A Method for High-Resolution Localization of Bio-magnetic Sources in the Brain

Hung-I Pai[1†], Chih-Yuan Tseng[1], H.C. Lee[1,2,3]

[1] *Department of Physics, National Central University, Zhongli, Taiwan 32001*
[2] *Graduate Institute of Systems Biology and Bioinformatics National Central University, Zhongli, Taiwan 32001*
[3] *National Center for Theoretical Sciences, Hsinchu, Taiwan 30043*

Abstract

Magnetoencephalography (MEG) provides dynamic spatial-temporal insight of neural activities in the cortex. Because the number of possible sources is far greater than the number of MEG detectors, the proposition to localize sources directly from MEG data is notoriously ill-posed. Here we develop a method based on data processing procedures including clustering, forward and backward filtering, and the method of maximum entropy. We show that taking as a starting point the assumption that the sources lie in the general area of the auditory cortex (an area of about 40 mm by 15 mm), our approach is capable of achieving reasonable success in pinpointing active sources with a resolution of a few mm and reducing the spatial distribution and number of false positives to a very low level.

Key words: MEG, ill-posed inverse problem, clustering, filtering, maximum entropy

1 Introduction

Magnetoencephalography (MEG) records magnetic fields generated from neurons when the brain is performing a specific function. Neural activities thus can be noninvasively studied through analyzing the MEG data. Since the number of neurons (unknowns) are far larger than the number of MEG sensors (knowns) outside the brain, the problem of identifying activated neurons from the magnetic data is ill posed. The problem becomes even more severe when noise is present.

A first and essential step in surpassing the obstacle of ill-posedness is to rely on prior knowledge of the general area of active current sources producing the

MEG data. Often, this prior knowledge is provided by functional magnetic resonance imaging (fMRI) experiments. The two kinds of experiments complement each other, MEG has high temporal resolution (about 10^{-3}s) but poor spatial resolution, whereas fMRI has high spatial resolution (C.T.W. Moonen, 1999) but poor temporal resolution (about 1 s). Consider auditory neural activity. FMRI shows that such activity is concentrate in the auditory cortex, two 40 mm by 15 mm areas respectively located on the two sides of the cortex (Fig. 1L). Because of poor temporal resolution, fMRI cannot resolve the rapid successive firing of the (groups of) neurons, instead it shows large areas of the auditory cortex lighting up. Under MEG, the same auditory activity will be detected by high time-resolution sensors (Fig. 1R) which however collectively cover an area (when projected down to the cortex) much greater than the auditory cortex.

Although many methods have been proposed to attack the ill-posed problem, including dipole fitting, minimum-norm-least-square (MNLS) (J.-Z. Wang, 1993), and the maximum entropy method (ME) (E. L. C. Amblard, 2004), improvements have been limited. One reason for this may be that these methods do not properly include anatomic constrains. In this work, we propose a novel approach to analyze noisy MEG data based on ME that pays special attention to obtaining better *priors* as input to the ME procedure by employing clustering and forward and backward filtering processes that take anatomic constrains into consideration. We demonstrate the feasibility of our approach by testing it on several simple cases.

2 The ill-posed inverse problem

Given the set of current sources (dipole strengths) $\{r_i|i = 1, 2, \cdots, N_r\}$ at sites $\{z_i|i = 1, 2, \cdots, N_r\}$, the magnetic field strength $m(\mathbf{x})$ measured by the sensor at spatial position $\mathbf{x}$ is given by,

$$m(\mathbf{x}) = \sum_{i=1}^{N_r} A(\mathbf{x})_i r_i + n(\mathbf{x}) \equiv \hat{A}(\mathbf{x}) \cdot \hat{r} + n(\mathbf{x}) \tag{1}$$

where the function $A(\mathbf{x})_i$, which is a function of z_i, is derived from the Biot-Savart law in vacuum, $n(\mathbf{x})$ is a noise term, and a hatted symbol denotes an N_r-component vector in source space. We consider the case where there are N_m sensors at locations $\mathbf{x}_\alpha$, α=1,2,$\cdots$, N_m. To simplify notation, we write $\mathbf{m}$ as an N_m-component vector whose α^{th} component is m^α=$m(\mathbf{x}_\alpha)$, similarly for $\hat{\mathbf{A}}$ and $\mathbf{n}$. Then Eq. (1) simplifies to

$$\mathbf{m} = \hat{\mathbf{A}} \cdot \hat{r} + \mathbf{n} \tag{2}$$

In practice, magnetic field strengths are measured at the N_m sensors and the inverse problem is to obtain the set of N_r dipole strengths r_i with $N_r >> N_m$. The presence of noise raises the level of difficulty of the inverse problem.

3 Methods

Human Cortex and MEG Sensors. In typical quantitative brain studies, the approximately 10^{10} neurons in the cortex are simulated by about 2.4×10^5 current dipoles whose directions are set parallel to the normal of the cortex surface (MRIdatabase, 2006; C.A. Cocosco, 1997; R.K.-S. Kwan, 1999, 1996; D.L. Collins, 1998; P. Nunez, 1981). For this study we focus on the auditory cortex, the area marked by the 40 mm$\times$ 15 mm rectangular shown in Fig. 1L that contains 2188 current dipoles.

In typical MEG experiments the human head is surrounded by a hemispheric tiling of magnetic field sensors called superconducting quantum interference devices (SQUIDS). In the experiment we consider, there are 157 sensors, each composed of a pair of co-axial coils of 15.5 mm diameter 50 mm apart, with the lower coil about 50 mm above the scalp (H. Kado, 1999). The centers of the sensors are represented by the gray dots in Fig. 1R. Details of geometry and of the A-matrixes in Eq. (1) are given by H.-I. Bai (2003).

Artificial MEG data and Noise. We use artificial MEG data generated by the forward equation, Eq. (1), from sets of current dipoles (to be specified below) in a small area (black circles in Figs. 2R) within the auditory cortex (the gray "rectangular" region in Figs. 2L, enlargement shown in Figs. 2R).

A site-independent white Gaussian noise is linearly superimposed on the MEG data. The signal to noise ratio (SNR) is defined as

$$SNR = -10log_{10}||\mathbf{n}_{max}||^2/||\mathbf{m}_{max}||^2 \tag{3}$$

where $\mathbf{m}_{max}$ is the amplitude at the sensor receiving the strongest noiseless MEG signal, and $\mathbf{n}_{max}$ is amplitude of the strongest simulated noise. In this study we have $\mathbf{m}_{max}$=7.4 fT (fT= femto-Tesla) and $\mathbf{n}_{max}$=0.05$\mathbf{m}_{max}$=0.37 fT, so that SNR=26 on each individual run. The artificial MEG data is generated by giving a current of 10 nA (nano-ampere) to each of the sources in a source set (see below), running the forward equation with noise 10 times and taking the averaged strengths at the sensors. The averaging has the effect of reducing the effective $\mathbf{n}_{max}$ by a factor of $\sqrt{10}$, and yielding an enhanced effective signal to noise ratio of SNR'=36.

