Title
Algorithms for Biomagnetic Source Imaging with Prior Anatomical Physiological Information

Author
Hughett, P.

Publication Date
1995-12-01

LBL-38115
UC-405

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

Algorithms for Biomagnetic Source Imaging with Prior Anatomical and Physiological Information

P.W. Hughett
(Ph.D. Thesis)

December 1995

Donner Laboratory

Biology & Medicine Division

Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098

DISCLAIMER

LBL-38115

Algorithms for Biomagnetic Source Imaging With Prior Anatomical and Physiological Information

Paul William Hughett
Ph.D. Thesis

Department of Electrical Engineering and Computer Sciences
University of California, Berkeley

and

Life Sciences Division
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

December 1995

This work was supported in part by the Advanced Technology Systems Division of the General Electric Company, the Lawrence Livermore National Laboratory through the University Collaborative Research Program of the Institute for Scientific Computing Research, a Faculty Research Grant from the University of California, the National Science Foundation 1995 Summer Institute in Japan, and the Director, Office of Energy Research, Office of Health and Environmental Research, Medical Applications and Biophysical Research Division, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Algorithms for Biomagnetic Source Imaging With Prior Anatomical and Physiological Information

by

Paul William Hughett

B.S. (Massachusetts Institute of Technology) 1986

M.S. (Massachusetts Institute of Technology) 1986

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Engineering—Electrical Engineering and Computer Sciences

in the

GRADUATE DIVISION

of the

UNIVERSITY of CALIFORNIA at BERKELEY

Committee in charge:

Professor Thomas F. Budinger, Chair
Professor Stephen E. Derenzo
Professor F. Alberto Grünbaum

1995

Abstract

Algorithms for Biomagnetic Source Imaging

With Prior Anatomical and Physiological Information

by

Paul William Hughett

Doctor of Philosophy

Engineering—Electrical Engineering and Computer Sciences

University of California at Berkeley

Professor Thomas F. Budinger, Chair

This dissertation derives a new method for estimating current source amplitudes in the brain and heart from external magnetic field measurements and prior knowledge about the probable source positions and amplitudes. The minimum mean square error estimator for the linear inverse problem with statistical prior information was derived and is called the optimal constrained linear inverse method (OCLIM). OCLIM includes as special cases the Shim-Cho weighted pseudoinverse and Wiener estimators but allows more general priors and thus reduces the reconstruction error. Efficient algorithms were developed to compute the OCLIM estimate for instantaneous or time series data. The method was tested in a simulated neuromagnetic imaging problem with five simultaneously active sources on a grid of 387 possible source locations; all five sources were resolved, even though the true sources were not exactly at the modelled source positions and the true source statistics differed from the assumed statistics.

The reconstruction quality can be characterized by the error (reconstructed versus true source distribution), residual (reconstructed versus actual measurements), and surprise (likelihood of the reconstruction given the priors). The mean and variance of these quality measures were derived for the case that only the source and noise covariances are known; the exact distribution was derived for Gaussian source and noise statistics. Confidence limits for the reconstruction were derived.

"Misinformative" statistical priors inconsistent with the actual source amplitudes, or source models not including the actual source locations, can yield inaccurate or incorrect reconstructions. OCLIM includes a statistical test for this possibility but it may not work reliably at the low signal-to-noise ratios of biomagnetic imaging. Thus an OCLIM reconstruction may suggest source locations but does not rule out the existence of sources not included in the model.

An additional contribution was to develop a generalized chi-squared distribution in which arbitrary means, variances, and covariances are allowed; the same distribution describes the squared norm of a Gaussian random vector. The mean, variance, and characteristic function of the generalized chi-squared distribution were found. A new FFT-based algorithm with error bounds was developed to compute this distribution, and is applicable to the calculation of other continuous distributions.

Professor Thomas F. Budinger, Chair

Acknowledgements

I am grateful to my research advisor and committee chair Prof. Thomas F. Budinger for suggesting the problem of using prior information in magnetic source imaging and to him and to my other committee members Prof. Stephen E. Derenzo and Prof. F. Alberto Grünbaum for the advice and criticism that helped me solve this problem and polish this written presentation. Dr. James G. Berryman pointed out the connections to Marquardt regularization and iterative nonlinear reconstruction algorithms. Prof. Solomon Eisenberg and Dr. Weiping Wang of Boston University made available their finite element model of the human torso and helped me to set up and run a set of simulations designed to determine the significance of the volume currents. Dr. Satoru Miyauchi made the facilities of his laboratory available to me during my 1995 summer stay in Tokyo; both he and Dr. Kensuke Sekihara helped me develop my ideas during discussions and introduced me to Japanese culture.

I am also grateful to my friends Richard and Linda Bull who helped to keep me in touch with the world outside academia and who provided an refuge when the pressures got too overwhelming.

This work was supported in part by the Advanced Technology Systems Division of the General Electric Company, in part by Lawrence Livermore Laboratory through the University Collaborative Research Program of the Institute for Scientific Computing Research, in part by a Faculty Research Grant from the University of California, in part by the National Science Foundation 1995 Summer Institute in Japan, and in part by the Director, Office of Energy Research, Office of Health and Environmental Research, Medical Applications and Biophysical Research Division of the Department of Energy under Contract No. DE-AC03-76SF00098.

Table of Contents

List of Figures

Chapter 1
Introduction

1.1 Background

The depolarization and repolarization of nerve and muscle cells is accompanied by ion currents through the cell membrane. Since the intercellular medium is conductive, these "primary currents" generate in turn "volume currents" in the body [37]. The total current consisting of both the primary and volume currents generates magnetic fields inside and outside the body. The magnetic fields resulting from the depolarization of a single neuron or myocyte can be detected in vitro but are not large enough to be detected outside the body. However, the coordinated activity of tens or hundreds of thousands of neurons in the brain or a similar volume of myocardium does produce magnetic fields which can be detected by arrays of SQUID (Superconducting QUantum Interference Device) detectors placed near the head or chest. The recording of these fields is known as magnetoencephalography (MEG) or magnetocardiography (MCG).

The generated magnetic fields are quite small (femtoTesla to picoTesla) in comparison to the earth's magnetic field (about 50 microTesla) and are easily concealed by naturally and humanly generated magnetic noise. Baule and McFee [6] appear to have been the first to detect the magnetic field of the heart, using a coil of 2 million turns in an open field to reduce the background magnetic noise. Cohen and his colleagues [17, 18, 19] introduced the use of magnetically shielded rooms and cryogenically cooled SQUIDs; this is still the only practical method of detecting and measuring these tiny fields in a laboratory or hospital environment.

Fagaly [28] is a fairly recent (1990) review of the state of the art in MEG and MCG instrumentation.

The volume currents also generate electric potential differences through Ohm's Law; these can be detected by electrodes placed in the body or on the body surface. The recording of such potentials from the brain and heart are known as electroencephalography (EEG) and electrocardiography (ECG). The potentials are in the microvolt to millivolt range; they are not difficult to measure, although some care is required to obtain good quality signals.

The recording and analysis of these electric and magnetic signals naturally generated by the brain and heart provides valuable information about the normal and pathological functioning of these organs and are useful in both clinical and research settings. The ECG in particular has proven invaluable in the diagnosis of cardiac arrhythmias. The magnetic and electric signals arise from the same sources and contain approximately the same information. Since it is easier to collect the electric data, it is more often used in practice.

If signals are collected by an array of detectors rather than a single detector, it is also possible to reconstruct the location, orientation, and amplitude of current sources within the body but high quality data is needed for accurate reconstruction. These techniques are known as magnetic source imaging (MSI) and electric source imaging (ESI). Magnetic and electric source imaging are complementary, since the magnetic field at the surface is most strongly determined by current sources directed tangentially to the surface and the electric potentials are determined by current sources directed perpendicular to the surface. It has not yet, however, proven practical to simultaneously collect both magnetic and electrical data. MSI should offer better reconstruction accuracy than ESI, because all body tissues are magnetically transparent and the magnetic fields propagate to the surface without distortion, whereas the electric potentials are distorted by conductivity inhomogeneities in the body. MSI is, however, more expensive because a shielded room and cyrogenically cooled detectors are needed to collect good magnetic data.

Because of its ability to locate the current sources, magnetic source imaging is useful in studies of heart [95] and brain [41] functioning. There are promising clin-

ical applications in cardiology and epilepsy. It is possible to non-invasively localize accessory pathways in Wolff-Parkinson-White syndrome prior to catheter ablation. MSI is also used in the surgical treatment of intractable epilepsy to locate both the epileptigenic focus to be resected and the functional areas of the brain to be conserved during surgery [81, 89, 87, 96]. There are other potential uses in neuroscience to delineate the spatial and temporal sequence of cognitive processing events in the brain in response to auditory, visual, and somatosensory stimuli [3, 27, 53, 83]. Biomagnetism also offers a research tool for schizophrenia [80] and for Parkinson disease. The function of peripheral sensory and motor nerves can be studied [84, 107] as can the prenatal MCG. A major strength of MSI is that it can resolve events separated by milliseconds, whereas other methods such as functional magnetic resonance imaging (fMRI), magnetic resonance spectroscopy (MRS), positron emission tomography (PET) and single-photon emission tomography (SPECT) have time resolutions of seconds to many minutes, depending on the information sought.

The central problem of magnetic source imaging is to reconstruct the current distribution inside some inaccessible volume from magnetic field measurements made outside that volume. It is an example of an inverse problem, as opposed to the forward problem of determining the magnetic field measurements from a known current distribution.

Magnetic source imaging shares with many other inverse problems two characteristics that make inverse problems generally more difficult to solve than the corresponding forward problems. First, the given data do not uniquely determine the solution. That is, there are infinitely many current distributions confined to the head or chest that generate the same external magnetic fields. Second, inverse problems are often ill-conditioned in that small changes in the given data can yield large changes in the reconstruction. As a result, the reconstruction often has high variance and poor resolution.

Both of these problems can be partially ameliorated by the use of prior information about the possible or probable solutions. The primary current sources are confined to certain types of tissue, principally the myocardium and the cortex of the brain; the location of these tissues in a particular patient can be determined

by a preliminary magnetic resonance or computed tomography scan. The source amplitudes are limited by the physiology of the heart and brain. Since coordinated activation of many neurons or myocytes is necessary to generate detectable signals, the source currents are spatially correlated. Requiring that the source current reconstruction be consistent with both the given data and these anatomical and physiological priors can be expected to improve the reconstruction accuracy.

Priors used in this way are intended to be "informative," to provide information not present in the magnetic field data. It is, however, possible that the priors are poorly chosen or that a particular data set is abnormal. If so, the assumed priors may be "misinformative" and lead to a less accurate reconstruction. Thus, the use of priors potentially carries both benefits and risks: informative priors should increase the reconstruction accuracy but misinformative priors may decrease it.

The purpose of this dissertation is to develop constrained reconstruction algorithms for magnetic source imaging (or other applications) for computing source current reconstructions that are consistent with both the given magnetic field data and with anatomical and physiological prior knowledge. A second purpose is to explore the problem of misinformative priors.

1.2 Prior Research

A variety of reconstruction algorithms have been used for the magnetic source imaging problem; see Sarvas [86] and van Oosterom [100] for reviews on this topic.

One approach is to model the unknown current distribution as one or more current dipoles with unknown position, orientation, and magnitude and then to find the unknown parameters by a least squares fit to the observed measurements [50]. This method is computationally expensive because it is nonlinear in the unknowns and iterative solution is required. Worse, the method is often numerically unstable for two or more dipoles.

Various other methods have been used for this problem including the method of alternating projections [65] and beamforming [104, 103].

A more recent approach models the unknown distribution as an array of dipoles with fixed positions but unknown magnitudes [45, 79, 92, 105, 106]. Then the

magnetic field measurements $\mathbf{b}$ can be written as a linear function $\mathbf{b} = \mathbf{Fq} + \mathbf{w}$ of the unknown current distribution $\mathbf{q}$ and measurement noise $\mathbf{w}$. The forward transfer matrix $\mathbf{F}$ is determined by solving the forward problem for unit sources. The inverse problem in this form can be solved directly, without iteration.

There are other medical imaging methods, including transmission tomography and magnetic resonance imaging, which provide high resolution on anatomical structure but which do not provide the same functional information that MSI does. A major theme of this dissertation is the use of prior information, obtained by anatomic imaging or other methods, to improve the resolution and accuracy of magnetic source imaging by constraining the set of possible solutions.

The prior knowledge must define at least the possible spatial locations of the sources and the locations of the detectors. This provides sufficient information to solve the forward problem and define the matrix $\mathbf{F}$. Given only this much prior information, the natural method for the inverse problem is the least-squares or minimum-norm least-squares (MNLS) method [105, 106]. This method (also known as Moore-Penrose inverse or pseudoinverse method) finds the current distribution that minimizes the squared difference between the measured fields and the fields generated by the reconstructed current distribution. Ferguson et al. [30] have developed a preconditioned pseudoinverse using a priori source amplitudes which, in their tests, gave better results than either truncation or Tikhonov regularization.

If the statistics of the measurement noise are available, maximum-likelihood (ML) methods are appropriate [90]. These methods maximize the likelihood of obtaining the measured fields given the reconstructed current distribution; if the noise is jointly Gaussian, they also minimize the squared difference between the measured and reconstructed fields, weighted to reflect the a priori noise variance.

If prior anatomical information is available from transmission tomography or magnetic resonance imaging, then it may be possible to define the a priori source amplitude as a function of position. Then minimum mean-square error (MMSE) methods are appropriate; they minimize the mean (average) squared difference between the true and reconstructed current distributions. Shim and Cho [91] have developed methods using a weighted pseudoinverse but their methods are optimal

and useful only when the a priori source variance is constant. Helstrom [42] has developed and applied an MMSE method for image restoration; Smith et al. [93] have developed and tested an MMSE method for magnetic source imaging with general a priori source variances; Franklin [34] has developed an MMSE estimator in the general context of Hilbert spaces.

1.3 An Approach to the Problem

The inverse problem with prior information can be represented symbolically as follows. We have some measured magnetic field data $\mathbf{b}$ which was generated by some unknown source current distribution $\mathbf{q}$ via some known functional dependence $\mathbf{b} = \mathbf{F}(\mathbf{q}) + \mathbf{w}$ with measurement noise $\mathbf{w}$. We are given in addition some prior knowledge $\mathcal{P}$ and seek an estimator $\mathbf{H}$ to compute an estimate $\hat{\mathbf{q}} = \mathbf{H}(\mathbf{b}, \mathcal{P})$ of $\mathbf{q}$ which satisfies some admissibility conditions $\mathcal{A}$ and which is maximally consistent with both the data $\mathbf{b}$ and the priors $\mathcal{P}$ as measured by some objective function $S(\hat{\mathbf{q}}, \mathbf{b}, \mathcal{P})$.

Developing an algorithm for constrained reconstruction requires making several interlocking choices: the form in which the unknown source distribution $\mathbf{q}$ is to be represented; the forward model $\mathbf{F}$ to be used; the admissibility conditions $\mathcal{A}$; the form in which the prior knowledge $\mathcal{P}$ is to be represented; the class of functions from which the estimator $\mathbf{H}$ is chosen; and the objective function S for choosing the "best solution" among those that are at least minimally consistent with the data and priors.

These choices represent a compromise among several desiderata: First, they should lead to an analytical or numerical solution method for the inverse problem; second, the method should involve an objective criterion for correctness or optimality; third, the parameters used in the source model should be physically meaningful, so that the user can understand and interpret the reconstruction; fourth, the required priors should be directly measureable (at least in principle) by some calibration experiment; fifth, the method should incorporate some test for misinformative priors; and sixth, the reconstruction should include statistical confidence limits.

Several of these desiderata relate to Popper's demarcation criterion for distinguishing scientific (or empirical) statements from non-scientific statements [72]. Popper asserts that a statement or hypothesis is scientific only if it is potentially falsifiable; that is, if there exists some logically possible experimental outcome that would disprove it. Scientific hypotheses generally involve reproducible experiments, since a statement about a non-reproducible experiment usually cannot be falsified. Scientific statements are corroborated to the extent that they have been subjected to serious attempts at falsification but have not been falsified. Corroboration does not guarantee truth but statements that have been corroborated are generally more trustworthy than uncorroborated statements.

In the contex of the inverse problem, the use of an objective optimality criterion helps to assure that the reconstructions are reproducible between analysts. Potentially measureable priors can be falsified by the calibration experiment and are thus scientific. If the reconstruction includes confidence limits, it can also be tested and perhaps falsified by more definitive (perhaps invasive) tests.

Many of these choices can be viewed as imposing constraints on the reconstruction. It is useful to distinguish hard and soft constraints. A hard constraint will never be violated by the reconstruction. For example, the source model is a hard constraint in that the reconstruction will never yield a source not contained in the source model. Conversely, a soft constraint may be violated, if the data support the violation. For example, the reconstruction will tend to assign larger source amplitudes at locations considered more likely to contain a source than at locations less likely, unless the data clearly indicate that a less likely source is responsible for the data.

Although each of the necessary choices will be discussed at length in the following chapters, it is helpful to consider them briefly now.

1.3.1 Source Model

The source model defines the form of the solution and limits the set of possible solutions. For example, expanding the unknown current distribution in a series of basis functions limits the possible solutions to the space spanned by that basis.

Similarly, requiring that the current density be zero outside the body limits the possible solutions. Choosing an appropriate source model can eliminate physically or physiologically impossible solutions and reduce the computational effort needed for reconstruction. An ill-chosen source model, of course, may not admit or even approximate the true distribution. Increasing the number of free parameters in the source model will generally increase both the computational cost and the precision with which the true distribution can be approximated.

We will see that the magnetic field is a linear function of the current distribution. That is, the field due to a weighted sum of current sources is equal to the same weighted sum of the fields due to the individual sources. Thus, it is convenient to represent the unknown current distribution as a weighted sum of elementary sources. The sources are chosen in advance to reflect a priori knowledge as to the location and orientation of the source currents; the weights are to be found by the inverse computation. If there is some prior information as to the magnitude of each source, it can be expressed as a probability distribution on the corresponding weight.

For example, in magnetoencephalography a set of current dipoles may be chosen as the elementary sources. The position and orientation of these dipoles are chosen to match the anatomy determined by magnetic resonance imaging. The volume currents induced by these dipoles may be included or ignored in computing the magnetic field due to each dipole.

More generally, a detailed electromagnetic finite element model [60, 97] of the brain and head could be used, choosing as the elementary sources a dipole current in each element that might contribute to the field. The volume currents are computed with the finite element model, as are the magnetic fields due to the elementary source and volume currents.

The source model is a hard constraint, since the reconstructed source distribution must be contained in the span of the selected elementary sources.

1.3.2 Forward Model

The forward model contains the physics of the problem, specifying exactly how

the measured magnetic fields depend on the current distribution. It is possible, for example, to choose a simple physical model in which the magnetic field outside the body is considered to depend on only the primary currents. This model is valid for certain simple geometries in which the volume currents have no net effect on the magnetic fields but it is only approximately true for more realistic body geometries. It does, however, have the advantage of theoretical simplicity and can be inexpensively computed. This simple model may be satisfactory when only low accuracy is required.

Another possible choice is to assume a simple spherical or cylindrical geometry which approximates the true shape of the body but for which it is possible to obtain explicit analytical solutions for the volume currents and magnetic fields. This has often been found satisfactory for magnetic source imaging of the brain, where the head is modelled as three concentric spheres with different conductivities.

A third choice is to use a detailed geometric model of the head or torso, possibly obtained from magnetic resonance imaging, and to compute the volume currents by finite element or boundary element methods. This choice is computationally expensive but produces the most accurate results.

The principal tradeoff involved in the choice of forward model is computational cost versus accuracy.

The choice of forward model generally imposes a hard constraint, often consistency with known physical laws. The requirement of charge conservation is an obvious example.

1.3.3 Prior Knowledge

The boundaries between prior knowledge and the other parts of the inverse problem are not always distinct. For example, the prior knowledge that all sources must be in the cortex or the myocardium may be most conveniently represented as part of the source model. Similarly, the requirement for charge conservation will be automatically incorporated in a finite element forward model.

A common form of prior knowledge that does not fall into any of the other categories is statistical knowledge about the source distribution and the measurement

noise. Statistical prior knowledge is a soft constraint, since it defines a preference for the more probable solution over the less probable but does not absolutely exclude any solution.

The imposition of priors into the inverse problem introduces a potential risk, since it is not guaranteed that the given data will in fact be consistent with the priors and it is imprudent at best force the reconstruction to be consistent with the priors when the data are not. This implies that a constrained reconstruction algorithm should—if possible—include some test to determine whether or not the given data are consistent with the assumed priors.

1.3.4 Admissibility Conditions

It is appropriate in many applications to impose admissibility conditions on the solution. For example, the test mentioned in the last section as to whether or not the data are consistent with the priors is just such an admissibility constraint. For another, it might be reasonable to require that the fields due to the reconstructed distribution match the measured fields to within some prespecified error. Such admissibility conditions can, of course, be incorporated in the objective function by adding a term which is zero if the conditions are satisfied and $+\infty$ if they are not, but it is often convenient to consider them separately. Admissibility conditions are hard constraints by definition.

1.3.5 Estimator Class

The estimator $\mathbf{H}$ may be restricted to a linear function or to a function of some particular parametric form. The class of allowable estimators is often chosen so that it is easy to determine the parameters defining the particular estimator or so that the estimate is easy to compute. As a general rule, the more restricted the class from which the estimator is drawn, the less good the resulting estimate is. That is, it is often possible to compute a poor estimate quickly, but a good estimate will require more time either in analysis or computation.

Given a particular class of estimators, there are often many different numerical algorithms for computing the estimate. For example, a linear estimator may

be defined as a system of linear equations and solved by any of a variety of algorithms, depending on the detailed structure of the linear system. The choice of an appropriate algorithm may have a major impact on the computational cost, but all algorithms for a given estimator should compute essentially the same solution, apart from roundoff and truncation errors.

1.3.6 Objective Function

There are many dimensions along which the quality of the reconstruction can be measured. For example, the reconstruction should match the true source distribution as closely as possible; it should be consistent with the priors assumed for the problem; and the reconstructed magnetic fields should match the measured fields as closely as possible. In addition, it may be desireable that the reconstruction be "smooth" in some defined sense.

It is not possible, in general, to achieve the best possible reconstruction on all dimensions simultaneously. The purpose of the objective function is to reduce all the conflicting requirements to a single measure of quality so that it is possible to decide which is the "best" estimator out of the class of all admissible estimators. Different choices of the objective function will, of course, lead to different "optimal" estimators and it may be useful to consider the tradeoffs between different objective functions. The objective function is a soft constraint.

Once the objective function is chosen, it may be of interest to examine how well the corresponding estimator does on other measures of reconstruction quality.

The objective function and admissibility conditions together define the most desireable estimator within the given class. For example, the "best" estimator may be the one that matches the true source distribution as closely as possible on average subject to the constraint that it must match the measured fields within a specified error.

1.4 Overview of the Dissertation

Chapter 2 discusses source modelling and the biomagnetic forward problem of determining the field measurements for a given source current distribution. Several

forward models are presented, from fast and simple to slow and accurate. Practical methods for computing the volume currents in realistic body geometries are also covered.

Chapter 3 discusses the nature and formalization of the prior knowledge, focussing on statistical prior knowledge of the source and noise amplitudes. Attention is paid to the physical interpretation of the priors and the way in which they could possibly be determined by experiment.

Chapter 4 presents a unified development of the linear inverse problem, perhaps with prior knowledge, including the MNLS, ML, weighted pseudoinverse, and MMSE methods and derives the optimal constrainted linear inverse method (OCLIM) as the most general and most accurate of these methods. It shows further that the weighted pseudoinverse, Gaussian maximum likelihood, and MNLS methods can all be obtained as special cases of OCLIM by an appropriate choice of priors.

Chapter 5 defines a generalization of the χ^2 distribution which will be needed to describe the statistics of reconstruction quality in the next chapter, and develops a numerical algorithm for computing it. As it turned out, the generalized χ^2 distribution is not essential to the use of OCLIM, but it appears to have other uses in statistics.

Chapter 6 discusses several useful measures of reconstruction quality and develops, for the estimators of Chapter 4, formulas for the mean reconstruction error, mean residual, a posteriori variance, and confidence limits. Several of these quality measures are random variables distributed according to the generalized χ^2 distribution.

Chapter 7 discusses efficient computational methods for OCLIM and shows the results from a Monte Carlo simulation to verify the theoretical predictions of the previous chapters and to compare reconstruction quality between methods with and without statistical priors for the sources. It also demonstrates that misinformative priors can, in fact, yield less accurate reconstructions than uninformative priors.

Chapter 8 attempts to address the problem of misinformative priors by developing and demonstrating by simulation a statistical test for misinformative priors

that detects some but not all data sets inconsistent with the priors. This test, unfortunately, does not appear to work reliably at the low signal-to-noise ratios characteristic of magnetic source imaging; furthermore, it appears that this inability is characteristic of the magnetic imaging problem rather than the specific test developed.

OCLIM provides an explicit estimate of the reconstruction error for given priors. Chapter 9 shows how this estimate can be used as an objective function for the optimal design of magnetometer arrays for MSI. That is, the geometry of the array can be chosen to provide the most accurate reconstruction for given priors, within the limits of technology.

OCLIM can be used to reconstruct dynamic sources distributions from time series of magnetic field measurements but a direct implementation is computationally expensive. Chapter 10 shows that the computational cost can be reduced by several orders of magnitude by making some assumptions about the nature of the spatiotemporal correlations between sources and gives specific computational algorithms. Also shown are some sample reconstructions using these algorithms.

The preceeding chapters investigated the characteristics of OCLIM in simplified problems. Chapter 11 demonstrates its performance in a realistic neuromagnetic source imaging problem; it also summarizes what appears to be the best implementation of OCLIM for the non-time-series case.

Finally, Chapter 12 summarizes the important results and indicates unresolved questions for future research.

Chapter 2

Biomagnetic Forward Problem

An inverse problem solver does not itself specify the physics of the problem to be solved but instead depends on a forward model. That forward model may be as simple or as detailed as the user feels is appropriate and the inverse problem solver will produce a solution consistent with the particular forward model used.

The first part of this chapter shows that, under a broad range of assumptions, the forward model for magnetic source imaging can be written in the linear form

$$\mathbf{b} = \mathbf{F}\mathbf{q} + \mathbf{w} \quad , \tag{1}$$

where $\mathbf{b}$ is a vector of field measurements, $\mathbf{q}$ is a vector of source amplitudes, $\mathbf{F}$ is a forward transfer matrix, and $\mathbf{w}$ is a vector of measurment errors or noise.

The forward transfer matrix $\mathbf{F}$ will, of course, depend on the particular assumptions made and the reconstructed source distribution for any given data set will depend on the forward model via the matrix $\mathbf{F}$. If, for example, the forward model assumes that the sources are current dipoles in an infinite homogeneous medium, that the volume currents can be ignored, and that the detectors measure the magnetic field at a point, then $\mathbf{F}$ will have one value and the inverse problem solver will reconstruct the source amplitudes under those assumptions. If, instead, the forward model uses the actual torso geometry, includes volume currents, and models magnetometer coils with finite area, then $\mathbf{F}$ will have a different value and the inverse problem solver will reconstruct the source amplitudes under these different assumptions.

The middle part of this chapter develops several specific forward models and shows how the forward transfer matrix $\mathbf{F}$ may be computed in each case, given a

forward problem solver for computing the magnetic field for a specified source current distribution. The models range in realism from a simple dipole model with no volume currents and point magnetometers to a sophisticated model with distributed current sources, volume currents in a realistic body shape, and magnetometer coils of arbitrary shape (including gradiometer configurations).

In the most sophisticated and accurate forward model considered, the volume currents and magnetic fields must be computed numerically using a finite element model (FEM) or boundary element model (BEM); the final part of the chapter discusses these methods in more detail.

2.1 Previous Research

The forward problem has received considerable attention. The bioelectric forward problem is to compute potentials and currents within the body and at the surface from the primary current sources and the conductivity distribution within the body. The biomagnetic forward problem adds to this the problem of computing the magnetic fields induced by the primary and volume currents.

The general area of electrophysiology is fundamental to the biomagnetic forward has received considerable research attention. Plonsey [71] discusses the general theory of bioelectric phenomena. Nunez [64] discusses the theory of electroencephalography in particular. Bayley [8] is a monograph on the theory of vector electrocardiography, primarily directed toward physicians but including some of the physics and mathematics.

Foster [32], Karlon [48] and Rush [85] give physiological values for the electrical conductivity and dielectric constants of tissue.

The following references discuss the formulation of the forward problems in general terms; methods of solution for specific versions of the problem will be discussed later.

McFee and Johnston [56, 57, 58] develop a theory of electrocardiography in terms of lead fields. Brody et al. [12] generalize this to compound leads and a lead tensor.

Plonsey [69] discusses applications of the method of images to the problem of the current flow due to a current dipole in a spherical or cylindrical conductor. In a companion paper [70], he shows how the Helmholtz reciprocity theorem can be applied to problems in electrocardiography and relates it to lead field theory.

Geselowitz [36, 35] gives explicit surface integrals for the magnetic field generated in a piecewise homogeneous volume conductor of arbitrary shape. Baule and McFee [7] give a similar analysis using lead field theory.

2.2 Source Models

It is convenient to represent the unknown current distribution $\vec{J}(\vec{r})$ as a weighted sum of N known elementary sources $\vec{J}_n(\vec{r})$ in the form

$$\vec{J}(\vec{r}) = \sum_{n=1}^{N} q_n \, \vec{J}_n(\vec{r}, P_n) \quad , \tag{2}$$

where q_n is a source amplitude and P_n is a (possibly empty) vector of other parameters describing the elementary source. Each elementary source $\vec{J}_n(\vec{r})$ is a vector-valued function giving the vector current density at any position $\vec{r}$. The inverse problem is to estimate the values of the unknown parameters q_n and P_n.

There are many possibilities for the set of elementary sources $\vec{J}_n(\vec{r})$, depending on the assumptions made about the unknown distribution $\vec{J}(\vec{r})$.

If the unknown distribution is assumed to be well-described by a few localized sources, then it is reasonable to use a few current dipoles whose positions, magnitudes, and orientations are to be determined [50]; then P_n includes the source positions and orientations. These "moving dipole" methods were the first developed for magnetic source imaging and yield a conceptually simple distribution containing only a few discrete sources. They are, however, computationally expensive because the nonlinear dependence of the measured fields on the source positions generally requires iterative solution.

The methods described in this report assume positions fixed a priori and are not directly useful in a moving diple formulation; they could possibly be used to find optimal source amplitudes at each step of an iterative scheme for improving the source position estimates.

The estimation of moving dipoles is expensive because the measured fields depend nonlinearly on the position and iterative solution is required; if only the source amplitudes are unknown, it becomes possible to solve the problem directly. But if the source model consists of only a few dipoles, it is not reasonable to assume that their positions are known a priori; there are too many possible source locations in the brain or heart.

An alternative is to assume that the unknown distribution is smooth and to expand it in a set of basis functions at fixed positions; these basis functions are the elementary current sources. A grid of current dipoles [106] or a finite element mesh [60, 97] define localized elementary sources; lead fields [106], multipole expansions [101], or Fourier basis functions define non-localized elementary sources. Then the unknown current distribution is completely defined by the weights $q_1, q_2, \ldots, q_N$ in the summation

$$\vec{J}(\vec{r}) = \sum_{n=1}^{N} q_n \, \vec{J}_n(\vec{r}) \quad , \tag{3}$$

and can be completely described by a vector $\mathbf{q} = [q_1, q_2, \ldots, q_N]^T$.

The choice of the source model generally involves a tradeoff between accuracy and computational cost, so it is preferable to use a source model no more complex than is necessary for accuracy. De Munck [25] has compared the potentials at the surface of the head due to dipoles versus more realistic models of physiological current sources for various conductivity models; he concludes that there is no significant difference between the realistic sources and dipole sources for EEG. Fender [29] reviews various source models used for dipole localization in the head and discusses their impact on the reconstruction accuracy.

2.3 Primary Currents and Volume Currents

The source models just described do not distinguish between the primary physiological current sources and the volume currents that they induce. In fact, the total current distribution consists of two distinct components [41]: a non-ohmic primary current $\vec{J}^p(\vec{r})$ directly related to neural activation and an ohmic volume current

$\vec{J}^v(\vec{r})$ required by charge conservation. That is, the total current density

$$\vec{J} = \vec{J}^p + \vec{J}^v \tag{4}$$

is the sum of two components: a primary current $\vec{J}^p$ which is non-ohmic and generated by physiological processes within the body; and a volume current $\vec{J}^v$ which obeys Ohm's law

$$\vec{J}^v = \sigma \vec{E} \tag{5}$$

and is induced by the primary current sources.

We will see shortly that the volume currents can be computed by analytic, finite element, or boundary element methods. Suppose that the primary source distribution is represented as a weighted sum

$$\vec{J}^p(\vec{r}) = \sum_n q_n \, \vec{J}^p_n(\vec{r}) \quad . \tag{6}$$

where $\vec{J}^p_n(\vec{r})$ is a convenient set of basis functions for the primary current $\vec{J}^p(\vec{r})$. Let $\vec{J}^v_n(\vec{r})$ be the volume current induced by the corresponding primary current source $\vec{J}^p_n(\vec{r})$. Then the elementary sources $\vec{J}_n(\vec{r})$ may be defined as the primary sources *plus* the corresponding volume currents

$$\vec{J}_n(\vec{r}) = \vec{J}^p_n(\vec{r}) + \vec{J}^v_n(\vec{r}) \quad . \tag{7}$$

The total current due to a primary current distribution $\vec{J}^p(\vec{r})$ is then by superposition

$$\begin{aligned} \vec{J}(\vec{r}) &= \sum_n q_n \, \vec{J}_n(\vec{r}) \\ &= \sum_n q_n \, \vec{J}^p_n(\vec{r}) + \sum_n q_n \, \vec{J}^v_n(\vec{r}) \\ &= \vec{J}^p(\vec{r}) + \vec{J}^v(\vec{r}) \end{aligned} \tag{8}$$

and so the primary current alone must be

$$\vec{J}^p(\vec{r}) = \sum_n q_n \, \vec{J}^p_n(\vec{r}) \quad . \tag{9}$$

That is, given estimates of q_n by some inverse method, it is possible to reconstruct either the total or the primary current. This is useful because the primary current is more directly related to the physiology of interest.

2.4 Significance of the Volume Currents

Accounting for the volume currents adds to the computational cost, so this should be done only if they have a significant effect on the magnetic fields. In fact, previous workers have found that the volume currents in realistic head [22] and torso [95] geometries contribute significantly to the total magnetic field and should be included in the forward model for the best accuracy. Purcell, Stroink, and Horacek [76] have also investigated the effects of the torso, lung, and blood pool boundaries in MCG and ECG. They find that both the isopotential and isofield maps are affected by the inclusion of tissue boundaries compared to a homogeneous infinite half-space.

Figure 2.1 shows the normal and tangential components of the magnetic field produced by the volume currents alone on a plane 1 cm above the chest wall for a 1 μA-m dipole source, which is typical of cardiac current dipoles; the source is in the anterior wall of the heart and directed almost tangential to the chest wall. These fields were computed using a very detailed finite element model with 186 536 elements developed by Eisenberg et al. at Boston University. Each contour represents 2500 fT, so the maximum field perturbation is about 22 500 fT, which is much larger than the measurement noise, typically about 50 fT. That is, the field perturbation cannot be ignored compared to the measurement noise and so the volume currents must be included in the forward model.

Section 11.2.2 shows that the volume currents can significantly affect the measured magnetic fields in neuromagnetic imaging.

2.5 Forward Problem

The biomagnetic forward problem is to compute, given any $\vec{J}^p$, the corresponding total current $\vec{J}$, magnetic field $\vec{B}$, and flux ψ_m threading each detector coil.

The frequencies involved in biomagnetic imaging are low (1–1000 Hz) compared to the times required for a disturbance to propagate through the body and it is sufficiently accurate to treat it as a quasistatic problem. Under quasistatic conditions, charge conservation requires

$$\nabla \vec{J} = 0 \quad . \tag{10}$$

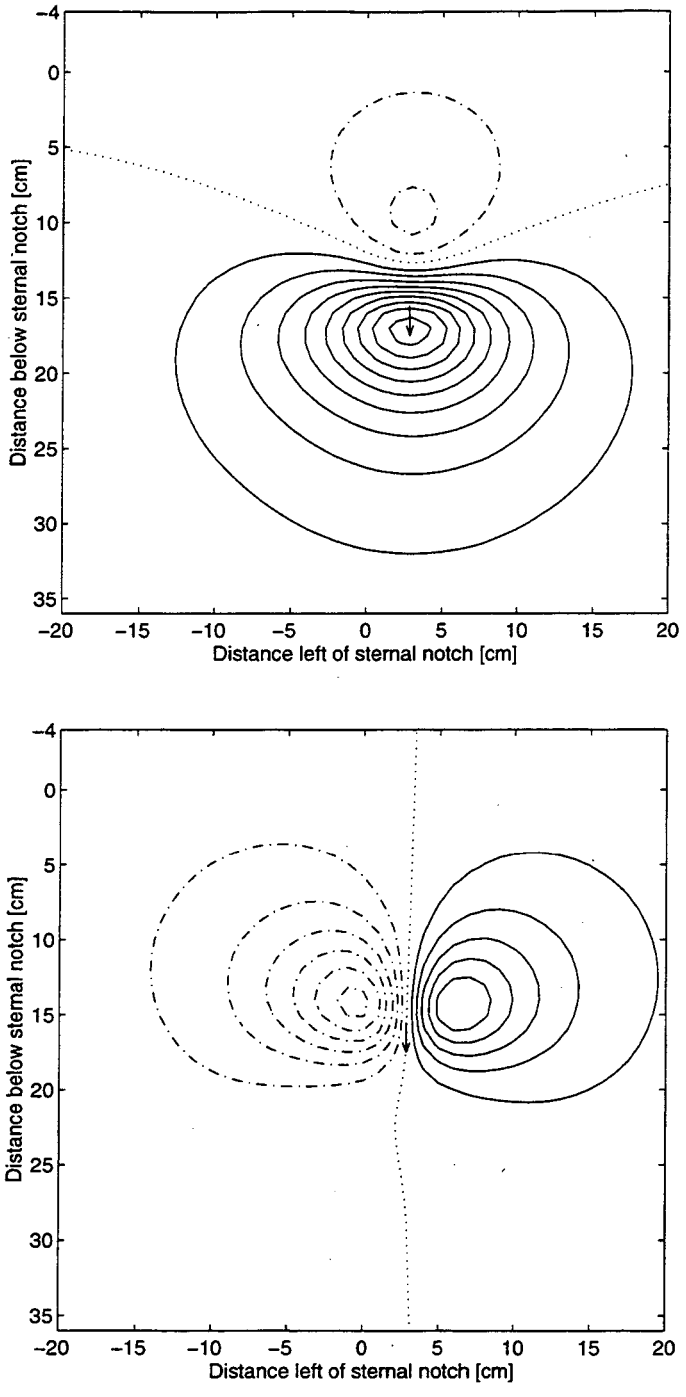


Figure 2.1 Cardiomagnetic field perturbations due to volume currents. These two plots show the magnetic fields on a plane 1 cm above the chest wall due to the volume currents alone for a typical cardiac current dipole of 1μA-m on the anterior wall of the heart, as shown by the arrow. The upper plot shows the lateral component of the field (across the chest); the lower plot shows the anterior component (normal to the chest). The dotted line indicates zero field; the solid line is positive; and the dot-dash line is negative. The contour interval is 2500 fT and a typical measurement noise level is 50 fT.

Introducing the scalar potential ϕ defined by

$$\vec{E} = -\nabla\phi \tag{11}$$

and substituting yields

$$\nabla\vec{J} = \nabla\vec{J}^p + \nabla\vec{J}^v = \nabla\vec{J}^p + \nabla(\sigma\vec{E}) = \nabla\vec{J}^p - \nabla(\sigma\nabla\phi) \tag{12}$$

and therefore

$$\nabla(\sigma\nabla\phi) = \nabla\vec{J}^p \quad , \tag{13}$$

which is Poisson's equation for the potential. Barr, Pilkington, et al. [5] give the appropriate boundary condition, which is that the normal component of $\nabla\phi$ vanishes at the surface of the body. Since this is a Neumann boundary condition, the solution ϕ is unique up to an additive constant. If it is required in addition that $\phi(\vec{r}_0) = 0$ at some fixed point $\vec{r}_0$, then the solution is unique and the mapping $\vec{J}^p \mapsto \phi$ is linear.

Once the potential ϕ is known, the total current is simply

$$\vec{J} = \vec{J}^p - \sigma\nabla\phi \quad . \tag{14}$$

The flux through a given coil is most easily computed through the vector potential

$$\vec{A}(\vec{r}) = \frac{\mu_0}{4\pi}\int_{R^3}\frac{\vec{J}(\vec{p})}{||\vec{r}-\vec{p}||}\,d^3\vec{p} \quad . \tag{15}$$

It is assumed that the permeability is constant and equal to the permeability of vacuum $\mu_0 = 4\pi \times 10^{-7}$ H/m. The magnetic flux density, if desired, is defined by

$$\vec{B} = \nabla \times \vec{A} \quad . \tag{16}$$

Suppose that the shape of the mth coil is defined by a closed curve C_m defining a surface segment S_m. Then the magnetic flux threading that coil is

$$\psi_m = \int_{S_m}\vec{B}\cdot d\vec{a} = \oint_{C_m}\vec{A}\cdot d\vec{r} \quad , \tag{17}$$

where Stoke's law has been used to obtain the last equality. Define ψ as the column vector $[\psi_m]$.

Now observe that each of the mappings $\mathbf{q} \mapsto \vec{J}^p \mapsto \phi \mapsto \vec{J} \mapsto \vec{A} \mapsto \psi$ is linear and therefore that the composite mapping $\mathbf{F} : \mathbf{q} \mapsto \psi$ must also be linear. Furthermore, $\mathbf{F}$ can be represented as a matrix with columns

$$\mathbf{f}_n = \psi|_{\mathbf{q}=\mathbf{e}_n} \quad , \tag{18}$$

where $\mathbf{e}_n$ is the vector with 1 in the nth position and 0 elsewhere. That is, the nth column of the forward transfer matrix $\mathbf{F}$ is computed by solving the forward problem with the nth source set to unit amplitude and all others set to zero.

Letting the mth measurement be $b_m = \psi_m + w_m$ with noise w_m, the measurement vector can be written in the canonical form

$$\mathbf{b} = \mathbf{F}\mathbf{q} + \mathbf{w} \quad . \tag{19}$$

Then, given a method to estimate the source amplitudes $\hat{q}_n$, the estimated primary current distribution is

$$\vec{J}^p(\vec{r}) = \sum_n \hat{q}_n \, \vec{J}^p_n(\vec{r}) \quad . \tag{20}$$

The total current distribution, if desired, is similarly

$$\vec{J}(\vec{r}) = \sum_n \hat{q}_n \, \vec{J}_n(\vec{r}) \quad , \tag{21}$$

where $\vec{J}_n(\vec{r})$ is the total current induced by the primary current $\vec{J}^p_n(\vec{r})$ and computed in the forward problem for $\mathbf{q} = \mathbf{e}_n$.

2.6 Forward Models without Volume Currents

It is possible to obtain a simple forward model by assuming that the volume currents have no appreciable effect on the field measurements and that the detectors sample the magnetic field at a point; such models are useful for testing inverse methods but are not usually accurate enough for real biomagnetic imaging. Then the current density can be represented as a weighted sum of N known elementary sources $\vec{J}_n(\vec{r})$ in the form

$$\vec{J}(\vec{r}) = \sum_{n=1}^{N} q_n \, \vec{J}_n(\vec{r}) \quad . \tag{22}$$

Using the Biot-Savart law [49, pages 151–153], the magnetic field due to $\vec{J}$ is

$$\begin{aligned}\vec{B}(\vec{r}) &= \frac{\mu_0}{4\pi}\int_{R^3}\frac{\vec{J}(\vec{p})\times(\vec{r}-\vec{p})}{\|\vec{r}-\vec{p}\|^3}\,d^3\vec{p} \\ &= \sum_n q_n\left[\frac{\mu_0}{4\pi}\int_{R^3}\frac{\vec{J}_n(\vec{p})\times(\vec{r}-\vec{p})}{\|\vec{r}-\vec{p}\|^3}\,d^3\vec{p}\right] \\ &= \sum_n q_n\vec{B}_n(\vec{r}) \end{aligned} \tag{23}$$

where

$$\vec{B}_n(\vec{r}) = \frac{\mu_0}{4\pi}\int_{R^3}\frac{\vec{J}_n(\vec{p})\times(\vec{r}-\vec{p})}{\|\vec{r}-\vec{p}\|^3}\,d^3\vec{p} \tag{24}$$

is the magnetic field at position $\vec{r}$ due to the nth source.

Now suppose that there are M detectors, the mth of which measures the component of the field in direction $\vec{s}_m$ at position $\vec{r}_m$. Furthermore, the measurement is contaminated by some noise w_m. Then that measurement b_m can be written as

$$\begin{aligned} b_m &= \vec{s}_m\cdot\vec{B}(\vec{r}_m) + w_m \\ &= \sum_n q_n\,\vec{s}_m\cdot\vec{B}_n(\vec{r}_m) + w_m \\ &= \sum_n q_n\left[\frac{\mu_0}{4\pi}\int_{R^3}\frac{\vec{s}_m\cdot\vec{J}_n(\vec{p})\times(\vec{r}_m-\vec{p})}{\|\vec{r}_m-\vec{p}\|^3}\,d^3\vec{p}\right] + w_m \\ &= \sum_n F_{mn}\,q_n + w_m \end{aligned} \tag{25}$$

where

$$F_{mn} = \frac{\mu_0}{4\pi}\int_{R^3}\frac{\vec{s}_m\cdot\vec{J}_n(\vec{p})\times(\vec{r}_m-\vec{p})}{\|\vec{r}_m-\vec{p}\|^3}\,d^3\vec{p} \tag{26}$$

is the response of the mth detector to the nth source. The forward transfer coefficient F_{mn} can also be regarded as the lead field [106] of the mth detector integrated over the current distribution of the nth source.

Rewriting these equations for b_m in matrix form yields

$$\mathbf{b} = \mathbf{F}\mathbf{q} + \mathbf{w} \quad . \tag{27}$$

2.7 Dipole Sources

In the particular case that each source $\vec{J}_n$ is a current dipole with moment $\vec{q}_n$ at position $\vec{p}_n$ (and still ignoring the volume currents), the field is

$$\vec{B}_n(\vec{r}) = \frac{\mu_0}{4\pi} \cdot \frac{\vec{q}_n \times (\vec{r} - \vec{p}_n)}{\|\vec{r} - \vec{p}_n\|^3} \tag{28}$$

and the forward transfer matrix $\mathbf{F}$ has entries

$$F_{mn} = \frac{\mu_0}{4\pi} \cdot \frac{\vec{s}_m \cdot \vec{q}_n \times (\vec{r}_m - \vec{p}_n)}{\|\vec{r}_m - \vec{p}_n\|^3} \quad . \tag{29}$$

A "rotating" dipole of unknown orientation at a given position may be represented as two or three orthogonal fixed dipoles. The fixed dipoles may be oriented along the coordinate axes; normal and tangential to the cerebral cortex; or along the principal axes of the assumed probability density for the rotating dipole. In the last case, the fixed dipoles will be uncorrelated.

In magnetoencephalography, a dipole source will ordinarily be oriented normal to the cerebral cortex and the direction of $\vec{J}_n$ may be chosen to match. If the surface is strongly curved, a normal dipole with large expected amplitude and two tangential dipoles with smaller expected amplitude may be used.

2.8 Forward Models with Volume Currents

The previous models are deficient in two major respects. First, in realistic torso geometries, the volume currents can significantly affect the measured fields. Second, practical magnetometer configurations do not sample the field at a point but rather measure the flux threading some coil of finite area.

As was shown above, the $\mathbf{F}$ matrix can be computed for the general case in terms of its columns

$$\mathbf{f}_n = \psi|_{\mathbf{q}=\mathbf{e}_n} \quad , \tag{30}$$

where $\mathbf{e}_n$ is the vector with 1 in the nth position and 0 elsewhere. That is, the nth column of the forward transfer matrix $\mathbf{F}$ is computed by solving the forward problem with the nth source set to unit amplitude and all others set to zero.

One path for that computation follows the sequence $\mathbf{e}_n \mapsto \vec{J}_n^p \mapsto \phi_n \mapsto \vec{J}_n \mapsto \vec{A}_n \mapsto \psi_{mn}$. The computation of the total current $\vec{J}_n$ and vector potential $\vec{A}_n$ from the primary current $\vec{J}_n^p$ may be done by finite element or boundary element methods and will be discussed at length in the remainder of this chapter.

The computation of flux ψ_{mn} from vector potential $\vec{A}_n$ may be done simply by discretizing the defining equation

$$\begin{aligned}\psi_m &= \oint_{C_m} \vec{A} \cdot d\vec{r} = \oint_{C_m} \left[\frac{\mu_0}{4\pi} \int_{R^3} \frac{\vec{J}(\vec{p})}{\|\vec{r}-\vec{p}\|} d^3\vec{p} \right] \cdot d\vec{r} \\ &= \frac{\mu_0}{4\pi} \int_{R^3} \vec{J}(\vec{p}) \cdot \left[\oint_{C_m} \frac{d\vec{r}}{\|\vec{r}-\vec{p}\|} \right] d^3\vec{p} \\ &= \frac{\mu_0}{4\pi} \int_{R^3} \vec{J}(\vec{p}) \cdot \vec{L}(\vec{r})\, d^3\vec{p} \quad , \end{aligned} \tag{31}$$

where the curve C_m defines the shape of the coil and $\vec{L}(\vec{r})$ is the vector lead field corresponding to that coil. If the coil is circular, or a sum of circles, the lead field reduces to an elliptic integral, or a sum of elliptic integrals. The problem of *efficiently* computing the flux from the total current will not be considered in this dissertation.

2.9 Analytic Methods

This section and the two following it discuss methods for computing the volume currents in various geometries from simplistic to realistic. This section reviews analytic results for such simple geometries as an infinite halfspace, sphere and cylinder.

Wilson and Bayley [108] give exact closed form solutions for the potential induced by current sources in a sphere, as does Frank [33]. Okada [66] gives a solution for a finite-length circular cylinder.

Sarvas [86] gives the magnetic fields due to a current dipole in a sphere or an infinite half-space. Cuffin and Cohen [23] give analytic solutions for the magnetic field due to a current dipole in an infinite half-space, sphere, prolate spheroid, or oblate spheroid. They show the effects of the dipole and volume currents separately and note that the normal component of the magnetic field is usually due entirely or mostly to the dipole source rather than the volume currents.

2.10 Finite Element Methods

Finite element methods solve Poisson's equation (13) for the potential by dividing the volume of interest into many polyhedral regions and assuming that the potential is constant within each region or varies according to a simple polynomial function. Requiring that the potential is continuous across the boundaries of each region yields a system of linear equations which can be solved by numerical methods. Given the potential, the total current and magnetic field can be determined directly.

Jin [46] is a monograph on the use of finite element methods for electromagnetics in various applications but does not specifically discuss biological applications.

Miller [60] is a review of finite element methods for bioelectric phenomena, including electrocardiology. Surprisingly, there is no discussion of electroencephalography.

Thevenet et al. [97] discuss the use of finite element modelling to obtain the potential distribution due to a current dipole in the head. They note that, unlike the boundary element method, it can model aniotropic conductors.

Yan, Nunez, and Hart [109] describe a three-dimensional finite element model of the head with a new formulation of dipole current sources. They test the finite element model by comparing it to analytic results for the three-spheres model of the head. Results for a more realistic head model are also presented.

Ragan [77, 78] uses a finite element model of the canine torso to estimate the electric fields and current densities produced by rapidly changing magnetic fields.

Finite element methods have the advantage that they can model anisotropic conductors such as muscle; but it is often difficult to model point or dipole sources. Finite element methods also are often more expensive to compute than are boundary element methods.

2.11 Boundary Element Methods

Boundary element methods are applicable when the conductivity is isotropic and piecewise constant. Then Poisson's equation can be transformed to an integral equation over the boundaries separating the regions of different conductivity. These

boundaries are divided into triangular patches of constant or linearly varying potential and requiring continuity again yields a system of linear equations.

Barr, Pilkington, et al. [5] have presented a boundary element method for computing the potentials on the surface of a piecewise homogeneous volume due to current sources within the volume. The surfaces bounding homogenous volumes are defined by triangular tesselations. They develop a set of linear equations relating the potentials of the actual geometry to the potentials at the same positions for an infinite homogeneous medium; these equations are singular but can be solved iteratively. The coefficients in the linear equations are related to the solid angle subtended by each triangle as seen from each other triangle; van Oosterom [102] gives an efficient algorithm for computing the necessary solid angles given the triangle vertices.

Horacek [43] modifies the Barr-Pilkington method to compute the magnetic fields produced by source and volume currents by introducing fictitious current layers at the boundaries that modify the currents induced in an infinite homogenous medium. Some of his results are identical to those of Geselowitz [36, 35], who did not explicitly show the fictitious layer currents. He handles the singularity of the linear equations by deflating them using a method due to Lynn and Timlake [51].

Purcell and Stroink [75] show that this linear system need not be solved iteratively. They perform a LU decomposition of the matrix, compute its inverse, and solve the linear system by multiplying by the inverse. This provides a significant speedup in moving dipole solutions of the inverse problem, where it is necessary to repeatedly solve the forward problem. In fact, it is not necessary to compute the matrix inverse explicitly; the LU decomposition can be used directly to solve the linear system at no increase in computational cost. In a later paper, Nenonen, Purcell, et al. [63] compute the matrix inverse using a singular value decomposition.

Ferguson, Zhang, and Stroink [31] generalize the boundary element method to provide for a linear variation of the potential across each triangle.

Meijs et al. [59] recommend several methods for improving the accuracy of the boundary element method in electro- and magnetoencephalography, where the strong conductivity contrast between the skull and soft tissues degrades the nu-

merical accurary. He does not address magnetocardiography and it is not known whether the same numerical problems occur there.

Hämäläinen and Sarvas [40] discuss the problems introduced by the low conductivity of the skull and suggest methods for solving those problems. They also conclude that for magnetoencephalography it suffices to consider only the compartments inside the skull; the volume currents outside the skull appear to contribute negligibly to the magnetic fields.

Budiman and Buchanan [13] introduce a variant of the boundary element method, which they call "weighted vertices."

Chapter 3
Prior Knowledge

The biomagnetic inverse problem is almost always ill-posed in that there are infinitely many current distributions that will yield the observed magnetic field values. The measured data alone provide no basis for distinguishing among all these possible solutions. But we usually know more than just the observed field values since we can draw on a general knowledge of biophysics and physiology as well as previous in vitro and invasive experiments.

The primary purpose of this dissertation is to develop and test algorithms for incorporating such anatomical and physiological prior knowledge into biomagnetic source imaging problems. This offers two advantages: First, the constraints imposed by the prior knowledge will restrict the set of possible solutions, possibly defining a unique solution consistent with both the measured data and the priors. Second, enforcing these constraints as part of the reconstruction guarantees that the solution will in fact satisfy these constraints. (Of course, it also means that the satisfaction of the constraints is no longer evidence for the plausibility of the reconstruction.)

3.1 Kinds of Prior Knowledge

For example, suppose that an experiment is intended to determine the regions of the brain that are activated in reading and the time history of activation at each of these locations. (Similar remarks will apply to other experimental and clinical applications.) Then it is reasonable to assume that:

1. The primary current sources are located within the cortex of the brain.

2. In fact, the primary sources for this cognitive task are probably located within the specific regions of the cortex previously implicated as involved in reading.

3. The primary sources are correlated over distances of several millimeters and over time intervals of tens to hundreds of milliseconds.

4. The amplitudes of the primary currents are consistent with the amplitudes estimated from a detailed cellular physiological model and with amplitudes observed in previous invasive experiments.

5. The volume currents satisfy Ohm's law for some determinable conductivity distribution within the head.

6. The total current satisfies charge conservation and is zero outside the body.

7. The magnetic fields at the detectors are determined by the the total current distribution via the Biot-Savart law.

8. The measurement noise due to the intrinsic noise of the detectors and due to external magnetic sources is similar to the noise observed in previous calibration experiments without a subject present.

9. The measurement noise due to current sources in the subject outside the brain (muscle twitches, eye motions, heart beat, and so on) is consistent from subject to subject if the experimental conditions are kept constant.

3.2 Reliability of Priors

Note that these various kinds of prior information differ significantly in reliability. There would seem to be little risk that the Biot-Savart law or Ohm's law will be violated. The assumptions about the probable source and noise amplitudes involve rather more risk; the generalization from previous experiments to the current one may be plausible but it is also possible that some new phenomena may be waiting to appear in this new experiment. The assumption that the brain will be active in only those areas previously implicated in reading is definitely a risky one, especially since one goal of the current experiment is to locate the active regions in this task; refusing to look outside the expected regions of activation may well blind a reconstruction algorithm to evidence for other active regions. Chapter 8 will show that this risk is real and discuss a method for detecting such "misinformative" priors.

