (11). The sub-matrix \( \hat{S}_M \) is the subgroup of explanatory signals associated with maternal cardiac activity. Likewise for \( \hat{S}_F \) and \( \hat{S}_B \).

However, we can eliminate \( \hat{S} \) altogether. Using (7), (11), and (13), we can write

\[ \hat{X}_i = M_i M_i^T (MM^T + \lambda I)^{-1}Y \]  

(14)

where \( \hat{X}_i \) is the estimate of the “ith” subgroup signal corresponding to the “ith” sub-model of \( M \) from (13). Although the minimum norm solution of (11) is the minimum norm fit for all signals, we find that in practice subgroups are well separated if
the signals are adequately modeled. Equation (14) defines the projection operator for the \(i\)th signal subgroup. Thus,

\[
P_i = M_i M_i^T (MM^T + \lambda I)^{-1}.
\]

Ideally, the projection operator \(P_i\) passes the \(X_i\) signal subgroup with near unity gain and attenuates all other signals. We also note that only the noise in \(Y\) that is aligned with the vectors of \(X_i\) can pass through the operator without attenuation. Because there are several variants to consider, we write our projector in general terms as shown below:

\[
P = B(A + I)^{-1}.
\]  

\(P\) is identified as needed to distinguish between the variants and the makeup of \(A\) and \(B\) define the variant of the operator.

### B. Projection Operators

#### 1) Null Projection Operator:

If \(PY\) isolates one or more signals from all others with good fidelity and unity gain, then a null projector can be computed that suppresses the isolated signal. The null projector, \(P_n\), is defined by

\[
P_n = I - P
\]

where \(I\) is the identity matrix. We point out that the signal rejection may not be perfectly complete and in some cases, the rejection is deliberately limited so that the spatial redistribution of the remaining signals is reduced.

In data such as MCG recordings, the maternal signal dominates all fetal signals, so that initially the only signal available for forming a model is the mMCG. For suppression of the mMCG, \(B = M_M M_M^T\) and \(A = B\). Using (15) and (17) gives

\[
P_n = I - (M_M M_M^T)(M_M M_M^T + I)^{-1}.
\]

Using our data, \(M_M M_M^T\) is singular and requires regularization. However, the principal consideration for the choice of \(\lambda\) is to suppress the action of undesired eigenvectors. For a practical explanation using the actual data, see Section IV.

#### 2) Orthogonal Projector:

The orthogonal projector, \(OP\), is a type of null projector and is described in [5]. It is presented here because of similarities with \(P_n\) and because we will compare the results obtained with both projectors. To compute \(OP\), one starts with the time-averaged signal of the maternal cardiac signal that we have already defined as \(M_M\). The time point where \(M_M\) has the largest amplitude is found and the data at that time point are taken as a column vector, \(v\). A null projector is computed using \((I - v(v^Tv)^{-1}v^T)\) and is then applied to the signal to annihilate the largest component in \(M_M\). \(M_M\) is recomputed using the projected data and the process is repeated until the maternal signal is suppressed to an acceptable level. The sequential process was shown in [5] to be equivalent to forming a single projection operator using a matrix, \(V\), such that the columns of \(V\) are the individual column vectors, \(v\), found at each step as described above. Then, a single projection vector, \(OP\), may be computed as follows:

\[
OP = I - V(V^TV)^{-1}V^T.
\]  

Applying the \(OP\) operator to the signal \(Y\) will suppress the major components of the maternal cardiac signal, leaving a small residual mixed with fetal signals. \(OP\) uses a set of linear independent vectors, typically 5 to 10 in number, to build the projector and there is no need to regularize the inverse in (19). On the other hand, \(P_n\) of (18) is built with several hundred vectors that are linearly dependent by nature. Although regularization of the inverse is required, the primary considerations in selecting \(\lambda\) are redistribution and the chosen level of mMCG suppression. The level of mMCG suppression for \(OP\) depends upon the number of iterations (the number columns vectors in \(V\)). In contrast, we will show that the level of mMCG suppression for \(P_n\) is continuously varied by the parameter \(\lambda\).

#### 3) Model Projectors:

The aim of these projectors is the extraction of the fetal heart signal or the fetal breathing signal. To find a model for the fetal heart or fetal breathing, one must first remove the maternal heart signals using either \(P_n\) or \(OP\). Once the maternal cardiac signal is projected out and the fetal R waves are found, \(M_F\) is computed and added to the model. If the fetus presents a significant fetal breathing signal and if it is possible to find the time-locked average of the fetal breathing signal over some window, then \(M_B\) can be added to the model. Then, let \(A = MM^T\), where \(M\) is the concatenation of \(M_M\), \(M_F\), and, depending on the data, \(M_B\). Inclusion of \(M_B\) is desirable since that will generally improve the performance of model-based projectors.

To form a projector that will extract the fMCG signal, let \(B = M_F M_F^T\). Using (15), we define the fetal cardiac projector, \(P_F\), as

\[
P_F = (M_F M_F^T)(MM^T + \lambda I)^{-1}.
\]

Likewise, we define the fetal breathing projector by letting \(B = M_B M_B^T\)

\[
P_B = (M_B M_B^T)(MM^T + \lambda I)^{-1}.
\]

Again, the choice of \(\lambda\) used in (20) and (21) is data dependent and is discussed in Section IV. It must be noted that both \(P_n\) and \(OP\) cause spatial redistribution of signals that can corrupt the fetal models and a scheme for reducing this effect is also presented in Section IV.