Using a threshold of T_S=14$\mathbf{n}_{max}$=5.1 fT we select a subset M_S of 7 "strong

signal" sensors. This implies a minimum value of SNR'=32.9 (with averaging) on each sensor in the set. Given the (assumed) normal distribution of noise intensity, this selection implies that at 99.99% confidence level the signals considered are not noise. Similarly we use a threshold of T_C=6$\mathbf{n}_{max}$=2.2 fT to select a subset M_C of 31 "clear signal" sensors, with minimum values of SNR'=25.6 on each sensor in the set. In actual computations below, we reduce the sensor space to one that include only those in the set M_C. In practice, the reduction replaces $\mathbf{m}$, $\mathbf{A}$, and $\mathbf{n}$ by $\mathbf{m}'$, $\mathbf{A}'$, and $\mathbf{n}'$, respectively.

Receiver Operating Characteristics Analysis. We evaluate the goodness of our results using receiver operating characteristics (ROC) analysis (T. Fawcett, 2004), in which the result is presented in the form of a plot of the true positive rate (S_n, sensitivity) versus the false positive rate (or 1-S_p, where S_p is the specificity). Let R be the total solution space of current sources, T the true solutions, or actual active sources, and P the positives, or the predicted active sources. Then F=R-T is the false solutions, TP=$(P \bigcap T)$ the true positives, FP=$(P \bigcap T)$-T the false positives, and FN=R-$(P \bigcap T)$ the false negatives. By definition S_n=TP/T=$(P \bigcap T)/T$ and 1-S_p=FP/F=$((P \bigcup T)-T)/(R-T)$. Intuitively, a good solution is one such that maximizes S_n while minimizing 1-S_p. In a null theory, the positives will fall randomly into R, hence TP/T=1-S_p, or S_n=1-S_p. Therefore, the merit of model producing a piece of data, (S_n, 1-S_p), showing on an ROC plot is measured by the difference between S_n and 1-S_p. In general, when a model is used to generate a curve in an ROC plot, the "area under the curve" (AUC), or the area between the curve and the S_n=1-S_p line, is a measure of merit of the model (E. DeLong, 1988).

Clustering and Sorting. Although implicit in the MRI head model introduced in Fig. 1 is a dramatic reduction of the number of neurons and the complexity of the cortex, the remaining number of effective neurons is still far greater than the number of detectors. We use a clustering algorithm (L. Kaufman, 1990) to further decrease the number of effective sources, in which the sources are partitioned according to spatial proximity and similarity in orientation into the set of N_C clusters $C = \{C_u | u = 1, 2, \cdots, N_C\}$, as follows. We require sources within a cluster to lie with a spatial radius of 5 mm and define ($N_u = \sum_{i \in C_u}$ is the number of current sources in cluster C_u)

$$\mathcal{A}'_u = \sum_{i \in C_u} \mathbf{A}'_i \tag{4}$$

as the "strength" of the $\mathbf{A}$-matrices in cluster C_u,

$$d_u = \sum_{i \in C_u} |N_u \mathbf{A}'_i - \mathcal{A}'_u| / N_u \tag{5}$$

as the "radius" – in the space of sensors – of cluster C_u, and

$$D_u = \sum_{C_v \in C} |\mathcal{A}'_v - \mathcal{A}'_u|/(N_C - 1) \tag{6}$$

as the average inter-cluster distance between C_u and all the other clusters. The clustering, including N_C, is determined by requiring that

$$d_u/D_u < \gamma_C, \quad \forall\ u = 1, 2, \cdots, N_C, \tag{7}$$

where γ_C is a parameter that controls the average cluster size; a smaller value of γ_C implies smaller and more numerous clusters. In the limit $\gamma_C \to 0$ every cluster will consist of a single source and $N_C \to N_r$, or 2188 in the present case. A clustering obtained with $\gamma_C = 1/7$ was used in this work. It partitions the 2188 sources into $N_C = 250$ clusters, whose size distribution is shown in Fig. 3.

The clustering results in the replacement of original source distribution by a coarse-grained distribution of virtual source-clusters whose N_r $\mathbf{A}'_i$-matrices are given by N_C $\mathcal{A}'_u$'s. The clustering reduces Eq. (2) to

$$\mathbf{m}' = \hat{\mathcal{A}}' \cdot \hat{r}_C + \mathbf{n}' \tag{8}$$

which has the same form as Eq. (2) except that here the hatted vectors have only N_C components and each of N_C components in $\hat{r}_C$ denotes the strength of the current dipole representing a cluster.

It is convenient to sort the cluster set C according to the field strength of the clusters. Since the field strength depends on the where it is measured, the sorted order will be sensor-dependent. We denote the sorted set for sensor α by $C^{\{\alpha\}}$. Thus we have:

$$C^{\{\alpha\}} = \{C_{u_\alpha} | u_\alpha = 1, 2, \cdots, N_C\}, \quad |\mathcal{A}'_{u_\alpha}| \geq |\mathcal{A}'_{v_\alpha}| \text{ if } u_\alpha < v_\alpha, \quad \forall \alpha \in M_C. \tag{9}$$

Forward Filtering. A key in improving the quality of the solution of an inverse problem is to reduce the number of false positives. In the MEG experiments under consideration, the plane of the sensors are generally parallel to the enveloping surface of the cerebral cortex. Such sensors are meant to detect signals emitted from current sources in sulci on the cortex, and are not sensitive to signals from sources in gyri. In practice, in our test cases T will be composed of sulcus sources. Therefore, if we simply remove those clusters having the weakest strengths, we will reduce FP at a higher rate than TP.

Given a positive fractional number $\xi < 1$, we use it to set an integer number $N_\xi < N_C$, and use N_ξ to define the truncated sets

$$R_\xi^{\{\alpha\}} = \{C_{u_\alpha} | u_\alpha = 1, 2, \cdots, N_\xi\}, \quad \forall\ \alpha \in M_S. \tag{10}$$

The integer N_ξ is determined by regression by demanding the union set

$$R_\xi = \bigcup_{\alpha \in M_S} R_\xi^{\{\alpha\}} = \xi R \tag{11}$$

to be a fraction ξ of R. We call this forward filtering process of reducing the pool of possible positives from R to R_ξ the "mostly sulcus model" (MSM). About 25% of current sources in R lie in gyri. Therefore, if we set ξ=0.75, very few potentially true sources will be left out. It turns out that in the region where 1-S_p is only slightly less than unity, setting P to R_ξ can offer the best result.

