Activation curves from optimally shaped regions

Hasson R and Swithenby SJ

The Open University, Milton Keynes, United Kingdom

Introduction

The dominant concern in MEG analysis has been to produce source maps and to co-register these to anatomical data [e.g. 1,2]. However, this may not be the most appropriate approach when there is a focus on specific source regions, e.g. the thalamus, fusiform gyrus etc. In these cases it may be more appropriate to generate an activation curve, a graph of the source power in a specified region as a function of time. Several methods of generating activation curves have been proposed [e.g. 3,4,5].

An activation curve can be found simply by integrating the current density over the region of interest [3] but this has two drawbacks. It is computationally inefficient and it leads to an unnecessarily high expected error. This can be traced to the use of small voxels with consequently high expected errors in the current density estimates. In this paper, an optimised method of generating an activation curve for a region of interest is presented. The method is based on the use of quadratic functionals [6]. It involves the calculation of the shape of a single voxel that approximates the region of interest and also minimizes the expected error.

Suppose that the current density, $\vec{j}$, is a square integrable field defined on the brain volume Ω (i.e. $\vec{j} \in L_2(\Omega)$) and that the region of interest is defined by an operator $\hat{X}$ on $\vec{j}$. Then the activation curve is defined to be $A(t) = \langle \vec{j}(t), \hat{X}\vec{j}(t)\rangle$ where $\langle\ ,\ \rangle$ is the following inner product on Ω [5]:

$$\langle \vec{j}_1, \vec{j}_2 \rangle = \int_\Omega \frac{\vec{j}_1 \cdot \vec{j}_2}{\omega(\vec{r})} \mathrm{d}\vec{r} \qquad (1)$$

If the MEG/EEG data vector is $b \in \mathbb{R}^N$ then the method may be expressed in terms of finding an $N \times N$ matrix Y such that

$$\langle \vec{j}, \hat{X}\vec{j} \rangle \simeq \langle b, Yb \rangle \qquad (2)$$

In the methods section, an expression for the matrix Y is derived which minimizes the expected error in $A(t)$. In the results section the method is applied to simulated data. It should be noted that no assumptions are made about the noise in the measurement channels other than it has zero mean and a well defined covariance matrix C, i.e. if the measurement noise is denoted by a vector e then the covariance matrix is defined by $C = \overline{e_i e_k}$ where $\overline{}$ denotes the time average.

Methods

The basic idea is to choose a matrix Y such that $\langle b, Yb\rangle$ is an approximation to the activation curve $A(t)$. The starting point is to discretize the source space (which may be the whole of the brain or a smaller volume) into M voxels of volume V_m where $m = 1, \ldots, M$. Similarly the current density $\vec{j}$ can be discretized on this set of voxels.

$$j_m = \frac{1}{|V_m|} \int_{V_m} \vec{j}(\vec{r})\, \mathrm{d}\vec{r} \qquad \text{where } m = 1, \ldots, M. \qquad (3)$$

In this discretized source space the biomagnetic forward problem can be expressed as $b = Lj + e$, where L is the lead field matrix (a discretized form of the lead fields $\vec{L}_j(\vec{r})$) and e is a vector which represents the measurement error. With these definitions the inner product $\langle b, Yb\rangle$ can be expanded:

$$\langle b, Yb \rangle = \langle Lj + e, Y(Lj + e) \rangle = j^T L^T Y L j + e^T Y L j + j^T L^T Y e + e^T Y e \qquad (4)$$

The task of making this an estimate for $A(t) = \langle \vec{j}, \hat{X}\vec{j}\rangle$ is made easier by reasoning that $A(t)$ is not required for an arbitrary current density $\vec{j}$ but only for the currents that can be measured by the experiment. It follows that Equation 4 should agree with $A(t)$ on the subspace spanned by the lead fields, i.e.

$$\vec{j}(t) = \sum_{i=1}^{N} a_i\, \omega(\vec{r})\, \vec{L}_i(\vec{r}) \qquad (5)$$

for some co-ordinate vector a. The factor $\omega(\vec{r})$ is a weighting factor which allows some flexibility in the procedure. The only restriction imposed on $\omega(\vec{r})$ is that the integral over each voxel is finite. In other papers the factor $\omega(\vec{r})$ has been interpreted as a probability weight [7].

