Chapter 26: Spatio-temporal models for EEG

W. Penny, N. Trujillo-Barreto and E. Aubert

May 25, 2006

Introduction

Imaging neuroscientists have at their disposal a variety of imaging techniques for investigating human brain function [Frackowiak et al. 2003]. Among these, the Electroencephalogram (EEG) records electrical voltages from electrodes placed on the scalp, the Magnetoencephalogram (MEG) records the magnetic field from sensors placed just above the head and functional Magnetic Resonance Imaging (fMRI) records magnetisation changes due to variations in blood oxygenation.

However, as the goal of brain imaging is to obtain information about the neuronal networks that support human brain function, one must first transform measurements from imaging devices into estimates of intracerebral electrical activity. Brain imaging methodologists are therefore faced with an inverse problem, 'How can one make inferences about intracerebral neuronal processes given extracerebral or vascular measurements ?'

We argue that this problem is best formulated as a model-based spatio-temporal deconvolution problem. For EEG and MEG the required deconvolution is primarily spatial, and for fMRI it is primarily temporal. Although one can make minimal assumptions about the source signals by applying 'blind' deconvolution methods [Makeig et al. 2002, McKeown et al. 1998], knowledge of the underlying physical processes can be used to great effect. This information can be implemented in a forward model that is inverted during deconvolution. In M/EEG, forward models make use of Maxwell's equations governing propagation of electromagnetic fields [Baillet et al. 2001], and, in fMRI hemodynamic models that link neural activity to 'Balloon' models of vascular dynamics [Friston et al. 2000].

To implement a fully spatio-temporal deconvolution, time-domain fMRI models must be augmented with a spatial component and spatial-domain M/EEG models with a temporal component. The previous Chapter showed how this could be implemented for fMRI. This Chapter describes a model-based spatio-temporal deconvolution method for M/EEG.

The underlying forward or 'generative' model incorporates two mappings. The first specifies a time-domain General Linear Model (GLM) at each point in source space. This relates effects of interest at each voxel to source activity at that voxel. This is identical to the 'mass-univariate' approach that is widely used in the analysis of fMRI [Frackowiak et al. 2003]. The second mapping relates source activity to sensor activity at each time point using the usual spatial-domain lead-field matrix (see Chapter 28).

Our model therefore differs from the standard generative model implicit in source reconstruction by having an additional level that embodies temporal

priors. There are two potential benefits of this approach. First, the use of temporal priors can result in more sensitive source reconstructions. This may allow signals to be detected that cannot be detected otherwise. Second, it provides an analysis framework for M/EEG that is very similar to that used in fMRI. The experimental design can be coded in a design matrix, the model fitted to data, and various effects of interest can be characterised using 'contrasts' [Frackowiak et al. 2003]. These effects can then be tested for statistically using Posterior Probability Maps (PPMs), as described in previous Chapters. Importantly, the model does not need to be refitted to test for multiple experimental effects that are potentially present in any single data set. Sources are estimated once only using a spatio-temporal deconvolution rather than separately for each temporal component of interest.

The Chapter is organised as follows. In the Theory section we describe the model and relate it to existing distributed solutions. The success of the approach rests on our ability to characterise neuronal responses, and task-related differences in them, using GLMs. We describe how this can be implemented for the analysis of ERPs and show how the model can be inverted to produce source estimates using Variational Bayes (VB). The framework is applied to simulated data and data from an EEG study of face processing.

Theory

Notation

Lower case variable names denote vectors and scalars. Whether the variable is a vector or scalar should be clear from the context. Upper case names denote matrices or dimensions of matrices. In what follows $\mathsf{N}(x;\mu,\Sigma)$ denotes a multivariate normal density over x, having mean μ and covariance Σ. The precision of a Gaussian variate is the inverse (co)variance. A gamma density over the scalar random variable x is written as $\mathsf{Ga}(x;a,b)$. Normal and Gamma densities are defined in Chapter 26. We also use $||x||^2 = x^T x$, denote the trace operator as $\mathsf{Tr}(X)$, X^+ for the pseudo-inverse, and use $\mathsf{diag}(x)$ to denote a diagonal matrix with diagonal entries given by the vector x.

Generative Model

The aim of source reconstruction is to estimate Primary Current Density (PCD) J from measured M/EEG measurements Y. If we have $m = 1..M$ sensors, $g = 1..G$ sources and $t = 1..T$ time points then J is of dimension $G \times T$ and Y is of dimension $M \times T$. The applications in this Chapter use a cortical source space in which dipole orientations are constrained to be perpendicular to the cortical surface. Each entry in J therefore corresponds to the scalar current density at particular locations and time points. Sensor measurements are related to current sources via Maxwell's equations governing electromagnetic fields [Baillet et al. 2001] (see Chapter 28).

Most established distributed source reconstruction or 'imaging' methods

[Darvas et al. 2004] implicitly rely on the following two-level generative model

$$\begin{aligned} p(Y|J,\Omega) &= \prod_{t=1}^{T} \mathsf{N}(y_t; Kj_t, \Omega) \\ p(J|\alpha) &= \prod_{t=1}^{T} \mathsf{N}(j_t; 0, \alpha^{-1}D^{-1}) \end{aligned} \tag{1}$$

where j_t and y_t are the source and sensor vectors at time t, K is the $[M \times G]$ lead-field matrix and Ω is the sensor noise covariance. The matrix D reflects the choice of spatial prior and α is a spatial precision variable. This generative model is shown schematically in Figure 1 and can be written as a hierarchical model

$$Y = KJ + E \tag{2}$$

$$J = Z \tag{3}$$

in which random fluctuations E correspond to sensor noise and the source activity is generated by random innovations Z. Critically, these assumptions provide empirical priors on the spatial deployment of source activity (see Chapter 29).

Because the number of putative sources is much greater than the number of sensors, $G >> M$, the source reconstruction problem is ill-posed. Distributed solutions therefore depend on the specification of a spatial prior for estimation to proceed. A common choice is the Laplacian prior used, for example, in Low Resolution Electromagnetic Tomography (LORETA) [Pascual-Marqui et al. 1994]. This can be implemented in the above generative model by setting D to compute local differences as measured by an L^2-norm, which embodies a belief that sources are diffuse and highly distributed. Other spatial priors, such as those based on L^1-norms [Fuchs et al. 1999], L^p-norms [Auranen et al. 2005], or Variable Resolution Electromagnetic Tomography (VARETA) [Valdes-Sosa et al. 2000] can provide more focal source estimates. These are all examples of schemes that use a single spatial prior and are special cases of a more general model [Mattout et al. 2006] that covers multiple priors. In this model the sensor noise and spatial prior covariances are modelled as mixtures of components Ω_i and Q_i respectively

$$\begin{aligned} p(Y|J,\Omega) &= \prod_{t=1}^{T} \mathsf{N}(y_t; Kj_t, \rho_1\Omega_1 + \rho_2\Omega_2 + ...) \\ p(J|\alpha) &= \prod_{t=1}^{T} \mathsf{N}(j_t; 0, \gamma_1 Q_1 + \gamma_2 Q_2 + ...) \end{aligned} \tag{4}$$

The advantage of this model is that multiple priors can be specified and are mixed adaptively by adjusting the covariance parameters ρ_i and γ_i, as described in Chapter 29. One can also use Bayesian model selection to compare different combinations of priors, as described in Chapter 35. For simplicity, we will deal with a single spatial prior component because we want to focus on temporal priors. However, it would be relatively simple to extend the approach to cover multiple prior covariance (or precision) components.

Also, in Chapter 35 we will describe a prior over a model class that, when used with Bayesian Model Averaging (BMA), can automatically provide either focal or distributed solutions depending on the reconstruction at hand. The applications in this Chapter use Laplacian priors.

Whatever the choice of spatial prior, the majority of source reconstruction applications follow a single-pass serial processing strategy. Either (i) spatial processing first proceeds to create source estimates at each time point and then (ii) temporal models are applied at these 'virtual depth electrodes' [Darvas et al. 2004, Kiebel and Friston 2004, Brookes et al. 2004]. Or (ii) time series methods are applied in sensor space to identify components of interest using eg. time windowing [Rugg and Coles 1995] or time-frequency estimation and then (ii) source reconstructions are then based on these components.

In this Chapter we review a multiple-pass strategy in which temporal and spatial parameter estimates are improved iteratively to provide an optimised and mutually constrained solution. It is based on the following three-level generative model

$$p(Y|J,\Omega) = \prod_{t=1}^{T} \mathsf{N}(y_t; Kj_t, \Omega) \tag{5}$$

$$p(J|W,\lambda) = \prod_{t=1}^{T} \mathsf{N}(j_t^T; x_t W, \lambda^{-1} I_G) \tag{6}$$

$$p(W|\alpha) = \prod_{k=1}^{K} \mathsf{N}(w_k; 0, \alpha^{-1} D^{-1}) \tag{7}$$

The first level, equation 5, is identical to the standard model. In the second level, however, source activity at each voxel is constrained using a $[T \times K]$ matrix of temporal basis functions, X. The tth row of X is x_t. The generative model is shown schematically in Figure 2.

The precision of the source noise is given by λ. In this Chapter λ is a scalar. In this Chapter we will apply the framework to analyse Event-Related Potentials (ERPs) [Rugg and Coles 1995]. Event-related source activity is described by the time domain GLM and remaining source activity will correspond to spontaneous activity. The quantity λ^{-1} can therefore be thought of as the variance of spontaneous activity in source space.

The regression coefficients W determine the weighting of the temporal basis functions. The third level of the model is a spatial prior that reflects our prior uncertainty about W. Each regression coefficient map, w_k (row of W), is constrained by setting D to correspond to the usual L^2-norm spatial prior. The spatial prior that is usually on the sources now, therefore, appears at a superordinate level.

Different choices of D result in different weights and different neighborhood relations. The applications in this paper use $D = L^T L$, where L is the surface Laplacian defined as

$$L_{ij} = \begin{cases} 1, & \textit{if } i{=}j \\ -\frac{1}{N_{ij}}, & \textit{if } i \textit{ and } j \textit{ are geodesic neighbors} \\ 0, & \textit{otherwise.} \end{cases}$$

where N_{ij} is the geometric mean of the number of neighbors of i and j. This prior has been used before in the context of fMRI with Euclidean neighbors [Penny and Flandin 2005, Woolrich et al. 2001].

The first level of the model assumes that there is Gaussian sensor noise, e_t, with zero mean and covariance Ω. This covariance can be estimated from prestimulus or baseline periods when such data are available [Sahani and Nagarajan 2004]. Alternatively, we assume that $\Omega = \mathsf{diag}(\sigma^{-1})$ where the mth element of σ^{-1} is the noise variance on the mth sensor. We provide a scheme for estimating σ_m, should this be necessary.

We also place Gamma priors on the precision variables σ, λ and α

$$\begin{aligned} p(\sigma) &= \prod_{m=1}^{M} \mathsf{Ga}(\sigma_m; b_{\sigma_{prior}}, c_{\sigma_{prior}}) \\ p(\lambda) &= \mathsf{Ga}(\lambda; b_{\lambda_{prior}}, c_{\lambda_{prior}}) \\ p(\alpha) &= \prod_{k=1}^{K} \mathsf{Ga}(\alpha_k; b_{\alpha_{prior}}, c_{\alpha_{prior}}) \end{aligned} \tag{8}$$

This allows the inclusion of further prior information into the source localisation. For example, instead of using baseline periods to estimate a full covariance matrix Ω we could use this data to estimate the noise variance at each sensor. This information could then be used to set $b_{\sigma_{prior}}$ and $c_{\sigma_{prior}}$, allowing noise estimates during periods of interest to be constrained softly by those from baseline periods. Similarly, we may wish to enforce stronger or weaker spatial regularisation on w_k by setting $b_{\alpha_{prior}}$ and $c_{\alpha_{prior}}$ appropriately. The applications in this Chapter, however, use uninformative gamma priors. This means that σ, λ and α will be estimated solely from the data Y.

In summary, the addition of the supraordinate level to our generative model induces a partioning of source activity into signal and noise. We can see this clearly by reformulating the probabilistic model as before

$$\begin{aligned} Y &= KJ + E \\ J^T &= XW + Z \\ W &= P \end{aligned} \tag{9}$$

Here we have random innovations Z which are 'temporal errors', ie. lack of fit of the temporal model, and P which are 'spatial errors', ie. lack of fit of a spatial model. Here the spatial model is simply a zero mean Gaussian with covariance $\alpha^{-1}D^{-1}$. We can regard XW as an empirical prior on the expectation of source activity. This empirical Bayes perspective means that the conditional estimates of source activity J are subject to bottom-up constraints, provided by the data, and top-down predictions from the third-level of our model. We will use this heuristic later to understand the update equations used to estimate source activity.