#### 4) Covariance Projectors:

Combining (1) and (5) gives

\[
\hat{X} = LCQ_L^T(LCQ_L^T + C_n)^{-1}Y.
\]

Using the definitions of (3) and (4),

\[
\hat{X} = C_X C_Y^{-1}Y.
\]

Since the subgroup signals in (2) are expected to be uncorrelated, the cross terms in the calculation of \(C_X\) can be ignored. Then,

\[
C_X = C_1 + C_2 + C_3 + C_4 + \ldots
\]

so that, in general, we can write

\[
\hat{X}_i = C_i C_Y^{-1}Y.
\]

Since \(C_Y\) is just the covariance of the data, \(EYY^T\), we only need to estimate \(C_i\), the covariance of the signal of interest in order to compute a projection operator. We assume that the
maternal and fMCG signals are sufficiently stationary so that we use the statistical model as a basis to estimate $C_i$ so that

$$C_i = E \{M_i M_i^T \}. \quad (26)$$

Alternatively, one may be able to isolate a signal in the frequency domain, and if so, $C_i$ can be computed using the bandwidth limited signal. Fetal breathing is a candidate signal for this technique. However, $C_i$ is estimated, and the projection operator defined in (27) is not bandwidth limited. Please consider that (26) may overestimate the amplitude of the covariance of discontinuous signals compared to the actual signal covariance in $C_Y$. Fetal breathing is an intermittent signal as opposed to $D = P$. In practice, $A + \lambda I$, so that

$$\mu_i = \lambda/(\lambda_i + \lambda). \quad (32)$$

Since each eigenvector is normalized, each $\mu_i$ is the gain associated with that eigenvector. Consequently, $\mu_i$ is the gain of the signal component that gave rise to that eigenvector. Equation (32) shows that for the larger maternal signals where $\lambda_i >> \lambda$, then $\mu_i \approx \lambda/\lambda_i$, so that the maternal signals that align with these eigenvectors are attenuated by the amount $\lambda/\lambda_i$. In practice, we choose $\lambda$ so that the nonmMCG eigenvalues fall below $\lambda$ which implies $\lambda_i << \lambda$, making $\mu_i \approx 1$. As suggested by (17), $P_n$ is similar to the identity matrix except that it is designed to suppress the mMCG eigenvectors while passing other signals.

In our data, the trace of $A$ is typically $\sim 10\%$ larger than the largest eigenvalue associated with the mMCG $R$ wave. We can utilize $t_\lambda$ as an easily computed approximate value for the largest eigenvalue, $d_{m\text{ax}}$. We now define a dimensionless scalar, $sf$, as follows:

$$\lambda = sf \cdot t_\lambda \quad (33)$$

where $t_\lambda$ is the trace of matrix $A$. From (33), we see that $sf \gg d_{m\text{ax}}$, so that $sf$ is the gain (attenuation) of the dominant mMCG signal. However, $sf$ is not the attenuation of the whole mMCG waveform since each eigenvector has a different gain according to (32).

Since the eigenvalues of $P_n$ are derived from the mMCG model, we assume, as a first approximation, that each eigenvector has an associated signal $x_i = \sigma_i \cdot u_i(t)$, where $u_i(t)$ is the normalized signal and $\sigma_i$ is the RMS amplitude of the signal. It then follows that the power of that component gives rise to the eigenvalue, so that $d_i = \sigma_i^2$. To estimate the RMS output for each component, we substitute $d_i = \sigma_i^2$ into (32) to estimate the gain of that signal and, then, multiply by the input $x_i$ to get the estimated output signal $y_i$:

$$y_i = \lambda x_i/\sigma_i^2 + \lambda = (\lambda \sigma_i/(\sigma_i^2 + \lambda)) \cdot u_i(t). \quad (34)$$

Setting $\sigma_i = \lambda^{1/2}$ maximizes (34), so that the RMS of all residual components should be less than or equal to $1/2 \lambda^{1/2}$. If we let $\sigma_i = \lambda^{1/2}$, and combine (33) and (34), we have the maximum mMCG residual component given by

$$y_{m\text{ax}} = (t_\lambda \times sf^{1/2}/4)^{1/2}. \quad (35)$$

The total residual mMCG is the sum of all residual components, but (34) and (35) suggests that the residual mMCG $\propto sf^{1/2}$. The proper selection of $sf$ is application dependent and details will be presented in Section IV.

We now consider the general projector described by (16), when $B \neq A$. First, we note that the most general model $M$ can have more than one component, for instance fMCG and fetal breathing models may be added. If so, then there should be some subset of the eigenvectors of $A$ associated with those models. Similar to (29), we write $B$ in terms of an eigenvalue

$$P_n = V(I+\lambda I)^{-1}V^T. \quad (31)$$

The null projector can be expressed in terms of the eigenvectors of $A$, which is derived from the model, but has modified eigenvalues given by the diagonal matrix $\lambda(D + \lambda I)^{-1}$. Let $\lambda_i$ be an eigenvalue of the diagonal matrix $D$ and let $\mu_i$ be the corresponding eigenvalue of $P_n$, then

$$\mu_i = \lambda/(\lambda_i + \lambda). \quad (32)$$

where the asterisk indicates the removal of the diagonal elements. Clearly, $R_e$ goes to zero as $P_n$ becomes more like the identity matrix, which causes no redistribution at all.

