Referee Consensus:
A Platform Technology for Nonlinear Optimization

Don Krieger
University of Pittsburgh
Presbyterian University Hospital
200 Lothrop Street; Suite B400
(412)648-9654
kriegerd@upmc.edu

Malcolm McNeil
University of Pittsburgh
Jinyin Zhang
Carnegie Mellon University
Walter Schneider
University of Pittsburgh

Xin Li
Carnegie Mellon University
David O. Okonkwo
University of Pittsburgh

ABSTRACT

Electrical current flow within populations of neurons is a fundamental constituent of brain function. The resulting fluctuating magnetic fields may be sampled noninvasively with an array of magnetic field detectors positioned outside the head. This is magnetoencephalography (MEG). Each source may be characterized by 5-6 parameters, the xyz location and the xyz direction. The magnetic field measurements are nonlinear in the location parameters; hence the source location is identifiable only via search of the brain volume. When there is one or a very few sources, this may be practical; solutions for the general problem have poor resolution and are readily defeated.

Referee consensus is a novel cost function which enables identification of a source at one location at a time regardless of the number and location of other sources. This "independence" enables solution of the general problem and insures suitability to grid computing. The computation scales linearly with the number of nonlinear parameters. Since the method is not readily disrupted by noise or the presence of multiple unknown source, it is applicable to single trial data.

MEG recordings were obtained from 26 volunteers while they performed a cognitive task The single trial recordings were processed on the Open Science Grid (≈300 CPU hours/sec of data) On average 500+ active sources were found throughout. Statistical analyses demonstrated 1-2 mm resolving power and high confidence findings ($p < 0.0001$) when testing for task specific information in the extracted virtual recordings.

Referee consensus is applicable to a variety of systems in addition to MEG, e.g. the connectivity problem, the blurred image, both passive and active SONAR, and seismic tomography. Applicability requires (1) that the measurements be linear in at least one of the source parameters and (2) that a sequence of measurements in time be obtained.

XSEDE '13, July 22 - 25 2013, San Diego, CA, USA

http://dx.doi.org/10.1145/2484762.2484789 .

Categories and Subject Descriptors

• Mathematics of computing~Solvers • Theory of computation~Massively parallel algorithms • Computing methodologies~Discrete space search • Applied computing~Physics

General Terms

Algorithms, Measurement, Experimentation.

Keywords

Nonlinear optimization, magnetoencephalography, MEG, open science grid, OSG, traumatic brain injury, concussion.

1. Introduction

Electrical current flow within populations of neurons is a fundamental constituent of brain function. The resulting fluctuating magnetic fields may be sampled noninvasively with an array of magnetic field detectors positioned outside the head. This is magnetoencephalography (MEG).

The signal at each MEG sensor is a weighted sum of the magnetic fields produced by sources within the brain, i.e. the relationship between the measurements and the amplitudes of the current sources is linear. But the number and locations of the sources are generally unknown and the relationship between the measurements and the source location coordinates is nonlinear [1]. These issues pose fundamental poorly solved problems with handling and interpreting MEG signals.

In restricted special cases the data are manipulated to reduce all significant contributors to the MEG to a single source. In these cases, the widely accepted Equivalent Current Dipole (ECD) localization is applicable. A single point source current dipole is assumed, requiring estimation of 5-6 parameters, 3 location coordinates and 2-3 current amplitude components. The parameter estimation is typically accomplished using an iterative gradient search method for all 5-6 parameters. The accuracy of this method is vulnerable both to extraneous sources and to instrument noise. This forces the investigator to average the data synchronized to an event and thereby reduce the quantity of extractable information by a factor of 100 or more.

For the general multiple source problem, many investigators use methods with which thousands of point sources are estimated in a

single operation, e.g. MNE [2], LORETA [3,4], VESTAL [5,6]. This approach provides a solution to the localization problem by including source locations with sufficient density to insure that no source is more than a few mm from one of them. But these methods produce estimates of thousands of parameters from hundreds of data points yielding estimates that are under fit. Because of this both localization accuracy and the ability to resolve sources which are near each other is poor. Furthermore these methods are readily defeated by noise in the data, again forcing up front averaging with consequent loss of information.

Referee consensus is used in a manner that is comparable to single source localization. But it is insensitive to the presence of either extraneous sources or noise. And it is dimensionally well-posed. These properties enable (1) application to single trial data and (2) resolution of sources within a few mm of each other.

A formal treatment of the simplest form of the method is presented followed by results of its application to the MEG problem. A generalized version of the method as well as directions for future work are discussed.

2. Methods

2.1 Referee Consensus Optimization

The objective of referee consensus optimization is to reliably determine the presence of sources whose actions are giving rise to measurements. It does so by providing a powerful statistical metric for deciding if a source is or is not present at a specific location. It enables accurate estimation of one or a few nonlinear variables at a time, regardless of the number of other variables whose values are contributing to the measurements.