Temporal priors

The usefulness of the spatio-temporal approach rests on our ability to characterise neuronal responses using GLMs. Fortunately, there is a large literature

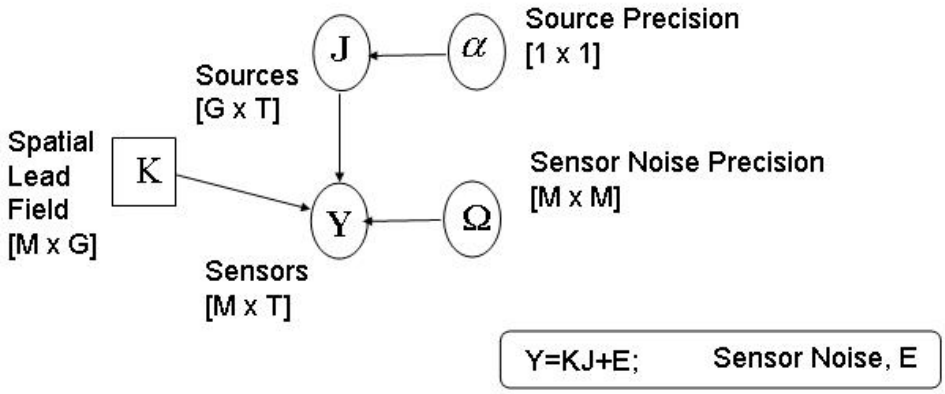

Figure 1: Generative model for source reconstruction. This is a graphical representation of the probabilistic model implicit in many distributed source solutions.

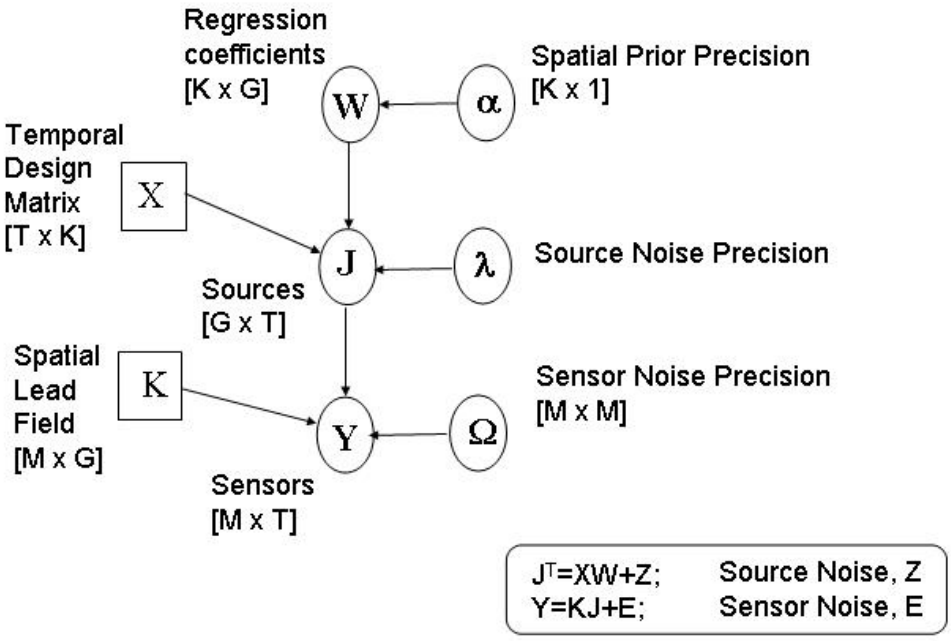

Figure 2: Generative model for source reconstruction with temporal priors. This is a hierarchical model with regression coefficients at the 'top' and M/EEG data at the 'bottom'.

that suggests this is possible. The type of temporal model necessary will depend on the M/EEG response one is interested in. These components could be (i) single trials, (ii) evoked components (steady-state or ERPs [Rugg and Coles 1995]) or (iii) induced components [Tallon-Baudry et al. 1996].

In this Chapter we focus on ERPs. We briefly review three different approaches for selecting an appropriate ERP basis set. These basis functions will form columns in the GLM design matrix, X (see equation 6 and figure 2).

Damped sinusoids

An ERP basis set can be derived from the fitting of Damped Sinusoidal (DS) components [Demiralp et al. 1998]. These are given by

$$\begin{aligned} j &= \sum_{k=1}^{K} w_k x_k \\ x_k &= \exp(i\phi_k)\exp(\alpha_k + i2\pi f_k)\delta_t \end{aligned} \tag{10}$$

where $i = \sqrt{-1}$, δ_t is the sampling interval and w_k, ϕ_k, α_k and f_k are the amplitude, phase, damping and frequency of the kth component. The $[T \times 1]$ vector x_k will form the kth column in the design matrix. Figure 3 shows how damped sinusoids can generate an ERP.

Fitting DS models to ERPs from different conditions allows one to make inferences about task related changes in constituent rhythmic components. For example, in [Demiralp et al. 1998], responses to rare auditory events elicited higher amplitude, slower delta and slower damped theta components than did responses to frequent events. Fitting damped sinusoids, however, requires a nonlinear estimation procedure. But approximate solutions can also be found using the Prony and related methods [Osborne and Smyth 1991] which require two-stages of linear estimation.

Once a DS model has been fitted, for example to the principal component of the sensor data, the components x_k provide a minimal basis set. Including extra regressors from a first-order Taylor expansion about phase, damping and frequency ($\frac{\partial x_k}{\partial \phi_k}, \frac{\partial x_k}{\partial \alpha_k}, \frac{\partial x_k}{\partial f_k}$) provides additional flexibility. Use of this expanded basis in our model would allow these attributes to vary with source location. Such Taylor series expansions have been particularly useful in GLM characterisations of hemodynamic responses in fMRI [Frackowiak et al. 2003].

Wavelets

ERPs can also be modelled using wavelets

$$j = \sum_{k=1}^{K} w_k x_k \tag{11}$$

where x_k are wavelet basis functions and w_k are wavelet coefficients. Wavelets provide a tiling of time-frequency space that gives a balance between time and frequency resolution. The Q-factor of a filter or basis function is defined as the central frequency to bandwidth ratio. Wavelet bases are chosen to provide constant Q [Unser and Aldroubi 1996]. This makes them good models of nonstationary signals, such as ERPs and induced EEG components

[Tallon-Baudry et al. 1996]. Wavelet basis sets are derived by translating and dilating a mother wavelet. Figure 4 shows wavelets from two different basis sets, one based on Daubechies wavelets and one based on Battle-Lemarie (BL) wavelets. These basis sets are orthogonal. Indeed the BL wavelets have been designed from an orthogonalisation of cubic B-splines [Unser and Aldroubi 1996].

If $K = T$, then the mapping $j \rightarrow w$ is referred to as a wavelet transform, and for $K > T$ we have an overcomplete basis set. More typically, we have $K \leq T$. In the ERP literature the particular subset of basis functions used is chosen according to the type of ERP component one wishes to model. Popular choices are wavelets based on B-splines [Unser and Aldroubi 1996].

In statistics, however, it is well known that an appropriate subset of basis functions can be automatically selected using a procedure known as 'wavelet shrinkage' or 'wavelet denoising'. This relies on the property that natural signals such as images, speech or neuronal activity can be represented using a sparse code comprising just a few large wavelets coefficients. Gaussian noise signals, however, produce Gaussian noise in wavelet space. This comprises a full set of wavelet coefficients whose size depends on the noise variance. By 'shrinking' these noise coefficients to zero using a thresholding procedure [Donoho and Johnstone 1994, Clyde et al. 1998], and transforming back into signal space, one can denoise data. This amounts to defining a temporal model. We will use this approach for the empirical work reported later on.

PCA

A suitable basis can also be derived from Principal Components Analysis (PCA). Trejo et al. [Trejo and Shensa 1999] for example, applied PCA and varimax rotation to the classification of ERPs in a signal detection task. They found, however, that better classification was more accurate with a Daubechies wavelet basis.

PCA decompositions are also used in the Multiple Signal Classification (MUSIC) approach [Mosher and Leahy 1998]. The dimension of the basis set is chosen to separate the signal and noise subspaces. Source reconstruction is then based on the signal, with information about the noise used to derive statistical maps based on pseudo-z scores. In [Friston et al. 2006], a temporal basis set is defined using the principal eigenvectors of a full-rank prior temporal covariance matrix. This approach makes the link between signal subspace and prior assumptions transparent.

Dimensionality

Whatever the choice of basis, it is crucial that the dimension of the signal subspace is less than the dimension of the original time series. That is, $K < T$. This is necessary for the temporal priors to be effective, both from a statistical and computational perspective.

Theoretically, one might expect the dimensionality of ERP generators to be quite small. This is because of the low-dimensional synchronisation manifolds that arise when nonlinear dynamical systems are coupled into an ensemble [Breakspear and Terry 2002].

In practice, the optimal reduced dimensionality can be found automatically using a number of methods. For wavelets this can be achieved using shrinkage

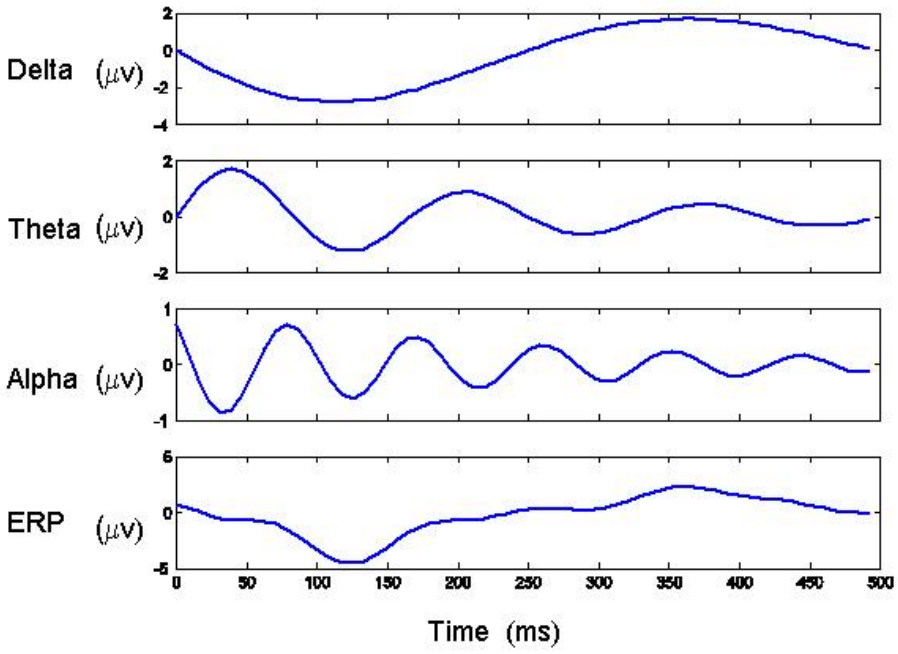

Figure 3: The figure shows how damped sinusoids can model ERPs. In this example damped delta, theta and alpha sinusoids, of particular phase, amplitude and damping, add together to form an ERP with an early negative component and a late positive component.

methods [Donoho and Johnstone 1994, Clyde et al. 1998], for PCA using various model order selection criteria [Minka 2000] and for damped sinusoids, Prony-based methods can use AR model order criteria [Roberts and Penny 2002]. Moreover, it is also possible to compute the model evidence of the source reconstruction model we have proposed, as shown in the following section. This can then be used to optimise the basis set.

