Dynamical EEG Inverse Problem and Causality Analysis of fMRI Data

Okito Yamashita

DOCTOR OF PHILOSOPHY

Department of Statistical Science School of Mathematical and Physical Science The Graduate University for Advanced Studies

> 2003 (School Year)

Contents

Lis	st of l	Figures	viii
Lis	st of 7	Tables	ix
Ab	ostrac	et	xiii
Ι	Dyı	namical EEG Inverse Problem	1
1	Intro 1.1 1.2	oduction Context and Motivation	3 3 6
2	Forw 2.1 2.2	vard problem and Inverse problemForward problemInverse problem2.2.1Dipole method2.2.2Distributed linear solution	7 7 9 10 10
3	Weig 3.1 3.2 3.3 3.4	ghted minimum L_2 -norm solutionSVD based solution.Bayesian interpretation.Estimation of the regularization parameter.3.3.1L-curveJ.3.2GCVGCV.3.3.3ABICABIC.Solution.	15 15 16 17 17 17 18 19 19
4	Dyna 4.1 4.2	amical Inverse ProblemDefinitionDefinitionFormulation4.2.1Penalized least squares form4.2.2State space representation	 23 23 24 25 27

5	Dyr	namical Inverse Solution	31			
	5.1	Kalman filtering algorithm	31			
	5.2	Recursive penalized least squares method	33			
		5.2.1 Algorithm \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots	33			
		5.2.2 Estimation of the regularization parameter λ	34			
		5.2.3 Relation between RPLS method and Kalman filtering	36			
	5.3	Projection Kalman filtering	36			
		5.3.1 Algorithm \ldots	36			
		5.3.2 Derivation \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots	38			
		5.3.3 Property	39			
	5.4	Partitioned Kalman filtering	46			
		5.4.1 Algorithm	46			
		5.4.2 Derivation \ldots	47			
		5.4.3 Remarks \ldots	49			
	5.5	Model	49			
		5.5.1 Dynamics	50			
		5.5.2 Observation noise	51			
		5.5.3 System noise \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots	51			
		5.5.4 Prewhitening	51			
		5.5.5 Parameter estimation	52			
		5.5.6 Model Comparison	52			
6	Bog	ulte	55			
U	6.1 Comparison of Filtering Algorithms					
	6.2	Simulation Study I	57			
	0.2 6 3	Simulation Study I	60			
	6.0	Beal Data Analysis	65			
	0.1		00			
7	Dise	cussion	71			
II	\mathbf{C}	ausality Analysis of fMRI Data	75			
~	.					
8	Intr	roduction	77			
	8.1	Context	77			
	8.2	Organization of Part II	78			
9	Causality analysis in time series					
	9.1	Granger Causality	81			
	9.2	Akaike's Noise Contribution Ratio	82			
		9.2.1 ANCR in frequency domain	82			
		9.2.2 ANCR in time domain	84			

10 Data analysis	87
10.1 Experiment \ldots	87
10.2 Aim	87
10.3 Procedure	88
10.3.1 Specify ROIs	89
10.3.2 Model fitting of time series in ROIs	90
	05
	95
11.1 Effective connectivity via MAR model	95
11.2 Effective connectivity via MEMAR model	97
11.3 Interpretation \ldots	98
12 Discussion 1	107
A Calculation of ABIC for WMN method	L 09
B Some lemmas on multivariate normal distributions	13
Acknowledgments 1	17
References 1	19

List of Figures

1.1	Schematic figure of the forward problem and inverse problem	4
2.1	The piecewise homogeneous volume conductor model of the head.	8
4.1	Schematic comparison between the instantaneous and the dy- namical inverse problem	25
5.1	Schematic of equation (5.60)	41
6.1	Time series of the local current vector on four specific voxels for the comparison of the filtering algorithms	58
6.2	Location of voxels in the analyses of the sections 6.2, 6.3 and 6.4	59
6.3	Spatial distributions of the current vectors	61
6.4	Time series of two sources	61
6.5	Simulated EEG observations at 19 standard electrode positions of the $10/20$ -system	63
6.6	Maximum-intensity projections of the distribution of current sources	64
6.7	Time series of the length of the estimated local current vector at four specific voxels	64
6.8	Clinical EEG recording at 19 standard positions of the $10/20$ - system versus time, obtained from a healthy 8-years old male	01
	child, awake with closed eyes	67
6.9	Logarithm of the parametric spectrum obtained from the esti- mated $AB(2)$ dynamics, versus frequency.	68
6 10	Spatial distributions of the component of the estimated current	00
0.10	vectors in the radial direction of spherical coordinates	69
6.11	Time series of the estimated local current vector in the radial	00
0.11	direction of spherical coordinates at four voxels	70
10.1	Block design of the random dot experiment	88
10.2	Schematic figure of T value truncation	93
10.3	T map via "SPM" and the locations of V1. V5 and PP	94
	\mathbf{r}	

11.1	Time series of V1, V5 and PP in the first session and their
	innovations by MAR model fitting
11.2	Time series of V1, V5 and PP in the second session and their
	innovations by MAR model fitting
11.3	Power spectra and the result of ANCR from the first session
	and the second session $(A1)$
11.4	Bar graph representation of ANCR for 0.018Hz, 0.031 Hz and
	of ANCR in time domain $(A1)$
11.5	Time series of V1, V5 and PP in the first session and their
	innovations by MEMAR model fitting
11.6	Power spectra and the result of ANCR on the task condition
	and the control condition $(A2)$
11.7	Bar graph representation of ANCR for 0.018Hz, 0.031 Hz and
	of ANCR in time domain $(A2)$

List of Tables

6.1	Comparison of the filtering algorithms	57
6.2	Comparison of DynLORETA and LORETA solutions	62
6.3	Model comparison for the real alpha data	66
11.1	Change of ANCR via the MAR model	97
11.2	Change of ANCR via the MEMAR model	98
11.3	Change of the effective connectivity at the frequency f_0 via the	
	MAR model	99

List of Symbols

ABIC	An (or Akaike) Bayesian Information Criterion
AIC	An (or Akaike) Information Criterion
EEG	ElectroEncepharoGrams
fMRI	functional Magnetic Resonance Imaging
GCV	Generalized Cross Validation
LORETA	LOw REsolution brain electromagnetic TomogrAphy
MAR	Multivariate AutoRegressive
RPLS	Recursive Penalized Least Squares
SVD	Singular Value Decomposition
WMN	Weighted Minimum L_2 -Norm
m (hold food fort)	a mater of operaine
\vec{r} (bold-face font)	a vector of any size
r	a 3-dimensional vector
A'	transpose of a matrix A
A	determinant of a matrix A
A^{-1}	inverse of a square matrix A
A^-	generalized pesudoinverse of a matrix A
A^+	Moore-Penrose pesudo inverse of a matrix ${\cal A}$
\overline{A}	a complex conjugate matrix of a matrix A
$A \otimes B$	Kronecker product of A and B
E(X)	expectation of a random variable X
$\operatorname{Var}(X)$	variance of a random variable X
$\operatorname{Cov}(X,Y)$	covariance between two random variables X and Y
$\operatorname{diag}(x_1,\ldots,x_n)$	diagonal matrix with elements x_1, \ldots, x_n
$N(\boldsymbol{\mu}, \Sigma)$	normal distribution with expectation $\pmb{\mu}$ and variance $\boldsymbol{\Sigma}$
$ m{x} _C^2$	L_2 norm of a vector \boldsymbol{x} with metric C ; $\boldsymbol{x}'C\boldsymbol{x}$
I_d	$d \times d$ identity matrix

К	lead field matrix	р.	10
L	3-dimensional Laplacian matrix	р.	12
N_e	number of EEG channels (20 \sim 100)		
N_s	dimension of the current vector (10000~)		

Abstract

Human being has long been challenging to understand functions and organizations of the brain. With striking developments of various measurement apparatus and methodology after twentieth century, we have accumulated not only the knowledge about the mechanism of our brain but also measurements of brain activities from various aspects. In order to make the best use of these data combined with a priori knowledge, the development of statistical methods is indispensable.

Nowadays the functional Magnetic Resonance Imaging (**fMRI**) technique and the electroencephalography (**EEG**) are two common tools for the understanding of human cognition as well as for the clinical diagnosis. By the fMRI technique, the change of regional cerebral blood flow, which is supposed to result from electrical neuronal activities on the corresponding local region, is measured as temporally successive images covering the whole brain volume with high spatial resolution but low temporal resolution. By the EEG, evoke potentials can be measured in several tens positions on the scalp surface with high temporal resolution as a consequence of the transmission of electric currents (a collection of electrical neuronal activities) inside the brain.

In this thesis, for the purpose of analyzing these two kind of the data sets, the methodology in the field of **time series analysis** will be applied and developed. Since these two data sets have distinct properties, the purpose and the tool for analysis are also distinct. Therefore this thesis consists of two parts, the inverse problem of the EEG and the causal analysis for the fMRI data.

In the first part of this thesis, the **dynamical inverse problem** of the EEG generation will be discussed. Since the EEG recording is an indirect observation of electrical sources inside the brain, the inference to localize the sources, called the 'inverse problem' are necessary. In general, in order to solve the inverse problem we have to combine additional information to the observation because it is impossible to uniquely determine the solution from the observation itself. In this thesis, we will consider the dynamical inverse problem so that general spatio-temporal constraints can be incorporated. This aspect has been neglected in many previous studies of the inverse problem of the EEG generation in spite of its importance.

Mathematically the dynamical inverse problem will be formulated as the

state estimation problem. The system equation in the state space representation describes general spatio-temporal constraints. By assuming a parametric model for the dynamics, we can choose in a sense the 'best fitting' constraints onto the solution. In principle both the parameter estimation and the state estimation (the solution) can be done by means of the celebrated Kalman filtering algorithm.

However due to high dimensionality of the state in the EEG application, the difficulty occurs in the computational aspect. As alternatives of ordinary Kalman filtering, the author will propose three approximate filtering algorithms; the recursive penalized least squares (RPLS) method, observable projection Kalman filtering and partitioned (spatio-temporal) Kalman filtering. The different ways of approximation of covariance matrices of the filtered and prediction states are employed in these algorithms. The simulation study will demonstrate similarity of the solutions via three methods in the case of simple dynamics. However the difference of three solutions could become larger when the dynamics becomes complex. It would be necessary to examine the situation of problems and validity of the assumption.

The data analysis of real alpha wave will show two sources located in the occipital region of both the left and right hemisphere, which has been reported in the previous studies. In addition, the estimated dynamics inside and outside the occipital region is observed to differ in periodicity using a regional AR model as the dynamics.

In the latter part of this thesis, the methodology to evaluate the effective connectivity of the fMRI data will be investigated. In the fMRI studies, recently, more attention has been paid to the analysis of the effective connectivity defined as "the influence that one neural system exerts over another" (Friston et al. 1995). In order to accomplish this purpose, the method developed in the multivariate time series analysis will be applied. It is a crucial advantage of this approach that no assumption about the direction of connectivity is required, whereas the structural equation model, the most common approach to evaluate the effective connectivity so far, requires to determine and to restrict the direction of connectivity apriori.

For this purpose, the author proposes to apply the Akaike's noise contribution ratio (ANCR), which quantifies the influence on one time series from another time series. Using the data from the random dot experiment, the change of the connectivity between two conditions will be evaluated by the ANCR as a measure. As a result, the increase of the connectivity on the task condition is observed compared with the connectivity on the control condition.

Part I

Dynamical EEG Inverse Problem

Chapter 1

Introduction

1.1 Context and Motivation

Measurements of electromagnetic fields on the scalp surface provide valuable information about the underlying brain dynamics. By measuring the electrical potential or the magnetic field on several locations of the scalp surface, electroencephalography (EEG) or magenetoencephalography (MEG) are obtained. It is commonly believed that these potentials are generated by electrical currents in the extracellular media, resulting from the electrical and chemical neuronal activity of the brain.

Currently there is considerable interest in localizing non-invasively such electrical sources of the EEG in the brain (source localization problems). The source localization problem can be separated to the forward problem and the inverse problem. The forward problem consists in determining the electromagnetic field at the scalp from a known source configuration. The inverse problem describes the opposite situation: given the volume conductor and the electromagnetic field at the scalp, the location and time course of the sources is sought.

The forward problem is a "simple electromagnetic problem" that can be expressed and solved with Maxwell's equation. Its difficulty lies in modeling the volume conductor and a human head, and lies in solving the resulting partial differential equation after modeling. Once the relationship between brain electrical activity and electromagnetic scalp fields has been established, the inverse problem can be approached. The main difficulty to solve the inverse problem lies in the fact that EEG/MEG observations do not contain a sufficient amount of information to precisely reproduce these current sources. For this reason, there will inevitably be an infinite number of possible inverse solutions consistent with the measurements given. In order to find out a unique and plausible inverse solution, we have to combine additional information in the guise of physiological or physical knowledge about the sources.



Figure 1.1: Schematic figure of the forward problem and inverse problem. Forward problem: Given a configuration of dipoles and a head model, electrical potentials on the head scalp is calculated. Inverse problem: given the volume conductor and the electromagnetic field at the scalp, the location and time course of the sources is sought.

Currently two common approaches have been employed for solving inverse problems; 'dipole method' and 'distributed linear solution' approach. In the 'dipole method' EEG measurements are generated by a relatively small number of focal dipole, each of which can be modeled as a single fixed dipole or re-oriented dipole (Scherg and Ebersole 1994). The idea is to render the inverse problem *overdetermined* by considering fewer unknown parameters than independent measurements available.

In the 'distributed linear solution' approach, a discretization of brain volume into a set of voxels is employed, each of which is considered to be the location of a current source. Corresponding to the number of voxels, the unknowns to be estimated increase and result in highly underdetermined problems but the EEG observation is described by a linear function of current sources. In order to obtain a unique solution, various constraints have been suggested in previous studies: as prominent examples I mention optimal resolution (Backus and Gilbert 1968; Grave de Peralta Menendez et al. 1997; Grave de Peralta Menendez and Gonzalez Andio 1999), L_2 minimum norm (Hämäläinen and Ilmoniemi 1984), L_1 minimum norm (called 'selective minimum norm') (Matsuura and Okabe 1995; Matsuura and Okabe 1996; Uutela et al. 1999) and maximum spatial smoothness (called 'low resolution brain electromagnetic tomography', LORETA) (Pascual-Marqui et al. 1994).

In several papers (Pascual-Marqui and Michel 1994; Pascual-Marqui 1999; Grave de Peralta Menendez and Gonzalez Andio 2002) relative advantages and disadvantages of these approaches have been discussed from a purely spatial point of view; however, these approaches exploit exclusively the information contained in one instantaneous measurement, i.e. the set of voltage measurements obtained from various electrodes at one single instant of time, whereas measurements of the EEG clearly have temporal structure.

Recently, temporal constraints have been taken into consideration for various applications related to inverse problems. For example, in the analysis of electrocardiograms (ECG) an algorithm for solving large-scale least squares problems based on multiple constraints, including explicit spatial and temporal constraints, has been proposed (Brooks et al. 1999). For the reconstruction of current distributions in the EEG/MEG inverse problem, other algorithms for solving the same large-scale least squares problem as mentioned above have been developed (Schmidt et al. 2001; Schmidt and Louis 2002a; Schmidt and Louis 2002b).

As another way of incorporating temporal constraints into the inverse problem, the state space representation has been employed in the analysis of data obtained by Electrical Impedance Tomography (EIT) and Single Photon Emission Tomography (SPET) (Karjalainen et al. 1997; Kaipio et al. 1999; Vauhkonen et al. 2001) and in the analysis of the geodetic data (Segall and Matthews 1997; McGuire and Segall 2003). The state estimation has been done by the use of (extended) Kalman filtering and Kalman smoothing.

In this thesis, the inverse problem of the EEG will be also formulated as the **state estimation problem** by considering time-dependent EEG measurements as a result of current sources evolving according to some dynamics. In the **state space representation**, the dynamics can be explicitly expressed as the transition matrix or function in the system equation. The spatio-temporal constraints can be expressed by combination of the model of the dynamics and the covariance structure of the system noise. In particular the inverse problem formulated in the state space representation would be referred to as the **dynamical inverse problem**.

As specific features of the EEG dynamical inverse problem, we can mention two points; the first is that there is not any known model describing the dynamics as well as the system noise, and the second is that the dimension of the state, corresponding to the number of voxels, is too high to directly apply the Kalman filtering algorithm.

So as to overcome the first point, exploration of the dynamics would be emphasized according to the established procedures of statistical modelling, i.e. by assuming a class of parametric models for the dynamics and comparing these model using the statistical criterion. Although the temporal constraints, as used in these studies so far, refer mainly to the aspect of temporal smoothness, it is important to explore models of the dynamics because a better model would lead to not only improved inverse solutions as well as the expression for time varying phenomenon of the sources.

As to the second point, three approximate filtering algorithms will be proposed as alternatives of the ordinal Kalman filtering algorithm; the recursive penalized least squares (RPLS) method, observable projection Kalman filtering and partitioned (spatio-temporal) Kalman filtering. All the algorithms approximate covariance matrices of the filtered and predicted states in the different ways; in the RPLS method, the previous filtered state variance is neglected, in observable projection Kalman filtering, the projection of the filtered and predicted matrices onto a subspace defined by the singular value decomposition (SVD) of the observation matrix (the lead field matrix) is utilized, and in partitioned (spatio-temporal) Kalman filtering, the elements in the near diagonal part of the filtered and predicted matrices are utilized.

Using both the simulation studies, the validity of the state estimation methods will be demonstrated. The outlined framework of exploring the dynamics will also be shown for the real data of alpha wave.

1.2 Organization of Part I

This part I is organized as follows. In Chap.2, previous studies of the forward problem and the inverse problem are reviewed. In Chap.3, detailed property of an important class of the inverse solution, called the weighted L_2 minimum norm solution (WMN), is discussed. The use of ABIC for estimating the regularization parameter is proposed. In Chap.4, after the definition of the dynamical inverse problem, an appropriate way of incorporating general spatio-temporal constraints is introduced, referring to the study of Schmidt et al. 2001. The state space representation of the dynamical inverse problem is presented. Then the solution of the dynamical inverse problem can be obtained by using one of three approximations of Kalman filtering proposed in Chap.5. The simulation studies and the application to the real data will be demonstrated in Chap.6 followed by the concluding discussion of the part I.

Chapter 2

Forward problem and Inverse problem

The source localization problem consists of two problems; the forward problem and the inverse problem. Mathematically the source localization problem in EEG can be summarized by the equation,

$$\boldsymbol{v} = \mathcal{F}(\vec{r}, \vec{j}) + \boldsymbol{\epsilon}$$

where \boldsymbol{v} , a vector of size $N_e \times 1$, is the potential at the N_e electrodes, \vec{r} and \vec{j} are the source location and moment, $\boldsymbol{\epsilon}$ is the additive noise, and \mathcal{F} is the function linking the source (\vec{r}, \vec{j}) and the potential \boldsymbol{v} . The function \mathcal{F} is the solution of the forward problem and depends only on the head model adopted (in principle). For multiple sources defined by $\vec{r_i}$ and $\vec{j_i}$ (with $i = 1, \dots, N_i$), the source localization problem is written as

$$\boldsymbol{v} = \sum_{i=1}^{N_i} \mathcal{F}(\vec{r}_i, \vec{j}_i) + \boldsymbol{\epsilon}.$$
(2.1)

The inverse problem is consist of the inference about $(\vec{r_i}, \vec{j_i})$ from the potential \boldsymbol{v} and known \mathcal{F} .

In this chapter, various methods proposed to solve the forward problem and the inverse problem of EEG are reviewed briefly. In the literature (Baillet and Mosher 2001), this topic has been reviewed compactly.

2.1 Forward problem

The EEG forward problem can be stated as follows: given the positions, orientations and magnitudes of dipole current sources, as well as the geometry and electrical conductivity of the head volume, calculate the distribution of the electrical potential on the surface of the head (scalp).



Figure 2.1: The piecewise homogeneous volume conductor model of the head.

For the biological signals of interest in EEG, we can assume the timederivatives of the associated electric field are sufficiently small that they can be ignored in Maxwell's equations. Under this quasi-static assumption, the forward problem can be described by Poisson's equation for electrical conduction in the head,

$$\nabla \cdot (\sigma \nabla \boldsymbol{V}) = \nabla \cdot \boldsymbol{J}^{p}, \quad \text{in } \Omega$$
(2.2)

and Neumann boundary conditions on the scalp,

$$\sigma(\nabla \boldsymbol{V}) \cdot \boldsymbol{n} = 0, \quad \text{on} \quad \Gamma_{\Omega} \tag{2.3}$$

where σ is a conductivity tensor and J^p is the primary current which is the quantity of interest in neuroscience. The head volume and the surface of the head are denoted by Ω and Γ_{Ω} , respectively and the unknown V is the electrical potential created by the distribution of primary current inside the brain.

The most commonly used head model (i.e. the geometry and electrical conductivity of the head) assumes that it is made up of a set of nested concentric volumes, each with homogeneous and isotropic conductivity. Under this assumption, Poisson's equation (2.2) will be simplified:

$$\sigma_i \Delta \boldsymbol{V} = \nabla \cdot \boldsymbol{J}^p, \quad \text{in} \ \ \Omega_i \tag{2.4}$$

with the boundary conditions:

$$\sigma_i \frac{\partial \boldsymbol{V}_i}{\partial \boldsymbol{n}} = \sigma_{i+1} \frac{\partial \boldsymbol{V}_{i+1}}{\partial \boldsymbol{n}}, \quad \text{on} \quad S_i$$
(2.5)

$$\boldsymbol{V}_i = \boldsymbol{V}_{i+1}, \quad \text{on } S_i \tag{2.6}$$

where Ω_i denote the volume of *i*th medium, S_i denotes the surface between the volumes Ω_i and Ω_{i+1} , σ_i and V_i denote the conductivity and potential in the volume Ω_i , respectively. Using Green's second identity, we can obtain a surface integral equation for $V(\vec{r})$ (see Chapter 2 of Phillipse 2001 for details),

$$\frac{\sigma_{j+1} + \sigma_j}{2} \boldsymbol{V}(\vec{r}) = \boldsymbol{V}_{\infty} + \frac{1}{4\pi} \sum_i (\sigma_{i+1} - \sigma_i) \int_{S_i} \boldsymbol{V}_i(\vec{r}') \frac{\vec{r} - \vec{r}'}{||\vec{r} - \vec{r}'||^3} \cdot dS', \vec{r} \in S_j$$
(2.7)

where we have assumed all surfaces are smooth, and V_{∞} is the primary potential (i.e. the solution for the infinite homogeneous medium due to the primary current J^p) obtained as

$$\boldsymbol{V}_{\infty} = \frac{1}{4\pi} \int_{\Omega} \boldsymbol{J}^{p}(\vec{r}') \cdot \frac{\vec{r} - \vec{r}'}{||\vec{r} - \vec{r}'||^{3}} d\vec{r}'.$$

Equation (2.7) is a explicit relationship between the primary current J^p and the surface potential on V. In the case of highly symmetrical geometry such as spheres, an analytical solution is possible (Ary et al. 1981; deMunck and Peters 1993; Riera and Fuentes 1998). For general geometry, however, this integral equation can not be solved in a closed-form. Hence the numerical method such as boundary elementary method (BEM) is required for this purpose (Geselowitz 1967; Sarvas 1987; Schlitt et al. 1995; Mosher et al. 1999). Note that the integral equation Eq.(2.7) has been derived with the piecewise homogeneous and isotropic conductivity assumptions. Even without these assumptions, the partial differential equation, Poisson's equation (2.2) can be solved numerically by finite element method (FEM) (Tong and Rosettos 1976; Awada et al. 1997; Klepfer et al. 1997) or finite difference method (FDM) (Rosenfeld et al. 1996; Salheen and Ng 1997) for general geometry and electrical conductivity of the head volume, though these methods demand higher computational cost and may cause higher numerical error.

2.2 Inverse problem

The general EEG inverse problem can be stated as follows: given a set of electric potentials from discrete sites on the surface of the head and associated positions of those measurements as well as the geometry and conductivity of the different regions within the head, calculate the locations, orientation and magnitudes of the electric current sources within the brain.

The difficulty of the attempt to solve this inverse problem arises from the fact that the EEG observations do not contain a sufficient amount of information to precisely reproduce these sources. For this reason, the solution of this inverse problem will be non-unique, i.e. there will be an infinite number of possible inverse solutions consistent with the measurements given. In order to identify a unique solution, we have to employ additional information in the guise of physiological or physical knowledge about the sources. In mathematical terms, this corresponds to imposing additional constraints onto the solution space. Currently two common approaches have been employed for solving inverse problems; 'dipole method' and 'distributed linear solution' approach.