Each source is characterized by 1 or more parameters[1]. The measurements must be linearly related to at least one of these. "Best" values for the "location," the nonlinear parameters[2], are identified using a "trial and error" procedure, a search. At each step of the search the values of the nonlinear parameters are fixed, enabling the linear formulation detailed below.

A referee consensus system may be formulated as a set of linear regression equations: $\vec{B}^{\mathbf{T}} = \mathbf{H}\vec{D}^{\mathbf{T}} + \vec{E}^{\mathbf{T}}$. The elements of $\vec{B}$ are the measurements (known). The elements of $\vec{D}$ are the linear source variables (to be estimated via ordinary least squares regression). The elements of $\vec{E}$ are the errors in the estimates. Each element of $\mathbf{H}$, the h_{jk}, are nonlinear functions of the parameters that characterize the j^{th} sensor and the k^{th} source. The equations are solvable only when these nonlinear parameters are fixed so that the h_{jk} are numbers. For the referee consensus method, the existence and the values of the parameters of one source at a time are determined. All other sources included in the system estimate are either at fixed dummy "referee" locations or at known source locations which are included to reduce the errors.

The formulation of a referee consensus system must include at least 2 sources: (1) the target (k^{th}) source, the one for which the nonlinear parameters are to be found by search and (2) at least one referee location. Each referee must be "correlated" to the target location, i.e. the dot products of the columns of $\mathbf{H}$ corresponding to the referees with the columns for the target location must be non-zero.

The ordinary least squares solution for $\vec{B}^{\mathbf{T}} = \mathbf{H}\vec{D}^{\mathbf{T}} + \vec{E}^{\mathbf{T}}$ is $\vec{D}^{\mathbf{T}} = \left(\left(\mathbf{H}^{\mathbf{T}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathbf{T}}\right)\vec{B}^{\mathbf{T}}$, i.e. the k^{th} element of $\vec{D}$ is a weighted sum of the measurements, $\vec{B}^{\mathbf{T}}$, where the weights are the entries in the k^{th} row of $\left(\mathbf{H}^{\mathbf{T}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathbf{T}}$. Each of these rows is then a linear filter. Effectiveness of the referee consensus method relies critically on the fact that the k^{th} filter has zero gain for contributions to the measurements from all of the entries in $\vec{D}$ except d_k. A proof is provided in Appendix I. This existence of zeroes enables (1) high fidelity estimation of the referee template, the time course of the value of d_k, $d_k[t], t = 1,\ldots,80$, and (2) calculation of the referee consensus cost function:

(1) The system is solved which includes sources at the target location, **A**, and a referee, **R**. The filter[3] for **R**, is applied to a sequence of measurements, 80 points in the work reported here. There is no contribution to this sequence from a source at **A** since the filter for **R** has zero gain for a source there. Hence we refer to this sequence as **R!A**, "R not A."

(2) The system is solved which includes the same referee but a different target location 1 mm from **A**, **A'**. Again the filter for **R** is applied to the sequence of measurements but this time there is a contribution from **A** since it was not included in the model. This sequence is **R!A'**.

(3) **R!A'-R!A** is an estimate of the time course of a source at **A,** $d_k[t], t = 1,\ldots,80$. If there is in fact a source at **A**, it should contribute to **R!A'** but not to **R!A**. This is tested as follows: If ((**R!A'-R!A**)•**R!A'**)^2 > ((**R!A'-R!A**)•**R!A**)^2, then there is a source present at **A** from the viewpoint of this referee. In that case the "referee consensus" is incremented by 1. Note that $\mathbf{H}$ differs with the inclusion of **A** vs **A'** only in the 2 columns for the target location. Since locations **A** and **A'** are close to each other, these columns differ only slightly. In the work here, 90 referee locations with 2 columns each were included in $\mathbf{H}$. Hence $\mathbf{H}$ and the resultant filters, the rows of $\left(\mathbf{H}^{\mathbf{T}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathbf{T}}$, were only slightly

[1] For the magnetic field due to a current dipole in a uniformly conducting sphere [1] each source is characterized by 5 parameters. 3 of these are nonlinear, the xyz coordinates of the location of the source. The 2 linear parameters are the 2 components of the amplitude of the current in the tangent plane. This physical model precludes detection of the radial component of an intracranial current source.

[2] One or more linear parameter may be identified by search along with the nonlinear ones. This could be required if there are too few sensors to enable inclusion of the all the linear parameters as linear source variables in the regression formulation.

[3] For the MEG problem, it is natural to handle these variables 2 at a time. For simplicity but without loss of generality we describe formation of the referee template for a single variable at a time.

perturbed with the major difference in gains being in the neighborhood of **A**. This isolates the estimation of the referee template from the influence of other sources since the gains elsewhere are nearly equal for **R!A** and **R!A'**. Hence the difference operation attenuates the contribution of other sources as well as the contributions of instrument noise.