Bayesian Inference

To make inferences about the sources underling M/EEG we need to invert our probabilistic model to produce the posterior density $p(J|Y)$. This is straightforward in principle and can be achieved using standard Bayesian methods [Gelman et al. 1995]. For example, one could use Markov Chain Monte Carlo (MCMC) to produce samples from the posterior. This has been implemented efficiently for dipole-like inverse solutions [Schmidt et al. 1999] in which sources are parameterised as spheres of unknown number, extent and location. It is,

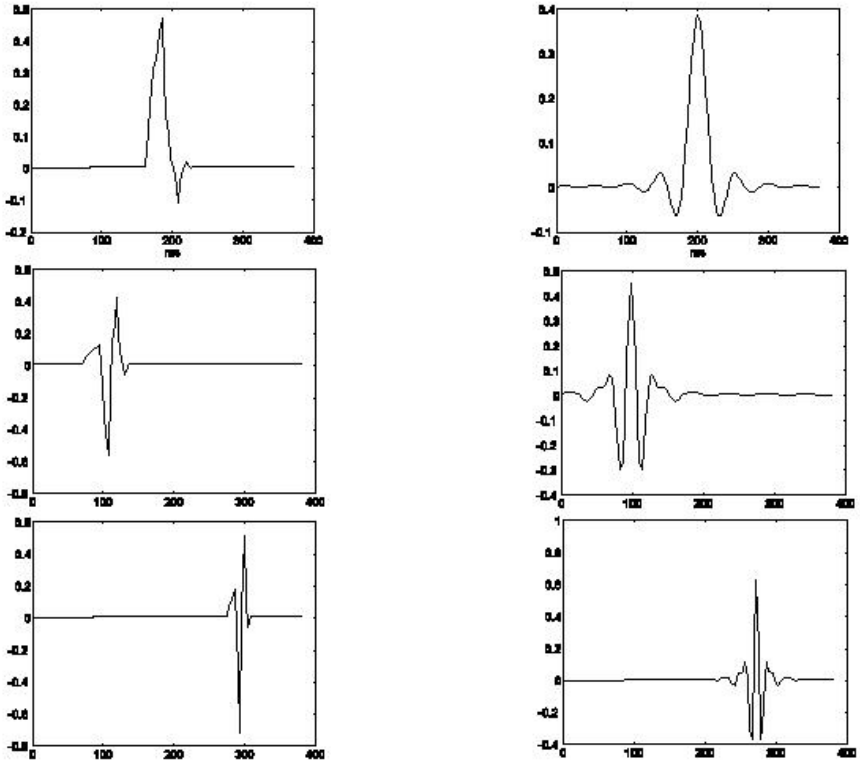

Figure 4: The graphs show wavelets from a Daubechies set of order 4 (left) and a Battle-Lemarie basis set of order 3 (right). The wavelets in the lower panels are higher frequency translations of the wavelets in the top panel. Each full basis set comprises multiple frequencies and translations covering the entire time domain.

however, computationally demanding for distributed source solutions, taking several hours for source spaces comprising $G > 1000$ voxels [Auranen et al. 2005]. In this work we adopt the computationaly efficient approximate inference framework called Variational Bayes (VB), that was reviewed in Chapter 26.

Approximate posteriors

For our source reconstruction model we assume the following factorisation of the approximate posterior

$$q(J, W, \alpha, \sigma, \lambda) = q(J)q(W)q(\alpha)q(\sigma)q(\lambda) \tag{12}$$

We also assume that the approximate posterior for the regression coefficients factorises over voxels

$$q(W) = \prod_{g=1}^{G} q(w_g) \tag{13}$$

This approximation was used in the spatio-temporal model for fMRI described in the previous Chapter.

Because of the spatial prior (equation 7), the regression coefficients in the true posterior $p(W|Y)$ will clearly be correlated. Our perspective, however, is that this is too computationally burdensome for current personal computers to take account of. Moreover, as we shall see in section , updates for our approximate factorised densities $q(w_g)$ do encourage the approximate posterior means to be similar at nearby voxels, thereby achieving the desired effect of the prior.

Now that we have defined the probabilistic model and our factorisation of the approximate posterior, we can use the procedure described in Chapter 26 to derive expressions for each component of the approximate posterior. We do not present details of these derivations in this Chapter. Similar derivations have been published elsewhere [Penny et al. 2005]. The following sections describe each distribution and the updates of its sufficient statistics required to maximise the lower bound on the model evidence, F.

Sources

Updates for the sources are given by

$$q(J) = \prod_{t=1}^{T} q(j_t) \tag{14}$$

$$q(j_t) = \mathsf{N}(j_t; \hat{j}_t, \hat{\Sigma}_{j_t}) \tag{15}$$

$$\hat{\Sigma}_{j_t} = \left(K^T \hat{\Omega} K + \hat{\lambda} I_G\right)^{-1} \tag{16}$$

$$\hat{j}_t = \hat{\Sigma}_{j_t} \left(K^T \hat{\Omega} y_t + \hat{\lambda} \hat{W}^T x_t^T\right) \tag{17}$$

where $\hat{j}_t$ is the tth column of $\hat{J}$ and $\hat{\Omega}$, $\hat{\lambda}$ and $\hat{W}$ are estimated parameters defined in the following sections. Equation 17 shows that our source estimates are the result of a spatio-temporal deconvolution. The spatial contribution to the estimate is $K^T y_t$ and the temporal contribution is $\hat{W}^T x_t^T$. From the

perspective of the hierarchical model, shown in Figure 5, these are the 'bottom-up' and 'top-down' predictions. Importantly, each prediction is weighted by its relative precision. Moreover, the parameters controlling the relative precisions, $\hat{\lambda}$ and $\hat{\Omega}$ are estimated from the data. This means that our source estimates derive from an automatically regularised spatio-temporal deconvolution. This property is shared by the spatio-temporal model for fMRI, described in the previous Chapter.

An alternative perspective on this computation is given by ignoring the regularisation term in 17. We then see that $\hat{\Sigma}_{j_t} K^T \hat{\Omega} = (K^T \hat{\Omega} K)^+ K^T \hat{\Omega} = B_w^T$, which is equivalent to a beamformer [Darvas et al. 2004]. Equation 17 then shows that our source estimates use beamformer predictions $B_w^T y_t$ that are modified using a temporal model. Beamformers cannot localise temporally correlated sources. But, as we shall see later, the spatio-temporal model can.

We end this section by noting that statistical inferences about current sources are more robust than point predictions. This property has been used to great effect with Pseudo-z beamformer statistics [Robinson and Vrba 1999], sLORETA [Pascual-Marqui 2002] and VARETA [Valdes-Sosa et al. 2000] source reconstructions, which divide current source estimates by their standard deviations. This approach can be adopted in the current framework as the standard deviations are readily computed from the diagonal elements of $\hat{\Sigma}_{j_t}$ using equation 24. Moreover, we can threshold these statistic images to create Posterior Probability Maps (PPMs), as introduced in Chapter 25.

Regression coefficients

Updates for the regression coefficients are given by

$$\begin{aligned} q(w_g) &= \mathsf{N}(w_g; \hat{w}_g, \hat{\Sigma}_{w_g}) \\ \hat{\Sigma}_{w_g} &= \left(\hat{\lambda} X^T X + d_{gg} \mathsf{diag}(\hat{\alpha})\right)^{-1} \\ \hat{w}_g &= \hat{\Sigma}_{w_g} \left(\hat{\lambda} X^T \hat{j}_g^T + \mathsf{diag}(\hat{\alpha}) r_g\right) \end{aligned} \tag{18}$$

where $\hat{\alpha}$ is defined in 22, d_{ij} is the i,jth element of D and r_g is given by

$$r_g = \sum_{g'=1, g' \neq g}^{G} d_{gg'} \hat{w}_{g'} \tag{19}$$

As shown in the previous Chapter, r_g is the weighted sum of neighboring regression coefficient estimators.

The update for $\hat{w}_g$ in equation 18 therefore indicates that the regression coefficient estimates at a given voxel regress towards those at nearby voxels. This is the desired effect of the spatial prior and it is preserved despite the factorisation in the approximate posterior. This equation can again be thought of in terms of the hierarchical model where the regression coefficient estimate is a combination of a bottom up prediction from the level below, $X^T \hat{j}_g^T$, and a top down prediction from the prior, r_g. Again, each contribution is weighted by its relative precision.

The update for the covariance in equation 18 shows that the only off-diagonal contributions are due to the design matrix. If the temporal basis functions are therefore chosen to be orthogonal then this posterior covariance will be

diagonal, thus making a potentially large saving of computer memory. One benefit of the proposed framework, however, is that non-orthogonal bases can be accomodated. This may allow for a more natural and compact description of the data.

Precision of temporal models

Updates for the precision of the temporal model are given by

$$
\begin{aligned}
q(\lambda) &= \mathsf{Ga}(\lambda; b_{\lambda_{post}}, c_{\lambda_{post}}) \\
\frac{1}{b_{\lambda_{post}}} &= \frac{1}{b_{\lambda_{prior}}} + \frac{1}{2}\sum_t \left(||\hat{j}_t - \hat{W}^T x_t^T||^2 + \mathsf{Tr}(\hat{\Sigma}_{j_t}) + \sum_{g=1}^{G} x_t \hat{\Sigma}_{w_g} x_t^T \right) \\
c_{\lambda_{post}} &= c_{\lambda_{prior}} + \frac{GT}{2} \\
\hat{\lambda} &= b_{\lambda_{post}} c_{\lambda_{post}}
\end{aligned}
\tag{20}
$$

In the context of ERP analysis, these expressions amount to an estimate of the variance of spontaneous activity in source space, $\hat{\lambda}^{-1}$, given by the squared error between the ERP estimate, $\hat{W}^T x_t^T$, and source estimate, $\hat{j}_t$, averaged over time and space and the other approximate posteriors.

Precision of forward model

Updates for the precision of the sensor noise are given by

$$
\begin{aligned}
q(\sigma) &= = \prod_{m=1}^{M} q(\sigma_m) \\
q(\sigma_m) &= \mathsf{Ga}(\sigma_m; b_{\sigma_{post}}, c_{\sigma_{post}}) \\
\frac{1}{b_m} &= \frac{1}{b_{\sigma_{prior}}} + \frac{1}{2}\sum_t \left(y_{mt} - k_m^T \hat{j}_t \right)^2 + \frac{1}{2} k_m^T \hat{\Sigma}_{j_t} k_m \\
c_m &= c_{\sigma_{prior}} + \frac{T}{2} \\
\hat{\sigma}_m &= b_m c_m \\
\hat{\Omega}^{-1} &= \mathsf{diag}(\hat{\sigma})
\end{aligned}
\tag{21}
$$

These expressions amount to an estimate of observation noise at the mth sensor, $\hat{\sigma_m}^{-1}$, given by the squared error between the forward model and sensor data, averaged over time and the other approximate posteriors.

Precision of spatial prior

Updates for the precision of the spatial prior are given by

$$\begin{aligned} q(\alpha) &= \prod_{k=1}^{K} q(\alpha_k) \\ q(\alpha_k) &= \mathsf{Ga}(\alpha_k; b_{\alpha_{post}}, c_{\alpha_{post}}) \\ \frac{1}{b_{\alpha_{post}}} &= \frac{1}{b_{\alpha_{prior}}} + ||D\hat{w}_k^T||^2 + \sum_{g=1}^{G} d_g s_{gk} \\ c_{\alpha_{post}} &= c_{\alpha_{prior}} + \frac{G}{2} \\ \hat{\alpha}_k &= b_{\alpha_{post}} c_{\alpha_{post}} \end{aligned} \tag{22}$$

where s_{gk} is the kth diagonal element of $\hat{\Sigma}_{w_g}$. These expressions amount to an estimate of the 'spatial noise variance', $\hat{\alpha}_k^{-1}$, given by the discrepancy between neighboring regression coefficients, averaged over space and the other approximate posteriors.

Implementation details

A practical difficulty with the update equations for the sources is that the covariance matrix $\hat{\Sigma}_{j_t}$ is of dimension $G \times G$ where G is the number of sources. Even low resolution source grids typically contain $G > 1000$ elements. This therefore presents a problem. A solution is found, however, with use of a Singular Value Decomposition (SVD). First, we define a modified lead field matrix $\bar{K} = \hat{\Omega}^{1/2} K$ and compute its SVD

$$\begin{aligned} \bar{K} &= USV^T \\ &= U\bar{V} \end{aligned} \tag{23}$$

where $\bar{V}$ is an $M \times G$ matrix, the same dimension as the lead field, K. It can then be shown using the matrix inversion lemma [Golub and Van Loan 1996] that

$$\begin{aligned} \hat{\Sigma}_{j_t} &= \hat{\lambda}^{-1} (I_G - R_G) \\ R_G &= \bar{V}^T (\hat{\lambda} I_M + SS^T)^{-1} \bar{V} \end{aligned} \tag{24}$$

which is simple to implement computationally, as it only requires inversion of an $M \times M$ matrix.