2.2.1 Dipole method

This approach is to consider that EEG measurements are generated by a relatively small number of focal dipole, each of which can be modeled as a single fixed dipole or re-oriented dipole (Scherg and Ebersole 1994). The idea is to render the inverse problem *overdetermined* by considering fewer unknown parameters than independent measurements available. The locations, orientations and strengths of these current dipoles, six parameters in all, can be estimated by minimizing the difference between the predicted and actual EEG measurements;

$$E(\vec{r}_1, \cdots, \vec{r}_{N_i}, \vec{j}_1, \cdots, \vec{j}_{N_i}) = ||\boldsymbol{v} - \sum_{i=1}^{N_i} \mathcal{F}(\vec{r}_i, \vec{j}_i)||^2, \qquad (2.8)$$

where N_i is the number of dipoles assumed, $\vec{r_i}$, $\vec{j_i}$ is the location and the moment (the orientation and the strength) of the *i*th dipole, respectively. The function \mathcal{F} is obtained by solving the forward problem. This minimization requires solving the forward problem for every possible configuration of dipoles. Thus the time required to solve the optimization problem grows exponentially with the number of dipoles, and the global optimum can be found only for models involving few dipoles. In all dipole methods, the solution depends heavily on the number of dipoles assumed but, in general, the actual number of dipoles cannot be determined a priori.

2.2.2 Distributed linear solution

This approach is based on a discretization of brain volume into a set of voxels, each of which is considered to be the location of current dipoles. Because the location of each current dipoles is now fixed, Eq.(2.1) becomes an *underdetermined* but linear problem,

$$\boldsymbol{v} = \mathbf{K}\boldsymbol{j} + \boldsymbol{\epsilon} \tag{2.9}$$

where where the vector $\mathbf{j} = (\vec{j}'_1, \cdots, \vec{j}'_{N_j})'$ represents (the strength and orientation of) the current dipoles at all the N_j locations, where $\vec{j}_i = (j_{x,i}, j_{y,i}, j_{z,i})$ encodes both orientation and strength of the current dipole on the *i*th voxel

(therefore the dimension of \boldsymbol{j} is three times of the number of voxels (denoted by N_s)). The known matrix K of size $N_e \times N_s$ is the **lead field matrix** linking the current sources \boldsymbol{j} to the electrical potential \boldsymbol{v} . In order to obtain a unique solution, various constraints have been suggested.

Weighted L_2 Minimum Norm

One approach of restricting the solution space of the current vector \mathbf{j} is to impose a constraint onto the current vector \mathbf{j} directly. The weighted minimum L_2 -norm (denoted by WMN) solution can be obtained by solving a constrained least squares problem,

$$\{\min_{\boldsymbol{j}} \boldsymbol{j}' W \boldsymbol{j}, \text{under constraint: } \boldsymbol{v} = \mathbf{K} \boldsymbol{j}\}.$$
(2.10)

for any given positive definite matrix W of size $N_s \times N_s$. The weight matrix W determines property of the solution, which is given by:

$$\hat{\boldsymbol{j}} = T\boldsymbol{v}$$

$$T = W^{-1} \mathbf{K}' [\mathbf{K} W^{-1} \mathbf{K}']^{+}$$
(2.11)

where $[KW^{-1}K']^+$ denotes the Moore-Penrose pseudoinverse of $KW^{-1}K'$. Because the solution (2.11) satisfies the measurement equation $\boldsymbol{v} = K\boldsymbol{j}$ exactly, the solution is not robust to the measurement contaminated by observation noise. For avoiding this effect as well as for numerical stabilization, we can consider the following **regularized** problem,

$$\min_{\boldsymbol{j}} \{ ||\boldsymbol{v} - \mathbf{K}\boldsymbol{j}||^2 + \lambda^2 \boldsymbol{j}' W \boldsymbol{j} \}.$$
(2.12)

where λ is the regularization parameter, which make a balance between observation fitting and the side constraint. Then the solution of this problem can be provided as:

$$\hat{\boldsymbol{j}} = (\mathbf{K}'\mathbf{K} + \lambda^2 W)^{-1}\mathbf{K}'\boldsymbol{v}$$

or

$$\hat{\boldsymbol{j}} = W^{-1} \mathbf{K}' (\mathbf{K} W^{-1} \mathbf{K}' + \lambda^2 \mathbf{I})^{-1} \boldsymbol{v}.$$

The WMN solution includes various kind of inverse solutions, corresponding to various choice of the weight matrix W.

- The minimum norm solution corresponds to the case, $W = I_{N_s}$ (Hämäläinen and Ilmoniemi 1984). Thereby the solution with the minimum power can be obtained.
- The 'LOw REsolution brain electromagnetic TomogrAphy' (LORETA) method corresponds to the case, W = L'L, (Pascual-Marqui et al. 1994).

The matrix L, called (3D-discretized) **Laplacian** matrix, is defined by:

$$[M]_{ij} = \begin{cases} 6 & (i = j) \\ -1 & (j \text{ is a neighbor of } i) \\ 0 & (\text{otherwise}) \end{cases}$$
$$L = \frac{1}{6} (M \otimes I_3) \qquad (2.13)$$

The *i*th row vector of L acts as a discrete differentiating operator by forming differences between the nearest neighbors of the *i*th voxel and *i*th voxel itself (i.e. the spatial second differencing operator). Thereby the solution with the maximum spatial smoothness can be obtained. In the literature Pascual-Marqui 1999, this solution was reported to be able to overcome the incapability of correct localization in 3-dimensional solution spaces which had been observed for the minimum norm solution.

• The column weighted minimum norm solution corresponds to the case, $W = \Omega$. The matrix Ω is a diagonal matrix with diagonal elements,

$$\Omega_{ii} = \mathbf{k}'_{\cdot i} \mathbf{k}_{\cdot i},$$

where $\mathbf{k}_{\cdot i}$ is the *i*th column vector of the lead field matrix K. This column normalization can be physically interpreted as the equal size of contribution to measurements from all the sources (Goronidnitsky and Rao 1997).

• Another choice of a diagonal weight matrix is possible by combining the information from another modality such as fMRI (Phillips et al. 2002)

In the next chapter, detailed topics of this WMN solution will be further discussed.

Optimal Resolution

Any linear inverse solution can be written in the form,

$$\hat{\boldsymbol{j}} = T\boldsymbol{v}$$

where a matrix T of size $N_s \times N_e$ is the inverse operator, which we want to find out. Substituting the forward equation (2.9) into this equation and taking expectation of both sides, we obtain

$$\mathbf{E}(\mathbf{\hat{j}}) = T\mathbf{K}\mathbf{j}.\tag{2.14}$$

The matrix $R \equiv T$ K in Eq.(2.14), which is referred to the **resolution matrix** describes how the estimated current source vector \hat{j} is reproduced by the true

current source vector \mathbf{j} . In the ideal case $R = I_{N_s}$, the estimate can completely reproduce the true source, however, this case could never happen because the rank of the resolution matrix is far from the full rank. There have been several studies of finding out an inverse operator T based on the optimal resolution matrix (Backus and Gilbert 1968; Grave de Peralta Menendez et al. 1997). The method of Backus-Gilbert is reviewed here.

The linear inverse solution at the uth voxel is:

$$\hat{j}^u = \boldsymbol{t}'_u \boldsymbol{v} \tag{2.15}$$

where t'_u is the *u*th row of *T*, a vector of size $1 \times N_e$. Substituting Eq.(2.9) into Eq.(2.15) and taking expectation gives:

$$\mathbf{E}(\hat{j}^u) = \mathbf{r}'_u \mathbf{j} \tag{2.16}$$

where:

$$\boldsymbol{r}'_u = \boldsymbol{t}'_u \mathbf{K} \tag{2.17}$$

corresponds to the *u*th row vector of the resolution matrix R. The vector \mathbf{r}'_u , called "averaging kernel" contains information about how the current estimates at *u*th voxel is influenced by all possible sources. Since it is preferable property for the averaging kernel \mathbf{r}'_u to have higher amplitude on the nearby-voxels of u, we form some measure of the width of \mathbf{r}'_u such as,

$$\mathcal{A}_{u} \equiv \sum_{s=1}^{N_{s}} r_{us}^{2} d_{us}^{2}$$

$$= \mathbf{r}'_{u} \operatorname{diag}(d_{u1}^{2}, \cdots, d_{uN_{s}}^{2}) \mathbf{r}_{u}$$

$$= \mathbf{t}'_{u} W_{u} \mathbf{t}_{u} \qquad (2.18)$$

where r_{us} is the sth element of the vector \mathbf{r}'_u , and d_{us} is the Euclidian distance between voxels u, s (defined by $||\vec{r}_u - \vec{r}_s||^2$, where \vec{r}_v is the position of a voxel v), and W_u is given by

$$W_u = \mathbf{K} \cdot \operatorname{diag}\left(d_{1s}^2, \cdots, d_{uN_s}^2\right) \cdot \mathbf{K}'.$$
(2.19)

We can also form a measure of the stability from variance of the estimate \hat{j}^u in Eq.(2.15) as

$$\mathcal{B}_u = \operatorname{Var}(\hat{j}^u) = \boldsymbol{t}'_u \operatorname{Var}(\boldsymbol{v}) \boldsymbol{t}_u = \boldsymbol{t}'_u C_{\boldsymbol{\epsilon}} \boldsymbol{t}_u$$
(2.20)

where $C_{\epsilon} \equiv \operatorname{Var}(\epsilon)$ is the observation noise covariance matrix.

Finally, the demand that r'_u has unit area, leads to the constraint

$$\sum_{s=1}^{N_s} r_{u,s} = \boldsymbol{t}_u \mathbf{k}_{\cdot*} = 1$$
 (2.21)

where $\mathbf{k}_{*} = (\sum_{s=1}^{N_s} \mathbf{K}_{1s}, \cdots, \sum_{s=1}^{N_s} \mathbf{K}_{N_es})$ is a vector of size $N_e \times 1$.

The Backus-Gilbert method consists of finding t_u which minimize

$$\mathcal{A}_u + \lambda \mathcal{B}_u = \mathbf{t}'_u (W_u + \lambda C_{\boldsymbol{\epsilon}}) \mathbf{t}_u \tag{2.22}$$

for some selected parameter λ , subject to the constraint $\mathbf{t}_u \mathbf{k}_{\cdot*} = 1$. The minimization problem has explicit analytic solutions

$$\boldsymbol{t}_{u}(\lambda) = \frac{[W_{u} + \lambda C_{\boldsymbol{\epsilon}}]^{-1} \mathbf{k}_{\cdot*}}{\mathbf{k}_{\cdot*}'[W_{u} + \lambda C_{\boldsymbol{\epsilon}}]^{-1} \mathbf{k}_{\cdot*}}$$
(2.23)

in terms of the parameter λ , and these different solutions, when combined with the data \boldsymbol{v} using Eq.(2.15), give different estimates $\hat{j}^u(\lambda)$. It should be emphasized that the $\boldsymbol{t}_u(\lambda)$ we obtain gives us the estimate at a single voxel. This means that in the simplest version of the Backus-Gilbert method we must perform the inversion for each voxel u. Some properties of the inverse solution via this method has been discussed particularly for the EEG inverse problem (Pascual-Marqui 1999).

Selective Minimum Norm

In the WMN method, a constraint based on L_2 norm is imposed on the current vector \boldsymbol{j} . Instead of employing a L_2 norm constraint, the inverse solution based on a L_1 norm constraint has been proposed by Matsuura and Okabe 1995; They call "Selective Minimum Norm" (SMN) solution. The solution can be obtained by solving the following minimization problem,

minimize
$$\sum_{u=1}^{N_s} |j^u|$$

subject to $K\boldsymbol{j} = \boldsymbol{v}.$ (2.24)

They have optimized this minimization problem by linear programming. As a prominent feature of this method, the solution vector obtained can be sparse. This means that focal current sources can be obtained without any assumption about the number of sources using the distributed source model. Note that the solution obtained from the minimization problem (2.24) is sensitive to the observation noise because of exact satisfaction of the observation. The solution allowing some observation noise has been also developed in Matsuura and Okabe 1996.

Note that in the statistical literature, this sparse property of the solution obtained from the least squares problem with a minimum L_1 norm penalty has been discussed in the context of shrinkage estimators of the regression analysis (Tibshirani 1996) and some efficient algorithms for solving this problem have been developed (Osborne et al. 2000; Efron et al. 2002).

Chapter 3

Weighted minimum L_2 -norm solution

A class of weighted minimum L_2 -norm (denoted by WMN) solution is one of the most popular method because of easiness of obtaining an analytical solution and flexibility of imposing various spatial constraints by adjusting a weight matrix. The detailed issues about WMN solution will be discussed in this chapter.

3.1 SVD based solution

Given a $N_s \times N_s$ weight matrix W (assumed to have an inverse) depending on spatial constraints, a WMN solution is obtained by minimizing a linear mixture of a weighted norm ||Wj|| and the residuals of the fit according to the observation equation. By assuming a Gaussian distribution $\boldsymbol{\epsilon} \sim N(0, \sigma^2 C_{\boldsymbol{\epsilon}})$ with a known covariance structure matrix $C_{\boldsymbol{\epsilon}}$ for the measurement noise, the objective function becomes

$$E(\boldsymbol{j};\boldsymbol{\lambda}) = ||\boldsymbol{v} - \mathbf{K}\boldsymbol{j}||_{C_{\boldsymbol{\epsilon}}^{-1}}^2 + \boldsymbol{\lambda}^2 ||W\boldsymbol{j}||^2.$$
(3.1)

where the parameter λ , called the regularization parameter expresses the balance between fitting the model and the prior constraint of minimizing ||WJ||. Minimization of this objective function for a given λ provides the WMN solution as

$$\hat{\boldsymbol{j}} = (\mathbf{K}' C_{\boldsymbol{\epsilon}}^{-1} \mathbf{K} + \lambda^2 W' W)^{-1} \mathbf{K}' C_{\boldsymbol{\epsilon}}^{-1} \boldsymbol{v}.$$
(3.2)

The computation of Eq.(3.2) is demanding due to the need to inverse $N_s \times N_s$ matrix. An amount of computation can be reduced by further simplification of Eq.(3.2) by means of singular value decomposition (SVD) of $\bar{K} \equiv C_{\epsilon}^{-1/2} K W^{-1}$:

$$\bar{\mathbf{K}} = USV',\tag{3.3}$$

where U and V are a $N_e \times N_e$ matrix and a $N_s \times N_e$ matrix, respectively, and S is a $N_e \times N_e$ diagonal matrix with the singular values s_i . Then, Eq.(3.2) can be arranged as follows,

$$\hat{\boldsymbol{j}} = W^{-1} (\bar{K}'\bar{K} + \lambda^2 I_{N_s})^{-1} \bar{K}' C_{\boldsymbol{\epsilon}}^{-1/2} \boldsymbol{v}.$$

$$= W^{-1} \frac{1}{\lambda^2} \bar{K}' (\frac{1}{\lambda^2} \bar{K}\bar{K}' + I_{N_e})^{-1} C_{\boldsymbol{\epsilon}}^{-1/2} \boldsymbol{v}$$

$$= W^{-1} V S' U' (U(S^2 + \lambda^2 I_{N_e})U')^{-1} C_{\boldsymbol{\epsilon}}^{-1/2} \boldsymbol{v}$$

$$= W^{-1} V \operatorname{diag} \left(\frac{s_i}{s_i^2 + \lambda^2}\right) U' C_{\boldsymbol{\epsilon}}^{-1/2} \boldsymbol{v}.$$
(3.4)

In the arrangement from the first to the second line, the identity,

$$(V^{-1} + H'R^{-1}H)^{-1}H'R^{-1} = VH'(HVH' + R)^{-1}$$

is employed (see p.85 of Kitagawa and Gersh 1996 in detail). This expression can be computed more efficiently than Eq.(3.2).

3.2 Bayesian interpretation

It is well known that penalized least squares problems can be interpreted from the view of Bayesian inference (Kitagawa and Gersh 1996). Dividing the righthand side of Eq.(3.1) by $-2\sigma^2$ and employing exponential yields

$$\exp\left\{-\frac{1}{2\sigma^2}||\boldsymbol{v} - \mathbf{K}\boldsymbol{j}||_{C_{\boldsymbol{\epsilon}}^{-1}}^2\right\} \cdot \exp\left\{-\frac{1}{2}\frac{\lambda^2}{\sigma^2}||W\boldsymbol{j}||^2\right\}$$
(3.5)

This equation corresponds to a posterior distribution $p(\mathbf{j}|\mathbf{v})$ in the case that

likelihood :
$$p(\boldsymbol{v}|\boldsymbol{j}) \sim N(\mathbf{K}\boldsymbol{j}, \sigma^2 C_{\boldsymbol{\epsilon}})$$
 (3.6)

prior distribution :
$$p(\mathbf{j}) \sim N(0, \frac{\sigma^2}{\lambda^2} (W'W)^{-1}).$$
 (3.7)

The WMN solution in Eq.(3.1) corresponds to MAP(Maximum A Posterior) estimates of Bayesian inference.

Eqs.(3.6),(3.7) can be represented by two following equations:

$$\boldsymbol{v} = \mathbf{K}\boldsymbol{j} + \boldsymbol{\epsilon} \quad \boldsymbol{\epsilon} \sim N(0, \sigma^2 C_{\boldsymbol{\epsilon}})$$
 (3.8)

$$W \boldsymbol{j} = \boldsymbol{\eta} \qquad \boldsymbol{\eta} \sim N(0, \frac{\sigma^2}{\lambda^2} \mathbf{I}).$$
 (3.9)

where $\boldsymbol{\epsilon}$ and $\boldsymbol{\eta}$ are uncertainty of observation (i.e observation noise) and prior knowledge, respectively. This form can be considered as an special example of **state space representation**. As will shown in the latter chapter, the state space representation gives a framework to formulate both instantaneous and dynamic inverse problems in an unified form.

3.3 Estimation of the regularization parameter

The regularization parameter λ should be chosen in an objective way, because the WMN solution will depend sensitively on this parameter. Various methods, such as Generalized Cross Validation ("GCV") criterion (Wahba 1990) and Lcurve method (Hansen 1992; Hansen 1994) has been employed for this purpose. In this work we propose employing ABIC (Akaike 1980a; Akaike 1980b) for estimating this parameter, because this criteria can be applied not only for selecting the regularization parameter but also for the model comparison (it is even possible to compare solutions of dynamical and instantaneous inverse problems). In this section L-curve, GCV and ABIC will be briefly reviewed or introduced for WMN solutions.

3.3.1 L-curve

Having realized that the regularization parameter λ controls the balance between the residual norm $||\boldsymbol{v} - \mathbf{K}\boldsymbol{j}||$ and the solution norm $||W\boldsymbol{j}||$, it is quite natural to plot points

$$(\log ||\boldsymbol{v} - \mathbf{K}\boldsymbol{j}_{\lambda}||, \log ||W\boldsymbol{j}_{\lambda}||)$$

for all valid regularization parameters. Here subscript λ is put on \boldsymbol{j} in order to emphasize that the solution \boldsymbol{j} depends on λ . The resulting curve is called 'Lcurve', because it turns out that the L-curve, when plotted in a log-log scale, has a characteristic L-shaped appearance with a distinct corner separating the vertical and the horizontal parts of the curve. The regularization parameter λ is chosen so that the point $(\log ||\boldsymbol{v} - \mathbf{K}\boldsymbol{j}_{\lambda}||, \log ||W\boldsymbol{j}_{\lambda}||)$ is located on the corner point of L-curve or the curvature of L-curve has a maximum value with respect to λ . The discussion of detailed properties of L-curve has appeared in Hansen 2000.

3.3.2 GCV

Cross Validation (CV) is a well-known technique to construct a certain predictive criteria when estimating the hyperparameters. In this technique, firstly a portion of the samples is used for estimating the parameters and then the remaining portion is used for calculating prediction error. The minimization of the prediction error with respect to the hyperparameters gives the estimation of the hyperparameters.

It has been shown that the value of this prediction error can be derived in a closed form for the linear least squares problems and this is generalized to so-called the generalized cross validation (GCV) (Wahba 1990). GCV is defined for the WMN solution as follows:

$$\operatorname{GCV}(\lambda) = \frac{||(\mathbf{I}_{N_e} - B(\lambda))C_{\boldsymbol{\epsilon}}^{-\frac{1}{2}}\boldsymbol{v}||^2}{\{\operatorname{tr}(\mathbf{I}_{N_e} - B(\lambda))\}^2}$$
(3.10)

where

$$B(\lambda) = C_{\epsilon}^{-\frac{1}{2}} \mathbf{K} (\mathbf{K}' C_{\epsilon}^{-1} \mathbf{K} + \lambda^2 W' W)^{-1} \mathbf{K}' C_{\epsilon}^{-\frac{1}{2}}.$$

Numerator and denominator of GCV criteria can be regarded as representing variance of the estimated observation error and bias resulting from the regularization term, respectively. By minimizing GCV a value for λ is obtained which achieves a compromise between these two quantities. In practice, GCV can be evaluated by applying the SVD as mentioned above:

$$\mathbf{I}_{N_e} - B(\lambda) = U(\mathbf{I} - S(S'S + \lambda^2 \mathbf{I}_{N_s})^{-1}S')U'.$$

Consequently, GCV can be expressed:

$$GCV(\lambda) = \frac{\sum_{i=1}^{N_e} \left(\frac{\lambda^2}{\lambda^2 + s_i^2} \tilde{v}_i\right)^2}{\left(\sum_{i=1}^{N_e} \frac{\lambda^2}{\lambda^2 + s_i^2}\right)^2}$$
(3.11)

where \tilde{v}_i is the *i*th component of the vector $U'C_{\epsilon}^{-1/2}\boldsymbol{v}$ and again s_i is *i*th diagonal element of S. This is the expression to be minimized.

3.3.3 ABIC

ABIC is the modification of AIC (Akaike 1973) to the empirical Bayesian framework as defined by

$$ABIC = -2\mathcal{L}^{(II)}(\theta) + 2N,$$

where N is the number of the hyperparameters θ in the model and $\mathcal{L}^{(II)}(\theta)$ is the log-likelihood of the hyperparameters in the context of empirical Bayesian inference, called the type-II log-likelihood or the marginal log-likelihood. In the case that there are unobservable variables in the model, the type-II likelihood is obtained by marginalizing the joint distribution with unobservable variables:

$$\mathcal{L}^{(II)}(\theta) = \log \int p(\boldsymbol{x}, \boldsymbol{y}; \theta) d\boldsymbol{y}$$

= $\log \int p(\boldsymbol{x} | \boldsymbol{y}; \theta_1) p(\boldsymbol{y}; \theta_2) d\boldsymbol{y}.$ (3.12)

where \boldsymbol{x} are the observable and \boldsymbol{y} the unobservable variables; $\boldsymbol{\theta} = (\theta_1, \theta_2)$ are hyperparameters to be estimated.

In the case of WMN solutions, using Bayesian interpretation (3.6), (3.7)and SVD of \bar{K} (3.3), the type-II log-likelihood can be simplified analytically as follows (see Appendix A for detailed calculation),

$$-2\mathcal{L}^{(II)}(\sigma,\lambda) = N_e \log \sigma^2 + \sum_{i=1}^{N_e} \log \frac{s_i^2 + \lambda^2}{\lambda^2} + \frac{1}{\sigma^2} \sum_{i=1}^{N_e} \tilde{v}_i^2 \frac{\lambda^2}{s_i^2 + \lambda^2}$$
(3.13)

where \tilde{v}_i is the *i*th component of the vector $U'C_{\epsilon}^{-1/2}\boldsymbol{v}$. Here a constant term has been ignored in Eq.(3.13). ABIC of the WMN solution is expressed:

ABIC
$$(\sigma, \lambda) = N_e \log \sigma^2 + \sum_{i=1}^{N_e} \log \frac{s_i^2 + \lambda^2}{\lambda^2} + \frac{1}{\sigma^2} \sum_{i=1}^{N_e} \tilde{v}_i^2 \frac{\lambda^2}{s_i^2 + \lambda^2} + 2N.$$
 (3.14)

The regularization parameter λ can be obtained by minimizing ABIC (or equivalently -2 times type-II log-likelihood). Differentiating the right side of Eq.(3.14) by σ^2 , the estimate of σ^2 can be obtained as

$$\hat{\sigma}^2 = \frac{1}{N_e} \sum_{i=1}^{N_e} \frac{\lambda^2}{s_i^2 + \lambda^2} \tilde{v}_i^2.$$
(3.15)

Substituting (3.15) into (3.14), the regularization parameter λ can be estimated by numerical minimization of the function

$$\mathcal{M}'(\lambda) = N_e \log \hat{\sigma}^2 + N_e + \sum_{i=1}^{N_e} \log \frac{s_i^2 + \lambda^2}{\lambda^2} + 2N.$$
 (3.16)

3.3.4 Remarks

Here are a few remarks on estimation of the regularization parameter λ obtained from ABIC and GCV from the author's little experience. GCV often has a very sharp curve with respect to λ around the minimizer whereas ABIC shows rather flat curve. On the other hands, GCV sometimes has no minimizer whereas ABIC almost always has the minimizer. Hence GCV is said to be more sensitive to λ than ABIC but less robust than ABIC. When both minimizer exist, the agreement between these two estimates are often observed. The author recommends to use ABIC because of its robustness and its availability for model comparison.