The process detailed in (1)-(3) was repeated in this work twice (2 components) for each of 90 referees. All 180 referee templates were combined using eigenvector analysis to generate a high fidelity estimate. The referee consensus is simply a count. The probability of getting a particular count is interpretable as one would interpret the flip of a fair coin. The expected value for K flips is K/2, e.g. the chance of getting 114 or more heads (1's) out of 180 is less than 1:100. This is the threshold for acceptance of the existence of a source at a particular location. The threshold had to be exceeded for 6 different **A'** locations chosen 1 mm from **A** along the x, y, and z axes in both direction. Hence the probability of accepting a source by chance was reduced to $1{:}10^{12}$. In addition to its use as a measure of confidence, the referee consensus may be used as a cost function to guide a gradient search as was done here.

2.2 Experimental Methodology

Under University of Pittsburgh IRB approval (PRO09040294), 26 participants included 16 with history of concussion were enrolled in this study. Written informed consent was obtained after which all volunteers sat for MEG recordings while performing the task. Either MRI (23) or CT (3) was used for anatomic localization.

MEG recordings were acquired in the UPMC Brain Mapping Center with a 306-channel sensor array (Neuromag VectorView, Elekta Inc., Stockholm, Sweden) in a magnetically shielded room (Imedco, Hagendorf, Switzerland). Data sampling rate was 1000 Hz with front end high and low pass filter settings: 0.1-330 Hz. Line noise was removed from the raw MEG at 60, …, 300 Hz [7].

With continuous MEG recording, each volunteer performed a visual choice task controlled by EPrime 2.0 (Psychology Software Tools; Pittsburgh, PA). Each trial consisted of one of 8 sentences followed by 3 consecutive test figures to which a rapid response was given. Each sentence consisted of 5 words: "The blue/green

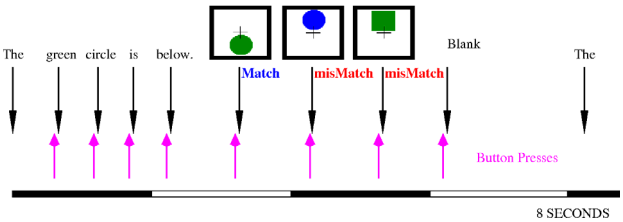

Figure 1. A sample trial is shown. The bar at the bottom is 4 sec long. The black arrows indicate stimuli. The magenta arrows indicate responses to the preceding stimulus. The 2^{nd}-5^{th} word stimuli were triggered by a button press with the index finger. If the test figure matched the sentence, the response was a button press with the index finger. If not, the response was a button press with the middle finger.

circle/square is above/below."

All presentations were placed on a white background. 8 blocks of 40 trials were presented. It was assumed that within each block of 40 trials (~4 min) the head was fixed. The task with all attendant events is schematized in Figure 1. The transition from each stimulus to the next was self-paced, i.e. triggered by a button press (Brain Logics Fiber Optic Button Response System, Psychology Software Tools, Pittsburgh, PA). With each button press, the preceding stimulus, word or test figure, disappeared for 2 screen refresh cycles (~33 msec) before presentation of the next stimulus. Following the button press indicating the match/mismatch choice for the 3^{rd} test figure, the screen was blank (white) for 1.0 sec with the fixation point appearing halfway through this interval.

2.3 Referee Consensus Data Processing

For each trial, 6 560 msec data segments were selected for extraction of virtual recordings using referee consensus processing. The search for active MEG sources was conducted for 80 msec time segments, one at a time, each overlapping the previous segment by 40 msec. The brain volume absent a sphere with 15 mm radius at the center of the head[4] was divided into ≈ 3000 ½ cm^3 cubes. Each instance of the search routine, mvrXS, searched the data from 40 target figures for one of these ½ cm^3 volumes. Each such job required about 1 hr of computing time. Processing the data for each volunteer required: 3000 (½ cc^3 volumes) x 8 (trials blocks) x 3 (target figures) ≈ 72,000 jobs. The calculations were hosted by the Open Science Grid (OSG). The overall usage over an 8 week period was 3.2 million hours of clock time.

mvrXS was written in Fortran 77 and compiled with gfortran. The dot product routines were optimized. The eigenvector decompositions and matrix inversions using Cholesky factorization were handled at 64 bit precision using LaPack [8]. All job control, results aggregation and transport, and housekeeping functions were handled by tcsh scripts.

Each instance of the executable required ≈ 300 Mbytes of core memory at run time. In addition each instance required network transport of 10-15 Mbytes of file data including the mvrXS static image, the MEG data segments, data specific tables, and results files. For each group of 3000 jobs, the MEG data segments were the same but accounted for ≈ 2/3 of the data transport demand. Since our jobs typically ran on only 5-20 facilities at a time, we were able to reduce this by using the http transportation layer provided by Condor which includes local file caching.