Backward Filtering. Another way of reducing the pool of positives is to limit them to those clusters, with unit current strength, whose $|\mathcal{A}'|$ value is greater than a threshold value A_0 at all the sensors in M_S. This yields, for each α, a reduced set with $N_\alpha < N_C$ clusters:

$$R_>^{\{\alpha\}} = \{C_{u_\alpha} | u_\alpha = 1, 2, \cdots, N_\alpha\}_\alpha, \quad \forall\ \alpha \in M_S. \tag{12}$$

Now we let R_{SHM}, the pool of positives for the "simple head model" (SHM), be the intersection of the reduced sets,

$$R_{SHM} = \bigcap_{\alpha \in M} R_>^{\{\alpha\}}. \tag{13}$$

R_{SHM} represents a coarse-grained, simplified cortex tailored to the MEG data at hand: every source-cluster in R_{SHM} has a relatively high probability of contributing significantly to all the sensors in M_S. A hypothetical case in which M_S is composed of the two sensors, α_1 and α_2, is depicted in Fig. 4.

For this paper A_0 is set to be 4.5 fT. Then the N_α's have values 113, 134, 102, 126, 103, 119, and 64, respectively, for the 7 sensors in M_S, and the simple head model set R_{SHM} contains 240 source currents, or about 11% of the total number of current dipoles in the auditory region.

The Maximum Entropy Method. The maximum entropy (ME) method is a method for deriving the "best" solution in ill-posed problems (E. L. C. Amblard, 2004; C. J. S. Clarke, 1989; L. K. Jones, 1989, 1990; I. Csiszar, 1991; F. N. Alavi, 1993; D. Khosla, 1997; G. Le Besnerais, 1999; R. He, 2000; H. Gzyl, 2002). Generally, the equation that admits multiple solutions is treated as a constraint and, given a *prior* probability distribution of solutions, the method finds a *posterior* probability distribution by maximizing the relative entropy of the probability distributions. When applied to the MEG case, Eq. (8) (or Eq. (2) without clustering) is used to constrain the posterior propbability distribution for a $\hat{r}$ that is the "best" ME solution, given $\mathbf{m}$ (measured) and $\mathbf{n}$ (presumed or otherwise obtained). The procedure is tedious but standard and

an implementation was reported by H.-I. Pai (2005). Here we only describe how the *prior* probability distribution of solutions are determined in this work.

We test several procedures ranked by their levels of complexity: (1) Simple ME (ME). All the 2188 individual dipoles are included in the *prior* P. Here as in all other cases, in ME iteration involving sensors, only those in the "clear signal set" M_C are included. (2) ME with clustering but not filtering (C-ME). Cluster are treated as units of sources and all clusters are included in the *prior* P. (3) ME with filtering but not clustering (F-ME). Individual dipoles are treated as units of sources but only those in R_ξ or R_{SHM} (whichever is the smaller set) are included in the *prior* P. (4) ME with clustering and filtering (C-F-ME). Cluster are treated as units of sources but only those in R_ξ or R_{SHM} are included in the *prior* P. Fig. 5 is the flowchart for computation for the above procedures. In each case the set of positives, P, hence S_p, is varied during the implementation of ME by a threshold on the strength of source dipoles for acceptance into the set. The exception is when the "F" procedure is taken. In this case, when $0<S_p\leq 0.25$, the set of positives, P, is directly set equal to R_ξ as described in Eq. (11), without going through the ME procedure. When $S_p>0.25$ the procedure is switched to SHM followed by ME. In addition, we compare ME and MNLS (J.-Z. Wang, 1993). When MNLS is involved the flowchart is the same as given in Fig. 5, with ME replaced by MNLS.

4 Results and discussions

We report three preliminary tests exploring the properties of the procedures.

Test 1. The true set T is one cluster containing 13 sources covering an area of 4 mm by 1 mm (Fig. 6L). The computed results, shown as an ROC plot, are given in Fig. 6R.

Not shown are results 1-S_p>0.20, when all four procedures give S_n=1. It is seen that C-F-ME gives the best result: S_n=1 for the entire range of values for 1-S_p. C-ME and F-ME are excellent when 1-$S_p\geq 0.025$, but completely fail to identify the true sources when 1-$S_p<0.025$. In comparison, ME performs not as well as C-ME and F-ME when 1-$S_p\geq 0.025$ but is better when smaller values of 1-S_p. Note that the R set contains 2188 sources and the T set 13 sources. So even at 1-S_p=0.025 there are still 54 false positive (FP) sources.

Test 2. The true set T is composed of two clusters containing 12 and 7 sources, respectively, covering an area 5.5 mm by 1 mm (Fig. 7L). The ROC plot (Fig. 7R) shows better results are obtained when F is involved (F-ME and C-F-ME). When clustering is involved (C-ME and C-F-ME) changes in S_n are discrete. This is because P may take only take four values, 1 when P

contains both true clusters, 0.63 or 0.37 when it contains of the two, and 0 when it contains none.

Recall that clustering simplifies the organization of the sources (R) but does not reduce the *prior* positives (P). F reduces the *prior* P but does not simplify R. C-F does both. That procedures with F are clearly better than those without highlights the paramount importance of a better *prior* in all but trivial situations when ME is employed.

Test 3. The true set T is composed of 42 active sources covering an area approximately 3 mm by 2 mm (Fig. 8L).

They are distributed in eight clusters (C_u, u=1 to 8) containing 18, 31, 15, 12, 9, 23, 18, and 17 sources, respectively, for a total of 143 sources. The intersection of these clusters with T are 4, 11, 4, 3, 4, 4, 6, and 6 sources, respectively. If clustering is applied, the minimum value for 1-S_p is 0.047 when S_n=1. The ROC shown in Fig. 8R shows that for this relatively complicated case simultaneous high S_n and S_p is difficult to achieve; we obtain S_n<0.6 for 1-S_p<0.20 in all procedures. Here again, procedures with F, which generate better *priors*, yield more accurate positives than those without. It is worth pointing out the accuracies of the positives given by F-ME and C-F-ME are very similar in almost the entire range of 1-S_p, but F-ME is decisively better than C-F-ME when 1-S_p is less than 0.01. The last effect brings out an inherent weakness of clustering. If the active sources (the T set) are spread out in more than one cluster and and if the union of the clusters is greater than the T set, then a solution with a non-null P ($S_n \neq 0$) and null FP (1-S_p=0) is not possible. In contrast, such an outcome is at least possible without clustering. As it happened, for the case at hand, at 1-$S_p \approx 0$, the P set for C-F-ME is cluster C_4 with 12 sources, of which 3 belongs to T, yielding S_n=3/42=0.071 and 1-S_p=9/2146=0.0042.