3.4 Posterior variance

Interval estimates for WMN solutions can be easily obtained from the posterior distribution (Eq.(A.3) in Appendix A) as

$$p(\boldsymbol{j}|\boldsymbol{v}) \propto \exp\left\{-\frac{1}{2\sigma^2}E(\hat{\boldsymbol{j}};\lambda)\right\} \cdot \exp\left\{-\frac{1}{2\sigma^2}(\boldsymbol{j}-\hat{\boldsymbol{j}})'(\mathbf{K}'C_{\boldsymbol{\epsilon}}^{-1}\mathbf{K}+\lambda^2W'W)(\boldsymbol{j}-\hat{\boldsymbol{j}})\right\}$$

Therefore, the posterior variance can be obtained as

$$\operatorname{Var}(\boldsymbol{j}|\boldsymbol{v}) = \sigma^{2} (\mathrm{K}' C_{\boldsymbol{\epsilon}}^{-1} \mathrm{K} + \lambda^{2} W' W)^{-1} \\ = \frac{\sigma^{2}}{\lambda^{2}} \Big\{ (W'W)^{-1} - W^{-1} V \operatorname{diag}\left(\frac{s_{i}^{2}}{s_{i}^{2} + \lambda^{2}}\right) V' W^{-1} \Big\}. \quad (3.17)$$

The inverse formula (p.85 of Kitagawa and Gersh 1996)

$$(V^{-1} + H'R^{-1}H)^{-1} = V - VH'(HVH' + R)^{-1}HV$$

and Eq.(3.3) are employed. In Eq.(3.17), the first term can be interpreted as prior variance (i.e. variance before any observation) and the second term can be interpreted as the amount of variance reduction due to the observation. Unfortunately the computation of this equation is impracticable for the size of our application ($N_s \ge 10000$), because the first term of Eq.(3.17) could be a $N_s \times N_s$ dense matrix which requires (temporal) huge memory, for example, 800Mb in our typical application. Instead, the posterior variance of Wj and the *i*th diagonal elements of that can be obtained as follows:

$$\operatorname{Var}(W\boldsymbol{j}|\boldsymbol{v}) = \frac{\sigma^2}{\lambda^2} \Big\{ \mathrm{I}_{N_s} - V \operatorname{diag}\left(\frac{s_i^2}{s_i^2 + \lambda^2}\right) V' \Big\},\tag{3.18}$$

and

$$\operatorname{Var}(wj_{i}|\boldsymbol{v}) = \frac{\sigma^{2}}{\lambda^{2}} \Big(1 - \sum_{j=1}^{N_{e}} V_{ij}^{2} \frac{s_{j}^{2}}{s_{j}^{2} + \lambda^{2}} \Big),$$
(3.19)

where wj_i and V_{ij} are the elements of Wj and V, respectively. Now these elements are feasible to keep in temporal memory.

In the literature (Pascual-Marqui 2002), Pascual has introduced another way to calculate the variance of the estimate \hat{j} . His derivation is summarized as follows. $S_{j}, S_{v}, S_{\epsilon}$ and $S_{\hat{j}}$ denote covariance matrices of j, v, ϵ and \hat{j} , respectively. From $v = Kj + \epsilon$,

$$S_{\boldsymbol{v}} = \mathbf{K} S_{\boldsymbol{j}} \mathbf{K}' + S_{\boldsymbol{\epsilon}} \tag{3.20}$$

is obtained. Since the estimate \hat{j} is given by

$$\hat{\boldsymbol{j}} = (\mathbf{K}' S_{\boldsymbol{\epsilon}}^{-1} \mathbf{K} + S_{\boldsymbol{j}}^{-1})^{-1} \mathbf{K}' S_{\boldsymbol{\epsilon}}^{-1} \boldsymbol{v}$$
$$= S_{\boldsymbol{j}} \mathbf{K}' S_{\boldsymbol{v}}^{-1} \boldsymbol{v}$$
(3.21)

taking the variance of both-hand side of Eq.(3.21) leads to

$$S_{\mathbf{j}} = S_{\mathbf{j}} \mathbf{K}' S_{\mathbf{v}}^{-1} \cdot S_{\mathbf{v}} \cdot S_{\mathbf{v}}^{-1} \mathbf{K} S_{\mathbf{j}}$$
$$= S_{\mathbf{j}} \mathbf{K}' S_{\mathbf{v}}^{-1} \mathbf{K} S_{\mathbf{j}}. \qquad (3.22)$$

This final expression (3.22) is employed as the variance of the estimate $\hat{\boldsymbol{j}}$ by Pascual. On the other hand, substituting $S_{\boldsymbol{j}} = \frac{\sigma^2}{\lambda^2} (W'W)^{-1}$ and $S_{\boldsymbol{\epsilon}} = \sigma^2 C_{\boldsymbol{\epsilon}}$ into Eq.(3.17), the posterior variance can be derived as

$$\begin{aligned} \operatorname{Var}(\boldsymbol{j}|\boldsymbol{v}) &= (\operatorname{K} S_{\boldsymbol{\epsilon}}^{-1} \operatorname{K}' + S_{\boldsymbol{j}})^{-1} \\ &= S_{\boldsymbol{j}} - S_{\boldsymbol{j}} \operatorname{K}' S_{\boldsymbol{v}}^{-1} \operatorname{K} S_{\boldsymbol{j}}. \end{aligned}$$
 (3.23)

Here from the second to the third equation the above-mentioned inverse formula is used. Note that this equation can be also derived using the formula of the multivariate normal distribution (see appendix B):

$$\operatorname{Var}(\boldsymbol{j}|\boldsymbol{v}) = \operatorname{Var}(\boldsymbol{j}) - \operatorname{Cov}(\boldsymbol{j}, \boldsymbol{v}) \operatorname{Var}(\boldsymbol{v})^{-1} \operatorname{Cov}(\boldsymbol{v}, \boldsymbol{j}).$$

Comparing Eq.(3.23) with Eq.(3.22), the Pascual's variance is consistent with the second term of the posterior variance. Hence it is not so obvious whether the variance as proposed by Pascual can be justified or not.
Chapter 4

Dynamical Inverse Problem

In the previous chapters, the notion of forward and inverse problems are introduced and various methods to solve inverse problems are reviewed. In this chapter, firstly **dynamical inverse problems** is defined in contrast to **instantaneous inverse problems**, then the state space representation is introduced as a mathematically suitable formulation of dynamical inverse problems.

4.1 Definition

As mentioned in the previous chapters, the EEG inverse problem is defined as the task to estimate the primary current density j from given measurement v. In particular, the inverse problem as formulated in Eq.(2.9) is called 'instantaneous inverse problems' or 'static inverse problems', because only the measurement at one single time point is used or no temporal relation is taken into consideration for the estimation of j.

An explicit discussion about the definition of 'dynamic(al) inverse problems' was first appeared in Schmidt and Louis 2002a. According to their definition,

• the properties **j** of the examined object do not change during the measuring process. Thus, we have to solve

$$\mathbf{K} \boldsymbol{j} = \boldsymbol{v}_i \quad \text{for all } i$$

This is called a static inverse problem.

• the examined object is allowed to change during the measuring process and we have to solve

$$\mathbf{K} \boldsymbol{j}_i = \boldsymbol{v}_i \quad \text{for all } i$$

This is called a dynamic inverse problem.

Since they employed 'dynamic inverse problem' for broadly describing any time-varying situation, but without reference to the particular case of variations in time which are the result of an actual dynamical evolution, the term 'nonstationary inverse problems' seems to be more suitable for their definition.

In this thesis, the term 'dynamical inverse problems' is used for a little more restricted situation as follows:

- Solutions of dynamical inverse problems have to be in agreement with two sets of information, which is represented by
 - the observation equations for all time points considered $\boldsymbol{v}_t = \mathbf{K} \boldsymbol{j}_t + \boldsymbol{\epsilon}_t \ (t = 1, 2, \cdots), \text{ and }$
 - some prespecified dynamics about $\boldsymbol{j}_t, \boldsymbol{j}_{t-1}, \boldsymbol{j}_{t-2}, \cdots$.
- Solutions of 'instantaneous inverse problems' have to be in agreement with two sets of information, namely
 - the observation equation $\boldsymbol{v} = \mathbf{K}\boldsymbol{j} + \boldsymbol{\epsilon}$, and
 - prior knowledge about \boldsymbol{j} .

In our definition, an explicit relationship for time course of \mathbf{j}_t is assumed and this assumption enable us to make mathematical formulation of the problem much easier as will mention in the next section.

In the case of the instantaneous inverse problem, the solution only reflects an instantaneous observation \boldsymbol{v}_t , whereas in the case of the dynamical inverse problem it reflects a sequence of temporally successive observations $\boldsymbol{v}_t, \boldsymbol{v}_{t-1}, \cdots$ such that some dynamics of the generators is imposed. In other word, as shown schematically in Fig.4.1, the dependence of the observation \boldsymbol{v}_t on the evolution of \boldsymbol{j}_t is explicitly considered in the dynamical inverse problem. If the evolution of \boldsymbol{j}_t does not follow any dynamics, the dynamical inverse problem becomes equivalent to the instantaneous problem, i.e. the dynamical problem is a generalization of the instantaneous problem.

4.2 Formulation

The dynamical inverse problem in the penalized least squares form and in the state space representation is discussed here. In order to consider general spatio-temporal constraints, we will choose to formulate the dynamical inverse problem in the state space representation; the dynamical inverse problem will be regarded as the **state estimation problem**.



Figure 4.1: Schematic comparison between the instantaneous inverse problem (top) and the dynamical inverse problem (bottom). Sources within brain and EEG observations are represented by rectangles and circles, respectively. Arrows represent the flow of information, as assumed by the underlying model.

4.2.1 Penalized least squares form

The objective function (3.1) of WMN method represents an example of the penalized least square (PLS) method, where the lack of spatial smoothness is penalized. It is a very convenient feature of the PLS method that we can obtain a solution with desired properties simply by adding suitable penalty terms. We can employ this virtue in order to extend a type of weighted minimum norm solutions such that the resulting solutions will also have the property of improved temporal smoothness.

A pioneering work of extending a instantaneous inverse problem employing spatial and temporal constraints has been introduced by Schmidt et al. 2001 for EEG inverse problems. They first considered a objective function as follows:

$$E(\boldsymbol{j}_{1},\cdots,\boldsymbol{j}_{T}) = \sum_{i=1}^{T} (||\boldsymbol{v}_{t} - \mathbf{K}\boldsymbol{j}_{t}||^{2} + \lambda_{1}^{2}||W\boldsymbol{j}_{t}||^{2}) + \sum_{t=2}^{T} \lambda_{2}^{2}||\boldsymbol{j}_{t} - \boldsymbol{j}_{t-1}||^{2} \quad (4.1)$$

where a weight matrix W is an identity matrix or Laplacian matrix according to corresponding a spatial constraint. The third term represents a penalty for temporal smoothness. By minimizing this objective function, they could obtain the solution having properties of temporal and spatial smoothness (in the case of W = L). In practical application, after transforming Eq.(4.1) to the next objective function,

$$E(\boldsymbol{J}) = ||\boldsymbol{V} - \mathbf{K}_T \boldsymbol{J}||^2 + \lambda_1^2 ||W_T \boldsymbol{J}||^2 + \lambda_2^2 ||(D \otimes \mathbf{I}_{N_s}) \boldsymbol{J}||^2$$
(4.2)

they introduced the normal equation for a large-scale least squares problem,

$$(\mathbf{K}'_T\mathbf{K}_T + \lambda_1^2 W'_T W_T + \lambda_2^2 (D'D \otimes \mathbf{I}_{N_s})) \boldsymbol{J} = \mathbf{K}'_T \boldsymbol{V}$$
(4.3)

where

$$J = (\boldsymbol{j}_1, \cdots, \boldsymbol{j}_T)'$$
$$V = (\boldsymbol{v}_1, \cdots, \boldsymbol{v}_T)'$$
$$K_T = I_T \otimes K, \quad W_T = I_T \otimes W$$
$$D = \begin{bmatrix} 1 & -1 & 0 & \cdots & \\ 0 & 1 & -1 & \ddots & \\ \vdots & \ddots & \ddots & \ddots & \\ \vdots & & 1 & -1 \end{bmatrix}$$

Note that the size of the matrices K_T , W_T , $D'D \otimes I_{N_s}$ is usually huge in the EEG application, for instance, K_T is a matrix of size $(T \cdot N_e) \times (T \cdot N_s)$. In order to solve the normal equation (4.3) without keeping the whole components of the matrices in memory, iterative least squares methods such as the block Jacobi method, the conjugate gradient algorithm (Schmidt et al. 2001) and the STR (spatio temporal regularizer) procedure (Schmidt and Louis 2002a; Schmidt and Louis 2002b) have been proposed in virtue of sparsity of the matrices.

By generalizing the objective function (4.1), we have a following form of the objective function with temporal and spatial constraints:

$$E(\boldsymbol{J}) = u(\boldsymbol{j}_1) + \sum_{t=2}^{T} (||\boldsymbol{v}_t - \mathbf{K}\boldsymbol{j}_t|| + \lambda_1^2 ||W_1 \boldsymbol{j}_t||^2 + \lambda_2^2 ||\boldsymbol{j}_t - f(\boldsymbol{j}_{t-1})||^2). \quad (4.4)$$

Here W_1 is a weight matrix depending on a spatial property of a solution, $f(\cdot)$ are a function representing a dynamics of current sources \mathbf{j}_t . A function $u(\mathbf{j}_1)$ is a cost function for estimating an initial source \mathbf{j}_1 , for example $u(\mathbf{j}_1) =$ $||\mathbf{v}_1 - \mathbf{K}\mathbf{j}_1|| + \lambda_1^2 ||W_1\mathbf{j}_1||^2$ as in the objective function (4.1). The choice of W_1 and $f(\cdot)$ depends on the prior knowledge or assumptions what kind of properties time series of sources \mathbf{j}_t is expected to have. For example, the above-mentioned "model" used in Schmidt et.al. 2001 corresponds to the choice $W_1 = \mathbf{L}$ and $f(\cdot) = \mathbf{I}$, therefore the solution will have spatial and temporal smoothness.

Eq.(4.4) shows a way to impose both spatial and temporal constraints on a solution. Another way of imposing two kind of constraints can be derived by introducing a parameter μ which make a balance of two constraints inside one penalty term as follows,

$$E(\boldsymbol{J}) = u(\boldsymbol{j}_{1}) + \sum_{t=2}^{T} (||\boldsymbol{v}_{t} - \mathbf{K}\boldsymbol{j}_{t}|| + \lambda_{1}^{2} ||W_{1}\boldsymbol{j}_{t} + \mu^{2}(\boldsymbol{j}_{t} - f(\boldsymbol{j}_{t-1}))||^{2})$$

$$= u(\boldsymbol{j}_{1}) + \sum_{t=2}^{T} (||\boldsymbol{v}_{t} - \mathbf{K}\boldsymbol{j}_{t}|| + \lambda_{1}^{2} ||W_{3}(\boldsymbol{j}_{t} - g(\boldsymbol{j}_{t-1}))||^{2})$$
(4.5)

where $W_3 = W_1 + \mu^2 I$ and $g(\cdot) = \mu^2 W_3^{-1} f(\cdot)$. The main idea of the objective function (4.5) is to impose spatio-temporal constraints in a penalty term at once, while to impose spatial and temporal constraints separately due to two penalty terms is an idea of the objective function (4.4).

Two objective functions for estimating time series of sources \mathbf{j}_t with general spatial temporal constraints have been introduced in this section. Minimization of Eq.(4.4) or (4.5) gives a corresponding solution. In the case of a nonlinear function $f(\cdot)$ (or $g(\cdot)$), however, the transformation to the large-scale least squares problem is impossible. Even in the case of $f(\cdot)$ being a linear function, we can not execute the iterative methods for solving the large-scale least squares problem after the transformation unless $f(\cdot)$ is a diagonal matrix whose diagonal components have the same values (i.e. dynamics with homogeneity and no interaction between voxels).

4.2.2 State space representation

In order to solve dynamical inverse problems, more appropriate formulation is needed because the minimization problem in the penalized least squares form can be solved only in a few case of temporal constraints. In other words, the penalized least squares form is too limited to obtain the solution for general spatial temporal constraints.

In section 3.2, by viewing the objective function of the WMN method from Bayesian statistics, the representation using random variables (state space representation) has been introduced. The objective functions Eqs.(4.4) and (4.5) can also be expressed by state space representation in the same way.

• The least squares form with three terms,

$$||\boldsymbol{v}_{t} - \mathbf{K}\boldsymbol{j}_{t}|| + \lambda_{1}^{2}||W_{1}\boldsymbol{j}_{t}||^{2} + \lambda_{2}^{2}||\boldsymbol{j}_{t} - f(\boldsymbol{j}_{t-1})||^{2}$$
(4.6)

can be expressed by following state space representation with an 'augmented' observation equation:

$$\begin{cases} \begin{bmatrix} \boldsymbol{v}_t \\ \boldsymbol{0} \end{bmatrix} = \begin{bmatrix} \mathbf{K} \\ \lambda_1 W_1 \end{bmatrix} \boldsymbol{j}_t + \tilde{\boldsymbol{\epsilon}}_t \quad \tilde{\boldsymbol{\epsilon}}_t \sim N(0, \sigma^2 \mathbf{I}_{N_s + N_e}) \\ \boldsymbol{j}_t = f(\boldsymbol{j}_{t-1}) + \boldsymbol{\eta}_t \qquad \boldsymbol{\eta}_t \sim N(0, \tau^2 \mathbf{I}_{N_s}) \end{cases}$$
(4.7)

Here a part of the augmented observation equation corresponds to the spatial constraint (the second term) in Eq.(4.6) and the system equation corresponds to the temporal constraint (the third term). Note that the covariance structure of the system noise equation in the representation (4.7) is always an identity matrix, due to separate imposition of a spatial and temporal constraint. This formulation has already been employed in other fields of inverse problems (Karjalainen et al. 1997; Vauhkonen et al. 1998).

• The least squares form with two terms,

$$||\boldsymbol{v}_t - \mathbf{K}\boldsymbol{j}_t|| + \lambda_1^2 ||W_3(\boldsymbol{j}_t - g(\boldsymbol{j}_{t-1}))||^2$$
(4.8)

can be expressed by following ordinary state space representation:

$$\begin{cases} \boldsymbol{v}_t = \mathbf{K}\boldsymbol{j}_t + \boldsymbol{\epsilon}_t & \boldsymbol{\epsilon}_t \sim N(0, \sigma^2 \mathbf{I}_{N_e}) \\ \boldsymbol{j}_t = g(\boldsymbol{j}_{t-1}) + \boldsymbol{\eta}_t & \boldsymbol{\eta}_t \sim N(0, \tau^2 (W'_3 W_3)^{-1}) \end{cases}$$
(4.9)

Here the covariance structure of system noise represents the spatial constraints in Eq.(4.8) and the system equation corresponds to the temporal constraint.

Once the dynamical inverse problem is formulated in these state space representations, to find out the solution of the dynamical inverse problem is equivalent to estimating the state j_t (i.e. the state estimation problem). In linear Gaussian case, the optimal state estimation is possible by the algorithm proposed by Kalman, so-called Kalman filtering or smoothing. In general, some extension of Kalman filtering (extended Kalman filter or particle filter) can be applied in order to estimate the state. Since the Kalman filter algorithm consists of a set of recursions along time axis, we can obtain the solution even when more flexible dynamics such as non-linear and time-varying dynamics is taken into consideration. In contrast, the penalized least squares form has the crucial limitation for the dynamics in transforming into the large-scale least squares problem.

The following points are the advantage of regarding the dynamical inverse problem as the state estimation problem:

- A solution can be obtained for wide range of dynamical models.
- The system and the observation equation in the state space representation can be interpreted intuitively, therefore the modeling, that is, imposing spatio-temporal constraints on the solution, can be done with flexibility and ease.
- In the recursions of Kalman filtering, likelihood can easily calculated. The evaluation of solutions, hence, is possible by likelihood.

Corresponding to two least squares form using multiple penalty terms and a spatio-temporal penalty term, two types of the state space representation (4.7),(4.9) have been also introduced. The advantage of the representation (4.7) over the representation (4.9) is that

• In modeling W_1 and $f(\cdot)$ of (4.7), the model can be easily interpreted from view of a spatial constraint and a temporal constraint onto \mathbf{j}_t , respectively, while modeling W_3 and $g(\cdot)$ of (4.9) is not corresponding to a spatial constraint and a spatial constraint onto \mathbf{j}_t .

On the other hand, the disadvantage of the representation (4.7) would be

• the cost of computation is too demanding because of high dimensionality of the observation equation.

In applications with low dimensional state, the representation (4.7) may be preferred, however, in the applications with high dimensional state, the representation (4.9) will be more appropriate because the priority should be put on the computational aspect rather than ease of the interpretation of the model. In the latter parts of this thesis, the dynamical inverse problem will be expressed in the representation (4.9).

Chapter 5

Dynamical Inverse Solution

As shown in the chapter 4, the dynamical inverse problem has been formulated as the state estimation problem in the state space representation,

$$\begin{cases} \boldsymbol{v}_t = \mathbf{K}\boldsymbol{j}_t + \boldsymbol{\epsilon}_t & \boldsymbol{\epsilon}_t \sim N(0, \sigma^2 C_{\boldsymbol{\epsilon}}) \\ \boldsymbol{j}_t = g(\boldsymbol{j}_{t-1}) + \boldsymbol{\eta}_t & \boldsymbol{\eta}_t \sim N(0, \tau^2 C_{\boldsymbol{\eta}}) \end{cases}$$
(5.1)

The procedure of solving the state estimation problem consists of three parts;

- (i) define the form of the dynamics $g(\cdot)$ and the noise covariance structure C_{ϵ}, C_{η} (Modeling).
- (ii) given the model, estimate the current source j_t (State Estimation).
- (iii) evaluate solutions resulting from different models by the statistical criterion (**Model Comparison**).

These topics are discussed in this chapter, starting from introduction of Kalman filtering algorithm for "State Estimation". Various approximate methods for high dimensional "State Estimation" will be proposed, followed by "Model-ing", "Model Comparison" and discussion.