The calculation of the referee consensus requires the 4 steps listed below. In order to measure the gradient of the referee consensus, these were carried out 6 times, once for each point adjacent to the target location 1 mm along the x, y, and z axes. Note that for the MEG problem, there are 2 linear variables for each source location.

1. Compute 180 referee templates, 2 for each of 90 referee locations.
2. Compute the first eigenvector of the 180 templates. This is the referee consensus template.
3. Compute 180 referee metrics, 2 for each of the 90 referee locations.
4. Sum the referee metrics to obtain the referee consensus.

[4] Due to geometric and other physical constraints, magnetic fields produced near the center of the head are undetectable. Efforts to measure them result in numeric instability in the computational algorithms.

If there is, in fact, an MEG source present at a location, the two members of the corresponding pair are likely correlated with each other. We can estimate the chance of getting 114 or more out of 180 in the worst case, i.e. when they are perfectly correlated, by reducing the referee consensus by ½: I.e. the chance of getting 57 or more heads out of 90 flips of a fair coin is 0.007; the chance of doing so 6 times in a row is $0.007^6 \approx 10^{-12}$. In addition the fairness of the "coin flips" is potentially compromised by correlations of the referees' transfer functions. Analysis of this potential bias is beyond the scope of this paper.

To identify the starting point for the search of each ½ cm^3 volume, referee consensus was applied once to 34 points scattered evenly through the cubic volume. A maximum of 6 steps guided by the gradient calculation was then allowed before terminating the search. Other efficiencies included applying laboriously generated tables piecemeal across time segments rather than storing their results or repeating them. Together these reduced the computational load by a factor of 40 compared to an exhaustive grid search.

3. Results

For each 80 msec time segment for each single trial, ≈500 active sources were found. For each identified source, the xyz coordinates, the value of the referee consensus, the 80 point time series referee template, and the two components of the amplitude were recorded. Figure 2 shows a typical distribution of values that were obtained for the referee consensus.

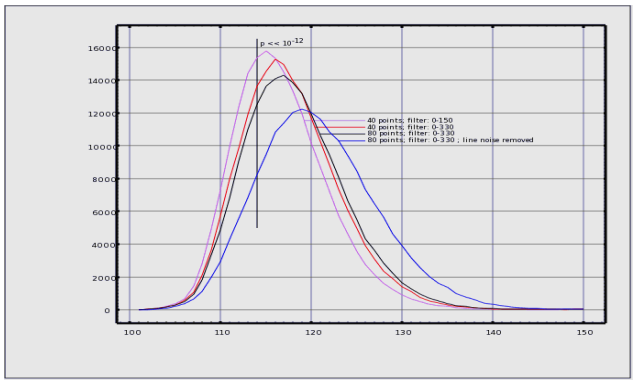

Figure 2. The distribution of values of the referee consensus metric is nearly symmetric and improves with both template length (40:red → 80:black), reduced low pass filtering (150 Hz:mauve → 330 Hz: red), and with removal of line noise (blue) from the MEG prior to processing . The vertical bar at 114 out of 180 for each of 6 tests has a nominal $p < 10^{-12}$ (see text).

The correlation was computed between all source pairs that occurred simultaneously as a function of the distance between the members of the pair. Typical results from 3 subjects are shown in Figure 3. The figure demonstrates that the method resolves sources as near as 2 mm apart.

In order to test for task specific information in the virtual recordings, a two stage process was used. The first stage enabled identification of a restricted volume within which identified virtual recordings were used for further analysis. This was needed to narrow the statistical testing to virtual recordings from a brain region that was likely involved in the task. Otherwise a global search would have been required which would have caused an intractable multiple comparison problem.

Stage I: A novel discriminant pattern source localization (DPSL) method [9] was applied to the MEG recordings. DPSL consists of two steps. (1) A classification algorithm is applied to find a spatial filter to distinguish different brain states. (2) The gain of the spatial filter is calculated for each voxel to reveal the locations of the task-related sources, i.e. the location(s) within the brain whose differential activation under the two conditions is most significant.

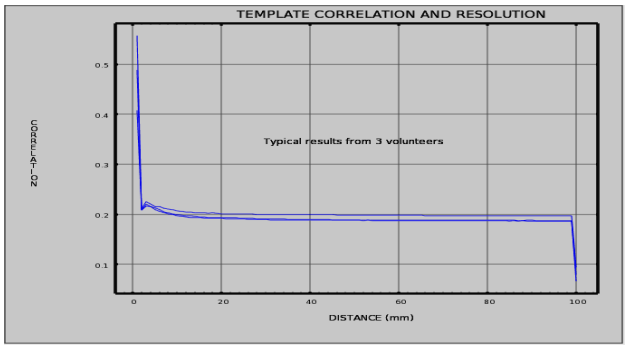

Figure 3. The correlation between all pairs of coincident source templates is shown as a function of distance between the sources from 1 mm to 10 cm. The absolute value of the correlation was used in each case. The variance in the measure across all pairs was ≈ 0.02. These results demonstrate resolving power of 2 mm. The results shown (3 subjects) were typical.