In terms of the discretized source space the relationship between the co-ordinates a and the current estimates j is given by

$$j = WL^T a \qquad \text{where } W \text{ is a diagonal matrix with } W_{mm} = \frac{1}{|V_m|} \int_{V_m} \omega(\vec{r}) \, \mathrm{d}\vec{r} \tag{6}$$

This can be substituted into Equation 4 to get:

$$\langle b, Yb \rangle = a^T LWL^T YLWL^T a + e^T YLWL^T a + a^T LWL^T Ye + e^T Ye \tag{7}$$

where some simplification has been achieved by noting that W is symmetric. This expression can be simplified further by defining the Gram-Schmidt matrix, $P = LWL^T$, to yield

$$\langle b, Yb \rangle = a^T PYPa + e^T YPa + a^T PYe + e^T Ye \tag{8}$$

The next step is to compare this with the operator $\hat{X}$ on the subspace spanned by the lead fields. A matrix X is defined to represent the operator $\hat{X}$ on the subspace spanned by the lead fields:

$$X_{ik} = \langle \omega(\vec{r}) \vec{L}_i(\vec{r}), \hat{X} \omega(\vec{r}) \vec{L}_k(\vec{r}) \rangle \qquad \text{where } i, k = 1, \ldots, N \tag{9}$$

With this definition the activation curve can be calculated as $A(t) = a^T Xa$. Comparing this with Equation 8 gives the condition for $\langle b, Yb \rangle$ to be a good estimate for $A(t)$:

$$a^T Xa \simeq a^T PYPa + e^T YPa + a^T PYe + e^T Ye \tag{10}$$

The task is to choose Y such that the above approximation is valid. One way of achieving this is to minimize the error function E defined by:

$$E = ||X - PYP||_2 + ||e^T YP||_2 + ||PYe||_2 + ||e^T Ye||_2 \tag{11}$$

where $|| \; ||_2$ is the L_2-norm. Equation 11 can be interpreted in physical terms. The first term is the error in approximating the operator $\hat{X}$ by Y. The second and third terms give a measure of the overlap between the measurement error and the imaging source space induced by Y. Note that these terms are equal for a symmetric Y. The fourth term is a measure of how Y magnifies the measurement error in all the signal channels.

To minimize E, $\partial E / \partial Y_{ik}$ is derived for each element of the matrix Y. This gives N^2 equations to solve for the N^2 unknowns Y_{ik}. These may be written as a single matrix equation. In order to illustrate the manipulations involved, the method will be elaborated for the fourth term in Equation 11. The fourth term is expanded using the definition of the L_2-norm:

$$||e^T Ye||_2 = \left(\sum_{\alpha,\beta} e_\alpha Y_{\alpha\beta} e_\beta \right)^2 \tag{12}$$

and is then differentiated:

$$\frac{\partial ||e^T Ye||_2}{\partial Y_{ik}} = 2 \left(\sum_{\alpha,\beta} e_\alpha Y_{\alpha\beta} e_\beta \right) e_i e_k = 2 \sum_{\alpha,\beta} e_i e_\alpha Y_{\alpha\beta} e_\beta e_k \tag{13}$$

At this point the random variables in this expression are replaced with their time averaged values, i.e. $\overline{e_i e_\alpha} = C_{i\alpha}$ and $\overline{e_\beta e_k} = C_{\beta k}$. So

$$\frac{\partial ||e^T Ye||_2}{\partial Y_{ik}} = 2 \sum_{\alpha,\beta} C_{i\alpha} Y_{\alpha\beta} C_{\beta k} \tag{14}$$

This is the ikth term of the matrix product CYC. Similarly, all of the other terms in Equation 11, when differentiated, give terms that can be written as the ikth elements of a product. So, the equations can be collected as:

$$-2PXP + 2P^2 YP^2 + 2P^2 YC + 2CYP^2 + 2CYC = 0 \tag{15}$$

This may be written in the form:

$$(P^2 + C)Y(P^2 + C) = PXP \tag{16}$$

To solve this equation let λ_i be the eigenvalue of the matrix P with eigenvector ϕ_i. Then the matrices X and C can be represented with respect to the basis $\{\phi_i\}$ as new matrices x and c, i.e.

$$X = \sum_{ik} \phi_i x_{ik} \phi_k^T \text{ where } x_{ik} = \phi_i^T X \phi_k, \qquad C = \sum_{ik} \phi_i c_{ik} \phi_k^T \text{ where } c_{ik} = \phi_i^T C \phi_k \tag{17}$$

With these definitions, the matrix Y can be finally expressed as:

$$Y = (P^2 + C)^{-1} P \left(\sum_{ik} \phi_i x_{ik} \phi_k^T \right) P (P^2 + C)^{-1} = \sum_{ik} \phi_i \frac{\lambda_i}{\lambda_i^2 + c_{ii}} x_{ik} \frac{\lambda_k}{\lambda_k^2 + c_{kk}} \phi_k^T \tag{18}$$

Error bars for the activation curve, $A(t)$, can be estimated using the last term in Equation 11 to give the amount of measurement noise reflected in the activation curve. The estimate is given by :

$$\text{noise} = \sum_{\alpha,\beta} C_{\alpha\beta} Y_{\alpha\beta} \tag{19}$$

Results

A simple simulated experimental system (Fig. 1) has been analysed to illustrate the method. The head is modelled as a homogeneously conducting sphere of radius 8.9 cm with its centre at (0,0,-0.07 cm). The source space is a 9 cm×9 cm square thin lamina consisting of 33×33 voxels in the plane z=-0.01 cm with centre (0,0,-0.01 cm). The measurement instrument is a hexagonal array of 37 second order axial gradiometers with baseline 5 cm with the lowest 'sensing' coils in the plane $z = 4$ cm.