Source estimates can be computed as shown in equation 17. In principle, this means the estimated sources over all time points and source locations are given by

$$\hat{J} = \hat{\Sigma}_{j_t} K^T \hat{\Omega} Y + \hat{\lambda} \hat{\Sigma}_{j_t} \hat{W}^T X^T$$

In practice, however, it is inefficient to work with such a large matrix during estimation. We therefore do not implement equations 16 and 17 but, instead, work in the reduced space $\hat{J}_X = \hat{J}X$ which are the sources projected onto the

design matrix. These projected source estimates are given by

$$\begin{aligned} \hat{J}_X &= \hat{J}X \\ &= \hat{\Sigma}_{j_t} K^T \hat{\Omega} Y X + \hat{\lambda} \hat{\Sigma}_{j_t} \hat{W}^T X^T X \\ &= A_{K\Omega} Y X + \hat{\lambda} A_W X^T X \end{aligned} \tag{25}$$

where YX and X^TX can be pre-computed and the intermediate quantities are given by

$$\begin{aligned} A_{K\Omega} &= \hat{\Sigma}_{j_t} K^T \hat{\Omega} \\ &= \hat{\lambda}^{-1} \left(K^T - R_G K^T \right) \\ A_W &= \hat{\Sigma}_{j_t} \hat{W}^T \\ &= \hat{\lambda}^{-1} \left(\hat{W}^T - R_G \hat{W}^T \right) \end{aligned} \tag{26}$$

Because these matrices are only of dimension $G \times M$ and $G \times K$ respectively, $\hat{J}_X$ can be efficiently computed. The term $X^T \hat{j}_g^T$ in equation 18 is then given by the gth row of $\hat{J}_X$.

The intermediate quantities can also be used to compute model predictions as

$$\begin{aligned} \hat{Y} &= K\hat{J} \\ &= K A_{K\Omega} Y + \hat{\lambda} K A_W X^T \end{aligned} \tag{27}$$

The m,tth entry in $\hat{Y}$ then corresponds to the $k_m^T \hat{j}_t$ term in equation 21. Other computational savings are as follows. For equation 21 we use the result

$$k_m^T \hat{\Sigma}_{j_t} k_m = \frac{1}{\hat{\sigma}_m} \sum_{m'=1}^{M} \frac{s_{m'm'}^2 u_{mm'}^2}{s_{m'm'}^2 + \hat{\lambda}} \tag{28}$$

where s_{ij} and u_{ij} are the i,jth entries in S and U respectively. For equation 20 we use the result

$$\mathsf{Tr}(\hat{\Sigma}_{j_t}) = \sum_{i=1}^{M} \frac{1}{s_{ii}^2 + \hat{\lambda}} + \frac{G - M}{\hat{\lambda}} \tag{29}$$

To summarise, our source reconstruction model is fitted to data by iteratively applying the update equations until the change in the negative free energy (see Chapter 26), F, is less than some user-specified tolerance. This procedure is summarised in the pseudo-code in Figure 6. This amounts to a process in which sensor data is spatially deconvolved, time series models are fitted in source space, and then the precisions (accuracy) of the temporal and spatial models are estimated. This process is then iterated and results in a spatio-temporal deconvolution in which all aspects of the model are optimised to maximise a lower bound on the model evidence.

Results

This section presents some preliminary qualitative results. In what follows we refer to the spatio-temporal approach as 'VB-GLM'.

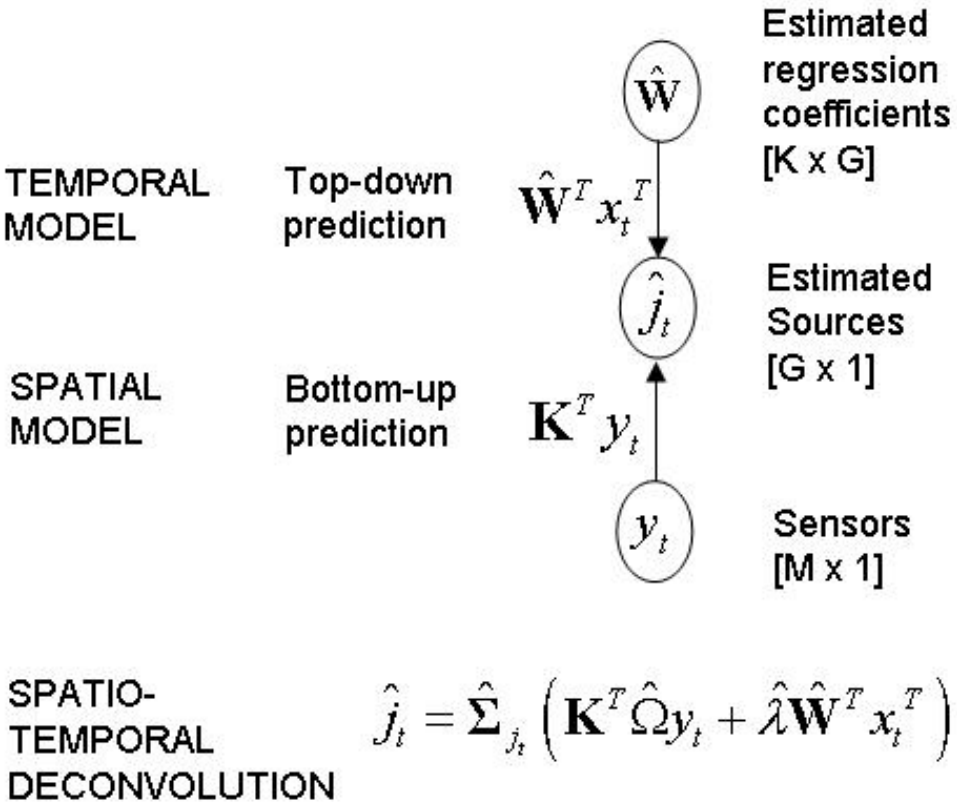

Figure 5: Probabilistic inversion of the generative model leads to a source reconstruction based on a spatio-temporal deconvolution in which bottom-up and top-down predictions, from sensor data and temporal priors, are optimally combined using Bayesian inference.

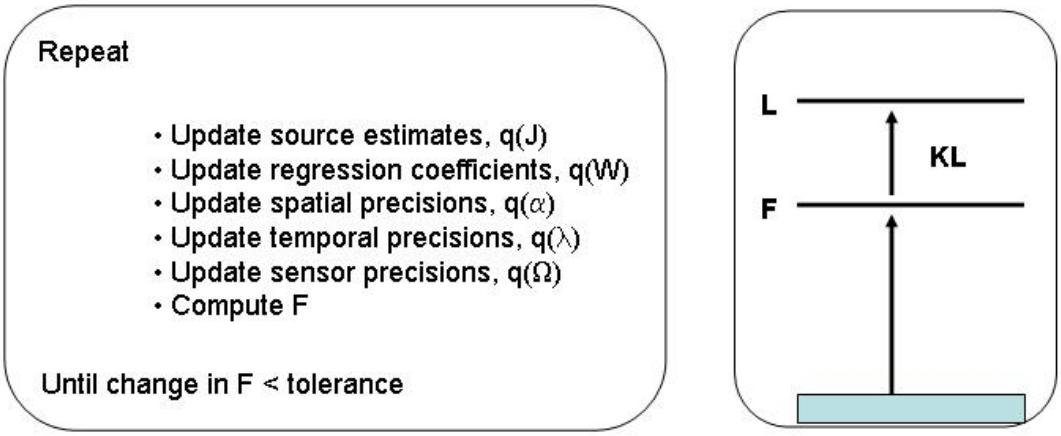

Figure 6: Pseudo code for spatio-temporal deconvolution of M/EEG. The parameters of the model $\theta = \{J, W, \Omega, \lambda, \alpha\}$ are estimated by updating the approximate posteriors until the negative free energy is maximised to within a certain tolerance (left panel). At this point, because the log evidence $L = \log p(Y)$ is fixed, the approximate posteriors will best approximate the true posteriors in the sense of KL-divergence (right panel), as described in Chapter 26. The equations for updating the approximate posteriors are given in the theory section.

Comparison with LORETA

We generated data from our model as follows. Firstly, we created two regressors consisting of a 10 Hz and 20Hz sinewave with amplitudes of 10 and 8 respectively. These formed the two columns of a design matrix shown in Figure 7. We generated 600ms of activity with a sample period of 5ms, giving 120 time points.

The sensor space was defined using $M = 32$ electrodes from the Brain Electrical Source Activity (BESA) system [Scherg and von Cramon 1986] shown in Figure 8. We used the three concentric sphere model to calculate the electric lead field [Rush and Driscoll 1969]. The center and radius of the spheres were fitted to the scalp, skull and cerebral tissue of a 'typical' brain from the Montreal Neurological Institute (MNI) data base [Evans et al. 1993]. The source space consisted of a mesh of points corresponding to the vertices of the triangles obtained by tessellation of the cortical surface of the same brain. A medium resolution spatial grid was used containing $G = 10,242$ points.

We define the Signal to Noise Ratio (SNR) as the ratio of the signal standard deviation to noise standard deviation and used sensor and source SNRs of 10 and 40 respectively. The spatial distribution of the two regression coefficients were identical, each of them consisting of two Gaussian blobs with a maximum amplitude of 10, and a Full Width at Half Maximum (FWHM) of 20mm.

Figure 9 shows the true and estimated sources at time point $t = 20$ms. The LORETA solution was found from an instantaneous reconstruction of the sensor data at that time point, using an L^2-norm and a spatial regularisation parameter $\hat{\alpha}$ (see equation 1) estimated using generalised cross-validation. The VB-GLM solution was found by applying the VB update equations described in the Theory section. As expected, VB provides a better solution both in terms of localisation accuracy and scaling.

ERP simulation

We then used our generative model to simulate ERP-like activity by using the regressors shown in Figure 10. The first regressor mimics an early component and the second a later component. These regressors were derived from a neural mass model describing activity in a distributed network of cortical areas [David and Friston 2003], which lends these simulations a degree of biological plausibility. These neural mass models are described at length in Chapter 32.

We then specified two source activations with the same amplitude and FWHM as in the previous example. The source space, sensor space and forward model were also identical to the previous example. Ten trials of sensor data were then generated using the same SNR as in the previous set of simulations. Signal epochs of 512ms were produced with a sampling period of 4ms giving a total of 5120ms of EEG. The data were then averaged over trials to calculate the sample ERP shown in Figure 11.

We then estimated the sources underlying the sample ERP with (i) a correctly specified model using the same two regressors used for generating the data and (ii) an over-specified model that also incorporated two additional spurious regressors shown in Figure 12. The design matrices for each of these models are shown in Figure 13. In the over-specified model, regressors 2 and 3 are highly correlated ($r = 0.86$). This can be seen most clearly in Figure 13.

The models were then fitted to the data using the VB update rules. As

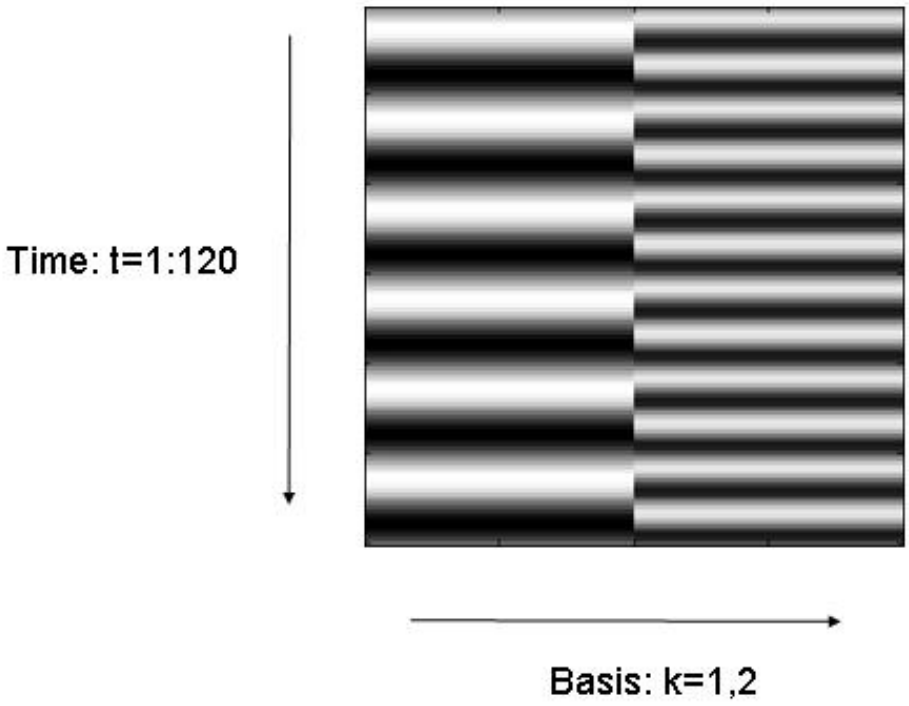

Figure 7: Simulations that compare VB-GLM with LORETA used the above design matrix, X. The columns in this matrix comprise a 10Hz and a 20Hz sinewave.