D. Regularization

We first consider the inverse term in (16): $(A + \lambda I)^{-1}$. $A$ is symmetric by definition because $A = M M^T$, and with minimal regularization we have found that $A$ is a positive definite. Under that condition, we can decompose $A$ into its eigenvectors, $V$, and associated eigenvalues that lie on the diagonal of matrix $D$, so that

$$A = V D V^T. \quad (29)$$

It is straightforward to show that

$$(A + I)^{-1} = V(D + \lambda I)^{-1}V^T. \quad (30)$$

Since the null projector defines $B = A$, we can write a general expression for the null projector in terms of eigenvectors and eigenvalues by combining (18), (29), and (30). After some basic simplification we have

$$P_n = V(I+\lambda I)^{-1}V^T. \quad (31)$$

The null projector can be expressed in terms of the eigenvectors of $A$, which is derived from the model, but has modified eigenvalues given by the diagonal matrix $\lambda(D + \lambda I)^{-1}$. Let $\lambda_i$ be an eigenvalue of the diagonal matrix $D$ and let $\mu_i$ be the corresponding eigenvalue of $P_n$, then

$$\mu_i = \lambda/(\lambda_i + \lambda). \quad (32)$$

where the asterisk indicates the removal of the diagonal elements. Clearly, $R_e$ goes to zero as $P_n$ becomes more like the identity matrix, which causes no redistribution at all.
matrix $G$ and eigenvectors $W$. Then,

$$\mathbf{B} = \mathbf{WGW}^T. \quad (36)$$

Using (16), (29), (30), and (36), the projection operator is expressed in terms of eigenvectors and eigenvalues

$$\mathbf{P} = (\mathbf{WGW}^T) \left( \mathbf{V(D + \lambda I)^{-1}V}^T \right). \quad (37)$$

We define $g_i$ as the eigenvalues of $G$. Using (30), we write an expression for the eigenvalues of the second term shown in parentheses. We define $\theta_i$ as the eigenvalue of the inverse term. Then,

$$\theta_i = 1/(d_i + \lambda). \quad (38)$$

For a projector designed to pass the fMCG, we would typically set $\lambda$ well below the eigenvalues associated with the major fMCG components in (38). Given that the model $\mathbf{M}_F$ is the basis of $\mathbf{B}$ and that $\mathbf{M}_F$ is included as part of $\mathbf{A}$, we assume that some eigenvectors and eigenvalues will be common to both (or at least similar). Given the condition that $d_i >> \lambda$ and there are joint eigenvectors, then the projector’s gain is $g_i / d_i \approx 1$ for each eigenvector common to $\mathbf{A}$ and $\mathbf{B}$. All other signals are suppressed by the selectivity of $\mathbf{B}$ (lack of a matching eigenvector) and the scaling from (38). The values of $s_f$ used for $\mathbf{P}_A$ and $\mathbf{P}_R$ range from $10^{-4}$ to $10^{-6}$ and must be determined by trial and error for each type of application.

The covariance formulation of (27) has two factors that are similar to the two factors in (37). $C_i$ is functionally associated with $\mathbf{B}$ and the inverse of $C_Y$ is functionally associated with $(\mathbf{A} + \lambda \mathbf{I})^{-1}$. As we will show in Section IV, our system noise essentially sets the minimum value for $\lambda$. Since the fetal signals are larger than the system noise, there is no need to regularize the inverse of $C_Y$.

III. DATA COLLECTION

Data presented in this paper were collected using the SARA system, a 151-channel magnetoencephalogram (MEG) system based on SQUID technology [24], [25]. In the SARA system, the magnetic sensors are radial gradiometers with a 2-cm-diameter coil and 8 cm coil separation (base line) [3], [6]. The instrument is installed in a magnetically shielded room (Vakuumschmelze Hanau, Germany), to reduce the effects of environmental noise. The gradiometers cover the whole maternal abdomen and capture maternal and fMCG, fetal breathing, and fetal brain activity. The patients sat in upright position during the recording. All signals were bandpass filtered with an eighth order, zero-phase, filter having a passband of 1 to 50 Hz unless stated otherwise. The sample rate was 312.5 samples per second. A total of 113 recordings of 10 min duration were taken from normal fetuses ranging in gestation age from 26 weeks to 38 weeks. Data collection protocols were approved by the IRB at University of Arkansas for Medical Sciences.

IV. RESULTS AND DISCUSSION

A. Performance of OP and $\mathbf{P}_n$

To calculate either OP or $\mathbf{P}_n$, $\mathbf{M}_M$ must first be computed. Since the mMCG dominates all other signals, it is easy to find the maternal R markers [27]. The median R-to-R time interval is taken as the duration of $\mathbf{M}_M$ (or window). The raw signal is time averaged using the maternal R markers so that 40% of the time window precedes each R marker and 60% follows. This basic procedure captures information about the P and T waves and the QRS complex and is used to compute either mMCG or fMCG cardiac models. The averaged mMCG, $\mathbf{M}_M$, is then used to compute either the $\mathbf{P}_n$ or OP null projector as specified by (18) or (19), respectively.