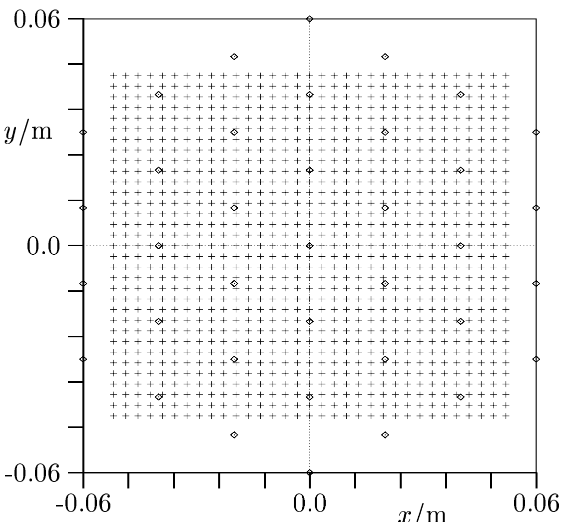

Fig. 1 A plan view of the experiment geometry. Crosses denote source space voxels and diamonds denote the projections of the centres of the detector coils.

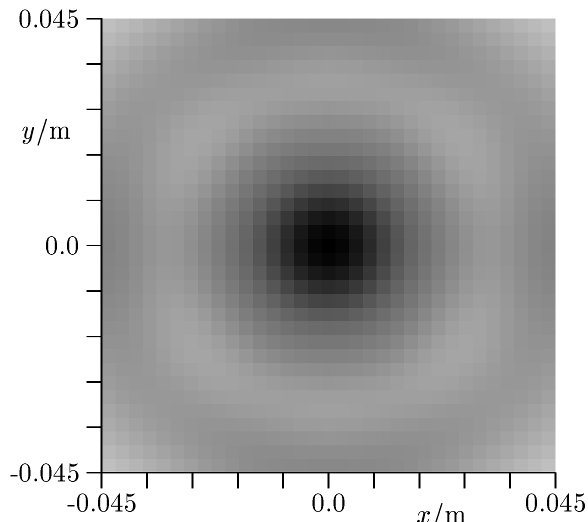

Fig. 2 The sensitivity profile in source space of the Y matrix which is derived from the operator $\hat{X} = \delta(\vec{r} - \vec{r}_0)$.

The spatial selectivity implicit in the use of the matrix Y can be shown by defining and plotting a sensitivity profile in source space. The definition requires the specification of a procedure which may be summarised as follows. Consider a point in source space and compute the idealised measurements b_1, b_2 and b_3 for a unit current dipole oriented along the x, y and x axes at that point. For a given Y, calculate the eigenvalues λ_1, λ_2 and λ_3 of the matrix $E_{ik} = \langle b_i, Y b_k \rangle$. Repeat this process across source space. The sensitivity profile of Y is defined by the variation of $\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}$. It may be thought of as an instrumental generalisation of the lead field of a single measurement channel.

Now consider, in the context of the simulated system, the simplest possible region of interest operator $\hat{X} = \delta(\vec{r} - \vec{r}_0)$ where $\delta(\ )$ denotes the Dirac delta function and $\vec{r}_0$=(0,0,-0.01 cm) is the centre of source space. This operator might be adopted if one simply wished to focus on a small volume of source space. The matrix Y used as an estimator from this operator is calculated using Equation 18. The sensitivity profile for this Y matrix is shown in Fig. 2.

The reconstruction of an activation curve has been tested on simulated data using this region of interest operator and simulated data from a time varying target dipole at (0,0,0 cm), i.e. 1 cm from the region of interest. The moment of the dipole varies sinusoidally at 10 Hz, with an envelope that rises linearly from zero at 200 ms to a maximum at 300 ms after which it remains constant. To show the insensitivity to dipole orientation the dipole moment was made to rotate smoothly in a tangential plane — this rotation is not discernible in the activation curve. In addition to the target dipole there is distractor dipole at (0,0.02,0 cm), which is active from 0 to 100 ms (triangular envelope) and again from 400 ms (square envelope).

In the period from 200 ms to 400 ms when only the target dipole is active, the calculated (power) activation curve matches closely that of the target. However, the existence of the distractor dipole within the sensitive region (see Fig. 2) gives rise to apparent activity between 0 ms and 100 ms and inaccuracy in the calculated activation curve for the period after 400 ms.