Through 5.1 to 5.4, the transition matrix A_t and the noise covariance matrices C_{ϵ}, C_{η} are assumed to be known; the model is represented as follows,

$$\boldsymbol{v}_t = \mathbf{K} \boldsymbol{j}_t + \boldsymbol{\epsilon}_t \ \boldsymbol{\epsilon}_t \sim N(0, \sigma^2 C_{\boldsymbol{\epsilon}})$$
 (5.2)

$$\boldsymbol{j}_t = A_t \boldsymbol{j}_{t-1} + \boldsymbol{\eta}_t \quad \boldsymbol{\eta}_t \sim N(0, \tau^2 C_{\boldsymbol{\eta}}).$$
(5.3)

5.1 Kalman filtering algorithm

The well-known Kalman filtering algorithm consists of the following prediction and filtering steps:

[Prediction]

$$\mathbf{j}_{t|t-1} = A_t \mathbf{j}_{t-1|t-1}$$
 (5.4)

$$\Sigma_{t|t-1} = A_t \Sigma_{t-1|t-1} A'_t + \tau^2 C_{\eta}$$
(5.5)

[Filter]

$$\boldsymbol{r}_t = \boldsymbol{v}_t - \mathbf{K} \boldsymbol{j}_{t|t-1} \tag{5.6}$$

$$\Lambda_t = \mathbf{K} \Sigma_{t|t-1} \mathbf{K}' + \sigma^2 C_{\boldsymbol{\epsilon}}$$
(5.7)

$$\mathcal{K}_t = \Sigma_{t|t-1} \mathbf{K}' \Lambda_t^{-1} \tag{5.8}$$

$$\boldsymbol{j}_{t|t} = \boldsymbol{j}_{t|t-1} + \mathcal{K}_t \boldsymbol{r}_t \tag{5.9}$$

$$\Sigma_{t|t} = (\mathbf{I} - \mathcal{K}_t \mathbf{K}) \Sigma_{t|t-1}$$
(5.10)

where $\mathbf{j}_{t|s} \equiv \mathrm{E}(\mathbf{j}_t|\mathbf{v}_1,\cdots,\mathbf{v}_s)$ and $\Sigma_{t|s} \equiv \mathrm{Var}(\mathbf{j}_t|\mathbf{v}_1,\cdots,\mathbf{v}_s)$ are expectation and variance of \mathbf{j}_t conditional on the observations $\mathbf{v}_1,\cdots,\mathbf{v}_s$. The **innovation** (one step-ahead prediction error) and the innovation variance are denoted by \mathbf{r}_t and Λ_t , respectively. By iterating the prediction step and the filtering step for each time point $t = 1, \cdots, T$, we could obtain the filtered estimate $\mathbf{j}_{t|t}$ $(t = 1, \cdots, T)$, which is known to be the optimal estimate of \mathbf{j}_t based on the observations up to current under the assumption of Gaussian noise. Following smoothing procedure provides the smoother estimate of \mathbf{j}_t , which is based on the whole data set $\mathbf{v}_1, \cdots, \mathbf{v}_T$:

[Fixed interval smoothing]

$$B_t = \Sigma_{t|t} A_t' \Sigma_{t+1|t}^{-1}$$
(5.11)

$$\mathbf{j}_{t|T} = \mathbf{j}_{t|t} + B_t(\mathbf{j}_{t+1|T} - \mathbf{j}_{t+1|t})$$
 (5.12)

$$\Sigma_{t|T} = \Sigma_{t|t} + B_t (\Sigma_{t+1|T} - \Sigma_{t+1|t}) B'_t.$$
(5.13)

The derivation of Kalman filtering and smoothing can be seen in Meinhold and Singpurwalla 1983; Kitagawa and Gersh 1996 from the view of Bayes statistics or in Kalman 1960; Ansley and Kohn 1982 using the orthogonal projection. Note that there are some variants of the fixed interval smoothing algorithm, and the algorithm as has been introduced here is referred to as 'classical' fixed interval smoother in some literatures (de Jong 1989; Durbin and Koopman 2001).

In the case of the EEG inverse problem, the dimension of the state \mathbf{j}_t is 3 times the number of voxels (in our example, around 10000), therefore the practical application of Kalman filtering and smoothing to such a high dimensional state vector will be very demanding (or even impossible) in terms of computational time and memory consumption. The main difficulty results

from the computation of each covariance matrix $\Sigma_{t|t}$ of size $N_s \times N_s$, which requires 800M byte memory for the matrix in our typical example. In order to overcome this difficulty it is necessary to design suitable approximations of the standard Kalman filtering algorithm.

In the following sections, three approaches to manage this problem will be discussed. Firstly a naive approach where only little modification of the WMN method is needed will be introduced. We choose to call this method **'recursive penalized least squares'** (RPLS) solution (Yamashita et al. 2002). Secondly the Kalman filtering algorithm using projection of the covariance matrices on an appropriate subspace will be proposed. Especially the algorithm where the subspace can be defined by the SVD of the observation matrix K will be derived (Yamashita and Ozaki 2003). A new approach to spatiotemporal Kalman filtering will be also presented which mainly exploits the information of the diagonal part of the covariance matrices (Galka et al. 2002). All of these approximations are based on the idea how to exploit important information in the covariance matrix without keeping the whole elements.

5.2 Recursive penalized least squares method

5.2.1 Algorithm

The algorithm of the RPLS method can be easily obtained by extending the WMN method. We will now discuss the practical estimation procedure in detail. An initial estimate (for t = 1) of the state \mathbf{j}_1 can be obtained from approach for solving any instantaneous inverse problem. For $t = 2, 3, \dots, T$, we can obtain an estimate of \mathbf{j}_t by recursively solving the penalized least squares problem

$$\hat{\boldsymbol{j}}_{t} = \underset{\boldsymbol{j}_{t}}{\operatorname{argmin}} \left\{ ||\boldsymbol{v}_{t} - \mathbf{K}\boldsymbol{j}_{t}||_{C_{\boldsymbol{\epsilon}}^{-1}}^{2} + \lambda^{2} ||\boldsymbol{j}_{t} - A_{t}\hat{\boldsymbol{j}}_{t-1}||_{C_{\boldsymbol{\eta}}^{-1}}^{2} \right\}$$
(5.14)

where \hat{j}_{t-1} is the estimate obtained in the previous step. The solution of (5.14) is easily obtained by

$$\hat{\boldsymbol{j}}_{t} = (\mathbf{K}' C_{\boldsymbol{\epsilon}}^{-1} \mathbf{K} + \lambda^{2} C_{\boldsymbol{\eta}}^{-1})^{-1} (\mathbf{K}' C_{\boldsymbol{\epsilon}}^{-1} \boldsymbol{v}_{t} + \lambda^{2} C_{\boldsymbol{\eta}}^{-1} A_{t} \hat{\boldsymbol{j}}_{t-1}).$$
(5.15)

However, direct computation of this expression is numerically impracticable due to the need of inverting a very big matrix. Instead, we will now show a convenient way to obtain a numerically easier accessible solution by appropriate variable transformation. In addition, this transformation will demonstrate clearly the relations between the RPLS method and Kalman filtering.

We start from the following variable transformations

$$\boldsymbol{\zeta}_t = \boldsymbol{j}_t - A_t \hat{\boldsymbol{j}}_{t-1} \tag{5.16}$$

$$\boldsymbol{r}_t = \boldsymbol{v}_t - \mathbf{K} A_t \hat{\boldsymbol{j}}_{t-1} \tag{5.17}$$

Here $\boldsymbol{\zeta}_t$ and \boldsymbol{r}_t correspond to system noise and innovation (1-step ahead prediction error), respectively. We can rewrite the objective function of Eq.(5.14) as follows:

$$E(\boldsymbol{\zeta}_t) = ||\boldsymbol{r}_t - \mathbf{K}\boldsymbol{\zeta}_t||_{C_{\boldsymbol{\epsilon}}^{-1}}^2 + \lambda^2 ||\boldsymbol{\zeta}_t||_{C_{\boldsymbol{\eta}}^{-1}}^2.$$
(5.18)

Then we can obtain an estimate of \boldsymbol{j}_t by

$$\hat{\boldsymbol{\zeta}}_t = T(\lambda) \boldsymbol{r}_t \tag{5.19}$$

$$\hat{\boldsymbol{j}}_t = A_t \hat{\boldsymbol{j}}_{t-1} + \hat{\boldsymbol{\zeta}}_t, \qquad (5.20)$$

where we have defined and arranged

$$T(\lambda) \equiv (\mathbf{K}' C_{\boldsymbol{\epsilon}}^{-1} \mathbf{K} + \lambda^2 C_{\boldsymbol{\eta}}^{-1})^{-1} \mathbf{K}' C_{\boldsymbol{\epsilon}}^{-1}$$
(5.21)

$$= C_{\eta}^{1/2} V \operatorname{diag}\left(\frac{s_i}{s_i^2 + \lambda^2}\right) U' C_{\epsilon}^{-1/2}.$$
 (5.22)

Here U, diag (s_i) , V are $N_e \times N_e$, $N_e \times N_e$ and $N_s \times N_e$ matrices obtained from the singular value decomposition (SVD) of $C_{\epsilon}^{-1/2} K C_{\eta}^{-1/2}$, respectively (Mardia et al. 1979). The computation of Eq.(5.22) is not as demanding as the computation of Eq.(5.15), because in Eq.(5.22) the large matrix to be inverted does not depend on λ , such that this inversion needs only to be carried out once, whereas in Eq.(5.15) the inversion has to be carried out repeatedly during the process of finding an optimal value of λ . A similar remark applies to the SVD of $C_{\epsilon}^{-1/2} K C_{\eta}^{-1/2}$, which also needs to be computed only once, since these three matrices are known.

In summary, the RPLS algorithm consists of the following two steps: Given \hat{j}_1 , for $t = 2, \dots, T$

(i) calculate the innovation \boldsymbol{r}_t :

$$\boldsymbol{r}_t = \boldsymbol{v}_t - \mathrm{K}A_t \hat{\boldsymbol{j}}_{t-1}$$

(ii) update the next estimate \hat{j}_t :

$$\hat{\boldsymbol{j}}_t = A_t \boldsymbol{j}_{t-1} + T(\lambda) \boldsymbol{r}_t$$

where $T(\lambda) = C_{\eta}^{1/2} V \operatorname{diag}\left(\frac{s_i}{s_i^2 + \lambda^2}\right) U' C_{\epsilon}^{-1/2}$.

5.2.2 Estimation of the regularization parameter λ

The regularization parameter λ should be chosen in an objective way, because the inverse solution will depend sensitively on this parameter. As shown in the previous section, various methods, such as GCV, L-curve and ABIC can provide the objective choice of λ in the case of the WMN method.

The ABIC criterion is used for the RPLS solution, though in the following derivation we will employ an approximation which might not be fully justified from the theoretical view.

The exact form of type II log-likelihood for the model (5.3) is given by

$$\mathcal{L}^{(II)}(\sigma,\tau) = \log \int p(\boldsymbol{v}_1,\cdots,\boldsymbol{v}_T,\boldsymbol{j}_1,\cdots,\boldsymbol{j}_T;\sigma,\tau) d\boldsymbol{j}_1\cdots d\boldsymbol{j}_T \qquad (5.23)$$

where \boldsymbol{v}_t are the observable and \boldsymbol{j}_t the unobservable variables; σ, τ are hyperparameters and the ratio σ/τ corresponds to λ .

Even if both distributions of the observation and system noise are assumed to be Gaussian, it is very difficult to calculate this multiple integral analytically, therefore we will approximate Eq.(5.23) by the sum of type-II log-likelihoods at each time point, $\mathcal{L}_t^{(II)}(\sigma, \tau)$. Since in the RPLS method we are basing the inference (if interpreted from the Bayesian viewpoint) on $p(\mathbf{r}_t | \boldsymbol{\zeta}_t; \sigma)$ and $p(\boldsymbol{\zeta}_t; \tau)$ as likelihood and prior distribution, respectively (see (5.18)), the pointwise type-II log-likelihood based on \mathbf{r}_t is given by

$$\mathcal{L}_{t}^{(II)}(\sigma,\tau) = \log \int p(\boldsymbol{r}_{t}|\boldsymbol{\zeta}_{t};\sigma)p(\boldsymbol{\zeta}_{t};\tau)d\boldsymbol{\zeta}_{t}.$$
(5.24)

Then $\mathcal{L}^{(II)}$ is approximated by the summation of $\mathcal{L}_t^{(II)}(\sigma, \tau)$. This approximation is reasonable, if the distribution of the innovations $p(\mathbf{r}_t), t = 1, 2, \cdots, T$ can be regard as serially independent.

Since $p(\mathbf{r}_t | \boldsymbol{\zeta}_t; \sigma)$ and $p(\boldsymbol{\zeta}_t; \tau)$ are assumed to be Gaussian, we can analytically calculate this integral and obtain (-2) times the point wise type-II log-likelihood:

$$-2\mathcal{L}_{t}^{(II)}(\sigma,\lambda) = N_{e}\log\sigma^{2} + \sum_{i=1}^{N_{e}}\log\frac{s_{i}^{2}+\lambda^{2}}{\lambda^{2}} + \frac{1}{\sigma^{2}}\sum_{i=1}^{N_{e}}\tilde{r}_{i,t}^{2}\frac{\lambda^{2}}{s_{i}^{2}+\lambda^{2}}$$
(5.25)

where $\tilde{r}_{i,t}$ is the *i*th component of the vector $U'C_{\epsilon}^{-1/2}\boldsymbol{r}_t$. Here we have replaced the parameter τ by $\lambda = \sigma/\tau$. A constant term has been ignored in Eq.(5.25). Then ABIC can be expressed by

$$ABIC(\sigma, \lambda) = TN_e \log \sigma^2 + T \sum_{i=1}^{N_e} \log \frac{s_i^2 + \lambda^2}{\lambda^2} + \frac{1}{\sigma^2} \sum_{t=1}^T \sum_{i=1}^{N_e} \tilde{r}_{i,t}^2 \frac{\lambda^2}{s_i^2 + \lambda^2} + 2N.$$
(5.26)

Estimates $\hat{\sigma}$ and λ can be obtained by minimizing this value and the minimum ABIC can be employed for the model comparison. Note that this expression is identical to Eq.(3.14) except \tilde{v} being replaced by \tilde{r} .

If we assume a parametric model for the dynamics A_t , the parameters $\boldsymbol{\theta}$ of the dynamics of which are unknown, need to be estimated also. This can be done again by minimizing ABIC, as given by Eq.(5.26), but now the innovations (1-step prediction errors) \boldsymbol{r}_t depend on these parameters, such that $ABIC(\sigma, \lambda)$ becomes $ABIC(\sigma, \lambda, \boldsymbol{\theta})$. In the implementation, this optimization is carried out by the simplex method, as provided by the "fminsearch" function of MATLAB.

5.2.3 Relation between RPLS method and Kalman filtering

The estimation procedure of RPLS method has the same structure as known from Kalman filtering : first the innovation is calculated (Eq.(5.17)) using the previous estimate (i.e. forming a prediction) and the current observation, then ζ_t is calculated (Eq.(5.19), corresponding to filtering) from the innovation. Let $j_{t-1|t-1}$ and $\Sigma_{t-1|t-1}$ denote the filtered state estimate and the filtered state error variance, respectively, as provided by Kalman filtering at time t - 1. At time t a new observation v_t becomes available, and the state estimate is updated according to

$$\boldsymbol{j}_{t|t} = A_t \boldsymbol{j}_{t-1|t-1} + \mathcal{K} \boldsymbol{r}_t^K \tag{5.27}$$

$$\mathcal{K} = \{ \mathbf{K}' C_{\epsilon}^{-1} \mathbf{K} + (\sigma^2 c_{t-1} A_t P_{t-1} A_t' + \lambda^{-2} C_{\eta})^{-1} \}^{-1} \mathbf{K}' C_{\epsilon}^{-1}$$
(5.28)

where \mathcal{K} denotes the Kalman gain, and the innovation \mathbf{r}_t^K is defined by $\mathbf{r}_t^K = \mathbf{v}_t - KA_t \mathbf{j}_{t-1|t-1}$. Here we denote the filtered error variance by $\Sigma_{t-1|t-1} = c_{t-1}P_{t-1}$. In the RPLS method, the equation corresponding to Eq.(5.27) is Eq.(5.20). Obviously, in Kalman filtering $\mathcal{K}\mathbf{r}_t^K$ is the estimator of system noise; the corresponding estimator is given by Eq.(5.19). By comparison with Eq.(5.21) it can readily be seen that the RPLS method becomes consistent with Kalman filtering if $c_{t-1} \rightarrow 0$. While, according to Eq.(5.28), the Kalman gain essentially depends on three components, representing system noise variance, observation noise variance and the uncertainty of the previous estimate, the RPLS method explicitly depends on only two of these components, namely system noise variance and observation noise variance.

5.3 Projection Kalman filtering

5.3.1 Algorithm

I shall introduce an approximation method of Kalman filtering by projecting the covariance matrices on certain appropriate subspace, which is the similar idea to the well-known dimension reduction method, Principle Component Analysis (PCA) in multivariate analysis. I would like to call this method **projection Kalman filtering**.

A projection matrix P is defined as a matrix of size $N_s \times N_p$, where $N_s \gg N_p$ and $P'P = I_{N_p}$. The first condition is for the dimension reduction and if the column vectors of P are linear independent each other, the orthogonalization is possible and gives the second condition without losing generality. By considering the projection of the original covariance matrices onto this space, for example, $C_{t|t-1} = P'\Sigma_{t|t-1}P$, then the update for the projected matrices $C_{t|t-1}$ and $C_{t-1|t-1}$ can be obtained as follows,

[Prediction]

$$\mathbf{j}_{t|t-1} = A_t \mathbf{j}_{t-1|t-1}$$
 (5.29)

$$C_{t|t-1} = \bar{A}_t C_{t-1|t-1} \bar{A}'_t + \tau^2 \bar{C}_{\eta}$$
(5.30)

[Filter]

$$\boldsymbol{r}_t = \boldsymbol{v}_t - \mathbf{K} \boldsymbol{j}_{t|t-1} \tag{5.31}$$

$$\Lambda_t = \bar{\mathbf{K}} C_{t|t-1} \bar{\mathbf{K}}' + \sigma^2 C_{\boldsymbol{\epsilon}}$$
(5.32)

$$\mathcal{K}_t = PC_{t|t-1}\bar{\mathbf{K}}'\Lambda_t^{-1} \tag{5.33}$$

$$\boldsymbol{j}_{t|t} = \boldsymbol{j}_{t|t-1} + \mathcal{K}_t \boldsymbol{r}_t \tag{5.34}$$

$$C_{t|t} = C_{t|t-1} - C_{t|t-1} \bar{K}' \Lambda_t^{-1} \bar{K} C_{t|t-1}$$
(5.35)

where $\bar{A}_t = P'A_tP$, $\bar{C}_{\eta} = P'C_{\eta}P$ and $\bar{K} = KP$. It is obvious that the choice of the projection matrix P is crucial for exploiting the important information in the filtered and predicted covariance matrices. In this thesis I shall propose to choose the matrix P based on the SVD of the lead field matrix K.

In concrete, at first the SVD of the lead field matrix K is considered:

$$\mathbf{K} = USW' \tag{5.36}$$

where U, S, W are matrices of size $N_e \times N_e, N_e \times N_e$ and $N_s \times N_e$, respectively. From the property of the SVD, $W'W = I_{N_e}$ is satisfied. If we consider a vector space spanned by the column vectors \boldsymbol{w}_i of W, that is, $\boldsymbol{j}_o = \sum_{i=1}^{N_e} a_i \boldsymbol{w}_i$, $(\exists i, a_i \neq 0)$, it can be easily seen that a vector \boldsymbol{j}_o satisfies

 $\mathbf{K}\boldsymbol{j}_{o}\neq0.$

This means the current vector on this space directly makes effect on the EEG observation through the lead field matrix K. In this sense the vector space spanned by the column vectors \boldsymbol{w}_i of W will be referred to the space as the **observable space** and can be considered to be one of the most influential subspace. Note that the observable space is the orthogonal complement of the kernel $\{\boldsymbol{j} \in R^{N_s} | K\boldsymbol{j} = 0\}$, sometimes referred to as the **null space**.

Motivated by this fact, here I would like to choose the matrix W as a projection matrix P, then the projection Kalman filtering algorithm becomes:

[Prediction]

$$\mathbf{j}_{t|t-1} = A_t \mathbf{j}_{t-1|t-1}$$
 (5.37)

$$C_{t|t-1} = \bar{A}_t C_{t-1|t-1} \bar{A}'_t + \tau^2 \bar{C}_{\eta}$$
(5.38)

[Filter]

$$\boldsymbol{r}_t = \boldsymbol{v}_t - \mathbf{K} \boldsymbol{j}_{t|t-1} \tag{5.39}$$

$$\Lambda_t = U(SC_{t|t-1}S + \sigma^2 U'C_{\epsilon}U)U'$$
(5.40)

$$\mathcal{K}_t = W C_{t|t-1} S U' \Lambda_t^{-1} \tag{5.41}$$

$$\boldsymbol{j}_{t|t} = \boldsymbol{j}_{t|t-1} + \mathcal{K}_t \boldsymbol{r}_t \tag{5.42}$$

$$C_{t|t} = C_{t|t-1} - C_{t|t-1}S(SC_{t|t-1}S + \sigma^2 U'C_{\epsilon}U)^{-1}SC_{t|t-1}$$
(5.43)

where $\bar{A}_t = W'A_tW$, $\bar{C}_{\eta} = W'C_{\eta}W$, matrices of size $N_e \times N_e$. This algorithm is mostly constituted by the computation of matrices of size $N_e \times N_e$. I choose to call this filtering algorithm **observable projection Kalman filtering**.

5.3.2 Derivation

In projection Kalman filtering, the expectation updating is consistent with that of ordinal Kalman filtering whereas the variance updating is done for the projected covariance matrix. Here the equations for the variance updating are derived.

We assume to have a projection filtered covariance matrix $C_{t-1|t-1}$ in the previous updating and a projection matrix P of size $N_s \times N_p$ with $P'P = I_{N_p}$ and $N_s \gg N_p$. Note that the relation between the projection filtered covariance $C_{t|t}$ and the full filtered covariance $\Sigma_{t|t} = \operatorname{Var}(\boldsymbol{j}_t|\boldsymbol{v}_1,\cdots,\boldsymbol{v}_t)$ is given by the equation

$$C_{t|t} = P' \Sigma_{t|t} P.$$

As in Eq.(5.5), the prediction variance updating of ordinal Kalman filtering is given by

$$\Sigma_{t|t-1} = A_t \Sigma_{t-1|t-1} A'_t + \tau^2 C_{\eta}.$$
(5.44)

By projecting both hand-side onto the space spanned by the column vectors of the projection matrix P, we have

$$P'\Sigma_{t|t-1}P = P'A_t\Sigma_{t-1|t-1}A_t'P + \tau^2 P'C_{\eta}P.$$
(5.45)

In order to obtain prediction updating of the projection covariance matrix $C_{t|t-1}$, the full covariance matrix is approximated by

$$\Sigma_{t-1|t-1} \approx (P')^{-} C_{t|t-1} P^{-},$$
 (5.46)

where P^- is any pseudoinverse matrix of P. Now the Moore-Penrose pseudo inverse matrix of P, which is equal to P' $(P^+ \equiv (P'P)^{-1}P' = P')$, is used for P^- and the substitution of resulting approximation $\Sigma_{t-1|t-1}$ into Eq.(5.45) leads to the prediction variance updating of $C_{t|t-1}$:

$$C_{t|t-1} = \bar{A}_t C_{t-1|t-1} \bar{A}'_t + \tau^2 \bar{C}_{\eta}.$$
(5.47)

where $\bar{A}_t = P'A_tP$, $\bar{C}_{\eta} = P'C_{\eta}P$ are the projected transition matrix and the projected system noise covariance matrix, respectively.

In the same way, the filtered variance updating can be obtained as

$$P'\Sigma_{t|t}P = P'\Sigma_{t|t-1}P - P'\Sigma_{t|t-1}\mathbf{K}'\Lambda_t^{-1}\mathbf{K}\Sigma_{t|t-1}P,$$

and is arranged:

$$C_{t|t} = C_{t|t-1} - C_{t|t-1} \bar{K}' \Lambda_t^{-1} \bar{K} C_{t|t-1}.$$
(5.48)

where $\bar{\mathbf{K}} = \mathbf{K}P$ and

$$\Lambda_t = K\Sigma_{t|t-1}K' + \sigma^2 C_{\epsilon}$$

= $\bar{K}C_{t|t-1}\bar{K}' + \sigma^2 C_{\epsilon}.$ (5.49)

When the matrix P is chosen as W of the SVD (5.36), the filtered variance updating can be further simplified. Since

$$\bar{\mathbf{K}} = \mathbf{K}W = USW'W = US,\tag{5.50}$$

and

$$\Lambda_t^{-1} = (USC_{t|t-1}S'U' + \sigma^2 C_{\epsilon})^{-1} = U(SC_{t|t-1}S' + \sigma^2 U'C_{\epsilon}U)^{-1}U'$$
(5.51)

then we obtain the filtered variance updating of observable projection Kalman filtering:

$$C_{t|t} = C_{t|t-1} - C_{t|t-1} \bar{K}' \Lambda_t^{-1} \bar{K} C_{t|t-1}$$

= $C_{t|t-1} - C_{t|t-1} S' U' U (S C_{t|t-1} S' + \sigma^2 U' C_{\epsilon} U)^{-1} U' U S C_{t|t-1}$
= $C_{t|t-1} - C_{t|t-1} S (S C_{t|t-1} S + \sigma^2 U' C_{\epsilon} U)^{-1} S C_{t|t-1}.$ (5.52)

Here $U'U = I_{N_e}$ is employed.

5.3.3 Property

Here some theoretical aspects of observable projection Kalman filtering are discussed. In this subsection the following linear model is only considered for simplicity

$$\boldsymbol{v}_t = \mathbf{K} \boldsymbol{j}_t + \boldsymbol{\epsilon}_t \quad \boldsymbol{\epsilon}_t \sim N(0, \sigma^2 C_{\boldsymbol{\epsilon}})$$
 (5.53)

$$\boldsymbol{j}_t = A \boldsymbol{j}_{t-1} + \boldsymbol{\eta}_t \quad \boldsymbol{\eta}_t \sim N(0, \tau^2 C_{\boldsymbol{\eta}}).$$
(5.54)

If C_{η} is a positive definite matrix, without loss of the generality, we assume $C_{\eta} = I_{N_s}$ by using the prewhitening approach as will introduced in section 5.5.