3.3 Sources and Forms of Priors

Many of the priors listed above may conveniently be treated as non-statistical or hard constraints that must be satisified for a possible solution; others may conveniently be treated as soft or statistical constraints that establish preferences between solutions but do not absolutely rule out any solutions. Furthermore, some of the priors listed above require some additional information about the specific subject or a suitable population of similar subjects.

This dissertation does not attempt to solve the problem of estimating apprpriate priors for any particular experiment. But the following remarks should make it plausible that it is possible to obtain suitable priors and to represent them in a form suitable for inclusion in an inverse problem.

Ohm's law, the Biot-Savart law, charge conservation, and the other laws of electromagnetics are incorporated into the forward model.

The anatomy of the head may be determined for the individual subject by magnetic resonance imaging, giving the spatial locations of the skull, scalp, cerebrospinal fluid, grey matter, and white matter. The electrical conductivities of these various tissues are known and may be combined with the spatial coordinates to yield the conductivity distribution within the head. The conductivity distribution may then be used in the forward model to account for the volume currents.

The possible (as opposed to probable) locations of the primary current sources may be determined from the known location of the cerebral cortex.

The probable amplitudes of the cortical sources and their spatial and temporal correlations may be represented as a probability distribution. The mean values, variances, covariances, and other parameters of this distribution may be estimated from several lines of investigation including detailed bioelectric models of the brain, invasive measurements made in animals (and when possible, in humans) and analysis of MEG measurements on a population of similar subjects.

The probable locations of the cortical regions activated by the reading task may be determined by locating previously implicated anatomical regions in the MR image. The relative probability of different possible source regions may be

represented as differences in the expected source variance (which is proportional to average source power). If desired, the most likely orientation of these sources may be set to be normal to the cortical sheet. Alternatively, the regions activated by the reading task could be indentified by functional MR imaging of blood flow differences in the brain between task and non-task conditions, or by PET imaging of metabolic differences.

The statistics of the measurement noise due to intrinstic detector noise and magnetic sources external to the subject can be estimated by tests of the detector system without a subject. The statistics of measurement noise due to sources in the subject but outside the brain may be estimated by testing a suitable population of subjects under non-task conditions.

Summarizing, the prior knowledge used to constrain the reconstruction may be expressed as (1) the laws of physics incorporated into the forward problem; (2) a source model defining a set of elementary current sources which are consistent with the known anatomy and physiology; (3) a probability distribution for the source amplitudes; and (4) a probability distribution for the detector noise. The forward model and source model have already been discussed; the remainder of this chapter will consider the representation of statistical prior information concerning the source and noise amplitudes.

3.4 Source and Noise Statistics

As already discussed, the forward problem may be represented in the general form

$$\mathbf{b} = \mathbf{F}\mathbf{q} + \mathbf{w} \tag{32}$$

for measurements $\mathbf{b}$, forward transfer matrix $\mathbf{F}$, source amplitudes $\mathbf{q}$, and measurement noise $\mathbf{w}$.

The vector $\mathbf{q}$ of source amplitudes q_n is assumed to be a random vector with mean zero and covariance matrix $\mathbf{A} = \mathsf{E}\,\mathbf{q}\mathbf{q}^T$ with entries $\alpha_{ij}^2 = \mathsf{E}\,q_i q_j$. If the expected amplitudes are not zero-mean, the shifted vector $\tilde{\mathbf{q}} = \mathbf{q} - \mathsf{E}\,\mathbf{q}$ is zero-mean and can be used instead. Since $\mathbf{A}$ is a covariance matrix, it is symmetric and positive semidefinite. The diagonal entries α_{nn}^2 are non-negative and α_{nn} is

Signal Source	Signal Strength (femtoTesla)	Frequency Range (Hz)
Cardiogram	50 000	0.1–100
Encephalogram (δ)	1 500	0.3–50
Auditory Evoked Response	500	0.2–40
Visual Evoked Response	200	0.1–50
Evoked Cortical Activity	60	0.2–50

Figure 3.1 Some typical values for biomagnetic signals. These data are taken from Fagaly [28, Figure 1]. For comparison, the earth's magnetic field is about 50 microTesla.

the root-mean-square (RMS) amplitude of the nth source. The off-diagonal entries $\alpha_{ij}^2, i \neq j$ are the correlations between sources and may be positive, negative, or zero. Electrical activity in the brain and heart is spatially coherent and sources separated by less than the coherence length will be correlated; these correlations correspond to non-zero off-diagonal elements in $\mathbf{A}$.

The noise vector $\mathbf{w}$ is assumed to be a random vector with mean zero and covariance matrix $\mathbf{\Sigma} = \mathsf{E}\ \mathbf{w}\mathbf{w}^T$ with entries $\sigma_{ij}^2 = \mathsf{E}\ w_i w_j$. Since $\mathbf{\Sigma}$ is a covariance matrix, it is symmetric and positive semidefinite. Each diagonal entry σ_{mm}^2 is the expected noise power of the mth detector and is non-negative; σ_{mm} is the RMS noise amplitude of the mth detector. The off-diagonal entries $\sigma_{ij}^2, i \neq j$ are the noise covariances and may be positive, negative, or zero. If the noise is uncorrelated between sensors, then $\mathbf{\Sigma}$ is diagonal.

The cross-covariance between the source and noise amplitude vectors is $\mathbf{\Gamma} = \mathsf{E}\ \mathbf{q}\mathbf{w}^T$ with entries $\gamma_{nm} = \mathsf{E}\ q_n w_m$. In many applications, there will be no source-noise correlation and $\mathbf{\Gamma}$ will be zero.

Fagaly [28, Figure 1] gives some typical values for signal strength as seen at the magnetometer; some of these are reproduced in Figure 3.1. In terms of the notation used in this dissertation, the listed values are typical values for the measured signals b_m. The source amplitudes can be estimated from these values but are less well constrained.

Fagaly also [pages 26–28] discusses the measurement noise. The SQUID mag-

Typical RMS noise (fT)		
	Noise spectral power ($fT/\sqrt{Hz}$)	
Bandwidth	100	10
100	1000	100
50	707	71
10	316	32
5	224	22
1	100	10

Figure 3.2 Some typical RMS noise values for biomagnetic signals. These data are computed from the data of Figure 3.1.

netometer itself has a spectral noise power of about 10 $fT/\sqrt{Hz}$ but environmental noise is often much larger. An eddy current shielded room can reduce the enviromental noise to less than 200 $fT/\sqrt{Hz}$; a magnetically shielded room to less than 3 $fT/\sqrt{Hz}$, which is less than the intrinsic noise of the detector itself. Figure 3.2 shows the RMS measurement noise for various conbinations of bandwidth and noise spectral power typical of MEG.

3.5 Measurement Statistics

Given these definitions, the measurement covariance is

$$\mathbf{B} = \mathsf{E}\,\mathbf{b}\mathbf{b}^T = \mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{F}\mathbf{\Gamma} + \mathbf{\Gamma}^T\mathbf{F}^T + \mathbf{\Sigma} \quad . \tag{33}$$

In the case that $\mathbf{\Gamma} = 0$, this simplifies to $\mathbf{B} = \mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma}$. The expected signal power at the mth detector is (where $\mathbf{e}_m$ is the vector with one in the mth position and zeros elsewhere)

$$\begin{aligned} \beta_m^2 &= \mathsf{E}(b_m - w_m)^2 = \mathsf{E}(\mathbf{e}_m^T\mathbf{F}\mathbf{q})^2 \\ &= \mathbf{e}_m^T\mathbf{F}\mathbf{A}\mathbf{F}^T\mathbf{e}_m \quad , \end{aligned} \tag{34}$$

which is just the mth diagonal element of $\mathbf{F}\mathbf{A}\mathbf{F}^T$. The total signal power over all detectors is $\sum_m \beta_m^2 = \mathrm{Tr}(\mathbf{F}\mathbf{A}\mathbf{F}^T)$. The total noise power is $\sum_m \sigma_{mm}^2 = \mathrm{Tr}(\mathbf{\Sigma})$ and

Signal-to-noise ratio (dB)		
	Noise spectral power ($fT/\sqrt{Hz}$)	
Signal Source	100	10
Cardiogram	34.0	54.0
Encephalogram (δ)	6.5	26.5
Auditory Evoked Response	–2.0	18.0
Visual Evoked Response	–11.0	9.0
Evoked Cortical Activity	–21.4	–1.4

Figure 3.3 Some typical SNR values for biomagnetic signals. These data are computed from the data of Figure 3.1, assuming that the full bandwidth is used and that no signal averaging is used.

so it is reasonable to define the signal-to-noise (power) ratio as

$$\mathrm{SNR} = \mathrm{Tr}(\mathbf{FAF}^T)/\,\mathrm{Tr}(\mathbf{\Sigma}) \quad . \tag{35}$$

Dividing the numerator and denominator by the number of detectors M yields the equivalent definition

$$\mathrm{SNR} = \frac{(1/M)\sum_m \beta_m^2}{(1/M)\sum_m \sigma_{mm}^2} = \frac{\beta_{\mathrm{avg}}^2}{\sigma_{\mathrm{avg}}^2} \tag{36}$$

where β_{avg}^2 and σ_{avg}^2 are the average expected signal and noise power per detector. Now suppose that the typical spectral noise power is n in $fT/\sqrt{Hz}$ and the measurement bandwidth is B in Hz. Then the rms noise is $\sigma^2 = Bn^2$ and the SNR can be written in the form

$$\mathrm{SNR} = \frac{\beta_{\mathrm{avg}}^2}{Bn^2} \quad . \tag{37}$$

The signal-to-noise ratio expressed in decibels (dB) is $10\,\log_{10}$ SNR.

The SNR for a particular system and experiment will depend on the signal sought (cardiogram, encephalogram, or evoked potential), instrument configuration (magnetometer or gradiometer), ambient magnetic noise, shielding, and signal averaging. Most practical experiments should fall in the range 0–40 dB.

Figure 3.3 shows the calculated SNR for the signal types described in Figure 3.1, for some typical values of noise. These values can be improved by signal averaging or narrowing the bandwidth.

3.6 Assumption A

We pause here to introduce a technical assumption—to be called assumption A—that will be useful later. Precisely stated, we will assume hereafter that that none of the source amplitudes q_n or the noise amplitudes w_m is almost surely equal to a linear combination of the remaining source and noise amplitudes. (Two random variables are "almost surely" equal if they are equal with probability one.) That is, every source and noise amplitude has some non-zero residual variance even after the effect of every other source and noise amplitude has been accounted for.

Assumption A is unlikely to be an issue in practice. About the only way to violate it is to set some a priori source variance to zero or to use some field measurement twice; the problem is easily fixed by omitting the source with zero variance (since it is known a priori), choosing a nonzero variance, or omitting the redundant measurement.

The value of assumption A is mathematical; it authorizes some algebraic manipulations that would otherwise be questionable. Specifically, we will use the following consequences:

1. The covariance matrices $\mathbf{A} = \mathsf{E}\,\mathbf{q}\mathbf{q}^T$, $\mathbf{\Sigma} = \mathsf{E}\,\mathbf{w}\mathbf{w}^T$, and $\mathbf{B} = \mathsf{E}\,\mathbf{b}\mathbf{b}^T$ are all invertible, symmetric positive definite, and have Cholesky decompositions [38, pages 141–142] in the forms

$$\mathbf{A} = \mathbf{R}\mathbf{R}^T, \qquad \mathbf{\Sigma} = \mathbf{S}\mathbf{S}^T, \quad \text{and} \qquad \mathbf{B} = \mathbf{L}\mathbf{L}^T \tag{38}$$

where $\mathbf{R}$, $\mathbf{S}$, and $\mathbf{L}$ are lower triangular and invertible. Furthermore, expressions of the form $\mathbf{R}^{-1}\mathbf{X}$, $\mathbf{S}^{-1}\mathbf{X}$, or $\mathbf{L}^{-1}\mathbf{X}$ can be efficiently computed by forward substitution [38, pages 86–90] without explicitly computing the matrix inverse.

2. The form $\mathsf{E}\,\|\mathbf{D}\mathbf{b}\|^2$ is strictly greater than zero for every non-zero matrix $\mathbf{D}$.

The remainder of this section is devoted to proving these consequences.

Definition: A set of random vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ is linearly dependent if there exist constant vectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$, not all zero, such that $\mathbf{u}_1^T \mathbf{x}_1 + \cdots + \mathbf{u}_n^T \mathbf{x}_n = 0$ almost surely. If the set $\mathbf{x}_1, \ldots, \mathbf{x}_n$ is not linearly dependent, it is linearly independent.

Proposition: Assumption A implies that the set $\{\mathbf{q}, \mathbf{w}\}$ is linearly independent.

Proof: Suppose that $\mathbf{q}$ and $\mathbf{w}$ are linearly dependent. Then there exist vectors $\mathbf{u}$ and $\mathbf{v}$, not both zero, such that $\mathbf{u}^T\mathbf{q} + \mathbf{v}^T\mathbf{w} = \sum_i u_i q_i + \sum_j v_j w_j = 0$ almost surely. But at least one component of $\mathbf{u}$ or $\mathbf{v}$ is non-zero. Say that component is u_i. Then solving for q_i yields an linear expression in the remaining components that is almost surely equal to q_i. Thus assumption A must be false. ■

Proposition: Suppose that a set of random vectors is linearly dependent. Then adding any additional random vector to the set yields a linearly dependent set.

Proof: Set the coefficient of the added vector to zero. ■

Corollary: Any subset of a linearly independent set is linearly independent.

Corollary: The vector $\mathbf{q}$ is linearly independent. So is $\mathbf{w}$.

Proposition: The measurement vector $\mathbf{b} = \mathbf{F}\mathbf{q} + \mathbf{w}$ is linearly independent.

Proof: Suppose that $\mathbf{b}$ is linearly dependent. Then there exists a non-zero constant $\mathbf{u}$ such that $\mathbf{u}^T\mathbf{b} = 0$ almost surely. But this implies that $\mathbf{u}^T\mathbf{F}\mathbf{q} + \mathbf{u}^T\mathbf{w} = 0$ almost surely. Thus $\mathbf{q}$ and $\mathbf{w}$ must be linearly dependent. ■

Proposition: Suppose that the random vector $\mathbf{x}$ is linearly independent. Then the correlation matrix $\mathbf{C}_\mathbf{x} = \mathsf{E}\,\mathbf{x}\mathbf{x}^T$ is nonsingular.

Proof: Suppose that $\mathbf{C}_\mathbf{x}$ is singular. Then there exists a constant $\mathbf{u} \neq 0$ such that $\mathbf{u}^T\mathbf{C}_\mathbf{x}\mathbf{u} = 0$. But $\mathbf{u}^T\mathbf{C}_\mathbf{x}\mathbf{u} = \mathbf{u}^T\,\mathsf{E}(\mathbf{x}\mathbf{x}^T)\mathbf{u} = \mathsf{E}(\mathbf{u}^T\mathbf{x}\mathbf{x}^T\mathbf{u}) = \mathsf{E}\,\|\mathbf{u}^T\mathbf{x}\|^2$ which can be zero only if $\mathbf{u}^T\mathbf{x} = 0$ almost surely. But this implies that $\mathbf{x}$ is linearly dependent. ■

Corollary: The covariance matrices $\mathbf{A} = \mathsf{E}\,\mathbf{q}\mathbf{q}^T$, $\mathbf{\Sigma} = \mathsf{E}\,\mathbf{w}\mathbf{w}^T$, and $\mathbf{B} = \mathsf{E}\,\mathbf{b}\mathbf{b}^T$ are all invertible.

Corollary: Since all covariance matrices are symmetric and positive semidefinite, $\mathbf{A}$, $\mathbf{\Sigma}$, and $\mathbf{B}$ are symmetric positive definite and have Cholesky decompositions in the forms

$$\mathbf{A} = \mathbf{R}\mathbf{R}^T, \qquad \mathbf{\Sigma} = \mathbf{S}\mathbf{S}^T, \quad \text{and} \qquad \mathbf{B} = \mathbf{L}\mathbf{L}^T \tag{39}$$

where $\mathbf{R}$, $\mathbf{S}$, and $\mathbf{L}$ are lower triangular and invertible.

Proposition: If $\mathbf{x}$ is linearly independent, then $\mathsf{E}\,||\mathbf{D}\mathbf{x}||^2 > 0$ for every matrix $\mathbf{D} \neq 0$.

Proof: Suppose that $\mathsf{E}\,||\mathbf{D}\mathbf{x}||^2 = 0$. Then, almost surely, $||\mathbf{D}\mathbf{x}|| = 0$, which implies that $\mathbf{D}\mathbf{x} = 0$, which implies that $\mathbf{x}$ is linearly dependent. ■

3.7 Gaussian Priors

Many of the results that will be presented hereafter are independent of the actual distributions of $\mathbf{q}$ and $\mathbf{w}$, provided that the covariances and cross-covariance are known. Some results, however, depend on the particular form of the distribution. For these results, it is convenient to assume that $\mathbf{q}$ and $\mathbf{w}$ are jointly Gaussian. More explicitly, we assume that the composite vector

$$\mathbf{v} = \begin{bmatrix} \mathbf{q} \\ \mathbf{w} \end{bmatrix} \tag{40}$$

is jointly Gaussian with mean 0 and covariance matrix

$$\mathbf{C}_\mathbf{v} = \begin{bmatrix} \mathbf{A} & \mathbf{\Gamma} \\ \mathbf{\Gamma}^T & \mathbf{\Sigma} \end{bmatrix} \quad . \tag{41}$$

Given this additional assumption of normality, it is possible and useful to compute the conditional distributions of $\mathbf{q}$ and $\mathbf{w}$ given any particular $\mathbf{b}$. Suppose that $\mathbf{x}$ and $\mathbf{y}$ are zero-mean, jointly Gaussian random vectors with covariance matrices $\mathbf{C}_{\mathbf{xx}}$ and $\mathbf{C}_{\mathbf{yy}}$ and cross-covariance $\mathbf{C}_{\mathbf{xy}} = \mathbf{C}_{\mathbf{yx}}^T$. Then [62, page 88, property 3.3], for given $\mathbf{y}$, $\mathbf{x}$ is normally distributed with mean

$$\mu_{\mathbf{x}|\mathbf{y}} = \mathbf{C}_{\mathbf{xy}}\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{y} \tag{42}$$

and variance

$$\mathbf{C}_{\mathbf{x}|\mathbf{y}} = \mathbf{C}_{\mathbf{xx}} - \mathbf{C}_{\mathbf{xy}}\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{C}_{\mathbf{yx}} \quad . \tag{43}$$

In this case, it is convenient to compute the mean and covariance of $\mathbf{v}$ given $\mathbf{b}$; then it is trivial to compute the means, covariances, and cross-covariance of $\mathbf{q}$ and $\mathbf{w}$. Thus,

$$\mathbf{C_{bb}} = \mathsf{E}\,\mathbf{b}\mathbf{b}^T = \mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{F}\boldsymbol{\Gamma} + \boldsymbol{\Gamma}^T\mathbf{F}^T + \boldsymbol{\Sigma} \tag{44}$$

and

$$\mathbf{C_{vb}} = \begin{bmatrix} \mathbf{C_{qb}} \\ \mathbf{C_{wb}} \end{bmatrix} = \begin{bmatrix} \mathsf{E}\,\mathbf{q}\mathbf{b}^T \\ \mathsf{E}\,\mathbf{w}\mathbf{b}^T \end{bmatrix} = \begin{bmatrix} \mathbf{A}\mathbf{F}^T + \boldsymbol{\Gamma} \\ \boldsymbol{\Gamma}^T\mathbf{F}^T + \boldsymbol{\Sigma} \end{bmatrix} \quad . \tag{45}$$

Furthermore,

$$\mathsf{E}(\mathbf{v} \mid \mathbf{b}) = \mathbf{C_{vb}}\mathbf{C_{bb}^{-1}}\mathbf{b} = \begin{bmatrix} \mathbf{C_{qb}}\mathbf{C_{bb}^{-1}}\mathbf{b} \\ \mathbf{C_{wb}}\mathbf{C_{bb}^{-1}}\mathbf{b} \end{bmatrix} \tag{46}$$

and

$$\begin{aligned} \mathrm{Cov}(\mathbf{v} \mid \mathbf{b}) &= \mathbf{C_{vv}} - \mathbf{C_{vb}}\mathbf{C_{bb}^{-1}}\mathbf{C_{vb}^T} \\ &= \begin{bmatrix} \mathbf{A} - \mathbf{C_{qb}}\mathbf{C_{bb}^{-1}}\mathbf{C_{qb}^T} & \boldsymbol{\Gamma} - \mathbf{C_{qb}}\mathbf{C_{bb}^{-1}}\mathbf{C_{wb}^T} \\ \boldsymbol{\Gamma}^T - \mathbf{C_{wb}}\mathbf{C_{bb}^{-1}}\mathbf{C_{qb}^T} & \boldsymbol{\Sigma} - \mathbf{C_{wb}}\mathbf{C_{bb}^{-1}}\mathbf{C_{wb}^T} \end{bmatrix} \quad . \end{aligned} \tag{47}$$

Thus, for given $\mathbf{b}$, the vectors $\mathbf{q}$ and $\mathbf{w}$ are jointly Gaussian with means

$$\begin{aligned} \mathsf{E}(\mathbf{q} \mid \mathbf{b}) &= \mathbf{C_{qb}}\mathbf{C_{bb}^{-1}}\mathbf{b} \\ &= (\mathbf{A}\mathbf{F}^T + \boldsymbol{\Gamma})(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{F}\boldsymbol{\Gamma} + \boldsymbol{\Gamma}^T\mathbf{F}^T + \boldsymbol{\Sigma})^{-1}\mathbf{b} \quad , \end{aligned} \tag{48}$$

$$\begin{aligned} \mathsf{E}(\mathbf{w} \mid \mathbf{b}) &= \mathbf{C_{wb}}\mathbf{C_{bb}^{-1}}\mathbf{b} \\ &= (\boldsymbol{\Gamma}^T\mathbf{F}^T + \boldsymbol{\Sigma})(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{F}\boldsymbol{\Gamma} + \boldsymbol{\Gamma}^T\mathbf{F}^T + \boldsymbol{\Sigma})^{-1}\mathbf{b} \quad , \end{aligned} \tag{49}$$

covariances

$$\begin{aligned} \mathrm{Var}(\mathbf{q} \mid \mathbf{b}) &= \mathbf{A} - \mathbf{C_{qb}}\mathbf{C_{bb}^{-1}}\mathbf{C_{qb}^T} \\ &= \mathbf{A} - (\mathbf{A}\mathbf{F}^T + \boldsymbol{\Gamma})(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{F}\boldsymbol{\Gamma} + \boldsymbol{\Gamma}^T\mathbf{F}^T + \boldsymbol{\Sigma})^{-1}(\mathbf{F}\mathbf{A} + \boldsymbol{\Gamma}^T) \quad , \end{aligned} \tag{50}$$

$$\begin{aligned} \mathrm{Var}(\mathbf{w} \mid \mathbf{b}) &= \boldsymbol{\Sigma} - \mathbf{C_{wb}}\mathbf{C_{bb}^{-1}}\,\mathbf{C_{wb}^T} \\ &= \boldsymbol{\Sigma} - (\boldsymbol{\Gamma}^T\mathbf{F}^T + \boldsymbol{\Sigma})(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{F}\boldsymbol{\Gamma} + \boldsymbol{\Gamma}^T\mathbf{F}^T + \boldsymbol{\Sigma})^{-1}(\mathbf{F}\boldsymbol{\Gamma} + \boldsymbol{\Sigma}) \quad , \end{aligned} \tag{51}$$

and cross-covariance

$$\begin{aligned} \mathrm{Cov}(\mathbf{q}, \mathbf{w} \mid \mathbf{b}) &= \boldsymbol{\Gamma} - \mathbf{C_{qb}}\mathbf{C_{bb}^{-1}}\mathbf{C_{wb}^T} \\ &= \boldsymbol{\Gamma} - (\mathbf{A}\mathbf{F}^T + \boldsymbol{\Gamma})(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{F}\boldsymbol{\Gamma} + \boldsymbol{\Gamma}^T\mathbf{F}^T + \boldsymbol{\Sigma})^{-1}(\mathbf{F}\boldsymbol{\Gamma} + \boldsymbol{\Sigma}) \quad . \end{aligned} \tag{52}$$

Chapter 4

The Linear Inverse Problem With Prior Knowledge

Given the forward problem and priors discussed in the previous chapters, the inverse problem is to find a "best" estimate $\hat{\mathbf{q}}$ of the unknown source amplitude vector $\mathbf{q}$ from given values for the forward transfer matrix $\mathbf{F}$, the field measurements $\mathbf{b}$, and perhaps other information such as the noise covariance $\mathbf{\Sigma}$ and source covariance $\mathbf{A}$.

This dissertation considers only linear inverse methods. That is, the best estimate $\hat{\mathbf{q}}$ is always computed in the form

$$\hat{\mathbf{q}} = \mathbf{H}\mathbf{b} \tag{53}$$

where $\mathbf{H}$ is a linear operator depending on $\mathbf{F}$, $\mathbf{A}$, $\mathbf{\Sigma}$, and $\mathbf{\Gamma}$.

The simple solution $\hat{\mathbf{q}} = \mathbf{F}^{-1}\mathbf{b}$ does not work in general for inverse problems; $\mathbf{F}$ is rarely invertible and usually rectangular. The inverse problem is often both overdetermined in the sense that no solution $\hat{\mathbf{q}}$ exactly solves $\mathbf{F}\mathbf{q} = \mathbf{b}$ and underdetermined in that many different values of $\hat{\mathbf{q}}$ provide equally close approximations. Different criteria for the "best" approximation lead to different inverse methods.

4.1 Minimum-Norm Least Squares Methods

The least squares (LS) criterion is to minimize the residual

$$\varepsilon^2 = \|\mathbf{b} - \mathbf{F}\hat{\mathbf{q}}\|^2 \quad , \tag{54}$$

which is a measure of the discrepancy between the measured and reconstructed field values. The least squares solution is not necessarily unique; there may be many different solutions that achieve the minimum ε^2. The minimum-norm least

squares (MNLS) criterion chooses from all these minimum-residual solutions the unique solution with smallest norm $||\hat{\mathbf{q}}||^2$. No prior information is necessary and the value of ε^2 can be computed for any given $\mathbf{b}$.

The maximum likelihood (ML) criterion is, assuming that $\mathbf{w}$ is jointly Gaussian, to minimize the weighted residual

$$\chi^2 = (\mathbf{b} - \mathbf{F}\hat{\mathbf{q}})^T \boldsymbol{\Sigma}^{-1} (\mathbf{b} - \mathbf{F}\hat{\mathbf{q}}) \quad , \tag{55}$$

which is a measure of the discrepancy between the measured and reconstructed field values, weighted by the a priori noise variance. As with least squares, the solution is not necessarily unique. Noise statistics are required but source statistics are not.

If the noise amplitudes are assumed to be independent and identically distributed ($\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}$), then the MNLS solution minimizes the χ^2 statistic and is also the maximum likelihood solution.

The Moore-Penrose generalized inverse (also known as the pseudoinverse) is a generalization of the matrix inverse to arbitrary rectangular matrices. It was first reported by Moore [61] in 1920 as the unique matrix satisfying certain algebraic conditions not relevant here, and rediscovered by Penrose [67] in 1955 using different but equivalent algebraic conditions. Penrose [68] also showed that the generalized inverse computes the MNLS solution.

The Moore-Penrose inverse can be computed from the singular value decomposition (SVD) [38, page 243] of the forward transfer matrix $\mathbf{F}$. To fix notation, the SVD theorem states that the $M \times N$ matrix $\mathbf{F}$ can be decomposed in the form

$$\mathbf{F} = \mathbf{U} \boldsymbol{\Lambda} \mathbf{V}^T = \sum_{k=1}^{K} \lambda_k \, \mathbf{u}_k \mathbf{v}_k^T \tag{56}$$

where the rank of $\mathbf{F}$ is $K \leq K' = \min(M, N)$; $\mathbf{U}$ is an $N \times N$ orthogonal matrix with orthonormal columns $\mathbf{u}_k$; $\mathbf{V}$ is an $M \times M$ orthogonal matrix with orthonormal columns $\mathbf{v}_k$; and $\boldsymbol{\Lambda}$ is an $M \times N$ diagonal matrix with diagonal entries $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_K > \lambda_{K+1} = \cdots = \lambda_{K'} = 0$. The values λ_k for $k = 1, \ldots, K$ are the singular values of $\mathbf{F}$; the vectors $\mathbf{u}_k$ and $\mathbf{v}_k$ for $k = 1, \ldots, K$ are the left and right singular vectors of $\mathbf{F}$. The additional vectors $\mathbf{u}_k, k = K+1, \ldots, M$, span the complement

of the range space of $\mathbf{F}$; the additional vectors $\mathbf{v}_k, k = K+1, \ldots, N$, span the null space of $\mathbf{F}$.

The Moore-Penrose inverse of $\mathbf{F}$ is then

$$\mathbf{F}^\dagger = \sum_{\lambda_k \neq 0} \lambda_k^{-1} \mathbf{v}_k \mathbf{u}_k^T \tag{57}$$

and is equal to the ordinary matrix inverse if $\mathbf{F}$ is non-singular. If $\mathbf{F}$ is singular, the system $\mathbf{Fq} = \mathbf{b}$ is ill-posed and has infinitely many possible solutions; $\hat{\mathbf{q}} = \mathbf{F}^\dagger \mathbf{b}$ is the particular solution with the smallest norm, or the minimum-norm least squares (MNLS) solution.

The Moore-Penrose inverse in its pure form is not generally suitable for inverse problems with measurement noise. Suppose that $\mathbf{b} = \mathbf{Fq} + \mathbf{w}$ and consider the estimate $\hat{\mathbf{q}} = \mathbf{F}^\dagger \mathbf{b} = \mathbf{F}^\dagger \mathbf{F}\mathbf{q} + \mathbf{F}^\dagger \mathbf{w}$. The error due to noise is

$$\mathbf{F}^\dagger \mathbf{w} = \sum_{\lambda_k \neq 0} \lambda_k^{-1} (\mathbf{u}_k^T \mathbf{w}) \mathbf{v}_k \quad , \tag{58}$$

which grows without bound as the singular values λ_k decrease toward zero. The error in $\hat{\mathbf{q}}$ is roughly proportional to the reciprocal of the smallest singular value and can easily swamp the correct answer.

One simple cure is to drop all singular values less than a threshold ϵ, defining a truncated pseudoinverse

$$\mathbf{F}^\epsilon = \sum_{\lambda_k > \epsilon} \lambda_k^{-1} \mathbf{v}_k \mathbf{u}_k^T \quad . \tag{59}$$

Wang, Williamson, and Kaufman [105, 106] have used a truncated pseudoinverse for magnetic source imaging, although they do not state their truncation criterion. The threshold ϵ must be chosen with care, as is illustrated in Figure 7.3; the weighted pseudoinverse method discussed below provides one systematic approach.

4.2 Weighted Pseudoinverse Methods

The minimum mean square error (MMSE) or mean square criterion is to minimize the average reconstruction error

$$\eta^2 = \mathsf{E}\, e^2 = \mathsf{E}\, \|\mathbf{q} - \hat{\mathbf{q}}\|^2 \quad , \tag{60}$$

which is a measure of the discrepancy between the reconstructed and the true current distributions. The error $e^2 = \|\mathbf{q} - \hat{\mathbf{q}}\|^2$ depends on the true distribution and cannot be computed for any specific $\mathbf{b}$; only its mean value η^2 can be determined. The mean square error η^2 depends on the source and noise statistics, so both of these must be known (or assumed) a priori.

To avoid the numerical problems of MNLS, Shim and Cho [91] have defined a weighted pseudoinverse (which they call the stochastic SVD pseudoinverse)

$$\mathbf{F}^w = \sum_{\lambda_k \neq 0} c_k \lambda_k^{-1} \mathbf{v}_k \mathbf{u}_k^T \quad , \tag{61}$$

where the weights c_k are chosen to yield the minimum mean square error. Restricting the weights to zero and one yields a truncated pseudoinverse. Their derivation and results unfortunately contain some typographical errors; the following derivation hopefully corrects those errors without introducing any new ones, and generalizes their result by allowing correlations between noise and source amplitudes ($\mathbf{\Gamma} \neq 0$).

Shim and Cho have applied the optimal truncated pseudoinverse method (OTPIM) and optimal weighted pseudoinverse method (OWPIM) to PET reconstruction; Jeffs, Leahy, and Singh [45] have used the optimal truncated pseudoinverse for magnetic source imaging of the brain.

To determine the optimal c_k, write the source and noise vectors in terms of the singular vectors of $\mathbf{F}$ to obtain

$$\mathbf{q} = \sum_{k=1}^{N} a_k \mathbf{v}_k \qquad \text{and} \qquad \mathbf{w} = \sum_{k=1}^{M} s_k \mathbf{u}_k \quad . \tag{62}$$

Note that $a_k = \mathbf{v}_k^T \mathbf{q}$ and $s_k = \mathbf{u}_k^T \mathbf{w}$, so that $\mathsf{E}\, a_k = 0$ and $\mathsf{E}\, s_k = 0$. Define the covariances

$$\alpha_k^2 = \mathsf{E}\, a_k^2 \quad = \mathbf{v}_k^T \mathbf{A} \mathbf{v}_k \quad , \qquad k = 1, \ldots, N \quad ; \tag{63a}$$

$$\sigma_k^2 = \mathsf{E}\, s_k^2 \quad = \mathbf{u}_k^T \mathbf{\Sigma} \mathbf{u}_k \quad , \qquad k = 1, \ldots, M \quad ; \text{and} \tag{63b}$$

$$\gamma_k = \mathsf{E}\, a_k\, s_k = \mathbf{v}_k^T \mathbf{\Gamma} \mathbf{u}_k \quad , \qquad k = 1, \ldots, K' \tag{63c}$$

which are the diagonal entries of the rotated covariance matrices $\mathbf{V}^T \mathbf{A} \mathbf{V}$, $\mathbf{U}^T \mathbf{\Sigma} \mathbf{U}$, and $\mathbf{V}^T \mathbf{\Gamma} \mathbf{U}$ respectively. The notation here may be a bit confusing; the symbols

α_{ij}^2, σ_{ij}^2, and γ_{ij} with *two* subscripts denote the entries of the *unrotated* covariance matrices $\mathbf{A}$, $\mathbf{\Sigma}$, and $\mathbf{\Gamma}$.

The optimal estimate of $\mathbf{q}$ is $\hat{\mathbf{q}} = \mathbf{F}^w\mathbf{b}$ with error

$$\begin{aligned}\mathbf{e} &= \mathbf{q} - \hat{\mathbf{q}} = \mathbf{q} - \mathbf{F}^w\mathbf{b} = \mathbf{q} - \mathbf{F}^w(\mathbf{F}\mathbf{q} + \mathbf{w}) \\ &= \mathbf{q} - \mathbf{F}^w\mathbf{F}\mathbf{q} - \mathbf{F}^w\mathbf{w} \\ &= \sum_{k=1}^{N} a_k\mathbf{v}_k - \mathbf{F}^w\mathbf{F}\sum_{k=1}^{N} a_k\mathbf{v}_k - \mathbf{F}^w\sum_{k=1}^{M} s_k\mathbf{u}_k \quad ;\end{aligned}$$

and since $\mathbf{v}_{K+1}, \ldots, \mathbf{v}_N$ are in the null space of $\mathbf{F}$ and $\mathbf{u}_{K+1}, \ldots, \mathbf{u}_M$ are in the null space of $\mathbf{F}^w$,

$$\begin{aligned}\mathbf{e} &= \sum_{k=1}^{N} a_k\mathbf{v}_k - \mathbf{F}^w\mathbf{F}\sum_{k=1}^{K} a_k\mathbf{v}_k - \mathbf{F}^w\sum_{k=1}^{K} s_k\mathbf{u}_k \\ &= \sum_{k=1}^{K}\left[(1-c_k)a_k - \frac{c_k\, s_k}{\lambda_k}\right]\mathbf{v}_k + \sum_{k=K+1}^{N} a_k\mathbf{v}_k \quad . \end{aligned} \tag{64}$$

Then the mean square error is

$$\begin{aligned}\eta^2 &= \mathsf{E}\,\|\mathbf{e}\|^2 = \mathsf{E}\,\mathbf{e}^T\mathbf{e} \\ &= \sum_{k=1}^{K}\left[(c_k-1)^2\alpha_k^2 + \frac{2c_k(c_k-1)\gamma_k}{\lambda_k} + \frac{c_k^2\sigma_k^2}{\lambda_k^2}\right] + \sum_{k=K+1}^{N}\alpha_k^2 \quad . \end{aligned} \tag{65}$$

Finding the minimum by the usual procedure yields the coefficients

$$c_k = \frac{\alpha_k^2\lambda_k^2 + \gamma_k\lambda_k}{\alpha_k^2\lambda_k^2 + 2\gamma_k\lambda_k + \sigma_k^2} \quad , \qquad k = 1, \ldots, K \tag{66}$$

for the optimal weighted pseudoinverse of $\mathbf{F}$.

Note that the off-diagonal covariances $\mathsf{E}\,a_k a_l$, $\mathsf{E}\,s_k s_l$, and $\mathsf{E}\,a_k s_l$ with $k \neq l$ do not appear in the above expression. In effect, they are assumed to be zero. Equivalently, the covariance matrices $\mathbf{V}^T\mathbf{A}\mathbf{V}$, $\mathbf{U}^T\mathbf{\Sigma}\mathbf{U}$, and $\mathbf{V}^T\mathbf{\Gamma}\mathbf{U}$ are assumed to be diagonal. This is called the case of "coaxial" priors and will be discussed later.

If the weights are restricted to $c_k = 0$ or $c_k = 1$, the weighted pseudoinverse becomes a truncated pseudoinverse. Thus, any truncated pseudoinverse is a special case of the weighted pseudoinverse and will, in general, have a larger error than the optimal weighted pseudoinverse.

To obtain the optimal truncated pseudoinverse, consider the kth error term

$$\eta_k^2 = (c_k - 1)^2 \alpha_k^2 + \frac{2c_k(c_k-1)\gamma_k}{\lambda_k} + \frac{c_k^2 \sigma_k^2}{\lambda_k^2}$$
$$= \begin{cases} \alpha_k^2 & \text{if } c_k = 0 \\ \sigma_k^2/\lambda_k^2 & \text{if } c_k = 1 \end{cases} \quad . \tag{67}$$

The minimum error is achieved when the smaller of the two possible values is taken; that is, the kth term should be retained if $\sigma_k^2/\lambda_k^2 < \alpha_k^2 \Leftrightarrow \lambda_k > \sigma_k/\alpha_k$ and dropped otherwise. Thus the optimal truncated pseudoinverse is

$$\mathbf{F}^t = \sum_{\lambda_k > \sigma_k/\alpha_k} \lambda_k^{-1} \mathbf{v}_k \mathbf{u}_k^T \quad . \tag{68}$$

The ratio $\alpha_k^2 \lambda_k^2/\sigma_k^2$ can be interpreted as the signal-to-noise ratio for the source vector $\mathbf{v}_k$, so that the optimal truncated pseudoinverse retains exactly those terms for which the SNR exceeds one. Setting $\sigma_k^2/\lambda_k^2 = \alpha_k^2$ in equation (66) yields $c_k = \frac{1}{2}$, so rounding c_k to 0 or 1 in the optimal weighted pseudoinverse yields the optimal truncated pseudoinverse. Note that the optimally truncated pseudoinverse does not necessarily comprise the largest singular values.

4.3 Optimal Constrained Linear Inverse Method

The Shim-Cho weighted pseudoinverse is optimal (in the MMSE sense) over all possible weighted pseudoinverses but is not, in general, optimal over all possible linear estimators of the form $\hat{\mathbf{q}} = \mathbf{Hb}$.

The optimal constrained linear estimator is given by the matrix $\mathbf{H}$ that minimizes

$$\eta^2 = \mathsf{E}\,\|\mathbf{q} - \hat{\mathbf{q}}\|^2 = \mathsf{E}\,\|\mathbf{q} - \mathbf{Hb}\|^2 \quad , \tag{69}$$

and since $\mathbf{b} = \mathbf{Fq} + \mathbf{w}$,

$$= \mathsf{E}\,\|\mathbf{q} - \mathbf{H}(\mathbf{Fq} + \mathbf{w})\|^2$$
$$= \mathsf{E}\,\|(\mathbf{I} - \mathbf{HF})\mathbf{q} - \mathbf{Hw}\|^2 \quad . \tag{70}$$

Now consider a variation $\mathbf{H} + \epsilon\,\delta\mathbf{H}$ where ϵ is a scalar and $\delta\mathbf{H}$ is a matrix to obtain

$$\eta^2 = \mathsf{E}\,\|(\mathbf{I} - \mathbf{HF} - \epsilon\,\delta\mathbf{HF})\mathbf{q} - \mathbf{Hw} - \epsilon\,\delta\mathbf{Hw}\|^2$$

$$
\begin{aligned}
&= \mathsf{E}\,\|(\mathbf{I}-\mathbf{HF})\mathbf{q}-\mathbf{Hw}-\epsilon\,\delta\mathbf{H}(\mathbf{Fq}+\mathbf{w})\|^2 \\
&= \mathsf{E}\,\|(\mathbf{I}-\mathbf{HF})\mathbf{q}-\mathbf{Hw}\|^2 \\
&\quad + 2\epsilon\ \mathsf{E}(\mathbf{Fq}+\mathbf{w})^T\,\delta\mathbf{H}^T[(\mathbf{I}-\mathbf{HF})\mathbf{q}-\mathbf{Hw}] \\
&\quad + \epsilon^2\ \mathsf{E}\,\|\delta\mathbf{H}(\mathbf{Fq}+\mathbf{w})\|^2 \\
&= c_0 + c_1\epsilon + c_2\epsilon^2 \qquad (71)
\end{aligned}
$$

for scalars c_0, c_1, and c_2 which depend on $\delta\mathbf{H}$.

Now $\mathbf{H}$ can minimize η^2 only if c_1 is zero for any value of $\delta\mathbf{H}$. To see this, assume on the contrary that there is some non-zero $\delta\mathbf{H}$ such that c_1 is not zero. Assumption A ensures that $c_2 = \mathsf{E}\,\|\delta\mathbf{H}(\mathbf{Fq}+\mathbf{w})\|^2 = \mathsf{E}\,\|\delta\mathbf{Hb}\|^2$ is positive whenever $\delta\mathbf{H}$ is not zero. Then η^2 has a unique minimum at $\epsilon_{\min} = -c_1/(2c_2)$. But this means that η^2 is smaller for $\mathbf{H}+\epsilon_{\min}\,\delta\mathbf{H}$ than for $\mathbf{H}$. Thus, contrary to our assumption, $\mathbf{H}$ is not the minimum. Therefore, c_1 must be zero for every $\delta\mathbf{H}$.

The positivity of c_2 also guarantees that η^2 has a minimum rather than a maximum. Furthermore, since η^2 is quadratic in ϵ, the minimum is unique and there is no maximum. That is, the optimal $\mathbf{H}$ exists and is unique.

Now observe that, for random vectors $\mathbf{x}$ and $\mathbf{y}$, the expectation $\mathsf{E}\,\mathbf{x}^T\mathbf{Dy}$ is zero for all $\mathbf{D}$ only if $\mathsf{E}\,\mathbf{xy}^T = 0$. (Consider the set of matrices $\mathbf{D}$ in which one entry takes the value one and all the other entries are zero.) Then the scalar c_1 can be zero for arbitrary $\delta\mathbf{H}$ only if

$$
\begin{aligned}
0 &= \mathsf{E}\,[(\mathbf{I}-\mathbf{HF})\mathbf{q}-\mathbf{Hw}](\mathbf{Fq}+\mathbf{w})^T \\
&= \mathsf{E}\,\mathbf{HFqq}^T\mathbf{F}^T - \mathsf{E}\,\mathbf{qq}^T\mathbf{F}^T + \mathsf{E}\,\mathbf{Hwq}^T\mathbf{F}^T \\
&\quad + \mathsf{E}\,\mathbf{HFqw}^T - \mathsf{E}\,\mathbf{qw}^T + \mathsf{E}\,\mathbf{Hww}^T \\
&= \mathbf{HFAF}^T - \mathbf{AF}^T + \mathbf{H\Gamma}^T\mathbf{F}^T \\
&\quad + \mathbf{HF\Gamma} - \mathbf{\Gamma} + \mathbf{H\Sigma} \quad , \qquad (72)
\end{aligned}
$$

which can be solved to yield

$$
\mathbf{H} = (\mathbf{AF}^T + \mathbf{\Gamma})(\mathbf{FAF}^T + \mathbf{F\Gamma} + \mathbf{\Gamma}^T\mathbf{F}^T + \mathbf{\Sigma})^{-1} \qquad (73)
$$

where assumption A guarantees the existence of the inverse. Franklin [34] has developed essentially the same formula in the more general context of Hilbert space.

4.4 Useful Special Cases

There are several special cases in which the OCLIM filter reduces to previously known inverse methods.

4.4.1 Noise Uncorrelated with Sources

If the noise is uncorrelated with the sources, then $\mathbf{\Gamma} = 0$ and the filter simplifies to

$$\mathbf{H} = \mathbf{A}\mathbf{F}^T(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1} \quad . \tag{74}$$

This mean square estimator has been previously used by Helstrom [42] for image restoration and by Smith et al. [93] for magnetic source imaging.

The preconditioned pseudoinverse of Ferguson [30] has the form

$$\tilde{\mathbf{H}} = \mathbf{D}\tilde{\mathbf{L}}^T(\tilde{\mathbf{L}}\tilde{\mathbf{L}}^T)^{-1} \tag{75}$$

where $\mathbf{D}$ is a diagonal matrix of the expected source amplitudes and $\tilde{\mathbf{L}} = \mathbf{F}\mathbf{D}$ is a preconditioned forward transfer matrix. Substituting $\mathbf{A} = \mathbf{D}\mathbf{D}^T$ and $\mathbf{\Sigma} = 0$ in equation (74) yields

$$\begin{aligned} \mathbf{H} &= \mathbf{D}(\mathbf{D}^T\mathbf{F}^T)((\mathbf{F}\mathbf{D})(\mathbf{D}^T\mathbf{F}^T) + 0)^{-1} \\ &= \mathbf{D}\tilde{\mathbf{L}}(\tilde{\mathbf{L}}\tilde{\mathbf{L}}^T)^{-1} \quad , \end{aligned} \tag{76}$$

which is exactly the preconditioned pseudoinverse.

The OCLIM filter (74) is also related to the filtered backprojection algorithm for emission tomography[pages 60–63] [47]. The factor $(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1}$ filters the projections $\mathbf{b}$, the factor $\mathbf{F}^T$ backprojects, and $\mathbf{A}$ weights the result according to the a priori probabilities.

Furthermore, observe that

$$\begin{aligned} & \mathbf{F}^T + \mathbf{F}^T\mathbf{\Sigma}^{-1}\mathbf{F}\mathbf{A}\mathbf{F}^T = \mathbf{F}^T\mathbf{\Sigma}^{-1}\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{F}^T \\ \Leftrightarrow \quad & (\mathbf{A}^{-1} + \mathbf{F}^T\mathbf{\Sigma}^{-1}\mathbf{F})\mathbf{A}\mathbf{F}^T = \mathbf{F}^T\mathbf{\Sigma}^{-1}(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma}) \\ \Leftrightarrow \quad & \mathbf{A}\mathbf{F}^T(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1} = (\mathbf{A}^{-1} + \mathbf{F}^T\mathbf{\Sigma}^{-1}\mathbf{F})^{-1}\mathbf{F}^T\mathbf{\Sigma}^{-1} \quad . \end{aligned} \tag{77}$$

Thus the optimal filter can also be written in the form

$$\mathbf{H} = (\mathbf{A}^{-1} + \mathbf{F}^T\mathbf{\Sigma}^{-1}\mathbf{F})^{-1}\mathbf{F}^T\mathbf{\Sigma}^{-1} \quad . \tag{78}$$

This form represents a variant backprojection algorithm in which filtering is done after backprojection; Budinger et al. call this the "filter of the backprojection algorithm" [14]. The factor $\mathbf{\Sigma}^{-1}$ normalizes the noise variance in the projections, $\mathbf{F}^T$ backprojects, and $(\mathbf{A}^{-1} + \mathbf{F}^T\mathbf{\Sigma}^{-1}\mathbf{F})^{-1}$ filters the backprojected image.

4.4.2 Coaxial priors

Suppose that the covariance matrices $\mathbf{A}$, $\mathbf{\Sigma}$, and $\mathbf{\Gamma}$ all become diagonal when they are rotated into the $\mathbf{U}$ and $\mathbf{V}$ coordinates defined by the singular vectors of $\mathbf{F}$. That is, suppose that the rotated matrices $\mathbf{V}^T\mathbf{A}\mathbf{V}$, $\mathbf{U}^T\mathbf{\Sigma}\mathbf{U}$, and $\mathbf{V}^T\mathbf{\Gamma}\mathbf{U}$ are all diagonal. Then the matrices $\mathbf{F}$, $\mathbf{A}$, $\mathbf{\Sigma}$, and $\mathbf{\Gamma}$ can *all* be written in terms of the singular vectors $\mathbf{u}_k$ and $\mathbf{v}_k$ defined by the singular value decomposition of $\mathbf{F}$, or

$$\mathbf{F} = \sum_{k=1}^{K} \lambda_k \mathbf{u}_k \mathbf{v}_k^T \quad , \tag{79a}$$

$$\mathbf{A} = \sum_{k=1}^{N} \alpha_k^2 \mathbf{v}_k \mathbf{v}_k^T \quad , \tag{79b}$$

$$\mathbf{\Sigma} = \sum_{k=1}^{M} \sigma_k^2 \mathbf{u}_k \mathbf{u}_k^T \quad , \text{ and} \tag{79c}$$

$$\mathbf{\Gamma} = \sum_{k=1}^{K'} \gamma_k \mathbf{v}_k \mathbf{u}_k^T \quad . \tag{79d}$$

In other words, these matrices all share the same singular axes; this may be called the case of "coaxial priors." Since $\mathbf{A}$ and $\mathbf{\Sigma}$ are positive definite, $\alpha_k^2 > 0$ for $k = 1, \ldots, N$, and $\sigma_k^2 > 0$ for $k = 1, \ldots, M$. Recall that $K = \text{rank}(\mathbf{F}) \leq K' = \min(M, N)$.

Then the OCLIM filter can be expressed in the same axes as

$$\begin{aligned} \mathbf{H} &= (\mathbf{A}\mathbf{F}^T + \mathbf{\Gamma})(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{F}\mathbf{\Gamma} + \mathbf{\Gamma}^T\mathbf{F}^T + \mathbf{\Sigma})^{-1} \\ &= \sum_{k=1}^{K'} \frac{\alpha_k^2\lambda_k^2 + \gamma_k\lambda_k}{\alpha_k^2\lambda_k^2 + 2\gamma_k\lambda_k + \sigma_k^2} \lambda_k^{-1} \mathbf{v}_k \mathbf{u}_k^T \quad , \end{aligned} \tag{80}$$

which is exactly the optimal weighted pseudoinverse of $\mathbf{F}$. That is, the optimal weighted pseudoinverse is obtained as a special case of the optimal constrained

linear filter when the covariance matrices are coaxial with the forward transfer matrix; it is inferior to OCLIM otherwise.

Except for the special case of uniform priors (discussed below), coaxial priors seem unlikely to occur in practice.

The truncated pseudoinverse can also be obtained as a special case of coaxial priors, although the assumptions required are rather perverse: $\sigma_k = 0 = \gamma_k$ implies $c_k = 1$; and either $\sigma_k = \infty$ or $\alpha_k = 0 = \gamma_k$ implies $c_k = 0$.

4.4.3 Stationary Priors

Suppose that the source and noise amplitudes are (spatially) stationary, that there are the same number of sources and detectors ($M = N$), and that the forward transfer function is shift-invariant. Then $\mathbf{A}$, $\mathbf{\Sigma}$, and $\mathbf{F}$ are all square and (approximately) circulant. Thus [44, p. 145], they are all diagonalized by the discrete Fourier transform represented as the unitary matrix

$$[\mathbf{W}]_{mn} = \frac{1}{\sqrt{N}} e^{-j2\pi mn/N} \quad . \tag{81}$$

That is, $\tilde{\mathbf{A}} = \mathbf{WAW}^*$ is diagonal and the diagonal entries, which we will denote $P_q(k)$, constitute the power spectrum of the source amplitudes q_n. Similarly, $\tilde{\mathbf{\Sigma}} = \mathbf{W\Sigma W}^*$ is diagonal and its diagonal entries $P_w(k)$ constitute the noise power spectrum. Furthermore, $\tilde{\mathbf{F}} = \mathbf{WFW}^*$ is diagonal and its diagonal entries $F(k)$ constitute the transfer function in frequency space. Since $\mathbf{F}$ is real, $\mathbf{F}^T = \mathbf{F}^*$ and the optimal filter can be written as $\mathbf{H} = \mathbf{AF}^*(\mathbf{FAF}^* + \mathbf{\Sigma})^{-1}$, or

$$\tilde{\mathbf{H}} = \mathbf{WHW}^* = \tilde{\mathbf{A}}\tilde{\mathbf{F}}^*(\tilde{\mathbf{F}}\tilde{\mathbf{A}}\tilde{\mathbf{F}}^* + \tilde{\mathbf{\Sigma}})^{-1} \quad ,$$

and, since all the matrices are diagonal,

$$H(k) = \frac{P_q(k)F^*(k)}{|F(k)|^2 P_q(k) + P_w(k)} \quad , \tag{82}$$

which is the Wiener deconvolution filter [15, pages 209–210].

4.4.4 Uniform Priors

In the case of "uniform priors," the approximate amplitude of the elementary sources is known but there is no basis for believing that any one is more active than any other; and the same is true of the noise amplitudes. That is, the source amplitudes are independent and identically distributed; so are the noise amplitudes. Then every source has the same expected activity α^2 and the covariance matrix $\mathbf{A}$ takes the form $\alpha^2\mathbf{I}$; the noise covariance $\mathbf{\Sigma}$ takes the form $\sigma^2\mathbf{I}$ and the cross-covariance $\mathbf{\Gamma}$ is zero. It follows that $\mathbf{A}$, $\mathbf{\Sigma}$, and $\mathbf{\Gamma}$ are all coaxial with any forward transfer matrix $\mathbf{F}$. Then both OCLIM and OWPIM simplify to

$$\begin{aligned}\mathbf{H} &= \mathbf{F}^T(\mathbf{F}\mathbf{F}^T + (\sigma^2/\alpha^2)\mathbf{I})^{-1} \\ &= \sum_{k=1}^{K} \frac{\lambda_k^2}{\lambda_k^2 + \sigma^2/\alpha^2}\, \lambda_k^{-1}\mathbf{v}_k\mathbf{u}_k^T \quad ,\end{aligned} \tag{83}$$

where we have used the fact that $\lambda_k = 0$ whenever $k > K$ to reduce the upper limit from K' to K. Since $\mathbf{F}\mathbf{F}^T$ is positive semidefinite, the indicated matrix inverse exists whenever $\sigma^2/\alpha^2 > 0$. The only prior knowledge required is the ratio α^2/σ^2; thus this special case is useful for noise-tolerant reconstruction given only rough estimates of source and noise amplitude.

This form can be regarded as a Marquardt [54, 55] or Tikhonov [99] regularization of the pseudoinverse $\mathbf{F}^\dagger = \mathbf{F}^T(\mathbf{F}\mathbf{F}^T)^{-1}$. One important difference is that the value of the regularization parameter is determined by the given value of σ^2/α^2 and need not be determined by experiment.

4.4.5 No Prior Information

The case of "no priors" may be approached by letting α^2 go to infinity. In this case, the OCLIM filter goes to the limit

$$\mathbf{H} = \sum_{k=1}^{K} \lambda_k^{-1}\, \mathbf{v}_k\mathbf{u}_k^T \quad , \tag{84}$$

which is just the Moore-Penrose inverse or the MNLS estimator. Taking the limit as the noise goes to zero ($\sigma^2 \to 0$) yields the same result.

4.5 Maximum A Posteriori Estimation

The OCLIM estimator can also be obtained as the maximum a posteriori estimator when $\mathbf{q}$ and $\mathbf{w}$ are Gaussian but uncorrelated. In maximum a posteriori (MAP) estimation, we seek the estimate $\hat{\mathbf{q}}$ which maximizes the a posteriori probability $p(\mathbf{q} \mid \mathbf{b})$ of $\mathbf{q}$ given a set of measurements $\mathbf{b}$.