Classification accuracy was tested separately for 7 task relevant binary conditions, 8 frequency band estimates, and 24 time segments, i.e. 1344 tests. The data acquired during the test figure presentations was used. The task relevant binary conditions were as follows: The test figure was (1) blue/green, (2) square/circle, (3) above/below the fixation point, (4) a match/mismatch to the sentence, (5) a match to the sentence preceded by a figure which was a match/mismatch, or (6) a mismatch preceded by a match/mismatch. The frequency band estimates were obtained with wavelet filters for 0-15 hz, 15-30, 30-45, 45-60, 60-75, 75-90, 90-105, and 105-120 hz. Adjacent 32 msec time segments were used beginning 200 msec prior to stimulus onset.

For the subject whose results are shown in Figure 4, DPSL achieved significant classification accuracy in the 0-15 hz band for 2 conditions: (1) above/below (248 msec post stimulus) and (2) match preceded by match/mismatch (152 msec post stimulus) (Figure 4). The gains for the 2 corresponding spatial filters were computed for voxels on a 24 mm cubic grid covering the brain. The single grid location with maximum gain for each classification was then used in stage II.

Stage II: Referee consensus optimization was used to extract virtual recordings. The coordinates of each recording was stored along with the consensus template, the total power and the average amplitude of the recording. Analysis of variance was used with a single binary factor, 3 covariates, and 26 dependent variables. Separate tests were computed for the 2 task relevant binary conditions and 12 time segments, i.e. 24 tests. The binary conditions were those identified in Phase I. For each only virtual recordings were included that fell within 11 mm of the selected grid location. 80 msec time segments were used beginning 40

msec prior to stimulus onset with each overlapping the previous by 40 msec.

The covariates were (1) which block of 40 trials was being presented, i.e. 1 to 8, (2) the time of the onset of the test figure within the trial block, and (3) the square of the time of the test figure onset. The 26 dependent variables were the total power and average amplitude of the recording and 24 frequency estimates of the template obtained with a Fourier transform. These latter were 12.5 hz wide running from 12.5 hz to 300 hz.

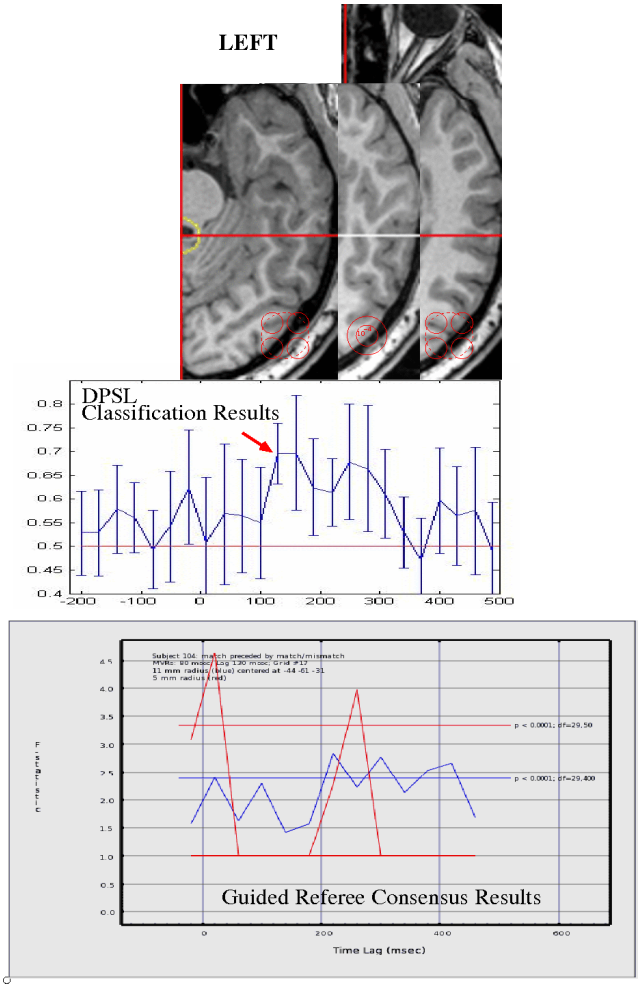

Figure 4. Task relevant information was found in single trial data with two methods in sequence. When the test figure matches the sentence the brain is differentially activated depending on whether the preceding test figure was a match or a mismatch. The middle panel was obtained from MEG recordings using discriminant pattern source localization (DPSL). The lower panel was obtained using MANCOVA from virtual recordings within 11 mm of a location guided by DPSL . The red arrows indicates the most significant DPSL result. The MR slices (upper panel) are 6 mm apart. The referee consensus results shown in blue are for the larger sphere (11 mm radius) centered at the grid location shown in the middle cut. The results shown in red are for the smaller sphere (5 mm radius). The 8 additional grid locations in the slices above and below (5 mm radii) were also tested.