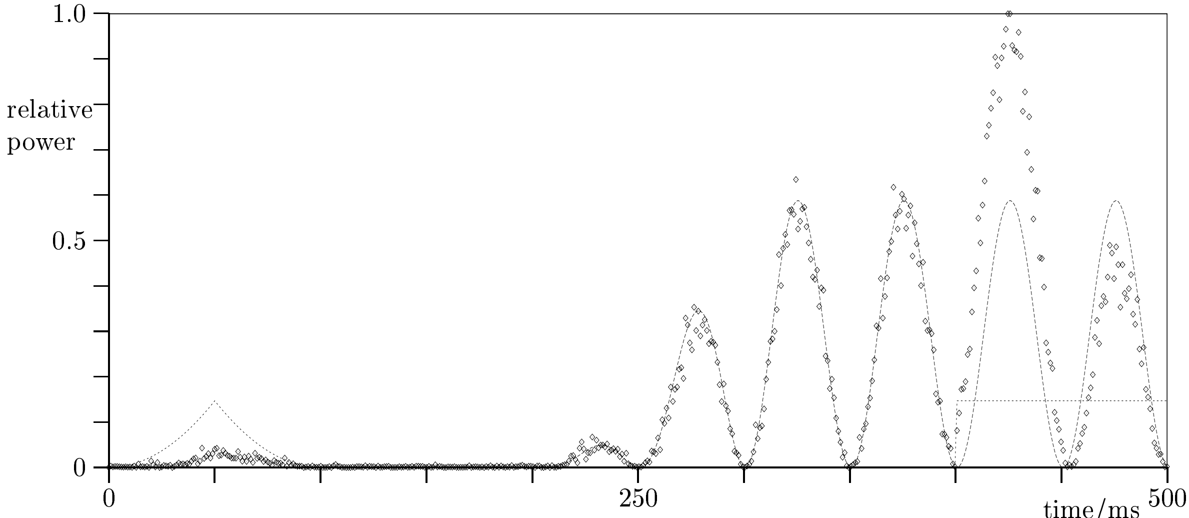

Fig. 3 Activation curves for a simulated experiment. The solid line and the dotted lines are the activation curves of the target and distractor dipoles. The diamonds are the calculated activation curve from the Y matrix shown in Fig. 2. The error bars, omitted for clarity, would be approximately twice the height of the diamonds.

Discussion

We have shown that is possible to directly compute activation curves from the measurement data without integration over a source image. The method is a computationally optimal way of tracking the power dissipated in a specific portion of the brain. It seems robust to noise and is not dependent on the noise having a Gaussian profile. Correlations between measurement channels are fully taken into account.

For evoked response experiments the covariance matrix, C, can be estimated from the prestimulus period. For other experiments it might be more suitable to make the *a priori* assumption that the noise is uncorrelated gaussian noise with variance α^2 which could be considered as a parameter of the method. As α increases, the more closely the Y matrix sensitivity pattern matches the region of interest, but the larger the error bars on the resulting activation curve.

References

[1] Sarvas J: Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. Phys. Med. Biol. 32(1): 11–22, 1987

[2] Schwartz D, Lemoine D, Poisot E and Barillot C: Registration of MEG/EEG data with 3D MRI: methodology and precision issues. Brain Topography 9(2): 101–116, 1996

[3] Singh KD, Ioannides AA, Hasson R., Ribary U, Lado F and Llinas R: Extraction of dynamic patterns from distributed current solutions of brain activity. In M. Hoke et al. (eds.) BIOMAGNETISM: Clinical Aspects, Amsterdam: Elsevier/IOS-Press, 1992, pp 767–773

[4] Tesche CD, Uusitalo MA, Ilmoniemi RJ, Huotilainen M, Kajola M and Salonen O: Signal space projections of MEG data characterise both distributed and well-localised neuronal sources. Electroenceph. Clin. Neurophysiol. 95: 189–200, 1995

[5] Robinson SE and Rose DF: Current source image estimation by spatially filtered MEG. In Hoke M, Erne SN, Okada YC et al (eds.): Biomagnetism: Clinical Aspects, Amsterdam: Elsevier/IOS-Press, 1992, pp 761–765

[6] Grave de Peralta Menendez R, Gonzalez Andino SL, Michel CM: How to deal with the non-uniqueness of the biomagnetic inverse problem: well-posed questions for ill-posed problems. Conference Digest for ISBET'97, 8th World Conf. of the Int. Soc. for Brain Electromagnetic. Tomography: Abstract P–2.16 1997

[7] Hasson R and Swithenby SJ: The theoretical basis of iterative distributed solutions to the biomagnetic inverse problem. In Aine C et al. (eds.) Advances in Biomagnetism Research: Biomag 96, Springer Verlag, Berlin, In press 1998