After projecting out the mMCG using either OP or $\mathbf{P}_n$, some residual mMCG will be leftover. Time averaging null projected data based on the maternal R markers will produce the averaged residual mMCG. We use the largest magnitude found in the averaged residual mMCG to judge the performance of a null projector. The same measure is normally used to either stop the OP nulling algorithm of (19) or to choose $\lambda$ (or $s_f$) in (18).

To demonstrate the general properties of the null projectors, a 10-min recording taken from a 36-week-old fetus was processed using the OP and $\mathbf{P}_n$ projectors and the results are shown in Fig. 1. Using the OP method of (19), null vectors were sequentially placed in the signal space until the magnitude of the residual mMCG fell below 0.1 pT (our typical stopping threshold). Seven null vectors were needed to satisfy the stopping threshold for the OP method. As presented in the theory section, the value of $s_f$ sets the degree of mMCG suppression of a $\mathbf{P}_n$ projector. We adjusted $s_f$ iteratively until the $\mathbf{P}_n$ projector produced to within 1% the same averaged residual mMCG as produced by OP. We hereafter refer to such a projector as a matched $\mathbf{P}_n$. Fig. 1(a) shows a 2 s window of the recorded data, $\mathbf{Y}$. Both mMCG and fMCG signals are visible in Fig. 1(a). The $\mathbf{M}_M$, OP, and matched $\mathbf{P}_n$ were derived from averaged mMCG data.
shown in Fig. 1(b). Figs. 1(c) and 1(e) show the output of the OP and matched \( P_n \) applied to the data shown in Fig. 1(a). Note that the mMCG is effectively suppressed in both Figs. 1(c) and 1(e). Figs. 1(d) and 1(f) show the resulting averaged residual mMCG plots for OP and matched \( P_n \).

In order to study the performance of OP and matched \( P_n \), a total of 113 datasets were processed and various measures were extracted. First, the maternal R markers for each dataset were found. Then, the fetal R markers were found after applying \( P_n \) with \( sf = 10^{-6} \). Without regard to a stopping threshold, OP projectors with 1 through 15 column vectors in \( V \) were computed for each dataset using (19). A corresponding matched \( P_n \) was computed for every OP projector. The peak magnitude of the averaged residual mMCG, the peak magnitude of the averaged fMCG (using fetal R markers), and \( sf \) were found for each of the 1695 combinations of data and projectors. Fig. 2 shows how \( P_n \) affected the magnitude of the residual mMCG and of the averaged fMCG as a function of \( sf \). The results presented in Fig. 2 show that the averaged residual mMCG is proportional to \((sf)^{1/2}\) over six decades as (35) suggests. Over the range \( 10^{-1} < sf < 10^{-4} \), the amplitude of the fMCG is approximately constant, but begins to drop when \( sf \) is smaller than \( 10^{-4} \). There is little point in using \( sf < 10^{-6} \) because the fMCG and mMCG are reduced equally.

We now describe a number of measures used to compare and evaluate the performance of the null projectors \( P_n \) and OP. We define windows for the fMCG QRS complex, pre-QRS, and post-QRS as follows:

1) the QRS window is 67.2 ms (±/− 10 samples @ 312.5 Hz) and centered on the fetal R marker;
2) the pre-QRS window is variable in length and includes all values that precede the QRS window; and
3) the post-QRS window is variable in length and includes all values that follow the QRS window, see Fig. 3(a) as an example.

We define the P-Q window as a 32 ms (10 samples @ 312.5 Hz) window that exhibits minimal change in signal level and is located between the end of the P wave and the start of the QRS complex. At each possible window position, the window mean for each channel is subtracted from the data and the L1 norm is computed. The window position with the lowest norm is selected. The position relative to the R marker and the P–Q window means are saved for later use. Fig. 3(b) shows an example of the P–Q window position after the means have been subtracted (baseline correction) [26].

To show how the null projectors affected fetal QRS amplitude, the amplitude before and after application of the projector was examined. To estimate the actual fetal QRS amplitude, we simply averaged the unprojected data (raw) using the fetal R markers and extracted the amplitude of the largest signal. See Fig. 4(a) for an example. Some remaining mMCGs are present in the fMCG signal. To minimize this influence, we rejected datasets where the RMS value of the QRS window was not at least ten times the RMS value of the concatenated pre-QRS and post-QRS windows. Ninety-five datasets met our criterion. We then determined the largest peak in the time averaged fetal QRS before and after applying OP or matched \( P_n \). The ratio gave the attenuation associated with each OP iteration and matched \( P_n \).

As an example, the averaged fMCG that results from applying two iterations of OP is shown in Fig. 4(b), and the averaged fMCG that results from applying the matched \( P_n \) is shown in Fig. 4(c). We also note that we used the same peak for each paired OP and matched \( P_n \) calculation so that there is no bias in favor one or the other because of the residual mMCG.