For a match preceded by match/mismatch joint use of all the dependent variables achieved significance for several lags at $p < 0.0001$ (Figure 4). There was no significant finding for any individual dependent variables on any of the tests.

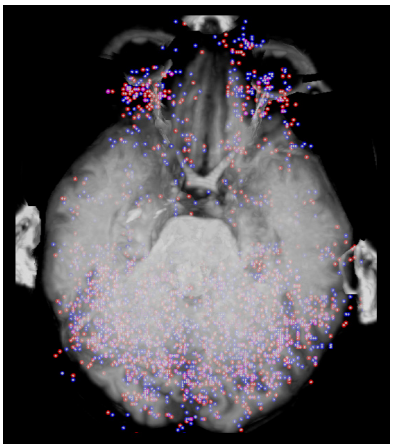

Figure 5. Tonic activation of the retina during visual stimulation. Results are from 120 40 msec single trial epochs processed with referee consensus optimization. Recordings were obtained 30-69 msec after test figure presentation. Activation of the cerebellum is also seen. Only the infra tentorial portion of the brain is shown. Gray matter is not seen. A 3D animation sequence may be found online at http://youtu.be/NStnmfCeZlQ .

4. Discussion

The referee consensus method is computationally expensive. But that expenditure produces intracranial current measurements from single trial data that retain task specific information (Figure 4) and that distinguishes sources 1-2 millimeters apart (Figure 3).

The referee consensus method utilizes linear filters as do all "source space" methods in current use for the MEG problem. And it uses a search as does equivalent current dipole source localization [1]. But in other ways it is quite different: (1) Rather than using a single filter with unit gain for a target location, 1000+ pairs of filters are used for dummy (referee) locations, each with zero gain at or very near the target location. (2) To decide if there is a source at a target location, a probabilistic measure of goodness of fit is constructed from the output of 1000+ filter pairs for that target rather than using a measurement error metric [1] or a post-hoc test on the outputs of the filters for all of the targets [4,6,10]. (3) The time course of the activity at the target location results from a joint estimation procedure applied to the output of all 1000+ filter pairs rather than from the output of a single filter.

Referee selection is a critical element in this method. Figure 6 shows the transfer function of a difference filter, **R!A'-R!A** for one referee. The figure shows why this approach works and that it does not produce ideal filters, i.e. filters with gain 1.0 at the target and 0.0 everywhere else. It seems likely that incremental but significant improvement could be gained by detailed analysis of referee selection. Note that the results of any such analysis will differ for different problems, i.e. different nonlinear functions used to compute the elements of **H**. Since the referee filter design requires at least one zero in each filters' transfer functions, it is