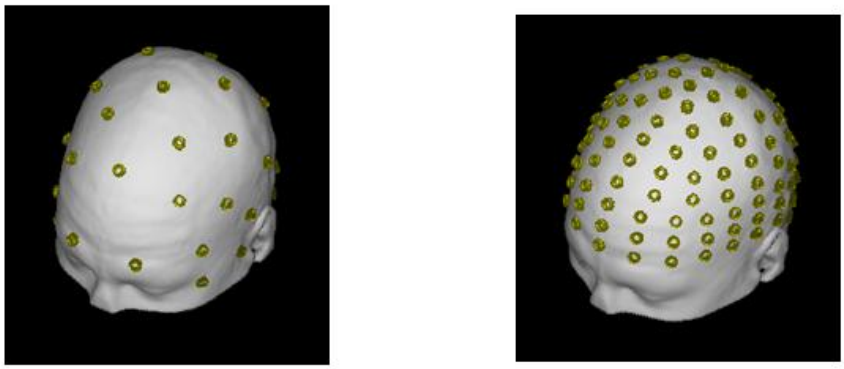

Figure 8: Electrode positions for the 32-sensor BESA system (left) and 128-sensor BioSemi system (right).

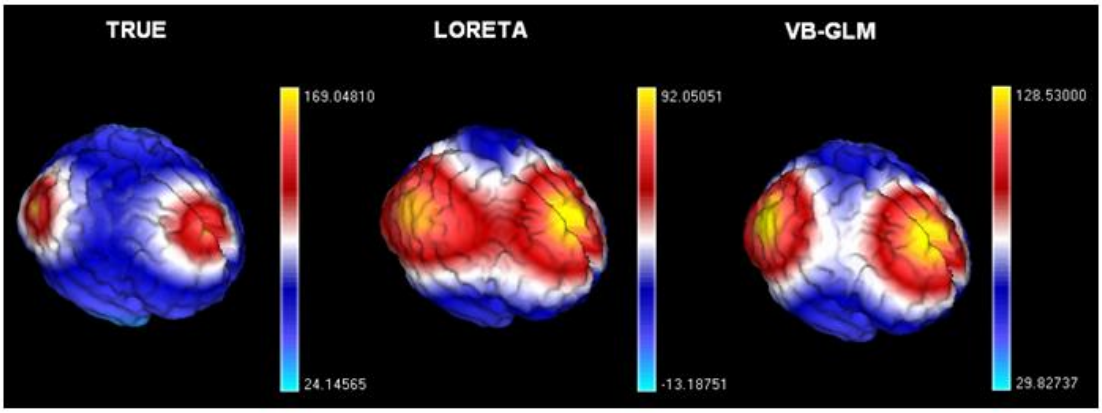

Figure 9: True and estimated source distributions at time $t = 20$ms. Note the scaling in the figures. The VB-GLM approach is better both in terms of spatial localisation and the scaling of source estimates.

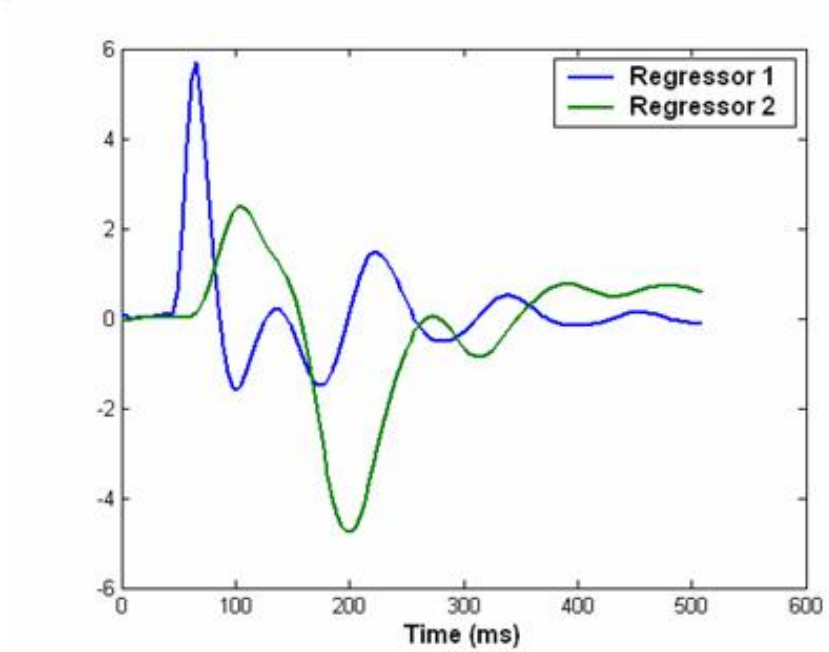

Figure 10: Two ERP components, derived from a biophysical model, used to generate simulated ERP data. These mimic an early component and a late component.

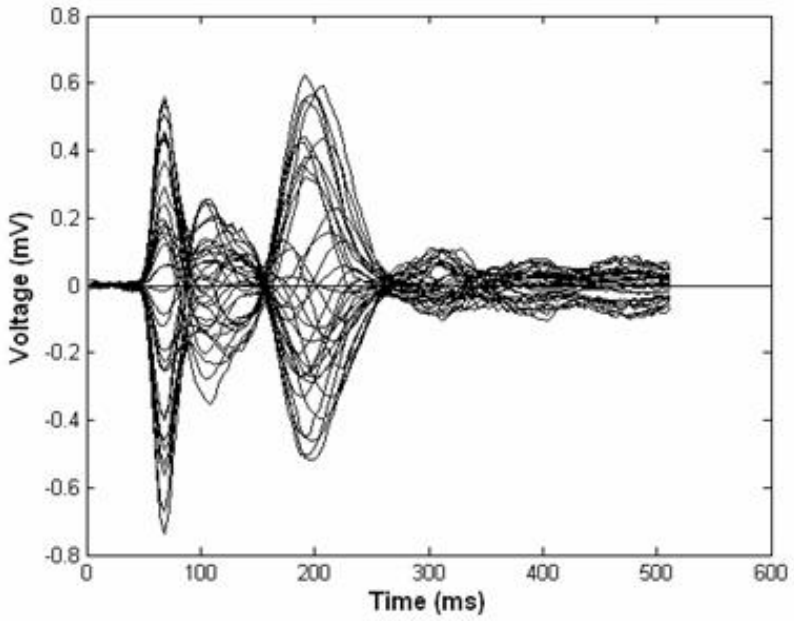

Figure 11: A butterfly plot of simulated ERPs at 32 sensors.

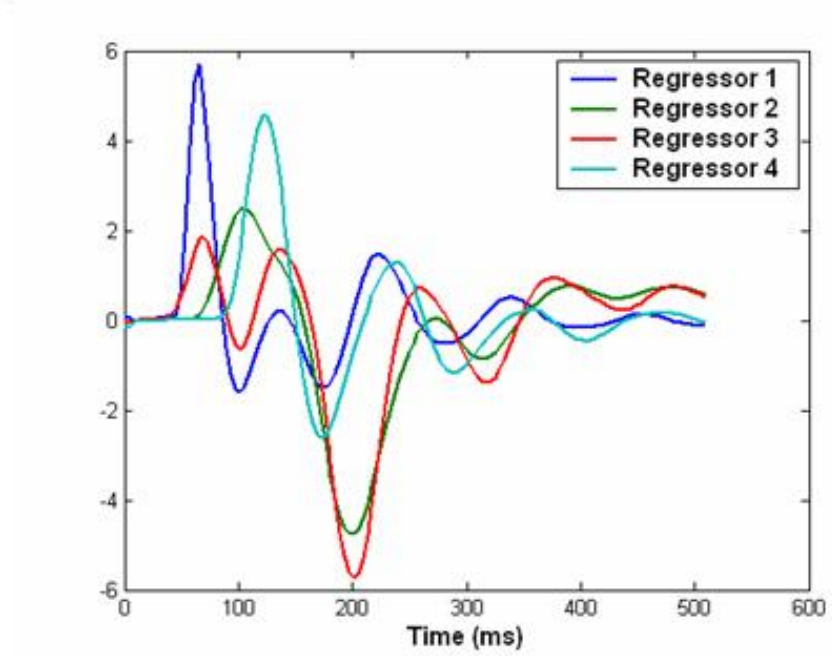

Figure 12: Four components, derived from a biophysical model, used in an over-specified ERP model.

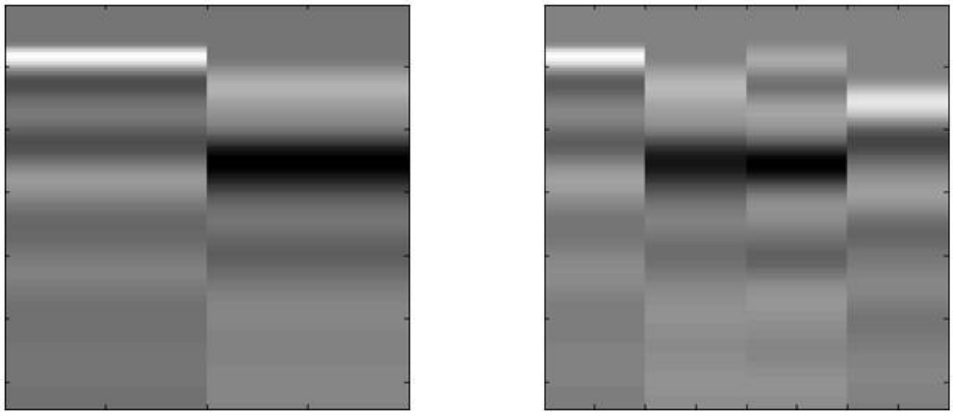

Figure 13: Design matrices, X, used for localisation of biophysical components. Model 1 (left) contains the regressors used to generate the data and Model 2 (right) contains two additional spurious regressors. These regressors have been plotted as time series in Figures 10 and 12.

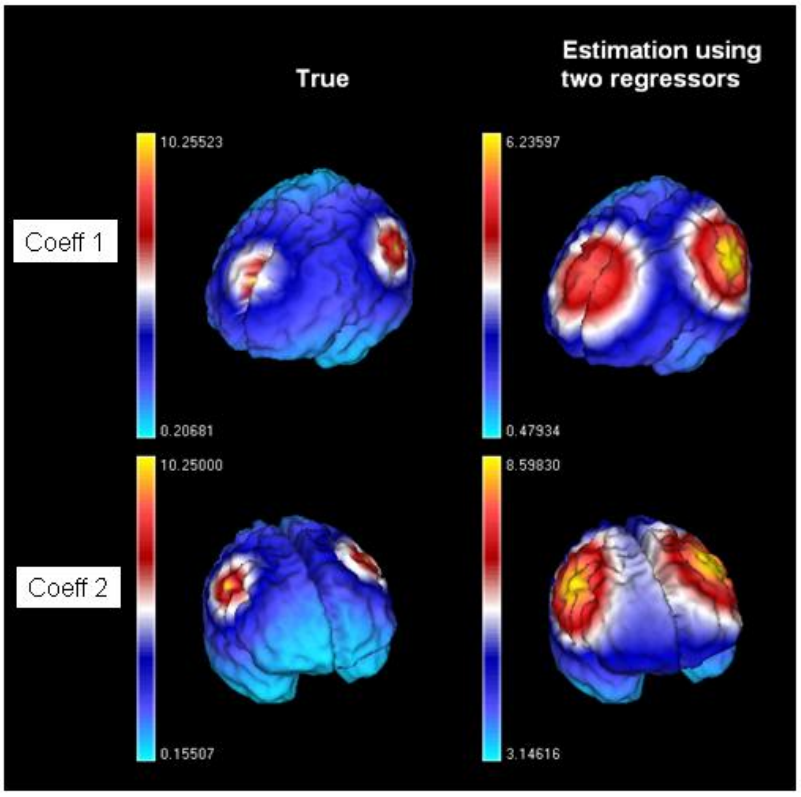

Figure 14: Regression coefficients, w_g, from ERP simulation. 'Coeff 1' and 'Coeff 2' denote the first and second entries in the regression coefficient vector w_g. True model (left) and estimates from correctly specified model (right).