Let us consider the SVD of the lead field matrix as follows:

$$\mathbf{K} = USW_1' \tag{5.55}$$

This form is same as Eq.(5.36) and U, S, W_1 are uniquely determined (if all the singular values are distinct). Then the augmented version of the SVD can be expressed as follows:

$$\mathbf{K} = U \begin{bmatrix} s_1 & & \\ & \ddots & \\ & & s_{N_e} \end{bmatrix} \begin{bmatrix} \mathbf{0} & \\ & & \end{bmatrix} \begin{bmatrix} & W_1 & W_2 \\ & & & \end{bmatrix}^{\prime}, \quad (5.56)$$

where s_i is the *i*th diagonal element of S and W_2 is a matrix of size $(N_s - N_e) \times N_s$ which constitutes the null space $\{\mathbf{j} \in \mathbb{R}^{N_s} | \mathbf{K}\mathbf{j} = 0\}$, the orthogonal complement of W_1 . W_2 can not be determined uniquely, however the column vector of W_2 can be constructed column by column using the Gram-Schmidt orthonomarization. The resulting matrix $W_A = [W_1 | W_2]$ becomes a orthogonal matrix, that is, $W'_A W_A = W_A W'_A = \mathbf{I}$ is satisfied. Note that the zero matrix of size $N_e \times (N_s - N_e)$ is added to S, corresponding to W_2 .

Now we apply the similarity transformation to the current vector \mathbf{j}_t by W'_A

$$\boldsymbol{h}_t = W'_A \boldsymbol{j}_t,$$

resulting in the new state space representation for the state h_t as

$$\boldsymbol{v}_t = US_A \boldsymbol{h}_t + \boldsymbol{\epsilon}_t \tag{5.57}$$

$$\boldsymbol{h}_t = A\boldsymbol{h}_{t-1} + \tilde{\boldsymbol{\eta}}_t. \tag{5.58}$$

where $\tilde{A} = W'_A A W_A$ and $\tilde{\eta}_t = W'_A \eta_t$ and $S_A = [S \mid \mathbf{0}]$. Here the covariance matrix of $\tilde{\eta}_t$ again becomes $\tau^2 \mathbf{I}_{N_s}$ because W_A is a orthogonal matrix. Note that the original state \mathbf{j}_t can be easily obtained from \mathbf{h}_t by $\mathbf{j}_t = W_A \mathbf{h}_t$.

Then we partition \mathbf{h}_t according to the partition of the matrix S_A , as denoted by $\mathbf{h}_t = (\mathbf{h}'_{t,o}, \mathbf{h}'_{t,n})'$. Since the matrix S_A has nonzero entries only in the first N_e columns, only the corresponding components of \mathbf{h}_t , that is, $\mathbf{h}_{t,o}$ can make



Figure 5.1: Schematic of equation (5.60). If there is no information flow from $h_{t,n}$ to $h_{t,o}$ (i.e. $\tilde{A}_{on} = 0$), the observations v_t have no information about $h_{t,n}$.

effect on the observation:

$$\boldsymbol{v}_{t} = U \begin{bmatrix} S \mid \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{h}_{t,o} \\ \\ \\ \\ \boldsymbol{h}_{t,n} \end{bmatrix} = US\boldsymbol{h}_{t,o}.$$
(5.59)

Relating to the partition of h_t , the system equation (5.58) can be rewritten as follows

$$\begin{bmatrix} \boldsymbol{h}_{t,o} \\ \boldsymbol{h}_{t,n} \end{bmatrix} = \begin{bmatrix} \tilde{A}_{oo} & \tilde{A}_{on} \\ \tilde{A}_{no} & \tilde{A}_{nn} \end{bmatrix} \begin{bmatrix} \boldsymbol{h}_{t-1,o} \\ \boldsymbol{h}_{t-1,n} \end{bmatrix} + \tilde{\boldsymbol{\eta}}_t.$$
(5.60)

If $\tilde{A}_{on} = 0$, then $\boldsymbol{h}_{t,o} = \tilde{A}_{on} \boldsymbol{h}_{t-1,o} + \tilde{\boldsymbol{\eta}}_{t,o}$. Therefore the components $\boldsymbol{h}_{t,o}$ has no information about $\boldsymbol{h}_{t',n}, t' < t$, resulting in no effect of $\boldsymbol{h}_{t',n}$ on the observation

 \boldsymbol{v}_t . Recall that only the component $\boldsymbol{h}_{t,o}$ of \boldsymbol{h}_t makes effect on the observation \boldsymbol{v}_t through the observation matrix. Intuitively speaking from the consideration above, the component $\boldsymbol{h}_{t,n}$ is redundant when $\tilde{A}_{on} = 0$. In other words, the dimension of the state can be reduced to $\boldsymbol{h}_{t,o}$ of small dimension N_e from the full state \boldsymbol{h}_t of huge dimension N_s .

In order to consider this argument more precisely, firstly let us partition the covariance matrices of the filtered (predicted or smoothed) state \boldsymbol{h}_t and the system noise $\tilde{\boldsymbol{\eta}}_t$ as well as the transition matrix A, corresponding to the partition of the state $\boldsymbol{h}_t = (\boldsymbol{h}'_{t,o}, \boldsymbol{h}'_{t,n})'$:

$$C_{t|s} = \begin{bmatrix} C_{t|s}^{oo} & C_{t|s}^{on} \\ \hline C_{t|s}^{no} & C_{t|s}^{nn} \end{bmatrix}, \quad C_{\tilde{\eta}} = \begin{bmatrix} C_{\tilde{\eta}}^{oo} & \mathbf{0} \\ \hline \mathbf{0} & C_{\tilde{\eta}}^{nn} \end{bmatrix}, \quad \tilde{A} = \begin{bmatrix} \tilde{A}_{oo} & \tilde{A}_{on} \\ \hline \tilde{A}_{no} & \tilde{A}_{nn} \end{bmatrix}$$

Without loss of generality $C_{\tilde{\eta}}$ assumes to be a diagonal matrix (by applying the similarity transformation $C_{\tilde{\eta}}^{-1/2}$).

Now we describe the Kalman filtering algorithm for the state space representation (5.57) and (5.58) using this partitioned matrix. The prediction variance updating of $C_{t|t-1}^{oo}$, $C_{t|t-1}^{no}$ can be written as:

$$C_{t|t-1}^{oo} = \tilde{A}_{oo} C_{t-1|t-1}^{oo} \tilde{A}'_{oo} + \Delta^{oo} + C_{\tilde{\eta}}^{oo}$$
(5.61)

$$C_{t|t-1}^{no} = \tilde{A}_{nn} C_{t-1|t-1}^{no} \tilde{A}'_{oo} + \Delta^{no}$$
(5.62)

where

$$\Delta^{oo} = \tilde{A}_{on} C^{nn}_{t-1|t-1} \tilde{A}'_{on} + \tilde{A}_{oo} C^{on}_{t-1|t-1} \tilde{A}'_{on} + \tilde{A}_{on} C^{no}_{t-1|t-1} \tilde{A}'_{oo} \Delta^{no} = \tilde{A}_{no} C^{oo}_{t-1|t-1} \tilde{A}'_{oo} + \tilde{A}_{no} C^{on}_{t-1|t-1} \tilde{A}'_{on} + \tilde{A}_{nn} C^{nn}_{t-1|t-1} \tilde{A}'_{on}.$$

The innovation covariance matrix Λ_t and the Kalman gain \mathcal{K}_t^h can be simplified as:

$$\Lambda_t = \begin{bmatrix} US & \mathbf{0} \end{bmatrix} C_{t|t-1} \begin{bmatrix} US & \mathbf{0} \end{bmatrix}' + C_{\boldsymbol{\epsilon}} = USC_{t-1|t-1}^{oo} S'U' + C_{\boldsymbol{\epsilon}}.$$
(5.63)

$$\begin{aligned}
\mathcal{K}_{t}^{h} &= C_{t|t-1}(US_{A})'\Lambda_{t}^{-1} \\
&= \begin{bmatrix} C_{t|t-1}^{oo} \\ C_{t|t-1}^{no} \end{bmatrix} S'U'(USC_{t-1|t-1}^{oo}S'U' + C_{\epsilon})^{-1}.
\end{aligned}$$
(5.64)

Finally the filtered variance updating can be calculated as follows,

$$C_{t|t} = (I - \mathcal{K}_{t}^{h}(US_{A})')C_{t|t-1}$$

= $C_{t|t-1} - \begin{bmatrix} C_{t|t-1}^{oo} \\ C_{t|t-1}^{no} \end{bmatrix} R \begin{bmatrix} C_{t|t-1}^{oo} & C_{t|t-1}^{no} \end{bmatrix},$

where $R = S'U'\Lambda_t^{-1}SU$. The corresponding portion of $C_{t|t}^{oo}$ and $C_{t|t}^{no}$ are obtained by

$$C_{t|t}^{oo} = C_{t|t-1}^{oo} - C_{t|t-1}^{oo} R C_{t|t-1}^{oo}$$
(5.65)

$$C_{t|t}^{no} = C_{t|t-1}^{no} - C_{t|t-1}^{no} R C_{t|t-1}^{oo}$$
(5.66)

The equations necessary for the prediction and filtered variance updating of \mathbf{h}_t are described in (5.61) ~ (5.66) using the partitioned matrices. It should be noted that only the portions of the covariance matrix relating to observable space, that is, $C_{t|t-1}^{oo}$, $C_{t|t-1}^{no}$ are required for the derivation of the Kalman gain (5.64). The variance updating of these parts are shown in Eqs.(5.61),(5.62),(5.65) and (5.66).

From the discussion above, the following proposition are satisfied.

Proposition 1 If $\tilde{A}_{on} = 0$ and $\tilde{A}_{no} = 0$, under some conditions about the initial state estimates $\mathbf{h}_{0|0}$ and covariance $C_{0|0}$, Kalman filtering for the state \mathbf{h}_t (see Eq.(5.58)) is equivalent to that for the state $\mathbf{h}_{t,o}$ of small dimension N_e of which state space representation is given by

$$\boldsymbol{v}_t = US\boldsymbol{h}_{t,o} + \boldsymbol{\epsilon}_t \tag{5.67}$$

$$\boldsymbol{h}_{t,o} = A_{oo}\boldsymbol{h}_{t-1,o} + \tilde{\boldsymbol{\eta}}_{t,o}.$$
(5.68)

Proof Since $\tilde{A}_{on} = 0$, $\tilde{A}_{no} = 0$ from the assumption, $\Delta^{oo} = \Delta^{no} = \mathbf{0}$. Thus Eqs. (5.61),(5.62) become

$$C_{t|t-1}^{oo} = \tilde{A}_{oo} C_{t-1|t-1}^{oo} \tilde{A}_{oo} + C_{\tilde{\eta}}^{oo}$$
(5.69)

$$C_{t|t-1}^{no} = \tilde{A}_{nn} C_{t-1|t-1}^{no} \tilde{A}'_{oo}.$$
(5.70)

Eqs.(5.69),(5.70),(5.64),(5.65) and (5.66) constitute Kalman filtering to the full state \mathbf{h}_t (under the assumptions). If the initial state variance is specified so that $C_{0|0}^{no} = 0$, then Eqs.(5.70),(5.66) lead to $C_{t|t}^{no} = 0$ for any t. Hence the exact Kalman filtering algorithms for \mathbf{h}_t is consisting of Eqs.(5.69),(5.64) and (5.65).

On the other hand, Kalman filtering to the state $h_{t,o}$ is consisting of Eqs.(5.69), (5.65) and

$$\mathcal{K}_{t}^{h_{o}} = C_{t|t-1}^{oo} S' U' (USC_{t-1|t-1}^{oo} S' U' + C_{\epsilon})^{-1}.$$
(5.71)

Since this Kalman gain is equivalent to the Kalman gain in Eq.(5.64) when $C_{t|t-1}^{no} = 0$, the updating of state covariance matrices are equivalent.

The original state \boldsymbol{j}_t is related to the state \boldsymbol{h}_t and $\boldsymbol{h}_{t,o}$ by $\boldsymbol{j}_t = W_A \boldsymbol{h}_t = W_1 \boldsymbol{h}_{t,o} + W_2 \boldsymbol{h}_{t,n}$, thus if $\boldsymbol{h}_{0|0,n} = \boldsymbol{0}$ then $W_A \boldsymbol{h}_{t|t} = W_1 \boldsymbol{h}_{t|t,o}$.

(Remark 1) The conditions against an initial state sound very strong. However in the absence of strong prior knowledge, the initial covariance matrix is often assumed to be a diagonal matrix thereby the assumption for the covariance matrix is satisfied. The assumption $\mathbf{h}_{0|0,n} = \mathbf{0}$ means the initial state is estimated so that the estimate \mathbf{j}_0 is on the space spanned by the column vectors of W_1 . Even if $\mathbf{h}_{0|0,n} \neq \mathbf{0}$, this difference might make little effect because $\mathbf{h}_{0|0,n}$ does not depend on the observations. As will be shown later, observable projection Kalman filtering is consisting of the variance updating of $C_{t|t-1}^{oo}$ and the mean updating of the full state \mathbf{h}_t .

(**Remark 2**) The Kalman gain (5.64) only requires $C_{t|t-1}^{oo}$, $C_{t-1|t-1}^{no}$ because the last $N_s - N_e$ columns of the observation matrix US_A is consisting of zeros. It is notable that huge matrices $C_{t|t-1}^{nn}$ need not be kept and computed.

(**Remark 3**) The assumption $\tilde{A}_{no} = 0$ is also needed because if not, Eq.(5.70) will become

$$C_{t|t-1}^{no} = \tilde{A}_{no} C_{t-1|t-1}^{oo} \tilde{A}'_{oo} + \tilde{A}_{nn} C_{t-1|t-1}^{no} \tilde{A}'_{oo}.$$

Thus $C_{t|t-1}^{no} \neq \mathbf{0}$ for some t even if $C_{0|0}^{no} = 0$.

(**Remark 4**) This proposition indicates that under the assumptions $\tilde{A}_{on} = 0$ and $\tilde{A}_{no} = 0$, the dimension of the state can be reduced considerably. Though the assumptions $\tilde{A}_{on} = 0$ and $\tilde{A}_{no} = 0$ seem to be strong, the most feasible choice of the dynamics in many applications, that is, the random walk model satisfies this assumption. It is easy to check that a transition matrix of the form $A = aI_{N_s}$, where a is a scalar constant, satisfies this condition, since

$$\tilde{A} = W'_A A W_A = a I_{N_s}. \tag{5.72}$$

Proposition 2 For the random walk model, it is sufficient to consider the reduced state space representation (5.67), (5.68) under the mild condition of the initial filtered variance.

Now we discuss observable projection Kalman filtering of the state space representation (5.53),(5.54). Firstly the similarity transformation W_A is applied to the original state \mathbf{j}_t so that the state space is represented by (5.57),(5.58). In this representation, the state variance is updated as in (5.69) ~ (5.66) whereas the state estimates (mean) is updated as follows

$$\boldsymbol{h}_{t|t} = \boldsymbol{h}_{t|t-1} + \mathcal{K}_t^h(\boldsymbol{v}_t - US_A \boldsymbol{h}_{t|t-1}).$$
(5.73)

The equation (5.73) can be represented by the original state \mathbf{j}_t and W_A :

$$\boldsymbol{j}_{t|t} = \boldsymbol{j}_{t|t-1} + W'_A \mathcal{K}^h_t (\boldsymbol{v}_t - \mathbf{K} \boldsymbol{j}_{t|t-1})$$
(5.74)

Thus it is possible to update the state variance of h_t while updating the state estimates of j_t . Note that the state variance updating (5.69) ~ (5.66) still demands huge memory consumption (\tilde{A} is a $N_s \times N_s$ matrix), which makes impracticable to employ Kalman filtering directly. Furthermore the similarity transformation W_A can not be determined uniquely.

Observable projection Kalman filtering can be obtained by projecting all the matrix relating to the variance updating on the space spanned by W_1 which is uniquely determined from the lead field matrix. In other words, instead of using (5.69) ~ (5.66), Eqs.(5.69),(5.65) and (5.64) with $C_{t|s}^{no} = 0$ are employed for updating the filtered and the predicted variance. In summary observable projection Kalman filtering updates the state estimation for the full state \mathbf{j}_t while the state variance is approximately updated in the observable space. It should be noted the filtering of $\mathbf{h}_{t,o}$ in (5.67),(5.68) corresponds to updating both the state estimates and variance in the observable space. If the dynamics A such that the assumptions in Prop.1 are satisfied is used, observable projection Kalman filtering is completely consistent with ordinal Kalman filtering. Therefore the computation and memory cost can be drastically reduced without loss of any information due to the approximation.

Proposition 3 For the random walk model, observable projection Kalman filtering can works without any approximation under the mild condition against the initial filtered variance.

(**Remark**) From Eq.(5.64), the bias of the filtered state estimate due to the approximation can be evaluated as

$$- \begin{bmatrix} \Delta^{oo} \\ C^{no}_{t|t-1} \end{bmatrix} S' U' \Lambda_t^{-1} \boldsymbol{r}_t$$

where $\Delta^{oo} = \tilde{A}_{on}C^{nn}_{t-1|t-1}\tilde{A}'_{on} + \tilde{A}_{oo}C^{on}_{t-1|t-1}\tilde{A}'_{on} + \tilde{A}_{on}C^{no}_{t-1|t-1}\tilde{A}'_{oo}$ and \mathbf{r}_t is the innovation.

It is obvious that projection Kalman filtering approaches to ordinary Kalman filtering when W_1 is augmented to W_A . However in order to appropriately choose new basis which is incorporated into W_1 , additional knowledge about the electrical sources is needed. The use of the transition matrix through the observation matrix is possible. For example, the SVD of the truncated observable matrix (see Kailath 1980 or Aoki 1987 for the observable matrix)

$$\mathcal{O}_i = (\mathbf{K}', (\mathbf{K}A)', \cdots, (\mathbf{K}A^i)')' \tag{5.75}$$

could be used as an immediate extension of observable projection Kalman filtering (corresponding to the case i = 0).

5.4 Partitioned Kalman filtering

5.4.1 Algorithm

This approximate filtering method has been first applied for the EEG dynamical inverse problem by Galka et al. 2002. In the literature the authors refer to "spatio-temporal Kalman filtering". However, I would like to call it "partitioned Kalman filtering" because it is a key idea to partition the state \mathbf{j}_t of large dimension into sets of the state of small dimension.

Let j_t^v and a_{ij} denote the *v*th element of \mathbf{j}_t and the *ij*th element of a transition matrix A, respectively. We further define the current source vector, the lead field matrix and the transition matrix removed the *v*th corresponding element as follows,

Here the vector $\mathbf{k}_{\cdot v}$ denotes the *v*th column vector of the lead field matrix K.

Then the model (5.3) for the state of one voxel j_t^v can be written as

$$\boldsymbol{v}_t = \boldsymbol{k}_{\cdot v} j_t^v + \mathbf{K}^{-v} \boldsymbol{j}_t^{-v} + \boldsymbol{\epsilon}_t \quad \boldsymbol{\epsilon}_t \sim N(0, \sigma^2 C_{\boldsymbol{\epsilon}})$$

$$j_t^v = a_{vv} j_{t-1}^v + \boldsymbol{a}^{-v} \boldsymbol{j}_{t-1}^{-v} + \eta_t^v \quad \eta_t^v \sim (0, \tau^2).$$

This is the state space representation for the **partitioned state** j_t^v . By regarding the variable \mathbf{j}_t^{-v} as the exogenous variable (i.e. a constant with respect to j_t^v), the following algorithm can be obtained.

[Prediction]

for each voxel $v, (v = 1, \cdots, N_s)$

$$j_{t|t-1}^{v} = a_{vv}j_{t-1|t-1}^{v} + \boldsymbol{a}^{-v}\boldsymbol{j}_{t-1|t-1}^{-v}$$
(5.76)

$$\nu_{t|t-1}^{v} = a_{vv}^{2} \nu_{t-1|t-1}^{v} + \sum_{i \neq v} a_{vi}^{2} \nu_{t-1|t-1}^{i} + \tau^{2}$$
(5.77)

[Filter]

$$\boldsymbol{r}_t = \boldsymbol{v}_t - \mathbf{K} \boldsymbol{j}_{t|t-1} \tag{5.78}$$

$$\Lambda_t = \sum_{i=1}^{N_s} \mathbf{k}_{\cdot i} \nu_{t|t-1}^i \mathbf{k}'_{\cdot i} + \sigma^2 C_{\boldsymbol{\epsilon}}$$
(5.79)

for each voxel $v, (v = 1, \cdots, N_s)$

$$\mathcal{K}_t^v = \nu_{t|t-1}^v \mathbf{k}'_{\cdot v} \Lambda_t^{-1} \tag{5.80}$$

$$j_{t|t}^{v} = j_{t|t-1}^{v} + \mathcal{K}_{t}^{v} \boldsymbol{r}_{t}$$
(5.81)

$$\nu_{t|t}^{v} = \nu_{t|t-1}^{v} - \mathcal{K}_{t}^{v} \mathbf{k}_{\cdot v} \nu_{t|t-1}^{v}$$
(5.82)

[Fixed interval smoothing]

for each voxel $v, (v = 1, \cdots, N_s)$

$$b_{t,vi} = \frac{\nu_{t|t}^v a_{iv}}{\nu_{t+1|t}^i}$$
(5.83)

$$j_{t|T}^{v} = j_{t|t}^{v} + b_{t,vv}(j_{t+1|T}^{v} - j_{t+1|t}^{v}) + \sum_{i \neq v} b_{t,vi}(j_{t+1|T}^{i} - j_{t+1|t}^{i})$$
(5.84)

$$\nu_{t|T}^{v} = \nu_{t|t}^{v} + b_{t,vv}^{2}(\nu_{t+1|T}^{v} - \nu_{t+1|t}^{v}) + \sum_{i \neq v} b_{t,vi}^{2}(\nu_{t+1|T}^{i} - \nu_{t+1|t}^{i}) \quad (5.85)$$

where $j_{t|s}^v, \nu_{t|s}^v$ is the conditional expectation and variance of j_t^v , which are scaler values. The value $b_{t,vi}$ is the *vi*th element of B_t in Eq.(5.11).

It is important to see that to regard the variable \mathbf{j}_t^{-v} as the exogenous variable against j_t^v corresponds to ignore the correlation between the elements in the vector \mathbf{j}_t^{-v} and the variable j_t^v . Since this technique mainly makes use of the information from the diagonal components of the original filtered and prediction covariance matrices and only the computation for small matrices is needed, the computation and memory consumption can crucially be reduced, compared with the original Kalman filter algorithm. It should be noted that it is possible to take the correlation within the partitioned state into consideration by a larger partition of the original state \mathbf{j}_t (the above-mentioned algorithm is in the case of the smallest partition i.e. the partitioned state corresponds to each element of the state \mathbf{j}_t).

5.4.2 Derivation

In order to make clear what kind of approximations are employed, the derivation of the above-mentioned algorithm will be shown.

Let $V_t = (v_1, v_2, \dots, v_t)$ denote the space spanned by the observations up to t. Given $j_{t-1|t-1}^v$ and $\nu_{t-1|t-1}$ for all voxels v, the predicted state expectation and variance can be calculated,

$$j_{t|t-1}^{v} = \mathrm{E}[j_{t}^{v}|\boldsymbol{V}_{t-1}]
 = \mathrm{E}[j_{t}^{v}|j_{t-1|t-1}^{1},\cdots,j_{t-1|t-1}^{N_{s}}]
 = a_{vv}j_{t-1|t-1}^{v} + \boldsymbol{a}^{-v}\boldsymbol{j}_{t-1|t-1}^{-v},$$
(5.86)

and

$$\nu_{t|t-1}^{v} = \mathrm{E}[(j_{t}^{v} - j_{t|t-1}^{v})^{2} | \boldsymbol{V}_{t-1}] \\
= \mathrm{E}[\{a_{vv}(j_{t-1}^{v} - j_{t-1|t-1}^{v}) + \boldsymbol{a}^{-v}(\boldsymbol{j}_{t-1}^{-v} - \boldsymbol{j}_{t-1|t-1}^{-v}) + \eta_{t}^{v}\}^{2} | \boldsymbol{V}_{t-1}] \\
= a_{vv}^{2} \mathrm{Var}(j_{t-1}^{v} | \boldsymbol{V}_{t-1}) + \mathrm{Var}(\boldsymbol{a}^{-v} \boldsymbol{j}_{t-1}^{-v} | \boldsymbol{V}_{t-1}) \\
+ \mathrm{Cov}(a_{vv} j_{t-1}^{v}, \boldsymbol{a}^{-v} \boldsymbol{j}_{t-1}^{-v} | \boldsymbol{V}_{t-1}) + \tau^{2} \\
\approx a_{vv}^{2} \nu_{t-1|t-1}^{v} + \sum_{i \neq v} a_{vi}^{2} \nu_{t-1|t-1}^{i} + \tau^{2}.$$
(5.87)

From the third equation to the last equation (5.87), the approximation

$$\operatorname{Cov}(j_{t-1}^{i}, j_{t-1}^{k} | \boldsymbol{V}_{t-1}) \approx 0 \quad i \neq k$$
 (5.88)

has been employed. It is obvious that the third term of the third line is vanished from this assumption and the second term can be simplified as

$$\operatorname{Var}(\boldsymbol{a}^{-v}\boldsymbol{j}_{t-1}^{-v}|\boldsymbol{V}_{t-1}) = \sum_{i \neq v} a_{vi}^2 \operatorname{Var}(j_{t-1}^i|\boldsymbol{V}_{t-1}) + \sum_{i,k \neq v} a_{vi}a_{vj} \operatorname{Cov}(j_{t-1}^i, j_{t-1}^k|\boldsymbol{V}_{t-1}),$$

and the terms of the covariance are vanished again.