Now suppose that $\mathbf{q}$ and $\mathbf{w}$ are uncorrelated zero-mean Gaussian random vectors with respective covariance matrices $\mathbf{A}$ and $\mathbf{\Sigma}$. Given these, the a priori probability of a given source vector $\mathbf{q}$ is

$$p(\mathbf{q}) = ((2\pi)^N |\mathbf{A}|)^{-1/2} \exp(-\tfrac{1}{2}\mathbf{q}^T \mathbf{A}^{-1} \mathbf{q}) \quad . \tag{85}$$

Similarly, the probability of a given noise vector $\mathbf{w}$ is

$$p(\mathbf{w}) = ((2\pi)^M |\mathbf{\Sigma}|)^{-1/2} \exp(-\tfrac{1}{2}\mathbf{w}^T \mathbf{\Sigma}^{-1} \mathbf{w}) \quad . \tag{86}$$

Since $\mathbf{w} = \mathbf{b} - \mathbf{F}\mathbf{q}$, the probability of a given measurement $\mathbf{b}$ given a source vector $\mathbf{q}$ must be

$$p(\mathbf{b} \mid \mathbf{q}) = ((2\pi)^M |\mathbf{\Sigma}|)^{-1/2} \exp(-\tfrac{1}{2}((\mathbf{b} - \mathbf{F}\mathbf{q})^T \mathbf{\Sigma}^{-1} (\mathbf{b} - \mathbf{F}\mathbf{q})) \quad . \tag{87}$$

Given these, Bayes' rule gives the a posteriori distribution of $\mathbf{q}$ as

$$p(\mathbf{q} \mid \mathbf{b}) = \frac{p(\mathbf{b} \mid \mathbf{q})\, p(\mathbf{q})}{p(\mathbf{b})} \quad . \tag{88}$$

Since we are seeking the maximum with respect to $\mathbf{q}$ and the denominator $p(\mathbf{b})$ is independent of $\mathbf{q}$, it suffices to maximize the product

$$\begin{aligned} p(\mathbf{b}|\mathbf{q})\, p(\mathbf{q}) = {} & ((2\pi)^M |\mathbf{\Sigma}|)^{-1/2} \exp(-\tfrac{1}{2}(\mathbf{b} - \mathbf{F}\mathbf{q})^T \mathbf{\Sigma}^{-1} (\mathbf{b} - \mathbf{F}\mathbf{q})) \\ & \times ((2\pi)^N |\mathbf{A}|)^{-1/2} \exp(-\tfrac{1}{2}\mathbf{q}^T \mathbf{A}^{-1} \mathbf{q}) \quad . \end{aligned} \tag{89}$$

But maximizing this is equivalent to maximizing

$$\log(p(\mathbf{b}|\mathbf{q})\, p(\mathbf{q})) = -\tfrac{1}{2}(\mathbf{b} - \mathbf{F}\mathbf{q})^T \mathbf{\Sigma}^{-1} (\mathbf{b} - \mathbf{F}\mathbf{q}) - \tfrac{1}{2}\mathbf{q}^T \mathbf{A}^{-1} \mathbf{q} \tag{90}$$

or minimizing

$$\tau^2 = (\mathbf{b} - \mathbf{F}\mathbf{q})^T \mathbf{\Sigma}^{-1} (\mathbf{b} - \mathbf{F}\mathbf{q}) + \mathbf{q}^T \mathbf{A}^{-1} \mathbf{q} \quad . \tag{91}$$

The minimum can be determined by a variational analysis. Suppose that $\hat{\mathbf{q}}$ minimizes τ^2 and write $\mathbf{q} = \hat{\mathbf{q}} + \delta\mathbf{q}$. Then

$$\begin{aligned}\tau^2 &= (\hat{\mathbf{q}} + \delta\mathbf{q})^T \mathbf{A}^{-1}(\hat{\mathbf{q}} + \delta\mathbf{q}) + (\mathbf{F}\hat{\mathbf{q}} - \mathbf{b} + \mathbf{F}\delta\mathbf{q})^T \boldsymbol{\Sigma}^{-1}(\mathbf{F}\hat{\mathbf{q}} - \mathbf{b} + \mathbf{F}\delta\mathbf{q}) \\ &= \hat{\mathbf{q}}^T \mathbf{A}^{-1}\hat{\mathbf{q}} + (\mathbf{F}\hat{\mathbf{q}} - \mathbf{b})^T \boldsymbol{\Sigma}^{-1}(\mathbf{F}\hat{\mathbf{q}} - \mathbf{b}) \qquad \text{(line 1)} \\ &+ 2\delta\mathbf{q}^T \mathbf{A}^{-1}\hat{\mathbf{q}} + 2\delta\mathbf{q}^T \mathbf{F}^T \boldsymbol{\Sigma}^{-1}(\mathbf{F}\hat{\mathbf{q}} - \mathbf{b}) \qquad \text{(line 2)} \\ &+ \delta\mathbf{q}^T \mathbf{A}^{-1}\delta\mathbf{q} + \delta\mathbf{q}^T \mathbf{F}^T \boldsymbol{\Sigma}^{-1}\mathbf{F}\delta\mathbf{q} \quad . \qquad \text{(line 3)}\end{aligned}$$

For $\hat{\mathbf{q}}$ to be a minimum, we must have line 3 non-negative and line 2 zero for any value of $\delta\mathbf{q}$. Since $\mathbf{A}$ and $\boldsymbol{\Sigma}$ are positive definite, line 3 is positive whenever $\delta\mathbf{q}$ is non-zero. Line 2 will be zero for any $\delta\mathbf{q}$ provided that

$$\mathbf{A}^{-1}\hat{\mathbf{q}} + \mathbf{F}^T \boldsymbol{\Sigma}^{-1}(\mathbf{F}\hat{\mathbf{q}} - \mathbf{b}) = 0 \quad . \tag{92}$$

Solving for $\hat{\mathbf{q}}$ yields

$$\hat{\mathbf{q}} = (\mathbf{A}^{-1} + \mathbf{F}^T \boldsymbol{\Sigma}^{-1}\mathbf{F})^{-1}\, \mathbf{F}^T \boldsymbol{\Sigma}^{-1}\mathbf{b} \quad , \tag{93}$$

which is one form of the OCLIM estimator.

4.6 Nuisance Sources

In some applications it happens that many sources are expected but only some of them are of interest to the experimenter. This section considers whether estimating the amplitudes of the uninteresting "nuisance" sources will improve the estimates for the interesting source amplitudes. Suppose that the source vector $\mathbf{q}$ is partitioned into a vector of N_1 interesting sources $\mathbf{q}_1$ and N_2 nuisance sources $\mathbf{q}_2$ arranged as

$$\mathbf{q} = \begin{bmatrix} \mathbf{q}_1 \\ \mathbf{q}_2 \end{bmatrix} \quad . \tag{94}$$

Then we wish to find a linear estimator $\mathbf{H}_1$ in the computation $\hat{\mathbf{q}}_1 = \mathbf{H}_1\mathbf{b}$ that will minimize the error

$$\eta_1^2 = \mathsf{E}\,\|\mathbf{q}_1 - \hat{\mathbf{q}}_1\|^2 \tag{95}$$

in the interesting sources alone.

The expected mean square error can be rewritten in the form

$$\eta_1^2 = \mathsf{E}(\mathbf{q}_1 - \mathbf{H}_1\mathbf{F}\mathbf{q} - \mathbf{H}_1\mathbf{w})^T(\mathbf{q}_1 - \mathbf{H}_1\mathbf{F}\mathbf{q} - \mathbf{H}\mathbf{w}) \tag{96}$$

and introducing a variation $\mathbf{H}_1 + \delta\mathbf{H}$ yields

$$\begin{aligned}\eta_1^2 + \delta\eta_1^2 = \mathsf{E}(\mathbf{q}_1 - \mathbf{H}_1\mathbf{F}\mathbf{q} - \mathbf{H}_1\mathbf{w})^T(\mathbf{q}_1 - \mathbf{H}_1\mathbf{F}\mathbf{q} - \mathbf{H}_1\mathbf{w}) \\ - 2\,\mathsf{E}(\mathbf{F}\mathbf{q} + \mathbf{w})^T\delta\mathbf{H}^T(\mathbf{q}_1 - \mathbf{H}_1\mathbf{F}\mathbf{q} - \mathbf{H}_1\mathbf{w}) \\ + \mathsf{E}(\mathbf{F}\mathbf{q} + \mathbf{w})^T\delta\mathbf{H}^T\delta\mathbf{H}(\mathbf{F}\mathbf{q} + \mathbf{w}) \quad .\end{aligned} \tag{97}$$

The existence of an extremum requires that

$$\mathsf{E}(\mathbf{q}_1 - \mathbf{H}_1\mathbf{F}\mathbf{q} - \mathbf{H}_1\mathbf{w})(\mathbf{F}\mathbf{q} + \mathbf{w})^T = 0 \quad , \tag{98}$$

which simplifies after some algebra to

$$\mathbf{H}_1 = (\mathbf{A}_1\mathbf{F}^T + \mathbf{\Gamma}_1)(\mathbf{F}\mathbf{A}\mathbf{F} + \mathbf{F}\mathbf{\Gamma} + \mathbf{\Gamma}^T\mathbf{F}^T + \mathbf{\Sigma})^{-1} \quad , \tag{99}$$

where

$$\mathbf{A}_1 = \mathsf{E}\,\mathbf{q}_1\mathbf{q}^T \qquad \text{and} \qquad \mathbf{\Gamma}_1 = \mathsf{E}\,\mathbf{q}_1\mathbf{w}^T \tag{100}$$

consist of the first N_1 rows of $\mathbf{A}$ and $\mathbf{\Gamma}$. In fact, the selective estimator $\mathbf{H}_1$ consists of the first N_1 rows of the complete estimator $\mathbf{H}$.

We can further write

$$\mathbf{A}_1 = \mathsf{E}\,\mathbf{q}_1\mathbf{q} = [\mathsf{E}\,\mathbf{q}_1\mathbf{q}_1^T \quad \mathsf{E}\,\mathbf{q}_1\mathbf{q}_2^T] \triangleq [\mathbf{A}_{11} \quad \mathbf{A}_{12}] \tag{101}$$

and

$$\mathbf{b} = \mathbf{F}\mathbf{q} + \mathbf{w} = [\mathbf{F}_1 \quad \mathbf{F}_2]\begin{bmatrix}\mathbf{q}_1 \\ \mathbf{q}_2\end{bmatrix} + \mathbf{w} \quad . \tag{102}$$

Then $\mathbf{A}_1\mathbf{F}^T + \mathbf{\Gamma}_1 = \mathbf{A}_{11}\mathbf{F}_1 + \mathbf{A}_{12}\mathbf{F}_2 + \mathbf{\Gamma}_1$ and the selective estimator becomes

$$\mathbf{H}_1 = (\mathbf{A}_{11}\mathbf{F}_1 + \mathbf{A}_{12}\mathbf{F}_2 + \mathbf{\Gamma}_1)(\mathbf{F}\mathbf{A}\mathbf{F} + \mathbf{F}\mathbf{\Gamma} + \mathbf{\Gamma}^T\mathbf{F}^T + \mathbf{\Sigma})^{-1} \quad . \tag{103}$$

If the interesting and nuisance sources are uncorrelated, then $\mathbf{A}_{12} = 0$ and the selective estimator is

$$\mathbf{H}_1 = (\mathbf{A}_{11}\mathbf{F}_1 + \mathbf{\Gamma}_1)(\mathbf{F}\mathbf{A}\mathbf{F} + \mathbf{F}\mathbf{\Gamma} + \mathbf{\Gamma}^T\mathbf{F}^T + \mathbf{\Sigma})^{-1} \quad . \tag{104}$$

That is, the nuisance sources must be included in the measurement covariance

$$\mathbf{B} = \mathbf{F}\mathbf{A}\mathbf{F} + \mathbf{F}\mathbf{\Gamma} + \mathbf{\Gamma}^T\mathbf{F}^T + \mathbf{\Sigma} \tag{105}$$

but may be ignored thereafter. An experimental determination of the priors thus requires estimating the measurement covariance $\mathbf{B}$ and the covariance matrix $\mathbf{A}_{11}$ of the interesting sources only.

4.7 Estimation Subject to a Linear Constraint

The inverse problem can be generalized by requiring in addition that the solution $\hat{\mathbf{q}}$ satisfies a linear equality constraint of the form $\mathbf{C}\hat{\mathbf{q}} = 0$.

For example, charge conservation under quasistatic conditions requires that $\nabla^2 \vec{J} = 0$; a finite difference discretization leads to a linear constraint of the form $\mathbf{DJ} = 0$ where $\mathbf{J}$ is a vector of current density values at the selected grid points. Although this approach to imposing charge conservation appears inferior to finite element or boundary element approaches and was not pursued further, it may be useful for other applications.

For $\mathbf{C}\hat{\mathbf{q}} = \mathbf{CHb} = 0$ for every $\mathbf{b}$ consistent with the priors, it is necessary and sufficient under Assumption A that $\mathbf{CH} = 0$. Sufficiency is obvious. For necessity, suppose that $\mathbf{CH} \neq 0$. Then there exists a $\mathbf{b}$ such that $\mathbf{CHb} \neq 0$. But, under Assumption A, the measurement covariance $\mathbf{B} = \mathsf{E}\,\mathbf{b}\mathbf{b}^T$ is nonsingular and it follows that $\mathbf{b}$ is not confined to any subspace. That is, there is always a $\mathbf{b}$ consistent with the priors such that $\mathbf{CHb} \neq 0$.

Then the method of Lagrange multipliers yields the augmented objective function

$$S = \mathsf{E}\,\|\hat{\mathbf{q}} - \mathbf{q}\|^2 + \sum_k \lambda_k^T \mathbf{C}\mathbf{H}\mathbf{e}_k \tag{106}$$

where each λ_k is a column vector of Lagrange multipliers and $\mathbf{e}_k$ is the kth standard basis vector. Then, since $\hat{\mathbf{q}} = \mathbf{Hb} = \mathbf{HFq} + \mathbf{Hw}$,

$$S = \mathsf{E}(\mathbf{HFq} + \mathbf{Hw} - \mathbf{q})^T(\mathbf{HFq} + \mathbf{Hw} - \mathbf{q}) + \sum_k \lambda_k^T \mathbf{C}\mathbf{H}\mathbf{e}_k \quad . \tag{107}$$

The first variation is

$$\begin{aligned} \delta S &= 2\,\mathsf{E}(\mathbf{HFq} + \mathbf{Hw} - \mathbf{q})^T(\delta\mathbf{HFq} + \delta\mathbf{Hw}) + \sum_k \lambda_k^T \mathbf{C}\delta\mathbf{H}\mathbf{e}_k \\ &= 2\,\mathsf{E}(\mathbf{HFq} + \mathbf{Hw} - \mathbf{q})^T \delta\mathbf{H}(\mathbf{Fq} + \mathbf{w}) + \sum_k \lambda_k^T \mathbf{C}\delta\mathbf{H}\mathbf{e}_k \end{aligned} \tag{108}$$

which must equal zero for arbitrary $\delta\mathbf{H}$ when $\mathbf{H}$ is optimal. This implies that

$$\mathsf{E}(\mathbf{Fq} + \mathbf{w})(\mathbf{HFq} + \mathbf{Hw} - \mathbf{q})^T + \frac{1}{2}\sum_k \mathbf{e}_k \lambda_k^T \mathbf{C} = 0 \quad . \tag{109}$$

Assuming for convenience that $\mathbf{q}$ and $\mathbf{w}$ are uncorrelated and absorbing the constant factor $1/2$ into the λ_k yields

$$\mathbf{FAF}^T\mathbf{H}^T - \mathbf{FA} + \mathbf{\Sigma H}^T + \sum_k \mathbf{e}_k \lambda_k^T \mathbf{C} = 0 \quad , \tag{110}$$

or, after transposing and extracting common terms,

$$\mathbf{H}(\mathbf{FAF}^T + \mathbf{\Sigma}) - \mathbf{AF}^T + \mathbf{C}^T\mathbf{\Lambda} = 0 \quad , \tag{111}$$

where

$$\mathbf{\Lambda} = \sum_k \lambda_k \mathbf{e}_k^T \tag{112}$$

is a matrix of Lagrange multipliers. Solving for $\mathbf{H}$ yields

$$\mathbf{H} = (\mathbf{AF}^T - \mathbf{C}^T\mathbf{\Lambda})(\mathbf{FAF}^T + \mathbf{\Sigma})^{-1} \quad . \tag{113}$$

Now the desired constraint is that $\mathbf{CH} = 0$, or

$$\mathbf{C}(\mathbf{AF}^T - \mathbf{C}^T\mathbf{\Lambda})(\mathbf{FAF}^T + \mathbf{\Sigma})^{-1} = 0 \quad . \tag{114}$$

Since $(\mathbf{FAF}^T + \mathbf{\Sigma})$ is non-singular, this is equivalent to

$$\mathbf{C}(\mathbf{AF}^T - \mathbf{C}^T\mathbf{\Lambda}) = 0 \quad . \tag{115}$$

Now assume without loss of generality that the rows of $\mathbf{C}$ are linearly independent. Otherwise $\mathbf{C}\hat{\mathbf{q}} = 0$ contains a redundant constraint that could be removed without changing the set of admissable $\hat{\mathbf{q}}$. Then it follows [94, pp. 157–158] that $\mathbf{CC}^T$ is nonsingular and so

$$\mathbf{\Lambda} = (\mathbf{CC}^T)^{-1}\mathbf{CAF}^T \quad . \tag{116}$$

Finally, substituting into the expression for $\mathbf{H}$ yields

$$\mathbf{H} = (\mathbf{I} - \mathbf{C}^T(\mathbf{CC}^T)^{-1}\mathbf{C})\mathbf{AF}^T(\mathbf{FAF}^T + \mathbf{\Sigma})^{-1} \quad . \tag{117}$$

Note that $(\mathbf{I} - \mathbf{C}^T(\mathbf{CC}^T)^{-1}\mathbf{C})$ is the projection matrix onto the null space of $\mathbf{C}$ and so the optimal solution subject to the constraint $\mathbf{C}\hat{\mathbf{q}} = 0$ can be obtained simply by projecting the unconstrained solution into the null space of the constraint matrix $\mathbf{C}$.

Chapter 5

The Generalized Chi-squared Distribution

This chapter defines and develops a generalization of the χ^2 distribution which will be necessary later to describe the statistics of various measures of reconstruction quality. The generalized χ^2 distribution is the distribution followed by the sum of the squares of Gaussian random variables with arbitrary means, variances, and covariances. Alternatively, it is the distribution followed by the squared norm of an arbitrary Gaussian random vector.

The generalized χ^2 distribution is important to this dissertation for the following reason: The measurement residual $\mathbf{r} = \mathbf{b} - \mathbf{F}\hat{\mathbf{q}}$, or the difference between the actual measurements and the measurements predicted by the reconstruction, gives an indication of the completeness of the source model used in the reconstruction. The squared residual $\chi^2 = ||\mathbf{r}||^2$ reduces the residual to a single number which, if too large, indicates that the source model is possibly incomplete. It is important to know the distribution of $||\mathbf{r}||^2$ to decide what is meant by "too large." If reconstruction is done using a least-squares method, then the squared residual $||\mathbf{r}||^2$ follows the standard χ^2 distribution. If, however, OCLIM is used for the reconstruction, then the squared residual follows a generalization of the χ^2 distribution. It is the purpose of this chapter to define that generalization and to develop an algorithm for computing it. The algorithm developed is applicable to any continuous random variable meeting certain regularity conditions and may be useful in its own right.

5.1 Some Mathematical Preliminaries

It is useful to collect here some elementary results about random vectors, characteristic functions, and Fourier transforms which will be used later in this chapter.

Proposition: Let $p(x)$ be the probability density of a real-valued random variable and $\phi(t) = \mathsf{E}\{e^{itx}\}$ its characteristic function. Let

$$\mathcal{F} : f(x) \mapsto F(u) = \int_{-\infty}^{\infty} e^{-i2\pi ux} f(x)\, dx \tag{118}$$

denote the Fourier transform. Then

$$\mathcal{F}\{p(x)\} = \phi(-2\pi u) \qquad \text{and} \qquad p(x) = \mathcal{F}^{-1}\{\phi(-2\pi u)\} \quad . \tag{119}$$

Proof: Substitute $t = -2\pi u$ into the definition of the characteristic function. Then take the inverse Fourier transform. ■

Proposition: Let $p(x)$ be the probability density of a real-valued random variable and $\phi(t) = \mathsf{E}\{e^{itx}\}$ its characteristic function. Then

$$x^n\, p(x) = \frac{i^{-n}}{2\pi} \int_{-\infty}^{\infty} \phi^{(n)}(t)\, e^{-itx}\, dt \tag{120}$$

and

$$|x^n\, p(x)| \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} |\phi^{(n)}(t)|\, dt \quad . \tag{121}$$

Proof: Differentiating the definition of $\phi(t)$ n times yields

$$\phi^{(n)}(t) = i^n \int_{-\infty}^{\infty} x^n\, p(x)\, e^{itx}\, dx \quad ,$$

from which the Fourier integral theorem yields the first result. The second result then follows immediately by taking the absolute value. ■

Definition: The Frobenius norm of a general matrix $\mathbf{G}$ is defined as

$$\|\mathbf{G}\|_F^2 = \sum_{ij} g_{ij}^2 \quad . \tag{122}$$

Given this definition, it is trivial to show the useful fact that $\|\mathbf{G}\|_F^2 = \mathrm{Tr}(\mathbf{G}\mathbf{G}^T)$.

Proposition: Let $\mathbf{x}$ be a random vector with mean $\bar{\mathbf{x}}$ and variance $\mathbf{C_x}$ and let $\mathbf{A}$ be a symmetric positive definite matrix with Cholesky decomposition $\mathbf{R}\mathbf{R}^T$. Then

$$\mathsf{E}(\mathbf{x}^T\mathbf{A}\mathbf{x}) = \mathrm{Tr}(\mathbf{R}^T\mathbf{C_x}\mathbf{R}) + \bar{\mathbf{x}}^T\mathbf{A}\bar{\mathbf{x}} \tag{123}$$

and, in particular,

$$\mathsf{E}\,\|\mathbf{x}\|^2 = \mathrm{Tr}(\mathbf{C_x}) + \|\bar{\mathbf{x}}\|^2 \quad . \tag{124}$$

Proof: Consider first the case that $\bar{\mathbf{x}} = 0$. Then $\mathsf{E}\,\mathbf{x}^T\mathbf{x} = \sum_j \mathsf{E}\,x_j^2 = \mathrm{Tr}(\mathbf{C_x})$. Then consider

$$\begin{aligned}\mathsf{E}\,\mathbf{x}^T\mathbf{A}\mathbf{x} &= \mathsf{E}(\mathbf{x}-\bar{\mathbf{x}}+\bar{\mathbf{x}})^T\mathbf{R}\mathbf{R}^T(\mathbf{x}-\bar{\mathbf{x}}+\bar{\mathbf{x}}) \\ &= \mathsf{E}(\mathbf{x}-\bar{\mathbf{x}})^T\mathbf{R}\mathbf{R}^T(\mathbf{x}-\bar{\mathbf{x}}) + \mathsf{E}(\mathbf{x}-\bar{\mathbf{x}})^T\mathbf{R}\mathbf{R}^T\bar{\mathbf{x}} \\ &\quad + \mathsf{E}\,\bar{\mathbf{x}}^T\mathbf{R}\mathbf{R}^T(\mathbf{x}-\bar{\mathbf{x}}) + \mathsf{E}\,\bar{\mathbf{x}}^T\mathbf{R}\mathbf{R}^T\bar{\mathbf{x}} \quad .\end{aligned}$$

Now observe that the random vector $\mathbf{R}^T(\mathbf{x}-\bar{\mathbf{x}})$ has mean zero and variance $\mathbf{R}^T\mathbf{C_x}\mathbf{R}$; then the first equality follows immediately. The second equality follows from the substitution $\mathbf{A} = \mathbf{I}$. ∎

5.2 Definition

We now consider the problem of computing the distribution of the norm of a Gaussian random vector.

This distribution can be viewed as a generalization of the well-known χ^2 distribution. A random variable Y is said to have a χ^2 distribution with ν degrees of freedom if it can be written as

$$Y = \sum_{j=1}^{\nu} x_j^2 \tag{125}$$

where the x_j are independent, identically distributed Gaussian random variables with mean zero and unit variance. In the generalized χ^2 distribution, the x_j may have arbitrary means, variances, and covariances.

More concisely, suppose that $\mathbf{x}$ is a Gaussian random vector with mean $\bar{\mathbf{x}}$ and covariance $\mathbf{C_x}$. Then the squared norm $\|\mathbf{x}\|^2 = \sum_j x_j^2$ will be said to have a generalized χ^2 distribution with parameters $\bar{\mathbf{x}}$ and $\mathbf{C_x}$. In symbols, $\|\mathbf{x}\|^2 \sim \chi^2(\bar{\mathbf{x}}, \mathbf{C_x})$.

Some previously known distributions are special cases of the generalized χ^2 distribution and will be discussed later.

5.2.1 Canonical Form

We begin with a narrower canonical form of the generalized χ^2 distribution and show that it defines exactly the same class of distributions as the more general form above.

Definition: Suppose that $z_j, j = 1, \ldots, \nu$ are independent normal random variables with respective means μ_j and variances $\sigma_j^2 > 0$. Let μ_0 be a real constant. Then the random variable

$$Y = \mu_0^2 + \sum_{j=1}^{\nu} z_j^2 \tag{126}$$

has a generalized χ^2 distribution. (Note that the requirement that $\sigma_j \neq 0$ involves no loss of generality; any such terms can be gathered into the constant term μ_0^2.)

The central and non-central χ^2 distributions with ν degrees of freedom and noncentrality parameter μ^2 are obtained as special cases by setting $\mu_0 = 0$, $\mu_1 = \mu$, $\mu_2 = \mu_3 = \cdots = \mu_\nu = 0$ and $\sigma_j = 1$. The Rayleigh distribution with ν degrees of freedom and variance σ^2 is obtained by setting $\mu_0 = \mu_j = 0$ and $\sigma_j = \sigma$. The computational algorithm described here will be applicable to these special cases but may or may not be more efficient than algorithms specific to the Rayleigh or non-central χ^2 distributions [20, 26, 73, 74, 82].

5.2.2 Characteristic Function

There does not appear to be an explicit expression for the probability density function for the generalized χ^2 distribution. However, the characteristic function can be found and provides a method for numerically evaluating the probability density.

We begin by developing the characteristic function for a single term ($\nu = 1$). Let X be a normal random variable with mean μ and variance σ^2 and define $Y = X^2$. Then the cumulative distribution function for Y is

$$\begin{aligned}
\text{Prob}(Y \leq y) &= \text{Prob}(X^2 \leq y) = \text{Prob}(-\sqrt{y} \leq X \leq \sqrt{y}) \\
&= \text{Prob}\left(\frac{-\sqrt{y} - \mu}{\sigma} \leq \frac{X - \mu}{\sigma} \leq \frac{\sqrt{y} - \mu}{\sigma}\right) \\
&= \frac{1}{\sqrt{2\pi}} \int_{\frac{-\sqrt{y}-\mu}{\sigma}}^{\frac{\sqrt{y}-\mu}{\sigma}} e^{-z^2/2}\, dz \quad .
\end{aligned}$$

Differentiating by Leibnitz' rule yields the density function

$$p(y) = \frac{1}{2\sigma\sqrt{2\pi y}}\left[\exp\left(-\frac{(\sqrt{y}-\mu)^2}{2\sigma^2}\right) + \exp\left(-\frac{(\sqrt{y}+\mu)^2}{2\sigma^2}\right)\right] \quad .$$

The characteristic function is

$$\begin{aligned}\psi(t) &= \mathsf{E}\{e^{itY}\} \\ &= \int_0^\infty \frac{e^{ity}}{2\sigma\sqrt{2\pi y}}\left[\exp\left(-\frac{(\sqrt{y}-\mu)^2}{2\sigma^2}\right) + \exp\left(-\frac{(\sqrt{y}+\mu)^2}{2\sigma^2}\right)\right] dy \quad .\end{aligned}$$

Substituting $y = x^2$ and simplifying,

$$\begin{aligned}&= \frac{1}{\sigma\sqrt{2\pi}}\int_0^\infty \exp\left(-\frac{(1-2i\sigma^2t)x^2 - 2\mu x + \mu^2}{2\sigma^2}\right) dx \\ &+ \frac{1}{\sigma\sqrt{2\pi}}\int_0^\infty \exp\left(-\frac{(1-2i\sigma^2t)x^2 + 2\mu x + \mu^2}{2\sigma^2}\right) dx\end{aligned}$$

and applying formula 7.4.2 of Abramowitz and Stegun [2],

$$\begin{aligned}&= \frac{1}{2\sigma\sqrt{2\pi}}\sqrt{\frac{2\pi\sigma^2}{1-i2\sigma^2t}}\exp\left(\frac{i\mu^2t}{1-i2\sigma^2t}\right)\times \\ &\quad\left[\operatorname{erfc}\left(\frac{-\mu}{\sqrt{1-i2\sigma^2t}}\right) + \operatorname{erfc}\left(\frac{\mu}{\sqrt{1-i2\sigma^2t}}\right)\right]\end{aligned}$$

and since

$$\begin{aligned}\operatorname{erfc}(x) + \operatorname{erfc}(-x) &= 1 - \operatorname{erf}(x) + 1 - \operatorname{erf}(-x) \\ &= 2 - \operatorname{erf}(x) + \operatorname{erf}(x) \\ &= 2 \quad ,\end{aligned}$$

we obtain the result

$$\psi(t) = (1-i2\sigma^2t)^{-1/2}\exp\left(\frac{i\mu^2t}{1-i2\sigma^2t}\right) \quad . \tag{127}$$

The moments of X and Y are conveniently computed via the characteristic function of X,

$$\phi_X(t) = \exp(i\mu t - \tfrac{1}{2}\sigma^2t^2) \quad ,$$

yielding

$$\mathsf{E}X = \phi_X'(0)/i = \mu \quad , \tag{128a}$$

$$\mathsf{E}Y = \mathsf{E}X^2 = \phi_X''(0)/i^2 = \sigma^2 + \mu^2 \quad , \tag{128b}$$

$$\mathsf{E}X^3 = \phi_X'''(0)/i^3 = \mu(3\sigma^2 + \mu^2) \quad \text{, and} \tag{128c}$$

$$\mathsf{E}Y^2 = \mathsf{E}X^4 = \phi_X^{\mathrm{iv}}(0)/i^4 = 3\sigma^4 + 6\sigma^2\mu^2 + \mu^4 \quad . \tag{128d}$$

Furthermore, the variance of Y is

$$\begin{aligned}\mathrm{Var}(Y) &= \mathsf{E}(Y - \mathsf{E}\,Y)^2 = \mathsf{E}\,Y^2 - (\mathsf{E}\,Y)^2 = \mathsf{E}\,X^4 - (\mathsf{E}\,X^2)^2 \\ &= 2\sigma^2(\sigma^2 + 2\mu^2) \quad . \end{aligned} \tag{129}$$

Then for the general case of $X_j, j = 1, \ldots, \nu$, independent and normally distributed with means μ_j and variances σ_j^2, the random variable

$$Y = \mu_0^2 + \sum_{j=1}^{\nu} X_j^2 \tag{130}$$

must have the characteristic function

$$\phi(t) = e^{i\mu_0^2 t} \prod_{j=1}^{\nu} (1 - i2\sigma_j^2 t)^{-1/2} \exp\left(\frac{i\mu_j^2 t}{1 - i2\sigma_j^2 t}\right) \quad , \tag{131}$$

the mean

$$\mathsf{E}\,Y = \mu_0^2 + \sum_{j=1}^{\nu} (\sigma_j^2 + \mu_j^2) \quad , \tag{132}$$

and the variance

$$\mathrm{Var}\,Y = \sum_{j=1}^{\nu} 2\sigma_j^2(\sigma_j^2 + 2\mu_j^2) \quad . \tag{133}$$

Note also that the mimimum possible value of Y is μ_0^2. If $\nu > 0$, there is no upper bound.

5.2.3 General Form

The distribution of $\|\mathbf{y}\|^2$ for a Gaussian random vector $\mathbf{y}$ can be expressed in terms of the canonical χ^2 distribution developed above. Suppose that $\mathbf{y}$ is a Gaussian random vector with mean $\bar{\mathbf{y}}$ and covariance matrix $\mathbf{C_{yy}}$. Then the squared vector magnitude $\|\mathbf{y}\|^2 = \mathbf{y}^T\mathbf{y}$ has a generalized χ^2 distribution. To see this, observe that the covariance matrix $\mathbf{C_{yy}}$ is positive semidefinite and can be written in the form $\mathbf{C_{yy}} = \mathbf{UDU}^T$ where $\mathbf{U}$ is orthogonal and $\mathbf{D}$ is diagonal with non-negative elements. Define $\mathbf{z} = \mathbf{U}^T\mathbf{y}$. Then the mean of $\mathbf{z}$ is $\bar{\mathbf{z}} = \mathsf{E}\,\mathbf{z} = \mathbf{U}^T\bar{\mathbf{y}}$ and the variance is

$$\mathsf{E}(\mathbf{z} - \bar{\mathbf{z}})(\mathbf{z} - \bar{\mathbf{z}})^T = \mathsf{E}\,\mathbf{U}^T(\mathbf{y} - \bar{\mathbf{y}})(\mathbf{y} - \bar{\mathbf{y}})^T\mathbf{U} = \mathbf{U}^T\mathbf{C_{yy}}\mathbf{U} = \mathbf{D} \quad .$$

That is, the components z_j of $\mathbf{z}$ are independent Gaussian random variables with means $\mu_j = [\mathbf{U}^T\bar{\mathbf{y}}]_j$ and variances $\sigma_j^2 = \mathbf{D}_{jj} = [\mathbf{U}^T\mathbf{C_{yy}}\mathbf{U}]_{jj}$. (Remember that if any variances σ_j^2 are zero, the corresponding μ_j must be lumped into the constant μ_0^2 of the canonical χ^2 distribution.) Then $\|\mathbf{z}\|^2 = \sum_j z_j^2$ has a generalized χ^2 distribution. Furthermore, since $\mathbf{U}$ is orthogonal, $\|\mathbf{y}\|^2 = \|\mathbf{U}\mathbf{z}\|^2 = \|\mathbf{z}\|^2$ has the same distribution.

The mean and variance of $\|\mathbf{y}\|^2$ can also be written in matrix form as

$$\mathsf{E}\,\|\mathbf{y}\|^2 = \mathrm{Tr}(\mathbf{C_{yy}}) + \|\bar{\mathbf{y}}\|^2 \tag{134}$$

and

$$\mathrm{Var}\,\|\mathbf{y}\|^2 = 2\,\mathrm{Tr}(\mathbf{C_{yy}^2}) + 4\bar{\mathbf{y}}^T\mathbf{C_{yy}}\bar{\mathbf{y}} \quad . \tag{135}$$

The expression for $\mathsf{E}\,\|\mathbf{y}\|^2$ holds for any random vector $\mathbf{y}$. To see that the expression for $\mathrm{Var}\,\|\mathbf{y}\|^2$ is valid, consider

$$\begin{aligned} 2\,\mathrm{Tr}(\mathbf{C_{yy}^2}) + 4\bar{\mathbf{y}}^T\mathbf{C_{yy}}\bar{\mathbf{y}} &= 2\,\mathrm{Tr}(\mathbf{U}\mathbf{D}\mathbf{U}^T\mathbf{U}\mathbf{D}\mathbf{U}^T) + 4\bar{\mathbf{y}}\mathbf{U}\mathbf{D}\mathbf{U}^T\bar{\mathbf{y}} \\ &= 2\,\mathrm{Tr}(\mathbf{D}^2) + 4\bar{\mathbf{z}}^T\mathbf{D}\bar{\mathbf{z}} \\ &= 2\sum_j \sigma_j^4 + \sum_j 4\mu_j^2\sigma_j^2 \\ &= \mathrm{Var}\,\|\mathbf{z}\|^2 = \mathrm{Var}\,\|\mathbf{y}\|^2 \quad . \end{aligned}$$

Now suppose that $\mathbf{A}$ is a symmetric positive definite matrix with Cholesky decomposition $\mathbf{A} = \mathbf{R}\mathbf{R}^T$ and consider the distribution of $\mathbf{y}^T\mathbf{A}\mathbf{y} = \|\mathbf{R}^T\mathbf{y}\|^2$. Since $\mathbf{R}^T\mathbf{y}$ is a linear combination of Gaussian random variables, it is Gaussian and must have mean $\mathbf{R}^T\bar{\mathbf{y}}$ and variance $\mathbf{R}^T\mathbf{C_{yy}}\mathbf{R}$. It follows that $\|\mathbf{y}^T\mathbf{A}\mathbf{y}\|^2$ must have a generalized χ^2 distribution, or

$$\mathbf{y}^T\mathbf{A}\mathbf{y} \sim \chi^2(\mathbf{R}^T\bar{\mathbf{y}}, \mathbf{R}^T\mathbf{C_{yy}}\mathbf{R}) \quad . \tag{136}$$

5.2.4 Asymptotic Behavior

We prove here the asymptotic properties of the characteristic function $\phi(t)$ and of the density $p(x)$ which are essential to proving the convergence of the numerical algorithm for $p(x)$ proposed later. For convenience, only the case $\nu > 4$ is considered.

Lemma: For all real values of t,

$$-\frac{\mu^2}{\sigma^2} \leq \Re\left(\frac{i\mu^2 t}{1-i2\sigma^2 t}\right) \leq 0 \quad , \tag{137}$$

from which it follows that

$$e^{-\mu^2/\sigma^2} \leq \left|\exp\left(\frac{i\mu^2 t}{1-i2\sigma^2 t}\right)\right| \leq 1 \quad . \tag{138}$$

In both inequalities, the upper limit is obtained at $t = 0$ and the lower limit at $t = \pm\infty$.

Proof: Consider

$$\begin{aligned}\Re\left(\frac{i\mu^2 t}{1-i2\sigma^2 t}\right) &= \Re\left(\frac{(i\mu^2 t)(1+i2\sigma^2 t)}{(1-i2\sigma^2 t)(1+i2\sigma^2 t)}\right) \\ &= \Re\left(\frac{i\mu^2 t - 2\mu^2\sigma^2 t^2}{1+4\sigma^4 t^2}\right) \\ &= \frac{-2\mu^2\sigma^2 t^2}{1+4\sigma^4 t^2} \quad .\end{aligned}$$

The derivative of this last quantity is $-4\sigma^2\mu^2 t/(1+4\sigma^2\mu^2 t^2)$, which is negative for t positive, zero for $t = 0$, and positive for t negative. It follows that the maximum must occur at $t = 0$ and minima at $t = \pm\infty$. Substitution then yields the limiting values given above. ■

Lemma: For all real values of t,

$$|1-i2\sigma^2 t|^{-1/2} \leq |2\sigma^2 t|^{-1/2} \quad . \tag{139}$$

Proof: $|1-i2\sigma^2 t|^{-1/2} = |1+4\sigma^4 t^2|^{-1/4} \leq |4\sigma^4 t^2|^{-1/4} = |2\sigma^2 t|^{-1/2}$. ■

Proposition: Let $\phi(t)$ be the characteristic function of a generalized χ^2 variable as given above. Then, for all real values of t,

$$|\phi(t)| \leq \left(2^{-\nu/2}\prod_{j=1}^{\nu}\sigma_j^{-1}\right)|t|^{-\nu/2} \quad . \tag{140}$$

Proof: Consider

$$\begin{aligned}|\phi(t)| &= \left|e^{i\mu_0^2 t}\right| \cdot \prod_{j=1}^{\nu}|1-i2\sigma_j^2 t|^{-1/2} \cdot \left|\exp\left(\frac{i\mu_j^2 t}{1-i2\sigma_j^2 t}\right)\right| \\ &\leq (1)\cdot\prod_{j=1}^{\nu}|2\sigma_j^2 t|^{-1/2}\cdot(1) \\ &= \left(2^{-\nu/2}\prod_{j=1}^{\nu}\sigma_j^{-1}\right)|t|^{-\nu/2} \quad .\end{aligned}$$

■

Proposition: The probability density function $p(x)$ for a generalized χ^2 variable exists and is differentiable everywhere for $\nu > 4$.

Proof: (1) For the existence of $p(x)$ it suffices to show than $\phi(t)$ is continuous on the real axis and that it is absolutely integrable. Continuity is obvious. For absolute integrability consider that

$$\int_{-\infty}^{\infty} |\phi(t)|\, dt \leq \int_{-\infty}^{\infty} B\, |t|^{-\nu/2}\, dt \quad ,$$

where $B = 2^{-\nu/2} \prod_j \sigma_j^{-1}$, and that the latter integral exists and is finite whenever $\nu > 2$.

(2) For differentiability of $p(x)$ consider its derivative

$$\begin{aligned} |p'(x)| &\leq \frac{1}{2\pi} \int_{-\infty}^{\infty} |t\, \phi(t)|\, dt \\ &\leq \int_{-\infty}^{\infty} |t|\, B\, |t|^{-\nu/2} dt \\ &= \int_{-\infty}^{\infty} B\, |t|^{1-\nu/2} dt \quad , \end{aligned}$$

which converges whenever $\nu > 4$. ∎

Remark: The case $\nu \leq 4$ should be accessible to a more sophisticated analysis. It appears that $p(x)$ exists for all $\nu > 0$ and is continuous for $\nu > 2$.

To determine the asymptotic behavior of the density $p(x)$ it is convenient to start with the simple case $\nu = 1$, for which the characteristic function is

$$\psi(t) = (1 - i2\sigma^2 t)^{-1/2} \exp\left(\frac{i\mu^2 t}{1 - i2\sigma^2 t}\right) \quad . \tag{141}$$

For notational convenience, define the polynomials

$$f(t) = 1 - i2\sigma^2 t \qquad \text{and} \qquad g(t) = i\mu^2 t \quad . \tag{142}$$

Proposition: The nth derivative of $\psi(t)$ can be written in the form

$$\psi^{(n)}(t) = q_n(g(t)/f(t))\, (f(t))^{-n-1/2}\, e^{g(t)/f(t)} \tag{143}$$

where $q_n(s)$ is the polynomial of degree n or less defined by the recursion

$$\begin{aligned} q_0(s) &= 1 \qquad \text{and} \\ q_{n+1}(s) &= q_n'(s)(i\mu^2 + i2\sigma^2 s) + q_n(s)(i\mu^2 - (n + \tfrac{1}{2}) + i2\sigma^2 s) \quad . \end{aligned} \tag{144}$$

Furthermore, suppose that $q_n(s) = \sum_{j=0}^{n} q_{nj}\, s^j$. Then

$$|\psi^{(n)}(t)| \le Q_n\, |1 - i2\sigma^2 t|^{-n-1/2} \le Q_n\, |2\sigma^2 t|^{-n-1/2} \quad , \tag{145}$$

where

$$Q_n = \sum_{j=0}^{n} |q_{nj}| \left(\frac{\mu^2}{2\sigma^2}\right)^j \quad . \tag{146}$$

Proof: (1) We consider first the claim that $\psi^{(n)}(t)$ can be written in the given form. The formula $q_0 = 1$ follows immediately from the equation $\psi^{(0)}(t) = \psi(t)$. For the general case we proceed by induction. Assume that

$$\psi^{(n)}(t) = q_n(g/f)\, f^{-n-1/2}\, e^{g/f} \quad .$$

Then differentiation yields

$$\begin{aligned} \psi^{(n+1)} &= \left[q_n'(i\mu^2 + i2\sigma^2(g/f) + q_n(i\mu^2 - (n+\tfrac{1}{2}) + i2\sigma^2(g/f))\right]\, f^{-n-3/2}\, e^{g/f} \\ &= q_{n+1}(g/f)\, f^{-(n+1)-1/2}\, e^{g/f} \quad , \end{aligned}$$

as required, given the definition

$$q_{n+1}(s) = q_n'(i\mu^2 + i2\sigma^2 s + q_n(i\mu^2 - (n+\tfrac{1}{2}) + i2\sigma^2 s) \quad .$$

(2) For the first bound, observe that

$$\left|\frac{g(t)}{f(t)}\right| = \left|\frac{i\mu^2 t}{1 - i2\sigma^2 t}\right| = \left(\frac{\mu^4 t^2}{1 + 4\sigma^4 t^2}\right)^{1/2} \le \frac{\mu^2}{2\sigma^2} \quad .$$

Since $|g/f|$ is bounded and $q_n(s)$ is a polynomial, the polynomial $q_n(g/f)$ must also be bounded, with the bound Q_n as defined above. Now recall that the factor $|e^{g/f}|$ is bounded by 1. Then the bound on $\psi^{(n)}$ follows immediately. The second form of the bound follows from the previous lemma that $|1 - i2\sigma^2 t|^{-1/2} \le |2\sigma^2 t|^{-1/2}$. ■

The first few values for q_n and Q_n can be easily computed and are:

$$q_0(s) = 1 \tag{147a}$$

$$q_1(s) = i\mu^2 - \tfrac{1}{2} + i2\sigma^2 s \tag{147b}$$

$$q_2(s) = \tfrac{3}{4} - 2\sigma^2\mu^2 - \mu^4 - i2\mu^2 + s(-4\sigma^4 - 2\sigma^2\mu^2 - i\sigma^2) - s^2\, 4\sigma^4 \tag{147c}$$

$$Q_0(\mu, \sigma) = 1 \tag{148a}$$

$$Q_1(\mu, \sigma) = (\tfrac{1}{4} + \mu^4)^{1/2} + \mu^4 \tag{148b}$$

$$Q_2(\mu, \sigma) = \left[(\tfrac{3}{4} - 2\sigma^2\mu^2 - \mu^2)^2 + 4\mu^4\right]^{1/2} + \left[(2\sigma^2 + \mu^2)^2 + \tfrac{1}{4}\right]^{1/4}\, \mu^2 + \mu^4 \tag{148c}$$

Proposition: Let $\phi(t)$ be the characteristic function of a generalized χ^2 random variable. Then

$$|\phi''(t)| \leq \left(\nu Q_{2\max} + \nu(\nu-1)Q_{1\max}^2\right) \left(1 + 4\sigma_{\min}^4 t^2\right)^{-1-\nu/4} \tag{149}$$

where

$$Q_{2\max} = \max_k Q_2(\mu_k, \sigma_k) \quad , \tag{150a}$$

$$Q_{1\max} = \max_k Q_1(\mu_k, \sigma_k) \quad , \text{ and} \tag{150b}$$

$$\sigma_{\min} = \min_k \sigma_k \quad . \tag{150c}$$

Proof: The characteristic function can be written in the form

$$\phi(t) = \prod_{k=1}^{\nu} \psi_k(t) \quad ,$$

where

$$\psi_k(t) = (1 - i2\sigma_k^2 t)^{-1/2} \exp\left(\frac{i\mu_k^2 t}{1 - i2\sigma_k^2 t}\right) \quad .$$

Differentiating twice yields

$$\phi'' = \sum_k \psi_k'' \prod_\ell \psi_\ell + \sum_k \sum_\ell \psi_k' \psi_\ell' \prod_m \psi_m \quad ,$$

where k ranges over $1, 2, \ldots, \nu$, ℓ covers the same range but omits k, and m covers the same range but omits both k and ℓ. Furthermore,

$$\begin{aligned} |\phi''| &\leq \sum_k |\psi_k''| \prod_\ell |\psi_\ell| + \sum_k \sum_\ell |\psi_k' \psi_\ell'| \prod_m |\psi_m| \\ &\leq \sum_k Q_{2k} |f_k|^{-5/2} \prod_\ell |f_\ell|^{-1/2} \\ &\quad + \sum_k \sum_\ell Q_{1k} Q_{1\ell} |f_k|^{-3/2} |f_\ell|^{-3/2} \prod_m |f_m|^{-1/2} \quad , \end{aligned}$$

where $Q_{nk} = Q_n(\mu_k, \sigma_k)$ and $f_k = (1 - i2\sigma_k^2 t)$. Now observe that for $\alpha < 0$,

$$|f_k|^{-\alpha} = |1 - i2\sigma_k^2 t|^{-\alpha} \leq |1 - i2\sigma_{\min}^2 t|^{-\alpha} \quad .$$

Then it follows that

$$\begin{aligned} |\phi''| &\leq \nu Q_{2\max} |1 - i2\sigma_{\min}^2 t|^{-2-\nu/2} + \nu(\nu-1) Q_{1\max}^2 |1 - i2\sigma_{\min}^2 t|^{-2-\nu/2} \\ &= \left(\nu Q_{2\max} + \nu(\nu-1) Q_{1\max}^2\right) |1 - i2\sigma_{\min}^2 t|^{-2-\nu/2} \\ &= \left(\nu Q_{2\max} + \nu(\nu-1) Q_{1\max}^2\right) (1 + 4\sigma_{\min}^4 t^2)^{-1-\nu/4} \quad . \end{aligned}$$

■

Proposition: Let $p(x)$ be the probability density for a generalized χ^2 variable. Then for all values of x,

$$0 \le p(x) \le Ax^{-2} \tag{151}$$

where the real constant A has the value

$$A = \frac{\nu Q_{2\max} + \nu(\nu-1)Q_{1\max}^2}{4\sqrt{\pi}\sigma_{\min}^2} \cdot \frac{\Gamma(\nu/4+\frac{1}{2})}{\Gamma(\nu/4+1)} \quad . \tag{152}$$

Proof: Since $p(x)$ is a probability density it is necessarily non-negative. For the upper limit it suffices to show that $|x^2\, p(x)| \le A$. Now consider

$$\begin{aligned}
|x^2\, p(x)| &\le \frac{1}{2\pi}\int_{-\infty}^{\infty} |\phi''(t)|\, dt \\
&\le \frac{1}{2\pi}(\nu Q_{2\max} + \nu(\nu-1)Q_{1\max}^2)\int_{-\infty}^{\infty}(1+4\sigma_{\min}^4 t^2)^{-1-\nu/4}\, dt \\
&= \frac{1}{2\pi}(\nu Q_{2\max} + \nu(\nu-1)Q_{1\max}^2)\cdot\frac{\sqrt{\pi}}{2\sigma_{\min}^2}\cdot\frac{\Gamma(\nu/4+\frac{1}{2})}{\Gamma(\nu/4+1)} \\
&= \frac{\nu Q_{2\max} + \nu(\nu-1)Q_{1\max}^2}{4\sqrt{\pi}\sigma_{\min}^2}\cdot\frac{\Gamma(\nu/4+\frac{1}{2})}{\Gamma(\nu/4+1)} = A \quad ,
\end{aligned}$$

where the integral has been evaluated by formula 615 in Beyer [10]. ■

5.3 Numerical Algorithm

The preceeding section has developed an explicit form for the characteristic function of the generalized χ^2 distribution, but there does not appear to be any explicit form for the density itself. In this section we consider the problem of numerically computing that density using the fast Fourier transform (FFT).

The problem can be stated in more general form as follows: Supposing that we know an explicit form for the Fourier transform $F(u)$ of a real continuous-time function $f(x)$, how can $f(x)$ be computed?

Provided that both $F(u)$ and $f(x)$ are smooth and decay rapidly, it should be possible to to estimate $f(x)$ by evaluating $F(u)$ at some finite number of points, taking the inverse FFT to obtain $f(x)$ at some finite number of points, and interpolating the result to yield an estimate of $f(x)$. Since $F(u)$ is evaluated only at a finite number of discrete points, the estimated $f(x)$ will be distorted by sampling and truncation errors.

This approach to computing probability distributions is only sparsely represented in the literature. Bohman [11] briefly describes an FFT method for probability density functions but does not address the errors introduced by sampling and truncation. Belov and Galkin [9] describe a similar method for computing the compound Poisson distribution; they estimate the accuracy in the specific cases that they compute but do not address the general case. Maindonald [52] and Thisted [98] mention the FFT only in the context of time series analysis; they do not discuss any application to the computation of probability distribution functions. Several other authors have applied the FFT to the calculation of discrete distributions; see, for example, Abate and Whitt [1], Cavers [16], Daigle [24], and Grübel [39].

The use of the FFT to approximate the continuous Fourier transform is commonplace in the practice of digital signal processing but an error analysis is rarely done. The estimation of a probability density function is equivalent to the problem of estimating the spectrum of a non-time-limited signal given only a finite number of samples. Cooley et al. [21] describe the use of the FFT to compute Fourier integrals and analyze the error as a sum of aliasing (or sampling) and truncation errors; they do not, however, give explicit error bounds.

5.3.1 CFT/DFT Correspondence

We begin with the following theorem which gives conditions under which the continuous and discrete Fourier transforms (CFT and DFT) are equivalent. Cooley et al. [21, Theorem 1] give a different correspondence theorem.

Theorem (CFT/DFT Correspondence): Suppose that $F(u)$ is a spectrum which has been sampled and truncated so that it can be written in the form

$$F(u) = \sum_{k=-(N/2)+1}^{(N/2)-1} F(k/T)\,\delta(u - k/T) \quad . \tag{153}$$

Then the $N-1$ values $F(-(N/2)+1), \ldots, F((N/2)-1)$ suffice to define $F(u)$ and the corresponding signal (via the continuous Fourier transform) is

$$f(x) = \sum_{k=-(N/2)+1}^{(N/2)-1} F(k/T)\,e^{j2\pi kx/T} \quad . \tag{154}$$

Now let $\tau = T/N$ and define the discrete-time function

$$f[n] = f(n\tau) \qquad \text{for } n = 0, \ldots, N-1 \quad , \tag{155}$$

and the discrete-frequency spectrum

$$F[k] = \begin{cases} F(k/T) & k = 0, \ldots, (N/2)-1 \\ 0 & k = N/2 \\ F((k-N)/T) & k = (N/2)+1, \ldots, N-1 \end{cases} \quad . \tag{156}$$

Then $f[n]$ and $F[k]$ are related through the discrete Fourier transform

$$f[n] = \sum_{k=0}^{N-1} F[k]\, e^{j2\pi kn/N} \quad , \tag{157}$$

and $f(x)$ and $f[n]$ are related by the band-limited interpolation

$$f(x) = \sum_{n=0}^{N-1} f[n] \frac{\sin \pi(N-1)(x-n\tau)/T}{\sin \pi(x-n\tau)/T} \quad . \tag{158}$$

Proof: (1) To see the connection between $f[n]$ and $F[k]$, consider

$$\begin{aligned} f[n] = f(n\tau) &= \left[\sum_{k=-(N/2)+1}^{(N/2)-1} F(k/T)\, e^{j2\pi xk/T} \right]_{x=nT/N} \\ &= \sum_{k=-(N/2)+1}^{(N/2)-1} F(k/T)\, e^{j2\pi nk/N} \\ &= \sum_{k=0}^{N/2-1} F(k/T)\, e^{j2\pi nk/N} + \sum_{k=-(N/2)+1}^{-1} F(k/T)\, e^{j2\pi nk/N} \quad . \end{aligned}$$

Substituting $k' = k + N$ or $k = k' - N$ in the last summation yields, after using the fact that $e^{-j2\pi nN/N} = 1$ for all integer n,

$$f[n] = \sum_{k=0}^{N/2-1} F(k/T)\, e^{j2\pi nk/N} + \sum_{k'=(N/2)+1}^{N-1} F((k'-N)/T)\, e^{j2\pi nk'/N} \quad ,$$

and since $F[N/2]e^{j2\pi n(N/2)/N} = 0 \cdot e^{j2\pi n(N/2)/N} = 0$,

$$f[n] = \sum_{k=0}^{N-1} F[k]\, e^{j2\pi nk/N} \quad .$$

(2) To see the connection between $f(x)$ and $f[n]$, note first that $F[k]$ is given by the inverse discrete Fourier transform of $f[n]$, or

$$F[k] = \sum_{n=0}^{N-1} f[n]\, e^{-j2\pi kn/N} \quad .$$

Then consider

$$\begin{aligned}
f(x) &= \sum_{k=-(N/2)+1}^{(N/2)-1} F[k]\, e^{j2\pi kx/T} \\
&= \sum_{k=-(N/2)+1}^{(N/2)-1} \left[\sum_{n=0}^{N-1} f[n]\, e^{-j2\pi kn/N} \right] e^{j2\pi kx/T} \\
&= \sum_{n=0}^{N-1} f[n] \left[\sum_{k=-(N/2)+1}^{(N/2)-1} e^{-j2\pi kn/N}\, e^{j2\pi kx/T} \right] \\
&= \sum_{n=0}^{N-1} f[n] \left[\sum_{k=-(N/2)+1}^{(N/2)-1} \left(e^{j2\pi(x-n\tau)/T} \right)^k \right] \\
&= \sum_{n=0}^{N-1} f[n] \left[\frac{\left(e^{j2\pi(x-n\tau)/T} \right)^{(N-1)/2} - \left(e^{j2\pi(x-n\tau)/T} \right)^{-(N-1)/2}}{\left(e^{j2\pi(x-n\tau)/T} \right)^{1/2} - \left(e^{j2\pi(x-n\tau)/T} \right)^{-1/2}} \right] \\
&= \sum_{n=0}^{N-1} f[n] \left[\frac{\sin \pi(N-1)(x-n\tau)/T}{\sin \pi(x-n\tau)/T} \right] \quad .
\end{aligned}$$

■

5.3.2 The Algorithm

In the general case, sampling and truncating $F(u)$ to meet the conditions of this theorem will introduce errors in the computation of $f(x)$. Truncating $F(u)$ corresponds to lowpass filtering $f(x)$, so provided that both $F(u)$ and $f(x)$ are reasonably smooth and decay rapidly, it should be possible to use the following algorithm to estimate $f(x)$.

1. Choose a value T which sets the range of x over which $f(x)$ is to be computed and a value N which sets the number of points at which $f(x)$ is to be computed. For efficiency, N should be a power of two.

2. Compute the $N-1$ values $F(k/T)$ for $k = -(N/2)+1, \ldots, (N/2)-1$ and append a zero value to get exactly N values. Note that when $f(x)$ is real, $F(u)$

has Hermitian symmetry and that in fact only $N/2$ distinct values actually must be computed.