Fig. 9 shows the 21 strongest sources (the "strong set"), not necessarily all in the positive set, given by the four procedures when 1-S_p<0.02 (the sets do not change much in this range of 1-S_p). In the case of C-F-ME (panel (a)), the strong set comes from clusters C_3 and C_4, both of which lie in the vicinity of T. When 1-S_p is lowered (by raising the P-acceptance threshold) beyond a certain point C_3 is eliminated, leaving only C_4 in P in the situation discussed above. In the case of F-ME (panel (b)), some individual sources in the strong set are not in the vicinity of T. However, these are eliminated when 1-S_p is lowered, and the remaining individual sources happen to have a FP that is smaller than that in C-F-ME.

ME versus MNLS. We compare the effectiveness of the MNLS procedure against ME using the true set of test 3. Fig. 10L shows the ROC plots for C-F-ME (same plot as in Fig. 8), C-F-MNLS and MNLS. We observe that MNLS

is worse than ME (Fig. 8), C-F-MNLS is better than MNLS, and C-F-ME is better than C-F-MNLS. This shows, at least for the case tested, clustering is also beneficial to MNLS and, other things being equal, ME is more effective than MNLS. In the last instance the FP's of MNLS cover a large area of the auditory cortex.

The strong sets in the three procedures are given in Fig. 10R. If we take these dipoles as the positive set then we obtain the results given in Table 1.

In summary, we have employed several simple cases to illustrate that even when we may use fMRI data to tell us where the general region of the source currents is, the nature of the inverse problem is such that the challenge to precisely pinpoint sources with a resolution of a few mm is still great. We showed that clustering tends to limit the area covered by false positives, and filtering is effective for generating better *priors* for ME. When we use ME implemented by these procedures, we can achieve partial success in pinpointing sources concentrated in an area the size of a few mm across, while reducing the spatial distribution and number false positives to a very low level. We note that the area of the active sources is miniscule compared to the auditory cortex, which is itself yet much smaller than cortical surface covered by the sensors receiving clear or even strong signals. Thus, at least in the case of one compact set of active sources, our method can achieve a spatial resolution comparable to that of fMRI (C.T.W. Moonen, 1999). On the other hand, there is still much room for improvement. For example, if the method is to be equally successful in the case of two sets of active sources separated by, say, a few cm, then the procedure of backward filtering may need to further refined.

5 Acknowledgements

This work is supported in part by grant nos. 95-2911-I-008-004 and 96-2112-M-008-025, National Science Council (ROC), and the Cathy General Hospital-National Central University Grant 96-CGH-NCU-A1.

References

C.T.W. Moonen and P.A. Bandettini, *Functional MR angiography using inflow and phase-contrast MR acquisition techniques*. Berlin: Springer, 1999.

J.-Z. Wang, "Minimum-norm least-squares estimation: Magnetic source images for a spherical model head". *IEEE Trans on Biomed. Eng.*, **40**, 387–396, 1993.

E. L. C. Amblard and J. M. Lina, "Biomagnetic source detection by maximum entropy and graphical models." *IEEE Trans. Biomed. Eng.*, **51**, 427–442, 2004.

MRIdatabase: The simulated brain database website, http://www.bic.mni.mcgill.ca/brainweb/

C.A. Cocosco, V. Kollokian, R.K.-S. Kwan, A.C. Evans : "BrainWeb: Online Interface to a 3D MRI Simulated Brain Database" NeuroImage, vol.5, no.4, part 2/4, S425, 1997 – Proceedings of 3-rd International Conference on Functional Mapping of the Human Brain, Copenhagen, May 1997.

R.K.-S. Kwan, A.C. Evans, G.B. Pike : "MRI simulation-based evaluation of image-processing and classification methods" IEEE Transactions on Medical Imaging. 18(11):1085-97, Nov 1999.

R.K.-S. Kwan, A.C. Evans, G.B. Pike : "An Extensible MRI Simulator for Post-Processing Evaluation" Visualization in Biomedical Computing (VBC'96). Lecture Notes in Computer Science, vol. 1131. Springer-Verlag, 1996. 135-140.

D.L. Collins, A.P. Zijdenbos, V. Kollokian, J.G. Sled, N.J. Kabani, C.J. Holmes, A.C. Evans : "Design and Construction of a Realistic Digital Brain Phantom" IEEE Transactions on Medical Imaging, vol.17, No.3, p.463–468, June 1998.

P. Nunez, *Electric Fields of the Brain*. New York: Oxford University, 1981.

H. Kado, et al., "Magnetoencephalogram systems developed at KIT." *IEEE Transactions on Applied Superconductivity*, **9**, 4057–62, 1999.

H.-I. Bai, "Detection of Biomagnetic Source by the the Method of Maximum Entropy", master thesis, Department of Physics, National Central University. 2003 (unpublished). The coordinates of the sensors are given at the website http://pooh.phy.ncu.edu.tw/MEG/Index.htm.

L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*. New York: John Wiley and Sons, 1990.

T. Fawcett (2004), "ROC Graphs: Notes and Practical Considerations for Researchers", Technical report, Palo Alto, USA: HP Laboratories.

E. DeLong, D. DeLong and D. Clarke-Pearson, "Comparing the areas under two or more correlated receiver operating characteristic (ROC) curves: a nonparametric approach", *Biometrics*, **44**, 837–845, 1988.

C. J. S. Clarke and B. S. Janday, *Inverse Problems*, **5**, 483–500, 1989; C. J. S. Clarke, *ibid.* 999–1012.

L. K. Jones and V. Trutzer, "Computationally feasible high-resolution minimum-distance procedures which extend the maximum-entropy

method," *Inv. Prob.*, **5**, 749-766, 1989.

L. K. Jones and C. L. Byrne, "General entropy criteria for inverse problems, with applications to data compression, pattern classification, and cluster analysis," *IEEE Trans. Inform. Theory*, **36**, 23-30, 1990.

I. Csiszar, "Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems," *The Annals of Statistics*, **19**, 2032-2066, 1991.

F. N. Alavi, J. G. Taylor and A. A. Ioannides, "Estimates of current density distributions: I. Applying the principle of cross-entropy minimization to electrographic recordings," *Inv. Prob.*, **9**, 623–639, 1993.

D. Khosla and M. Singh, "A maximum entropy method for MEG source imaging," *IEEE Trans. Nuclear Sci.*, **44**, 1368–1374, 1997.

G. Le Besnerais, J.-F. Bercher and G. Demoment, "A new look at entropy for solving linear inverse problems," *IEEE Tran. Inf. Theory*, **45**, 1565–1578, 1999.

R. He, L. Rao, S. Liu, W. Yan, P. A. Narayana and H. Brauer, "The method of maximum mutual information for biomedical electromagnetic inverse problems," *IEEE Trans. Magnetics*, **36**, 1741–1744, 2000.

H. Gzyl, "Maximum entropy in the mean: A useful tool for constrained linear problems," in *Bayesian Inference and Maximum Entropy in Science and Engineering*, C. Williams, Ed. AIP Conf. Proc. 659, Am. Inst. Phys., New York, 2002, pp.361-385.