The innovation \mathbf{r}_t and innovation variance Λ_t can be calculated:

$$\boldsymbol{r}_t = \boldsymbol{v}_t - \mathbf{k}_{\cdot v} j_{t|t-1}^v - \mathbf{K}^{-v} \boldsymbol{j}_{t|t-1}^{-v}, \qquad (5.89)$$

and

$$\Lambda_{t} = \mathrm{E}[\{\mathbf{k}_{\cdot v}(j_{t}^{v} - j_{t|t-1}^{v}) + \mathrm{K}^{-v}(\mathbf{j}_{t}^{-v} - \mathbf{j}_{t|t-1}^{-v}) + \boldsymbol{\epsilon}\}^{2}]$$

$$= \mathbf{k}_{\cdot v} \mathrm{Var}(j_{t}^{v} | \mathbf{V}_{t-1}) \mathbf{k}_{\cdot v}' + \mathrm{Var}(\mathrm{K}^{-v} \mathbf{j}_{t-1}^{-v} | \mathbf{V}_{t-1})$$

$$+ \mathrm{Cov}(\mathbf{k}_{\cdot v} j_{t}^{v}, \mathrm{K}^{-v} \mathbf{j}_{t}^{-v} | \mathbf{V}_{t-1}) + \sigma^{2} C_{\boldsymbol{\epsilon}}$$

$$\approx \mathbf{k}_{\cdot v} \nu_{t|t-1}^{v} \mathbf{k}_{\cdot v}' + \sum_{i \neq v} \mathbf{k}_{\cdot i} \nu_{t|t-1}^{i} \mathbf{k}_{\cdot i}' + \sigma^{2} C_{\boldsymbol{\epsilon}}.$$
(5.90)

Again the similar arrangement as in the Eq.(5.87) has been employed from the third equation to the last equation (5.90).

Then the filtered state expectation can be calculated:

$$\begin{aligned} j_{t|t}^{v} &= \mathrm{E}[j_{t}^{v}|\boldsymbol{V}_{t-1}, \boldsymbol{r}_{t}] \\ &= \mathrm{E}[j_{t}^{v}|\boldsymbol{V}_{t-1}] + \mathrm{E}[j_{t}^{v}|\boldsymbol{r}_{t}] \\ &= j_{t|t-1}^{v} + \mathrm{Cov}(\boldsymbol{r}_{t}, j_{t}^{v})\mathrm{Var}(\boldsymbol{r}_{t})^{-1}\boldsymbol{r}_{t} \\ &= j_{t|t-1}^{v} + \mathcal{K}_{t}^{v}\boldsymbol{r}_{t} \end{aligned}$$

$$(5.91)$$

where the Kalman gain \mathcal{K}_t^v is defined and simplified as

$$\begin{aligned}
\mathcal{K}_{t}^{v} &= \operatorname{Cov}(\boldsymbol{r}_{t}, j_{t}^{v})\operatorname{Var}(\boldsymbol{r}_{t})^{-1} \\
&= \operatorname{E}\left[\operatorname{E}\left[\left(\mathbf{k}_{\cdot v}(j_{t}^{v} - j_{t|t-1}^{v}) + \operatorname{K}^{-v}(\boldsymbol{j}_{t}^{-v} - \boldsymbol{j}_{t|t-1}^{-v}) + \boldsymbol{\epsilon}\right)(j_{t}^{v} - j_{t|t-1}^{v})|\boldsymbol{V}_{t-1}\right]\right]\Lambda_{t}^{-1} \\
&\approx \nu_{t|t-1}^{v}\mathbf{k}_{\cdot v}\Lambda_{t}^{-1}.
\end{aligned}$$
(5.92)

Finally the filtered state variance can be calculated as

$$\nu_{t|t}^{v} = \operatorname{Var}(j_{t}^{v}|\boldsymbol{V}_{t-1},\boldsymbol{r}_{t})$$

$$= \operatorname{Var}(j_{t}^{v}|\boldsymbol{V}_{t-1}) - \operatorname{Cov}(j_{t}^{v},\boldsymbol{r}_{t})\operatorname{Var}(\boldsymbol{r}_{t})^{-1}\operatorname{Cov}(j_{t}^{v},\boldsymbol{r}_{t})'$$

$$= \nu_{t|t-1}^{v} - \mathcal{K}_{t}^{v}\Lambda_{t}\mathcal{K}_{t}^{v'}$$

$$= \nu_{t|t-1}^{v} - \mathcal{K}_{t}^{v}\mathbf{k}_{\cdot v}\nu_{t|t-1}^{v} \qquad (5.93)$$

In the derivation of the filtered state expectation and variance (5.91),(5.93), Lemma 2 in appendix B are employed. The derivation of fixed interval smoothing is direct from the standard Kalman smoothing algorithm in Sec.5.1.

5.4.3 Remarks

As shown in the previous subsection, the main assumption of this filtering technique is

$$\operatorname{Cov}(j_{t-1}^i, j_{t-1}^k | \boldsymbol{V}_{t-1}) \approx 0 \quad i \neq k.$$

The assumption that there is no correlation between the different voxels, seems to be unrealistic for this application. However if we combine the state of nearby voxels into one partitioned state, for example, in guise of knowledge of brain anatomy, this technique can exploit the information of the covariance within each partitioned state and will be more applicable. In my opinion, whether this method is appropriately employed or not depends on how to partition the state, therefore the knowledge of the specific application should be carefully taken into consideration for the partition. I would like to mention some additional remarks:

- For the case of the transition matrix A being sparse, this technique could be more suitable in point of computation cost because the summations in the prediction, filtering and smoothing steps are necessary only for a few terms.
- Positive definiteness of the predicted, filtered and smoothed covariance matrices is not guaranteed even in theory, while it is guaranteed in the Standard Kalman filter. This point will be a challenging future work.
- The system noise should be uncorrelated (i.e. C_{η} can be expressed by a diagonal (or near diagonal) matrix), because this technique mainly exploits the information of the covariance structure on the diagonal components. Hence the prewhitening (see Sec.5.5) is preferable before the filtering algorithm.
- The similar algorithm has already been employed in the field of data assimilation in geophysics (Fukumori 2001; Fukumori 2002).

5.5 Model

In the EEG source localization problem, no equation describing the dynamics of the dipoles on the cortex has been proposed so far. Hence our strategy in this thesis is an traditional approach in time series analysis, that is, we explore most appropriate dynamics in a class of the parametric model and evaluate the goodness of the model using the likelihood.

In the state space representation, the dynamics, the system and observation (instrumental) noise must be specified before the filtering algorithm.

5.5.1 Dynamics

In this thesis, only a class of the linear dynamics is explored so that the dynamics can be represented by a (transition) matrix of size $N_s \times N_s$. Since the dynamics represented by a dense matrix requires huge memory for keeping as well as parameter estimation of huge dimension, the dynamics being parameterized by rather small number of parameters and represented by a sparse matrix of size $N_s \times N_s$ is used.

• Random walk model:

$$\boldsymbol{j}_t = \boldsymbol{j}_{t-1} + \boldsymbol{\eta}_t, \tag{5.94}$$

This model corresponds to the case $A_t = I_{N_s}$ and is interpreted as imposing the temporal smoothness constraint on \mathbf{j}_t . The random walk model is one of the most feasible choice of the dynamics when no apriori knowledge is available. Because there is no parameter to be estimated, this model can be seen in many literatures of the dynamical inverse problem.

• Neighbor interaction AR model:

$$\boldsymbol{j}_t = A(\boldsymbol{\theta})\boldsymbol{j}_{t-1} + \boldsymbol{\eta}_t, A(\boldsymbol{\theta}) = a_0 \mathbf{I} + a_1 \mathbf{L}$$
 (5.95)

where $\boldsymbol{\theta} = (a_0, a_1)$. Since the Laplacian matrix L has entries only in the diagonal and their neighbor parts, this parameterization models the AR process with neighbor interactions. We can extend the nearest neighbor interactions to the second neighbor interactions by a parameterization,

$$A(\boldsymbol{\theta}) = a_0 \mathbf{I} + a_1 \mathbf{L} + a_2 \mathbf{L}^2.$$

Note that these dynamics are spatially homogeneous with isotropic interactions.

• Regional AR model:

$$\boldsymbol{j}_t = A(\boldsymbol{\theta})\boldsymbol{j}_{t-1} + \boldsymbol{\eta}_t, A(\boldsymbol{\theta}) = \text{diag}(a_1, \cdots, a_1, a_2, \cdots, a_2, \cdots, \cdots, a_R)$$
 (5.96)

where $\boldsymbol{\theta} = (a_1, a_2, \cdots, a_R)$. In Eq.(5.96), the same value of a parameter is taken in a certain region partitioned, for example, based on anatomical knowledge. This models the AR process with regional heteroscedasticity.

The random walk model is contained in the neighbor interaction AR model or the regional AR model as a special case. If more parameters are allowed, more complex models, for example, regional AR model with neighbor interactions can be considered. It should be noted that so far the number of parameters more than 10 is impracticable from the author's experience,

Note that if the past states up to p are incorporated into the dynamics, the dynamics is modeled by a matrix of size $(N_s \cdot p) \times (N_s \cdot p)$, which may cause computational difficulty even in these realistic parameterization of the dynamics.

5.5.2 Observation noise

The observation noise may reflect the property of electrolodes and the environment in measuring. These information, however, is generally difficult to be involved in the model. In this thesis, the simplest form, $C_{\eta} \propto I$, is assumed.

5.5.3 System noise

The LORETA solution with the spatial smoothness has been reported to attain the best localization between several instantaneous inverse solutions (Pascual-Marqui 1999). Motivated by this fact, we would like to impose the spatial smoothness on the source \mathbf{j}_t . However it is difficult to exactly impose the spatial smoothness constraint on the sources because in the state space representation employed through this thesis, the dynamics and the structure of the system noise are combined to constitute a spatio-temporal constraint onto the sources.

Our compromise for the system noise is $\eta_t = L^{-1} \boldsymbol{\xi}_t$ where $\boldsymbol{\xi}_t \sim N(0, \tau^2 I)$. By this, we assume the system noise $\boldsymbol{\eta}_t$ with spatial smoothness. How this system noise works as a constraint on the source \boldsymbol{j}_t can be seen from an example,

$$\boldsymbol{j}_t = \boldsymbol{j}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim N(0, \tau^2 (\mathrm{LL}')^{-1}). \tag{5.97}$$

If the previous source \mathbf{j}_{t-1} has the spatial smoothness, then \mathbf{j}_t also has the spatial smoothness because $\mathbf{\eta}_t$ is a random noise with spatial smoothness. Thus the model (5.97) imposes both the temporal and spatial smoothness on the sources \mathbf{j}_t , if the initial source vector \mathbf{j}_0 has the spatial smoothness. Unless the dynamics changes neighborhoods of the next future source rapidly, this system noise indirectly works as a spatial smoothness constraint on the source.

5.5.4 Prewhitening

Because $(LL')^{-1}$ is a dense matrix of size $N_s \times N_s$, the computational difficulty also arises. In practical, the **prewhitening** of the system noise covariance is employed before the filtering algorithm. By multiplying L on both hand-sides of the system equation, we obtain

$$\mathbf{L}\boldsymbol{j}_t = \mathbf{L}A_t \mathbf{L}^{-1} \mathbf{L} \boldsymbol{j}_{t-1} + \boldsymbol{\xi}_t \tag{5.98}$$

where A_t is any transition matrix and $\boldsymbol{\xi}_t \sim N(0, \tau^2 \mathbf{I})$. Then the state space representation for the new state $\tilde{\boldsymbol{j}}_t = \mathbf{L} \boldsymbol{j}_t$ can be considered:

1

$$\boldsymbol{v}_t = \tilde{\mathbf{K}} \tilde{\boldsymbol{j}}_t + \boldsymbol{\epsilon}_t \tag{5.99}$$

$$\tilde{\boldsymbol{j}}_t = \tilde{A}_t \tilde{\boldsymbol{j}}_{t-1} + \boldsymbol{\xi}_t \tag{5.100}$$

where $\tilde{K} = KL^{-1}$ and $\tilde{A}_t = LA_tL^{-1}$. Note that in the new representation, the system noise covariance matrix is diagonal. This approach can be also employed in the observation equation if the observation noise covariance structure is known.

5.5.5 Parameter estimation

The parameters Θ (including the dynamical parameters and the amplitude of noise variance) is estimated by the maximum likelihood estimate (MLE) method. As a virtue of Kalman filtering, the likelihood can easily obtained during the filtering process using the innovations (Schweppe 1965):

$$-2\log p(\boldsymbol{v}_1,\cdots,\boldsymbol{v}_T;\boldsymbol{\Theta}) = \sum_{t=1}^T \log |\Lambda_t(\boldsymbol{\Theta})| + \sum_{t=1}^T \boldsymbol{r}_t'(\boldsymbol{\Theta})\Lambda_t(\boldsymbol{\Theta})^{-1}\boldsymbol{r}_t(\boldsymbol{\Theta}).$$
(5.101)

where the innovation \mathbf{r}_t and the innovation variance Λ_t can be calculated as in Eqs.(5.6),(5.7) in Kalman filtering, as in Eqs.(5.39),(5.40) in observable projection Kalman filtering, and as in Eqs.(5.78),(5.79) in partitioned Kalman filtering. A constant term is ignored in Eq.(5.101). The MLE can be obtained by minimizing the -2 times log-likelihood (5.101).

In the RPLS method, the parameters are estimated from the type II loglikelihood, which is equivalent to the above-mentioned log-likelihood:

$$-2 \log p(\boldsymbol{v}_{1}, \cdots, \boldsymbol{v}_{T}; \boldsymbol{\Theta}) = T N_{e} \log \sigma^{2} + T \sum_{i=1}^{N_{e}} \log \frac{s_{i}^{2} + \lambda^{2}}{\lambda^{2}} + \frac{1}{\sigma^{2}} \sum_{t=1}^{T} \sum_{i=1}^{N_{e}} \tilde{r}_{i,t}^{2}(\boldsymbol{\Theta}) \frac{\lambda^{2}}{s_{i}^{2} + \lambda^{2}}.(5.102)$$

Here due to the approximation using in the RPLS method, the form (5.102) is simpler than Eq.(5.101). (see Sec.5.2 for detail).

5.5.6 Model Comparison

If there are many candidates of the model, the best candidate should be chosen using some objective criteria. In statistics, the goodness of solutions resulting from corresponding models is evaluated by likelihood so that a model with the highest likelihood is chosen as the best.

Unfortunately in the parametric approach the log-likelihood at the MLE $\hat{\Theta}$ has bias depending on the number of parameters in a model, therefore the

quantity to correct (or reduce) this bias has been introduced in various literatures; as examples, AIC (Akaike 1973), SIC (Schwarz 1978), AICc (Hurvich and Tsai 1989; Bengtsson and Cavanaugh).

In this study we choose to employ AIC because of its simple form and wide use of this criterion in various applications. AIC is defined as follows:

$$AIC = -2\mathcal{L}(\hat{\Theta}) + 2N_p \tag{5.103}$$

where N_p is the number of parameters in the model for fitting and $\mathcal{L}(\cdot)$ is the log-likelihood function.

Chapter 6

Results

In this chapter the results for the simulation studies and for the analysis of real EEG data will be presented. Firstly, for comparison of three approximate filtering algorithms and Kalman filtering algorithm, the small number of voxels (282 voxels) is used. After that, two simulation studies and the real data analysis will be demonstrated using the following practical settings in common:

- The lead field matrix K was calculated by using the boundary element method for a three-shell head model (Riera and Fuentes 1998).
- A brain model, derived from the Average Probabilistic MRI Atlas produced by the Montreal Neurological Institute (Mazziotta, Toga, Evans, Fox, and Lancaster 1995), was employed.
- The resolution of the voxel discretization was 7 mm, resulting in a total number of 8723 voxels. Generators are assumed to be located only within gray matter, therefore the number of voxels which have to be considered, reduces to 3433 (see Fig.6.2).
- The number and locations of EEG electrodes follows the standard 10-20 system.

6.1 Comparison of Filtering Algorithms

In order to compare the solutions obtained by three proposed filtering algorithms and Kalman filtering, the small number of voxels is used in this section. In this voxel set, the number of voxels is 282, therefore the dimension of the current vector to be estimated becomes 846. EEG measurements are observed at 12 electrodes on the scalp. A time series of T = 256 observations is generated from the model,

$$\boldsymbol{v}_{t} = \mathbf{K}\boldsymbol{j}_{t} + \boldsymbol{\epsilon}_{t}$$

$$\boldsymbol{j}_{t} = (a_{1}\mathbf{I}_{N_{s}} + b_{1}\mathbf{L})\boldsymbol{j}_{t-1} + (a_{2}\mathbf{I}_{N_{s}} + b_{2}\mathbf{L})\boldsymbol{j}_{t-2} + \boldsymbol{\eta}_{t}$$

$$\operatorname{Var}(\boldsymbol{\epsilon}_{t}) = \sigma^{2}\mathbf{I}_{N_{e}} \quad \operatorname{Var}(\boldsymbol{\eta}_{t}) = \tau^{2}(\mathbf{L}\mathbf{L}')^{-1}.$$

(6.1)

where $(a_1, a_2, b_1, b_2, \sigma, \tau) = (1.65, -1.00, 0.05, 0.00, 0.02, 0.01)$. Small amplitude of the system noise is chosen compared with the amplitude of the current source, whereas rather high amplitude of the observation noise is chosen so that the standard deviation of the noise corresponds to 30% as much as that of the noise-free observations.

Using the true initial states, the states $\mathbf{j}_1, \dots, \mathbf{j}_T$ is estimated by the RPLS method (denoted by RPLS), observable projection Kalman filtering (denoted by O-Kalman), partitioned Kalman filtering (denoted by P-Kalman) and ordinary Kalman filtering. In applying RPLS, O-Kalman, and P-Kalman, all the parameters are estimated. In applying ordinary Kalman filtering, the parameters estimated by O-Kalman is used because as shown in table 6.1, Kalman filtering takes too much time to optimize the parameters. In table 6.1, for each four algorithms, computational time for running the filtering algorithm for 256 observations and the values of root mean square error (RMSE) are listed. All the computation was done on a PC with a clock rate of 2GHz and with ROM of 2GB.

It can be seen that Kalman filtering attains the least RMSE in sacrifice of computational time. On the other hand, RPLS and O-Kalman attains reasonable accuracy with reasonable computational time. The performance of P-Kalman is a little worse both in accuracy and in computational efficiency in this case.

In Fig.6.1, for four specific voxels, the time series estimated by four algorithms and the true time series are presented. It can be observed that the estimated time series obtained by RPLS and O-Kalman are very similar. We can also see that the estimated time series of P-Kalman and Kalman are noisy especially on the voxels [B] and [C]. This is probably due to the inaccurate estimates of the observation and system noise variance. Since the ratio of these variance determines reliability for observations, the misspecification of this ratio tends to cause too noisy estimates. In addition again due to the inaccurate estimates of the dynamical parameters, on the voxel [D] the estimated time series by P-Kalman fails to chase the true time series. It should be noted that on the voxels with high amplitude such as [A] and [D], the time series estimated by four algorithms are not so different, whereas on the voxels with rather small amplitude such as [B] and [C], the difference can be observed as have discussed above. Therefore in total the difference of the solutions obtained by four algorithms is not so large. Actually when we see the results as movies, the difference is too small to distinguish.

	Computational Time (sec.)	RMSE
RPLS	1.24	0.74
O-Kalman	2.28	0.88
P-Kalman	27.1	4.49
Kalman	2580	0.45

 Table 6.1: Comparison of the filtering algorithms

6.2 Simulation Study I

In this simulation study, the dynamical inverse solution will be compared with the instantaneous solution. As the simulated sources, two extended current sources are set; one in the occipital area, the other in the frontal area (see the left panel of Fig.6.3). These two sources are designed to evolve in a deterministic and periodic way using trigonometric functions. In Fig.6.4, the time series of these two sources are plotted by the succession of circles. By multiplying the lead field matrix to the current source vectors, a time series of the EEG measurement without any noise was generated. Then, four different EEG observations were generated by adding Gaussian white noise of four level (5%, 10%, 20%, 33%).

For each four sets of the observation,

- the instantaneous inverse solution is obtained via LORETA.
- the dynamical inverse solution is obtained via DynLORETA.

Here DynLORETA consists of a dynamical model with a maximum spatial smoothness constraint and the RPLS method as a state estimation algorithm (Yamashita et al. 2004). As the dynamical model, the regional AR(2) model, where a whole brain is separated into three regions (occipital region, frontal region, the remaining region) is employed.

The goodness of the solutions is evaluated using the following four criterion;

• Root Mean Square Error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{N_e \cdot T} \sum_{t=1}^{T} ||\boldsymbol{j}(t) - \boldsymbol{\hat{j}}(t)||^2}$$

where $\mathbf{j}(t)$ and $\hat{\mathbf{j}}(t)$ are the true and the estimated current vector at time point t, respectively.

• Akaike Bayes Information Criteria (ABIC):

$$ABIC = -2\mathcal{L}^{(II)}(\hat{\lambda}, \hat{\theta}) + 2 \times N_p$$



Figure 6.1: Time series of the local current vector on four specific voxels. In four figures of the left-top panel [A], from the top to the bottom. the time series of one of four voxels estimated by RPLS, O-Kalman, P-Kalman and Kalman are plotted by '+', respectively. In each four figures, the true time series is also plotted by real line. In another three panels [B],[C] and [D], the estimated time series and the true time series of another three voxels are shown in the same way. Note that the scale of time series is different voxel by voxel.


Figure 6.2: Location of voxels employed in the simulation studies and the real data analysis. The position of each voxel is denoted by a dot. The number of voxels is 3433.

where N_p is the number of the parameters. A solution with a smaller value is better.

• Localization Error (LE) (Pascual-Marqui 1999) (the distance between the location of the true source and an estimated source) :

$$LE = ||\vec{r}(v) - \vec{r}(\hat{v})||$$

where $\vec{r}(u)$ is the coordinate of a voxel u. Voxels on which the amplitude of the true and estimated current source vector attains (local) maximum is denoted by v and \hat{v} , respectively.

• Visibility (VI) (Pascual-Marqui 1999) (the ratio of the amplitude of an estimated source to the amplitude of the true source :

$$\mathrm{VI} = \frac{\hat{j}_{\hat{v}}}{j_v}$$

where j_v is the vth component of the true current vector \boldsymbol{j} . A solution with VI close to 1 is preferable.

The results of comparison are summarized in table 6.2, . From this table, the following observations can be pointed out:

- The values of RMSE for the DynLORETA solution are a half of those for LORETA solution.
- The values of RMSE for the DynLORETA solution do not change irrespective of the observation noise level, whereas those for the LORETA solution increase as the observation noise level increases. This weakness of instantaneous inverse solutions has already been noticed by Schmidt et al. 2001.
- As to LE for the occipital source, the DynLORETA solutions sometimes get worse than the LORETA solutions.
- As to LE for the frontal source, the DynLORETA solutions always attain smaller (better) values than the LORETA solutions.
- As to Vi, for both the occipital and frontal sources, the DynLORETA solutions are better than the LORETA solutions. This fact can be observed in Fig.6.3 as the spatial distribution of the current vectors and in Fig.6.4 as the time series of the sources on two voxels.
- In each case, the ABIC value of the DynLORETA solution is smaller than that of ABIC. This indicates the DynLORETA solution is better than the LORETA solution. This result is consistent with the results of another criterion. It should be noted that ABIC is available even without knowing the *TRUE* current vector.

As the advantage of DynLORETA over LORETA, we mention higher visibility, less ghost, the temporal smoothness and robustness for the observation noise. However DynLORETA sometimes results in worse localization if the localization of the initial state estimate is not good.