The attenuation at each OP step is presented in Fig. 5 in a boxplot format. Normally, the OP algorithm terminates between 5 and 10 iterations. Considering that range, the median attenuation for OP was less than the matched \( P_n \), but gain variation was more. By these measures, the two methods appear to be roughly equivalent with both methods showing increased...
In practice, we use the peak magnitude of the averaged mMCG residual as a metric for measuring the suppression of the mMCG waveform, but it can be argued that a better metric would be the RMS of the mMCG residual. To explore that possibility, we computed the RMS of the unaveraged data using the previously defined windows: QRS, pre+post QRS, and P–Q (with baseline removed). After applying a $P_n$ projector, QRS data around each fetal R marker were collected and combined, and then the RMS value was computed. Data not marked as QRS were classified as pre+post QRS data and the RMS value was computed. Using the averaged fMCG, the position of the P–Q window and the channel means were found as described above. The P–Q window means were subtracted from the raw data on a per channel basis. The P–Q data relative to each R marker was then combined and the RMS value was computed. Computing the RMS during the P–Q interval with the mean removed is our best attempt to measure the RMS residual mMCG without fMCG contamination. Even though the QRS and P–Q windows are short, there were sufficient samples for computing the RMS because each dataset had at least 1000 fetal R markers. Finally, the averaged residual mMCG was computed. All 113 recordings were processed. The median values obtained using the $P_n$ operator are shown in Fig. 6. The values for iteration “0” were computed on the original data without application of a projector. We do not show the OP results because they are similar to the $P_n$ results.

The three RMS values computed using the raw data plotted at position “0” have almost identical median values, indicating that the short window lengths do not bias the results and that the mMCG is the dominant RMS signal in the raw data. Over the first few OP iterations, the fetal RMS values and the residual mMCG parallel each other. However, by the fifth iteration, the change in the gain shown in Fig. 5(a) fully accounts for the monotonic drop in the fetal RMS median values. In contrast, the median residual mMCG data closely follows (35). Clearly the RMS values of the P–Q window set the upper limit for the RMS of the residual mMCG for all OP iterations. But the projection operator will pass some system noise so that we are not certain of the makeup of the P–Q signal especially since the median values track the QRS and do not follow (35). On this basis, we adopted the use of the residual mMCG (based on averaging) for our metric of mMCG suppression.

**B. Regularization**

Understanding the role of the eigenvalues of matrix $A$ in (16) is a key to selecting the proper value of $\lambda$ (or $sf$) for
regularization. We first consider the regularization of $P_n$. The model, $M_M$, for the null projector, $P_n$, is simply the time-averaged mMCG computed as described previously. Then, $A = B = M_M M_T^T$ in (16). Using the same dataset used in Fig. 1, the eigenvalues of $(A + \lambda I)$ were computed using four values of $\lambda$. Fig. 7(a) is a plot of the first 30 eigenvalues ranked from largest to smallest. The solid line is for $\lambda = 10^{-35}$ which is far below any useful value but satisfied the requirements of our matrix inversion algorithm. The dashed lines are curves obtained using the indicated values of $\lambda$. In Fig. 7(b), the 30 smallest eigenvalues of $P_n$, ranked from smallest to largest, are plotted. Curves for the three values of $\lambda$ are shown, but the curve for $\lambda = 10^{-35}$ is completely outside the range of the graph. The data shown in Fig. 8 illustrate the effect $\lambda$ has upon the performance of $P_n$. Fig 8(a) shows a second window of recorded data $Y$, and 8(b)—(d) show the effect of the three indicated values of $\lambda$ on the properties of $P_n$. Note the scale changes in panels B, C, and D. In this example, there are three maternal and four fetal QRS complexes. The fMCG is barely visible in 8(a) and the mMCG is not visible in 8(d). Note also that the fMCG is significantly attenuated going from 8(c) and (d). We did not present the data corresponding to the case where $\lambda = 10^{-35}$ because no signal remained. According to (32), in the extreme, when $\lambda \to 0$, then $P_n \to 0$.

If $\lambda$ is significantly less than an eigenvalue found in $A$, then any signal associated with that eigenvector will be suppressed by $P_n$ according to (32). Since the eigenvectors of $A$ come directly from the averaged mMCG, the projector will effectively suppress the major signal components of the mMCG. $P_n$ is similar to the identity matrix, except for the suppressed mMCG eigenvectors, and as expected, most eigenvalues are very close to unity as indicated by Fig. 7(b). As $\lambda$ is reduced, more and more eigenvalues are attenuated, and consequently, all signals including the fMCG are affected accordingly.

Since the mMCG is dominant in all of our data, plots of the eigenvalues of $(A + \lambda I)$ are always similar to Fig. 7(a) in that one eigenvalue dominates the power in the mMCG. We found that the trace of $A$ is always just larger than the largest eigenvalue. For the data shown in Fig. 7(a), the largest eigenvalue is $4.94 \cdot 10^{-20}$ and the trace is $5.33 \cdot 10^{-20}$. Computing the trace is faster and simpler than computing an inverse and, then, finding the largest eigenvalue. If we use the trace of $A$ as an approximate value for the largest eigenvalue, according to (32) we have a convenient method to suppress the dominant component ($d_{\text{max}}$) of the mMCG to a known fraction, specifically $\mu_{\text{max}} \approx 0.16$.