H.-I. Pai, C.-Y. Tseng, and H. C. Lee, "Identifying bio-magnetic sources by maximum entropy approach", in Bayesian Inference and Maximum entropy methods in Science and Engineering. Ed. K. Knuth, A. E. Abbas, R. D. Moris, and J. P. Castle, *AIP Conf. Proc.* **803**, 527–534, 2005 (copy available at xxx.lanl.gov/abs/qbio. NC/050804).

6 Tables

			C(luster)	F	C-F
S_n	ME	0.000	0.048	0.119	0.167
	MNLS	0.000	0.000	0.048	0.000
$1 - S_p$	ME	0.0096	0.0087	0.0073	0.0064
	MNLS	0.0096	0.0096	0.0087	0.0096

Table 1

7 Figures

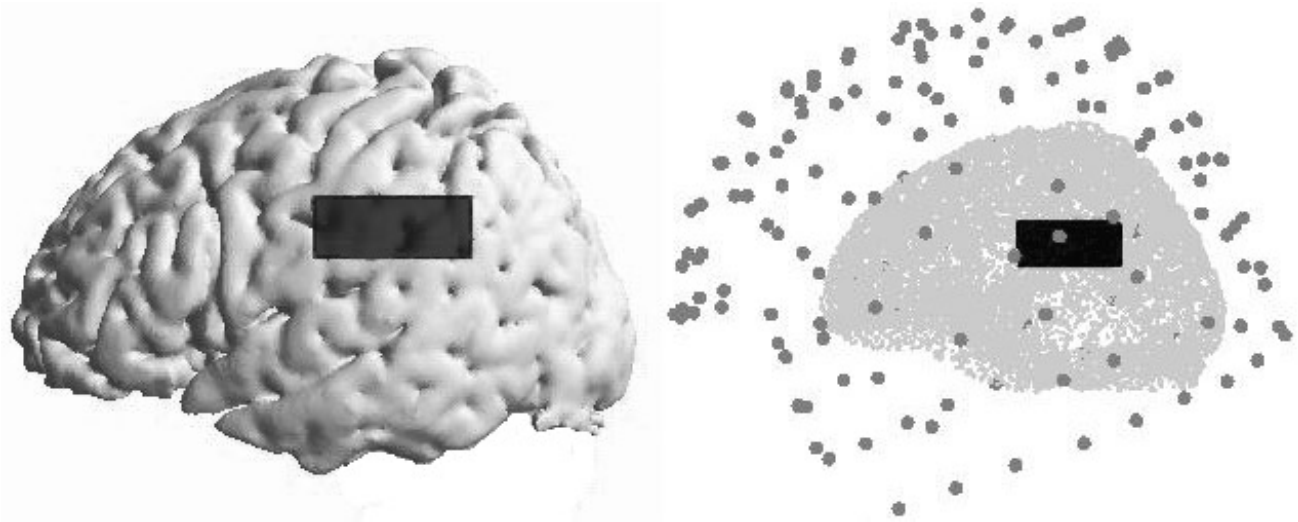
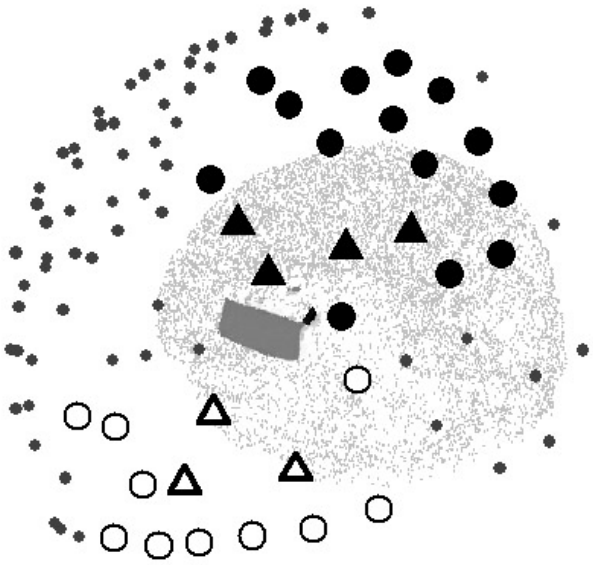
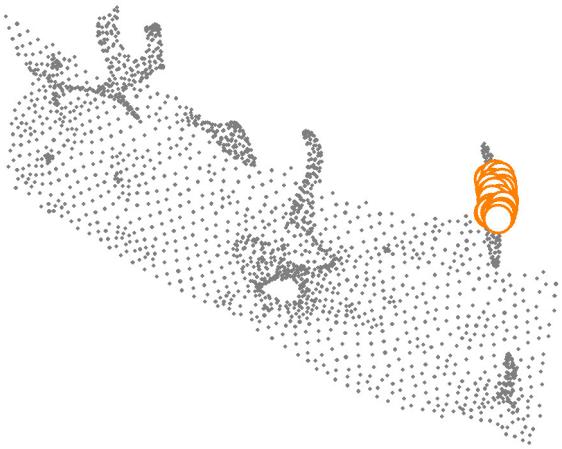


Fig. 1.





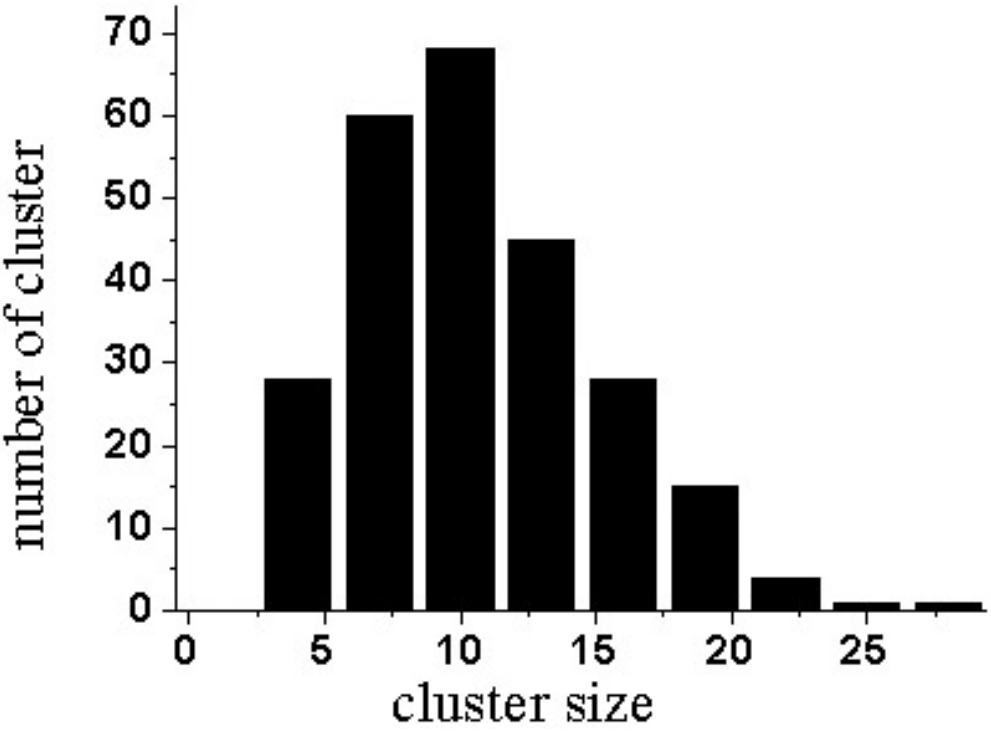

number of cluster
70
60
50
40
30
20
10
0
0
5
10
15
20
25
cluster size

Fig. 3.