3. Take the inverse FFT of these N values to obtain the N values $f(n)$ for $n = 0, \ldots, N-1$.

4. These values define the points $f(nT/N)$ of the function $f(x)$. If desired, interpolate these data to obtain f as a continuous function.

The interesting question now, is whether or not this procedure converges to the correct answer as N and T increase, and what the error bounds might be for a specified N and T.

5.3.3 Error Bounds

This section gives error bounds for the algorithm and states conditions under which the algorithm will converge as N and T increase toward infinity. Cooley at al. [21] discuss the sources of error but do not give explicit error bounds. Condition (1) of the following theorem is imposed because we require an error bound for a non-negative random variable; it could be replaced by a condition similar to (2) for a general real random variable. Condition (2) requires that $f(x)$ is "almost" time-limited in the sense that its tails decay rapidly enough that they can be ignored for N and T large enough; this allows us to bound the sampling error introduced by evaluating $F(u)$ only at discrete points. Condition (3) requires that $f(x)$ is "almost" band-limited in the sense that the tails of its spectrum $F(u)$ decay rapidly enough that they can be ignored; this allows us to bound the truncation error introduced by considering $F(u)$ over only a finite interval.

Theorem (Error Bounds): Let $f(x)$ be a function and $F(u)$ its Fourier transform and suppose that there exist constants A, B, $\alpha > 1$, and $\beta > 1/2$ such that

(1) $f(x) = 0$ for all $x < 0$;

(2) $|f(x)| < A|x|^{-\alpha}$ for all $x > T$; and

(3) $|F(u)| < B|u|^{-\beta}$ for all $|u| > N/2$.

Let $f_2(x)$ be the approximation to $f(x)$ computed by the procedure given above. Then the approximation error defined by

$$\|f - f_2\| = \left[\int_0^T |f(x) - f_2(x)|^2 \, dx\right]^{1/2} \tag{159}$$

is bounded by

$$\|f - f_2\| \leq \frac{A}{T^{\alpha-1/2}} \cdot \frac{\alpha}{\alpha - 1} + \frac{B}{\sqrt{\beta - 1/2}} T^{\beta+1/2} \left(\frac{2}{N-2}\right)^{\beta-1/2} . \tag{160}$$

Proof: The diagram below summarizes the connections among $F(u)$, $f_2(x)$, and $f(x)$. The abbreviations are ICFT = inverse continuous Fourier transform; IFS = inverse Fourier series; IDFT = inverse discrete Fourier transform. To periodify a signal is to convolve it with an infinite sequence of impulses (the "comb" function).

$$\begin{array}{ccccccc} F(u) & \xrightarrow{\text{sample}} & F_1[k] & \xrightarrow{\text{truncate}} & F_2[k] & = & F_2[k] \\ \downarrow \text{ICFT} & & \downarrow \text{IFS} & & \downarrow \text{IFS} & & \downarrow \text{IDFT} \\ f(x) & \xrightarrow{\text{periodify}} & f_1(x) & \xrightarrow{\text{lowpass}} & f_2(x) & \xleftarrow{\text{interpolate}} & f_2[k] \end{array}$$

The claim that $f_2(x)$ is is obtained from $f_2[k]$ by interpolation is justified by the correspondence theorem just proven. Now since $\|\cdot\|$ is a norm, we have that

$$\|f - f_2\| \leq \|f - f_1\| + \|f_1 - f_2\|$$

and it suffices to evaluate each of the error terms $\|f - f_1\|$ and $\|f_1 - f_2\|$.

(1) For the first term, consider

$$\begin{aligned} \|f - f_1\|^2 &= \int_0^T |f(x) - f_1(x)|^2 \, dx \\ &= \int_0^T |f(x) - \sum_{\ell=0}^{\infty} f(x + \ell T)|^2 \, dx \\ &= \int_0^T \left|\sum_{\ell=1}^{\infty} f(x + \ell T)\right|^2 dx \\ &\leq \int_0^T \left[\sum_{\ell=1}^{\infty} |f(x + \ell T)|\right]^2 dx \end{aligned}$$

$$\leq \int_0^T \left[\sum_{\ell=1}^{\infty} A\,|x+\ell T|^{-\alpha}\right]^2 dx$$
$$\leq \int_0^T \left[\sum_{\ell=1}^{\infty} A\,|\ell T|^{-\alpha}\right]^2 dx$$
$$= A^2 T^{1-2\alpha}\left[\sum_{\ell=1}^{\infty} \ell^{-\alpha}\right]^2 .$$

But given that $\alpha > 1$,

$$\sum_{\ell=1}^{\infty} \ell^{-\alpha} = 1 + \sum_{\ell=2}^{\infty} \ell^{-\alpha}$$
$$\leq 1 + \int_1^{\infty} \ell^{-\alpha}\, d\alpha$$
$$= 1 + \left[\frac{1}{1-\alpha}\ell^{1-\alpha}\right]_1^{\infty}$$
$$= \frac{\alpha}{\alpha - 1} ,$$

and so it follows that

$$\|f - f_1\|^2 \leq A^2 T^{1-2\alpha}\left[\frac{\alpha}{\alpha-1}\right]^2 .$$

That is, the error term $\|f - f_1\|$ can be reduced below any given value by choosing T large enough, provided that $1 - 2\alpha < 0$, which is always true if $\alpha > 1$.

(2) For the second error term, consider

$$\|f_1 - f_2\|^2 = \int_0^T |f_1 - f_2|^2\, dx$$
$$= T \sum_{k=-\infty}^{\infty} |F_1[k] - F_2[k]|^2 \qquad \text{by Parseval's theorem}$$
$$= T \sum_{|k|\geq N/2} |F_1[k]|^2 \qquad \text{since } F_1[k] = F_2[k] \text{ for } |k| < N/2$$
$$= T \sum_{|k|\geq N/2} |F(k/T)|^2 \qquad \text{by definition of } F_1$$
$$\leq T \sum_{|k|\geq N/2} B^2\,|k/T|^{-2\beta}$$
$$= 2B^2 T^{1+2\beta} \sum_{k=N/2}^{\infty} k^{-2\beta} ,$$

where the series converges provided that $\beta > 1/2$,

$$\leq 2B^2T^{1+2\beta}\int_{(N/2)-1}^{\infty} k^{-2\beta}\,dk \quad ,$$

$$= 2B^2T^{1+2\beta}\left[\frac{1}{2\beta-1}\left(\frac{N-2}{2}\right)^{1-2\beta}\right]$$

$$= \frac{2B^2}{2\beta-1}T^{1+2\beta}\left(\frac{N-2}{2}\right)^{1-2\beta} \quad ,$$

which, provided that $\beta > 1/2$, converges to zero for any fixed T as $N \to \infty$. ■

Corollory: For any $\epsilon > 0$, choosing

$$T > \left(\frac{A\alpha\epsilon}{2(\alpha-1)}\right)^{2/(2\alpha-1)} \tag{161}$$

and

$$N > 2 + 2\left(\frac{B\epsilon}{2\sqrt{\beta-1/2}}T^{\beta+1/2}\right)^{2/(2\beta-1)} \tag{162}$$

is sufficient to guarantee that

$$\|f - f_2\| = \left[\int_0^T |f(x) - f_2(x)|^2\,dx\right]^{1/2} < \epsilon \quad . \tag{163}$$

Proof: It suffices to require that $\|f - f_1\| < \epsilon/2$ and $\|f_1 - f_2\| < \epsilon/2$. Simple algebra then yields the claimed results. ■

Proposition: The algorithm given above converges for the generalized χ^2 distribution with $\nu > 4$. The parameters that appear in the error bound have the values

$$A = \frac{\nu Q_{2\max} + \nu(\nu-1)Q_{1\max}^2}{4\sqrt{\pi}\sigma_{\min}^2} \cdot \frac{\Gamma(\nu/4+\frac{1}{2})}{\Gamma(\nu/4+1)} \tag{164a}$$

$$\alpha = 2 \tag{164b}$$

$$B = 2^{-\nu}\pi^{-\nu/2}\left(\prod_{j=1}^{\nu}\sigma_j^{-1}\right) \tag{164c}$$

$$\beta = \nu/2 \tag{164d}$$

Proof: The first two results follow immediately substituting the asymptotic limits for the generalized χ^2 distribution in the error bounds just derived. For B

and β, it is necessary to restate the asymptotic limit of the characteristic function in terms of the Fourier transform by making the substitution $t = -2\pi u$ to obtain

$$\begin{aligned} |F(u)| &= |\phi(-2\pi u)| \\ &\leq 2^{-\nu/2} \left(\prod_{j=1}^{\nu} \sigma_j^{-1} \right) |2\pi u|^{-\nu/2} \\ &= 2^{-\nu} \pi^{-\nu/2} \left(\prod_{j=1}^{\nu} \sigma_j^{-1} \right) |u|^{-\nu/2} \end{aligned}$$

or

$$|F(u)| \leq B\,|u|^{-\beta}$$

for B and β defined as above. ■

Given the generalized χ^2 distribution, we now have the tools to statistically characterize the reconstruction quality; this will be the subject of the next chapter.

Chapter 6

Reconstruction Quality

There are various ways in which a reconstruction can be "good." First, the reconstructed source distribution can closely match the true source distribution. That is, the reconstruction error $\mathbf{e} = \mathbf{q} - \hat{\mathbf{q}}$ can be close to zero. Second, the reconstructed fields can closely match the measured fields. That is, the measurement residual $\mathbf{r} = \mathbf{b} - \mathbf{F}\hat{\mathbf{q}}$ can be close to zero. Third, the reconstructed source distribution can closely match the assumed priors. That is, the quantity $\hat{\mathbf{q}}$ can be close to zero, which is the assumed mean of the prior source distribution. Each of these quality measures is a vector quantity and may be reduced to a scalar by taking some appropriate vector norm.

The purpose of this chapter is to characterize the meaning and statistical distribution of each of these measures and their norms. In the general case, only the means and covariances of $\mathbf{q}$ and $\mathbf{w}$ are known and it is possible to determine only the means and variances of the quality measures and the means of their norms. But if $\mathbf{q}$ and $\mathbf{w}$ are assumed to be jointly Gaussian, then these quality measures follow a generalized χ^2 distribution.

6.1 Reconstruction Error

The reconstruction error $\mathbf{e} = \mathbf{q} - \hat{\mathbf{q}}$ measures how closely the reconstructed current distribution $\hat{\mathbf{q}}$ matches the true current distribution $\mathbf{q}$. Although it is not possible to compute $\mathbf{e}$ in a real reconstruction problem (since it depends on the unknown true distribution $\mathbf{q}$), it is possible to compute the statistical distribution of $\mathbf{e}$, which is a useful metric for the quality of a reconstruction algorithm.

The mean and variance of $\mathbf{e}$ may be computed in straightforward fashion from its definition. The mean is

$$\mathsf{E}\,\mathbf{e} = \mathsf{E}(\mathbf{q} - \hat{\mathbf{q}}) = \mathsf{E}(\mathbf{q} - \mathbf{H}(\mathbf{F}\mathbf{q} + \mathbf{w})) = 0 \quad . \tag{165}$$

The variance is

$$\begin{aligned} \mathrm{Var}(\mathbf{e}) &= \mathsf{E}(\mathbf{q} - \mathbf{H}\mathbf{b})(\mathbf{q} - \mathbf{H}\mathbf{b})^T \\ &= \mathsf{E}\,\mathbf{q}\mathbf{q}^T - \mathsf{E}\,\mathbf{q}(\mathbf{H}\mathbf{b})^T - \mathsf{E}\,\mathbf{H}\mathbf{b}\mathbf{q}^T + \mathsf{E}\,\mathbf{H}\mathbf{b}\mathbf{b}^T\mathbf{H} \quad , \end{aligned}$$

and since $\mathsf{E}\,\mathbf{q}\mathbf{b}^T = \mathsf{E}\,\mathbf{q}(\mathbf{F}\mathbf{q} + \mathbf{w})^T = \mathbf{A}\mathbf{F}^T + \mathbf{\Gamma}$,

$$\begin{aligned} \mathrm{Var}(\mathbf{e}) &= \mathbf{A} - (\mathbf{A}\mathbf{F}^T + \mathbf{\Gamma})\mathbf{H} - \mathbf{H}^T(\mathbf{A}\mathbf{F}^T + \mathbf{\Gamma})^T \\ &\quad + \mathbf{H}(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{F}\mathbf{\Gamma} + \mathbf{\Gamma}^T\mathbf{F}^T + \mathbf{\Sigma})\mathbf{H}^T \\ &= (\mathbf{I} - \mathbf{H}\mathbf{F})\mathbf{A}(\mathbf{I} - \mathbf{H}\mathbf{F})^T + \mathbf{H}\mathbf{\Sigma}\mathbf{H}^T \\ &\quad - (\mathbf{I} - \mathbf{H}\mathbf{F})\mathbf{\Gamma}\mathbf{H}^T - \mathbf{H}\mathbf{\Gamma}^T(\mathbf{I} - \mathbf{H}\mathbf{F})^T \quad . \end{aligned} \tag{166}$$

If $\mathbf{H}$ is the OCLIM estimator, then the variance simplifies to

$$\mathrm{Var}(\mathbf{e}) = \mathbf{A} - (\mathbf{A}\mathbf{F}^T + \mathbf{\Gamma})(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{F}\mathbf{\Gamma} + \mathbf{\Gamma}^T\mathbf{F}^T + \mathbf{\Sigma})^{-1}(\mathbf{A}\mathbf{F}^T + \mathbf{\Gamma})^T \quad . \tag{167}$$

In the common case that $\mathbf{\Gamma} = 0$, the error variance for OCLIM reduces further to

$$\begin{aligned} \mathrm{Var}(\mathbf{e}) &= \mathbf{A} - \mathbf{A}\mathbf{F}^T(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1}\mathbf{F}\mathbf{A} \\ &= (\mathbf{A}^{-1} + \mathbf{F}^T\mathbf{\Sigma}^{-1}\mathbf{F})^{-1} \quad . \end{aligned} \tag{168}$$

The mean squared norm $\eta^2 = \mathsf{E}\,\|\mathbf{e}\|^2$ of the reconstruction error is the quality metric optimized by both the OCLIM and OWPIM estimators. Its value is simply

$$\eta^2 = \mathrm{Tr}(\mathrm{Var}(\mathbf{e})) \quad . \tag{169}$$

For the weighted pseudoinverse with coaxial priors, the mean square error η^2 can also be computed via equation (65).

6.2 Gaussian Priors and Confidence Limits

In the case of Gaussian priors, it is possible to say a little more about the distributions of $\mathbf{e}$ and $\|\mathbf{e}\|^2$. Note first that $\mathbf{e} = \mathbf{q} - \mathbf{H}\mathbf{F}\mathbf{q} - \mathbf{H}\mathbf{w}$ is a linear combination

of Gaussian random vectors and so $\mathbf{e}$ must also be Gaussian with the mean and variance derived above. Furthermore, the squared norm $\|\mathbf{e}\|^2$ must be generalized χ^2 with parameters 0 and $\mathrm{Var}(\mathbf{e})$, or, in symbols

$$\|\mathbf{e}\|^2 \sim \chi^2(0, \mathrm{Var}(\mathbf{e})) \quad . \tag{170}$$

Given the distribution of $\mathbf{e}$, it is also possible to establish confidence limits on the reconstructed source amplitude $\hat{\mathbf{q}}$. For a given measurement data set $\mathbf{b}$, the expected reconstruction error is

$$\begin{aligned} \mathsf{E}(\mathbf{e} \mid \mathbf{b}) &= \mathsf{E}(\mathbf{q} - \mathbf{H}\mathbf{b} \mid \mathbf{b}) = \mathsf{E}(\mathbf{q} \mid \mathbf{b}) - \mathbf{H}\mathbf{b} \\ &= (\mathbf{A}\mathbf{F}^T + \mathbf{\Gamma})(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{F}\mathbf{\Gamma} + \mathbf{\Gamma}^T\mathbf{F}^T + \mathbf{\Sigma})^{-1}\mathbf{b} - \mathbf{H}\mathbf{b} \\ &= \left((\mathbf{A}\mathbf{F}^T + \mathbf{\Gamma})(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{F}\mathbf{\Gamma} + \mathbf{\Gamma}^T\mathbf{F}^T + \mathbf{\Sigma})^{-1} - \mathbf{H}\right)\mathbf{b} \quad . \end{aligned} \tag{171}$$

If $\mathbf{H}$ is the OCLIM estimator, then $\mathsf{E}(\mathbf{e} \mid \mathbf{b}) = 0$. Thus, OCLIM has no systematic error but all other estimators do (unless they happen to coincide with the OCLIM estimator).

Now for any given $\mathbf{b}$, $\hat{\mathbf{q}} = \mathbf{H}\mathbf{b}$ is a constant and so the variance of the reconstruction error is

$$\begin{aligned} \hat{\mathbf{A}} &\triangleq \mathrm{Var}(\mathbf{e} \mid \mathbf{b}) = \mathrm{Var}(\mathbf{q} - \hat{\mathbf{q}} \mid \mathbf{b}) = \mathrm{Var}(\mathbf{q} \mid \mathbf{b}) \\ &= \mathbf{A} - (\mathbf{A}\mathbf{F}^T + \mathbf{\Gamma})(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{F}\mathbf{\Gamma} + \mathbf{\Gamma}^T\mathbf{F}^T + \mathbf{\Sigma})^{-1}(\mathbf{F}\mathbf{A} + \mathbf{\Gamma}^T) \quad , \end{aligned} \tag{172}$$

which is also the a posteriori variance of $\mathbf{q}$ given $\mathbf{b}$. Using the Cholesky decomposition

$$(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{F}\mathbf{\Gamma} + \mathbf{\Gamma}^T\mathbf{F}^T + \mathbf{\Sigma})^{-1} = \mathbf{L}\mathbf{L}^T \quad , \tag{173}$$

we obtain a form which is convenient for computation

$$\hat{\mathbf{A}} = \mathbf{A} - \left(\mathbf{L}^{-1}(\mathbf{F}\mathbf{A} + \mathbf{\Gamma}^T)\right)^T \left(\mathbf{L}^{-1}(\mathbf{F}\mathbf{A} + \mathbf{\Gamma}^T)\right) \quad . \tag{174}$$

If the noise is uncorrelated with the sources, then $\mathbf{\Gamma} = 0$ and the error variance becomes

$$\begin{aligned} \hat{\mathbf{A}} &= \mathbf{A} - \mathbf{A}\mathbf{F}^T(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1}\mathbf{F}\mathbf{A} \\ &= \mathbf{A} - (\mathbf{L}^{-1}\mathbf{F}\mathbf{A})^T(\mathbf{L}^{-1}\mathbf{F}\mathbf{A}) \quad , \end{aligned} \tag{175}$$

using the Cholesky decomposition

$$\mathbf{FAF}^T + \mathbf{\Sigma} = \mathbf{LL}^T \quad . \tag{176}$$

Note that in either case the error variance does not depend on the actual measurements $\mathbf{b}$ and thus can be computed in advance.

Each diagonal entry $\hat{\alpha}_{nn}^2$ of $\hat{\mathbf{A}}$ is the error variance of the corresponding estimate $\hat{q}_n$; hence each $\hat{q}_n$ has a standard (1σ) error equal to $\hat{\alpha}_{nn}$ and the 90% confidence limits are $\hat{q}_n \pm 1.645\,\hat{\alpha}_{nn}$.

6.3 Measurement Residual

The measurement residual $\mathbf{r} = \mathbf{b} - \mathbf{F}\hat{\mathbf{q}}$ measures how closely the reconstruction matches the measured fields. It depends only on known qualities and thus can be computed in a real reconstruction problem. If the residual is large, there may be some reason to doubt that the source model and reconstruction are adequate to explain the meaured fields. This section discusses the statistical distribution of the residual. The residual can be written in the form

$$\mathbf{r} = (\mathbf{I} - \mathbf{FH})\mathbf{b} \tag{177}$$

from which it is obvious that it has mean value

$$\mathsf{E}(\mathbf{r}) = (\mathbf{I} - \mathbf{FH})\,\mathsf{E}\,\mathbf{b} = 0 \tag{178}$$

and variance

$$\begin{aligned} \mathrm{Var}(\mathbf{r}) &= (\mathbf{I} - \mathbf{FH})\,\mathsf{E}\,\mathbf{b}\mathbf{b}^T(\mathbf{I} - \mathbf{FH})^T \\ &= (\mathbf{I} - \mathbf{FH})(\mathbf{FAF}^T + \mathbf{F\Gamma} + \mathbf{\Gamma}^T\mathbf{F}^T + \mathbf{\Sigma})(\mathbf{I} - \mathbf{FH})^T \quad . \end{aligned} \tag{179}$$

For the OCLIM estimator with $\mathbf{\Gamma} = 0$ we have

$$\mathbf{I} - \mathbf{FH} = \mathbf{I} - \mathbf{FAF}^T(\mathbf{FAF}^T + \mathbf{\Sigma})^{-1} = \mathbf{\Sigma}(\mathbf{FAF}^T + \mathbf{\Sigma})^{-1} \tag{180}$$

and the variance simplifies to

$$\begin{aligned} \mathrm{Var}(\mathbf{r}) &= \mathbf{\Sigma}(\mathbf{FAF}^T + \mathbf{\Sigma})^{-1}\mathbf{\Sigma} \\ &= (\mathbf{L}^{-1}\mathbf{\Sigma})^T(\mathbf{L}^{-1}\mathbf{\Sigma}) \quad . \end{aligned} \tag{181}$$

The appropriate norm for $\mathbf{r}$ should account for variations in noise or measurement accuracy between different detectors. The natural choice is

$$\begin{aligned}\chi^2 &= (\mathbf{b}-\mathbf{F}\hat{\mathbf{q}})^T\boldsymbol{\Sigma}^{-1}(\mathbf{b}-\mathbf{F}\hat{\mathbf{q}}) \\ &= \mathbf{b}^T(\mathbf{I}-\mathbf{F}\mathbf{H})^T\boldsymbol{\Sigma}^{-1}(\mathbf{I}-\mathbf{F}\mathbf{H})\mathbf{b} \quad ,\end{aligned} \tag{182}$$

which is also the traditional χ^2 statistic of linear regression.

The mean squared norm $\mathsf{E}\,\chi^2$ of the residual reveals how closely a reconstruction matches the given measurements on average and is useful as a figure of merit for a reconstruction filter $\mathbf{H}$. It depends only on the mean and variance of $\mathbf{r}$ and can be computed as

$$\begin{aligned}\mathsf{E}\,\chi^2 &= \mathsf{E}\,\mathbf{b}^T(\mathbf{I}-\mathbf{F}\mathbf{H})^T\boldsymbol{\Sigma}^{-1}(\mathbf{I}-\mathbf{F}\mathbf{H})\mathbf{b} \\ &= \mathsf{E}\,\mathbf{b}^T(\mathbf{I}-\mathbf{F}\mathbf{H})^T\mathbf{S}^{-T}\mathbf{S}^{-1}(\mathbf{I}-\mathbf{F}\mathbf{H})\mathbf{b} \\ &= \mathsf{E}\,\|\mathbf{S}^{-1}(\mathbf{I}-\mathbf{F}\mathbf{H})\mathbf{b}\|^2 \\ &= \mathrm{Tr}(\mathbf{S}^{-1}(\mathbf{I}-\mathbf{F}\mathbf{H})\mathbf{B}(\mathbf{I}-\mathbf{F}\mathbf{H})^T\mathbf{S}^{-T}) \\ &= \mathrm{Tr}((\mathbf{S}^{-1}(\mathbf{I}-\mathbf{F}\mathbf{H})\mathbf{L})(\mathbf{S}^{-1}(\mathbf{I}-\mathbf{F}\mathbf{H})\mathbf{L})^T) \\ &= \|\mathbf{S}^{-1}(\mathbf{I}-\mathbf{F}\mathbf{H})\mathbf{L}\|_F^2 \quad .\end{aligned} \tag{183}$$

For the weighted or truncated pseudoinverse,

$$\mathbf{I}-\mathbf{F}\mathbf{H} = \sum_{k=1}^{M}(1-c_k)\mathbf{u}_k\mathbf{u}_k^T \quad , \tag{184}$$

which yields the form

$$\mathsf{E}\,\chi^2 = \left\|\sum_{k=1}^{M}(1-c_k)(\mathbf{S}^{-1}\mathbf{u}_k)(\mathbf{L}^T\mathbf{u}_k)^T\right\|_F^2 \quad . \tag{185}$$

For the OCLIM estimator with $\boldsymbol{\Gamma}=0$, observe that

$$\begin{aligned}\mathbf{I}-\mathbf{F}\mathbf{H} &= \mathbf{I}-\mathbf{F}\mathbf{A}\mathbf{F}^T(\mathbf{F}\mathbf{A}\mathbf{F}^T+\boldsymbol{\Sigma})^{-1} \\ &= \boldsymbol{\Sigma}(\mathbf{F}\mathbf{A}\mathbf{F}^T+\boldsymbol{\Sigma})^{-1} \\ &= \mathbf{S}\mathbf{S}^T\mathbf{L}^{-T}\mathbf{L}^{-1} \quad .\end{aligned} \tag{186}$$

It then follows that

$$\mathsf{E}\,\chi^2 = \|\mathbf{S}^T\mathbf{L}^{-T}\|_F^2 = \|\mathbf{L}^{-1}\mathbf{S}\|_F^2 \quad . \tag{187}$$

In the case of Gaussian priors, it is possible to determine also the distribution of $\mathbf{r}$ and χ^2. Since $\mathbf{r}$ is a weighted sum of Gaussian random vectors, it is also a Gaussian random vector with the mean and variance given above. Furthermore, $\chi^2 = \mathbf{r}^T\mathbf{\Sigma}^{-1}\mathbf{r}$ must have a generalized χ^2 distribution with parameters 0 and $\mathbf{S}^{-1}\,\mathrm{Var}(\mathbf{r})\,\mathbf{S}^{-T}$. In symbols,

$$\chi^2 = (\mathbf{b} - \mathbf{F}\hat{\mathbf{q}})^T\mathbf{\Sigma}^{-1}(\mathbf{b} - \mathbf{F}\hat{\mathbf{q}}) \sim \chi^2(0,\, \mathbf{S}^{-1}\,\mathrm{Var}(\mathbf{r})\mathbf{S}^{-T}) \quad . \tag{188}$$

For the OCLIM estimator with $\mathbf{\Gamma} = 0$, this simplifies to

$$\chi^2 \sim \chi^2(0, (\mathbf{L}^{-1}\mathbf{S})^T(\mathbf{L}^{-1}\mathbf{S})) \quad . \tag{189}$$

6.4 Surprise

The solution $\hat{\mathbf{q}} = \mathbf{H}\mathbf{b}$ is also a random vector for which the mean and variance can be computed. The mean is

$$\mathsf{E}(\hat{\mathbf{q}}) = \mathsf{E}(\mathbf{H}\mathbf{b}) = 0 \tag{190}$$

and the variance is

$$\mathrm{Var}(\hat{\mathbf{q}}) = \mathrm{Var}(\mathbf{H}\mathbf{b}) = \mathbf{H}(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{F}\mathbf{\Gamma} + \mathbf{\Gamma}^T\mathbf{F}^T + \mathbf{\Sigma})\mathbf{H}^T \quad . \tag{191}$$

For the OCLIM estimator, the variance simplifies to

$$\mathrm{Var}(\hat{\mathbf{q}}) = (\mathbf{A}\mathbf{F}^T + \mathbf{\Gamma})(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{F}\mathbf{\Gamma} + \mathbf{\Gamma}^T\mathbf{F}^T + \mathbf{\Sigma})^{-1}(\mathbf{A}\mathbf{F}^T + \mathbf{\Gamma})^T \quad , \tag{192}$$

or, if $\mathbf{\Gamma} = 0$,

$$\mathrm{Var}(\hat{\mathbf{q}}) = \mathbf{A}\mathbf{F}^T(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1}\mathbf{F}\mathbf{A} = (\mathbf{L}^{-1}\mathbf{F}\mathbf{A})^T(\mathbf{L}^{-1}\mathbf{F}\mathbf{A}) \quad . \tag{193}$$

The natural norm for $\hat{\mathbf{q}}$ is

$$\rho^2 = \hat{\mathbf{q}}^T\mathbf{A}^{-1}\hat{\mathbf{q}} = \|\mathbf{R}^{-1}\hat{\mathbf{q}}\|^2 \quad , \tag{194}$$

which is weighted to reflect the a priori uncertainty in the source amplitudes and is thus a measure of the goodness of fit between the reconstucted source distribution and the a prior current distribution. Since ρ^2 increases as the reconstruction becomes less likely relative to the priors, it will be called the "surprise." The mean surprise $\mathsf{E}\,\rho^2$ indicates to what extent the reconstruction filter uses the prior information rather than the given measurements; it has the value

$$\begin{aligned}\mathsf{E}\,\rho^2 &= \mathsf{E}\,\|\mathbf{R}^{-1}\mathbf{H}\mathbf{b}\|^2 \\ &= \mathrm{Tr}(\mathbf{R}^{-1}\mathbf{H}\mathbf{B}\mathbf{H}^T\mathbf{R}^{-T}) \\ &= \mathrm{Tr}((\mathbf{R}^{-1}\mathbf{H}\mathbf{L})(\mathbf{R}^{-1}\mathbf{H}\mathbf{L})^T) \\ &= \|(\mathbf{R}^{-1}\mathbf{H}\mathbf{L})\|_F^2 \quad . \end{aligned} \tag{195}$$

For the OCLIM estimator with $\mathbf{\Gamma} = 0$, the mean surprise simplifies to

$$\mathsf{E}\,\rho^2 = \|\mathbf{R}^T\mathbf{L}^{-T}\|_F^2 = \|\mathbf{L}^{-1}\mathbf{R}\|_F^2 \quad . \tag{196}$$

For the weighted pseudoinverse, the mean surprise simplifies to

$$\begin{aligned}\mathsf{E}\,\rho^2 &= \|\mathbf{R}^{-1}\mathbf{H}\mathbf{L}\|_F^2 \\ &= \left\| \mathbf{R}^{-1} \left(\sum_{k=1}^{K} c_k \lambda_k^{-1} \mathbf{v}_k \mathbf{u}_k^T \right) \mathbf{L} \right\|_F^2 \\ &= \left\| \sum_{k=1}^{K} c_k \lambda_k^{-1} (\mathbf{R}^{-1}\mathbf{v}_k)(\mathbf{L}^T\mathbf{u}_k)^T \right\|_F^2 \quad . \end{aligned} \tag{197}$$

In the case of Gaussian priors, it is also possible to determine the distributions of $\hat{\mathbf{q}}$ and ρ^2. Since $\hat{\mathbf{q}}$ is a weighted sum of Gaussian random vectors, it is also a Gaussian random vector with the mean and variance given above. Furthermore, $\rho^2 = \hat{\mathbf{q}}\mathbf{A}^{-1}\hat{\mathbf{q}}$ must be distributed as a generalized χ^2 random variable, or

$$\rho^2 \sim \chi^2(0, \mathbf{R}^{-1}\,\mathrm{Var}(\hat{\mathbf{q}})\mathbf{R}^{-T}) \quad . \tag{198}$$

For the OCLIM estimator with $\mathbf{\Gamma} = 0$, the distribution simplifies further to

$$\rho^2 \sim \chi^2(0, (\mathbf{L}^{-1}\mathbf{F}\mathbf{R})^T(\mathbf{L}^{-1}\mathbf{F}\mathbf{R})) \quad . \tag{199}$$

Chapter 7

Inverse Methods Compared

This chapter compares the inverse methods discussed in the previous chapters using a Monte Carlo simulation.

7.1 Algorithms for Numerical Linear Algebra

Much research effort has been devoted to the efficient computer implementation of operations in linear algebra. The following three paragraphs briefly summarize the most useful results for the present purpose; more extensive discussions may be found in Golub [38] or another textbook on numerical linear algebra. Many of these algorithms are available in the LAPACK and BLAS libraries [4], which are written in FORTRAN 77; these two libraries were used to implement OCLIM.

The operations of computing the product of a matrix and a vector, the product of two matrices, the sum or difference of two matrices or vectors, and the product of a scalar and a vector or matrix can all be implemented in the obvious fashion from their definitions. A product of the form $\mathbf{y} = \mathbf{A}^T\mathbf{x}$ can be directly implemented as $y_i = \sum_j A_{ji}x_j$ without explicitly computing the transpose; similarly, products such as $\mathbf{X}^T\mathbf{X}$ and $\mathbf{A}\mathbf{F}^T$ can be done directly.

The operation of solving the linear system $\mathbf{A}\mathbf{y} = \mathbf{x}$ for given $\mathbf{A}$ and $\mathbf{x}$ benefits from a less direct approach. If $\mathbf{A}$ is symmetric and positive definite (which holds for all the linear systems considered in this dissertation), it has a Cholesky decomposition $\mathbf{A} = \mathbf{R}\mathbf{R}^T$ where $\mathbf{R}$ is lower triangular. Efficient algorithms for computing the Cholesky decomposition are known and may be found in any textbook on numerical linear algebra. Furthermore, there are efficient algorithms, known as back

substitution and forward substitution, for computing expressions of the form $\mathbf{A}^{-1}\mathbf{x}$, $\mathbf{R}^{-1}\mathbf{x}$, and $\mathbf{R}^{-1}\mathbf{S}$ once the Cholesky decomposition is known.

The operations of computing the norm of a vector and the Frobenius norm of a matrix can also be implemented in the obvious fashion from their definitions and are included in BLAS. Computing the trace of a matrix is not included but can be done trivially from the definition.

7.2 Computational Algorithms for OCLIM

The implementation of OCLIM breaks into two parts: a setup or initialization which is independent of the measured values $\mathbf{b}$ and can be done in advance, and the computation of the estimate $\hat{\mathbf{q}}$ for one or more measurement vectors $\mathbf{b}$. Only the case $\mathbf{\Gamma} = 0$ is considered here; that is, the noise is assumed to be uncorrelated with the sources. This section describes the computational algorithm actually used for these simulations; Section 11.1 describes an improved algorithm.

7.2.1 Initialization

Initialization includes the following steps. The values of $\mathbf{F}$, $\mathbf{A}$, and $\mathbf{\Sigma}$ are required as inputs. Operation counts are given only for the more expensive steps; a flop is a floating point operation such as the addition or multiplication of two floating point numbers.

1.0 Compute the matrix $\mathbf{F}$ using equation (26) or (29) in $O(MN)$ flops. Alternatively, compute the elements of $\mathbf{F}$ via equation (30) using a detailed finite element model as discussed in Chapter 2; this approach incorporates realistic body geometry and volume currents but is not used in this chapter.

1.1 Compute and save the product $\mathbf{FA}$ for later use. This requires $O(MN^2)$ flops.

1.2 Compute the matrix $\mathbf{B} = \mathbf{FAF}^T + \mathbf{\Sigma}$ by direct matrix multiplication and addition in $O(MN^2)$ flops. Compute and save $\mathrm{Tr}(\mathbf{FAF}^T)$.

1.3 Compute the Cholesky decomposition $\mathbf{LL}^T$ of $\mathbf{B}$ in $O(M^3)$ flops.

1.4 Compute $\mathbf{X} = \mathbf{L}^{-1}\mathbf{FA}$ by back substitution in $O(M^3)$ flops.

1.5 Compute $\hat{\mathbf{A}} = \mathbf{A} - \mathbf{X}^T\mathbf{X}$ directly in $O(N^3)$ flops. The diagonal entries $\hat{\alpha}_{nn}^2$ of $\hat{\mathbf{A}}$ are the squares of the standard errors for the corresponding estimates $\hat{q}_n$.

1.6 Compute the mean reconstruction error $\eta^2 = \mathrm{Tr}(\hat{\mathbf{A}})$.

1.7 Compute the Cholesky decomposition $\mathbf{SS}^T$ of the matrix $\mathbf{\Sigma}$ in $O(M^3)$ flops.

1.8 Compute $\mathsf{E}\,\chi^2 = \|\mathbf{L}^{-1}\mathbf{S}\|_F^2$ in $O(M^3)$ flops.

1.9 Compute the signal-to-noise power ratio $\mathrm{SNR} = \mathrm{Tr}(\mathbf{FAF}^T)/\,\mathrm{Tr}(\mathbf{\Sigma})$.

The values of $\mathbf{F}$, $\mathbf{L}$, $\mathbf{S}$, and $\mathbf{FA}$ must be saved for use in inverting particular data sets. The values of $\hat{\mathbf{A}}$, η^2, $\mathsf{E}\,\chi^2$, and SNR provide information to help the user interpret the reconstruction results. The total cost is $O(M^3) + O(N^3)$ flops; some minor improvements are possible if $\mathbf{A}$ and $\mathbf{\Sigma}$ are diagonal.

7.2.2 Reconstruction

Computing the estimate for each input $\mathbf{b}$ includes:

2.1 Compute $\mathbf{p} = \mathbf{L}^{-T}\mathbf{L}^{-1}\mathbf{b}$ by back substitution in $O(M^2)$ flops.

2.2 Compute the solution $\hat{\mathbf{q}} = (\mathbf{FA})^T\mathbf{p}$ in $O(MN)$ flops.

2.3 Compute the residual $\mathbf{r} = \mathbf{b} - \mathbf{F}\hat{\mathbf{q}}$ in $O(MN)$ flops.

2.4 Compute $\chi^2 = \|\mathbf{S}^{-1}\mathbf{r}\|^2$ in $O(M^2)$ flops.

The value of $\hat{\mathbf{q}}$ is the desired reconstruction. The values of $\mathbf{r}$ and χ^2 are intended to provide insight into the quality of the reconstruction. The total cost per reconstruction is $O(M^2) + O(MN)$ flops.

7.3 Methods

The four reconstruction methods discussed in Chapter 4 have been implemented in FORTRAN 77 for noise uncorrelated with sources ($\mathbf{\Gamma} = 0$) and independent priors ($\mathbf{A}$ and $\mathbf{\Sigma}$ diagonal). The LAPACK and BLAS libraries [4] were used for linear algebra computations. All computations were done in double precision.

More specifically, the methods implemented were the minimum-norm least squares method (MNLS) as defined by equation (59), with the threshold set to 10^{-10} times the largest singular value; the optimally truncated pseudoinverse method (OTPIM) defined by equation (68); the optimally weighted pseudoinverse method (OWPIM) defined by equation (61) and using the weights defined by equation (66);

and the optimal constrained linear inverse method (OCLIM) defined by equation (74).

This computer implementation was used in a Monte Carlo simulation to verify the theoretical results of Chapter 4. Three different geometrical configurations were tested, each at five different signal-to-noise ratios. All sources were modelled as current dipoles.

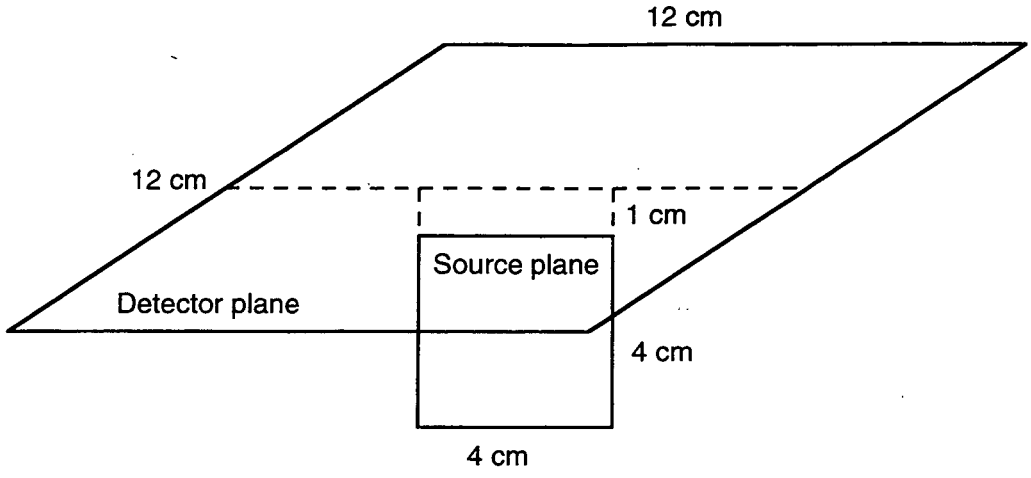
Figure 7.1 shows one of the test configurations. The sources are arranged in a 4 $\times$ 4 cm^2 planar array perpendicular to the detector plane, and centered below that plane with its nearest edge 1 cm away. The source plane contains an 8 $\times$ 8 array of current dipoles directed perpendicular to the plane. For non-uniform priors, the 28 sources in the central cruciform region are assigned a source variance $\alpha_A^2 = 1.0$; the remaining sources are assigned a different source variance $\alpha_B^2 = 0.01$. For uniform priors, all 64 sources are assigned the same source variance $\alpha^2 = 1$.

Figure 7.2 shows the detector array, which has a 12 $\times$ 12 cm^2 planar array of 144 detectors arranged in a 12 $\times$ 12 grid. Each detector measures the magnetic field perpendicular to the plane of the array. Noise amplitudes taken from independent normal distributions with mean zero and variance σ^2 are added to each field measurement.

This test configuration does not accurately model any real magnetometer or clinical application; it is intended to provide a fair comparison of the methods but not to demonstrate the results possible in any particular application.

7.4 Simulation Results

As was mentioned in Section 4.1, the choice of threshold in the truncated pseudoinverse can have a dramatic effect on the reconstruction error. Figure 7.3 shows the reconstruction error η^2 with uniform priors for a truncated pseudoinverse as a function of the signal-to-noise ratio and smallest singular value retained (denoted λ_{min}). Curves are shown for SNR = 127 dB (leftmost), 87 dB, 47 dB, and 7 dB (rightmost); the corresponding noise variances are $\sigma^2 = 10^{-20}$, 10^{-16}, 10^{-12}, and 10^{-8}. The error rises sharply as λ_{min} is decreased below the optimal value σ/α; thus it is generally better to retain too few terms than too many.



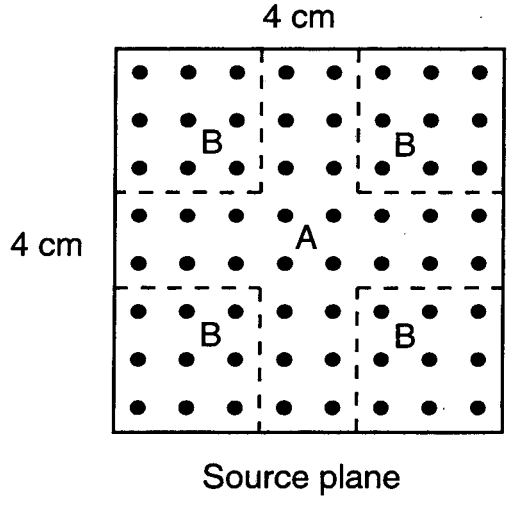


Figure 7.1 A simple test configuration. The source plane is perpendicular to the detector plane and contains 64 dipoles perpendicular to the source plane. In the case of uniform priors, all dipoles have the same expected power. In the case of non-uniform priors, the dipoles in region A have higher expected power.

Table 1 shows the theoretical and experimental results for the test configuration with uniform priors ($\alpha_A^2 = \alpha_B^2$) as the SNR varies from 128 to -32 dB. Note that the theoretical and experimental values all agree within 2 standard errors.

As predicted, OCLIM and OWPIM show identical results for uniform priors. OTPIM typically has slightly larger values of η^2, $\mathsf{E}\,\chi^2$, and $\mathsf{E}\,\rho^2$ than either. MNLS

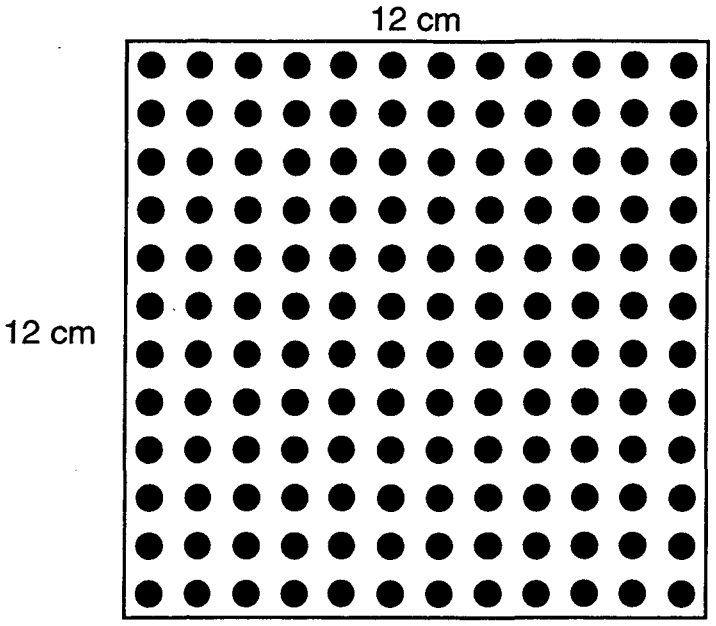


Figure 7.2 Detector grid. The detector grid contains a 12 by 12 array of sensors that sample the field perpendicular to the array. All sensors have the same prior noise variance.

has better $\mathsf{E}\,\chi^2$ values than any of the other methods but its error is orders of magnitude worse. The mean error η^2 and mean residual $\mathsf{E}\,\chi^2$ both increase as the noise increases; the mean surprise $\mathsf{E}\,\rho^2$ decreases.

Table 2 shows the theoretical and experimental results for the test configuration with non-uniform priors ($\alpha_A^2 \neq \alpha_B^2$) as the SNR varies from -37 dB to 123 dB (and noise variance σ^2 from 10^{-20} to 10^{-4}). The theoretical and experimental values again agree within 2 standard errors. Most of the other observations on Table 1 remain true. The big change is that OCLIM now has a smaller mean error η^2 than any of the other methods. The difference is largest when the noise is small and decreases to insignificance when the noise is large enough.

Figure 7.4 shows the reconstruction error from Table 2. The MNLS method is much worse than the other three, with reconstruction error orders of magnitude larger than the true solution.

Figure 7.5 shows the reconstruction error for OTPIM, OWPIM, and OCLIM only. The optimally truncated and weighted pseudoinverses give similar results,

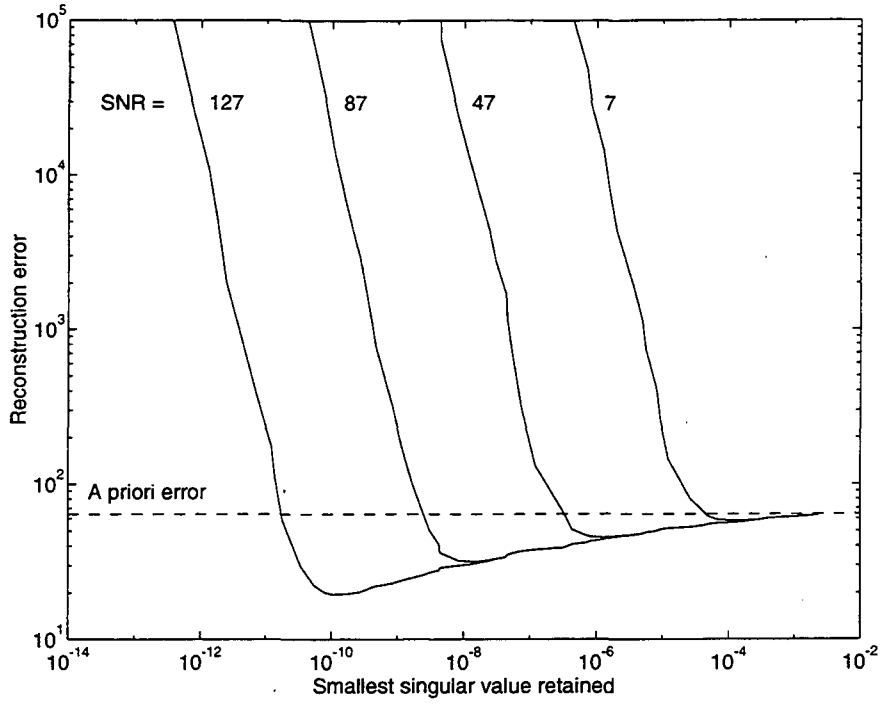


Figure 7.3 Reconstruction error for a truncated pseudoinverse. The reconstruction error η^2 is plotted as a function of the smallest singular value $\lambda_{\min}$ retained in the truncated pseudoinverse, for the test configuration with uniform priors. The four curves are for SNR = 127 dB (leftmost), 87 dB, 47 dB, and 7 dB (rightmost), with corresponding noise variances $\sigma^2 = 10^{-20}, 10^{-16}$, 10^{-12}, and 10^{-8}. The source variance is $\alpha^2 = 1.0$ for all sources and all curves. The horizontal dashed line shows the a priori error $\sum \alpha^2$, which is the error obtained by setting all source estimates to zero.

with OWPIM slightly better. OCLIM is definitely better than either, with the difference increasing as the SNR increases.

Figure 7.6 shows the mean residual $\mathsf{E}\,\chi^2$ for all four methods under the same conditions. MNLS has the smallest residual, which is independent of the SNR. The other three methods have a moderately larger residual which decreases as the SNR increases; differences among OTPIM, OWPIM, and OCLIM are slight.

Uniform priors give similar results, except that OCLIM and OWPIM become identical in this case.

Note that the knowledge of approximate source and noise amplitudes provided by the priors has allowed OCLIM, OWPIM, and OTPIM to generate much better

Table 1. This table shows how the reconstruction quality varies with noise level when the four different methods are used on the test configuration with uniform priors.

		OCLIM	OWPIM	OTPIM	MNLS
Configuration 1 with $\sigma^2 = 1.000 \times 10^{-20}$, $\alpha_A^2 = 1.000$, $\alpha_B^2 = 1.000$, $SNR = 127.5$ db, and 10000 data sets:					
η^2	theo	18.193	18.194	19.543	2.417×10^5
	expr	18.157 ± 0.059	18.157 ± 0.059	19.533 ± 0.062	$(2.391 \pm 0.021) \times 10^5$
$\mathsf{E}\,\chi^2$	theo	98.187	98.191	99.144	84.999
	expr	98.012 ± 0.140	98.006 ± 0.140	98.986 ± 0.143	84.797 ± 0.130
$\mathsf{E}\,\rho^2$	theo	45.807	45.806	47.543	2.417×10^5
	expr	45.919 ± 0.094	45.918 ± 0.094	47.636 ± 0.099	$(2.391 \pm 0.021) \times 10^5$
Configuration 1 with $\sigma^2 = 1.000 \times 10^{-16}$, $\alpha_A^2 = 1.000$, $\alpha_B^2 = 1.000$, $SNR = 87.5$ db, and 10000 data sets:					
η^2	theo	30.762	30.762	31.860	2.417×10^9
	expr	30.729 ± 0.077	30.729 ± 0.077	31.862 ± 0.079	$(2.391 \pm 0.021) \times 10^9$
$\mathsf{E}\,\chi^2$	theo	110.762	110.762	112.324	85.000
	expr	110.666 ± 0.148	110.666 ± 0.148	112.254 ± 0.151	84.797 ± 0.130
$\mathsf{E}\,\rho^2$	theo	33.238	33.238	33.860	2.417×10^9
	expr	33.352 ± 0.080	33.352 ± 0.080	33.983 ± 0.083	$(2.391 \pm 0.021) \times 10^9$
Configuration 1 with $\sigma^2 = 1.000 \times 10^{-12}$, $\alpha_A^2 = 1.000$, $\alpha_B^2 = 1.000$, $SNR = 47.5$ db, and 10000 data sets:					
η^2	theo	43.890	43.890	45.452	2.417×10^{13}
	expr	43.918 ± 0.092	43.918 ± 0.092	45.503 ± 0.094	$(2.391 \pm 0.021) \times 10^{13}$
$\mathsf{E}\,\chi^2$	theo	123.890	123.890	125.658	85.000
	expr	123.730 ± 0.157	123.730 ± 0.157	125.476 ± 0.160	84.797 ± 0.130
$\mathsf{E}\,\rho^2$	theo	20.110	20.110	21.452	2.417×10^{13}
	expr	20.135 ± 0.061	20.135 ± 0.061	21.454 ± 0.069	$(2.391 \pm 0.021) \times 10^{13}$
Configuration 1 with $\sigma^2 = 1.000 \times 10^{-8}$, $\alpha_A^2 = 1.000$, $\alpha_B^2 = 1.000$, $SNR = 7.5$ db, and 10000 data sets:					
η^2	theo	56.880	56.880	57.803	2.417×10^{17}
	expr	56.878 ± 0.105	56.878 ± 0.105	57.824 ± 0.107	$(2.391 \pm 0.021) \times 10^{17}$
$\mathsf{E}\,\chi^2$	theo	136.880	136.880	137.989	85.000
	expr	136.667 ± 0.165	136.667 ± 0.165	137.739 ± 0.167	84.797 ± 0.130
$\mathsf{E}\,\rho^2$	theo	7.120	7.120	7.803	2.417×10^{17}
	expr	7.161 ± 0.035	7.161 ± 0.035	7.883 ± 0.043	$(2.391 \pm 0.021) \times 10^{17}$
Configuration 1 with $\sigma^2 = 1.000 \times 10^{-4}$, $\alpha_A^2 = 1.000$, $\alpha_B^2 = 1.000$, $SNR = -32.5$ db, and 10000 data sets:					
η^2	theo	63.923	63.923	64.000	2.417×10^{21}
	expr	63.995 ± 0.113	63.995 ± 0.113	64.068 ± 0.113	$(2.391 \pm 0.021) \times 10^{21}$
$\mathsf{E}\,\chi^2$	theo	143.923	143.923	144.080	85.000
	expr	143.651 ± 0.170	143.651 ± 0.170	143.806 ± 0.170	84.797 ± 0.130
$\mathsf{E}\,\rho^2$	theo	0.077	0.077	0.000	2.417×10^{21}
	expr	0.076 ± 0.001	0.076 ± 0.001	0.000 ± 0.000	$(2.391 \pm 0.021) \times 10^{21}$

Table 2. This table shows how the reconstruction quality varies with noise level when the four different methods are used on the test configuration with non-uniform priors.