We now examine the effect of $\lambda$ on model-based projectors. Starting with the maternal and fetal models, $M_M$ and $M_F$, respectively, we concatenated them according to (10) to get $M$ and set $A = MM^T$ and $B = M_F M_F^T$. The eigenvalues of $(A + \lambda I)$ were computed using four values of $\lambda$ using the concatenated model. Each set of eigenvalues was ranked in descending order and the first 30 from each set are shown in Fig. 9. The solid curve is the data using $\lambda = 10^{-35}$ and the dashed curves result from using the indicated values of $\lambda$. For reference, the eigenvalues from the individual models $A = M_M M_M^T$ (dot-dash, “mMCG”) and $A = M_F M_F^T$ (dot-dash, “fMCG”) are provided for comparison with the full model. The solid line and the two dash-dot lines below the solid line use $\lambda = 10^{-35}$ to satisfy the requirements of our matrix inversion algorithm.

Observe that the first eigenvector of the fMCG model (bottom trace) is approximately $1 \cdot 10^{-22}$. To demonstrate the effect that $\lambda$ has on performance, two values of $\lambda$ just above $(2 \cdot 10^{-22})$ and then below $(4 \cdot 10^{-24})$ were selected. A value much less $(2 \cdot 10^{-27})$ was also included in the analysis. The three values Pn*Y, $\lambda = 10^{-20}$

Fig. 8. (a) Two seconds of recorded raw data from the dataset associated with the projectors described in Fig. 7. There are four fetal heartbeats and three maternal heartbeats in all panels. (b)–(d) show the same data as (a), after the application of $P_n$ projectors computed using the indicated values of $\lambda$. Note that different scales were selected for clarity.
Fig. 9. Plots of the 30 largest eigenvalues computed using three models and four values of $\lambda$. The three dashed lines and the solid line are values computed using the concatenated model (mMCG + fMCG). The two dot-dash lines are values computed using the individual models mMCG and fMCG. The values of $\lambda$ used to compute the top three curves are indicated. For the bottom three curves, $\lambda = 10^{-35}$ was used.

Fig. 10. (a)—(c) are plots of 4 s of processed output from three different $P_n$ projectors. All three projectors used the same mMCG + fMCG model $M$, but differed in the value of $\lambda$, $2 \cdot 10^{-22}$, $4 \cdot 10^{-24}$, and $2 \cdot 10^{-27}$, respectively.

Fig. 11. Covariance and a model projector were computed from the same data as shown in Fig. 10. The solid line is the plot of the eigenvalues of $(A + \lambda I)$ from the model (mMCG + fMCG) with $\lambda = 10^{-35}$. The dash-dot line is the plot of the eigenvalues found using the same model but setting $sf = 10^{-5}$. The dashed line is the plot of the eigenvalues found from the covariance of the recorded data, $Y$.

The results verify (38). For all eigenvalues smaller than the value of $\lambda$ chosen well below the principle eigenvalues of the fMCG components and the resulting fMCG waveforms have almost identical QRS amplitudes. Fig. 10(c) shows more details, corresponding to the inclusion of more eigenvalues. In Fig. 9, the first eigenvalue found in the fMCG model is the largest eigenvalue of (38). The fourth eigenvalue from the combined model of $A$ is almost identical with the first eigenvalue of $B$, and the respective eigenvectors are nearly identical. In reference to (38), we predicted unity gain for eigenvectors common to $A$ and $B$ because $g_i/d_i \approx 1$ when $d_i >> \lambda$. By comparing the R wave amplitude of the averaged fMCG (RAW) shown in Fig. 4(a) to the R wave amplitudes seen in Fig. 10(b) and 10(c), we see that the projector gain is approximately unity. Finally, there is no discernible mMCG in Fig. 10. As stated, (38) will pass the mMCG at a reduced level relative to the fMCG and there is no eigenvector associated with the mMCG in the $B$ component of the projector.

The behavior of a covariance projector is similar to a model projector. Using the same dataset, we computed the eigenvalues of the inverse terms from the model-based projector of (20) and the covariance based projector of (27). The 50 largest eigenvalues of the inverse terms are plotted in Fig. 11. The solid line is from the combined model (mMCG and fMCG concatenated) and $\lambda = 10^{-35}$. The dot-dash line is from the same model but $sf = 10^{-5}$. The dash line is from the covariance of $C_Y$ with no regularization. In our datasets, the system noise serves to regularize the inverse of the data covariance matrix $C_Y$. While one might choose to add a $\lambda I$ term, we have never found it to be advantageous.

In Fig. 11, we see that the covariance noise floor is roughly five decades below the largest eigenvalue. To compare the performance of the covariance projector versus the performance
of the model projector, we set \( sf = 10^{-5} \), so that the regularization floor is also five decades below the largest eigenvalue. Signals produced by the two projectors are presented in Fig. 12. The output of the model projector is shown in Fig. 12(a), while the output of the covariance projector is shown in Fig. 12(b). Note that the data shown in Fig. 10 correspond to the same time window as the data shown in Fig. 12. Further, the regularization selected for Fig. 10(b) appears to produce a signal similar to the covariance projector of Fig. 12(b). In Fig. 12(c), \( \lambda \) was set about two orders of magnitude above the system noise floor. At this level, some fMCG components are being suppressed.