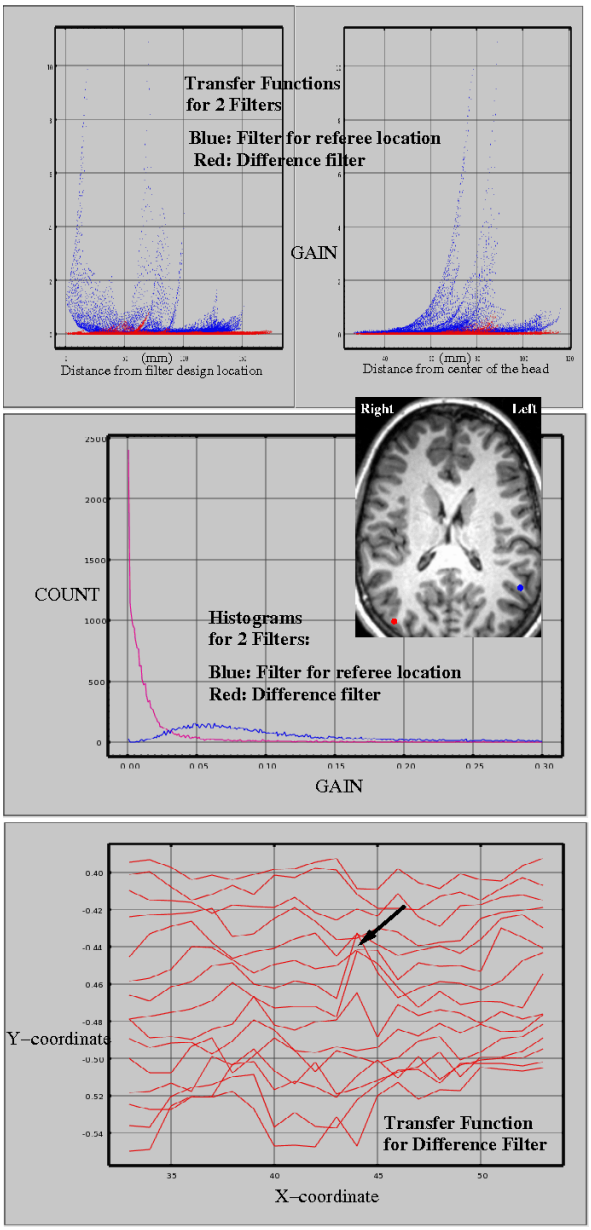

Figure 6. Upper panel: Transfer functions are shown for the filter at a referee location, **R!A,** (blue) and a difference filter, **R!A'-R!A** (red). Both were computed at points on a 1 mm grid in the axial plane 19 mm superior to the center of the head. The corresponding MR slice is shown with referee (blue dot) and target (red dot) locations. Middle panel: These histograms are for the same transfer functions. The difference filter produces a preponderance of low gains and shows a simple relationship to the distance from the center of the head (upper right panel). Lower panel: The transfer function of the same difference filter is shown for the neighborhood of the target location, **A** (arrow). Note that the peak spans 2 points on the y-axis, **A** and **A'**.

natural to consider beam space filter design [11] as an additional approach to improvement.

The formulation presented here is for the "instantaneous" method. The presumption is that there is no delay between the action of a source and the resultant effect at the sensor. There are numerous systems for which this is not the case, i.e. when there is a significant time lag between the action of a source and the resultant measurement. For instance seismic tomography is a problem that deals with time lags of minutes or more since a shock to the earth may be detected and localized using an array of accelerometers 1000's of miles away.

The general referee consensus model which is used to handle lagged relationships between sources and measurements is the set of equations which describe the causal relationship between the sequence of actions in time of many discrete sources and the resultant sequence in time of the measurements. Each source is defined by 2 or more variables, 1 of which is called the amplitude[5]. The key departure from the instantaneous version is that here the measurement and error vectors are sequences of measurements and errors over time and here the time course of each source amplitude is handled as a weighted sum of basis functions.

The connectivity or functional connectome is such a problem since it deals with time lags of 1-100's of msec. For that we consider the virtual recordings extracted from MEG measurements at specific locations to be due to (1) axial currents in nearby axon bundles due to passage of volleys of actions potentials and (2) population post synaptic currents in nearby gray matter due to the arrival of volleys of action potentials. We plan to use High Definition Fiber Tractography [12] to identify fiber tracts and neural populations likely coupled to them. The virtual recordings from sources found near these structures will be used as the measurements to determine the coupling strength of the neural populations to the fiber tract and the propagation times/velocities of the action potential volleys

All variables, parameters, and matrices in the formalism presented here are over the real numbers only. The key properties of these mathematical objects on which the formalism relies are (1) the existence of an inverse for any non-zero number and any non-singular matrix and (2) distributivity of multiplication over addition. These properties hold for the complex numbers, the quaternions, and the octonions, hence referee consensus optimization will likely work over these number systems. Thorough analysis of this assertion remains to be done.

5. Acknowledgements

This work was supported by the Pittsburgh Foundation, the Aberdeen Army Research Laboratory, the Clinical and Translation Science Institute at the University of Pittsburgh, and the Extreme Computing Consortium. Computing resources were provided by the Open Science Grid, which is supported by the National Science Foundation and the U.S. Department of Energy's Office of Science. We thank Bedda Rosario-Rivera and Catalin Trenchea for their invaluable critical comments.

[5] In the more general formulation, more than 1 amplitude variable may be used for each source. 2 or 3 are used in the MEG problem.

6. References

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

[2] Hämäläinen MS, Ilmoniemi R., 1984. Interpreting measured magnetic fields of the brain: Estimates of current distributions. Techical Report TKK-F-A559.

[3] Pascual-Marqui RD; Michel CM; Lehmann D. October 1994. Low-Resolution electromagnetic tomography – A new method for localizing electrical activity in the brain. Intl J Psychophys 18(1):49-65.

[4] Pascual-Marqui RD. 2002. Standardized low resolution brain electromagnetic tomography (sLORETA): technical details. Methods & Findings in Experimental & Clinical Pharmacology 2002, 24D:5-12. URL last visited Dec 2010: http://www.uzh.ch/keyinst/NewLORETA/sLORETA/sLORETA.htm

[5] Huang MX, Dale AM, Song T, Halgren E, Harrington DL, Podgorny I, Canive JM, Lewis S, Lee RR. 2006. Vector-based spatial–temporal minimum L1-norm solution for MEG NeuroImage 31 1025 – 1037.

[6] Huang MX, Theilmann RJ, Robb A, Angeles A, Nichols S, Drake A, D'Andrea J, Levy M, Holland M, Song T, Ge S, Hwang E, Yoo K, Cui L, Baker DG, Trauner D, Coimbra R, Lee RR. August 2009. Integrated Imaging Approach with MEG and DTI to Detect Mild Traumatic Brain Injury in Military and Civilian Patients J Neurotrauma 26:1213–1226.