		OCLIM	OWPIM	OTPIM	MNLS
Configuration 1 with $\sigma^2 = 1.000 \times 10^{-20}$, $\alpha_A^2 = 1.000$, $\alpha_B^2 = 0.010$, $SNR = 122.9$ db, and 10000 data sets:					
η^2	theo	4.486	10.727	11.419	2.417×10^5
	expr	4.456 ± 0.025	10.709 ± 0.041	11.400 ± 0.043	$(2.391 \pm 0.021) \times 10^5$
$\mathsf{E}\,\chi^2$	theo	101.352	98.999	100.316	84.999
	expr	101.172 ± 0.141	98.809 ± 0.140	100.117 ± 0.143	84.797 ± 0.130
$\mathsf{E}\,\rho^2$	theo	42.651	231.248	256.449	8.441×10^6
	expr	42.763 ± 0.090	232.159 ± 0.821	257.459 ± 0.943	$(8.395 \pm 0.081) \times 10^6$
Configuration 1 with $\sigma^2 = 1.000 \times 10^{-16}$, $\alpha_A^2 = 1.000$, $\alpha_B^2 = 0.010$, $SNR = 82.9$ db, and 10000 data sets:					
η^2	theo	10.209	17.039	17.519	2.417×10^9
	expr	10.205 ± 0.041	17.002 ± 0.054	17.492 ± 0.055	$(2.391 \pm 0.021) \times 10^9$
$\mathsf{E}\,\chi^2$	theo	114.253	111.774	111.585	85.000
	expr	114.161 ± 0.150	111.688 ± 0.149	111.483 ± 0.150	84.797 ± 0.130
$\mathsf{E}\,\rho^2$	theo	29.747	225.663	257.338	8.443×10^{10}
	expr	29.747 ± 0.075	225.201 ± 0.775	257.600 ± 0.927	$(8.395 \pm 0.081) \times 10^{10}$
Configuration 1 with $\sigma^2 = 1.000 \times 10^{-12}$, $\alpha_A^2 = 1.000$, $\alpha_B^2 = 0.010$, $SNR = 42.9$ db, and 10000 data sets:					
η^2	theo	17.366	22.307	22.939	2.417×10^{13}
	expr	17.403 ± 0.057	22.313 ± 0.064	22.936 ± 0.065	$(2.391 \pm 0.021) \times 10^{13}$
$\mathsf{E}\,\chi^2$	theo	128.069	125.447	126.082	85.000
	expr	127.872 ± 0.159	125.267 ± 0.158	125.880 ± 0.160	84.797 ± 0.130
$\mathsf{E}\,\rho^2$	theo	15.931	169.524	211.813	8.443×10^{14}
	expr	15.905 ± 0.053	168.929 ± 0.697	211.027 ± 0.971	$(8.395 \pm 0.081) \times 10^{14}$
Configuration 1 with $\sigma^2 = 1.000 \times 10^{-8}$, $\alpha_A^2 = 1.000$, $\alpha_B^2 = 0.010$, $SNR = 2.9$ db, and 10000 data sets:					
η^2	theo	24.393	26.655	27.093	2.417×10^{17}
	expr	24.404 ± 0.069	26.644 ± 0.072	27.089 ± 0.073	$(2.391 \pm 0.021) \times 10^{17}$
$\mathsf{E}\,\chi^2$	theo	139.623	138.394	141.165	85.000
	expr	139.386 ± 0.167	138.158 ± 0.166	140.937 ± 0.170	84.797 ± 0.130
$\mathsf{E}\,\rho^2$	theo	4.377	74.427	88.174	8.443×10^{18}
	expr	4.327 ± 0.027	73.841 ± 0.430	87.658 ± 0.622	$(8.395 \pm 0.081) \times 10^{18}$
Configuration 1 with $\sigma^2 = 1.000 \times 10^{-4}$, $\alpha_A^2 = 1.000$, $\alpha_B^2 = 0.010$, $SNR = -37.1$ db, and 10000 data sets:					
η^2	theo	28.333	28.349	28.360	2.417×10^{21}
	expr	28.303 ± 0.075	28.320 ± 0.075	28.331 ± 0.075	$(2.391 \pm 0.021) \times 10^{21}$
$\mathsf{E}\,\chi^2$	theo	143.973	143.973	144.028	85.000
	expr	143.704 ± 0.170	143.705 ± 0.170	143.760 ± 0.170	84.797 ± 0.130
$\mathsf{E}\,\rho^2$	theo	0.027	0.653	0.000	8.443×10^{22}
	expr	0.027 ± 0.000	0.649 ± 0.008	0.000 ± 0.000	$(8.395 \pm 0.081) \times 10^{22}$

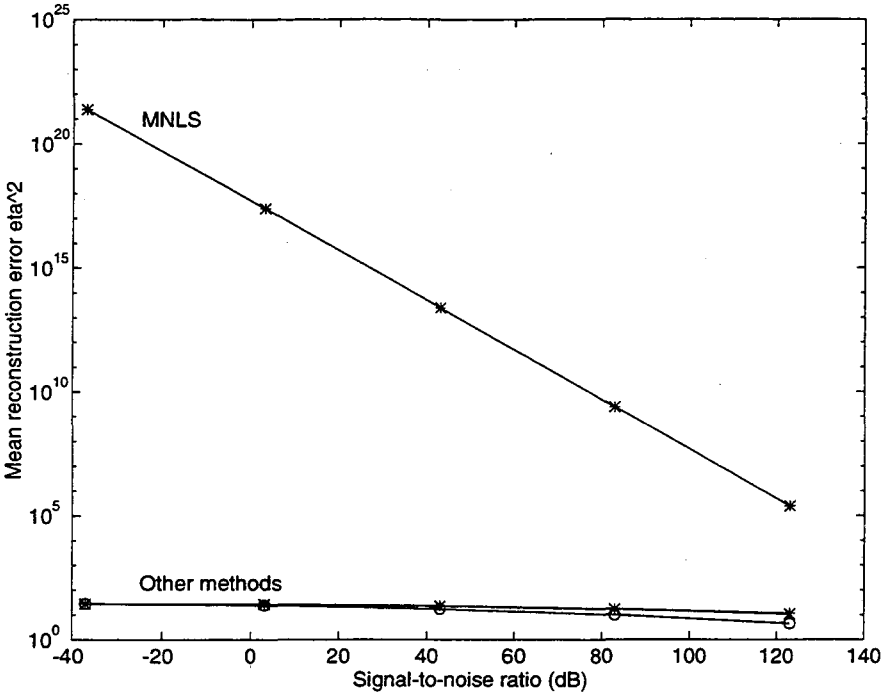


Figure 7.4 Reconstruction error (log scale). The reconstruction error η^2 is plotted as a function of signal-to-noise ratio for all four reconstruction methods with non-uniform priors. The source variance is held constant while the noise variance is changed. MNLS has by far the worst reconstruction error; the other three methods are barely distinguishable on this plot. Figure 7.5 shows the other three methods in an expanded plot.

solutions than MNLS. That is, knowing the expected signal and noise amplitudes regularizes the pseudoinverse and tames an otherwise ill-conditioned problem.

In summary, MNLS provides the best mean residual $\mathsf{E}\,\chi^2$ but does so only by allowing an extremely large reconstruction error η^2. Truncating the pseudoinverse reduces the error but increases the residual. For uniform priors, OTPIM, OWPIM, and OCLIM all give similar errors and residuals. For non-uniform priors, OCLIM has smaller error than any of the other methods, with residuals similar to OTPIM and OWPIM.

For this test configuration, initialization (steps 1.0–1.9) and reconstruction (steps 2.1–2.4) for a hundred data sets take about a second on a workstation class computer. There are only minor differences in the execution time required for the four different reconstruction methods.

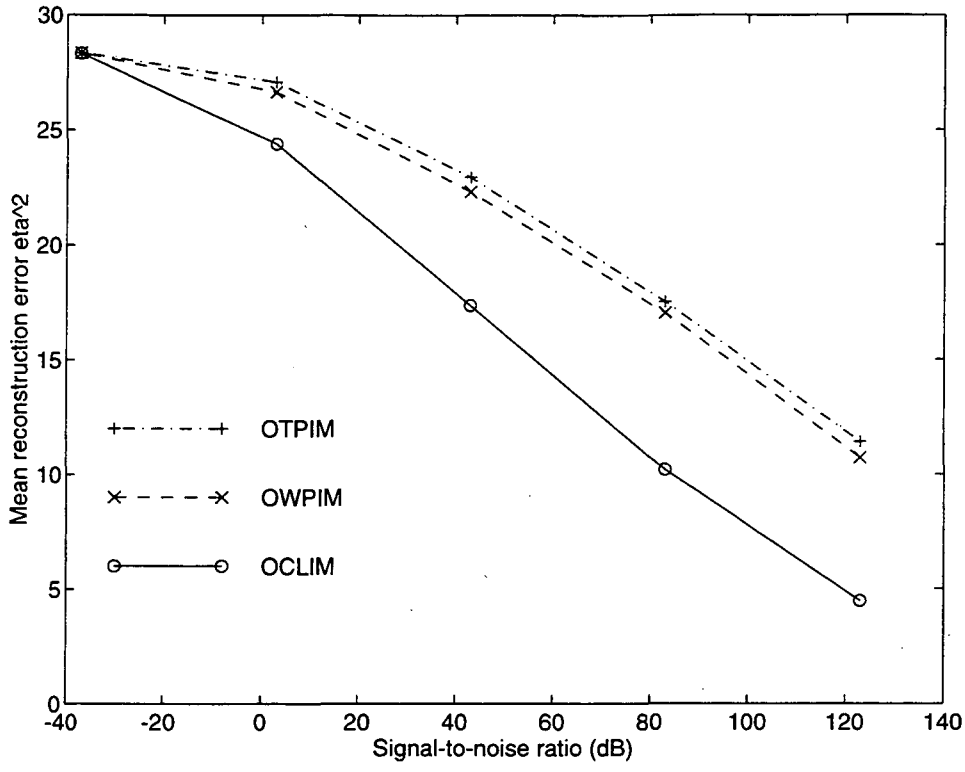


Figure 7.5 Reconstruction error (linear scale). The reconstruction error η^2 is plotted as a function of signal-to-noise ratio under the same conditions as Figure 7.4. The three curves shown are for OTPIM (+), OWPIM (×), and OCLIM (○). OTPIM and OWPIM provide similar results. OCLIM is better than either, with the difference increasing as the SNR increases.

Figures 7.7 and 7.8 show some sample reconstructions from the same test configuration, with $\alpha_A^2 = 1.0$, $\alpha_B^2 = 0.1$, $\sigma^2 = 3.16 \times 10^{-10}$, and SNR = 19 dB. A single source dipole was active with amplitude 8 and all other sources were zero. The lefthand plot of each pair was reconstructed with OCLIM; the righthand plot with OWPIM. The dipole is at the position indicated by the dot and is pointing out of the page. Grey shading indicates areas of current flow into the page.

Figure 7.7 shows three reconstructions of sources consistent with the priors (that is, sources in region A). The top pair shows a source relatively near the detector array. Both reconstructions correctly localize the true source, but the OCLIM reconstruction has higher peak amplitude and is more narrowly localized in space. The artifacts in the OCLIM reconstruction are all less than the a posteriori standard errors.

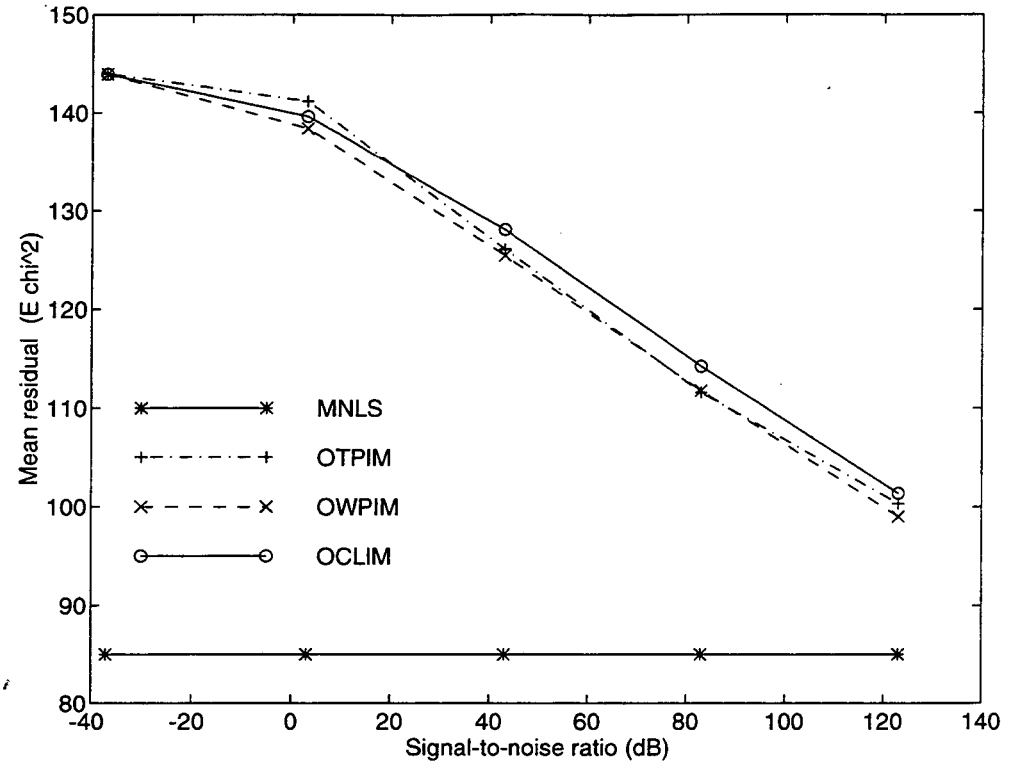


Figure 7.6 Mean residual. The mean residual $E\,\chi^2$ is plotted as a function of signal-to-noise ratio under the same conditions as Figure 7.4. All four methods are shown. MNLS has the lowest residual, which is independent of SNR. The other three methods have similar residuals, which decrease as the SNR increases.

The middle pair shows a source further from the detectors. Both reconstructions are displaced toward the detectors. OCLIM is still higher and narrower than OWPIM.

The bottom pair shows a source distant from the source array. Neither method gives a good reconstruction. OCLIM appears to localize the source but the amplitudes are less than the standard error. OWPIM hardly localizes the source at all.

In additional simulations done at a high SNR (83 dB), both shallow and deep sources were accurately localized without any displacement toward the surface. One of the reviewers commented that displacements toward the surface have been seen by other researchers; it appears that this effect occurs only for certain combinations of source depth and signal-to-noise ratio.

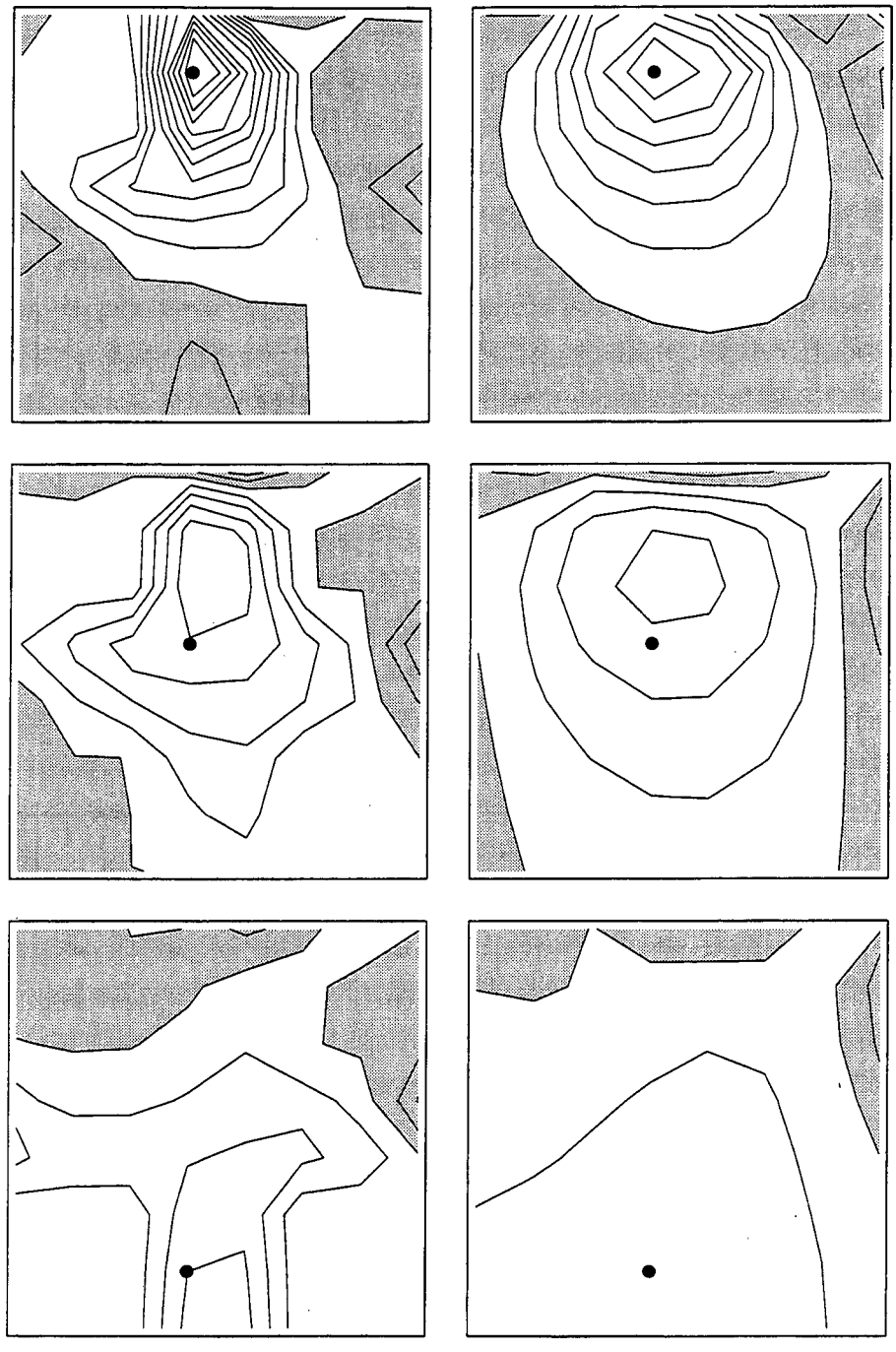


Figure 7.7 Reconstruction of sources consistent with the prior knowledge. The lefthand image of each pair shows the reconstruction using OCLIM with prior information; the righthand image shows the reconstruction using OWPIM with uniform priors. The true distribution is a single dipole out of the page at the position indicated by the dot. Grey shading indicates areas of current flow into the page. The signal-to-noise ratio is 19 dB.

Figure 7.8 shows two reconstructions of sources inconsistent with the priors. OCLIM forces the reconstruction to fit the assumed priors and incorrectly localizes both shallow and deep sources. OWPIM produces a reasonable reconstruction for a shallow source but again hardly localizes a deep source.

7.5 Discussion

If the unknown current distribution is expressed as a linear combination of elementary current distributions in fixed positions, then the magnetic field measurements are linear in the unknown source amplitudes. If, in addition, the cost function to be minimized is either the mean square error (reconstructed minus true currents) or the square residual (measured minus reconstructed fields), then the unknown source amplitudes may be found by solving a linear problem. This offers several advantages: The problem is well understood theroretically and software for its solution is readily available. There is only a single, global minimum. Efficient and reliable computer codes for linear algebra are readily available.

The minimum-norm least squares (MNLS) method, also known as the Moore-Penrose inverse and the generalized inverse, provides a lower residual than any other method but does not exploit prior knowledge. But if the problem is poorly conditioned and there is measurement noise, its reconstruction error can be orders of magnitude larger than the true current distribution. Magnetic imaging is both noisy and poorly conditioned, so MNLS is not generally suitable. Truncating the pseudoinverse can yield better results, but it is usually still inferior to the newer methods described in this dissertation.

The weighted pseudoinverse developed by Shim and Cho generalizes MNLS by including an arbitrary weight in each term of the outer product or spectral expansion of the Moore-Penrose inverse; a truncated pseudoinverse is obtained by restricting the weights to zero and one. Choosing the weights to minimize the mean square error yields an optimally weighted pseudoinverse method (OWPIM) or optimally truncated pseudoinverse method (OTPIM). The source and noise covariance matrices determine the optimum but only the diagonal entries are used; that is, no account is taken of spatial correlations between sources. Prior knowledge

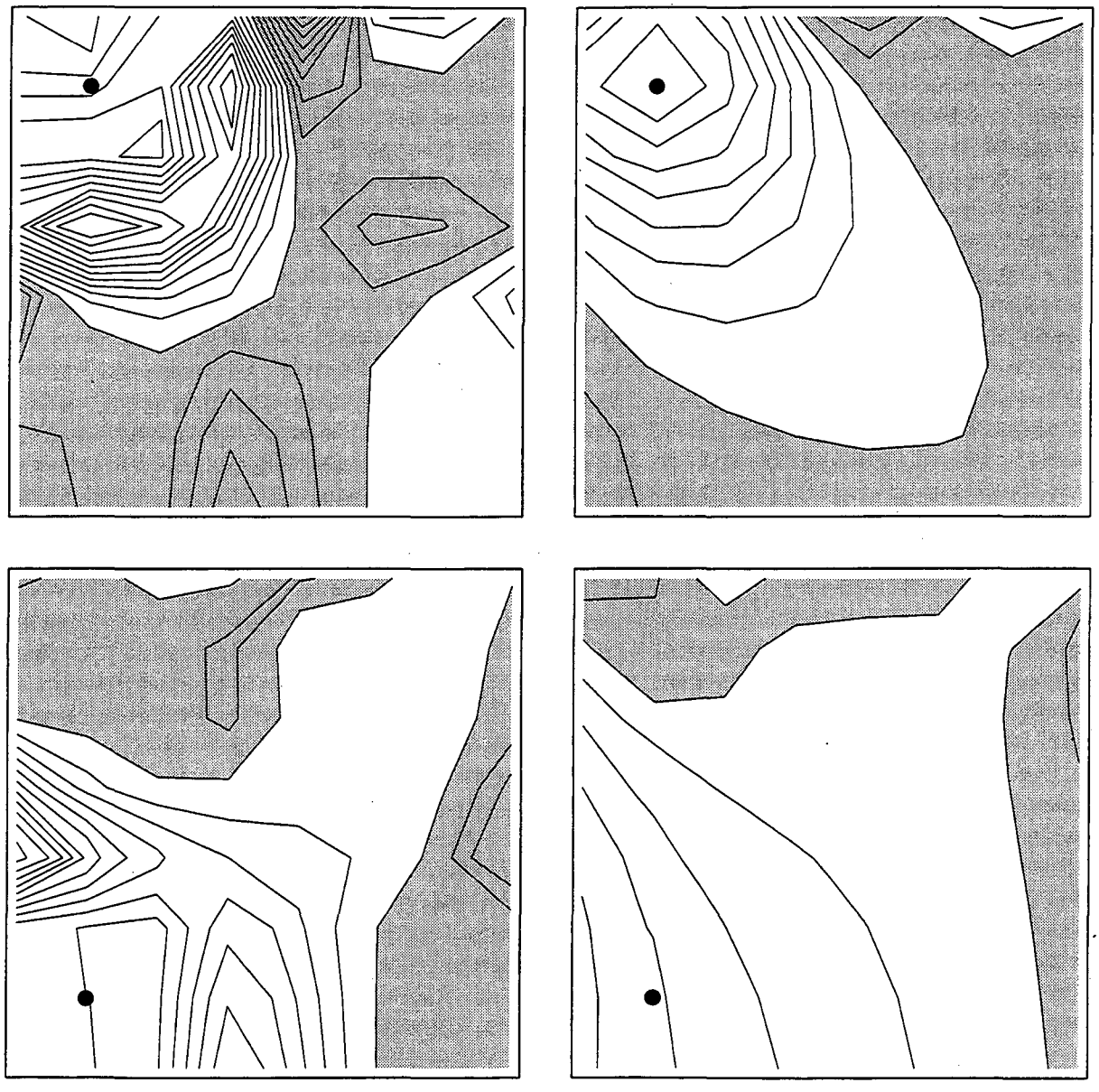


Figure 7.8 Reconstruction of sources inconsistent with the prior knowledge. The lefthand image of each pair shows the reconstruction using OCLIM with prior information; the righthand image shows the reconstruction with uniform priors. The true distribution is a single dipole out of the page at the position indicated by the dot. Grey shading indicates areas of current flow into the page. The signal-to-noise-ratio is 19 dB.

of the source and noise covariance is required but non-uniform priors are permitted only in special cases. OWPIM and OTPIM generally have mean square error larger than OCLIM and residuals larger than MNLS.

The optimal constrained linear inverse method (OCLIM) derived in this chapter uses prior knowledge to obtain a minimum mean square error estimate of the current distribution; OCLIM can be efficiently computed using a Cholesky decomposition. Any source and detector configuration is allowed as long as their positions are fixed a priori. Any correlations between source and noise amplitudes are permitted, including spatial correlations between sources or between detectors. OCLIM locates point sources more precisely than OWPIM but is prone to artifacts when the true sources are inconsistent with the priors.

OCLIM reduces to the optimally weighted pseudoinverse method when the source amplitudes are independent and identically distributed and to the minimum-norm least squares estimate in the limit of no measurement noise or no prior knowledge of the source amplitudes.

All four methods are fast to compute, taking about a second on a workstation for a problem with 64 sources and 144 detectors.

Of these methods, OCLIM appears the best suited to magnetic imaging, since it exploits prior information, provides the minimum reconstruction error, and is no more expensive to compute than the others.

The use of prior information, however, can be risky. If the priors are misinformative, they can distort the solution and lead to inaccurate reconstructions. The next chapter will discuss this issue.

Chapter 8

Priors Informative And Misinformative

The existence of misinformative priors in inverse problems, as demonstrated in the previous chapter, places the analyst in a unpleasant dilemma. The reconstruction accuracy can be improved by informative priors, thereby increasing the useful information obtained from the given measurements. But the accuracy can also be degraded by misinformative priors, leading to false reconstructions. Since the priors encode our expectations about the solution, a choice of priors that are inappropriate for the problem may lead directly to a solution which is neat, plausible—and wrong! The problem of priors is to somehow sail a middle course between ignorance and delusion. Put another way, how much trust should a reasonable person repose in an inverse problem solution obtained using prior knowledge? And how should the analyst choose the prior knowledge to maximize the trustworthy information gained? More specifically, given a set of priors and a set of measured data, is it reasonable to reconstruct the given data using the given priors?

This chapter briefly considers the first two questions and then concentrates on the third question in considering a statistical test for consistency between data and priors.

8.1 Epistemology of Prior Knowledge

An inverse problem solution is often a surrogate for knowledge that could be obtained by tests that are more definitive but also more expensive, more dangerous, or just impractical. The use of MEG rather than invasive electrodes to localize an

epileptic focus prior to surgery is a good example. Provided that the MEG reconstruction is accurate, it reduces the risk to the patient. Similarly, the use of MEG to trace the activation sequence in the brain during a reading task is a surrogate for invasive electrodes that impose an unacceptable risk in a healthy subject.

The claim that the inverse method is a useful surrogate is thus the claim that, in some particular case, it provides the same results that would be obtained by the more definitive but less practical method. This claim may be supported by the following three more general claims:

1. The assumed priors are true for some defined population.

2. The reconstruction method and assumed priors give solutions consistent with the definitive test in this population.

3. This particular case is drawn from that population.

The first two of these claims may be supported by experimental and statistical tests. It is the purpose of this chapter to investigate a statistical test for the third claim: that this particular data set is consistent with the assumed priors. In fact, the constraints defined by the prior knowledge will be regarded as a *hypothesis,* to be confirmed or refuted by the observed data.

8.2 Are the Data Consistent with the Prior Knowledge?

Consider the null hypothesis that the given data are taken from a population in which the source amplitudes $\mathbf{q}$ and the noise amplitudes $\mathbf{w}$ are jointly normal with covariances $\mathbf{A} = \mathsf{E}\,\mathbf{q}\mathbf{q}^T$ and $\mathbf{\Sigma} = \mathsf{E}\,\mathbf{w}\mathbf{w}^T$ and cross-covariance $\mathsf{E}\,\mathbf{q}\mathbf{w}^T = 0$. That is, the null hypothesis is that the data are taken from the population defined by the priors and thus that the priors are informative. Given this hypothesis, the distribution of the measurement vector $\mathbf{b}$ is defined, and it is possible to ask how likely the observed data are relative to this distribution. If the data lies in some critical region which is considered too unlikely, the hypothesis should be rejected; that is, the data do not appear to be from a population with the assumed priors, the priors are probably misinformative rather than informative, and so a reconstruction using those priors is not trustworthy. This critical region is normally chosen to include some fraction α of the total probability, where α is the false positive rate,

or the probability that a consistent data set will be rejected as inconsistent. Since $\mathbf{b}$ is a vector, it is convenient to condense it into a scalar test statistic and define the critical region by a range of values for the test statistic.

If the null hypothesis is that the observed data come from the population defined by the assumed statistics, then a positive result (of the statistical test) indicates that the data do not appear to come from the populated assumed in the priors. Thus the more desireable result (from the viewpoint of reconstruction with priors) is that the statistical test yield the negative result that the priors appear to be consistent with the data. The true negative outcome is that the data actually come from the population defined by the priors and the test correctly indicates this; the false negative outcome is that the data do not actually come from the population defined by the priors but the test incorrectly indicates that they do. Conversely, the true positive outcome is that the data set does not actually come from the assumed priors and the statistical test correctly claims that they do not; the false positive outcome is that the data actually come from the assumed population but the test claims that they do not.

One approach to a test statistic is to observe that, under the assumed prior source and noise statistics, the measured data $\mathbf{b}$ are distributed as a Gaussian random vector with mean zero and variance $\mathbf{B} = \mathbf{FAF}^T + \mathbf{\Sigma} = \mathbf{LL}^T$ where $\mathbf{L}$ is a nonsingular lower triangular matrix. Then the probability of a given data set $\mathbf{b}$ is given by the Gaussian distribution

$$p(\mathbf{b}) = ((2\pi)^M |\mathbf{B}|)^{-1/2} \exp(-\tfrac{1}{2}\mathbf{b}^T\mathbf{B}^{-1}\mathbf{b}) \quad , \tag{200}$$

where $|\mathbf{B}|$ is the determinant of $\mathbf{B}$, and the normalized magnitude

$$\tau^2 = \mathbf{b}^T\mathbf{B}^{-1}\mathbf{b} = \|\mathbf{L}^{-1}\mathbf{b}\|^2 \tag{201}$$

indicates how unlikely the data set $\mathbf{b}$ is relative to the assumed priors. If the τ^2 computed from the data actually observed is too large, then the hypothesis that the observed data are consistent with the assumed priors should be rejected, and the source distribution should *not* be reconstructed using those priors. Note that τ^2 is a sufficient statistic for this hypothesis and thus that there can be no more effective test for misinformative priors using only the available data set.

Now define $\mathbf{x} = \mathbf{L}^{-1}\mathbf{b}$ and observe that $\mathsf{E}\,\mathbf{x} = 0$ and

$$\mathsf{E}\,\mathbf{x}\mathbf{x}^T = \mathsf{E}(\mathbf{L}^{-1}\mathbf{b})(\mathbf{b}^T\mathbf{L}^{-T}) = \mathbf{L}^{-1}\mathbf{B}\mathbf{L}^{-T} = \mathbf{I} \tag{202}$$

and so $\tau^2 = \|\mathbf{x}\|^2$ is distributed as a χ^2 random variable with M degrees of freedom, where M is the number of detectors, or the number of distinct measurements in $\mathbf{b}$. Thus, τ^2 has mean M and variance $2M$.

Now suppose that the data set is actually drawn from the assumed population and that τ^2_{crit} is taken as the critical value. That is, the priors will be rejected for this data set if the corresponding $\tau^2 > \tau^2_{\text{crit}}$, even though the data were actually drawn from the assumed population. That is, the false positive rate must be

$$\alpha = 1 - P(\tau^2_{\text{crit}} \mid M) \tag{203}$$

where $P(\chi^2 \mid \nu)$ is the cumulative distribution function for the standard χ^2 distribution with ν degrees of freedom and M is the number of detectors. In practice, an acceptable false positive rate α will be chosen first and τ^2_{crit} selected to yield the desired false positive rate. Figure 8.1 shows the false positive rate α as a function of τ^2_{crit} for various values of M; these curves were computed using the algorithm described in section 5.3.

A second approach is to consider how well the reconstruction fits the observed data and the assumed source priors.

The traditional metric for goodness of fit to the measurements is

$$\chi^2 = (\mathbf{b} - \mathbf{F}\hat{\mathbf{q}})^T\mathbf{\Sigma}^{-1}(\mathbf{b} - \mathbf{F}\hat{\mathbf{q}}) \quad , \tag{204}$$

which was defined in Section 6.3 and is the squared sum of the measurement residuals $\mathbf{r} = \mathbf{b} - \mathbf{F}\hat{\mathbf{q}}$, weighted by the noise covariance $\mathbf{\Sigma}$. If $\hat{\mathbf{q}}$ is computed by the minimum-norm least square (MNLS) method [106], then χ^2 has an ordinary χ^2 distribution; but if $\hat{\mathbf{q}}$ is computed by OCLIM, χ^2 has a generalized χ^2 distribution as was shown in Section 6.3.

The goodness of fit between the a priori source statistics and the reconstruction may be defined by the surprise

$$\rho^2 = \hat{\mathbf{q}}^T\mathbf{A}^{-1}\hat{\mathbf{q}} \quad , \tag{205}$$

which was defined in Section 6.4 and is the negative log likelihood of the reconstructed source distribution relative to the prior source statistics.

The sum $\chi^2 + \rho^2$ is a plausible metric for the goodness of fit to the data and priors together. Now observe that, for the OCLIM estimator,

$$\begin{aligned}
\chi^2 + \rho^2 &= (\mathbf{b} - \mathbf{F}\hat{\mathbf{q}})^T \mathbf{\Sigma}^{-1} (\mathbf{b} - \mathbf{F}\hat{\mathbf{q}}) + \hat{\mathbf{q}}^T \mathbf{A}^{-1} \hat{\mathbf{q}} \\
&= \mathbf{b}^T (\mathbf{I} - \mathbf{FH})^T \mathbf{\Sigma}^{-1} (\mathbf{I} - \mathbf{FH}) \mathbf{b} + \mathbf{b}^T \mathbf{H}^T \mathbf{A}^{-1} \mathbf{H} \mathbf{b} \\
&= \mathbf{b}^T (\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1} \mathbf{\Sigma} \mathbf{\Sigma}^{-1} \mathbf{\Sigma} (\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1} \mathbf{b} \\
&\quad + \mathbf{b}^T (\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1} \mathbf{F}\mathbf{A}\mathbf{A}^{-1}\mathbf{A}\mathbf{F}^T (\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1} \mathbf{b} \\
&= \mathbf{b}^T (\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1} \mathbf{b} \\
&= \|\mathbf{L}^{-1}\mathbf{b}\|^2 = \tau^2 \quad ! \qquad (206)
\end{aligned}$$

Thus it turns out that this sum is exactly the statistic τ^2 already discussed. That is, the test statistic τ^2 can also be viewed as a goodness of fit to the data and the priors jointly, expressed as the sum $\tau^2 = \chi^2 + \rho^2$ of a goodness of fit χ^2 to the data and a goodness of fit ρ^2 to the priors. Furthermore, τ^2 is (apart from scale), the negative logarithm of the joint probability

$$\begin{aligned}
p(\mathbf{b}, \hat{\mathbf{q}}) &= p(\mathbf{b} \mid \hat{\mathbf{q}})\, p(\hat{\mathbf{q}}) \\
&= ((2\pi)^M |\mathbf{\Sigma}|)^{-1/2} \exp(-\tfrac{1}{2}((\mathbf{b} - \mathbf{F}\hat{\mathbf{q}})^T \mathbf{\Sigma}^{-1} (\mathbf{b} - \mathbf{F}\hat{\mathbf{q}})) \\
&\quad \times ((2\pi)^N |\mathbf{A}|)^{-1/2} \exp(-\tfrac{1}{2}\hat{\mathbf{q}}^T \mathbf{A}^{-1} \hat{\mathbf{q}}) \quad . \qquad (207)
\end{aligned}$$

of the measurements $\mathbf{b}$ and the reconstructed source amplitudes $\hat{\mathbf{q}}$. Note, however, that even though τ^2 is distributed as a standard χ^2 random variable, both χ^2 and ρ^2 individually are distributed as generalized rather than standard χ^2 random variables.

8.3 Validation Protocol

Considering the previous analysis, the following is a suitable protocol for testing whether a given data set $\mathbf{b}$ is consistent with the assumed priors:

1. Choose a acceptable false positive rate α; this is the fraction of the data sets which are actually consistent with the priors but which will nevertheless be

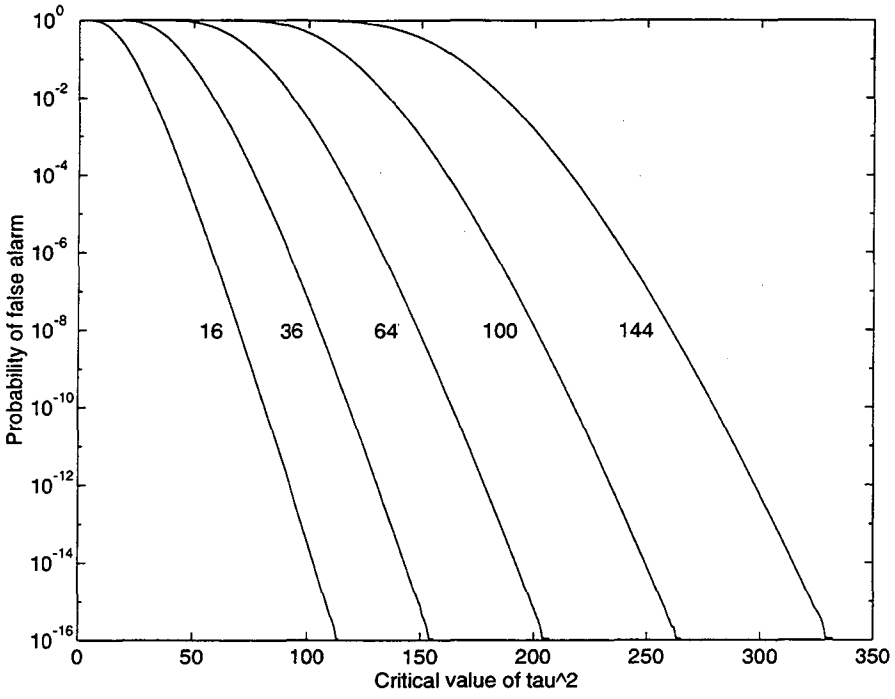


Figure 8.1 False positive rate. These curves show the false positive rate α as a function of of the critical value τ^2_{crit} for various values of M, the number of detectors. The false positive (or false alarm) rate is the fraction of source distributions that are actually consistent with the priors but that are reported as inconsistent.

classified as inconsistent. Figure 8.1 shows the false positive rate α as a function of τ^2_{crit} for various values of M.

2. Compute the corresponding critical value τ^2_{crit} such that

$$P(\tau^2_{\text{crit}} \mid M) = 1 - \alpha \tag{208}$$

where $P(\chi^2 \mid \nu)$ is the cumulative distribution function for the χ^2 distribution with ν degrees of freedom and M is the number of detectors.

3. For the given data set, compute

$$\tau^2 = \|\mathbf{L}^{-1}\mathbf{b}\|^2 = \chi^2 + \rho^2 \quad . \tag{209}$$

4. If $\tau^2 > \tau^2_{\text{crit}}$, classify the given data as inconsistent with the priors. Otherwise, classify it as consistent and proceed with the reconstruction.

8.4 Detectability of an Unexpected Source

It is also of interest to know how reliably the τ^2 test detects the presence of a source inconsistent with the assumed priors. To be specific, suppose that there is an extra source $\vec{J}_{\mathrm{X}}(\vec{r})$, in addition to those sources allowed by the priors. That extra source can be observed only through its effect on the measurement vector $\mathbf{b}$. Suppose that $\vec{J}_{\mathrm{X}}(\vec{r})$ acting alone gives the measurement vector $\mathbf{b}_{\mathrm{X}}$. Then the total measurement vector $\mathbf{b}$ including the prior sources, extra source, and noise will be normally distributed with mean $\mathbf{b}_{\mathrm{X}}$ and variance $\mathbf{B} = \mathbf{FAF}^T + \mathbf{\Sigma} = \mathbf{LL}^T$.

It follows that the test statistic τ^2 is distributed as a non-central χ^2 random variable with M degrees of freedom and non-centrality parameter

$$\tau_{\mathrm{X}}^2 = \|\mathbf{L}^{-1}\mathbf{b}_{\mathrm{X}}\|^2 \quad . \tag{210}$$

Intuitively, τ_{X}^2 measures the mismatch between the extra field $\mathbf{b}_{\mathrm{X}}$ and the assumed priors. An extra source cannot be detected if its effects are consistent with the measurement statistics defined by the priors, no matter how much its source distribution may differ from the expected source distribution. Note also that any source distribution in the null space of $\mathbf{F}$ has no effect on the measurements and cannot be detected.

Any given values of τ_{X}^2 and τ_{crit}^2 define a false positive probability α that a consistent source will be misclassified as inconsistent and a false negative probability β that an inconsistent source will be misclassified as consistent. Plotting curves of α versus the detection probability $1-\beta$ yields a set of receiver operating characteristic (ROC) curves for different values of τ_{X}^2; these curves define the possible tradeoffs between false positive (or false alarm) and false negative errors as a function of τ_{X}^2. The particular curves shown in Figure 8.2 are for $M = 144$.

8.5 Simulation Results

To test the reliability with which the τ^2 test distinguishes informative and misinformative priors, a series of simulations was done using the same test configuration as in Section 7.3. The 28 sources in the central region A were assigned a source

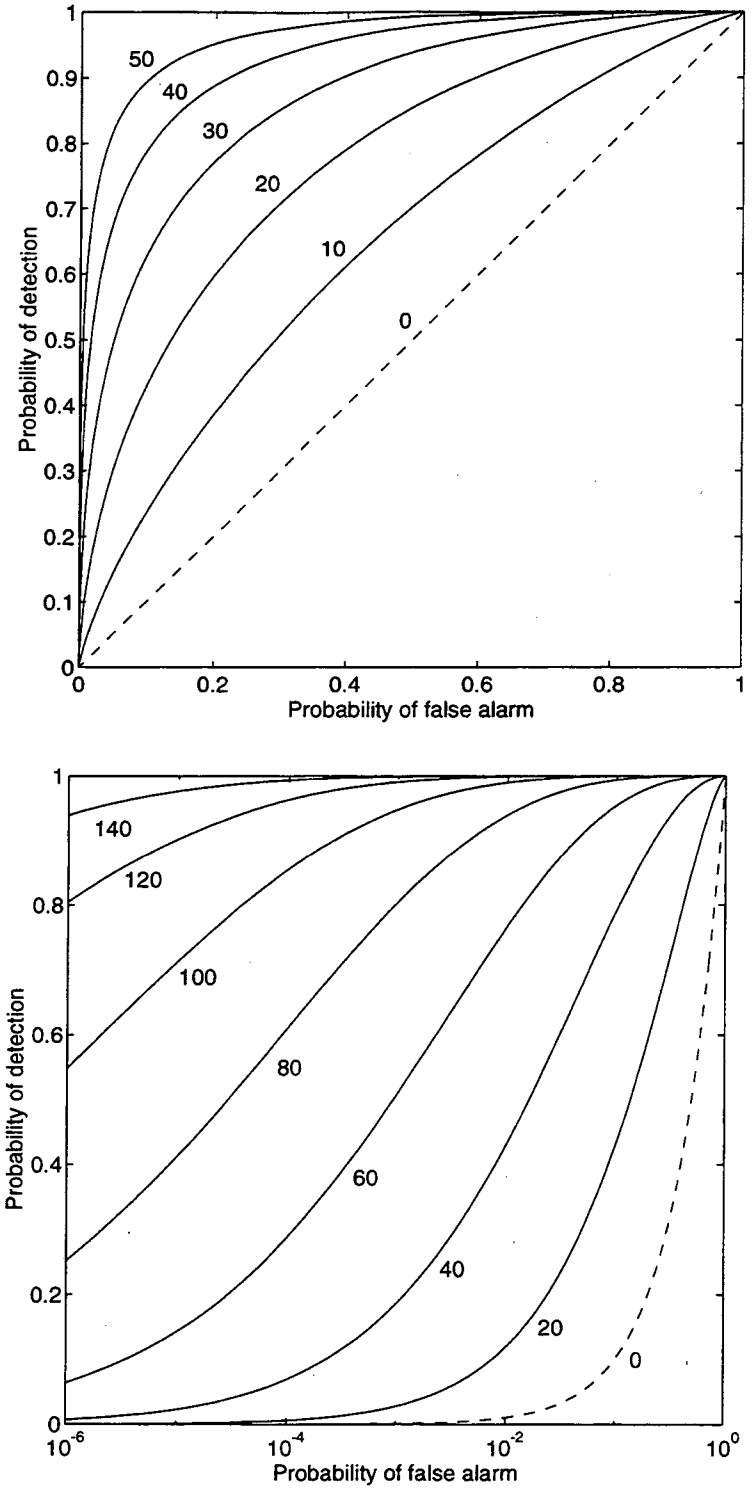


Figure 8.2 Typical ROC curves. These curves show the possible tradeoffs between false positive and false negative errors for various values of τ_x^2. The curves in the bottom plot use the same data as the top plot but use a logarithmic scale for the false positive probability to improve the visibility of false positive rates below 1%.

variance $\alpha_A^2 = 1$; sources in this region were considered consistent with the priors. The remaining sources, in region B, were assigned a different source variance $\alpha_B^2 = 0.1$; they were considered inconsistent with the priors.

Noise amplitudes taken from independent normal distributions with mean zero and variance σ^2 were added to each field measurement and two different noise models were considered. In the low-noise model $\sigma^2 = 10^{-16}$ and the signal-to-noise ratio is about 167 dB; this is much better than most biomagnetic experiments. In the high-noise model $\sigma^2 = 10^{-8}$ and the signal-to-noise ratio is about 7 dB; this is a more typical number for biomagnetic imaging.

The scatter diagrams in Figures 8.3 and 8.4 show the χ^2 and ρ^2 values for a series of consistent and inconsistent reconstructions, each with a single dipole source of magnitude 8. The pluses (+) correspond to sources in region A, which are consistent with the priors. The circles (o) correspond to sources in region B, which are inconsistent with the priors.

Figure 8.3 uses the low-noise model. The diagonal line corresponds to the critical value $\tau^2 = 200$ or $\alpha = 0.001$, and almost perfectly distinguishes the consistent sources from the inconsistent ones. The two exceptions in the lower lefthand corner are actually inconsistent with the priors but are misclassified as consistent; the source locations and reconstructions of these misclassified sources are shown in the next section. Both of the corresponding sources are distant from the detector array and close to consistent sources; it is possible that the detector array simply lacks the resolving power to distinguish these inconsistent sources from the nearby consistent sources.

Figure 8.4 uses the high-noise model. The diagonal line corresponds to $\tau^2 = 170$, or $\alpha = 0.1$, but does not separate the consistent and inconsistent sources; thus the detector array cannot resolve the different sources at this SNR.

It is also possible to plot the value of τ_{x}^2 as a function of position for an extra source of some specified strength. Figure 8.5 shows two such plots for a dipole of magnitude 8 in the model problem. These plots indicate the positions at which such a source would probably be detected, or not detected. The value $\tau_{\mathrm{x}}^2 = 100$ provides a false positive rate $\alpha = 0.01$ and false negative rate $\beta = 0.02$ and can be

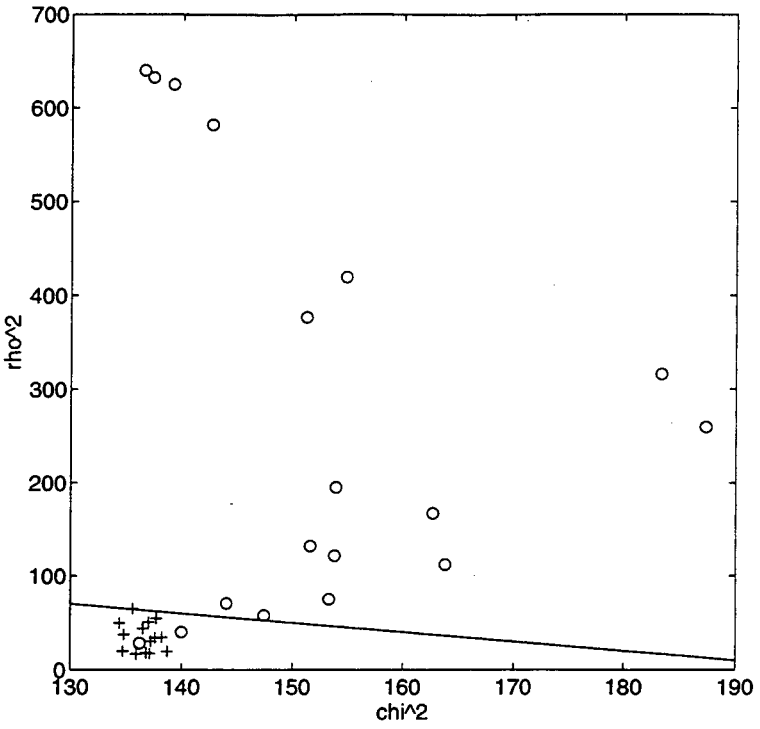


Figure 8.3 Residual and surprise for model problem—low noise. The residual (χ^2) and surprise (ρ^2) are plotted for several sources consistent (+) and inconsistent (o) with the priors in the model problem with SNR = 167 dB. The diagonal line corresponds to the critical value $\tau^2 = 200$ or $\alpha = 0.001$, and almost perfectly distinguishes the consistent sources from the inconsistent ones.

taken to distinguish the detectable extra sources from the undetectable.

The top plot uses the low-noise model. Most or all of the unexpected sources in the upper portions of region B would be detected, and some of those in the lower portions of region B.

The bottom plot uses the high-noise model. Only unexpected sources near the top corners of the source plane are likely to be detected; deeper unexpected sources are not distinguishable from the expected ones.

8.6 Sample Reconstructions

Figures 8.6 through 8.8 show some sample reconstructions using the low-noise model. In each reconstruction, the true source distribution consists of a single source dipole pointing out of the page with amplitude 8 at the location marked

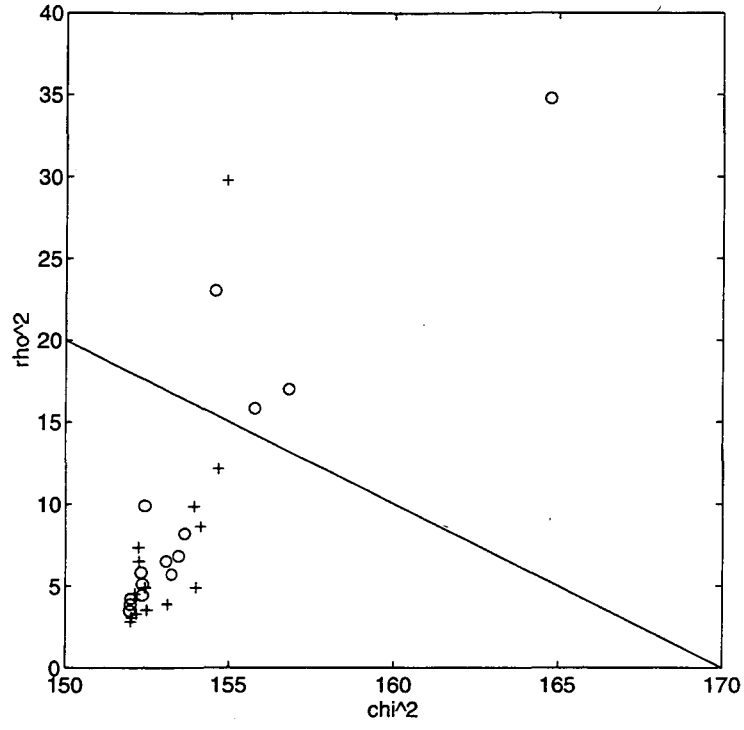


Figure 8.4 Residual and surprise for model problem—high noise. The residual (χ^2) and surprise (ρ^2) are plotted for several sources consistent (+) and inconsistent (o) with the priors in the model problem with SNR = 7 dB. The diagonal line corresponds to $\tau^2 = 170$, or $\alpha = 0.1$, but does not separate the consistent and inconsistent sources; thus the detector array cannot resolve the different sources at this SNR.

with the black dot. Grey shading indicates areas of current flow into the page. Figure 8.6 shows two examples of consistent sources correctly classified as consistent; figure 8.7 shows two inconsistent sources correctly classified as inconsistent; and figure 8.8 shows two consistent sources incorrectly classified as inconsistent. Note that in both of the misclassified cases, the unexpected source is both far from the detectors (which means a poor SNR) and close to one of the expected sources; a possible interpretation of this result is that the magnetometer array simply lacks the resolving power to distinguish the expected and unexpected sources.

8.7 Conclusions

The preceeding results have not provided a completely satisfactory solution to the problem of misinformative priors in biomagnetic source imaging. We have seen that

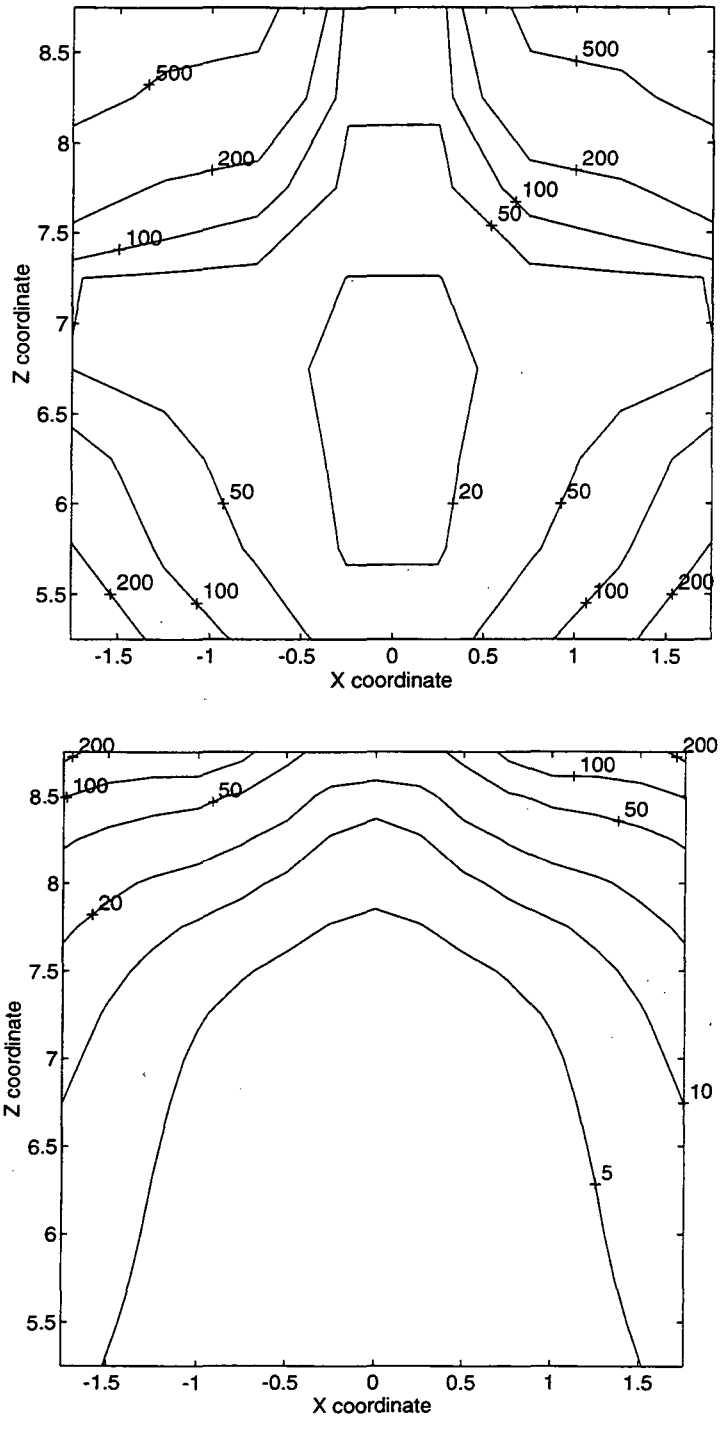


Figure 8.5 Source detectability as a function of position. Both these plots show the value of τ_x^2 as a function of position for an unexpected dipole source of magnitude 8. The top plot is for the low-noise model and the bottom plot is for the high-noise model. If $\tau_x^2 > 100$ at some position, then an unexpected dipole source of magnitude 8 at that position can be detected with better than 98% probability and no more than 1% false positives.

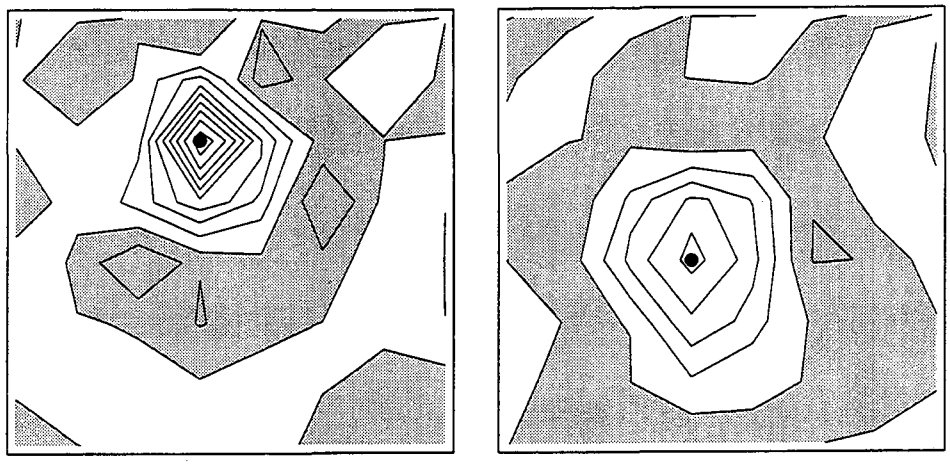


Figure 8.6 Examples of sources correctly classified as consistent.

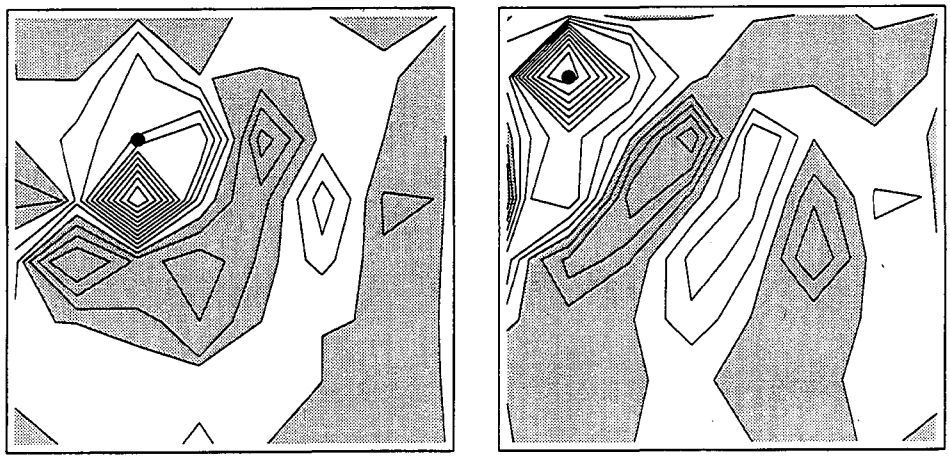


Figure 8.7 Examples of sources correctly classified as inconsistent.