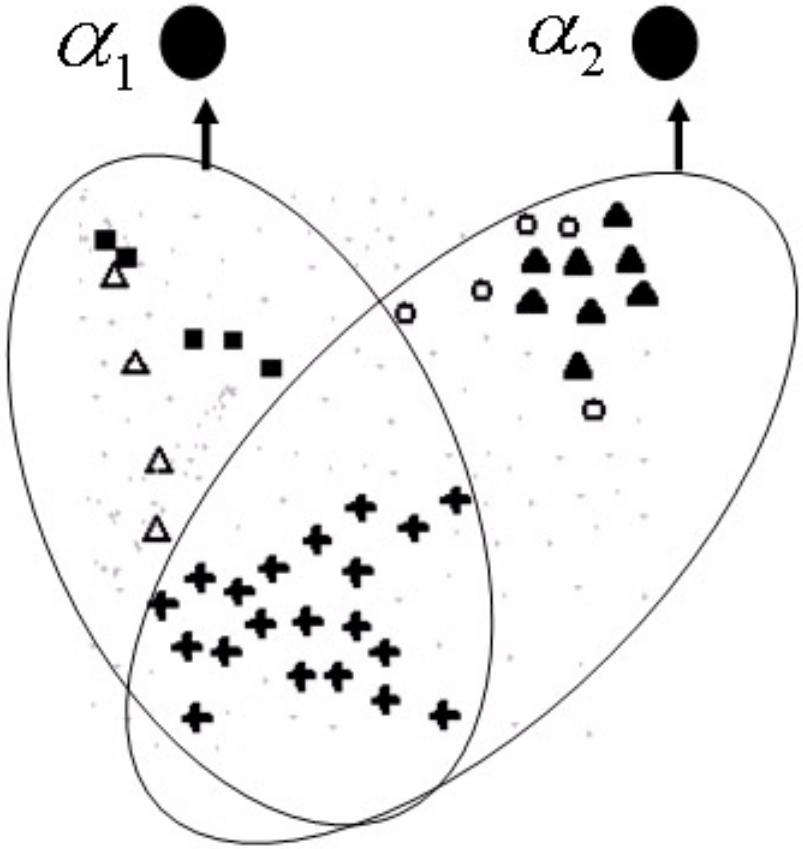


Fig. 4.

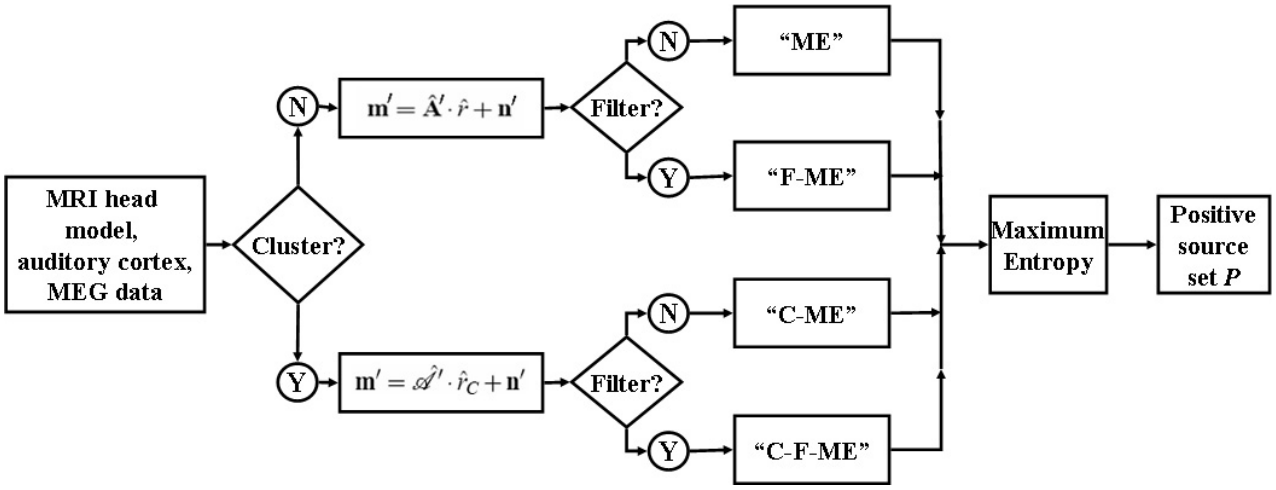

MRI head model, auditory cortex, MEG data
Cluster?
N
Y
$\mathbf{m}' = \hat{\mathbf{A}}' \cdot \hat{r} + \mathbf{n}'$
$\mathbf{m}' = \hat{\mathscr{A}}' \cdot \hat{r}_C + \mathbf{n}'$
Filter?
Filter?
N
Y
N
Y
"ME"
"F-ME"
"C-ME"
"C-F-ME"
Maximum Entropy
Positive source set *P*

Fig. 5.