[7] Krieger D, S Onodipe, PJ Charles, RJ Sclabassi. Real time signal processing in the clinical setting. Annals of Biomedical Engineering Vol 26: 462-472, 1998.

[8] Anderson E, Bai Z, Bischof C, Blackford S, Demmel J, Dongarra J, Du Croz J, Greenbaum A, Hammarling S, McKenney A, Sorensen D. 22 Aug 1999. LAPACK Users' Guide Release 3.0. http://www.netlib.org/lapack/lug/ (Last visited Dec2010).

[9] Zhang J, G Sudre, X Li, W Wang, D Weber, A Bagic. Task-related MEG source localization via discriminant analysis. Proceedings 33rd Intl. Conf. IEEE EMBS: 2351-2354, 2011.

[10] Douglas Cheyne, Leyla Bakhtazad, and William Gaetz 2006. Spatiotemporal Mapping of Cortical Activity Accompanying Voluntary Movements Using an Event-Related Beamforming Approach. Human Brain Mapping 27:213–229.

[11] Robinson SE, Vrba J 1999. Functional neuroimaging by synthetic aperture magnetometry. In: Yoshimine T, Kotani M, Kuriki S, Karibe H, Nakasato N, editors. Recent Advances in Biomagnetism: Proceedings From the 11th International Conference on Biomagnetism. Sendai: Tokoku University Press. p 302–305.

[12] Verstynen T, Jarbo K, Pathak S, Schneider W. 2011 In vivo mapping of microstructural somatotopies in the human corticospinal pathways. J Neurophys 105(1): 336-346.

7. Appendix I

Zeroes of the inverse solution for the instantaneous referee consensus system: Suppose we have measurements, $\vec{B} = \langle b_1 \ \cdots \ b_j \ \cdots \ b_J \rangle$, which contain a contribution from source $\mathbf{d_k}$, i.e.: $\vec{B}^{\mathsf{T}} = \begin{matrix} b_1 = \\ b_j = \\ b_J = \end{matrix} \begin{matrix} \mathbf{h_{1,k}d_k} + \delta_1 \\ \vdots \\ \mathbf{h_{j,k}d_k} + \delta_j \\ \vdots \\ \mathbf{h_{J,k}d_k} + \delta_J \end{matrix}$. The $\delta' s$ are due to the contributions of all other source, known and unknown. The inverse solution is $\vec{D}^{\mathsf{T}} = \left(\left(\mathbf{H}^{\mathsf{T}} \mathbf{H} \right)^{-1} \mathbf{H}^{\mathsf{T}} \right) \vec{B}^{\mathsf{T}} = \Xi \vec{B}^{\mathsf{T}}$.

We expand $\vec{D}^{\mathsf{T}}$ and substitute for $\vec{B}^{\mathsf{T}}$ in the inverse solution:

$$\begin{pmatrix} d_1 \\ \vdots \\ d_k \\ \vdots \\ d_K \end{pmatrix} = \Xi \begin{pmatrix} \mathbf{h_{1,k}d_k} + \delta_1 \\ \vdots \\ \mathbf{h_{j,k}d_k} + \delta_j \\ \vdots \\ \mathbf{h_{J,k}d_k} + \delta_J \end{pmatrix} = \Xi \begin{pmatrix} \mathbf{h_{1,k}d_k} \\ \vdots \\ \mathbf{h_{j,k}d_k} \\ \vdots \\ \mathbf{h_{J,k}d_k} \end{pmatrix} + \Xi \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_j \\ \vdots \\ \delta_J \end{pmatrix}.$$

Since $\begin{pmatrix} \mathbf{h_{1,k}d_k} \\ \vdots \\ \mathbf{h_{j,k}d_k} \\ \vdots \\ \mathbf{h_{J,k}d_k} \end{pmatrix}$ is the k^{th} column of $\mathbf{H}$ and $\Xi \mathbf{H} = \left(\left(\mathbf{H}^{\mathsf{T}} \mathbf{H} \right)^{-1} \mathbf{H}^{\mathsf{T}} \right) \mathbf{H} = \mathbf{I}$, the identity,

$$\Xi \begin{pmatrix} \mathbf{h_{1,k}d_k} \\ \vdots \\ \mathbf{h_{j,k}d_k} \\ \vdots \\ \mathbf{h_{J,k}d_k} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ \mathbf{d_k} \\ \vdots \\ 0 \end{pmatrix}, \text{ and } \begin{pmatrix} d_1 \\ \vdots \\ \mathbf{d_k} \\ \vdots \\ d_K \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ \mathbf{d_k} \\ \vdots \\ 0 \end{pmatrix} + \Xi \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_j \\ \vdots \\ \delta_J \end{pmatrix}.$$

In other words, the only source for which the inverse solution includes a contribution from the measurements due to $\mathbf{d_k}$ is $\mathbf{d_k}$. All others have a contribution from the measurements due to $\mathbf{d_k}$ of 0.