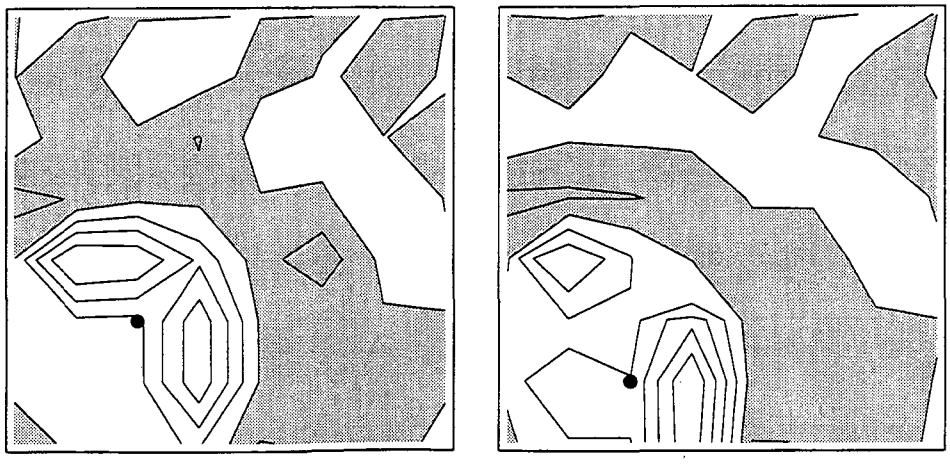


Figure 8.8 Examples of sources incorrectly classified as consistent.

true source distributions inconsistent with the assumed priors can generate artifactual reconstructions. Under favorable conditions of low noise and good separation between the unexpected and assumed source locations, the τ^2 test will often detect misinformative priors. Under unfavorable conditions, the τ^2 test does not reliably distinguish informative and misinformative priors.

According to Figure 3.3, the SNR in biomagnetic imaging can range from -21 dB to 54 dB, depending on the signal source, noise level, and bandwidth. More work is needed to determine whether or not the τ^2 test reliably discriminates between informative and misinformative priors in realistic biomagnetic imaging problems. Since the source model used in the chapter was too simplified to be realistic, it is also necessary to study the τ^2 test in realistic source models.

In the worst case, the τ^2 test will not reliably detect misinformative priors. The best SNR predicted in Figure 3.3 for the visual evoked response and evoked cortical activity is only 9 dB, so it is likely that the τ^2 test will in fact fail for these signal sources, although more work is needed to confirm this. In this case, the inability to detect misinformative priors, combined with the non-uniqueness of the biomagnetic inverse problem, means that while biomagnetic source reconstructions may be used to suggest source locations to be tested by other means, they cannot be used by themselves to prove the existence of a source, since there may be another reconstruction which fits the given data equally well but which does not include that source. Similarly, the absense of a source at some location in a biomagnetic source reconstruction does not prove that no such source is present.

An alternative possibility is to use prior information more assertively. Suppose that by fMRI or PET imaging we find that cerebral metabolic activity during a cognitive task is elevated in only a few regions. Then the hypothesis that activity is restricted to these regions becomes a hypothesis that can be tested using the τ^2 test on MEG data collected for the same task. The fact that only a few sources are allowed in the source model should improve the power of the τ^2 test even if the SNR remains poor.

Chapter 9

Optimal Magnetometer Design

As was noted in Chapter 6, the mean reconstruction error $\eta^2 = \mathsf{E}\,\|\mathbf{q} - \hat{\mathbf{q}}\|^2$ and the a posteriori source variances $\hat{\alpha}^2_{nn}$ do not depend on the actual field measurements $\mathbf{b}$ but only on the source and detector configuration and on the assumed source and noise priors. Thus the mean reconstruction error η^2 can be computed directly from the given values of the forward transfer matrix $\mathbf{F}$, the source covariance $\mathbf{A}$, and the noise covariance $\mathbf{\Sigma}$. This direct calculation of the reconstruction error is much faster than Monte Carlo simulation and can be done on a workstation class computer in a few minutes. The facts that η^2 depends on the detector configuration (and other priors), that it relates directly to the reconstruction accuracy, that it does not depend on the actual measured data, and that it can be computed inexpensively suggest that the reconstruction error η^2 will be useful as a cost function in the optimal design of magnetometer arrays.

This chapter demonstrates the use of the reconstruction error to select the optimal array width in a model problem and investigates the effects of number of detectors and source-to-detector distance on the reconstruction accuracy.

9.1 An Example Source/Detector Configuration

We will analyze the example source and magnetometer configuration shown in Figure 9.1. The source volume is $12 \times 12 \times 12$ cm^3 and contains a $4 \times 4 \times 4$ cubical grid of horizontal dipole pairs; all sources have the same RMS source amplitude $\alpha = 1$ μA-m. Each horizontal plane in the source volume contains a 4×4 array of in-plane dipole pairs as shown in Figure 9.2; the use of two orthogonal dipoles allows the representation of a dipole with arbitrary orientation and magnitude within

the plane. The detector plane contains a 12×12 grid of detectors sampling the vertical component of the field; all detectors have the same expected RMS noise amplitude $\sigma = 30$ fT. The nominal array width is 23 cm; the nominal source depth, or distance from the detector array to the top of the source volume, is 2 cm.

The following sections will explore the effects of varying the array width, the source depth, and the number of detectors used.

9.2 Error versus Array Width and Source Depth

The reconstruction error η^2 is the sum over all sources of the expected value of the squared reconstruction error for that source. That is,

$$\eta^2 = \sum_{n=1}^{N} \hat{\alpha}_{nn}^2 \tag{211}$$

where $\hat{\alpha}_{nn}^2 = \mathsf{E}(q_n - \hat{q}_n)^2$ is the expected squared error for the nth source. It is convenient to divide by the a priori source variance α^2 and consider the relative error η^2/α^2.

Figure 9.3 shows how the total relative reconstruction error η^2/α^2 varies as a function of the array width and the source depth. For every source depth considered, there is an optimal array width and that width increases as the source depth increases. The particular optimal widths found here apply only to the particular number of detectors, source variance, and noise variance used to compute them.

9.3 Standard Error for Each Source

The standard error $\hat{\alpha}_{nn}$ is the RMS value of the reconstruction error for the nth source; it shows how the total error η^2 is distributed over the various sources.

Figure 9.4 shows the standard error $\hat{\alpha}_{nn}$ for each source, using the nominal values for array width, source depth, and number of detectors. The a priori standard deviation α_{nn} is shown as the dotted line across the top of the figure. For this configuration, the reconstruction errors for the sources in the top plane are about 5% of their a prior uncertainties. For the next lower plane, the errors are about 40–80% of the a priori uncertainties; for the bottom two planes, the errors are hardly

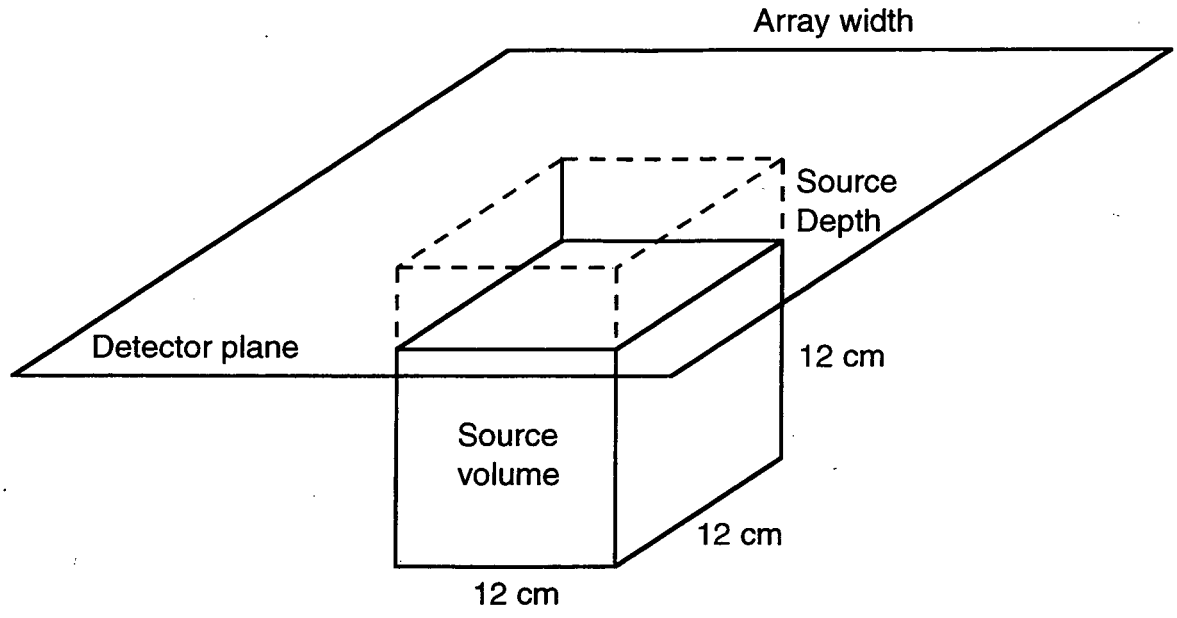


Figure 9.1 Example source and detector configuration. The source volume contains a $4 \times 4 \times 4$ cubical array of horizontal dipoles sources, each represented as a pair of orthogonal components. The detector plane is square with a nominal width of 23 cm and contains a nominal 12×12 array of magnetometers, each of which measures the vertical component of the field. The nominal source depth, or distance from the detector plane to the top plane of sources, is 2 cm. The array width, number of detectors, and source depth will be varied from their nominal values to examine the effects.

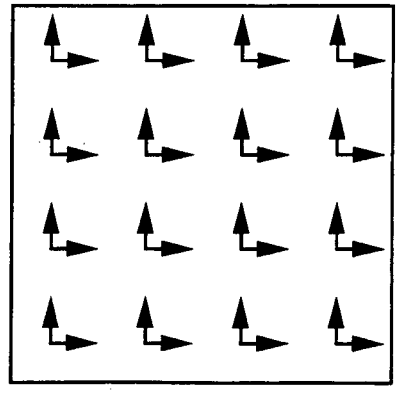


Figure 9.2 Typical source plane. Typical horizontal plane of sources in the source volume. Each source position contains a horizontal current dipole represented by two orthogonal components.

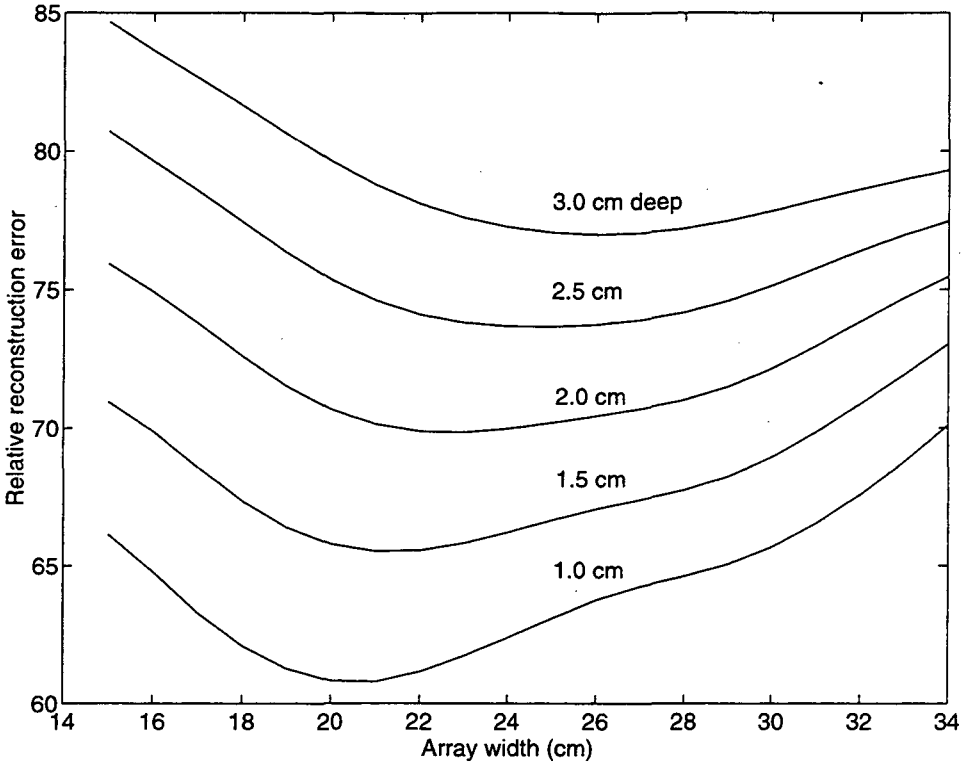


Figure 9.3 Total reconstruction error. Total reconstruction error versus array width and source depth for a 12×12 array of detectors. The source depth is the distance from the detector array to the top plane of sources.

less than the a priori uncertainties. That is, this detector array (for the assumed priors) gives good reconstructions for the topmost plane of sources, mediocre results for the second plane, and poor results for the bottom two planes.

9.4 Error vs Width and Number of Detectors

Figure 9.5 shows the total relative reconstruction error versus array width for five different numbers of detectors ranging from an 8×8 square grid to a 16×16 square grid. There is an optimal array width for each size of grid; that width increases as the number of detectors increases.

Note, however, that both the 8×8 and 10×10 curves display a bimodal characteristic. This is due to the accidental alignment of source and detector positions as shown in the plan views of Figure 9.6. The plus signs mark detector positions; the open circles mark source positions. The lefthand plot corresponds to the minimum

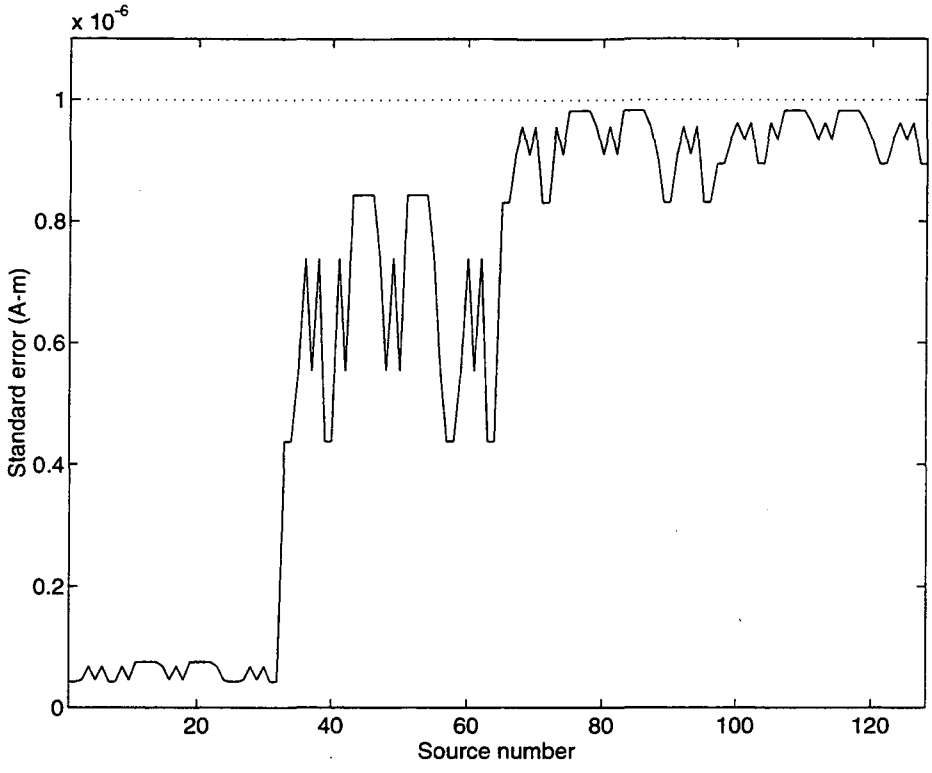


Figure 9.4 Standard error of each source. Standard error for each source using the nominal detector configuration and priors. Sources 1–32 are in the top plane, 33–64 in the next plane down, 65–96 in the third plane down, and 97–128 in the bottom plane. (Each source position includes two orthogonal components, each of which carries a distinct source number in this figure.) Sources near the detector plane have smaller error, as do source near the boundaries of the source cube.

error (at width 19 cm) for an 8×8 array; the sources generally fall between the detectors. The righthand plot corresponds to the local maximum at 23 cm; the sources fall almost directly under the detector positions.

Since dipole sources are used only for convenience in approximating a continuous unknown current distribution, this variation in reconstruction error must be considered an artifact of the simple dipole source model. The use of point detectors also contributes to the problem. More work is needed to create source and detector models that are immune to these accidental variations.

Figure 9.7 shows the optimal 8×8 and 16×16 arrays. The optimal width increases as the number of detectors increases but less than proportionally to the number of detectors per side; the optimal detector spacing decreases.

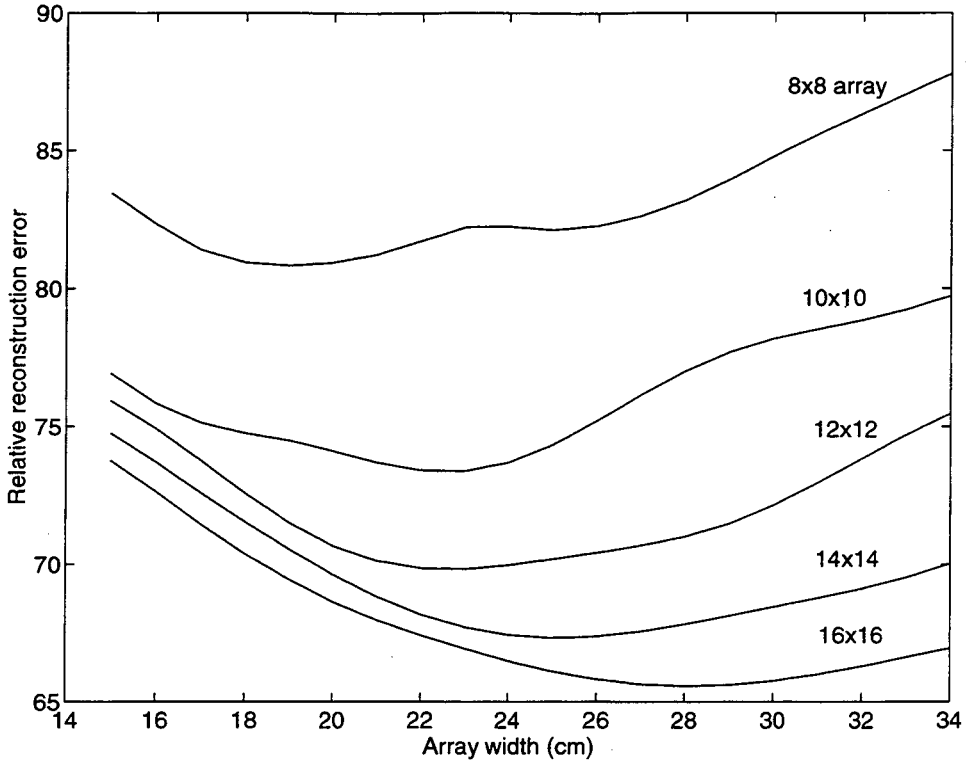


Figure 9.5 Total reconstruction error versus array width and number of detectors.. The optimal array width increases as the number of detectors increases. The bimodal characteristic of the 8×8 and 10×10 curves is due to accidental alignments between the source and detector locations. See Figure 9.6.

9.5 Discussion

The reconstruction error is potentially useful as a quality metric for magnetometer array design. Arbitrary source and detector positions are allowed. Correlated noise can be used; this allows modelling the effects of external magnetic interference. The method should generalize to distributed (non-dipole) current sources and arbitrary detector coil shapes.

Accidental alignments between dipole current sources and point detectors can distort the computed reconstruction error.

The optimal width increases as the array-to-source distance increases. The optimal width of a planar magnetometer array is subproportional to the number of detectors used per side.

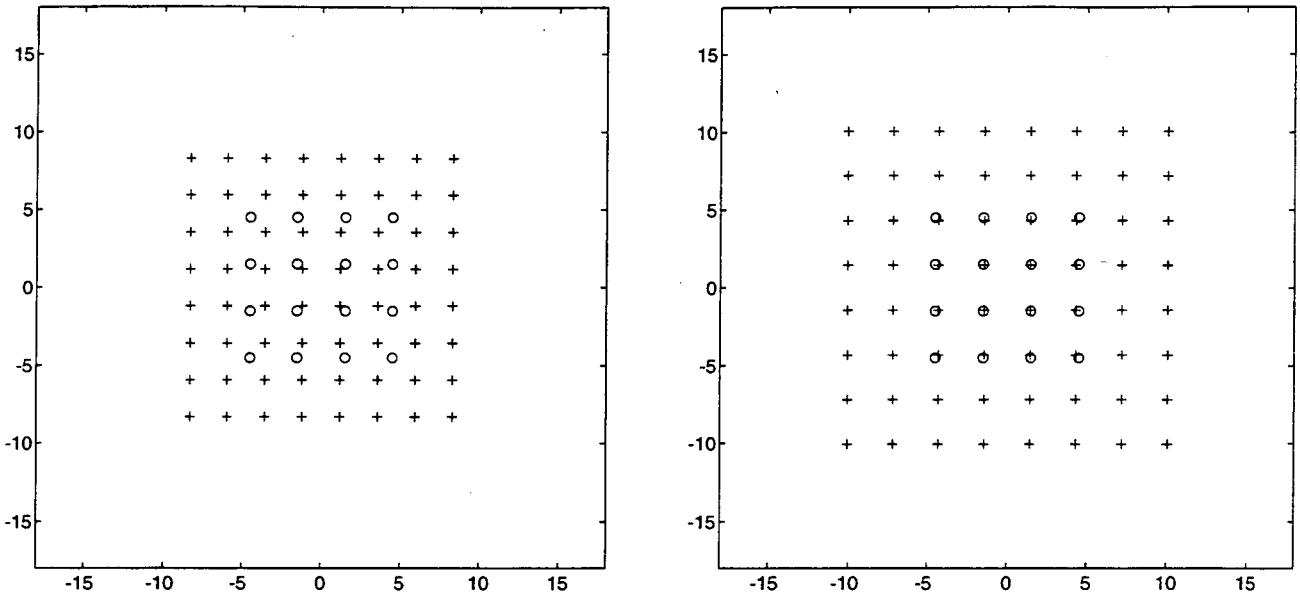


Figure 9.6 Source and detector alignments. The lefthand plot shows the source (o) and detector (+) locations in plan view for the minimum-error width in the 8×8 configuration. The sources fall between the detector locations. The righthand plot shows the source and detector locations for a configuration with local maximum error. The sources fall directly below the detector locations.

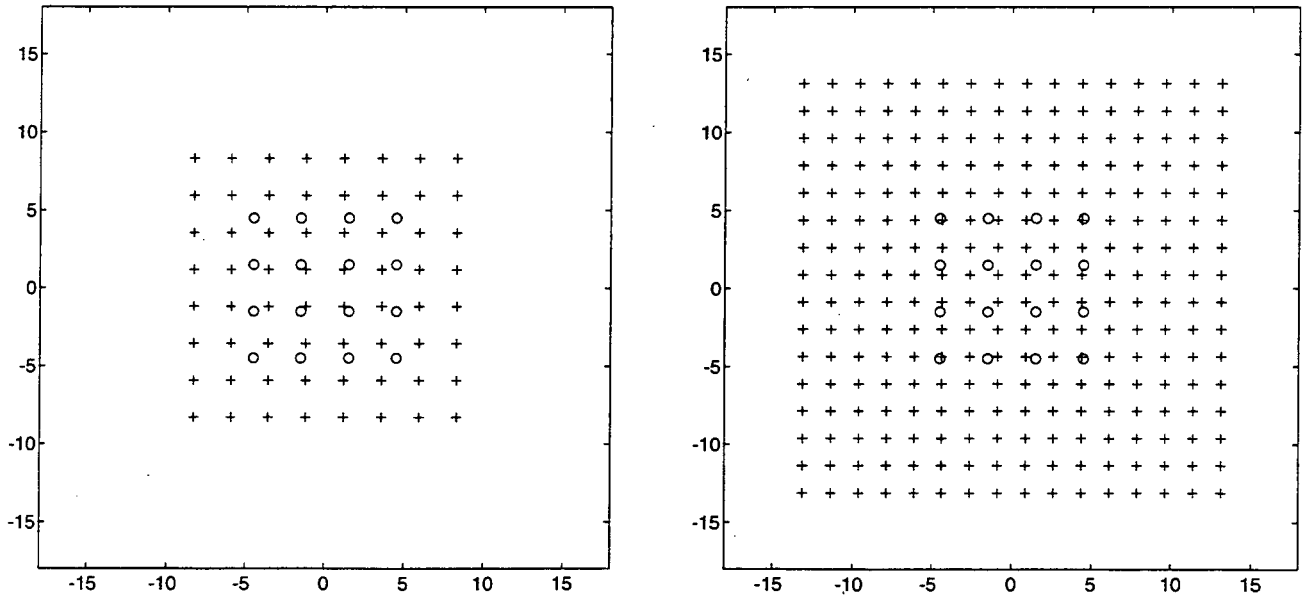


Figure 9.7 Optimal array width versus number of detectors. The lefthand plot shows the optimal 8×8 array; the righthand plot the optimal 16×16 array. The optimal width increases less than proportionally to the number of detectors per side and the optimal detector spacing decreases.

Chapter 10

Dynamic Reconstruction

The previous chapters have developed and demonstrated the OCLIM algorithm for the instantaneous biomagnetic source imaging problem. The general case, however, is that the source currents and generated magnetic fields vary with time.

If the current distribution changes with time, then a sequence of field measurements must be inverted to yield a sequence of current distributions. OCLIM generalizes easily in principle to this dynamic problem but requires large data sets to describe the priors and long computations to reconstruct the dynamic current distribution.

The purpose of this chapter is to discuss methods for simplifying the dynamic reconstruction problem by assuming and exploiting special structure in the priors and sometimes by approximating the optimal reconstruction filter.

We will be making heavy use of the Cholesky decomposition and so it is convenient to introduce the following notation rather than to introduce a new symbol each time we take the Cholesky decomosition of some matrix. Let $\mathbf{A}$ be a symmetric positive definite matrix. Then $\mathbf{A}^C$ will denote its Cholesky factor; that is, $\mathbf{A}^C$ is the unique lower triangular matrix such that $\mathbf{A} = \mathbf{A}^C(\mathbf{A}^C)^T$.

10.1 Dynamic Imaging of Magnetic Sources

Suppose that we have a sequence of field measurements on multiple detectors and wish to reconstruct a sequence of source amplitudes at multiple source locations. The field measurements can be represented by concatenating the instantaneous measurement vectors into a long column vector $\tilde{\mathbf{b}}$; a long source vector $\tilde{\mathbf{q}}$ and a long noise vector $\tilde{\mathbf{w}}$ are defined similarly. The spatiotemporal source covariance is

defined by $\tilde{\mathbf{A}} = \mathsf{E}\,\tilde{\mathbf{q}}\tilde{\mathbf{q}}^T$ and the noise covariance by $\tilde{\mathbf{\Sigma}} = \mathsf{E}\,\tilde{\mathbf{w}}\tilde{\mathbf{w}}^T$, where E denotes the expectation. Assuming that the instantaneous forward transfer matrix $\mathbf{F}$ is constant in time and that propagation delays are negligible, the forward equation becomes

$$\tilde{\mathbf{b}} = (\mathbf{I} \otimes \mathbf{F})\tilde{\mathbf{q}} + \tilde{\mathbf{w}} \quad , \tag{212}$$

where $\otimes$ is the Kronecker product. This has the same form as the instantanteous forward problem, and OCLIM yields the minimum mean square error estimator

$$\tilde{\mathbf{H}} = \tilde{\mathbf{A}}(\mathbf{I} \otimes \mathbf{F})^T[(\mathbf{I} \otimes \mathbf{F})\tilde{\mathbf{A}}(\mathbf{I} \otimes \mathbf{F})^T + \tilde{\mathbf{\Sigma}}]^{-1} \quad , \tag{213}$$

the optimal estimate

$$\hat{\mathbf{q}} = \tilde{\mathbf{H}}\tilde{\mathbf{b}} = \tilde{\mathbf{A}}(\mathbf{I} \otimes \mathbf{F})^T[(\mathbf{I} \otimes \mathbf{F})\tilde{\mathbf{A}}(\mathbf{I} \otimes \mathbf{F})^T + \tilde{\mathbf{\Sigma}}]^{-1}\tilde{\mathbf{b}} \quad , \tag{214}$$

and the test statistic for misinformative priors

$$\tau^2 = \tilde{\mathbf{b}}^T[(\mathbf{I} \otimes \mathbf{F})\tilde{\mathbf{A}}(\mathbf{I} \otimes \mathbf{F})^T + \tilde{\mathbf{\Sigma}}]^{-1}\tilde{\mathbf{b}} \quad . \tag{215}$$

For a reasonably sized problem—100 sources, 100 detectors, and 100 time instants—the prior covariance matrices $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{\Sigma}}$ alone require 400 MB of memory and the direct computation uses about 0.3 trillion floating point operations. In addition, the covariance matrices $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{\Sigma}}$ must be determined experimentally and it will be quite difficult to experimentally determine so many distinct parameters.

The standard errors of the reconstructed source amplitudes are computed from the diagonal of the a posteriori covariance matrix

$$\hat{\mathbf{A}} = \tilde{\mathbf{A}} - \tilde{\mathbf{A}}(\mathbf{I} \otimes \mathbf{F})^T[(\mathbf{I} \otimes \mathbf{F})\tilde{\mathbf{A}}(\mathbf{I} \otimes \mathbf{F})^T + \tilde{\mathbf{\Sigma}}]^{-1}(\mathbf{I} \otimes \mathbf{F})\tilde{\mathbf{A}} \quad . \tag{216}$$

A direct computation of $\hat{\mathbf{A}}$ requires an additional 1.3 trillion floating point operations.

Making some assumptions about the form of the priors can reduce the computational and storage cases significantly and the purpose of this chapter is to develop possible ways of doing so. The following cases are considered.

1. The source and noise amplitudes are stationary and uncorrelated in time.

2. The spatiotemporal source and noise covariances are separable into the products of spatial and temporal covariances.

3. The source and noise covariances are separable and stationary.

4. The source statistics are stationary; the noise statistics are stationary, uncorrelated between detectors, and uniform across detectors.

10.2 Stochastic Estimation of Errors

The cost of computing the standard errors may be reduced by using Monte Carlo simulation. Consider first the instantaneous case. Let $\mathbf{z}_1$ and $\mathbf{z}_2$ be independent standard normal random vectors and define

$$\mathbf{x} = \mathbf{F}\mathbf{A}^C\mathbf{z}_1 + \mathbf{\Sigma}^C\mathbf{z}_2 \qquad \text{and} \qquad \mathbf{y} = \mathbf{H}\mathbf{x} \quad . \tag{217}$$

Then

$$\mathsf{E}\,\mathbf{x} = 0 \qquad \text{and} \qquad \mathsf{E}\,\mathbf{x}\mathbf{x}^T = \mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma} \tag{218}$$

and so $\mathbf{x}$ has exactly the assumed statistics of the measurement vector $\mathbf{b}$ and $\mathbf{y}$ has exactly the statistics of the solution $\hat{\mathbf{q}}$. Thus the mean square residual

$$\mathsf{E}\,\chi^2 = \mathsf{E}(\mathbf{b} - \mathbf{F}\hat{\mathbf{q}})^T\mathbf{\Sigma}^{-1}(\mathbf{b} - \mathbf{F}\hat{\mathbf{q}}) \tag{219}$$

can be estimated by generating some number R of realizations of $\mathbf{z}_1$ and $\mathbf{z}_2$, computing the corresponding realizations $\mathbf{x}_r$ and $\mathbf{y}_r$, and computing the estimate

$$\mathsf{E}\,\chi^2 \cong \frac{1}{R}\sum_r (\mathbf{x} - \mathbf{F}\mathbf{y})^T\mathbf{\Sigma}^{-1}(\mathbf{x} - \mathbf{F}\mathbf{y}) \tag{220}$$

Furthermore, it is possible to estimate the standard errors from the same set of realizations. Consider

$$\begin{aligned}
\mathbf{A} - \mathsf{E}\,\mathbf{y}\mathbf{y}^T &= \mathbf{A} - \mathbf{H}(\mathsf{E}\,\mathbf{x}\mathbf{x}^T)\mathbf{H}^T \\
&= \mathbf{A} - \mathbf{A}\mathbf{F}^T(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1}(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1}\mathbf{F}\mathbf{A} \\
&= \mathbf{A} - \mathbf{A}\mathbf{F}^T(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1}\mathbf{F}\mathbf{A} \\
&= \hat{\mathbf{A}} \quad .
\end{aligned} \tag{221}$$

That is, the a posteriori covariance $\hat{\mathbf{A}}$ may be estimated by computing

$$\hat{\mathbf{A}} \cong \mathbf{A} - \frac{1}{R}\sum_r \mathbf{y}_r \mathbf{y}_r^T \quad . \tag{222}$$

In many cases only the diagonal elements $\hat{\alpha}_{nn}^2$ of $\hat{\mathbf{A}}$ are required; they can be computed even more simply by

$$\hat{\alpha}_{nn}^2 \cong \alpha_{nn}^2 - \frac{1}{R}\sum_r y_{rn}^2 \tag{223}$$

where α_{nn}^2 is the nth diagonal element of $\mathbf{A}$ and y_{rn} is the nth element of the rth realization of $\mathbf{y}$. Since high accuracy is not needed for the standard errors, a few realizations should suffice.

The extension to the time series case involves only extending the vectors $\mathbf{z}_1$, $\mathbf{z}_2$, $\mathbf{x}$, and $\mathbf{y}$ to the appropriate length and computing

$$\mathbf{x} = (\mathbf{I} \otimes \mathbf{F})\tilde{\mathbf{A}}^C \mathbf{z}_1 + \tilde{\mathbf{\Sigma}}^C \mathbf{z}_2 \tag{224a}$$

$$\mathbf{y} = \tilde{\mathbf{H}}\mathbf{x} \tag{224b}$$

$$\mathsf{E}\,\chi^2 = \frac{1}{R}\sum_r (\mathbf{x} - (\mathbf{I} \otimes \mathbf{F})\mathbf{y})^T \tilde{\mathbf{\Sigma}}^{-1} (\mathbf{x} - (\mathbf{I} \otimes \mathbf{F})\mathbf{y}) \tag{224c}$$

$$\hat{\mathbf{A}} = \tilde{\mathbf{A}} - \frac{1}{R}\sum_r \mathbf{y}_r \mathbf{y}_r^T \quad . \tag{224d}$$

The form of $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{\Sigma}}$ should, of course, be chosen to facilitate computation of their Cholesky factors. Under assumptions to be discussed later, it is sufficient to take the time average of a single realization.

10.3 Instantaneous Reconstructions

One simple model for the priors is to assume that the source and noise covariances are constant over time and that there are no correlations among source (or noise) amplitudes at different times. Then the spatiotemporal covariances can be written in the simpler forms

$$\tilde{\mathbf{A}} = \mathbf{I} \otimes \mathbf{A} \qquad \text{and} \qquad \tilde{\mathbf{\Sigma}} = \mathbf{I} \otimes \mathbf{\Sigma} \tag{225}$$

where $\mathbf{A}$ is the instantaneous source covariance matrix and $\mathbf{\Sigma}$ is the instantaneous noise covariance matrix. The optimal spatiotemporal reconstruction filter simplifies to

$$\tilde{\mathbf{H}} = \mathbf{I} \otimes \mathbf{A}\mathbf{F}^T(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1} \quad , \tag{226}$$

which means simply that an instantaneous OCLIM reconstruction is to be done independently at each time instant. The test statistic is

$$\tau^2 = \tilde{\mathbf{b}}^T[\mathbf{I} \otimes (\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1}]\tilde{\mathbf{b}} \tag{227}$$

which is simply a summation of the test statistics for each instant. The a posteriori covariance is

$$\hat{\mathbf{A}} = \mathbf{I} \otimes [\mathbf{A} - \mathbf{A}\mathbf{F}^T(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1}\mathbf{F}\mathbf{A}] \quad , \tag{228}$$

which is the instantaneous a posteriori convariance repeated at each time instant.

This approach is fast but inaccurate, since biomagnetic sources are ordinarily correlated both in time and in space. Sources in the heart and brain often remain active for tens or hundreds of milliseconds; ignoring these correlations reduces the accuracy of the reconstruction.

10.4 Priors Separable in Time and Space

A less drastic simplification is to assume that the spatiotemporal source covariance is separable into the product of a spatial covariance and a temporal covariance. That is,

$$\tilde{A}(x, x', t, t') = \mathsf{E}\, q(x,t)\, q(x',t') = A(x,x') \cdot S(t,t') \quad . \tag{229}$$

A (simplified) physical interpretation is that the shape of the temporal power spectrum is the same for all sources, although the magnitude may vary. Similarly, suppose that the spatiotemporal noise covariance is separable. Then the priors can be written as

$$\tilde{\mathbf{A}} = \mathbf{S} \otimes \mathbf{A} \qquad \text{and} \qquad \tilde{\mathbf{\Sigma}} = \mathbf{N} \otimes \mathbf{\Sigma} \quad , \tag{230}$$

where $\mathbf{A}$ is an instantaneous source correlation matrix (between sources) which is independent of time, $\mathbf{\Sigma}$ is an instantaneous noise correlation matrix (between detectors) which is independent of time, $\mathbf{S}$ is a temporal correlation matrix which

is the same for all sources, and $\mathbf{N}$ is a temporal correlation matrix which is the same for all detectors. Then the optimal source estimate becomes

$$\hat{\mathbf{q}} = (\mathbf{S} \otimes \mathbf{AF}^T)[\mathbf{S} \otimes \mathbf{FAF}^T + \mathbf{N} \otimes \mathbf{\Sigma}]^{-1}\tilde{\mathbf{b}} \quad . \tag{231}$$

The test statistic is

$$\tau^2 = \tilde{\mathbf{b}}^T[(\mathbf{S} \otimes \mathbf{FAF}^T + \mathbf{N} \otimes \mathbf{\Sigma}]^{-1}\tilde{\mathbf{b}} \tag{232}$$

and the a posteriori covariance is

$$\hat{\mathbf{A}} = \mathbf{S} \otimes \mathbf{A} - (\mathbf{S} \otimes \mathbf{AF}^T)[\mathbf{S} \otimes \mathbf{FAF}^T + \mathbf{N} \otimes \mathbf{\Sigma}]^{-1}(\mathbf{S} \otimes \mathbf{FA}) \quad . \tag{233}$$

Using the stochastic estimation method for the a posteriori covariance and arranging the computations to exploit repeated expressions yields the steps

$$\mathbf{p} = [\mathbf{S} \otimes \mathbf{FAF}^T + \mathbf{N} \otimes \mathbf{\Sigma}]^{-1}\tilde{\mathbf{b}} \tag{234a}$$

$$\tau^2 = \tilde{\mathbf{b}}^T\mathbf{p} \tag{234b}$$

$$\hat{\mathbf{q}} = (\mathbf{S} \otimes \mathbf{AF}^T)\mathbf{p} \tag{234c}$$

$$\mathbf{x} = (\mathbf{S}^C \otimes \mathbf{FA}^C)\mathbf{z}_1 + (\mathbf{N}^C \otimes \mathbf{\Sigma}^C)\mathbf{z}_2 \tag{234d}$$

$$\mathbf{y} = (\mathbf{S} \otimes \mathbf{AF}^T)[\mathbf{S} \otimes \mathbf{FAF}^T + \mathbf{N} \otimes \mathbf{\Sigma}]^{-1}\mathbf{x} \tag{234e}$$

$$\hat{\mathbf{A}} = \mathbf{S} \otimes \mathbf{A} - \mathsf{E}\,\mathbf{y}\mathbf{y}^T \quad , \tag{234f}$$

where $\mathbf{p}$ is an intermediate variable of no physical significance. The values of $\mathbf{p}$ and $\mathbf{y}$ should, in principle, be efficiently computable by some iterative method. Unfortunately, however, none of the iterative methods tried so far converges reliably.

10.5 Stationary Priors

The stronger assumption that the priors are also constant in time may be appropriate in some applications. Then the spatiotemporal correlation can be written in the form

$$\tilde{A}(x, x', t, t') = A(x, x') \cdot S(t - t') \tag{235}$$

and the source and noise covaraince matrices in the form

$$\tilde{\mathbf{A}} = \mathbf{S} \otimes \mathbf{A} \qquad \text{and} \qquad \tilde{\mathbf{\Sigma}} = \mathbf{N} \otimes \mathbf{\Sigma} \quad , \tag{236}$$

where the temporal covariance matrices $\mathbf{S}$ and $\mathbf{N}$ are Toeplitz and approximately circulant. But if $\mathbf{S}$ and $\mathbf{N}$ are circulant, they can be diagonalized by the Fourier transform, or

$$\mathbf{S} = \mathbf{W}^H \mathbf{D}_S \mathbf{W} \qquad \text{and} \qquad \mathbf{N} = \mathbf{W}^H \mathbf{D}_N \mathbf{W} \tag{237}$$

where the matrices $\mathbf{D}_S$ and $\mathbf{D}_N$ are diagonal and $\mathbf{W}$ is the matrix corresponding to the fast Fourier transform, as given in equation (81).

Then the optimal source estimate and test statistic can be computed via the FFT as

$$\hat{\mathbf{q}} = (\mathbf{W}^H \mathbf{D}_S \otimes \mathbf{A}\mathbf{F}^T)[\mathbf{D}_S \otimes \mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{D}_N \otimes \mathbf{\Sigma}]^{-1}(\mathbf{W} \otimes \mathbf{I})\tilde{\mathbf{b}} \tag{238}$$

and

$$\tau^2 = ((\mathbf{W} \otimes \mathbf{I})\tilde{\mathbf{b}})^H [(\mathbf{D}_S \otimes \mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{D}_N \otimes \mathbf{\Sigma}]^{-1}(\mathbf{W} \otimes \mathbf{I})\tilde{\mathbf{b}} \quad . \tag{239}$$

These last two equations do not obviously lend themselves to a fast computational algorithm, though they might be computable by an iterative approach.

More importantly, the a posteriori covariance becomes block circulant and the standard errors can now be estimated from the time average of a single realization. To see this, manipulate the a posterior covariance for separable priors

$$\begin{aligned} \hat{\mathbf{A}} &= \mathbf{S} \otimes \mathbf{A} - (\mathbf{S} \otimes \mathbf{A}\mathbf{F}^T)[\mathbf{S} \otimes \mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{N} \otimes \mathbf{\Sigma}]^{-1}(\mathbf{S} \otimes \mathbf{F}\mathbf{A}) \\ &= (\mathbf{W}^H \otimes \mathbf{I}) \Big[\mathbf{D}_S \otimes \mathbf{A} - (\mathbf{D}_S \otimes \mathbf{A}\mathbf{F}^T) \cdot \\ &\qquad [\mathbf{D}_S \otimes \mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{D}_N \otimes \mathbf{\Sigma}]^{-1}(\mathbf{D}_S \otimes \mathbf{F}\mathbf{A}) \Big] (\mathbf{W} \otimes \mathbf{I}) \quad . \end{aligned} \tag{240}$$

Now observe that the expression inside the outer square brackets consists only of sums, products, and inverses of block diagonal matrices. It follows that $\hat{\mathbf{A}}$ must be block circulant. Then all the blocks on the diagonal are identical and the a posteriori covariances do not depend on time. Furthermore, the standard errors given by the diagonal of the diagonal block are also independent of time and can be estimated by averaging over time in a single realization. That is, the instantaneous a posteriori covariance can be estimated as

$$\hat{\mathbf{A}} \cong s^2 \mathbf{A} - \frac{1}{T} \sum_t \mathbf{y}_t \mathbf{y}_t^T \tag{241}$$

where s^2 is the common diagonal element of $\mathbf{S}$, T is the number of time instants, and y_t is the tth instant of the vector $\mathbf{y}$. If only the standard errors are necessary, they can be estimated as

$$\hat{\alpha}_{nn}^2 \cong s^2 \alpha_{nn}^2 - \frac{1}{T} \sum_t y_{tn}^2 \tag{242}$$

where y_{tn} is the nth element of the tth instant of $\mathbf{y}$ and α_{nn}^2 is the nth diagonal element of $\mathbf{A}$.

10.6 Sequential Time and Space Reconstructions

The optimal reconstruction filter for separable priors is

$$\tilde{\mathbf{H}} = (\mathbf{S} \otimes \mathbf{A}\mathbf{F}^T)[\mathbf{S} \otimes \mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{N} \otimes \mathbf{\Sigma}]^{-1} \tag{243}$$

which can be rewritten as

$$\begin{aligned}
\tilde{\mathbf{H}} &= (\mathbf{S} \otimes \mathbf{A}\mathbf{F}^T)[(\mathbf{S}+\mathbf{N}) \otimes (\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma}) - \mathbf{S} \otimes \mathbf{\Sigma} - \mathbf{N} \otimes \mathbf{F}\mathbf{A}\mathbf{F}^T]^{-1} \\
&= (\mathbf{S} \otimes \mathbf{A}\mathbf{F}^T)[(\mathbf{S}+\mathbf{N}) \otimes (\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})]^{-1} \cdot \\
&\quad \Big[\mathbf{I} - \mathbf{S}(\mathbf{S}+\mathbf{N})^{-1} \otimes \mathbf{\Sigma}(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1} \\
&\quad - \mathbf{N}(\mathbf{S}+\mathbf{N})^{-1} \otimes \mathbf{F}\mathbf{A}\mathbf{F}^T(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1}\Big]^{-1} \quad .
\end{aligned}$$

If the terms

$$\mathbf{S}(\mathbf{S}+\mathbf{N})^{-1} \otimes \mathbf{\Sigma}(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1}$$

and

$$\mathbf{N}(\mathbf{S}+\mathbf{N})^{-1} \otimes \mathbf{F}\mathbf{A}\mathbf{F}^T(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1}$$

are small compared to $\mathbf{I}$, the reconstruction filter can be approximated by the Kronecker product

$$\tilde{\mathbf{H}} \cong [\mathbf{S}(\mathbf{S}+\mathbf{N})^{-1}] \otimes [\mathbf{A}\mathbf{F}^T(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1}] \quad , \tag{244}$$

which is equivalent to an instantaneous reconstruction at each instant, followed by a temporal filter applied independently to each source.

Assuming in addition that the priors are stationary yields the approximation

$$\tilde{\mathbf{H}} \cong [\mathbf{W}^H \mathbf{D}_S(\mathbf{D}_S + \mathbf{D}_N)^{-1} \mathbf{W}] \otimes [\mathbf{A}\mathbf{F}^T(\mathbf{F}\mathbf{A}\mathbf{F}^T + \boldsymbol{\Sigma})^{-1}] \quad . \tag{245}$$

The corresponding algorithm first does an instantaneous OCLIM reconstruction applied independently at each instant in time, followed by a temporal filter using the FFT applied independently to each source. Alternatively, a temporal filter may be applied first to each detector independently, followed by an instantaneous OCLIM reconstruction at each instant independently. A disadvantage of this approach is that the approximation introduces an unknown error into the computed values of τ^2 and $\hat{\mathbf{A}}$.

The order of the Kronecker factors may be reversed, in which case the temporal filter is applied first to the sequence of measured field values independently for each detector, followed by an instantaneous OCLIM reconstruction independently at each instant.

One suitable set of computations is as follows

$$\mathbf{c} = (\mathbf{W} \otimes \mathbf{I})\tilde{\mathbf{b}} \tag{246a}$$

$$\mathbf{p} = [(\mathbf{D}_S + \mathbf{D}_N) \otimes (\mathbf{F}\mathbf{A}\mathbf{F}^T + \boldsymbol{\Sigma})]^{-1}\mathbf{c} \tag{246b}$$

$$\tau^2 = \mathbf{c}^H \mathbf{p} \tag{246c}$$

$$\hat{\mathbf{q}} = (\mathbf{W}^H \mathbf{D}_S \otimes \mathbf{A}\mathbf{F}^T)\mathbf{p} \tag{246d}$$

$$\mathbf{x} = (\mathbf{D}_S^{1/2}\mathbf{W} \otimes \mathbf{F}\mathbf{A}^C)\mathbf{z}_1 + (\mathbf{D}_N^{1/2}\mathbf{W} \otimes \boldsymbol{\Sigma}^C)\mathbf{z}_2 \tag{246e}$$

$$\mathbf{y} = [\mathbf{W}^H \mathbf{D}_S(\mathbf{D}_S + \mathbf{D}_N)^{-1} \otimes \mathbf{A}\mathbf{F}^T(\mathbf{F}\mathbf{A}\mathbf{F}^T + \boldsymbol{\Sigma})^{-1}]\mathbf{x} \tag{246f}$$

$$\hat{\alpha}_{nn}^2 = s^2 \alpha_{nn}^2 - \frac{1}{T}\sum_t y_{tn}^2 \quad , \tag{246g}$$

10.7 Spatial SVD Reconstruction

Another implementation is obtained if we assume that, in addition to separability and stationarity, that the noise statistics are the same for all detectors and uncorrelated between detectors. Then the noise covariance can be written in the form

$$\tilde{\boldsymbol{\Sigma}} = \mathbf{N} \otimes (\sigma^2 \mathbf{I}) \tag{247}$$

where σ^2 is a positive scalar constant. Since $\mathbf{FAF}^T$ is symmetric, it has a singular value decomposition

$$\mathbf{FAF}^T = \mathbf{UD}_A\mathbf{U}^T \tag{248}$$

where $\mathbf{U}$ is orthogonal and $\mathbf{D}_A$ is diagonal. But then $\sigma^2\mathbf{I} = \sigma^2\mathbf{UU}^T$ and the optimal estimate becomes

$$\hat{\mathbf{q}} = (\mathbf{W}^H\mathbf{D}_S \otimes \mathbf{AF}^T\mathbf{U})[\mathbf{D}_S \otimes \mathbf{D}_A + \mathbf{D}_N \otimes \sigma^2\mathbf{I}]^{-1}(\mathbf{W} \otimes \mathbf{U}^T)\tilde{\mathbf{b}} \tag{249}$$

which can be efficiently computed since the matrix being inverted is diagonal.

The required calculations can thus be done in the steps

$$\mathbf{c} = (\mathbf{W} \otimes \mathbf{U}^T)\tilde{\mathbf{b}} \tag{250a}$$

$$\mathbf{p} = [\mathbf{D}_S \otimes \mathbf{D}_A + \mathbf{D}_N \otimes \sigma^2\mathbf{I}]^{-1}\mathbf{c} \tag{250b}$$

$$\tau^2 = \mathbf{c}^H\mathbf{p} \tag{250c}$$

$$\hat{\mathbf{q}} = (\mathbf{W}^H\mathbf{D}_S \otimes \mathbf{AF}^T\mathbf{U})\mathbf{p} \tag{250d}$$

$$\mathbf{x} = (\mathbf{D}_S^{1/2}\mathbf{W} \otimes \mathbf{FA}^C)\mathbf{z}_1 + (\mathbf{D}_N^{1/2}\mathbf{W} \otimes \sigma\mathbf{I})\mathbf{z}_2 \tag{250e}$$

$$\mathbf{y} = (\mathbf{W}^H\mathbf{D}_S \otimes \mathbf{AF}^T\mathbf{U})[\mathbf{D}_S \otimes \mathbf{D}_A + \mathbf{D}_N \otimes \sigma^2\mathbf{I}]^{-1}(\mathbf{I} \otimes \mathbf{U}^T)\mathbf{x} \tag{250f}$$

$$\hat{\alpha}_{nn}^2 = s^2\alpha_{nn}^2 - \frac{1}{T}\sum_t y_{tn}^2 \quad , \tag{250g}$$

10.8 A Model Problem

These methods were tested by computer simulation on the same test configuration used in Chapter 7. The 28 sources in the central region A were assigned a source variance $\alpha_A^2 = 1$; the remaining sources were assigned a different source variance $\alpha_B^2 = 0.1$. There were no correlations between sources. Noise amplitudes taken from independent normal distributions with mean zero and variance $\sigma^2 = 3.16 \times 10^{-16}$ were added to each field measurement.

Figure 10.1 shows the assumed prior source and noise power spectra ($\mathbf{D}_S$ and $\mathbf{D}_N$) for the model problem, in arbitrary units. The source power was assumed to be concentrated at low frequencies and the noise was assumed to be white.

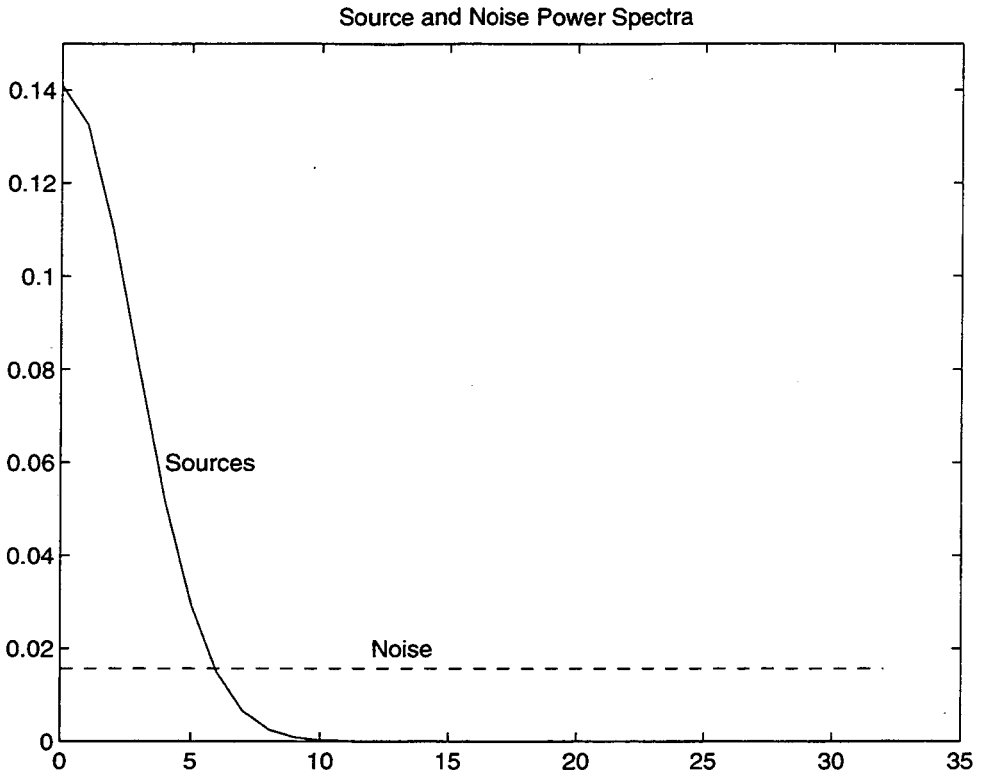


Figure 10.1 Assumed source and noise temporal power spectra. The source power was assumed to be concentrated at low frequencies and the noise was assumed to be white. The units are arbitrary.

10.9 Time-Space and Space-Time Plots

Both the true and reconstructed source distributions are difficult to display since they involve two (or three) space dimensions, one time dimension, and one amplitude dimension. The time-space plot illustrated at the top of Figure 10.2 is one useful way to display these data and consists of an 8 by 8 array of subplots corresponding to the 8 by 8 array of sources on the source plane. Each subplot shows how the amplitude of that source varies with time. For reconstructed data, two curves will be shown in each subplot; they indicate the upper and lower 1σ limits of the reconstructed amplitude. In this example, the source in row 5 and column 3 exhibits a sinusoidal variation with time. All the other sources have zero amplitude.