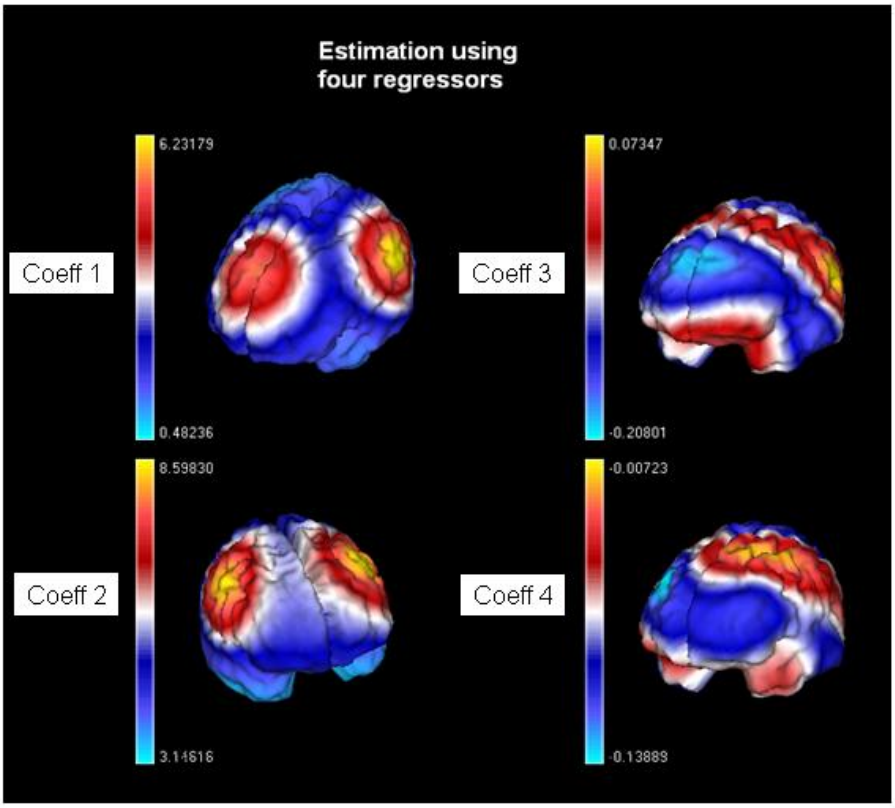

Figure 15: Estimated regression coefficients, $\hat{w}_g$, from over-specified model. The true coefficients are shown in Figure 14. Note the scaling of coefficients 3 and 4 (the true values are zero). Despite the high temporal correlation between regressors 2 and 3, the coefficients for regressor 3 have been correctly shrunk towards zero. This is a consequence of the spatial prior and the iterative nature of the spatio-temporal deconvolution (see Figure 6).

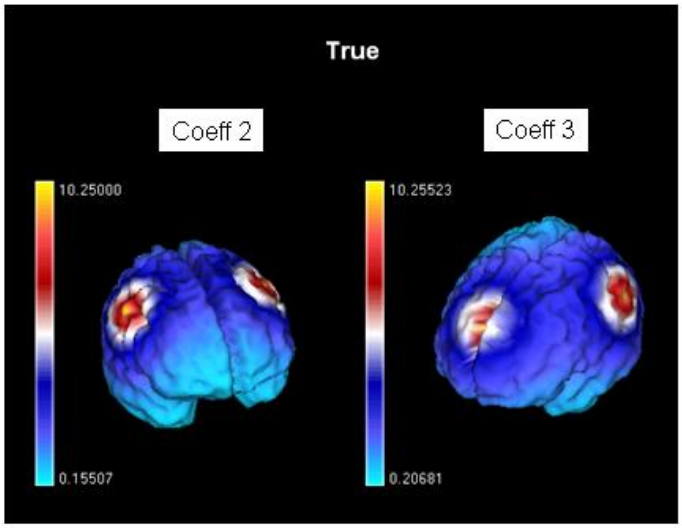

Figure 16: True regression coefficients for ERP simulation with correlated sources. This simulation used a design matrix comprising the regressors shown in Figure 12, with the first and fourth coefficients set to zero and the second and third set as shown in this figure.

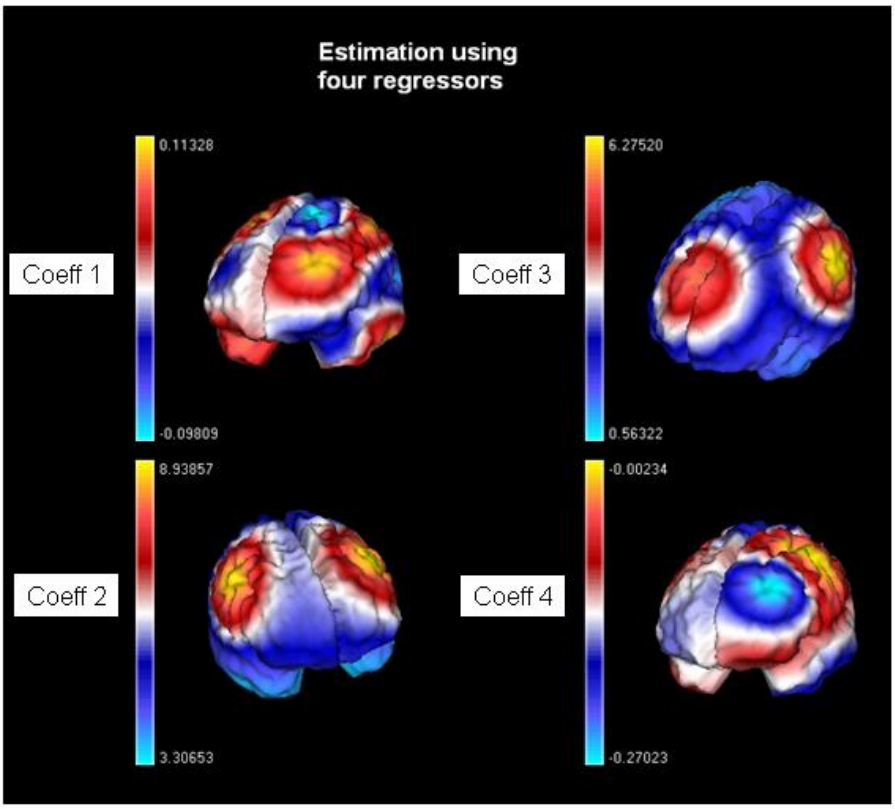

Figure 17: Estimated regression coefficents, $\hat{w}_g$, for ERP simulation with correlated sources. Coefficients 2 and 3 resemble the true values shown in Figure 16 whereas regressors 1 and 4 have been correctly shrunk towards zero by the spatio-temporal deconvolution algorithm.

shown in Figures 14 and 15, the true effects (regression coefficients) are accurately recovered even for the over-specified model. The spurious regression coefficients are shrunk towards zero. This is a consequence of the spatial prior and the iterative spatio-temporal deconvolution. This also shows that source reconstruction with temporal priors is robust to model mis-specification.

We then performed a second simulation with the set of regressors shown in Figure 12, and identical specifications of source space, sensor space, forward model and SNR. But in this example we generated data from the regression coefficients shown in Figure 16, regression coefficients one and four being set to zero. This data therefore comprises three distributed sources (i) a right-lateralised source having time series given by a scaled, noise-corrupted regressor 2, (ii) a frontal source given by a scaled, noise-corrupted regressor 3 and (iii) a left-lateralised source comprising a noisy, scaled mixture of regressors 2 and 3. These sources are therefore highly correlated.

The VB-GLM model, using a full design matrix comprising all four regressors, was then fitted to this data. The estimated regression coefficients are shown in Figure 17. Regressors 1 and 4 have been correctly estimated to be close to zero whereas regressors 2 and 3 bear a close resemblance to the true values. This shows that VB-GLM, in contrast to eg. beamforming approaches, is capable of localising temporally correlated sources.

Face ERPs

This section presents an analysis of a face processing EEG data set from Henson et al. [Henson et al. 2003]. The experiment involved presentation of images of faces and scrambled faces, as described in Figure 18.

The EEG data were acquired on a 128-channel BioSemi ActiveTwo system, sampled at 1024 Hz. The data were referenced to the average of left and right earlobe electrodes and epoched from -200ms to +600ms. These epochs were then examined for artifacts, defined as timepoints that exceeded an absolute threshold of 120 microvolts. A total of 29 of the 172 trials were rejected. The epochs were then averaged to produce condition specific ERPs at each electrode.

The first clear difference between faces and scrambled faces is maximal around 160ms, appearing as an enhancement of a negative component (peak 'N160') at occipito-temporal channels (eg. channel 'B8'), or enhancement of a positive peak at Cz (eg channel 'A1'). These effects are shown as a differential topography in Figure 19 and as time series in Figure 20.

A temporal model was then fitted using wavelet shrinkage [Donoho and Johnstone 1994]. Before applying the model, the data were first downsampled and the 128 samples following stimulus onset were extracted. These steps were taken as we used WaveLab [1] to generate the wavelet bases. This uses a pyramid algorithm to compute coefficients, so requiring the number of samples to be a power of two.

We then extracted the first eigenvector of the sensor data using a Singular Value Decomposition (SVD) and fitted wavelet models to this time series. A number of wavelet bases were examined, two samples of which are shown in Figure 4 . These are the Daubechies-4 and Battle-Lemarie-3 wavelets. Figure 21 shows the corresponding time series estimates. These employed $K = 28$ and $K = 23$ basis functions respectively, as determined by application of the wavelet

[1] WaveLab is available from http://www-stat.stanford.edu/wavelab.

shrinkage algorithm [Donoho and Johnstone 1994]. We used the smaller Battle-Lemarie basis set in the source reconstruction that follows.

ERPs for faces and scrambled faces were then concatenated to form a vector of 256 elements at each electrode. The overall sensor matrix Y was then of dimension 256×128. The design matrix, of dimension 256×46, was created by having identical block diagonal elements each comprising the Battle-Lemarie basis. This is shown in Figure 22. The source space was then defined using a medium resolution cortical grid defined using the typical MNI brain, as in the previous sections. Current source orientations were assumed perpendicular to the cortical surface. Constraining current sources based on a different individuals anatomy is clearly sub-optimal, but nevertheless allows us to report some qualitative results.

We then applied the source reconstruction algorithm and obtained a solution after twenty minutes of processing. Figure 23 shows differences in the source estimates for faces minus scrambled faces at time $t = 160$ms. The images show differences in absolute current at each voxel. They have been thresholded at 50% of the maximum difference at this time point. The maximum difference is plotted in red and 50% of the maximum difference in blue. At this threshold four main clusters of activation appear at (i) right fusiform, (ii) right anterior temporal, (iii) frontal and (iv) superior centro-parietal.

These activations are consistent with previous fMRI [Henson et al. 2003] and MEG analyses of faces minus scrambled faces in that face processing is lateralised to the right hemisphere and in particular to fusiform cortex. Additionally, the activations in temporal and frontal regions, although not significant in group random effects analyses, are nonetheless compatible with observed between- subject variability on this task.

Discussion

This Chapter has described a model-based spatio-temporal deconvolution approach to source reconstruction. Sources are reconstructed by inverting a forward model comprising a temporal process as well as a spatial process. This approach relies on the fact that EEG and MEG signals are extended in time as well as in space.

It rests on the notion that MEG and EEG reflect the neuronal activity of a spatially distributed dynamical system. Depending on the nature of the experimental task, this activity can be highly localised or highly distributed and the dynamics can be more, or less, complex. At one extreme, listening for example to simple auditory stimuli produces brain activations that are highly localised in time and space. This activity is well described by a single dipole located in brainstem and reflecting a single burst of neuronal activity at eg. t=20ms post-stimulus. More complicated tasks, such as oddball paradigms, elicit spatially distributed responses and more complicated dynamics that can appear in the ERP as damped sinusoidal responses. In this Chapter we have taken the view that by explictly modelling these dynamics one can obtain better source reconstructions.

This view is not unique within the source reconstruction community. Indeed, there have been a number of approaches that also make use of temporal priors. Baillet and Garnero [Baillet and Garnero 1997], in addition to consid-

ering edge-preserving spatial priors, have proposed temporal priors that penalise quadratic differences between neighboring time points. Schmidt et al. [Schmidt et al. 2000] have extended their dipole-like modelling approach using a temporal correlation prior which encourages activity at neighboring latencies to be correlated. Galka et al. [Galka et al. 2004] have proposed a spatiotemporal Kalman filtering approach which is implemented using linear autoregressive models with neighborhood relations. This work has been extended by Yamashita et al. [Yamashita et al. 2004] who have developed a 'Dynamic LORETA' algorithm in which the Kalman filtering step is approximated using a recursive penalised least squares solution. The algorithm is, however, computationally costly, taking several hours to estimate sources in even low-resolution source spaces.

Compared to these approaches, our algorithm perhaps embodies stronger dynamic constraints. But the computational simplicity of fitting GLMs, allied to the efficiency of variational inference, results in a relatively fast algorithm. Also, the GLM can accomodate damped sinusoidal and wavelet approaches that are ideal for modelling transient and nonstationary responses.

The dynamic constraints implicit in our model help to regularize the solution. Indeed, with M sensors, G sources, T time points and K temporal regressors used to model an ERP, if $K < MT/G$ the inverse problem is no longer underdetermined. In practice, however, spatial regularisation will still be required to improve estimation accuracy.