C. Redistribution of Null Projectors

Null projection operators redistribute signals that are not removed from signal space. To compare the redistribution of OP and \( \mathbf{P}_n \), we defined a redistribution estimator (28). Larger values of \( R_e \) indicate a departure from the identity matrix and likely increase in redistribution. Using 113 datasets, the value for \( R_e \) was computed for 1695 OP and matched \( \mathbf{P}_n \), see Fig. 13. The results are presented as box plots for the respective projectors. We reiterate that the redistribution estimator, \( R_e \), is only an estimate of the potential for redistribution because the actual redistribution of fetal signals also depends upon the non-mMCG signals. The potential for redistribution is almost always less for the performance matched \( \mathbf{P}_n \) projector. Further, \( R_e \) increases with increasing number of OP vectors (and therefore decreasing \( sf \)).

Fig. 13 shows that there is to be a tradeoff between suppression of the mMCG and redistribution of fetal signals used to construct our models. To minimize redistribution errors in the fetal models, we use a combination of averaging and a low redistribution null projector (\( R_e \approx 0.004 \)) to suppress the mMCG. If the fMCG is large with respect to the mMCG and there are many fetal R markers, then only averaging is needed so that no redistribution is introduced into the model. See Fig. 4(a) for an example of incomplete averaging and Fig. 4(c) after application of \( \mathbf{P}_n \).

D. Computation of Models

In contrast to the dataset in the previous sections that did not have fetal breathing, a dataset thought to include significant periods of fetal breathing was selected for analysis. Characteristics that identify fetal breathing in magnetic recordings have been reported earlier (28). Fetal breathing occurs about 30% of the time near term (29) and has the following characteristics: quasi periodic sinusoidal of about 1 Hz, amplitude comparable with the fMCG, in sensor space it presents close to the fMCG signal, and it is commonly intermittent. The signal identified has all of the above characteristics and for the purposes of this paper, we assume that the observed signal is fetal breathing. Fig. 14 shows a three-component statistical model extracted from the selected dataset that includes mMCG (\( \mathbf{M}_M \)), fMCG (\( \mathbf{M}_F \)), and a signal that is likely fetal breathing (\( \mathbf{M}_B \)).

We first extracted the maternal R markers and then used them to compute the averaged mMCG (\( \mathbf{M}_M \)) as shown in Fig. 14 column numbers 1 to 284. Computing the fetal model was a multistep procedure. Using \( \mathbf{M}_M \), a \( \mathbf{P}_n \) projector with \( sf = 4.5 \cdot 10^{-5} \), was applied to the data to reveal the fetal R waves. See Fig. 15(a) for an example that shows fMCG and fetal breathing. Note that the null projectors primarily suppress the mMCG, but other signals remain, although attenuated. The fetal R markers were extracted using the algorithm described in (27). Referring back to the discussion concerning Fig. 4, most of the mMCG was removed from the fMCG and fetal breathing models by averaging. To complete the suppression of the mMCG,
Fig. 14. Plot of a three component statistical model from a 148 channel recording. The mMCG model occupies columns 1 to 284, the fMCG model occupies columns 285 to 428, and the fetal breathing model occupies columns 429 to 729.

Fig. 15. Dataset with fetal breathing was selected for analysis. (a) Signal from a three second window produced by the null projector, \( P_n \) (\( sf = 4.5 \times 10^{-5} \)). This signal is free of mMCG and was used to find fetal R markers. It shows significant fetal breathing. (b) The output of a second \( P_n \) (\( sf = 10^{-1} \)) for the same three second window. The fMCG model and the breathing model shown in Fig. 14 were computed by averaging this signal around the fetal R markers and the zero crossing of the breathing pattern, respectively.

Fig. 16. Three seconds of output from three different projection operators designed to pass only the fMCG signal and applied to the same data. (a) Output of a two-model fetal mMCG projector. In this case, the fetal breathing component was not included in the model. (b) Output of an fMCG projector using the three-component model shown in Fig. 14. (c) Output of a covariance fMCG projector.

\( M_B \) was found by manually identifying the zero crossing point of 19 cycles of the breathing signal and averaging 150 time samples (0.48 s) on either side of the zero crossings. The fetal breathing model, \( M_B \), is shown in columns 429 to 729.

E. Model and Covariance Projectors

The models in Fig. 14 were used to construct three different projectors designed to isolate either the fMCG or the fetal breathing signal. The results are shown in Figs. 16 and 17. For reference, the time windows of Figs. 15, 16, and 17 are the same.

Fig. 16(a) shows the output of an fMCG projector, \( P_F \), computed using (20) but with only \( M_M \) and \( M_F \) included in \( M \). Fig. 16(b) shows the output of \( P_F \), but in this case \( M_M \), \( M_F \), and \( M_B \) are included in \( M \). In Fig. 16(a) and 16(b), \( sf = 10^{-5} \).

Fig. 16(c) shows the output of a covariance-based projector \( P_C \) computed using (27). In this case, the covariance, \( C_i \), of the fMCG signal was estimated from (26) using the fMCG model, \( M_F \), shown in Fig. 14. All three projectors passed the fMCG signal and suppressed the mMCG signal. However, the two-model projector of Fig. 16(a), also passed some of the fetal breathing signal, while the other two projectors attenuated the fetal breathing signal. The three-model projector produced comparable results to the \( P_C \) projector when \( 10^{-4} > sf > 10^{-6} \) while the covariance projector did not require regularization.