The space-time plot illustrated at the bottom of Figure 10.2 is an alternative representation of the same data. In this plot there are 64 images arranged left-to-

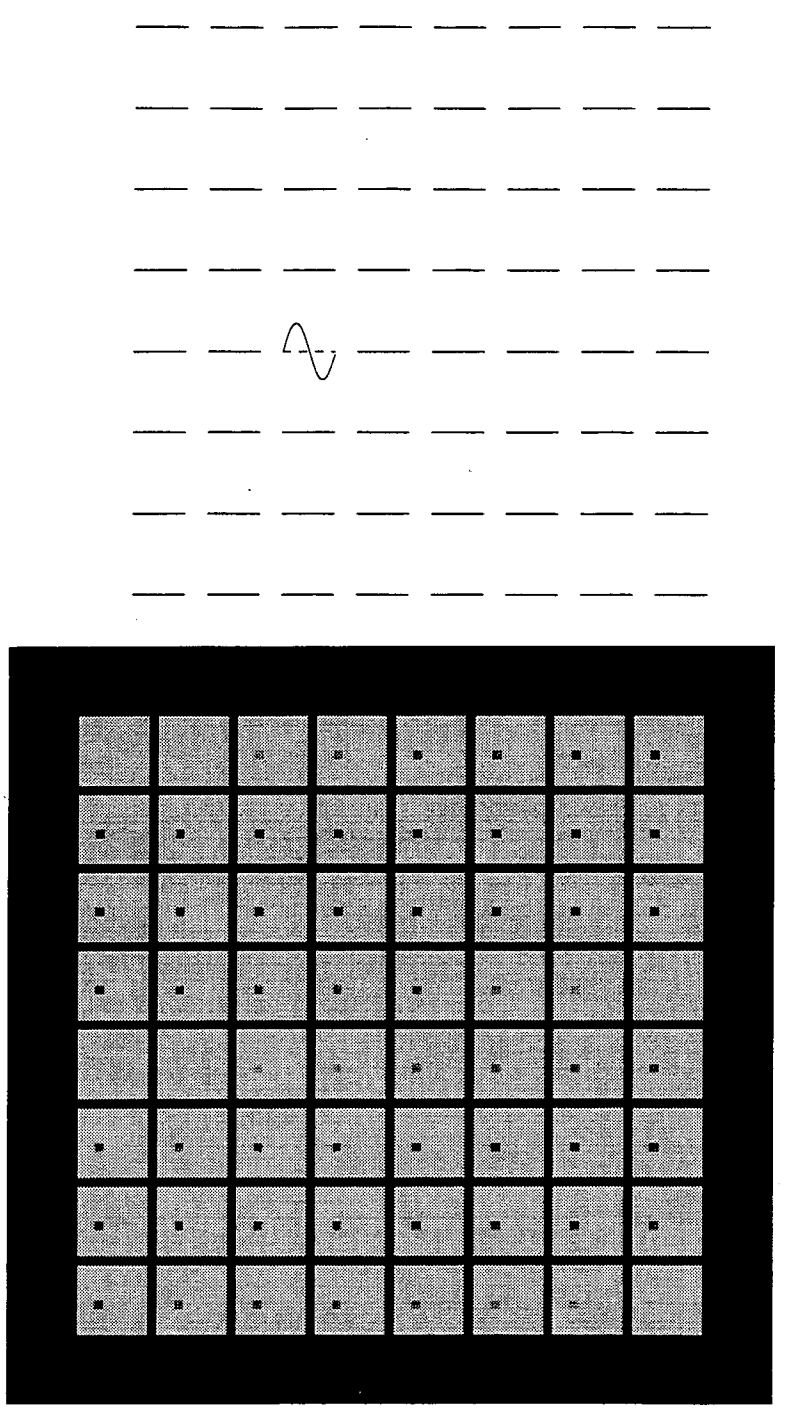


Figure 10.2 Time-space and space-time plots for reconstructed data. Both plots show a source distribution consisting of a single active source with a sinusoidal amplitude variation in time. The top "time-space plot" shows, for each source location, the amplitude variation in time. The bottom "space-time plot" shows, at each instant of time, the amplitude at each position with red representing positive and blue negative amplitudes. The source configuration used in these plots is shown in Figure 7.1.

right and top-to-bottom. (The fact that they form an 8 by 8 array is coincidental.) Each image corresponds to one instant of time and shows the amplitude of all 64 sources at that instant. Each image is an 8 by 8 array of pixels corresponding to the 8 by 8 array of sources and the color of the pixel denotes the amplitude of the source. Grey means zero amplitude, blue is current flow toward the viewer (positive), red is away from the viewer (negative), and the color saturation indicates the amplitude. Both red and blue become a dark grey in a grey scale reproduction. This example shows the same source distribution used above.

10.10 Simulation Results

Figure 10.3 shows a true source distribution to be reconstructed by the methods discussed in this paper. A moving source originates at row 7 and column 5 and moves along a curvilinear path to row 2 and column 5. Figure 10.4 shows the detector measurements (including noise) generated by this source distribution.

Figure 10.5 shows the results obtained by the method of instantaneous reconstruction on the model problem. The time-space plot at the top shows the upper and lower 1σ error limits on the amplitude. Figure 10.6 shows the results obtained by the method of sequential reconstruction on the model problem. Figure 10.7 shows the results obtained by the method of spatial SVD reconstruction on the model problem.

10.11 Discussion and Conclusions

All of the reconstruction methods track the apparent movement of the active source from bottom to top, although there is a substantial loss of spatial resolution, especially for source locations near the bottom. The instantaneous reconstruction has more high frequency noise than the others, since it has no temporal filtering. The sequential and SVD methods give similar results, though the SVD method appears to give slightly better spatial and temporal resolution; this improvement, however, seems to depend on the choice of problem and the criteria used for evaluation. All three of the methods tested here are quite fast, taking no more than a few minutes

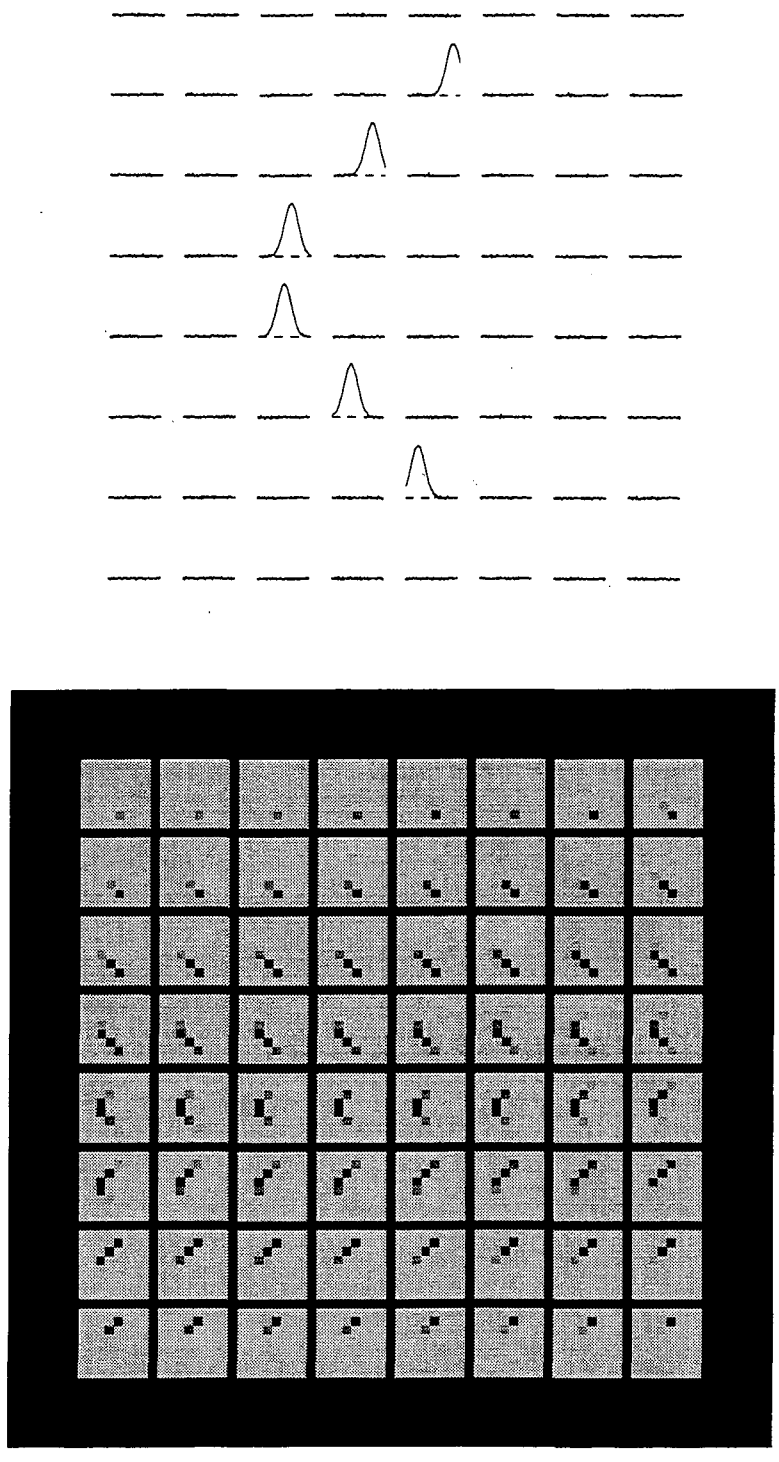


Figure 10.3 True source distribution for the model problem. A current source originates at row 7 and column 5 and moves along a curvilinear path to row 2 and column 5.

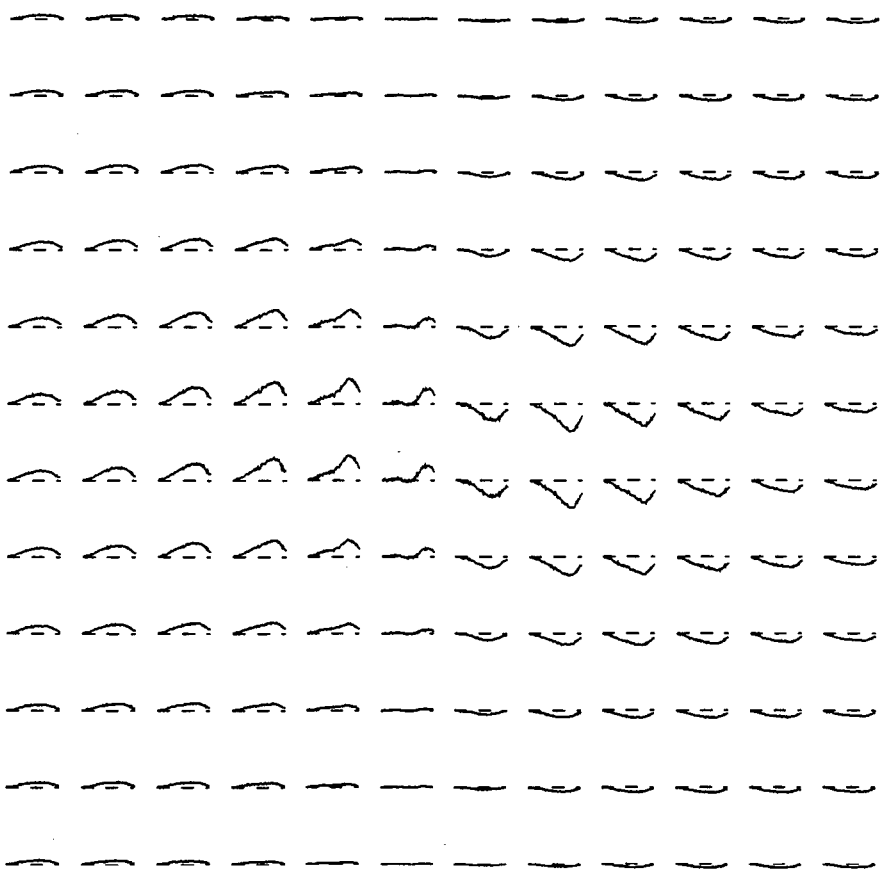


Figure 10.4 Detector measurements for the model problem. This plot shows, for each detector location, the measured field, including noise, as a function of time. The detector configuration is shown in Figures 7.1 and 7.2.

to compute; the direct computation is 100 to 1000 times slower and has not been implemented or tested. Instantaneous reconstruction is the fastest, sequential next, and SVD slowest. Since it appears to have the best resolution, the SVD method may be best for any problem that satisfies its conditions on the form of the priors. The sequential method has less restrictive conditions and is a good second choice.

All of these methods restrict the form of the priors to obtain computational efficiency. Whether the restrictions are sufficiently loose to permit the solution of real biomagnetic imaging problems is a question for future research.

The availability of standard errors for the reconstruction and a means for testing for violations of the priors provides a necessary statistical check of the validity and significance of the reconstructions obtained.

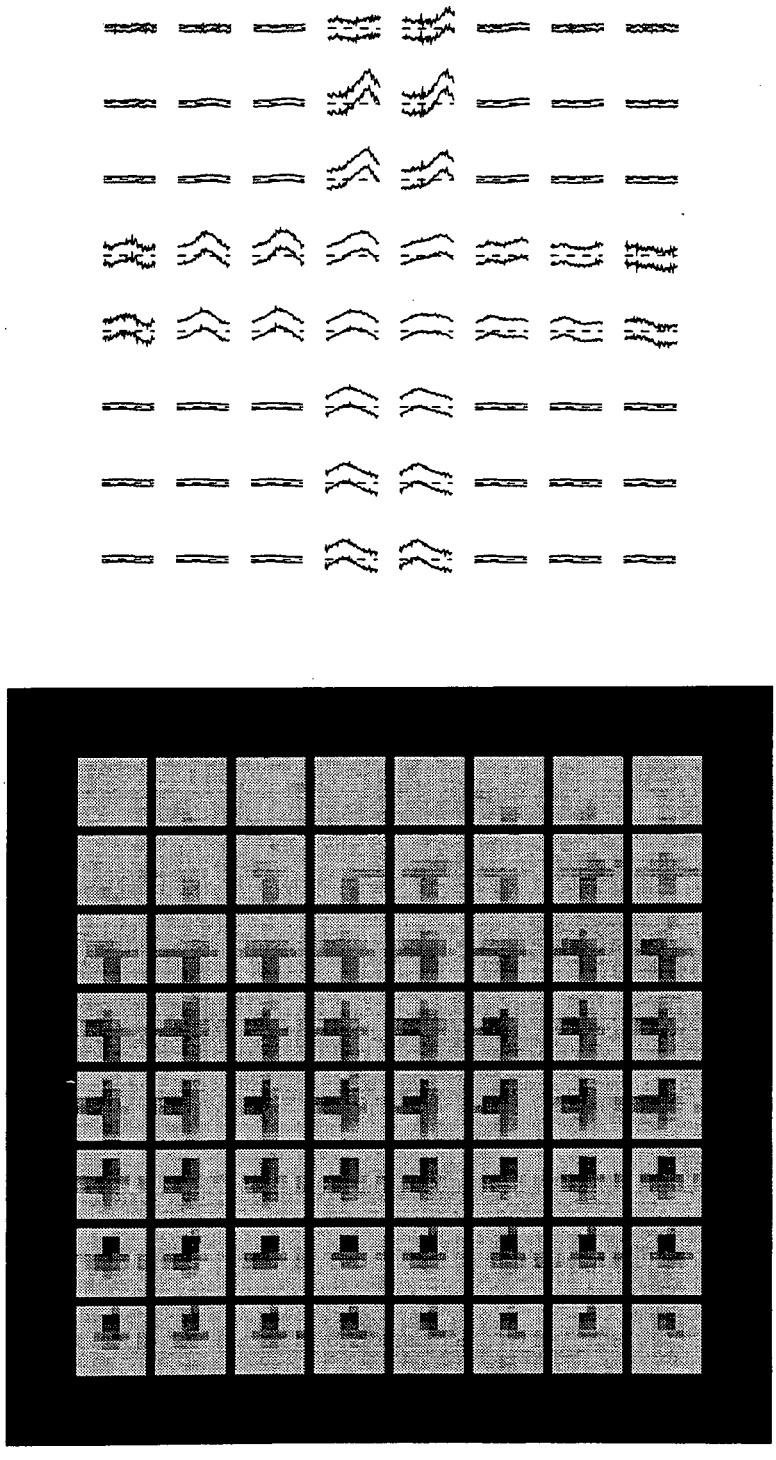


Figure 10.5 Results of instantaneous reconstruction. Reconstructing independently at each instant of time yields the source distribution shown above. The upper and lower curves in the upper plot indicate the upper and lower 1σ confidence limits.

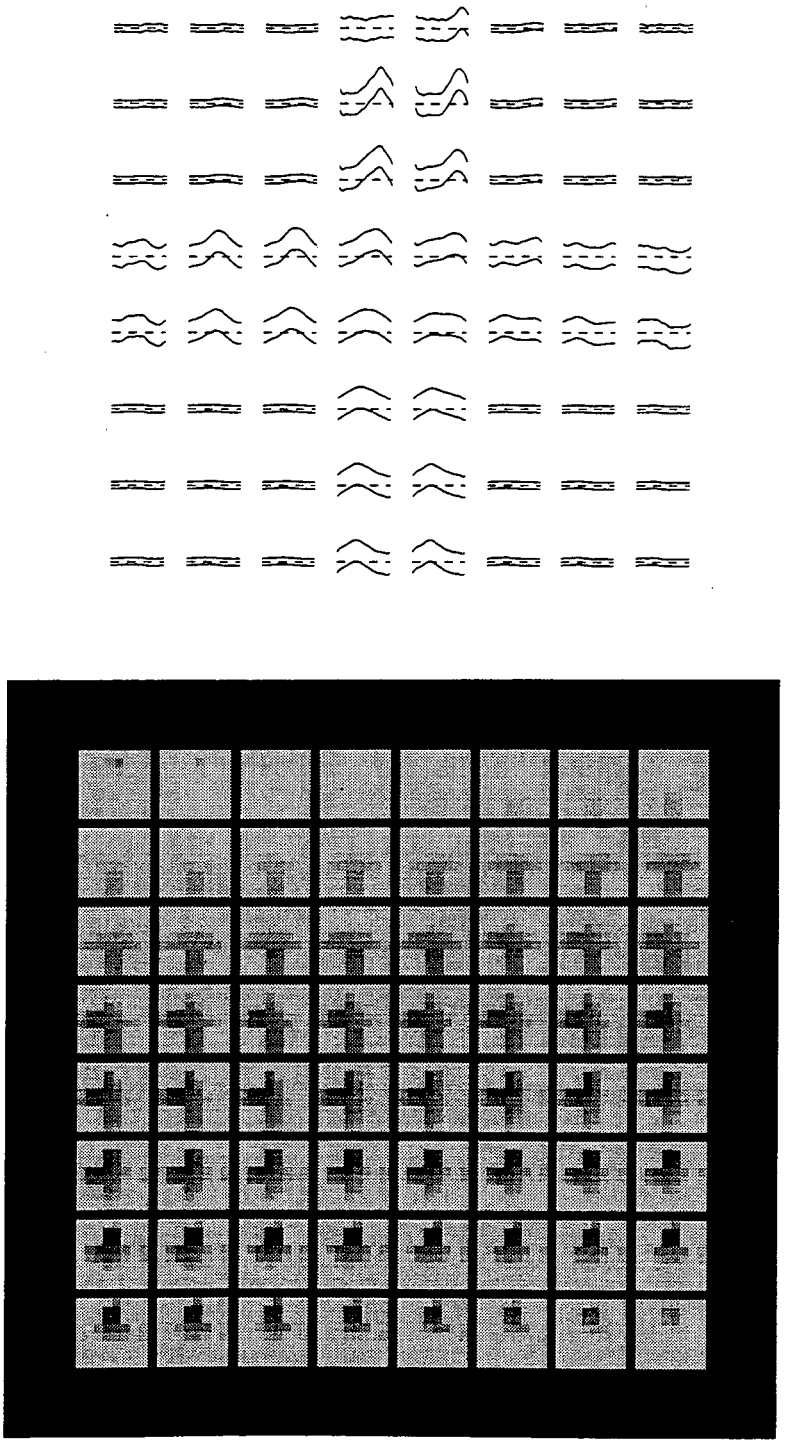


Figure 10.6 Results of sequential reconstruction. Reconstructing sequentially in time and space yields the source distribution shown above. The reconstruction is less noisy and slightly sharper than obtained by instantaneous reconstruction.

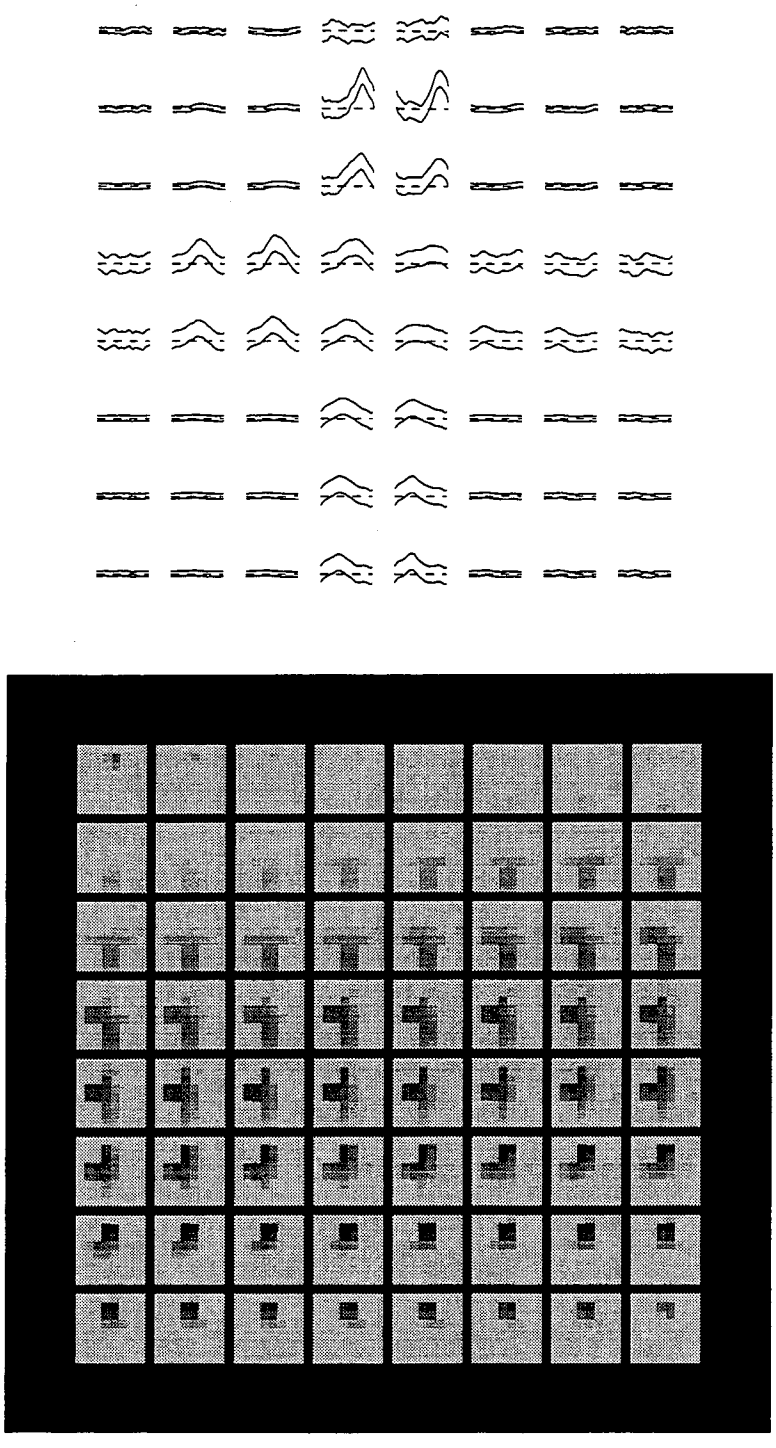


Figure 10.7 Results of spatial SVD reconstruction. Reconstruction using the SVD method yields the current distribution shown above. This reconstruction is the sharpest and least noisy but makes the strongest assumptions on the form of the spatiotemporal covariance.

Chapter 11

Simulations Using A Spherical Head Model

The previous chapters of this dissertation have developed an inverse method using prior information and explored various aspects of its performance in simplified MEG problems. It appears from these simulations that prior information about the location and amplitude of the primary current sources can be used to improve the reconstruction. The results of the previous chapters, however, were obtained using simplified magnetic source models that are unrealistic in that: The SNR used in many of the previous simulations is higher than is usually available in magnetoencephalography. Real sources in the head do not lie exactly on the modelled source locations. There are usually only a few relatively strong focal sources rather than many weaker sources with independent Gaussian amplitudes. And densely sampling the possible source volume to achieve high spatial resolution will usually generate more possible sources than there are detectors.

The primary purpose of this chapter is to present the results of a more realistic MEG reconstruction problem and to demonstrate that these differences between theory and reality do not prevent OCLIM from providing useful reconstructions. A second purpose is to demonstate that the inclusion of volume currents in the forward problem improves source localization when vector magnetometers are used.

11.1 An Improved Algorithm for OCLIM

The algorithm presented in Section 7.2 can be speeded up by using some new identities connecting the values to be computed and perhaps by using Monte Carlo

estimation for the confidence limits. We consider the case of sources uncorrelated with noise ($\mathbf{\Gamma} = 0$), which is the most common in practice. Then the measurement residual is

$$\begin{aligned}\mathbf{r} = \mathbf{b} - \mathbf{F}\hat{\mathbf{q}} &= (\mathbf{I} - \mathbf{F}\mathbf{A}\mathbf{F}^T(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1})\mathbf{b} \\ &= \mathbf{\Sigma}(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1}\mathbf{b} = \mathbf{\Sigma}\mathbf{p} \quad ,\end{aligned} \tag{251}$$

where $\mathbf{p} = (\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1}\mathbf{b}$ is already computed as an intermediate result. Furthermore, consider

$$\begin{aligned}\mathbf{p}^T\mathbf{r} = \mathbf{p}^T\mathbf{\Sigma}\mathbf{p} &= \mathbf{b}^T(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1}\mathbf{\Sigma}(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1}\mathbf{b} \\ &= \mathbf{b}^T[\mathbf{\Sigma}^{-1} - \mathbf{\Sigma}^{-1}\mathbf{F}(\mathbf{F}^T\mathbf{\Sigma}^{-1}\mathbf{F} + \mathbf{A}^{-1})^{-1}\mathbf{F}^T\mathbf{\Sigma}^{-1}]\mathbf{\Sigma}\mathbf{\Sigma}^{-1}\mathbf{\Sigma} \\ &\quad [\mathbf{\Sigma}^{-1} - \mathbf{\Sigma}^{-1}\mathbf{F}(\mathbf{F}^T\mathbf{\Sigma}^{-1}\mathbf{F} + \mathbf{A}^{-1})^{-1}\mathbf{F}^T\mathbf{\Sigma}^{-1}]\mathbf{b} \\ &= \mathbf{b}^T[\mathbf{I} - \mathbf{\Sigma}^{-1}\mathbf{F}(\mathbf{F}^T\mathbf{\Sigma}^{-1}\mathbf{F} + \mathbf{A}^{-1})^{-1}\mathbf{F}^T]\mathbf{\Sigma}^{-1} \\ &\quad [\mathbf{I} - \mathbf{F}(\mathbf{F}^T\mathbf{\Sigma}^{-1}\mathbf{F} + \mathbf{A}^{-1})^{-1}\mathbf{F}^T\mathbf{\Sigma}^{-1}]\mathbf{b} \\ &= (\mathbf{b} - \mathbf{F}\hat{\mathbf{q}})^T\mathbf{\Sigma}^{-1}(\mathbf{b} - \mathbf{F}\hat{\mathbf{q}}) \\ &= \chi^2 \quad ,\end{aligned} \tag{252}$$

where the second line was obtained by using the well-known ABCD lemma [44, p. 24]

$$(\mathbf{A} - \mathbf{B}\mathbf{C}\mathbf{D})^{-1} = \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1} - \mathbf{D}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{D}\mathbf{A}^{-1} \quad . \tag{253}$$

Similarly, notice that

$$\tau^2 = \mathbf{b}^T(\mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{\Sigma})^{-1}\mathbf{b} = \mathbf{b}^T\mathbf{p} \quad . \tag{254}$$

Finally, the surprise may be computed using the identity $\tau^2 = \chi^2 + \rho^2$, or $\rho^2 = \tau^2 - \chi^2$; the expected value of the surprise is $\mathsf{E}\,\rho^2 = \mathsf{E}\,\tau^2 - \mathsf{E}\,\chi^2 = M - \mathsf{E}\,\chi^2$.

11.1.1 Initialization

Initialization includes those computational steps that are independent of the actual data and which can be done in advance of data collection. The required inputs are the forward transfer matrix $\mathbf{F}$, the source covariance matrix $\mathbf{A}$, the noise covariance matrix $\mathbf{\Sigma}$, and the number of detectors M.

1. Compute and save the product $\mathbf{FA}$ for later use. This requires $O(MN^2)$ flops.

2. Compute the matrix $\mathbf{B} = \mathbf{FAF}^T + \mathbf{\Sigma}$ by direct matrix multiplication and addition in $O(MN^2)$ flops. Compute and save $\mathrm{Tr}(\mathbf{FAF}^T)$.

3. If desired, compute the signal-to-noise ratio $\mathrm{SNR} = \mathrm{Tr}(\mathbf{FAF}^T)/\mathrm{Tr}(\mathbf{\Sigma})$, using the intermediate value $\mathbf{FAF}^T$ computed in the last step.

4. Compute and save the Cholesky decomposition $\mathbf{LL}^T$ of $\mathbf{B}$ in $O(M^3)$ flops.

5. Choose a acceptable false positive rate α; this is the fraction of the data sets which are actually consistent with the priors but which will nevertheless be classified as inconsistent. Figure 8.1 shows the false positive rate α as a function of τ^2_{crit} for various values of M. Compute the corresponding critical value τ^2_{crit} such that $P(\tau^2_{\mathrm{crit}} \mid M) = 1 - \alpha$ where $P(\chi^2 \mid M)$ is the cumulative distribution function for the χ^2 distribution with M degrees of freedom.

The values of $\mathbf{L}$, $\mathbf{FA}$, and τ^2_{crit} must be saved for use in inverting particular data sets. The value of SNR gives some indication of the reconstruction quality that can be expected. The total cost is $O(M^3) + O(MN^2)$ flops; some minor improvements are possible if $\mathbf{A}$ or $\mathbf{\Sigma}$ are diagonal.

11.1.2 Confidence Limits: Direct Method

The direct computation of the confidence limits is exact (to within roundoff error) but may be slower than the Monte Carlo method when $N > M$. In either case, the computation may be done prior to data collection. The computational steps are:

1. Compute the Cholesky decomposition $\mathbf{SS}^T$ of the matrix $\mathbf{\Sigma}$ in $O(M^3)$ flops if $\mathbf{\Sigma}$ is dense and $O(M)$ flops if it is diagonal.

2. Compute $\mathsf{E}\,\chi^2 = \|\mathbf{L}^{-1}\mathbf{S}\|_F^2$ in $O(M^3)$ flops if $\mathbf{S}$ is dense and $O(M^2)$ if $\mathbf{S}$ is diagonal.

3. Compute the mean surprise by $\mathsf{E}\,\rho^2 = M - \mathsf{E}\,\chi^2$.

4. Compute $\mathbf{X} = \mathbf{L}^{-1}\mathbf{FA}$ by back substitution in $O(M^3)$ flops. This can be reduced to $O(M^2)$ flops if $\mathbf{A}$ is diagonal.

5. Compute $\hat{\mathbf{A}} = \mathbf{A} - \mathbf{X}^T\mathbf{X}$ directly in $O(N^3)$ flops. The diagonal entries $\hat{\alpha}^2_{nn}$ of $\hat{\mathbf{A}}$ are the squares of the standard errors for the corresponding estimates $\hat{q}_n$.

6. Compute the mean reconstruction error $\eta^2 = \mathrm{Tr}(\hat{\mathbf{A}})$.

The values of $\hat{\mathbf{A}}$, η^2, $\mathsf{E}\,\chi^2$, and $\mathsf{E}\,\rho^2$ provide information to help the user interpret the reconstruction results. The total computational cost is $O(M^3) + O(N^3)$ flops if $\mathbf{A}$ is dense or $O(M^2) + O(N^3)$ flops if $\mathbf{A}$ is diagonal.

11.1.3 Confidence Limits: Monte Carlo Method

Computing the confidence limits by the Monte Carlo method may be advantageous if $\mathbf{A}$ and $\mathbf{\Sigma}$ are diagonal, if only low accuracy is required, or if $N \gg M$. This version assumes that the a posteriori covariances $\hat{\alpha}_{ij}^2$ for $m \neq n$ are not required but only the a posteriori variances $\hat{\alpha}_{mm}^2$.

1. Compute the Cholesky decomposition $\mathbf{R}\mathbf{R}^T$ of the source covariance matrix $\mathbf{A}$. This takes $O(N)$ flops if $\mathbf{A}$ is diagonal and $O(N^3)$ flops if dense.

2. Compute the Cholesky decomposition $\mathbf{S}\mathbf{S}^T$ of the noise covariance matrix $\mathbf{\Sigma}$. This takes $O(M)$ flops if $\mathbf{\Sigma}$ is diagonal and $O(M^3)$ flops if dense.

3. Repeat steps 4–6 R times, where the number of realizations R is chosen to yield sufficient accuracy.

4. Generate two independent standard normal random vectors $\mathbf{z}_1$ and $\mathbf{z}_2$, where $\mathbf{z}_1$ has length N and $\mathbf{z}_2$ has length M. This takes $O(M+N)$ flops per realization.

5. Compute a sample measurement vector $\mathbf{x} = \mathbf{F}\mathbf{R}\mathbf{z}_1 + \mathbf{S}\mathbf{z}_2$. This takes $O(MN) + O(N^2)$ flops per realization.

6. Compute the intermediate value $\mathbf{p} = \mathbf{L}^{-1}\mathbf{x}$ and the corresponding reconstruction $\mathbf{y} = (\mathbf{F}\mathbf{A})^T\mathbf{L}^{-T}\mathbf{p}$. This takes $O(M^2) + O(MN)$ flops per realization.

7. Compute the approximate mean square residual $\mathsf{E}\,\chi^2 = \frac{1}{R}\sum_r \mathbf{p}_r^T\mathbf{\Sigma}\mathbf{p}_r$. This requires $O(RM^2)$ flops if $\mathbf{\Sigma}$ is dense but only $O(RM)$ flops if it is diagonal.

8. Compute the approximate a posteriori variances $\hat{\alpha}_{nn}^2 \cong \alpha_{nn}^2 - \frac{1}{R}\sum_r y_{rn}^2$ where y_{rn} is the nth element of the rth realization of $\mathbf{y}$. This requires $O(RN)$ flops.

The total computational cost is $O(RM^2) + O(RN^2)$ provided that both $\mathbf{A}$ and $\mathbf{\Sigma}$ are diagonal or are sparse enough that their Cholesky decompositions can be computed in not more than $O(M^2)$ and $O(N^2)$ flops.

11.1.4 Reconstruction

Computing the estimate for each input $\mathbf{b}$ requires the values of $\mathbf{FA}$, $\mathbf{L}$, and τ^2_{crit} computed during initialization. The compuational steps are:

1. Compute $\mathbf{p} = \mathbf{L}^{-T}\mathbf{L}^{-1}\mathbf{b}$ by back substitution in $O(M^2)$ flops.

2. Compute $\tau^2 = \mathbf{p}^T\mathbf{b}$. If $\tau^2 > \tau^2_{\text{crit}}$, classify the given data as inconsistent with the priors; this means that the data set $\mathbf{b}$ does not appear to be taken from the statistical population defined by the priors and should not be reconstructed using those priors. Otherwise, classify it as consistent and proceed with the reconstruction.

3. Compute the solution $\hat{\mathbf{q}} = (\mathbf{FA})^T\mathbf{p}$ in $O(MN)$ flops.

4. Compute the residual $\mathbf{r} = \mathbf{\Sigma p}$ in $O(MN)$ flops if $\mathbf{\Sigma}$ is dense and $O(M)$ if $\mathbf{\Sigma}$ is diagonal.

5. Compute $\chi^2 = \mathbf{p}^T\mathbf{r}$ in $O(M)$ flops.

6. Compute $\rho^2 = \tau^2 - \chi^2$.

The value of $\hat{\mathbf{q}}$ is the desired reconstruction. The values of $\mathbf{r}$, χ^2, τ^2, and ρ^2 are intended to provide insight into the quality of the reconstruction. The total cost per reconstruction is $O(M^2) + O(MN)$ flops.

11.2 Methods

The following MEG model problem was used in individual reconstructions and in Monte Carlo simulations to investigate the performance of the OCLIM algorithm under realistic conditions.

11.2.1 Detector Configuration

The detector configuration used in the model problem is intended to model a generic next-generation whole-head MEG system but does not exactly model any particular system. The detector array consists of 82 vector magnetometers more or less uniformly distributed on a spherical shell of radius 13 cm and extending 105° from the top of the head, thus covering slightly more than a hemisphere. Each vector magnetometer measures the three components of the field normal and tangential

to the head, giving a total of 246 measurements. Each measurement has zero-mean Gaussian noise of 50 fT RMS amplitude and the noise is uncorrelated between measurements.

11.2.2 Assumed Source Models

The head was modeled as a spherical conductor of radius 10.5 cm at the inside boundary of the skull. Since the skull has very low conductivity relative to the brain and cerebrospinal fluid, currents flowing in the skull and scalp have a neglible effect on the measured magnetic fields and were not simulated. The brain was taken as the upper hemisphere of radius 9 cm and the cortex as the upper hemispherical shell of inner radius 7 cm and outer radius 9 cm. The primary current distribution was modelled as a rectangular grid of three-component current dipoles at a 1.8 cm spacing.

Note that in this geometry, radially oriented current dipoles produce no external magnetic field and thus cannot be detected at all. In addition, the volume currents produce only magnetic fields tangent to the surface of the sphere and have no effect on radial field measurements.

Two different source configurations were simulated. In the whole brain (CP−) source model, all sources were assumed to lie within the brain but not necessarily in the cortex; there were 277 distinct sites and 831 distinct source amplitudes to be reconstructed, as shown in Figure 11.1. In the cortical priors model (CP+), all sources were assumed to lie within the cortex only; there were 129 sites and 387 independent source amplitudes, as shown in Figure 11.2.

In both source models, each independent source amplitude was assumed to be Gaussian with mean zero and standard deviation 2.5 nA-m. This source variance was chosen to yield an RMS signal amplitude at the nearest detector of about 200 fT, which is typical of visual evoked field measurements.

The cortical priors model was chosen to facilitate comparison with the whole brain model. If the purpose was to obtain the best reconstructions rather than to compare the effects of the source models, it would probably be appropriate

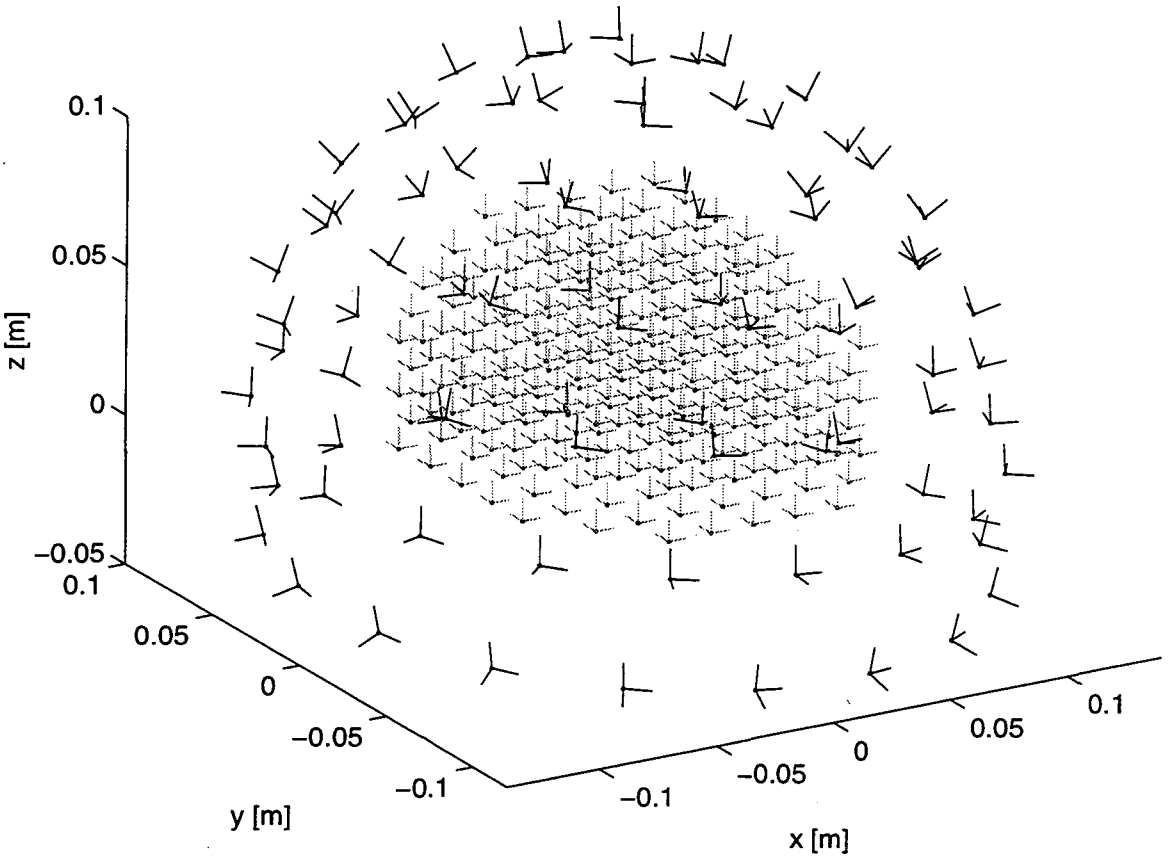


Figure 11.1 Source and detector configuration without priors (CP−). The detectors are vector magnetometers more or less uniformly distributed over a spherical cap; the large black triads show their location and orientation. The sources were placed on a rectangular grid throughout the entire brain and are shown by the smaller gray triads.

to increase the number of sources, use only tangential sources, and space them uniformly over the surface of several hemispherical shells.

Each of these models was used both with and without volume currents. In the no volume current (VC−) case, the magnetic field at the detctors is assumed to depend only on the primary current. In the volume current (VC+) case, the magnetic field at the detectors is computed using the formulae given by Sarvas [86] for dipole sources in a spherical conductor. This yields four distinct models to be compared, as summarized in Figure 11.3.

Figure 11.4 shows the RMS signal amplitude at each detector for all four of the source models; the detectors are numbered to place all the radial detectors first,

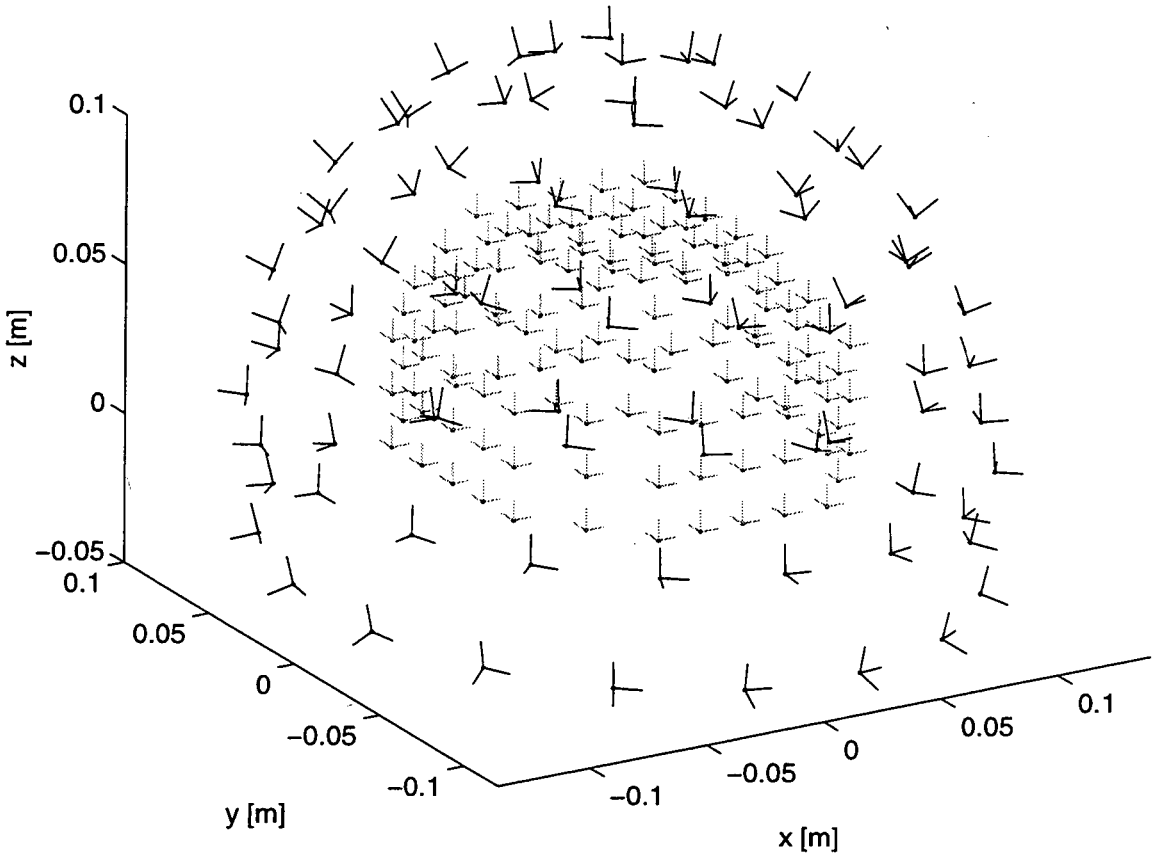


Figure 11.2 Source and detector configuration with priors (CP+). The detector configuration is identical to the previous figure, but the sources were restricted to the cortex.

SM0	CP− VC−	No priors, no volume currents;
SM1	CP+ VC−	Cortical priors, no volume currents;
SM2	CP− VC+	No priors, volume currents included; and
SM3	CP+ VC+	Both cortical priors and volume currents included.

Figure 11.3 Source models to be tested.

followed by the tangential detectors. Note that the signal ranges between 100–200 fT for the radial detectors, which is typical of visual evoked field imaging. Including the volume currents for this spherical head model reduces the tangential fields by about two-thirds but has no effect on the radial fields; thus we expect that including

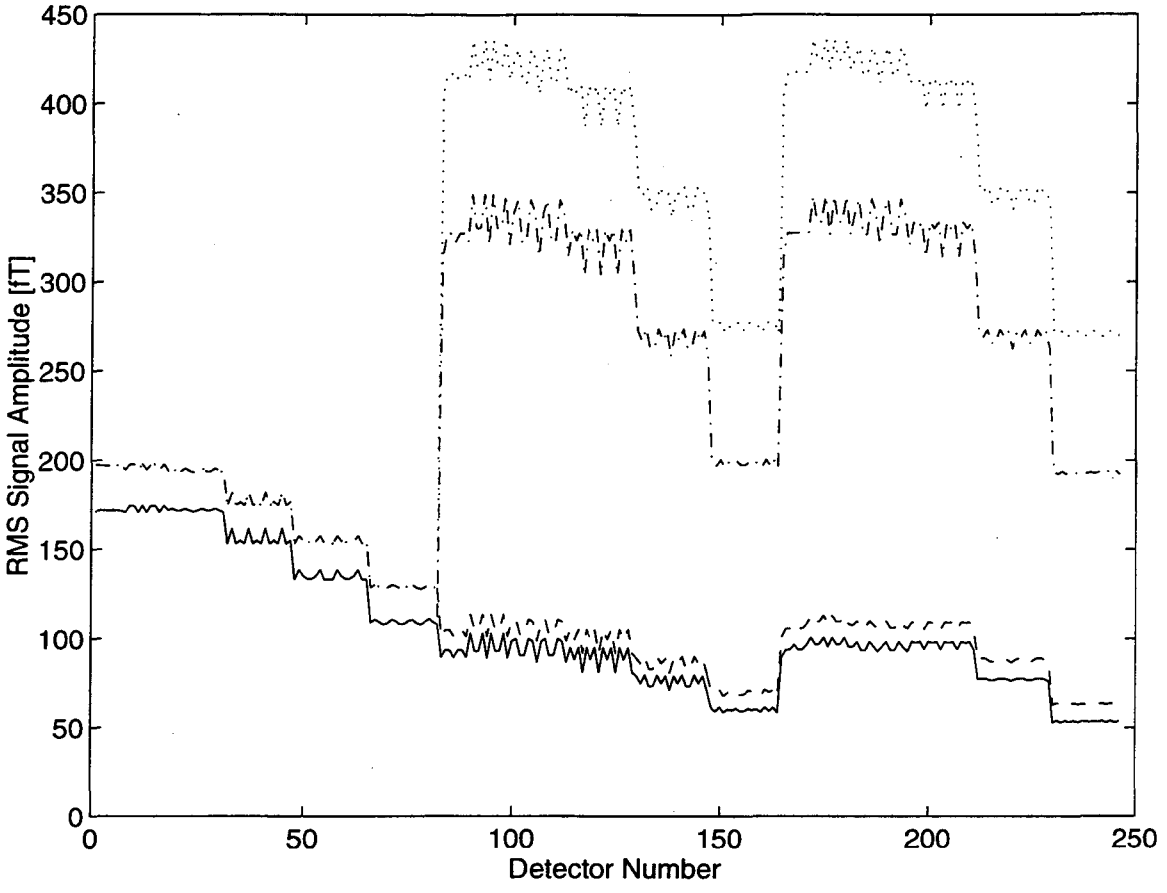


Figure 11.4 RMS signal amplitude for model problem. This plot shows the RMS signal amplitude for each detector for all four of the source models; the models are, from top to bottom, CP−VC−, CP+VC−, CP−VC+, and CP+VC+. Detectors 1–82 measure the radial component of field, detectors 83–164 the eastward component, and detectors 165–246 the northern component. The inclusion of volume currents has no effect on the radial field, but reduces the tangential fields by about two-thirds. The radial field in the CP+VC+ source model ranges from 100–180 fT, which is typical for visual evoked response measurements.

the volume currents with significantly improve the reconstruction accuracy when tangential field measurements are used. The signal-to-noise ratios for the CP−VC+ and CP+VC+ source models are 8.01 dB and 6.89 dB.

11.2.3 True Source Model

The assumed source model is that all of the modelled source locations are active simultaneously but each with a relatively low amplitude. A more realistic model of the human brain is that only a few sources are active at any given time but

have higher amplitudes; furthermore, the positions of these sources will not exactly match any of the assumed source locations.

One of the purposes of this simulation is to explore whether this simplified Gaussian model is adequate to allow reconstruction of realistic source distributions. Thus the statistical model actually used to generate source distributions in this simulation is as follows: Choose at random 5 locations within the cortex as defined above (without regard to whether or not they fall on the modelled source locations). At each of these random locations, choose a direction randomly chosen over the surface of a sphere and assign a current dipole moment of magnitude 20 nA-m at that location and in that direction.

To compare this to the assumed source distribution, suppose that we have the 129 sites of the CP+ model, of which a randomly chosen 5 are active. The expected variance at a random site is then $(5/129) \cdot (20 \text{ nA-m})^2$ which divides equally over the three components to yield a variance of $(1/3) \cdot (5/129) \cdot (20 \text{ nA-m})^2$ or a standard deviation of 2.27 nA-m for each component, which approximately matches the assumed value of 2.5 nA-m.

11.2.4 Energy Concentration

A common measure of the localization accuracy of a reconstruction algorithm is the full-width half-maximum (FWHM) of its point spread function; that is, the diameter of the circle or sphere which connects the points at which the point spread function has half its maximum value. This is not appropriate in the present problem, since the point spread function is asymmetric and since the reconstruction grid is rather sparse. An alternative is to compute how much of the reconstructed source energy appears within a specified radius σ_0 of the true source position. Given the sparse grids used here, however, it is not appropriate simply to count all the reconstructed source energy within the specified radius, since this sum will change discontinuously as the true source location moves relative to the modelled source locations. A better approach is to weight this sum by some smooth decreasing function of distance. A Gaussian is a reasonable choice, although some investigation of alteratives might be useful. Since the reconstructed amplitudes may be

either positive or negative, they are squared to yield energy, which is then weighted and summed. The actual value of the energy is probably less meaningful than the fraction of energy near the true source location relative to the total energy of the reconstruction. Thus, the energy concentration may be formally defined by

$$\text{Energy concentration} = \frac{\sum_{n=1}^{N} \hat{q}_n^2 \cdot \exp(-\|\vec{p} - \vec{p}_n\|^2 / 2\sigma_0^2)}{\sum_{n=1}^{N} \hat{q}_n^2} \quad , \tag{255}$$

where N is the total number of modelled sources, $\hat{q}_n$ is the reconstructed amplitude of the nth modelled source, $\vec{p}$ is the true position of the source, $\vec{p}_n$ is the position of the nth modelled source, and σ_0 is the standard deviation of the Gaussian selected.

The choice of radius σ_0 should be related to the resolution of the reconstruction or to the spacing between the modelled sources. One convenient choice is set it to the maximum possible distance between the true source location and the nearest modelled source location. This is achieved when the true source is located at the center of a cube formed by eight modelled source locations; the distance is then $\sigma_0 = \frac{1}{2}\sqrt{3} \cdot 1.8 \text{ cm} = 1.559 \text{ cm}$. Fortunately, the precise value chosen for σ_0 does not seem to affect the relative performance of the source models.

11.3 Simulation Results and Discussion

The following specific hypotheses were investigated by computer simulation of the model problem:

1. The inclusion of volume currents and prior information on possible source locations decreases the width of the point spread function, presumably improving resolution.

2. It is possible to distinguish true sources of moment 20 nA-m from background noise in the reconstruction.

3. OCLIM will provide useful reconstructions even if there are only a few active sources which do not lie exactly on the modelled source locations and there are more modelled source locations than field measurements.

4. OCLIM can resolve several simultaneously active sources within the brain.

In all of these simulations, volume currents were included in computing the measured fields, regardless of the assumed source model.

11.3.1 Source Models Compared

The Monte Carlo simulation was done to test Hypothesis 1 by comparing the source energy localization performance of the four source models. Fifty realizations were computed, each of which has a single tangential source of moment 20 nA-m at a random location in the cortex. Each of the source models was then used to reconstruct the source location and the energy localization computed; the results were plotted to yield Figure 11.5, where + denotes CP−VC− vs CP+VC+, × denotes CP+VC− vs CP+VC+, and o denotes CP−VC+ vs CP+VC+. If a point is above the diagonal line, it means that CP+VC+ provides better energy localization than the alternative model. All these simulations included measurement noise.

Note in Figure 11.5 that almost all of the sample points are above the diagonal, indicating the that CP+VC+ model has better source localization than any of the other three source models. CP−VC+ is slightly worse than CP+VC+ but consistently so, indicating that priors do provide an improvement in source localization. Both CP−VC− and CP+VC− are usually worse than CP−VC+ and CP+VC+, indicating that the inclusion of volume currents in the forward model improves the source localization; note that the difference between either CP−VC− or CP+VC− and CP+VC+ is quite variable, so that neglecting volume currents will both increase the reconstruction error and add its variance.

11.3.2 True Sources versus Background Noise

Hypothesis 2 claims that true sources can be distinguished from background noise in the reconstruction. To estimate the background levels of the energy concentration when no source is present, the previous simulation was repeated with the dipole moment set to zero; the resulting mean and standard deviation of the energy concentration are shown in Figure 11.6. Note that almost all of the realizations plotted in Figure 11.5 exceed the upper 2σ threshold for the zero-signal energy concentration; thus it appears that either this or perhaps a 3σ threshold would reliably distinguish sources of 20 nA-m moment or higher from reconstruction noise.

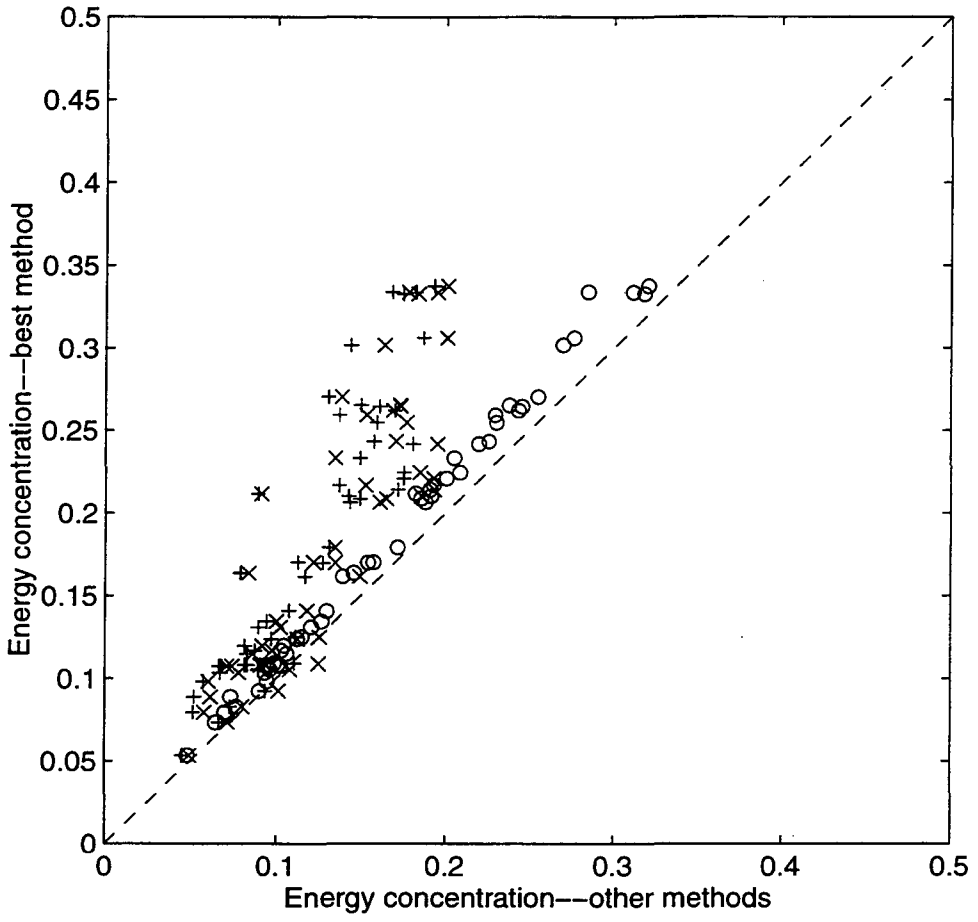


Figure 11.5 Energy concentration compared for source models. Fifty realizations of a source location in the cortex were generated and reconstructed with each of the four source models. This scatter plot compares the energy concentration for the four source models, where + denotes CP−VC− vs CP+VC+, × denotes CP+VC− vs CP+VC+, and o denotes CP−VC+ vs CP+VC+. If a point is above the diagonal line, it means that CP+VC+ provides better energy localization than the alternative model indicated by the symbol.

11.3.3 Sample Point Spread Functions

Hypothesis 3 claims that OCLIM is robust to certain discrepancies between the assumed and actual source models. The simulations in this section verify that claim by showing samples of the point spread functions for sources at various positions; they also supplement the quantitative measures of localization by giving a visual picture of the point spread function.