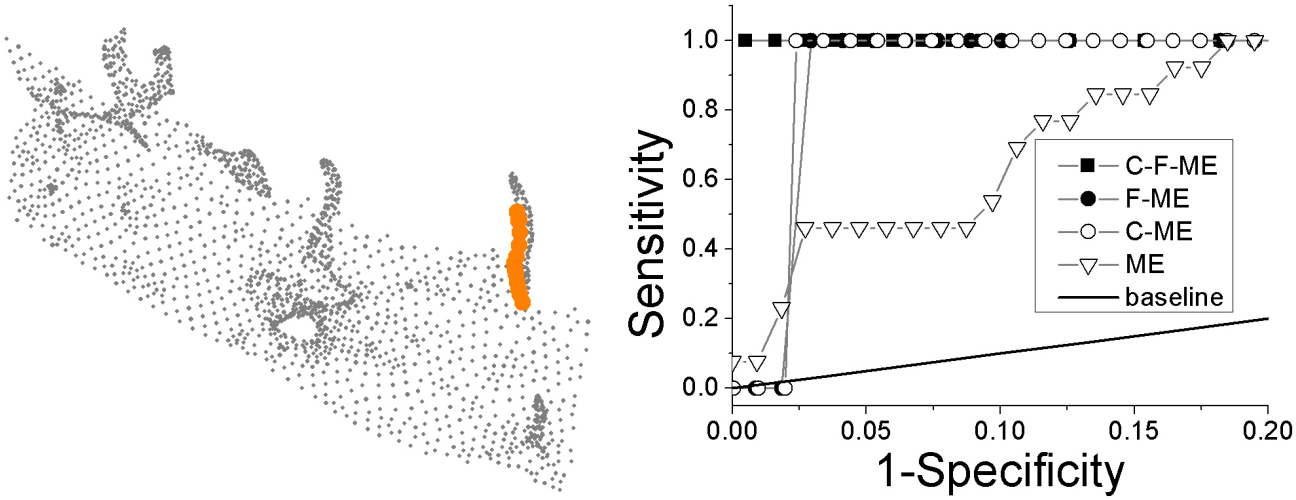

1.0
0.8
0.6
0.4
0.2
0.0
Sensitivity
0.00
0.05
0.10
0.15
0.20
1-Specificity
C-F-ME
F-ME
C-ME
ME
baseline

Fig. 6.

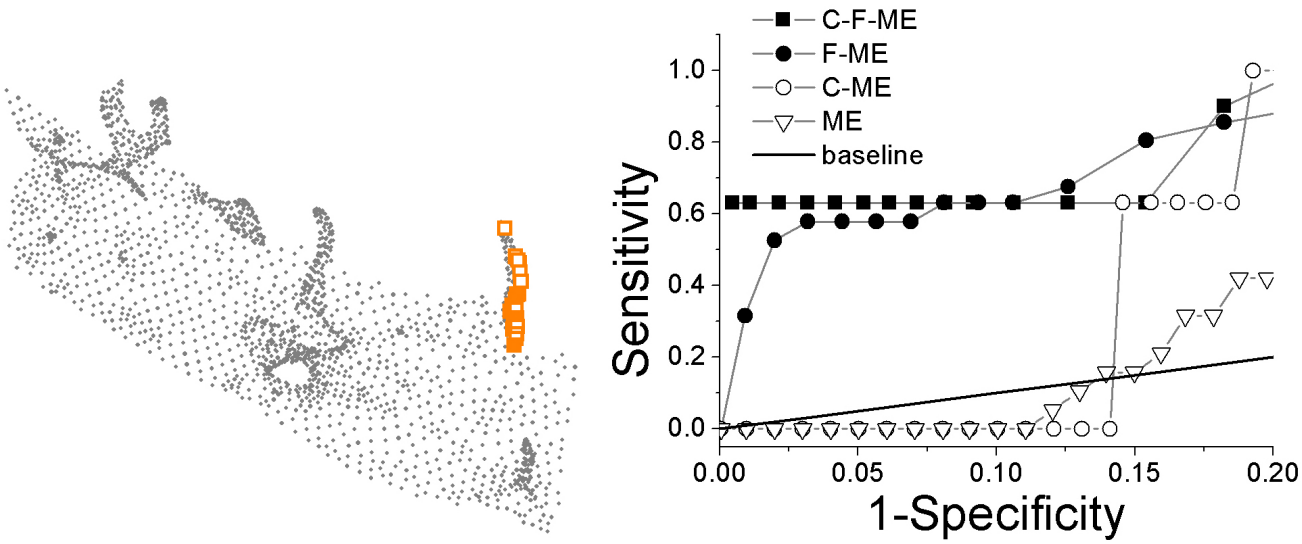

C-F-ME
F-ME
C-ME
ME
baseline
Sensitivity
1-Specificity
0.0
0.2
0.4
0.6
0.8
1.0
0.00
0.05
0.10
0.15
0.20

Fig. 7.

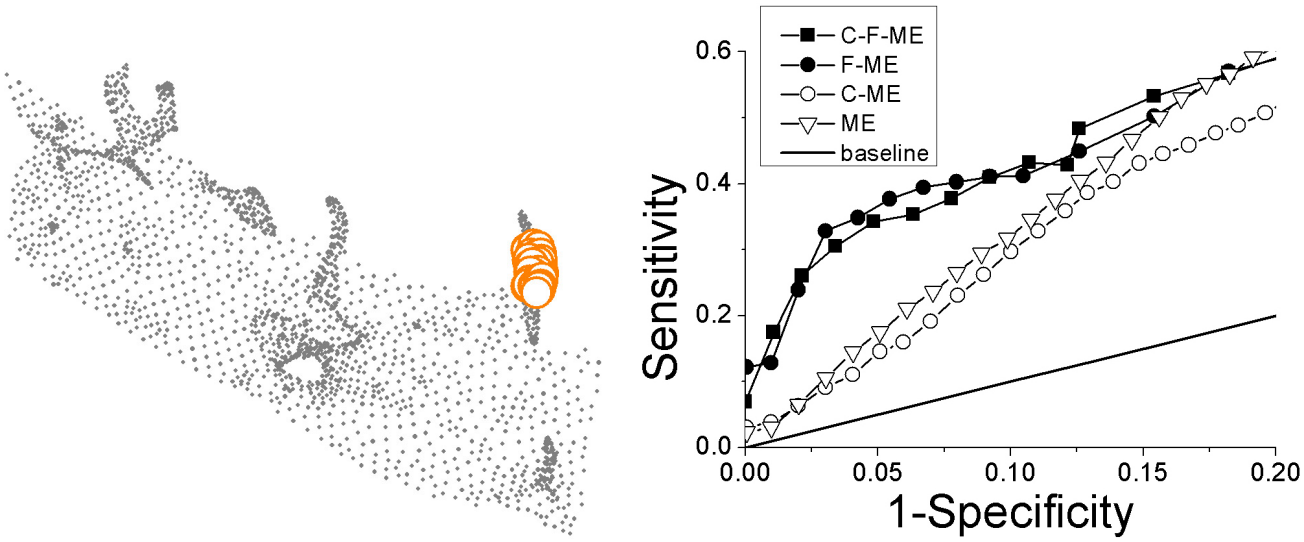

C-F-ME
F-ME
C-ME
ME
baseline
Sensitivity
1-Specificity
0.0
0.2
0.4
0.6
0.00
0.05
0.10
0.15
0.20

Fig. 8.