This paper has described a spatio-temporal source reconstruction method embodying well known phenomenological descriptions of ERPs. A similar method has recently been proposed in [Friston et al. 2006] (see also Chapter 30), but the approaches are different in a number of respects. First, in [Friston et al. 2006] scalp data Y are (effectively) projected onto a temporal basis set X and source reconstructions are made in this reduced space. This results in a computationally efficient procedure based on Restricted Maximum Likelihood (ReML), but one in which the fit of the temporal model is not taken into account. This will result in inferences about W and J which are over-confident. If one is simply interested in population inferences based on summary statistics (ie. $\hat{W}$) from a group of subjects, then this does not matter. If, however, one wishes to make within-subject inferences then the procedure described in this chapter is the preferred approach. Second, in [Friston et al. 2006] the model has been augmented to account for trial-specific responses. This treats each trial as a 'random effect' and provides a method for making inferences about induced responses. The algorithm described in this chapter, however, is restricted to treating trials as fixed effects. This mirrors standard first-level analyses of fMRI in which multiple trials are treated by forming concatenated data and design matrices.

A further exciting recent development in source reconstruction is the application of Dynamic Causal Models (DCMs) to M/EEG. DCMs can also be viewed as providing spatio-temporal reconstructions, but ones where the temporal priors are imposed by biologically informed neural mass models. This offers the possibility of making inferences about task-specific changes in the synaptic efficacy of long range connections in cortical hierarchies, directly from imaging data. These developments are described in Chapter 43.

References

[Auranen et al. 2005] T. Auranen, A. Nummenmaa, M. Hammalainen, I. Jaaskelainen, J. Lampinen, A. Vehtari, and M. Sams. Bayesian analysis of the neuromagnetic inverse problem with l^p norm priors. *Neuroimage* 26(3),870–884, 2005.

[Baillet and Garnero 1997] S. Baillet and L. Garnero. A Bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem. *IEEE Transactions on Biomedical Engineering*, pages 374–385, 1997.

[Baillet et al. 2001] S. Baillet, J.C. Mosher, and R.M. Leahy. Electromagnetic Brain Mapping. *IEEE Signal Processing Magazine*, pages 14–30, November 2001.

[Tallon-Baudry et al. 1996] C. Tallon Baudry, O. Bertrand, C. Delpuech, and J. Pernier. Stimulus specificity of phase-locked and non phase-locked 40hz visual responses in human. *The Journal of Neuroscience*, 16(13):4240–4249, 1996.

[Breakspear and Terry 2002] M. Breakspear and J.R. Terry. Nonlinear interdependence in neural systems: motivation, theory and relevance. *International Journal of Neuroscience*, 112:1263–1284, 2002.

[Brookes et al. 2004] M. Brookes, A. Gibson, S. Hall, P. Furlong, G. Barnes, A. Hillebrand, K. Singh, I. Halliday, S. Francis, and P. Morris. A general linear model for MEG beamformer imaging. *Neuroimage*, 23(3):936–946, 2004.

[Clyde et al. 1998] M. Clyde, G. Parmigiani, and B. Vidakovic. Multiple shrinkage and subset selection in wavelets. *Biometrika*, 85:391–402, 1998.

[Darvas et al. 2004] F. Darvas, D. Pantazis, E. Kucukaltun Yildirim, and R. Leahy. Mapping human brain function with MEG and EEG: methods and validation. *Neuroimage*, 2004.

[David and Friston 2003] O. David and K.J. Friston. A neural mass model for MEG/EEG: coupling and neuronal dynamics. *NeuroImage*, 20(3):1743–1755, 2003.

[Demiralp et al. 1998] T. Demiralp, A. Ademoglu, Y. Istefanopoulos, and H.O. Gulcur. Analysis of event-related potentials (ERP) by damped sinusoids. *Biological Cybernetics*, 78:487–493, 1998.

[Donoho and Johnstone 1994] D.L. Donoho and I.M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81:425–455, 1994.

[Evans et al. 1993] A. Evans, D. Collins, S. Mills, E. Brown, R. Kelly, and T. Peters. 3d statistical neuroanatomical models from 305 mri volumes. *Proc. IEEE Nuclear Science Symposium and Medical Imaging Conference*, 95:1813–1817, 1993.

[Frackowiak et al. 2003] R.S.J. Frackowiak, K.J. Friston, C. Frith, R. Dolan, C.J. Price, S. Zeki, J. Ashburner, and W.D. Penny. *Human Brain Function.* Academic Press, 2nd edition, 2003.

[Friston et al. 2000] K.J. Friston, A. Mechelli, R. Turner, and C.J. Price. Nonlinear responses in fMRI: The Balloon model, Volterra kernels and other hemodynamics. *NeuroImage*, 12:466–477, 2000.

[Friston et al. 2006] K.J. Friston, R.N.A. Henson, C. Phillips, and J. Mattout. Bayesian estimation of evoked and induced responses. Neuroimage, accepted for publication, 2006.

[Fuchs et al. 1999] M. Fuchs, M. Wagner, T. Kohler, and H. Wischman. Linear and nonlinear current density reconstructions. *Journal of Clinical Neurophysiology*, 16(3):267–295, 1999.

[Galka et al. 2004] A. Galka, O. Yamashita, T. Ozaki, R.Biscay, and P. Valdes-Sosa. A solution to the dynamical inverse problem of EEG generation using spatiotemporal Kalman filtering. *NeuroImage*, 23(2):435–453, 2004.

[Gelman et al. 1995] A. Gelman, J.B. Carlin, H.S. Stern, and D.B. Rubin. *Bayesian Data Analysis*. Chapman and Hall, Boca Raton, 1995.

[Golub and Van Loan 1996] G.H. Golub and C.F. Van Loan. *Matrix Computations*. John Hopkins University Press, 3rd edition, 1996.

[Henson et al. 2003] R.N.A. Henson, Y. Goshen-Gottstein, T. Ganel, L.J. Otten, A. Quayle, and M.D. Rugg. Electrophysiological and hemodynamic correlates of face perception, recognition and priming. *Cerebral Cortex*, 13:793–805, 2003.

[Kiebel and Friston 2004] S.J. Kiebel and K.J. Friston. Statistical Parametric Mapping for Event-Related Potentials II: A Hierarchical Temporal Model. *NeuroImage*, 22(2):503–520, 2004.

[Makeig et al. 2002] S. Makeig, M. Westerfield, T.-P. Jung, S. Enghoff, J. Townsend, E. Courchesne, and T.J. Sejnowski. Dynamic brain sources of visual evoked responses. *Science*, 295:690–694, 2002.

[Pascual-Marqui et al. 1994] R. Pascual Marqui, C. Michel, and D. Lehman. Low resolution electromagnetic tomography: a new method for localizing electrical activity of the brain. *International Journal of Psychophysiology*, pages 49–65, 1994.

[Pascual-Marqui 2002] R. Pascual Marqui. Standardized low resolution electromagnetic tomography (sLORETA): technical details. *Methods and Findings in Experimental and Clinical Pharmacology*, 24:5–12, 2002.

[Mattout et al. 2006] J. Mattout, C. Phillips, W.D. Penny, M. Rugg, and K.J. Friston. MEG source localisation under multiple constraints: an extended Bayesian framework. *NeuroImage*, 30(3):753, 2006.

[McKeown et al. 1998] M.J. McKeown, S. Makeig, G.G. Brown, T.P. Jung, S.S. Kindermann, A.J. Bell, and T.J. Sejnowski. Analysis of fMRI data by blind separation into independent spatial components. *Human Brain Mapping*, 6:160–188, 1998.

[Minka 2000] T.P. Minka. Automatic choice of dimensionality for PCA. Technical Report 514, MIT Media Laboratory, Perceptual Computing Section, 2000.

[Mosher and Leahy 1998] J.C. Mosher and R.M. Leahy. Recursive MUSIC: a framework for eeg and meg source localization. *IEEE Transactions on Biomedical Engineering*, 47:332–340, 1998.

[Osborne and Smyth 1991] M.R. Osborne and G. K. Smyth. A modified Prony algorithm for fitting functions defined by difference equations. *Journal of Scientific and Statistical Computing*, 12:362–382, 1991.

[Penny and Flandin 2005] W.D. Penny and G. Flandin. Bayesian analysis of single-subject fMRI: SPM implementation. Technical report, Wellcome Department of Imaging Neuroscience, 2005.

[Penny et al. 2005] W.D. Penny, N. Trujillo-Barreto, and K.J. Friston. Bayesian fMRI time series analysis with spatial priors. *NeuroImage*, 24(2):350–362, 2005.

[Roberts and Penny 2002] S.J. Roberts and W.D. Penny. Variational Bayes for Generalised Autoregressive models. *IEEE Transactions on Signal Processing*, 50(9):2245–2257, 2002.

[Robinson and Vrba 1999] S. Robinson and J. Vrba. Functional neuroimaging by Synthetic Aperture Magnetometry (SAM). In *Recent Advances in Biomagnetism*, Sendai, Japan, 1999. Tohoku University Press.

[Rugg and Coles 1995] M.D. Rugg and M.G.H. Coles. *Electrophysiology of Mind: Event-Related Potentials and Cognition*. Oxford University Press, 1995.

[Rush and Driscoll 1969] S. Rush and D. Driscoll. EEG electrode sensitivity – an application of reciprocity. *IEEE Transactions on Biomedical Engineering*, 16(1):15–22, 1969.

[Sahani and Nagarajan 2004] M. Sahani and S.S. Nagarajan. Reconstructing MEG sources with unknown correlations. In L. Saul S. Thrun and B. Schoelkopf, editors, *Advances in Neural Information Processing Systems*, volume 16. MIT, Cambridge, MA, 2004.

[Scherg and von Cramon 1986] M. Scherg and D. von Cramon. Evoked dipole source potentials of the human auditory cortex. *Electroencephalography and Clinical Neurophysiology*, 65:344–360, 1986.

[Schmidt et al. 1999] D.M. Schmidt, J.S. George, and C.C. Wood. Bayesian inference applied to the electromagnetic inverse problem. *Human Brain Mapping*, 7:195–212, 1999.

[Schmidt et al. 2000] D.M. Schmidt, D.M. Ranken J.S. George, and C.C. Wood. Spatial-temporal Bayesian inference for MEG/EEG. In *12th International Conference on Biomagnetism*, Helsink, Finland, August 2000.

[Trejo and Shensa 1999] L. Trejo and M.J. Shensa. Feature extraction of event-related potentials using wavelets: an application to human performance monitoring. *Brain and Language*, 66:89–107, 1999.

[Unser and Aldroubi 1996] M. Unser and A. Aldroubi. A review of wavelets in biomedical applications. *Proceedings of the IEEE*, 84:626–638, 1996.

[Valdes-Sosa et al. 2000] P. Valdes-Sosa, F. Marti, F. Garcia, and R. Casanova. Variable Resolution Electric-Magnetic Tomography. In *Proceedings of the 10th International Conference on Biomagnetism*, volume 2, pages 373–376, 2000.

[Woolrich et al. 2001] M.W. Woolrich, B.D. Ripley, M. Brady, and S.M. Smith. Temporal autocorrelation in univariate linear modelling of fMRI data. *NeuroImage*, 14(6):1370–1386, December 2001.

[Yamashita et al. 2004] O. Yamashita, A. Galka, T. Ozaki, R. Biscay, and P. Valdes-Sosa. Recursive penalised least squares solution for dynamical inverse problems of EEG generation. *Human Brain Mapping*, (21):221–235, 2004.