Figures 11.7 through 11.11 show a series of reconstructions in which a single

Model	Mean ± S.D.	Mean + 2 S.D.
CP−VC−	0.0266 ± 0.0114	0.0494
CP+VC−	0.0279 ± 0.0115	0.0509
CP−VC+	0.0297 ± 0.0148	0.0593
CP+VC+	0.0302 ± 0.0150	0.0602

Figure 11.6 Zero-signal energy concentration versus source model. The mean and standard deviation of energy concentration when no source is present are shown for each source model. The upper 2σ deviation could be taken to distinguish true signals from background noise in the reconstruction.

true source is active and there is no measurement noise. Thus they show the point spread function of the reconstruction algorithm for several possible source locations. The first column shows the true location of the source; the next four columns show the reconstructions obtained using the four source models, all at the same intensity scale. Each row of images represents one transverse slice through the head. The solid circle is the inner surface of the skull (and thus the boundary of the spherical conductor); the two dotted circles indicate the inner and outer boundaries of the cortex. The capital letters in the top left image denote Anterior, Posterior, Right, and Left. Successive rows present successive slices through the head from top to bottom.

These reconstructions confirm the quantitative results on energy concentration shown in the last section. The CP+VC+ source model generally provides a higher peak intensity than the other source models and there is less energy outside the peak. The width of the point spread function varies depending on the location of the source, but usually has most of its energy within 1–2 pixels or 1.8–3.6 cm of the peak. In Figures 11.7 through 11.9 the reconstructed peak is at the same location as the true source; in Figures 11.10 and 11.11 it is displaced by one pixel. Comparing the models does not show any clear pattern as to whether priors or volume currents provide a larger improvement.

Another set of simulations was done for the same sources but including measurement noise. The patterns observed are much the same except that there is

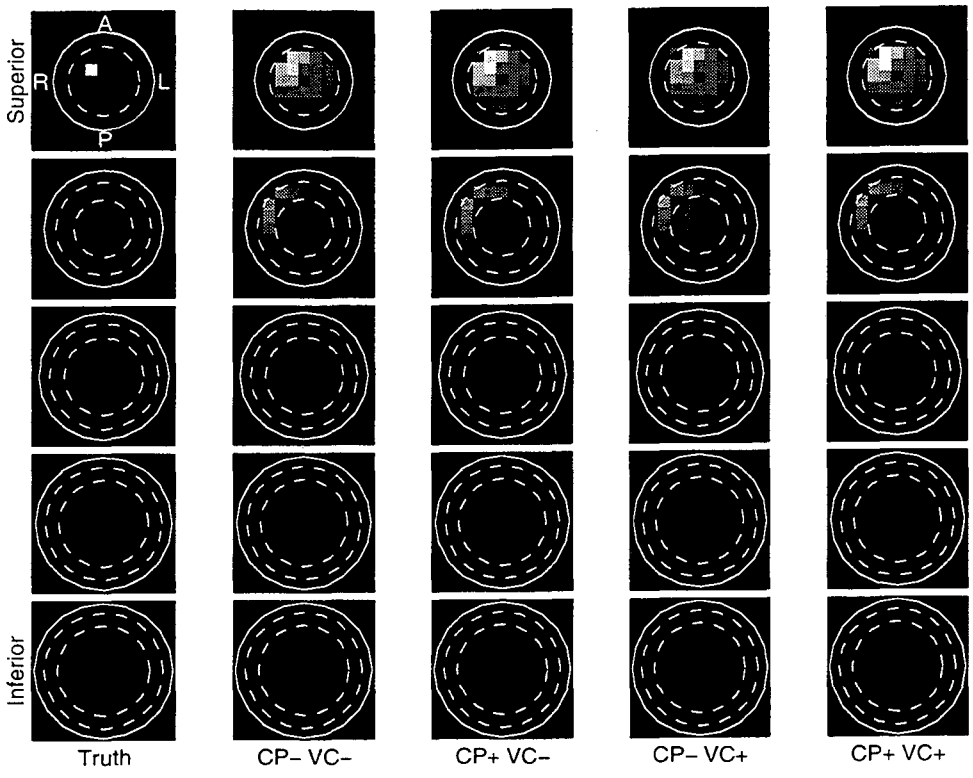


Figure 11.7 Sample point spread function #1.

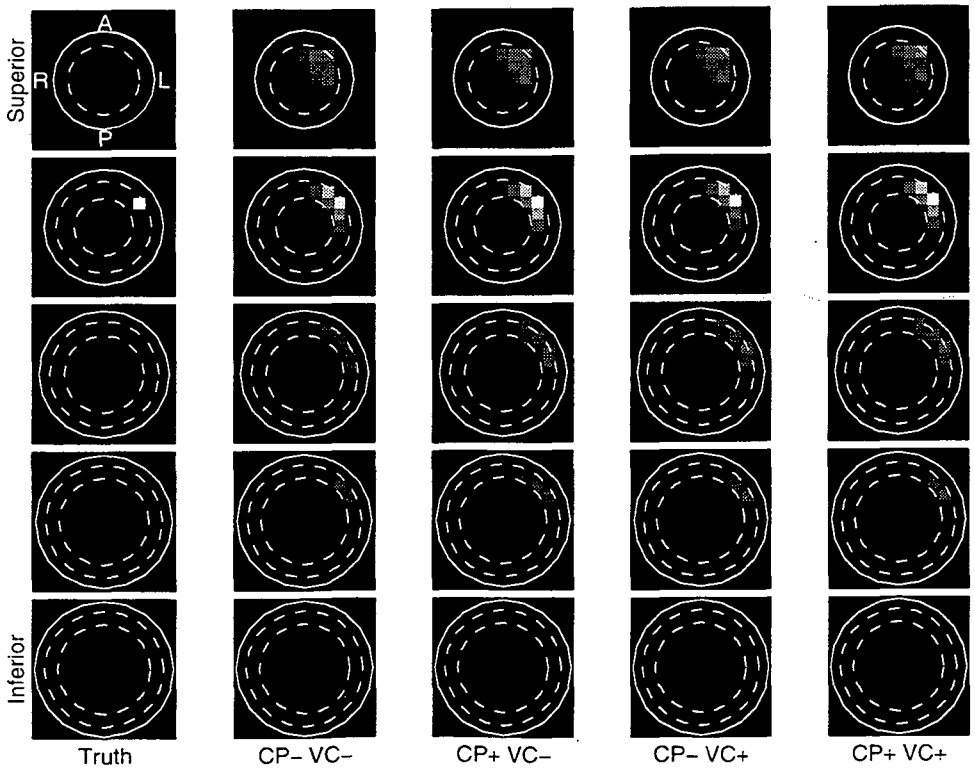


Figure 11.8 Sample point spread function #2.

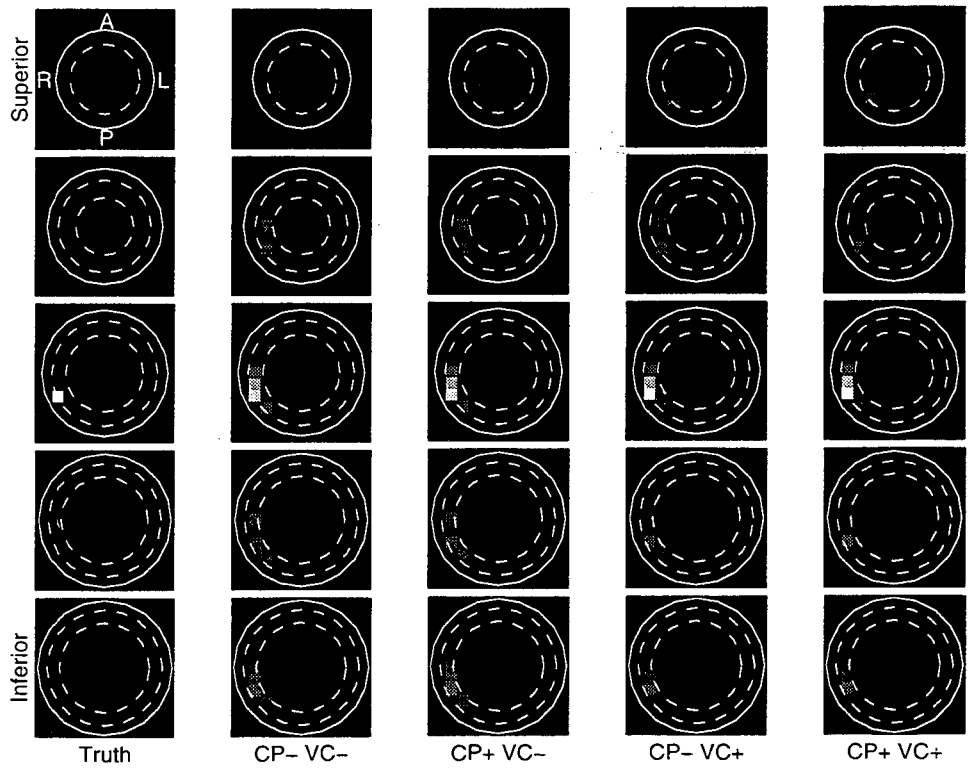


Figure 11.9 Sample point spread function #3.

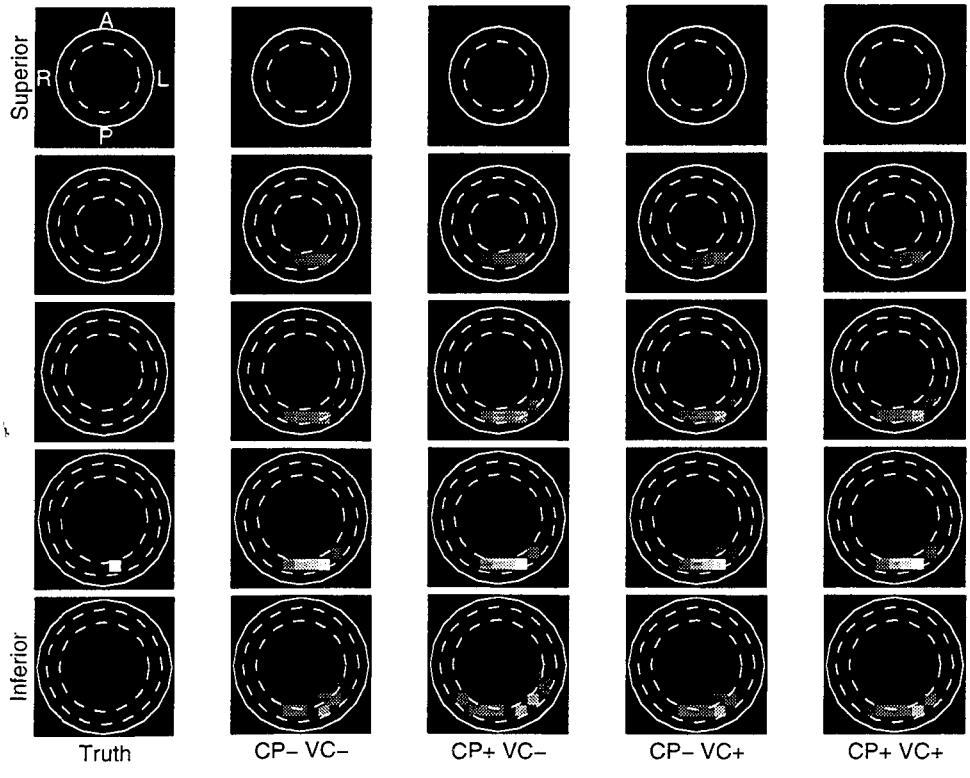


Figure 11.10 Sample point spread function #4.

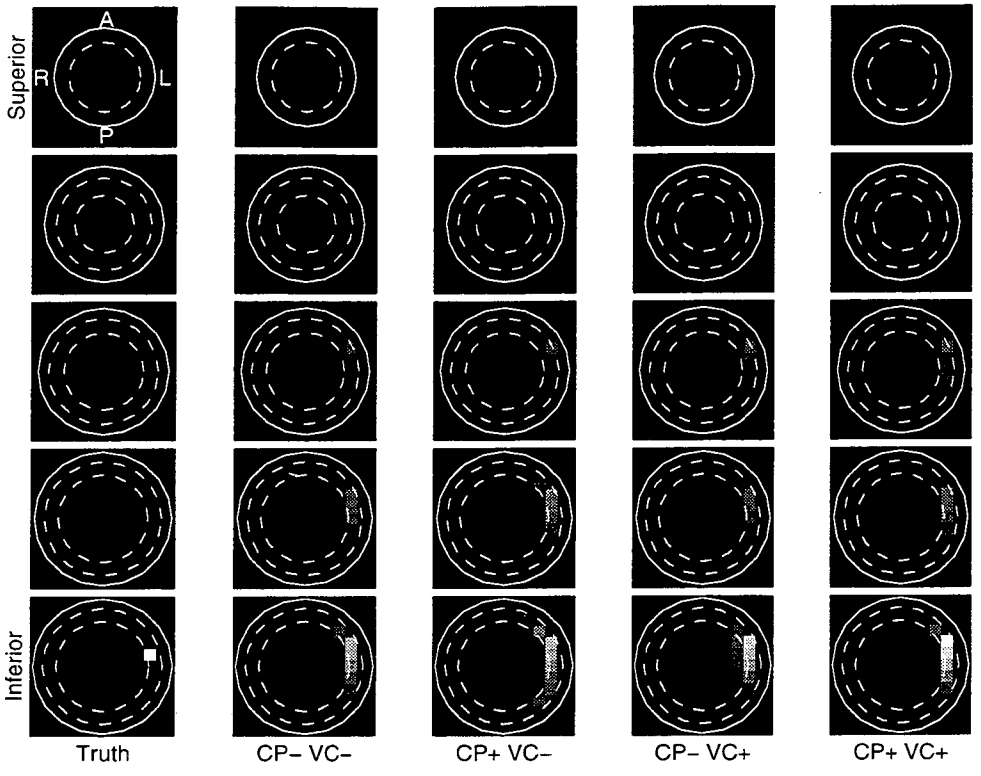


Figure 11.11 Sample point spread function #5.

noticeably more background noise in the reconstructions. The sources, however, remain clearly distinguishable from the background.

11.3.4 Reconstruction of Multiple Sources

Hypothesis 4 claims that OCLIM can resolve several simultaneously active sources in the brain.

Figure 11.12 shows the reconstruction for 5 simultaneously active sources in the cortex, generally tangential with dipole moments ranging from 18 to 25 nA-m. These sources were *not* placed directly over the modelled source locations. Two of the sources, in slices 3 and 5 counting from the top, are clearly visible in the reconstruction. The sources in slices 2 and 4 are marginally visible but appear to have been displaced to slice 3. The source in slice 1 is visible as a blurred reconstruction in the same slice; it is perhaps arguable that a source is present but it is not clearly localized.

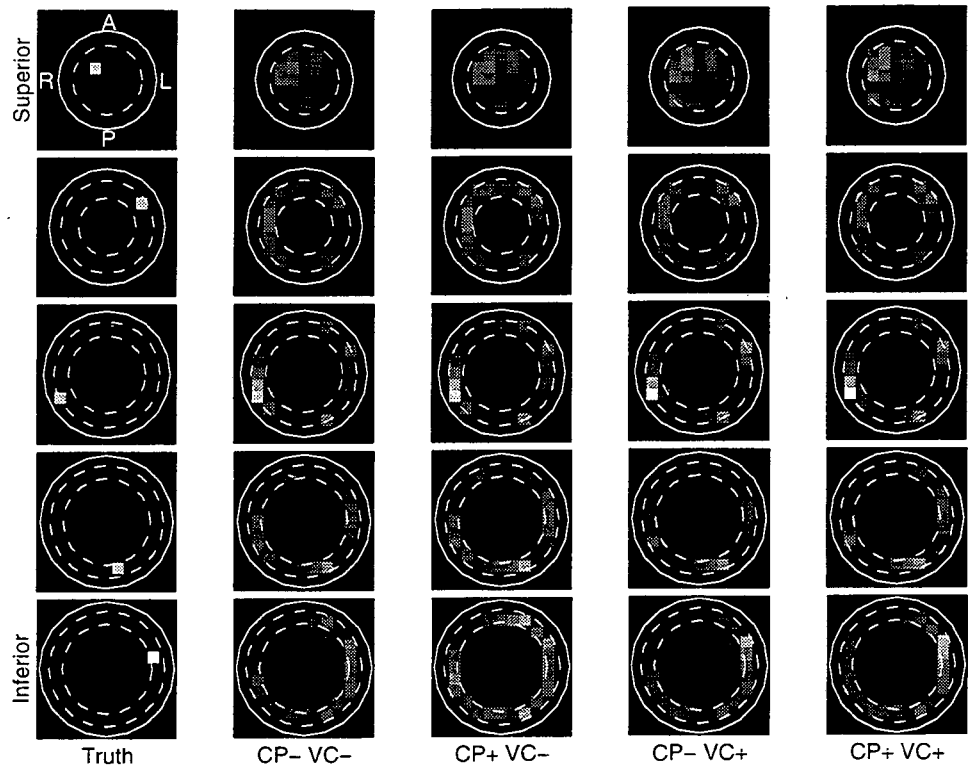


Figure 11.12 Reconstruction of multiple sources—high noise.

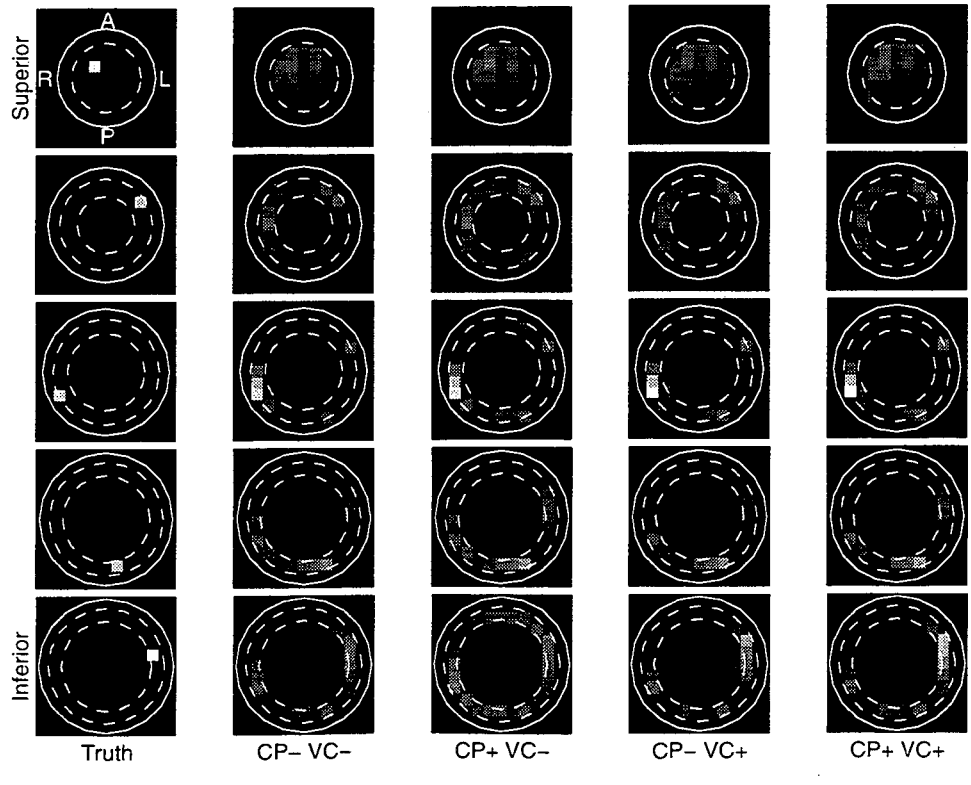


Figure 11.13 Reconstruction of multiple sources—low noise. Signal averaging has been used to reduce the measurement noise by a factor of five.

Figure 11.13 shows the reconstruction for the same five sources, except that the noise has been reduced by a factor of 5 to simulate repeating the experiment 25 times ard averging the results; the assumed noise level has not been changed. In this case all five of the sources are found in the reconstruction, although not all of them are well localized. The fact that reducing the noise improves the reconstruction only slightly indicates that the primary contribution to the reconstruction error is the width of the point spread function rather than measurement noise.

The source locations used in these simulations were well separated but were still not always sharply localized. It is likely that sources closer together would not be clearly resolved.

11.4 Conclusions

An OCLIM reconstruction will provide useful results even when the true source locations do not exactly match the modelled source locations, when there are more modelled sources than detectors, and when there are only a few true sources of large amplitude rather than many small sources.

If the true sources are strong and not too close together, the OCLIM reconstruction will often resolve several simultaneous true sources, although this appears to require a somewhat better SNR that is available in the model problem. The ability to detect multiple simultaneous sources is a useful improvement over moving dipole methods.

The point spread functions of the true sources are generally rather broad, with a FWHM of perhaps 1–3 cm; thus it is not possible to resolve sources placed any closer together.

The inclusion of volume currents in the forward problem provides a definite improvement in the energy concentration in the reconstruction of a dipole source, and restricting the source model to the cortex provides a small but consistent improvement. Thus including volume currents and source priors can be expected to provide better source localization. Note, however, that if radial magnetometers are used with a spherical conductor head model, that the volume currents have no effect on the reconstructions.

Chapter 12

Summary and Conclusions

This final chapter summarizes the contributions made by this dissertation, offers some conclusions as to their value in biomagnetic imaging, indicates some unresolved problems, and presents some recommendations for further research.

12.1 Contributions

This dissertation makes the following contributions to the solution of the linear inverse problem with statistical priors, including the biomagnetic source imaging problem:

- A proof that, under the usual assumptions for the biomagnetic forward problem, fixing the position of the sources makes the measured flux values linear functions of the unknown source amplitudes. The usual assumptions are that it is quasistatic (propagation delays and displacement currents can be ignored) and that Ohm's law holds for biological tissues (current is proportional to electric field). In consequence, the biomagnetic inverse problem becomes a linear inverse problem and it is possible to compute the forward transfer matrix for any physical model by solving the forward problem for unit sources. Furthermore, it is possible to reconstruct the primary current sources alone rather than the total current.

- The derivation of the optimal constrained linear inverse method (OCLIM), which is the minimum-mean square error estimator for the linear inverse problem with statistical prior information.

- The demonstration that the OCLIM estimator contains as special cases the minimum-norm least squares (MNLS), optimally truncated pseudoinverse (OTPIM), optimally weighted pseudoinverse (OWPIM), first-order Tikhonov regular-

ized pseudoinverse, and the Wiener estimator for certain classes of priors and that it provides smaller reconstruction error in the general case. Furthermore, the optimal truncation, weighting, or regularization parameter is directly related to the signal-to-noise ratio. The maximum a posterior estimator is also a special case under the assumption of Gaussian statistics.

- The definition of a new measure of reconstruction quality, called the surprise, which is the log likelihood of the reconstructed source distribution relative to the assumed priors; the surprise indicates how implausible the reconstruction is relative to the priors.

- The statistical characterization of the surprise, the reconstruction error (reconstruction versus truth), and the residual (reconstructed versus actual measurements) for both the general and Gaussian cases.

- The derivation of confidence limits for the reconstructed source amplitudes under the assumption of Gaussian priors and the discovery that in many cases it is more efficient to estimate the confidence limits by Monte Carlo methods than to compute them directly.

- The development of efficient algorithms for the numerical computation of the OCLIM estimate and its quality estimates for both instantaneous and time series data. The identification of assumptions on the spatiotemporal structure of the source and noise priors which allow the computation to be speeded up by several orders of magnitude.

- A demonstration that the OCLIM reconstruction yields useful results in a realistic neuromagnetic source imaging problem, even when the true sources do not fall exactly on the modelled sources and when the true source statistics do not exactly match the assumed statistics. Furthermore, OCLIM can often resolve multiple simultaneously active sources in the brain.

- The discovery that misinformative priors can yield less accurate reconstructions than uninformative priors, and the development of a statistical test that can sometimes detect misinformative priors. This test, however, may not appear to work reliably at the low signal-to-noise ratios characteristic of biomagnetic imaging. If the test is not reliable, one important consequence is that an OCLIM

reconstruction cannot be taken to rule out the existence of sources not included in the model. That is, the reconstruction may be taken to suggest source locations but should not be used to prove the existence or non-existence of a source.

The problem of statistically characterizing the reconstruction quality has led to two other contributions which may be more widely applicable.

- The definition of a generalized χ^2 distribution in which arbitrary means, variances, and covariances are allowed, and the derivation of the mean, variance, and characteristic function of that distribution. The same distribution suffices to describe the distribution of $\mathbf{x}^T\mathbf{A}\mathbf{x}$ or $||\mathbf{x}||^2$ for any Gaussian random vector $\mathbf{x}$ and symmetric positive semidefinite matrix $\mathbf{A}$.

- The development of an FFT-based algorithm to compute the probability density function of a general continuous non-negative univariate random variable given its characteristic function. The derivation of an L_2 error bound for this algorithm and the demonstration that the algorithm converges provided that the characteristic function and the density function decay sufficiently rapidly.

Some less important contributions include:

- A demonstration that including volume currents in the forward problem yields better reconstructions.

- A precise physical interpretation of the common assumption that the covariance matrix of a random vector is nonsingular.

- A proof that estimating the amplitudes of uninteresting "nuisance" sources will not improve the estimates of the desired sources, provided that the nuisance sources are taken into account in computing the measurement covariance.

- A proof that the optimal estimate subject to a linear constraint can be obtained by projecting the unconstrained optimal estimate into the null space of the constraint matrix.

- A demonstration that the estimates of reconstruction quality provided by OCLIM can be used to evaluate and compare possible magnetometer array designs.

12.2 Unresolved Problems

The main thrust of this dissertation has been to develop the mathematical theory needed to incorporate statistical prior information in the linear inverse problem and, to a lesser extent, to develop efficient computational algorithms for this problem. Except for Chapter 11, the various simulations serve primarily to validate the mathematical theory. The simulations of Chapter 11 are best regarded as a pilot experiment to determine whether or not the use of OCLIM and prior information is sufficiently promising to justify further development. It appears that they are, but there still remain some serious practical and theoretical difficulties in the way of clinical or research application.

12.2.1 Estimating the Priors

This dissertation has focused on the use of prior information to improve reconstruction accuracy, *given* the appropriate prior information, but it does not address the problem of obtaining the necessary prior information. Obviously, any practical application of the methods developed here will require the development of methods for actually estimating the source and noise statistics. Given that misinformative priors can yield an erroneous reconstruction, it will be important to be able to demonstrate that the priors are in fact correct for some defined population. It is necessary in addition to explore the effects of slight errors in the priors.

12.2.2 Misinformative Priors

The use of prior information, carefully estimated or not, carries with it the risk that the actual data set is not drawn from the population defined by the priors. If so, then any reconstruction using such misinformative priors is apt to be misleading, if not outright wrong. Under favorable conditions, such as a high signal-to-noise ratio, strong unexpected sources well separated from the modelled sources, and a restrictive source model, the τ^2 test often detects the presence of the unexpected sources and warns against trusting the reconstruction with priors. It is not yet clear whether or not magnetoencephalography offers sufficiently favorable conditions that

it is possible to consistently detect unexpected sources. Given that an unexpected source might not be detected (and the non-uniqueness of the solution), source reconstructions from MEG data cannot be regarded as definitive in showing the existence or non-existence of a source at some specified location. That is, they may be used to suggest a source hypothesis to be tested by other methods but cannot, at least in their present form, be used to critically test a source hypothesis.

12.2.3 Hypothesis Testing

The biomagnetic source imaging problem is essentially singular in that there are many different source distributions consistent with the data. Using prior information to choose one of these possible distributions as "the" solution does not provide a satisfactory means of testing a hypothesis that there is or is not an active source at a specified location. This inability to prove or disprove the existence of a particular source limits the applicability of biomagnetic source imaging in medicine and cognitive research.

12.2.4 Computational Cost

The computation effort required to initialize an OCLIM reconstruction grows as the ninth power of the desired resolution, since the number of possible sources in a given volume is proportional to the inverse cube of the source spacing, and the cost of the Cholesky decomposition used in OCLIM grows as the cube of the number of sources. Thus it becomes very expensive to do high-resolution reconstructions. It is also necessary to solve the forward problem for each possible source; if a finite element or boundary element model is used and there are many sources, this may also add substantially to the computational cost.

12.3 Conclusions

OCLIM is potentially valuable in biomagnetic imaging, since it allows the reconstruction of multiple simultaneously active sources under realistic conditions; it still needs to be demonstrated with real experimental data rather than purely by simulation. The inclusion of volume currents in the forward model and of prior

knowledge as to the possible location and amplitude of the sources improves the resolution and accuracy of the reconstruction.

A disadvantage of prior information, however, is that it can introduce reconstruction errors if the assumed priors do not truly represent the population from which the data set is taken or the actual data set is an outlier in this population. This dissertation develops a statistical test for this possibility but it does not detect all such cases. Until this problem of misinformative priors is better understood, OCLIM and prior information should be used conservatively, to suggest source hypotheses but not to critically test them. The fact that many different source distributions can yield the same external field measurements also suggests than any magnetic source reconstruction should be used conservatively.

An alternative approach to the use of prior information might be applicable in cognitive research. Other functional imaging methods such as PET and fMRI can localize regions of elevated metabolism in the brain, which are presumed to be the regions involved in processing stimuli, but cannot resolve the temporal course of activation. If the source model used in OCLIM includes only the relatively few regions that appear to be active, then OCLIM can be used to test the hypothesis that only these regions are active and determine the time course of activity. This approach has not be explored in this dissertation but might prove effective.

The mathematical approach taken in OCLIM is quite general and can perhaps be extended to other inverse problems in medical imaging including positron emission tomography (PET) and single photon emission computed tomography (SPECT).

References

[1] Abate, J. and W. Whitt. Numerical inversion of probability generating functions. *Operations Research Letters*, 12(4):245–251, October 1992.

[2] Abramowitz, M. and I. A. Stegun, editors. *Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables*, volume 55 of *Applied Mathematics Series*. National Bureau of Standards, 1964.

[3] Aine, C. J., I. Bodis-Wollner, and J. S. George. Generators of visually evoked neuromagnetic responses: Spatial-frequency segregation and evidence for multiple sources. In Sato [88], pages 141–155.

[4] Anderson, E., Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, and D. Sorenson. *LAPACK Users' Guide*. Society for Industrial and Applied Mathematics, Philadelphia, 1992.

[5] Barr, R. C., T. C. Pilkington, J. P. Boineau, and M. S. Spach. Determining surface potentials from current dipoles, with application to electrocardiography. *IEEE Trans. Biomed. Eng.*, BME–13(2):88–92, April 1966.

[6] Baule, G. and R. McFee. Detection of the magnetic field of the heart. *Am. Heart J.*, 66(1):95–96, July 1963.

[7] Baule, G. and R. McFee. Theory of magnetic detection of the heart's electrical activity. *J. Appl. Physics*, 36(6):2066–73, June 1965.

[8] Bayley, R. H. *Biophysical Principles of Electrocardiology*, volume I of *Electrocardiographic Analysis*. Paul Hoeber, Inc., New York, 1958.

[9] Belov, A. G. and V. Y. Galkin. Algoritmy vychisleniya funktsii veroyatnosteĭ obobshchennogo puassonovskogo raspedeleniya [Algorithms for computing the probability function of the generalized Poisson distribution]. In Tikhonov, A. N. and A. A. Samarskiĭ, editors, *Aktual'nye Voprosy Prikladnoĭ Matematiki*, pages 18–28. Izdatel'stvo Moskovskogo Universiteta, 1989.

[10] Beyer, W. H. *CRC Standard Mathematical Tables.* CRC Press, 1981.

[11] Bohman, H. Numerical Fourier inversion. In Kahn, P. M., editor, *Computational Probability*, pages 257–260. Academic Press, Inc., 1980.

[12] Brody, D. A., J. C. Bradshaw, and J. W. Evans. A theoretical basis for determining heart-lead relationships of the equivalent cardiac multipole. *IRE Trans. Biomed. Eng.*, BME–8(2):139–143, April 1961.

[13] Budiman, J. and D. S. Buchanan. An alternative to the biomagnetic forward problem in a realistically shaped head model, the "weighted vertices". *IEEE Trans. Biomed. Eng.*, 40(10):1048–1053, October 1993.

[14] Budinger, T. F., G. T. Gullberg, and R. H. Huesman. Emission computed tomography. In Herman, G. T., editor, *Image Reconstruction from Projections*, volume 32 of *Topics in Applied Physics*, chapter 5, pages 147–246. Springer-Verlag, 1979.

[15] Castleman, K. R. *Digital Image Processing.* Prentice-Hall, Inc., 1979.

[16] Cavers, J. K. On the fast Fourier transform inversion of probability generating functions. *J. Inst. Math. Appl.*, 22:275–282, 1978.

[17] Cohen, D. Magnetoencephalography: Evidence of magnetic fields produced by alpha-rhythm currents. *Science*, 161:784, 1968.

[18] Cohen, D. Magnetoencephalography: Detection of the brain's electrical activity with a superconducting magnetometer. *Science*, 175:664, 1972.

[19] Cohen, D., E. A. Edelsack, and J. E. Zimmerman. Magnetocardiograms taken inside a shielded room with a superconducting point-contact magnetometer. *Appl. Phys. Lett.*, 16:278, 1970.

[20] Cohen, J. D. Noncentral chi-square: Some observations on recurrence. *The American Statistician*, 42:120–122, 1988.

[21] Cooley, J. W., P. A. W. Lewis, and P. D. Welch. Application of the fast Fourier transform to computation of Fourier integrals, Fourier series, and convolution integrals. *IEEE Trans. Audio Electroacoustics*, 15(2):79–84, June 1967.

[22] Cuffin, B. N. Effects of head shape on EEG's and MEG's. *IEEE Trans. Biomed. Eng.*, 37(1):44–52, Jan 1990.

[23] Cuffin, B. N. and D. Cohen. Magnetic fields of a dipole in special volume conductor shapes. *IEEE Trans. Biomed. Eng.*, BME–24:372–381, July 1977.

[24] Daigle, J. N. Queque length distributions from probability generating functions via discrete Fourier transforms. *Operations Research Letters*, 8(4):229–236, August 1989.

[25] De Munck, J. C., B. W. van Dijk, and H. Spekreijse. Mathematical dipoles are adequate to describe realistic generators of human brain activity. *IEEE Trans. Biomed. Eng.*, 35(11):960–966, November 1988.

[26] Ding, C. G. Computing the non-central χ^2 distribution function. *Applied Statistics*, 41(2):478–482, 1992.

[27] Erné, S. N. and M. Hoke. Short-latency evoked magnetic fields from the human auditory brainstem. In Sato [88], pages 167–176.

[28] Fagaly, R. L. Neuromagnetic instrumentation. In Sato [88], pages 11–32.

[29] Fender, D. H. Models of the human brain and the surrounding media: Their influence on the reliability of source localization. *Journal of Clinical Neurophysiology*, 8(4):381–90, October 1991.

[30] Ferguson, A. S., D. Vardy, R. Hren, and G. Stroink. A regularized minimum norm method for calculating distributions of source currents on epicardial surfaces. Poster presentation at the 9th International Conference on

Biomagnetism, Vienna, August 1993. Expected to be published by Elsevier Science Publishers, Amsterdam, in 1994.

[31] Ferguson, A. S., X. Zhang, and G. Stroink. A complete linear discretization for calculating the magnetic field using the boundary element method. *IEEE Trans. Biomed. Eng.*, 41(5):455–460, May 1994.

[32] Foster, K. R. and H. P. Schwan. Dielectric properties of tissues and biological materials: A critical review. *CRC Crit. Rev. Biomed. Eng.*, 17:25–104, 1989.

[33] Frank, F. Electric potential produced by two point sources in a homogeneous conducting sphere. *J. Appl. Phys.*, 23:1225–1228, November 1952.

[34] Franklin, J. N. Well-posed extensions of ill-posed linear problems. *J. Math. Anal. Appl.*, 31:682–716, 1970.

[35] Geselowitz, D. B. On the magnetic field generated outside an inhomogeneous volume conductor by internal current sources. *IEEE Trans. Magn.*, MAG–6(2):346–347, 1970.

[36] Geselowitz, D. B. On bioelectric potentials in an inhomogenous volume conductor. *Biophys. J.*, 7(1):1–11, January 1967.

[37] Gloor, P. Neuronal generators and the problem of localization in electroencephalography: Application of volume conductor theory to electroencephalography. *J. Clin. Neurophysiol.*, 2(4):327–354, 1985.

[38] Golub, G. H. and C. F. Van Loan. *Matrix Computations.* Johns Hopkins University Press, second edition, 1989.

[39] Grübel, R. The fast Fourier transform algorithm in applied probability theory. *Nieuw Archief voor Wiskunde—Vierde Serie*, 7(3):289–300, November 1989.

[40] Hämäläinen, M. S. and J. Sarvas. Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. *IEEE Trans. Biomed. Eng.*, 36(2):165–171, February 1989.

[41] Hari, R. and R. J. Ilmoniemi. Cerebral magnetic fields. *CRC Crit. Rev. Biomed. Eng.*, 14(2):93–126, 1986.

[42] Helstrom, C. W. Image restoration by the method of least squares. *J. Opt. Soc. Am.*, 57(3):297–303, 1967.

[43] Horacek, B. M. Digital model for studies in magnetocardiography. *IEEE Trans. Magn.*, MAG–9(3):440–444, September 1973.

[44] Jain, A. K. *Fundamentals of Digital Image Processing.* Prentice-Hall, Inc., 1989.

[45] Jeffs, B., R. Leahy, and M. Singh. An evaluation of methods for neuromagnetic image reconstruction. *IEEE Trans. Biomed. Eng.*, BME–34(9):713–723, September 1987.

[46] Jin, J. *The Finite Element Method in Electromagnetics.* John Wiley & Sons, 1993.

[47] Kak, A. C. and M. Slaney. *Principles of Computerized Tomographic Imaging.* IEEE Press, 1988.

[48] Karlon, W. J. Defibrillation current density distributions: A three-dimensional finite element model of the canine thorax. Master's thesis, Boston University, 1991.

[49] Kraus, J. D. *Electromagnetics.* McGraw-Hill Book Company, Inc., third edition, 1984.

[50] Lütkenhöner, B., K. Lehnertz, M. Hoke, and C. Pantev. On the biomagnetic inverse problem in the case of multiple dipoles. *Acta Oto-Laryngologica*, Suppl 491:94–105, 1991.

[51] Lynn, M. S. and W. P. Timlake. The use of multiple deflations in the numerical solution of singular systems of equations with applications to potential theory. *SIAM J. Numer. Anal.*, 5(2):303–322, 1968.

[52] Maindonald, J. M. *Statistical Computation.* John Wiley & Sons, 1984.

[53] Mäkelä, J. and R. Hari. Long-latency auditory evoked magnetic fields. In Sato [88], pages 177–191.

[54] Marquardt, D. W. An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Indust. Appl. Math.*, 11(2):431–441, June 1963.

[55] Marquardt, D. W. Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation. *Technometrics*, 12(3):591–612, August 1970.

[56] McFee, R. and F. D. Johnston. Electrocardiographic leads, I. *Circulation*, 8:554–568, October 1953.

[57] McFee, R. and F. D. Johnston. Electrocardiographic leads, II. *Circulation*, 9:255–266, February 1954.

[58] McFee, R. and F. D. Johnston. Electrocardiographic leads, III. *Circulation*, 9:868–880, June 1954.

[59] Meijs, J. W., O. W. Weier, M. J. Peters, and A. van Oosterom. On the numerical accuracy of the boundary element method. *IEEE Trans. Biomed. Eng.*, 36(10):1038–49, October 1989.

[60] Miller, C. E. and C. S. Henriquez. Finite element analysis of bioelectric phenomena. *CRC Crit. Rev. Biomed. Eng.*, 18(3):207–33, 1990.

[61] Moore, E. H. On the reciprocal of the general algebraic matrix. *Bull. Am. Math. Soc.*, 26:389 and 394–395, 1920.

[62] Morrison, D. F. *Multivariate Statistical Methods.* McGraw-Hill Book Company, Inc., 1967.

[63] Nenonen, J., C. J. Purcell, B. M. Horacek, G. Stroink, and T. Katila. Magnetocardiographic functional localization using a current dipole in a realistic torso. *IEEE Trans. Biomed. Eng.*, 38(7):658–664, July 1991.

[64] Nunez, P. L. *Electric Fields of the Brain: The Neurophysics of Electroencephalography.* Oxford University Press, New York, 1981.

[65] Oh, S., C. Ramon, R. J. Marks II, A. C. Nelson, and M. G. Meyer. Resolution enhancement of biomagnetic images using the method of alternating projections. *IEEE Trans. Biomed. Eng.*, BME–40(4):323–328, April 1993.

[66] Okada, R. H. Potentials produced by an eccentric current dipole in a finite-length circular conducting cylinder. *IRE Trans. on Medical Electronics*, ME–7:14–19, December 1956.

[67] Penrose, R. A generalized inverse for matrices. *Proc. Cambridge Philos. Soc.*, 51:406–413, 1955.

[68] Penrose, R. On best approximate solutions of linear matrix equations. *Proc. Cambridge Philos. Soc.*, 52(1):17–19, 1956.

[69] Plonsey, R. Current dipole images and reference potentials. *IEEE Trans. Biomed. Eng.*, BME–10(1):3–8, January 1963.

[70] Plonsey, R. Reciprocity applied to volume conductors and the ECG. *IEEE Trans. Biomed. Eng.*, BME–10(1):9–12, January 1963.

[71] Plonsey, R. and D. G. Fleming. *Bioelectric Phenomena.* McGraw-Hill Book Company, Inc., 1969.

[72] Popper, K. R. *The Logic of Scientific Discovery.* Harper & Row, second edition, 1968.

[73] Posten, H. O. Computer algorithms for the classical distribution functions. In Grey, D. R., P. Holmes, V. Barnett, and G. M. Constable, editors, *Proceedings of the First International Conference on the Teaching of Statistics*, volume 1, pages 313–330, Sheffield, U. K., 1982. University of Sheffield Printing Unit.

[74] Posten, H. O. An effective algorithm for the noncentral chi-squared distribution function. *American Statistician*, 43(4):261–263, November 1989.

[75] Purcell, C. J. and G. Stroink. Moving dipole inverse solutions using realistic torso models. *IEEE Trans. Biomed. Eng.*, 38(1):82–84, January 1991.

[76] Purcell, C. J., G. Stroink, and B. M. Horacek. Effect of torso boundaries on electric potential and magnetic field of a dipole. *IEEE Trans. Biomed. Eng.*, 35(9):671–8, September 1988.

[77] Ragan, P. Magnetically induced electric fields and current densities in the canine heart: A finite element study. Master's thesis, Boston University, 1994.

[78] Ragan, P. M., W. Wang, and S. R. Eisenberg. Magnetically induced currents in the canine heart: A finite element study, 1994. Preprint.

[79] Ramon, C., M. G. Meyer, A. C. Nelson, F. A. Spelman, and J. Lamping. Simulation studies of biomagnetic computed tomography. *IEEE Trans. Biomed. Eng.*, BME–40(4):317–322, April 1993.

[80] Reite, M. Magnetoencephalography in the study of mental illness. In Sato [88], pages 207–222.

[81] Ricci, G. B. Italian contributions the magnetoencephalograhic studies of the epilepsies. In Sato [88], pages 247–260.

[82] Robertson, G. H. Computation of the noncentral chi-square distribution. *Bell System Technical Journal*, 48:201–207, 1969.

[83] Rossini, P. M. and R. Traversa. Somatosensory evoked fields in magnetoencephalography: Basic principles and applications. In Sato [88], pages 157–166.

[84] Roth, B. J. Biomagnetic studies of peripheral nerves and skeletal muscle. In Sato [88], pages 101–117.

[85] Rush, S., J. Abildskov, and R. McFee. Resistivity of body tissues at low frequencies. *Circulation Research*, 12:40–50, 1963.

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

[87] Sato, S. Epilepsy research: NIH experience. In Sato [88], pages 223–230.

[88] Sato, S., editor. *Magnetoencephalography*, volume 54 of *Advances in Neurology*. Raven Press, New York, 1990.

[89] Sato, S., M. Balish, and R. Muratore. Principles of magnetoencephalography. *J. Clin. Neurophysiol.*, 8(2):144–156, April 1991.

[90] Shepp, L. A. and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Trans. Med. Imaging*, MI-1(2):113–122, October 1982.

[91] Shim, Y. S. and Z. H. Cho. SVD pseudoinversion image reconstruction. *IEEE Trans. ASSP*, ASSP-29(4):904–909, August 1981.

[92] Singh, M., D. Doria, V. W. Henderson, G. C. Huth, and J. Beatty. Reconstruction of images from neuro-magnetic fields. *IEEE Trans. Nuclear Sci.*, NS–31(1):585–589, 1984.

[93] Smith, W. E., W. J. Dallas, W. H. Kullman, and H. A. Schlitt. Linear estimation theory applied to the reconstruction of a 3-D vector current distribution. *Appl. Opt.*, 29(5):658–667, February 1990.

[94] Strang, G. *Linear Algebra and Its Applications*. Harcourt Brace Jovanovich, 1988.

[95] Stroink, G. Cardiomagnetic imaging. In Zaret, B. L., L. Kaufman, A. S. Berson, and R. A. Dunn, editors, *Frontiers in Cardiovascular Imaging*, chapter 11, pages 161–177. Raven Press, New York, 1993.

[96] Sutherling, W. W. and D. S. Barth. Magnetoencephalography in clinical epilepsy studies: The UCLA experience. In Sato [88], pages 231–245.

[97] Thevenet, M., O. Bertrand, F. Perrin, T. Dumont, and J. Pernier. The finite element method for a realistic head model of electrical brain activities: Preliminary results. *Clin. Phys. Physiol. Meas.*, 12(Suppl A):89–94, 1991.

[98] Thisted, R. A. *Elements of Statistical Computation: Numerical Computation*. Chapman & Hall, 1988.

[99] Tikhonov, A. N. and V. Y. Arsenin. *Solutions of Ill-posed Problems*. Halsted Press, 1977.

[100] van Oosterom, A. History and evolution of methods for solving the inverse problem. *J. Clin. Neurophysiol.*, 8(4):371–380, October 1991.

[101] van Oosterom, A. Mathematical aspects of source modeling. *Acta Oto-Laryngologica*, Suppl 491:70–79, 1991.

[102] van Oosterom, A. and J. Strackee. The solid angle of a plane triangle. *IEEE Trans. Biomed. Eng.*, BME–30(2):125–126, February 1983.

[103] Van Veen, B., J. Joseph, and K. Hecox. Localization of intra-cerebral sources of electrical activity via linearly constrained minimum variance spatial filtering. In *IEEE Sixth SP Workshop on Statistical Signal and Array Processing Conference Proceedings*, pages 526–529, Victoria, BC, Canada, 1992. IEEE Press.

[104] Van Veen, B. D. and K. M. Buckley. Beamforming: A versatile approach to spatial filtering. *IEEE ASSP Magazine*, pages 4–24, April 1988.

[105] Wang, J.-Z. Minimum-norm least-squares estimation: Magnetic source images for a spherical model head. *IEEE Trans. Biomed. Eng.*, BME–40(4):387–396, April 1993.

[106] Wang, J.-Z., S. J. Williamson, and L. Kaufman. Magnetic source images determined by a lead-field analysis: The unique minimum-norm least-squares estimation. *IEEE Trans. Biomed. Eng.*, BME–39(7):665–675, July 1992.

[107] Weinberg, H., D. Cheyne, and D. Crisp. Electroencephalographic and magnetoencephalographic studies of motor function. In Sato [88], pages 193–205.

[108] Wilson, F. N. and R. H. Bayley. The electric field of an eccentric dipole in a homogeneous spherical conducting medium. *Circulation*, 1:84–92, January 1950.

[109] Yan, Y., P. L. Nunez, and R. T. Hart. Finite-element model of the human head: scalp potentials due to dipole sources. *Medical and Biological Engineering and Computing*, 29(5):475–481, September 1991.

List of Symbols

Here are listed and defined the mathematical symbols used repeatedly in this dissertation. Nonce symbols used only briefly are omitted. The symbols are listed in alphabetical order, with Greek letters listed as if spelled out.

$\mathbf{A}$ $= \mathsf{E}\,\mathbf{q}\mathbf{q}^T$. The a priori covariance matrix of the source amplitudes.

$\hat{\mathbf{A}}$ $= \mathrm{Var}(\mathbf{q} \mid \mathbf{b})$. The a posteriori covariance matrix of the source amplitudes.

$\vec{A}(\vec{r})$ The magnetic vector potential as a function of position $\vec{r}$.

a_k $= \mathbf{v}_k^T\mathbf{q}$. The component of $\mathbf{q}$ in the direction of the kth right singular vector of $\mathbf{F}$.

α_{ij}^2 $= \mathsf{E}\,q_iq_j$. The a priori cross-covariance of source amplitudes q_i and q_j. The (i,j)th entry of $\mathbf{A}$.

$\hat{\alpha}_{ij}^2$ $= \mathsf{E}\,\hat{q}_i\hat{q}_j$. The a posteriori cross-covariance of the estimated source amplitudes $\hat{q}_i$ and $\hat{q}_j$. The (i,j)th entry of $\hat{\mathbf{A}}$.

α_k^2 $= \mathsf{E}\,a_k^2 = \mathsf{E}\,\|\mathbf{v}_k^T\mathbf{q}\|^2$. The variance of a_k.

α False positive rate.

$\vec{B}(\vec{r})$ Magnetic field as a function of position.

$\vec{B}_n(\vec{r})$ Magnetic field due to the nth source.

b_m Magnetic field measurement (including noise) at the mth detector.

$\mathbf{b}$ The vector of field measurements b_m.

$\mathbf{B}$ $= \mathsf{E}\,\mathbf{b}\mathbf{b}^T = \mathbf{F}\mathbf{A}\mathbf{F}^T + \mathbf{F}\mathbf{\Gamma} + \mathbf{\Gamma}^T\mathbf{F}^T + \mathbf{\Sigma}$. The covariance of the measurement vector $\mathbf{b}$.

β_n^2 $= \mathsf{E}(b_m - w_m)^2$. The variance of the signal at the mth detector.

β False negative rate.

$\mathbf{C_{xy}}$ The covariance matrix of the (zero-mean) random vectors $\mathbf{x}$ and $\mathbf{y}$, defined as $\mathsf{E}\,\mathbf{x}\mathbf{y}^T$.

c_k The kth weighting coefficient in the weighted pseudoinverse.

$\mathbf{A}^C$ The Cholesky factor of the symmetric positive definite matrix $\mathbf{A}$. That is, the unique lower triangular matrix $\mathbf{R}$ such that $\mathbf{A} = \mathbf{R}\mathbf{R}^T$.

χ^2 $= (\mathbf{b} - \mathbf{F}\hat{\mathbf{q}})^T\mathbf{\Sigma}^{-1}(\mathbf{b} - \mathbf{F}\hat{\mathbf{q}})$. The residual difference between the measured and reconstructed magnetic fields.

E Expectation of a random variable or a random vector.

$\vec{E}(\vec{r})$ Electric field as a function of position $\vec{r}$.

$\mathbf{e}$ $= \mathbf{q} - \hat{\mathbf{q}}$. The measurement error, defined as the difference between the true and reconstructed current distributions.

$\mathbf{e}_n$ The nth standard basis vector, which contains 1 in position n and 0 everywhere else.

e^2 $= \|\mathbf{e}\|^2 = \|\mathbf{q} - \hat{\mathbf{q}}\|^2$. The reconstruction error, defined as the squared difference between the true and reconstructed current distributions.

η^2 $= \mathsf{E}\,\|\mathbf{e}\|^2 = \mathsf{E}\,\|\mathbf{q} - \hat{\mathbf{q}}\|^2$. The mean reconstruction error.

F_{mn} $= \partial b_m / \partial q_n$. Coefficient relating the field measurement b_m to the source amplitude q_n.

$\mathbf{F}$ The forward transfer matrix with entries $F_{mn} = \partial b_m / \partial q_n$ relating the field measurement vector $\mathbf{b}$ to the source amplitude vector $\mathbf{q}$.

$\mathbf{F}^\dagger$ The Moore-Penrose inverse of $\mathbf{F}$.

$\|\mathbf{A}\|_F$ The Frobenius norm of the matrix $\mathbf{A}$, defined as $(\sum_{mn} A_{mn}^2)^{1/2}$.

$\mathbf{\Gamma}$ The cross-covariance of the source amplitude vector $\mathbf{q}$ and noise amplitude vector $\mathbf{w}$. Its entries are $\gamma_{ij} = \mathsf{E}\, q_i w_j$.

γ_{ij} $= \mathsf{E}\, q_i w_j$. The covariance of the source amplitude q_i and the noise amplitude w_j.

γ_k $= \mathsf{E}\, a_k s_k$. The covariance of a_k and s_k.

$\mathbf{H}$ A matrix representing a linear estimator of $\mathbf{q}$.

$\mathbf{A}^H$ The complex conjugate (or Hermitian) transpose of the matrix $\mathbf{A}$.

$\vec{J}(\vec{r})$ Current density as a function of position. Equals the sum of a primary current $\vec{J}^p(\vec{r})$ and a volume current $\vec{J}^v(\vec{r})$.

$\vec{J}_n(\vec{r})$ Current density contributed by the nth source. Equals the sum of a primary current $\vec{J}_n^p(\vec{r})$ and a volume current $\vec{J}_n^v(\vec{r})$.

K Rank of $\mathbf{F}$.

K' The lesser of M and N; the maximum possible rank of $\mathbf{F}$.

k Index over the singular vectors of $\mathbf{F}$.

$\mathbf{L}$ $\mathbf{L}\mathbf{L}^T$ is the Cholesky decomposition of $\mathbf{B}$.

λ_k The kth singular value of $\mathbf{F}$.

M Total number of detectors used.

m Index over detectors.

μ_0 Magnetic permeability of vacuum. Equal to $4\pi \times 10^{-7}$ H/m.

N Total number of current sources in model.

$\mathbf{N}$ The temporal noise covariance matrix in time series reconstruction.

n Index over the current sources.

ν *Number* of degrees of freedom in an ordinary or generalized χ^2 distribution.

$\vec{p}$ Position.

$\vec{p}_n$ Position of the nth current source.

$p(x)$ Probability density function of a random variable.

$\phi(\vec{r})$ Electric potential as a function of position $\vec{r}$.

$\phi(t)$ The characteristic function of a random variable.

ψ_m The magnetic flux threading the mth detector coil.

q_n Amplitude of the nth source.

$\mathbf{q}$ The vector of source amplitudes q_n.

$\hat{\mathbf{q}}$ $= \mathbf{Hb}$. An estimate of the source amplitude vector $\mathbf{q}$.

$\rho(\vec{r})$ Resistivity as a function of position.

ρ^2 The surprise, or negative log likelihood of the reconstruction relative to the assumed priors. Defined as $\hat{\mathbf{q}}^T \mathbf{A}^{-1} \hat{\mathbf{q}}$.

$\vec{r}$ Position.

$\vec{r}_m$ Position of the mth detector.

$\mathbf{r}$ The measurement residual $\mathbf{b} - \mathbf{F}\hat{\mathbf{q}}$.

$\mathbf{R}$ $\mathbf{R}\mathbf{R}^T$ is the Cholesky decomposition of $\mathbf{A}$.

$\vec{s}_m$ Unit vector in direction of the field component sampled by the mth detector.

$\mathbf{S}$ $\mathbf{S}\mathbf{S}^T$ is the Cholesky decomposition of $\mathbf{\Sigma}$.

$\mathbf{S}$ The temporal source covariance matrix in time series reconstruction.

s_k $= \mathbf{u}_k^T\mathbf{w}$. The component of $\mathbf{w}$ in the direction of the kth left singular vector of $\mathbf{F}$.

$\mathbf{\Sigma}$ $= \mathsf{E}\,\mathbf{w}\mathbf{w}^T$. The covariance matrix of the noise amplitude vector $\mathbf{w}$.

σ_{ij}^2 $= \mathsf{E}\,w_i w_j$. The covariance of noise amplitudes w_i and w_j. The (i,j)th entry of $\mathbf{\Sigma}$.

σ_k^2 $= \mathsf{E}\,s_k^2$. The variance of s_k.

$\sigma(\vec{r})$ Electrical conductivity as a function of position $\vec{r}$.

τ^2 The negative log likelihood of a given data set $\mathbf{b}$ relative to the assumed priors.

$\mathbf{A}^T$ Transpose of the matrix $\mathbf{A}$.

$\mathrm{Tr}(\mathbf{A})$ Trace of the matrix $\mathbf{A}$.

$\mathbf{u}_k$ The kth left singular vector of $\mathbf{F}$.

$\mathbf{v}_k$ The kth right sigular vector of $\mathbf{F}$.

$\mathbf{W}$ The unitary matrix representing the discrete Fourier transform and defined by $[\mathbf{W}]_{mn} = \frac{1}{\sqrt{N}}\, e^{-j2\pi mn/N}$.

w_m Noise amplitude in the mth detector.

$\mathbf{w}$ The vector $[w_m]$ of noise amplitudes.

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
TECHNICAL INFORMATION DEPARTMENT
BERKELEY, CALIFORNIA 94720