6.3 Simulation Study II

In order to confirm whether observable projection Kalman filtering works well in large number of voxels, a simulation experiment will be performed. For this purpose a time series of T = 500 observations from a AR(2) model of voxel dynamics, with nearest-neighbor interactions, as described by

$$\boldsymbol{v}_{t} = \mathbf{K}\boldsymbol{j}_{t} + \boldsymbol{\epsilon}_{t}$$

$$\boldsymbol{j}_{t} = (a_{1}\mathbf{I}_{N_{s}} + b_{1}\mathbf{L})\boldsymbol{j}_{t-1} + (a_{2}\mathbf{I}_{N_{s}} + b_{2}\mathbf{L})\boldsymbol{j}_{t-2} + \boldsymbol{\eta}_{t}$$

$$\operatorname{Var}(\boldsymbol{\epsilon}_{t}) = \sigma^{2}\mathbf{I}_{N_{e}} \quad \operatorname{Var}(\boldsymbol{\eta}_{t}) = \tau^{2}\mathbf{I}_{N_{s}}$$

(6.2)



SIMULATED DATA

DYNLORETA ESTIMATION

LORETA ESTIMATION

Figure 6.3: Spatial distributions of the current vectors for the simulation (left panel) and for the inverse solutions obtained by DynLORETA (middle panel), and by LORETA (right panel). The solutions were obtained from the EEG observation at time about 0.08 second with 10% observation noise.



Figure 6.4: Time series of two sources. In the top panel, for the frontal source, the time series of true, estimated by DynLORETA and estimated by LORETA are plotted by circles 'o', thin line and thick line, respectively. In the bottom panel, for the occipital source, the time series of true, estimated by DynLORETA and estimated by LORETA are plotted by circles 'o', thin line and thick line, respectively.

		5%	10%	20%	33%
RMSE	LOR	14.2	14.3	14.6	14.9
	DLOR	7.2	7.5	7.3	7.3
ABIC	LOR	-44604	-43817	-42049	-40164
	DLOR	-59595	-54345	-48770	-44915
LE-O	LOR	6.0	7.6	8.2	9.4
(mm)	DLOR	5.5	9	10.8	8.8
LE-FR	LOR	15.8	15.5	15.2	15.6
(mm)	DLOR	11.3	11.0	11.8	11.9
VI-O	LOR	0.32	0.32	0.3	0.29
	DLOR	1	1.02	1.03	0.99
VI-FR	LOR	0.10	0.10	0.09	0.08
	DLOR	0.51	0.49	0.47	0.50

Table 6.2: Comparison of DynLORETA and LORETA solutions. For each level of the observation noise, the goodness of the LORETA solution and the DynLORETA solution are compared using four criterion - RMSE (Root Mean Square Error), ABIC (Akaike's Bayes Information Criteria), LE (Localization Error) and VI (Visibility). 'O' and 'Fr' denote 'Occipital' and 'Frontal', respectively ('LE-O' denotes localization error of the occipital source).

where the parameters are chosen as $(a_1, a_2, b_1, b_2, \sigma, \tau) = (1.82, -1.00, 0.07, 0.00, 20, 10)$. The initial distribution of the current source is chosen from a LORETA solution of the real data of alpha wave at single time point so as to let this simulation approach a realistic situation (Fig.6.6). The EEG observations $v_1, \ldots v_T$ are shown in Fig.6.5. In the figure a stationary oscillation can be seen around the channel O2 like alpha activity.

Then the following AR(2) model is fitted to the generated observations,

$$\boldsymbol{v}_{t} = \mathbf{K}\boldsymbol{j}_{t} + \boldsymbol{\epsilon}_{t}$$

$$\boldsymbol{j}_{t} = (a_{1}\mathbf{I}_{N_{s}} + b_{1}\mathbf{L})\boldsymbol{j}_{t-1} + (a_{2}\mathbf{I}_{N_{s}} + b_{2}\mathbf{L})\boldsymbol{j}_{t-2} + \boldsymbol{\eta}_{t}$$

$$\operatorname{Var}(\boldsymbol{\epsilon}_{t}) = \sigma^{2}\mathbf{I}_{N_{e}} \quad \operatorname{Var}(\boldsymbol{\eta}_{t}) = \tau^{2}(\mathbf{L}\mathbf{L}')^{-1}.$$

(6.3)

using observable projection Kalman filtering. For starting the filtering algorithm, LORETA solutions from the first two observations $\boldsymbol{v}_1, \boldsymbol{v}_2$ are employed. The parameters are estimated as $(\hat{a}_1, \hat{a}_2, \hat{b}_1, \hat{b}_2, \hat{\sigma}, \hat{\tau}) = (1.82, -1.00, 0.04, 0.05, 19, 104)$ by maximizing the likelihood function Eq.(5.101). Except τ , the estimated parameters are close to the true ones. It should be noted that the system noise covariance matrix of the fitted model and that of the true model is different. Due to misspecification of the system noise covariance structure, the estimate of τ is considered to be far from the true value. In Fig.6.7 for four specific voxels, chosen from the frontal, the right temporal, the left temporal and the occipital regions, the corresponding time series of the length of current vectors (i.e. $|j| = \sqrt{j_x^2 + j_y^2 + j_z^2}$) for the simulation ("truth") and for the obtained inverse solutions are plotted. In each figure we can see that the inverse solutions via observable projection Kalman filtering reproduce the true time series very well though in the beginning of time series, the deviation of the estimated series from the true series is large due to inaccuracy of the initial state estimates. In this simulation study, this good reconstruction can be seen in all voxels.



Figure 6.5: Simulated EEG observations at 19 standard electrode positions of the 10/20-system, according to the model given by Eq.(6.2).



Figure 6.6: Maximum-intensity projections of the distribution of current sources, used as initial value for the simulation: front view (left panel), top view (middle panel) and side view (right panel). Grayscales correspond to the length of the local current vectors. This distribution was obtained from a LORETA solution of the real alpha wave at single time point.



Figure 6.7: Time series of the length of the estimated local current vector at four specific voxels. In each panel, the estimated time series and the true time series are plotted by thick line and thin line, respectively.

6.4 Real Data Analysis

In Fig.6.8 a clinical EEG recording is shown, recorded from a healthy child in awake state with closed eyes. At the occipital electrodes O1 and O2 an oscillation is visible which represents the characteristic alpha rhythm. For the analysis of this data set a regional homogeneous AR(2) model is employed, given by

$$\boldsymbol{Y}_{t} = K\boldsymbol{J}_{t} + \boldsymbol{\epsilon}_{t}$$
$$\boldsymbol{J}_{t}^{v} = \begin{cases} a_{1}\boldsymbol{J}_{t-1}^{v} + a_{2}\boldsymbol{J}_{t-2}^{v} + \boldsymbol{\eta}_{t}^{v} & v \in G \\ b_{1}\boldsymbol{J}_{t-1}^{v} + b_{2}\boldsymbol{J}_{t-2}^{v} + \boldsymbol{\eta}_{t}^{v} & v \notin G \end{cases}$$
$$\operatorname{Var}(\boldsymbol{\epsilon}_{t}) = \sigma^{2}\operatorname{I}_{N_{e}}, \quad \operatorname{Var}(\boldsymbol{\eta}_{t}) = \tau^{2}(\mathrm{L}'\mathrm{L})^{-1}.$$

Here the dynamics within a certain region G is assumed to differ from the dynamics within the remaining part of brain. We have chosen the region G as a sphere of radius 30mm centered within the occipital lobes; the center was chosen according to the result of a LORETA solution of the same data.

Estimation of the dynamical parameters (a_1, a_2, b_1, b_2) by numerical optimization provides the estimates (1.95, -0.99, 1.83, -0.83). From these estimated parameter, the parametric spectrum of the AR model can be calculated (Shumway 2000). In Fig.6.9 the parametric spectra both inside G and out side G are shown. As can be seen, the parametric spectrum inside G displays a clear peak around 8.4Hz, whereas the power spectrum outside G does not display a clear peak, but just a drop of power towards higher frequencies. The peak at 8.3 Hz falls well within the known range for alpha activity. These results illustrate that by assuming a parametric model it is possible to make detailed inference about the dynamics of the sources.

In Fig.6.10, we illustrate for 12 consecutive points of time (with a time shift of 0.012 seconds = every 3 time point) the evolution of the spatial distribution of a component of the estimated current vectors. We choose to display the component of the current vectors which corresponds to the radial direction of spherical coordinates, with the origin being located at the center of the head. The solution provides two main sources oscillating in opposite phase in the left and right occipital region. These two sources can be considered to be generators of alpha rhythm (Valdés-Sosa, Bosch, Grave de Peralta-Menendez, Hernández, Pascual, and Biscay 1992; Rodin and Rodin 1995).

In Fig.6.11, the time series of the estimated local current vector in the radial direction of spherical coordinates at four voxels are shown. In the top figure, two time series chosen from the voxels locating on the center of two sources as shown in Fig.6.10 are plotted. We can see clear oscillation of alpha activity for both two sources. Furthermore, we can see the change of the phase of two sources during this two seconds. In the bottom figure, two time series chosen from (arbitrary) voxels outside G are plotted. These two time series do

not reveal any clear oscillation reflecting the feature of the estimated dynamics outside G.

In Tab.6.3, the goodness of fitting using another models are demonstrated. 'LORETA' is the instantaneous inverse solution with spatial smoothness, 'Random walk(2)' is the dynamical inverse solution with $\mathbf{j}_t = 2\mathbf{j}_{t-1} - \mathbf{j}_{t-2} + \mathbf{\eta}_t$, and 'Regional AR(2)' is the dynamical inverse solution using the model mentioned above in this section. The result indicates that the dynamical inverse solutions are much better than the instantaneous inverse solution and also indicates that the log-likelihood can be improved significantly with the introduction of a few parameters in the 'Regional AR(2)' model compared with the 'Random walk(2)' model.

	-2 log-likelihood	Number of parameters
LORETA	117995	2
Random walk (2)	107880	2
Regional $AR(2)$	100943	6

Table 6.3: Model comparison for the real alpha data



Figure 6.8: Clinical EEG recording at 19 standard positions of the 10/20-system versus time, obtained from a healthy 8-years old male child, awake with closed eyes. The vertical axis represents observed voltages relative to the average reference. In the right panel, the EEG recordings amplified around 3 sec. (marked by a diamond) are demonstrated.



Figure 6.9: Logarithm of the parametric spectrum obtained from the estimated AR(2) dynamics, versus frequency. Top panel shows the parametric spectrum inside region G and bottom panel shows that outside region G. The region G is defined in the text of section 6.3.



Figure 6.10: Spatial distributions of the component of the estimated current vectors in the radial direction of spherical coordinates. These 12 figures correspond to the inverse solutions of the EEG time series in the right panel of Fig.6.8.



Figure 6.11: Time series of the estimated local current vector in the radial direction of spherical coordinates at four voxels. In the top panel, two time series chosen from the voxels locating on the center of two sources as shown in Fig.6.10 are plotted. The time series of the right source and of the left source are plotted with thick line and with thin line, respectively. In the bottom panel, two time series chosen from arbitrary voxels outside G are plotted.

Chapter 7

Discussion

In the part I, the inverse problem of estimating generators of EEG recordings have been addressed with particular emphasis on the use of dynamical constraints. The following issues were discussed:

- The "dynamical inverse problem" of the EEG have been discussed. By formulating the dynamical inverse problem as the state space representation, we can consider a general dynamical constraint in the system equation. In particular, we consider a class of parametric models for the dynamics and choose the best candidate by optimizing the parameters.
- As a parametric model for the spatio-temporal brain dynamics to be used in the simulation study, we have employed a AR(2) model including nearest-neighbor interaction. This particular class of parametric models is expected to be useful for two reasons; firstly, these models can be interpreted as discretizations of partial differential equations describing spatio-temporal dynamical phenomena; secondly, they can be formulated by using highly sparse matrices which renders them appropriate for application to high-dimensional problems.
- In principle, the optimum solution of this state estimation problem is given by Kalman filtering and Kalman smoothing; however, due to the high dimensionality of the state in the EEG inverse problem, the direct application of Kalman filtering is very demanding (or even impossible) in terms of computational time and memory consumption. As alternatives, three approximate algorithms of Kalman filtering were proposed; the recursive penalized least squares (RPLS) method, observable projection Kalman filtering and partitioned (spatio-temporal) Kalman filtering. From their assumptions and properties, the RPLS method is appropriate for a preliminary analysis in virtue of its modest computational time and its easiness of the implementation. Observable projection Kalman filtering is appropriate when the dynamics is specified as a simple model

such as the random walk model. Partition Kalman filtering could be employed when correlation among partitioned states can be ignored.

- In the simulation study of the small number of voxels, the solutions obtained by these three algorithms and Kalman filtering are compared. As a result, the RPLS method and observable projection Kalman filtering provide the considerably accurate solutions with modest computational time. Although Kalman filtering attains the best solution, Kalman filtering is impracticable even in this small number of voxels in terms of computational time. The solution obtained by partitioned Kalman filtering is not so good as the solutions obtained by the RPLS method or observable projection Kalman filtering, because of the inaccurate estimates of the parameters. In this simulation, the superiority of the RPLS method or observable projection Kalman filtering was demonstrated. However it is dangerous to generalize this result, since this simulation is too simple and far from realistic. Further investigation is necessary in the future.
- In the simulation study for comparison of the instantaneous solution (LORETA) and the dynamical inverse solution (DynLORETA), Dyn-LORETA shows superior performance in terms of the visibility, RMSE, and the robustness to the observation noise. While the spatial features of the LORETA solutions are inherited to DynLORETA, additional improvements of the solution become possible through incorporation of temporal information. On the other hand, if the dynamical model has not been well-chosen, the solutions of DynLORETA tend to be very similar to the corresponding LORETA solutions, because inappropriate dynamical constraints result in very weak regularization. In addition the better estimate of the initial state for starting RPLS is an important point for substantial improvements of whole estimates.
- In an analysis of clinical EEG data we have employed a regional AR(2) model, characterized by the presence of different dynamics inside and outside the occipital area. As a result of observation projection Kalman filtering we have observed two occipital sources which are opposite in phase during some interval. Both from the parameter estimates and from the estimated time series at occipital voxels these two sources can be consider to be the generator of the alpha wave.

In the future the ideas and methods presented in this thesis should be developed further:

• Information from other brain-imaging modalities (such as fMRI, NIRS) should be incorporated into the model identification. This will render it possible to explore physiologically more

meaningful dynamics and ultimately also result in better inverse solutions.

- More realistic situation such as the non-linear dynamics and misspecification of the model should be taken into consideration in the future simulation.
- In order to reduce the amount of computation as well as stabilizing the computation, it is desirable to develop the method for the dimension reduction of the state.
- The theoretical aspect of three proposed algorithms should be further investigated. By this, it would be possible to develop more efficient algorithm in time or accuracy. Furthermore filtering algorithms developed in another research field such as in the data assimilation (Evensen 1994; Evensen 1997; Evensen 2003) will be examined and incorporated into the EEG inverse problem.

Part II

Causality Analysis of fMRI Data

Chapter 8

Introduction

8.1 Context

The brain appears to adhere to two fundamental principles of functional organization, **functional specialization** and **functional integration**. The distinction relates to that between *localisationism* and *connectionism* that dominated thinking about cortical function in the nineteenth century. Functional specialization suggests that a cortical area is specialized for some aspects of perceptual or motor processing. The cortical infrastructure supporting a single function may then involve many specialized areas whose union is mediated by the functional integration among them.

In early 1990s, the outstanding methodology of measuring activities of human brain, which is called functional Magnetic Resonance Imaging (**fMRI**) technique, has been developed. This technique enables us to get the information of the activity inside human brains noninvasively. In virtue of this advantage, various cognitive experiments (for example, visual, auditory and motor) of human brains have been done largely in order to clarify the relationship between functions and related areas.

The fMRI technique provides temporally successive images reflecting the change of regional cerebral blood flow, which is believed to result from electrical neuronal activities on the corresponding local area, with high spatial resolution but rather low temporal resolution. It seems that so far in fMRI studies the emphasis has been put on the statistical analysis of how to localize cortical areas related to a specific cognitive experiment (specialization) (Friston et al. 1995; Büchel et al. 1996; Worsley et al. 2002). Obviously it is more important and challenging work for deeper understanding of human brains to develop the statistical analysis of how to connect these localized areas (integration).

Integration within a distributed system is well understood in terms of 'effective connectivity'. Effective connectivity is defined as "the influence that one neural system exerts over another" (Friston et al. 1995). For the purpose of evaluating the effective connectivity, several statistical methods have been applied; structural equation model (McIntosh and Gonzalez-Lima 1994), regression based analysis (Friston et al. 1995) and so on. Although these methods work very well for specifying magnitude of the influence one exerts another, they can work only in a highly constrained way, typically exploring the connections between a network of brain areas whose anatomy is explicitly defined by a model (i.e. the direction of the connectivity needs to be prespecified).

This disadvantage, however, can be overcome by **causal analysis** which has been developed in the field of **time series analysis**. The pioneering work about causality has been done in the literature of Granger 1969, in which the comprehensive definition of causality based on multivariate time series has been discussed. His definition is entirely based on 'the predictability of some series for the others', and here the flow of time plays a central role. Following the work of Granger, various measures and tests of causality has been proposed (Geweke 1982; Geweke 1984; Hosoya 1991; Kaminski et al. 2001). Apart from these work, Akaike has proposed his measure, Akaike's Noise Contribution Ratio (ANCR), in order to characterize and analyze feedback systems (Akaike 1968). His measure is based on the fact that the powerspectrum of a time series from multivariate AR (MAR) model can be decomposed to the sum of power of the other series and itself. A crucial advantage of ANCR is that it is possible to evaluate causal relations between many time series (more than three) whereas the measures followed by the work of Granger can be applied to essentially bivariate or block bivariate time series.

Because of low temporal resolution of the fMRI, there have not been so many works emphasizing times series analysis except very recent works, for example, an application of a multivariate time series model has been proposed for determining only directions of the connectivity (Harrison et al. 2003). In this thesis I would like to propose to apply ANCR as an extension of Harrison's method so that both direction and magnitude of the effective connectivity can be evaluated.

The ANCR will be applied to the data from the random dot experiment. The 'change' of the connectivity on the task condition compared with that on the control condition will be evaluated using the feature of ANCR that the causal relations can be measured for each periodic component. The result shows the possibility of this analysis. It should be noted that for this connectivity analysis, the data was acquired in higher sampling rate of 1 sec. than in typical sampling rate of 3 sec.

8.2 Organization of Part II

In Chap.9, previous works of the causality analysis in time series field are reviewed briefly. The definition of Granger causality and the derivation of Akaike's noise contribution ration (ANCR) are described here. In Chap.10, the whole procedure for evaluating the change of the effective connectivity is explained in detail. Chap.11 presents the results for the data from the random dot experiment and the discussion about the results is followed.

Chapter 9

Causality analysis in time series

The term, **effective connectivity** has been defined as "the influence that one neural system exerts over another" (Friston et al. 1995) or "the direct effect one region has on another" (McIntosh 2000). In order to evaluate the effective connectivity, we have to determine the direction and the magnitude of the influence (effect). In fMRI studies, the structural equation model (SEM), where the direction of the influence needs to be determined apriori, has mainly been employed for this purpose. The crucial information of the direction, however, can be provided by the **causality analysis** on multivariate time series data.

In this chapter, the important works about the causality analysis are reviewed. At first the notion of the causality in time series is presented according to the paper of Granger 1969. Secondly the main tool for evaluating the effective connectivity in this thesis, which is referred to as Akaike's noise contribution ratio (ANCR) (Akaike 1968), is introduced.

9.1 Granger Causality

It could be the acceptable statement "the cause precedes the results" for everyone. Based on this statement, Granger has explicitly defined the notion of the "causality" in time series analysis.

Let's define the following notations at first.

- A_t : a a stationary stochastic process
- \overline{A}_t : the set of *past* values $\{A_{t-1}, A_{t-2}, \cdots\}$.
- $P_t(A|B)$: the optimum, unbiased, least-squares predictor of A_t using the set of values B_t .
- $\epsilon_t(A|B)$: the predictive error series, that is, $A_t P_t(A|B)$.
- $\sigma^2(A|B)$: the variance of $\epsilon_t(A|B)$.

Furthermore let U_t be all the information in the universe and $U_t - Y_t$ be all this information apart from the specified series Y_t . Then He has defined the causality as,

Definition 1 (Causality) If $\sigma^2(X|\overline{U}) < \sigma^2(X|\overline{U-Y})$, we say that Y is causing X, denoted by $Y_t \Rightarrow X_t$. We say that Y_t is causing X_t if we are better able to predict X_t using all available information than the information apart from Y_t had been used.

In practical, the unrealistic availability of the universal information U is replaced by a set of related timeseries in an application.

9.2 Akaike's Noise Contribution Ratio

This ratio was developed to characterize the feedback system consisting of several related time series (Akaike and Nakagawa 1988). The linear feedback can be modeled by the following Multivariate AutoRegressive (MAR),

$$\boldsymbol{Z}_{t} = \sum_{i=1}^{p} A(i) \boldsymbol{Z}_{t-i} + \boldsymbol{\epsilon}_{t}$$
(9.1)

where $\boldsymbol{\epsilon}_t$ is a white noise process with the distribution $N(0, C_{\boldsymbol{\epsilon}})$, called an "innovation" process. The $d \times 1$ vector, $\boldsymbol{Z}_t = (z_{1,t}, \cdots, z_{d,t})'$ is comprised of the ensemble of d time series. Each time series in the system is driven by the innovation $\boldsymbol{\epsilon}_t$ $(t = 1, 2, \cdots)$ and is evolving through the $d \times d$ coefficient matrices, A(i)s.

Through this section, we assume the "causality condition" for the MAR model, which can be checked by examining whether all solutions of the characteristic function $c(\lambda)$,

$$c(\lambda) = \det\left(I - \sum_{i=1}^{p} A(i)\lambda^{i}\right) = 0$$

are located outside the unit circle. This condition guarantees that the present data can be described only with the terms of the past innovations, but not influenced by the future innovations.

9.2.1 ANCR in frequency domain

ANCR has originally been developed so as to measure the contribution power of another time series on the power of one time series in frequency domain. ANCR can be computed from the parametric spectrum of MAR model. The parametric spectrum of the MAR model (9.1) can be calculated from the AR coefficients A(i), $(i = 1, \dots, p)$ and the innovation covariance matrix C_{ϵ} as follows,

$$P(f) = H(f)C_{\epsilon}\overline{H}(f)', \qquad (9.2)$$

where f denotes the frequency ranging from 0 to 0.5 and P(f) denotes the power spectrum matrix, where the diagonal entry $P_{ii}(f)$ and the non-diagonal entry $P_{ij}(f)$ represent the power spectrum of the *i*th time series $z_{i,t}$ and the cross spectrum between $z_{i,t}$ and $z_{j,t}$, respectively. The matrix H(f) describing the transfer functions (frequency response) from a set of the innovation series ϵ_t to a set of the time series \mathbf{Z}_t is given by

$$H(f) = \left(I - \sum_{j=1}^{p} A(j) \exp\{-i2\pi f j\}\right)^{-1},$$
(9.3)

where $i = \sqrt{-1}$.

Here we further assume that the innovations are uncorrelated each other, that is, the covariance matrix C_{ϵ} is diagonal,

$$C_{\epsilon} = \operatorname{diag}(\sigma_1^2, \cdots, \sigma_d^2). \tag{9.4}$$

Then we obtain the following equations by directly calculating Eq.(9.2),

$$P_{ii}(f) = \sum_{j=1}^{d} |H_{ij}(f)|^2 \sigma_j^2, \quad i = 1, \cdots, d.$$
(9.5)

This equation shows that the power spectrum of $z_{i,t}$ of the frequency components f can be decomposed to d terms $|H_{ij}(f)|^2 \sigma_j^2$, $j = 1, \dots, d$, each of which can be interpreted as the power of the jth innovation $\epsilon_{j,t}$ transferring to $z_{i,t}$ via the transfer function $H_{ij}(f)$. Thus $|H_{ij}(f)|^2 \sigma_j^2$ can be regarded as the power contribution of the innovation $\epsilon_{j,t}$ on the power spectrum of $z_{i,t}$. Finally ANCR is defined as a ratio of each contribution to the power spectrum $P_{ii}(f)$,

$$R_{j \to i}(f) = \frac{|H_{ij}(f)|^2 \sigma_j^2}{P_{ii}(f)}.$$
(9.6)

ANCR is computed for every combination of i and j and for every frequency f, therefore the contribution can be evaluated for each periodic component.