Fig. 17(a) shows the output of an fMCG projector, \( P_F \), computed using (21), \( M_M \), \( M_F \), and \( M_B \), and \( sf = 10^{-5} \).

Fig. 17(b) and 17(c) shows the output of two different covariance-based projectors computed using (27) but utilizing two different methods to compute \( C_i \). In Fig. 17(b), \( C_i \) was estimated using \( M_B \) and (26). In this example, the amplitude of the projector’s output was scaled...
by 0.2 to match the amplitude seen in Fig. 17(a). The scaling by 0.2 reflects the intermittent nature of the fetal breathing as indicated in the discussion following (26). In Fig. 17(c), C_1 was estimated as follows: The signal Y was filtered with an eighth order, Butterworth, zero-phase, digital filter with a passband of 0.5 to 1.25 Hz. The narrow bandwidth filter separated the fetal breathing signal from other signals, and the covariance, C_1, was calculated directly from the bandwidth-limited Y. There was no need to rescale the amplitude. The important point to be made by this example is that any appropriate method can be utilized to estimate the covariance of such an isolated signal.

**F. Other Considerations**

Conceptually, a covariance-based mMCG projector could be used to form a null projector, but this approach fails to achieve the hoped for results, usually leaving residual signals that are larger than the null projector defined in (18). As pointed out, a covariance projector may typically affect the amplitude of a projected signal. A mismatch in amplitude of a few percent between the input and output is inconsequential for an isolation projector, but will limit the rejection when that projector is used to form a null projector. Attempts to scale the gain of multimodel and covariance projectors for use as a null projector fell short of the performance of P_n.

The primary scope of this paper is to define and present effective projectors and to demonstrate their utility when applied to a moderate length recording of 10 min. Longer recordings make it more likely that either the mother or fetus will move, violating the requirement that the signals remain quasi-stationary. To some extent, nonstationarity can be dealt with by finding periods of maternal and fetal inactivity from which to build the desired projector. Since nonstationarity is data dependent, it can only be addressed in a statistical study, which is beyond the scope of this paper.

Finally, (12) introduces the model resolution matrix [30], [31] for linear inverse problems, which might provide a means to evaluate crosstalk between signal subgroups.

**V. Conclusion**

Both the fetal and maternal cardiac signals are generated by multiple current sources that are temporally correlated and spread over many centimeters. Even so, we demonstrate that sensor level projection operators can be designed that will effectively suppress the mMCG signal or selectively pass either the fMCG signal or the fetal breathing signal. Our time-averaged model-based projectors utilize the unweighted minimum norm formalism, while our covariance projectors utilized the weighted minimum norm formalism.

In one dataset, where we suspected a fetal breathing signal, we build a three-model projector. The suppression of fetal breathing in the fMCG output was noticeably improved. We also demonstrated the feasibility of using either a time-locked averaged model or a bandwidth-limited signal for computation of a covariance projector. Even though the fMCG and the fetal breathing signal overlapped spatially, the two signals were separated with minimal crosstalk.

We developed equations that predict the gain of our projectors as a function of their eigenvalues. The equations provide an improved understanding of the effects of regularization in the minimum-norm formalism when applied to our projectors.

We demonstrate the effectiveness of a null projector, P_n, that may be used to suppress the mMCG signal and that the P_n projector is very similar in performance to the OP projector. Further, the degree of suppression may be set by the regularization parameter. The method was automated and used to find null projectors for 113 datasets without any user intervention.

Finally, we address the issue of redistribution in a quantifiable way and describe how to reduce the effects of the redistribution.

**REFERENCES**


James D. Wilson received the B.S. degree in physics from Arkansas Polytechnic College, Russellville, Arkansas, in 1972, and the M.S. degree in instrumental sciences from the University of Arkansas, Fayetteville, AR, USA, in 1984. From 1972 to 1978, he was an Electronics Design Engineer for Fantron Corp. in Little Rock, AR, USA. He joined the University of Arkansas for Medical Sciences in 1978 as a Research Assistant researching medical aerosols. In 1985, he began teaching electronics in the Graduate Institute of Technology in Little Rock, AR, USA. He is currently an Assistant Director for Research, Graduate Institute of Technology, University of Arkansas at Little Rock campus. His research interests include fetal MEG studies, application of ultrasound for clot lysis, aerosol physics, electronics for instrumentation, and signal processing. He has 66 journal publications, two book chapters, and four patents.

Mr. Wilson served six years on the Arkansas Highway and Transportation Department Research Advisory Council. He is a Member of the Sigma XI Research Society.

Jens Haueisen received the M.S. and Ph.D. degrees in electrical engineering from the Technical University Ilmenau, Ilmenau, Germany, in 1992 and 1996, respectively. From 1996 to 1998, he was a Postdoc and, from 1998 to 2005 was the Head of the Biomagnetic Center, Friedrich-Schiller-University, Jena, Germany. Since 2005, he is a Professor of biomedical engineering and directs the Institute of Biomedical Engineering and Informatics at the Technical University Ilmenau, Germany.

His main research interests include the numerical computation of bioelectric and biomagnetic fields and biological signal analysis.