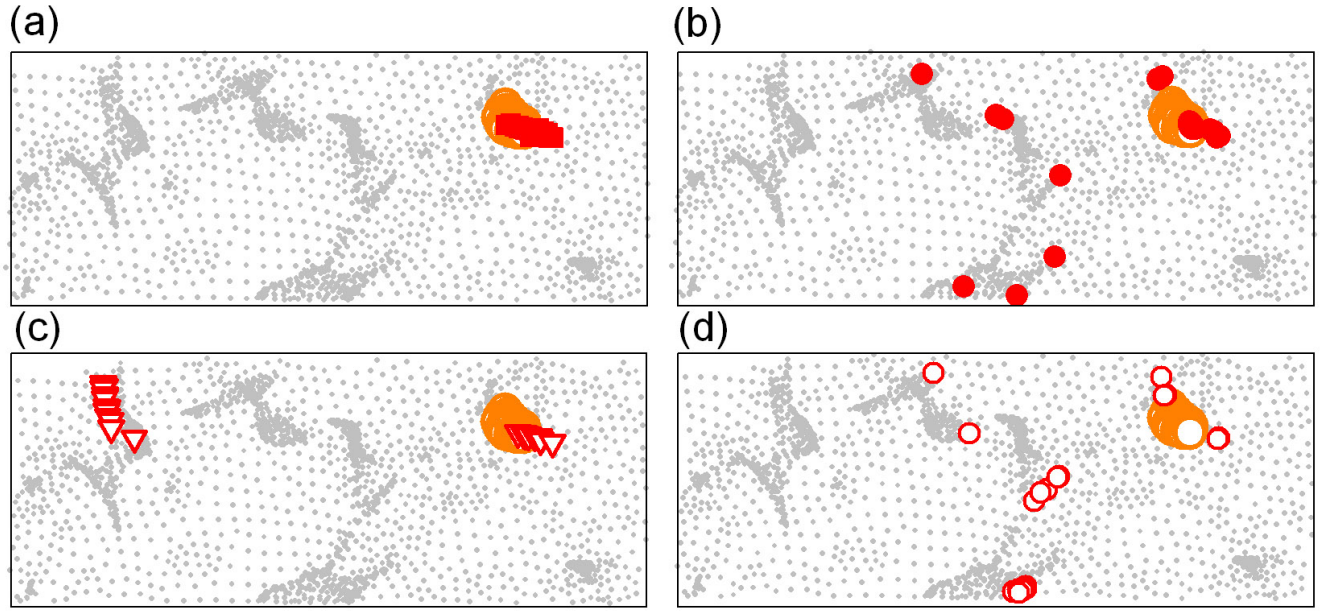


Fig. 9.

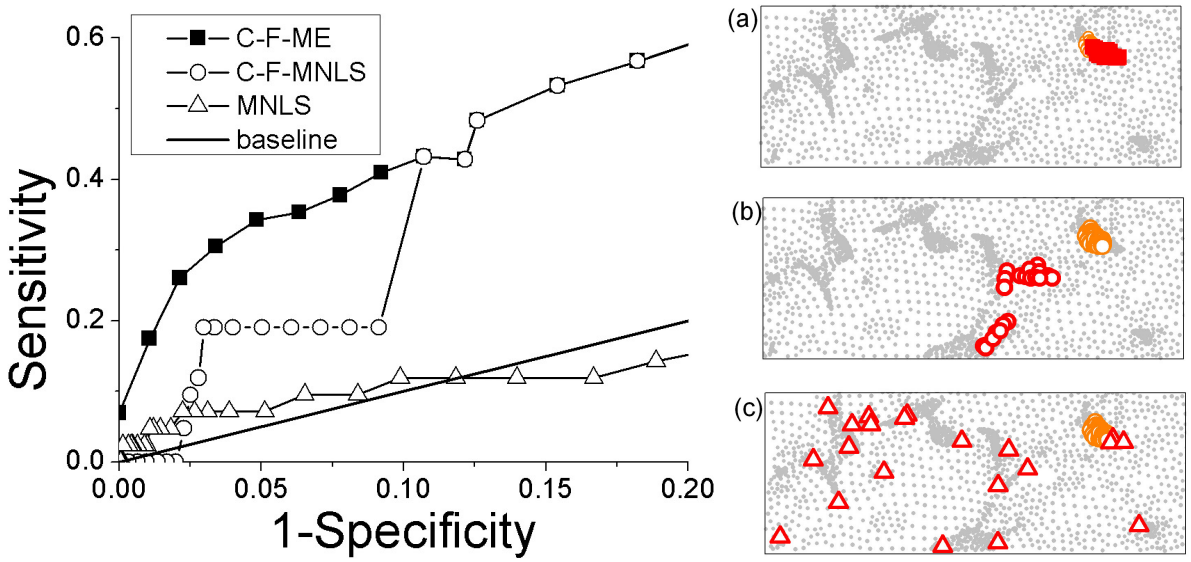

C-F-ME
C-F-MNLS
MNLS
baseline
Sensitivity
1-Specificity
0.0
0.2
0.4
0.6
0.00
0.05
0.10
0.15
0.20
(a)
(b)
(c)

Fig. 10.

8 Legends

Table.1:Comparisons of ME and MNLS using the set of 42 active sources of test 3. In each procedure the positive set is composed of the 21 current dipoles with the strongest currents. The third column gives results with neither clustering nor F.

Figure.1: Left, side view of the human cortex; the frontal lobe points to the left. The region of interest, the auditory cortex, is marked by the dark rectangular. Right, schematic setup of an MEG experiment. The occipital lobe (right side in graph) is closer to the sensors because the person being tested is lying face-up.

Figure.2: Left, the circular and triangular symbols are the positions of the 31 sensors with detectable signals, including the 7 (triangular) with signals above threshold. Sensors with magnetic flux going into (out) the page are solid (hollow). Right, detail of auditory cortex. The dark (orange in color) circles in the top-right corner indicate the general area of active sources used to generate artificial MEG data.

Figure.3: Distribution of cluster size in the clustering of 2188 sources into 250 clusters.

Figure.4: Illustration of backward filtering, when M_S is composed of the two sensors, α_1 and α_2. Both $R_{\{\alpha_1\}}$ and $R_{\{\alpha_2\}}$ contain 3 clusters while R_{SHM} contains only one cluster.

Figure.5: Data processing flowchart for producing a *prior* input for the maximum entropy method.

Figure.6: Left, true cluster set of sources (solid circles) in test 1. Right, ROC plot for test 1.

Figure.7: Left, the two true cluster sets of sources (solid circles and hollow squares) in test 2. Right, ROC plot for test 2.

Figure.8: Left, the 42 source dipoles in test 3. Right, ROC plot for test 3; only results with 1-S_p<0.2 are shown.

Figure.9: The strong sets (21 strongest sources) in test 3 at 1-S_p<0.02. In each case the gray (orange in color) blotch near the top-right corner is the T set of 42 active sources. (a) C-F-ME (solid squares). (b) F-ME (bullets). (c) C-ME (triangles). (d) ME (open circles).

Figure.10: Left, ROC plots for MNLS, C-F-MNLS, and C-F-ME. Set T is

the 42 active sources of test 3. Right, the strong set (dark symbols) at 1-S_p<0.02. (a) C-F-ME (solid square); (b) C-F-MNLS (open circle); (c) CMNLS (triangle). T is the set of grey (orange) circles near the top-right corner in panels (a-c).