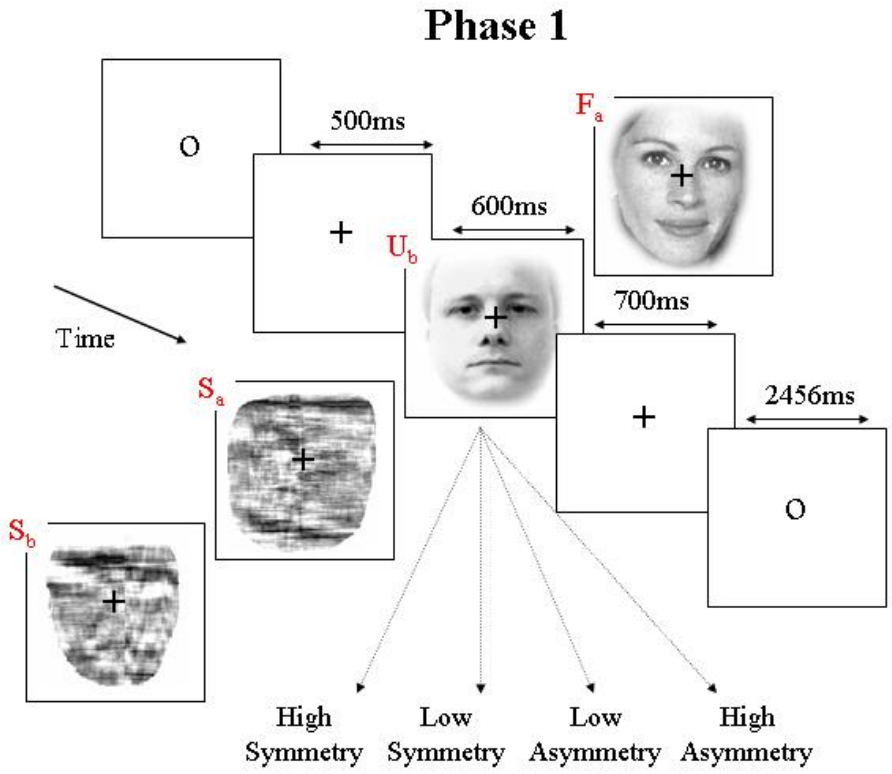

Figure 18: Face paradigm. The experiment involved randomised presentation of images of 86 faces and 86 scrambled faces. Half of the faces belong to famous people, half are novel, creating 3 conditions in total. In this Chapter we consider just two conditions (i) faces (famous or not) and (ii) scrambled faces. The scrambled faces were created by 2D Fourier transformation, random phase permutation, inverse transformation and outline-masking. Thus faces and scrambled faces are closely matched for low-level visual properties such as spatial frequency. The subject judged the left-right symmetry of each image around an imaginary vertical line through the centre of the image. Faces were presented for 600ms, every 3600ms.

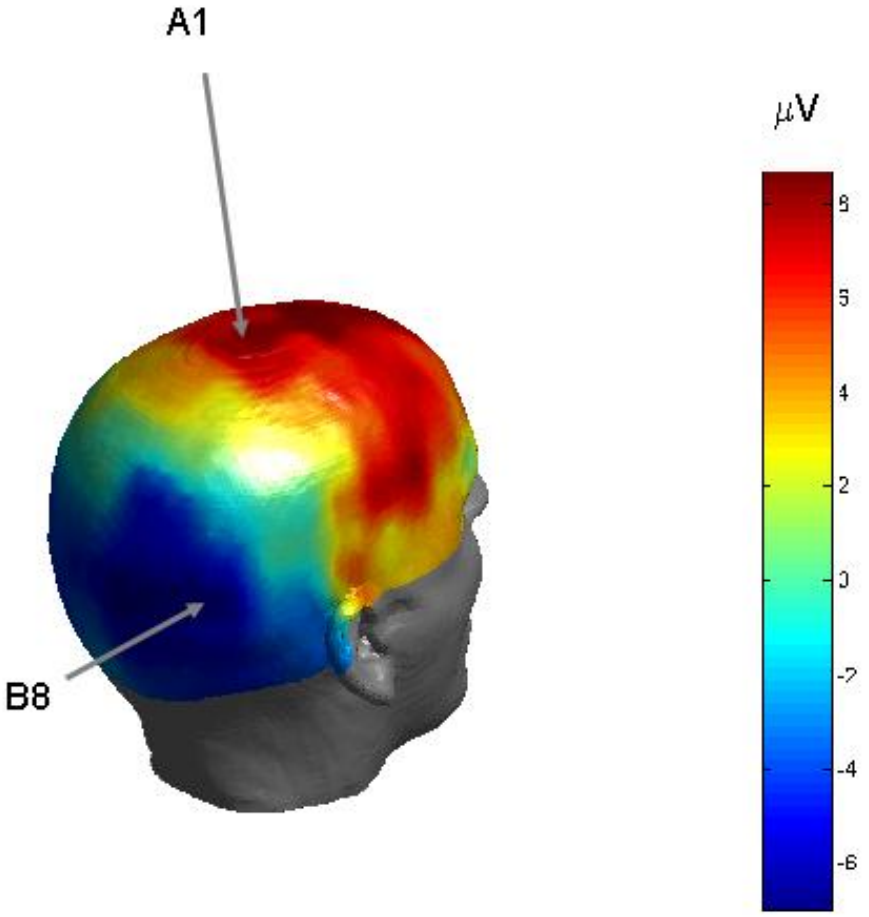

Figure 19: The figure shows differential EEG topography for faces minus scrambled faces at $t = 160$ms poststimulus.

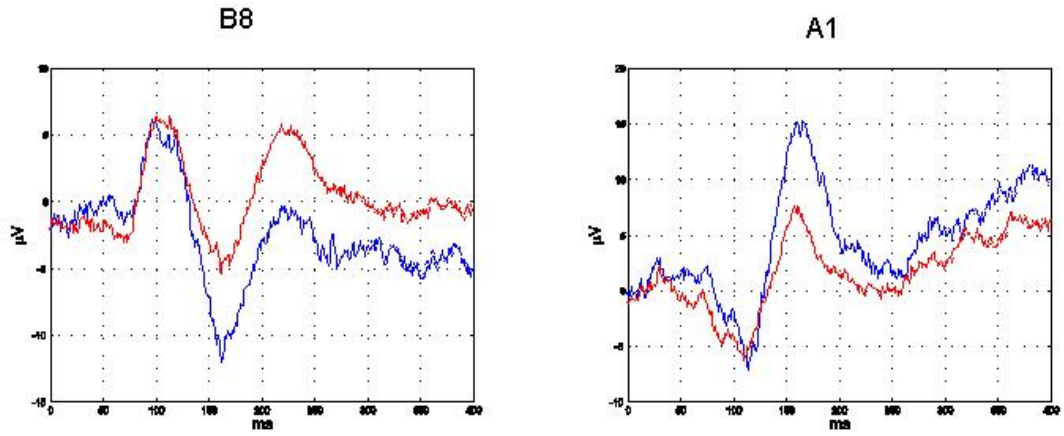

Figure 20: Sensor time courses for face data at occipito-temporal electrode B8 (left) and vertex A1 (right) for faces (blue) and scrambled faces (red).

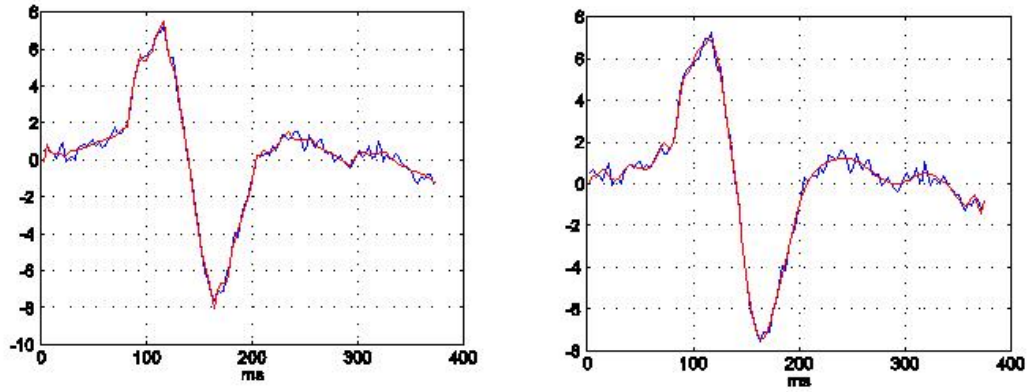

Figure 21: First eigen-timeseries of downsampled ERP for unfamiliar faces (blue lines in both plots) with wavelet shrinkage approximations using Daubechies basis (left) and Battle-Lemarie basis (right).

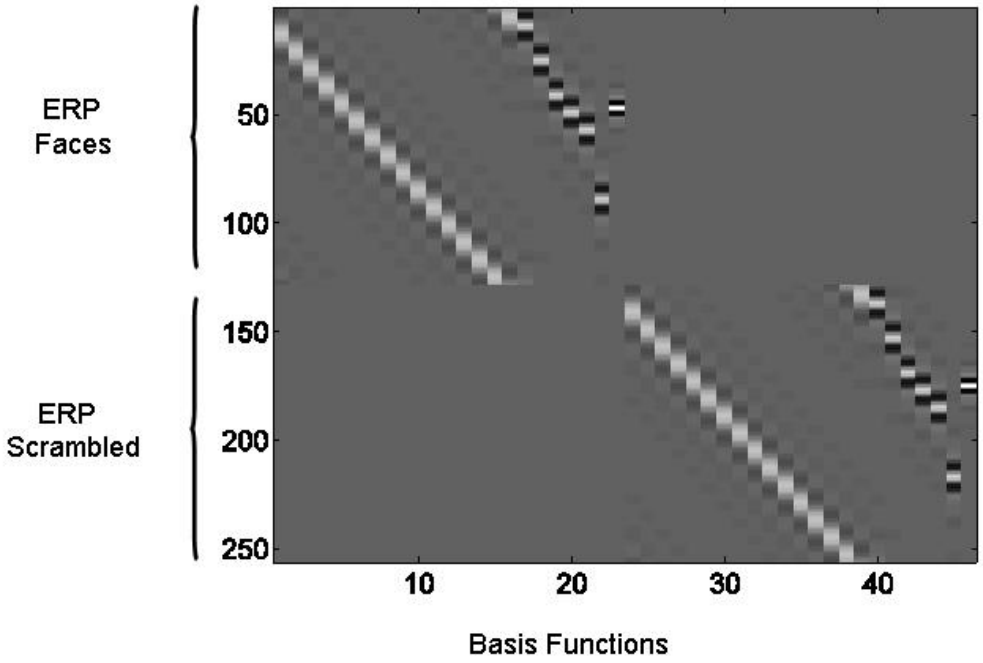

Figure 22: Design matrix for source reconstruction of ERPs from Face data. Each block contains a 23-element Batte-Lemarie basis set. The first components, forming diagonals in the picture, are low frequency wavelets. The high frequency wavelets are concentrated around the N160, where the signal is changing most quickly.

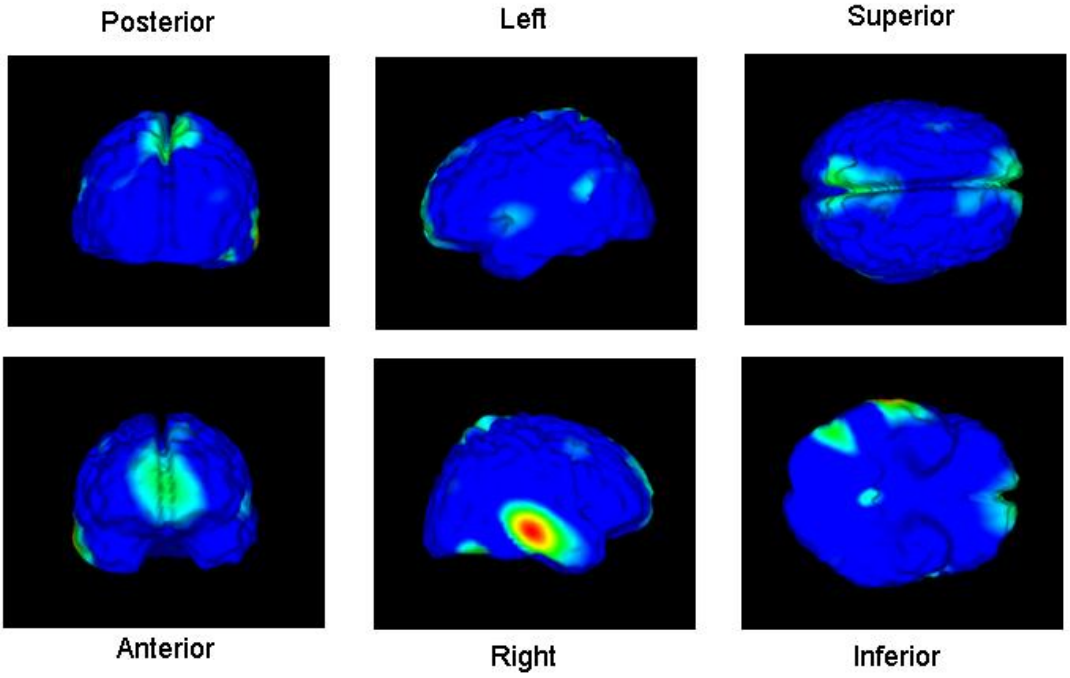

Figure 23: These images are derived from the source reconstruction of ERPs in response to faces and scrambled faces. The plots show absolute differences between faces and scrambled faces at t=160ms post-stimulus. The maps have been thresholded such that the largest difference appears in red and 50% of the largest difference appears in blue.