I would like to remark that since the diagonal innovation covariance matrix is a crucial assumption in the derivation, the diagonality of an estimated innovation covariance matrix should be checked in the application of this method. A further remark is that the extension of ANCR has been proposed so that the contribution can be measured without the assumption of a diagonal innovation covariance matrix (Tanokura and Kitagawa 2003).

9.2.2 ANCR in time domain

The amount of the information containing in ANCR as have defined in Eq.(9.6) is huge and the result is sometimes sensitive to the estimates. These fact cause the difficulty of interpreting the result, especially in the case that the comparison of two results of ANCR is necessary as in this study. Motivated by the need of some summary of ANCR, ANCR in time domain will be introduced here. From the author's limited knowledge, there is not any report about this issue, though the introduction of this measure is straightforward.

Under the causality condition, the MAR model (9.1) can be represented by the infinite sum of the past innovations as

$$\boldsymbol{Z}_{t} = \sum_{s=0}^{\infty} M(s)\boldsymbol{\epsilon}_{t-s}$$
(9.7)

where M(s), so-called the impulse response function, can be calculated from

$$M(s) = \sum_{k=0}^{\min(p,s)} A(k)M(s-k)$$
(9.8)

with M(0) = I (Brockwell and Davis 1996). Taking variance of both sides of Eq.(9.7) gives

$$\operatorname{Var}(\boldsymbol{Z}_t) = \sum_{s=0}^{\infty} M(s) C_{\boldsymbol{\epsilon}} M'(s).$$
(9.9)

In the same way as the definition of ANCR in frequency domain, the variance of the time series $z_{i,t}$, which is the *i*th diagonal element of $Var(\mathbf{Z}_t)$, can be represented by the sum of each innovation contribution as,

$$S_i^2 = \sum_{j=1}^d \left(\sigma_j^2 \sum_{s=0}^\infty M_{ij}^2(s) \right)$$
(9.10)

where S_i^2 is the variance of $z_{i,t}$ and $M_{i,j}(s)$ is the *ij*th element of M(s). Here the innovation covariance C_{ϵ} is assumed to be diagonal again.

Then, ANCR in time domain can be defined as follows,

$$r_{i \to j} = \frac{\sigma_j^2 \sum_{s=0}^{\infty} M_{ij}^2(s)}{S_i^2}.$$
(9.11)

In practical, the numerical computation of M(s), $s = 0, 1, \cdots$ is necessary for the evaluation of ANCR Eq.(9.11). The summation of M(s) in Eq.(9.11), hence, is truncated up to a certain finite number. When any characteristic roots of $c(\lambda)$ are not located near the unit circle, the truncation of a moderate number causes almost no numerical error.

The relationship between ANCR in frequency domain and ANCR in time domain can be seen from the spectral representation of the autocovariance function as provided by,

$$\Gamma(h) = \int_{-0.5}^{0.5} P(f) e^{i2\pi f h} df.$$
(9.12)

Here $\Gamma(h)$ is the autocovariance function of \mathbf{Z}_t .

By setting h = 0 and substituting the representation of the parametric spectrum (9.2) in Eq.(9.12), the following equation is obtained:

$$\operatorname{Var}(\boldsymbol{Z}_t) = \int_{-0.5}^{0.5} H(f) C_{\boldsymbol{\epsilon}} \overline{H}(f)' df.$$
(9.13)

Comparing the diagonal element in both sides, we obtain the equation as follows:

$$S_i^2 = \sum_{j=1}^d \left(\sigma_j^2 \int_{-0.5}^{0.5} |H_{ij}(f)|^2 df \right)$$
(9.14)

From this equation and Eq.(9.10), ANCR in time domain can be interpreted as the averaged ratio over all the frequency components, whereas ANCR in frequency domain provides the contribution ratios for each frequency.

Chapter 10

Data analysis

Firstly the experiment and the aim of data analysis are shown and then the entire procedure to accomplish this aim is explained.

10.1 Experiment

In this thesis, we employ the connectivity analysis for the data acquired from a simple random dot experiment. In this experiment, a screen is set above a face of a subject lying in the MRI scanner. During the data acquisition, the screen displays two different objects corresponding to two conditions of the experiment as follows:

- Task condition : many random dots are moving radial.
- Control condition : only a red point is fixed on the center.

The timing of the task condition of the data acquisition is shown in Fig.10.1. Two sessions have been executed: the first session is constituted of the repetition of the task and control condition of duration 30 sec. and the second session is constituted by the continuous control condition of duration 300 sec.

The data, with temporal resoluti on of 1 sec., has been acquired using T2^{*}-weighted, gradient echo, echo planer imaging sequences with the 3 Tesra MRI scanner (Allegra; Simens, Erlangen, Germmany).

10.2 Aim

In previous phygiological experiments, it is well confirmed that three regions, called the primary visual cortex (V1), visual motion detection area (V5) and the posterior parietal cortex (PP) for integration of spatial visual information, are associated with perception of visual motion.



Figure 10.1: Block design of the random dot experiment.

The aim of this study is to evaluate the 'change' of the connectivity among these regions on the task and control condition. For this purpose, two strategies are taken in this thesis.

- A1. By fitting the MAR model to the data of the first session, the connectivity on the task condition can be evaluated from ANCR at the frequency f_0 which corresponds to the period of the task condition. In contrast, by fitting the MAR model to the data of the second session, ANCR at the frequency f_0 may not reflect any task related behavior.
- A2. Fit the multiplicative exogenous-dependent multivariate AR (MEMAR) model for the data only in the first session as follows,

$$\boldsymbol{Z}_{t} = \sum_{i=1}^{p} (A(i) + B(i)s_{t-d})\boldsymbol{Z}_{t-i} + \boldsymbol{\epsilon}_{t}$$
(10.1)

where s_t models the visual stimulus ($s_t = 1$ when task condition, $s_t = 0$ otherwise). From $A(1), \dots, A(p)$ and C_{ϵ} , ANCR on the control condition is computed, whereas from $A(1) + B(1), \dots, A(p) + B(p)$ and C_{ϵ} , ANCR on the task condition is computed.

10.3 Procedure

The procedure for evaluating the effective connectivity is outlined is as follows:

(i) determine the regions of interest (ROIs) and pick up the time series from each region.

- (ii) fit the model for the time series obtained above.
- (iii) evaluate the effective connectivity among ROIs using ANCR as defined in the previous section.

In the following subsections, the detailed issues in each step of the procedure will be explained.

10.3.1 Specify ROIs

The fMRI data provides huge dimensional time series, whose dimension corresponds to the number of voxels in the brain (typically one hundred thousand of voxels). Before fitting the time series model, a set of time series must be selected from the regions of interest related to an experiment.

In this study, the software 'SPM' (http://www.fil.ion.ucl.ac.uk/spm/) is used for this purpose, because this software is the most common in the fMRI community and this software is fully equipped with well-established image processing tools such as for the realignment of head movement and for the registration to the standard brain.

The following preprocess have been done:

- P1. "Slice timing" for adjusting different sampling timing in the acquisition of each slice.
- P2. "Realignment" for adjusting the movement of the head.
- P3. "Normalizing" for the registration to the standard brain.

Then, the ROIs are determined by the following statistical procedure,

- S1. Making the t-map by fitting the general linear model voxel by voxel (Friston et al. 1995; Worsley et al. 2002).
- S2. Selecting the regions so that each region consists of voxels within a sphere of radius 6mm, where the center voxels are chosen as the local maximizers of t-values.
- S3. Pick up the representative time series from each region specified in the step S2. The first principle component of all the time series within each region has been chosen as the representative.

The specified regions are shown in Fig.10.3 and a set of time series chosen are shown in top three panels in Figs.11.1 and 11.2 for the first and the second session, respectively.

10.3.2 Model fitting of time series in ROIs

After choosing the time series from ROIs, the multivariate time series model is fitted for the evaluation of the effective connectivity. The detailed issues for estimating the parameters in the MEMAR model (10.1) are discussed here. From now we assume that the distribution of innovation follows the normal distribution.

Maximum Likelihood Estimation

Let's define Θ as the collection of the AR coefficients as in (10.4). The loglikelihood is represented as,

$$\mathcal{L}(\boldsymbol{\Theta}, C_{\boldsymbol{\epsilon}}) = \log p(\boldsymbol{Z}_{1}, \cdots, \boldsymbol{Z}_{T}; \boldsymbol{\Theta}, C_{\boldsymbol{\epsilon}})$$

$$= \sum_{t=p+1}^{T} \log p(\boldsymbol{Z}_{t} | \boldsymbol{Z}_{t-1}, \cdots, \boldsymbol{Z}_{t-p}; \boldsymbol{\Theta}, C_{\boldsymbol{\epsilon}}) + \log p(\boldsymbol{Z}_{1}, \cdots, \boldsymbol{Z}_{p}; \boldsymbol{\Theta}, C_{\boldsymbol{\epsilon}}).$$
(10.2)

The conditional distribution in Eq.(10.2) is given by

$$p(\boldsymbol{Z}_t | \boldsymbol{Z}_{t-1}, \cdots, \boldsymbol{Z}_{t-p}; \boldsymbol{\Theta}, C_{\boldsymbol{\epsilon}}) \sim N(\bar{\boldsymbol{Z}}_t, C_{\boldsymbol{\epsilon}}),$$
 (10.3)

where

$$\bar{\boldsymbol{Z}}_{t} = \sum_{i=1}^{p} (A(i) + B(i)s_{t-d})\boldsymbol{Z}_{t-i} \equiv \boldsymbol{\Theta}' \boldsymbol{\mathcal{X}}_{t}$$

$$\boldsymbol{\mathcal{X}}_{t} \equiv (\boldsymbol{Z}'_{t-1}, \cdots, \boldsymbol{Z}'_{t-p}, s_{t-d}\boldsymbol{Z}'_{t-1}, \cdots, s_{t-d}\boldsymbol{Z}'_{t-p})'$$

$$\boldsymbol{\Theta}' \equiv (A(1), \cdots, A(p), B(1), \cdots, B(p)). \quad (10.4)$$

The log-likelihood (10.2) can be rewritten in the form,

$$\mathcal{L}(\boldsymbol{\Theta}, C_{\boldsymbol{\epsilon}}) = C + \frac{(T-p)}{2} \log |C_{\boldsymbol{\epsilon}}^{-1}| - \frac{1}{2} \sum_{t=p+1}^{T} (\boldsymbol{Z}_t - \boldsymbol{\Theta}' \boldsymbol{\mathcal{X}}_t)' C_{\boldsymbol{\epsilon}}^{-1} (\boldsymbol{Z}_t - \boldsymbol{\Theta}' \boldsymbol{\mathcal{X}}_t)$$
(10.5)

where C is a constant term and the term $\log p(\mathbf{Z}_1, \dots, \mathbf{Z}_p; \mathbf{\Theta}, C_{\epsilon})$ has been neglected (when T is large, this term is small compared with the second and third terms in Eq.(10.5)). The maximum likelihood (ML) estimator can be obtained as the maximizer of $\mathcal{L}(\mathbf{\Theta}, C_{\epsilon})$, which is resulting in

$$\hat{\boldsymbol{\Theta}}' = \left[\sum_{t=p+1}^{T} \boldsymbol{Z}_{t} \boldsymbol{\mathcal{X}}_{t}'\right] \left[\sum_{t=p+1}^{T} \boldsymbol{\mathcal{X}}_{t} \boldsymbol{\mathcal{X}}_{t}'\right]^{-1}$$
(10.6)

$$\hat{C}_{\boldsymbol{\epsilon}} = \frac{1}{T-p} \sum_{t=p+1}^{T} \hat{\boldsymbol{\epsilon}}_t \hat{\boldsymbol{\epsilon}}_t'$$
(10.7)

where the innovation estimate is given by

$$\hat{\boldsymbol{\epsilon}}_t = \boldsymbol{Z}_t - \hat{\boldsymbol{\Theta}}' \boldsymbol{\mathcal{X}}_t. \tag{10.8}$$

The *j*th row of $\hat{\Theta}'$ is

$$\hat{\boldsymbol{\theta}}'_{j} = \left[\sum_{t=p+1}^{T} z_{j,t} \mathcal{X}'_{t}\right] \left[\sum_{t=p+1}^{T} \mathcal{X}_{t} \mathcal{X}'_{t}\right]^{-1}$$
(10.9)

where $\boldsymbol{\theta}_j = (A_{j1}(1), \cdots, A_{jd}(1), \cdots, A_{jd}(p), \cdots, B_{jd}(p))'$ is a vector of size $2pd \times 1$. The variance of $\hat{\boldsymbol{\theta}}_j$ is given by

$$\operatorname{Var}(\hat{\boldsymbol{\theta}}_{j}) = \hat{\sigma}_{j}^{2} \Big[\sum_{t=p+1}^{T} \mathcal{X}_{t} \mathcal{X}_{t}' \Big]^{-1}, \qquad (10.10)$$

where $\hat{\sigma}_j^2$ is the *j*th diagonal element of \hat{C}_{ϵ} . Note that the ML estimators of the coefficients (10.6) is equivalent to the ordinary least squares estimators (see the chapter 11 of Hamilton 1994 for detail).

Selection of p and d

The parameters, p and d are preferably determined by some information criteria such as AIC (Akaike 1973), SIC (Schwarz 1978), AICc (Hurvich and Tsai 1989) and so on. In this study we choose to employ AIC because of its simple form and wide use of this criterion in various applications. AIC is defined as bellows:

$$AIC = -2\mathcal{L}(\hat{\Theta}, \hat{C}_{\epsilon}) + 2N_p \tag{10.11}$$

where N_p is the number of parameters in the model for fitting. AIC is the asymptotically unbiased estimates of Kullback-Leibler (KL) discrepancy between the true model and the fitted model. Therefore the minimization of these criteria corresponds to choosing the nearest model from the true model in the sense of KL discrepancy. The estimators of p and d are determined by calculating AIC for each p and d in some range and by choosing the minimizers.

T-value truncation

As has shown in Chap.9, rather complex computation is necessary for calculating ANCR. If we employ the estimated parameters for this purpose, ANCR may give very sensitive result due to the estimation error. In order to obtain more reliable result, the *t*-value truncation technique is proposed to employ. In this technique, roughly speaking, we regard AR coefficients of which estimates are approximately 0 as exact 0. This can be done by judging whether the t-value of each estimator is less than the pre-defined threshold.

Here this technique is not employed for each element of Θ_{ij} , but for a collection of the coefficients describing the causality from $z_{j,t}$ to $z_{i,t}$, that is, $\hat{a}_{ij} \equiv (\hat{A}_{ij}(1), \dots, \hat{A}_{ij}(p))'$ for $i, j = 1, \dots, d$, because the estimates of these coefficients are usually highly correlated. Note that *T*-value truncation to the collection of these coefficients corresponds to the test of Granger's non-causality from $z_{j,t}$ to $z_{i,t}$.

Let us define an indicator matrix C_{ij}^a consisting of 0 or 1 entries so that

$$\hat{\boldsymbol{a}}_{ij} = C^a_{ij} \hat{\boldsymbol{\theta}}_i \tag{10.12}$$

is satisfied. From Eq.(10.10), the variance of \hat{a}_{ij} can be obtained as

$$\hat{V}_{ij} = C^a_{ij} \operatorname{Var}(\hat{\boldsymbol{\theta}}_j) C^{a\prime}_{ij}.$$
(10.13)

Since the distribution of the statistics $\hat{a}'_{ij}\hat{V}^{-1}_{ij}\hat{a}_{ij}$ is according to χ^2_d , where d is the freedom, we judge that the vector \hat{a}_{ij} is large from $(0, \dots, 0)'$ with significance level α when

$$\hat{a}'_{ij}V_{ij}^{-1}\hat{a}_{ij} > c_{1-\alpha}^{\chi_d^2}$$
(10.14)

is satisfied, where $c_{1-\alpha}^{\chi_d^2}$ is the upper $\alpha\%$ point of the χ_d^2 distribution. In case of Eq.(10.14) being satisfied, the estimators $\hat{\boldsymbol{a}}_{ij}$ are kept, otherwise $\hat{\boldsymbol{a}}_{ij}$ is set to $(0, \dots, 0)'$. For every $i, j = 1, \dots, d$, Eq.(10.14) is examined.

This technique could provide more robust result. Furthermore Eq.(10.14) is essentially equivalent to the Wald test under the null hypothesis $H_0: \mathbf{a}_{ij} = (0, \dots, 0)'$. Because it is difficult to constitute the test on the ratios, this test may help to judge the significance of the connectivity. In this study, the level $\alpha = 0.01$ is used.



Figure 10.2: Schematic figure of Eq.(10.14). The ellipse represents a contour of $1 - \alpha\%$. If the origin is located in this ellipse, \hat{a}_{ij} is set to $(0, \dots, 0)$.



Figure 10.3: Result of t-map (top) and the locations of three ROIs (V1, V5, PP) (bottom) using the software "SPM".
Chapter 11

Result

The results of the connectivity analysis among three ROIs (V1, V5 and PP) are shown in this chapter. As have mentioned in Sec.10.2, for evaluating the change of the effective connectivity, two following strategies are taken and compared.

- A1. By fitting the MAR model to the data of the first session, the connectivity on the task condition can be evaluated from ANCR at the frequency f_0 which corresponds to the period of the task condition. In contrast, by fitting the MAR model to the data of the second session, ANCR at the frequency f_0 may not reflect any task related behaviors.
- A2. Fit the multiplicative exogenous-dependent multivariate AR model for the data only in the first session as follows,

$$\boldsymbol{Z}_{t} = \sum_{i=1}^{p} (A(i) + B(i)s_{t-d})\boldsymbol{Z}_{t-i} + \boldsymbol{\epsilon}_{t}$$

where s_t models the visual stimulus ($s_t = 1$ when task condition, $s_t = 0$ otherwise). From $A(1), \dots, A(p)$ and C_{ϵ} , ANCR on the control condition is computed, whereas from $A(1) + B(1), \dots, A(p) + B(p)$ and C_{ϵ} , ANCR on the task condition is computed.

The results of the analysis [A1] and [A2] will be shown in Sec.11.1 and 11.2, respectively and the discussion of the results will be followed.

11.1 Effective connectivity via MAR model

The time series of the first session and the second session from three ROIs are shown on three top panels in Figs.11.1 and 11.2, respectively. The MAR model of order p = 6, which is determined by AIC, is fitted to the time series of each session. The time series, the histogram and the autocorrelation function of the innovations are also shown on another panels in the same figures. In addition, the sample correlation of the innovations can be calculated as

$$\begin{bmatrix} 1 & 0.27 & -0.01 \\ 0.27 & 1 & 0.17 \\ -0.01 & 0.17 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0.00 & -0.18 \\ 0.00 & 1 & 0.16 \\ -0.18 & 0.16 & 1 \end{bmatrix},$$

for the first and the second sessions. From the diagnosis of the innovations, the MAR model can fit to the data in both two sessions reasonably well and the assumption for calculating ANCR, Eq.(9.4), is acceptable. Therefore the evaluation of the effective connectivity via ANCR could be applicable.

In Fig.11.3, the parametric spectrum and ANCR calculated from the estimated model is shown. The top three figures illustrate the power spectrum of three ROIs. On the task condition, the power spectrum of V1 has a clear peak at the frequency $f_0 = 0.018$ HZ, which almost corresponds to the period of experiment 60 sec., whereas the power spectrum of PP has a clear peak at the frequency $f_1 = 0.031$ HZ, which is a harmonic component of f_0 (two vertical dotted lines denotes the frequencies f_0, f_1). We can observe that at the frequency f_0 , the strength of the connectivity in bottom-up direction is increasing on the task condition, whereas that in top-down direction is increasing at the higher frequency f_1 .

The same result of ANCR at the frequencies f_0 , f_1 and in time domain is represented as the bar graphs in Fig.11.4. It is easily observed that at the frequency f_0 the area occupied by V1 on the task condition is larger than that on the continuous control condition, which means the bottom-up connectivity is strengthened. In contrast, at the frequency f_1 the area occupied by PP on the task condition is increasing, which means the top-down connectivity is strengthened. ANCR in time domain reflects the above-mentioned observations (i.e. a part of both the bottom-up and top-down connectivity increases), because this can be regarded as averaged ANCR over all the frequencies. In Tab.11.1, the differences of the ratios between two conditions are summarized. Only values larger than 0.1 are shown in the table.

	f_0	f_1	time	
$V1 \rightarrow V5$	0.65	-	0.35	
$V1 \rightarrow PP$	0.2	-	-	
$V5 \rightarrow PP$	0.15	-	-	
PP→V1	-	0.3	-	
$PP \rightarrow V5$	-	0.55	0.2	
$V5 \rightarrow V1$	-	-	-	

Table 11.1: Change of ANCR via the MAR model.

11.2 Effective connectivity via MEMAR model

The time series of the first session are only employed for this analysis [A2], and are shown on three top panels in Fig.11.5. Then the MEMAR model of order p = 6 and d = 6, determined by AIC, is fitted. The time series, the histogram and the autocorrelation function of the innovations are also shown on another panels in the same figure. In addition, the sample correlation of the innovations can be calculated as

$$\left[\begin{array}{rrrr} 1 & 0.28 & 0.00 \\ 0.28 & 1 & 0.17 \\ 0.00 & 0.17 & 1 \end{array}\right].$$

From the diagnosis of the innovations, the model fitting is reasonably good and the assumption for calculating ANCR, Eq.(9.4), is acceptable. Therefore the evaluation of the effective connectivity via ANCR could be applicable.

In Fig.11.6, the parametric spectrum and ANCR calculated from the estimated model is shown. The top three figures illustrate the power spectrum of three ROIs. Similar to the previous analysis [A1], on the task condition, the power spectrum of V1 has a clear peak at the frequency $f_0 = 0.017$ HZ, whereas the power spectrum of PP has a clear peak at the frequency $f_1 = 0.033$ HZ (two vertical dotted lines denotes the frequencies f_0, f_1). We can observe that at the frequency f_0 , the strength of the connectivity in bottom-up direction is increasing on the task condition, whereas almost no change can be observed at the higher harmonic frequency f_1 .

The same result of ANCR at the frequencies f_0 , f_1 and in time domain is represented as the bar graphs in Fig.11.4. It is easily observed that at the frequency f_0 the area occupied by V1 on the task condition is larger than that on the control condition, which means the bottom-up connectivity is strengthened. ANCR in time domain shows the complete dominance of V1 on all the ROIs on the task condition, which can not be observed in ANCR of the frequencies f_0, f_1 . This may be due to the numerical error in computation or indicates that ANCR in time domain may not be an appropriate measure in some cases. In Tab.11.2, the differences of the ratios between two conditions are summarized. Only values larger than 0.1 are shown in the table.

	f_0	f_1	time	
$V1 \rightarrow V5$	0.4	-	0.7	
$V1 \rightarrow PP$	0.3	-	0.8	
$V5 \rightarrow PP$	-	-	-0.15	
PP→V1	-	-	-	
$PP \rightarrow V5$	-	-	-0.2	
$V5 \rightarrow V1$	-	-	-	

 Table 11.2: Change of ANCR via the MEMAR model.

11.3 Interpretation

As to the change of the effective connectivity around the same periodic component as the visual stimulus, both the analyses [A1] and [A2] demonstrate similar increase of the bottom-up connectivity on the task condition. However there is some discrepancy in the interpretation.

The MEMAR model fitting in [A2] can be considered as separately fitting a set of two different AR models to the time series on the task condition and on the control condition, respectively. Therefore each AR model should not describe the behavior of the whole series such as the peak of the slow component f_0 , but should describe the behavior of rather fast component. In this sense, it is very difficult to interpret the meaning of the peak f_0 of the power spectrum, consequently, the meaning of ANCR at f_0 .

On the other hand, in [A1], the AR model is fitted to the whole time series of the first session. As expected the parametric power spectrum shows a clear peak at the frequency f_0 same as the designed stimulus period and ANCR at this frequency can be interpreted as the task related connectivity. As a contrast to the task related connectivity, ANCR at the same frequency f_0 has been calculated by fitting the AR model to the whole time series of the second session when no stimulus. Thus the difference between ANCRs at the frequency f_0 of the first session and of the second session can be interpreted as the change of the connectivity caused by the task. This interpretation may be understandable if we imagine the electromagnetic wave of the radio broadcast which conveys the information using a certain frequency band. In Tab.11.3, for 5 subjects, the change of the effective connectivity, is summarized using the difference of ANCR at the frequency f_0 obtained by the analysis [A1] as a measure. It can be observed that the bottom-up connectivity tend to increase on the task condition. It should be noted that ANCR for the first session shows the similarity among the subjects, whereas ANCR for the second session differs subject by subject very much.

		subjecti	subject2	subjects	subject4	subjecto
	$V1 \rightarrow V5$	0.61	0.55	0.13	0.68	-
	$V1 {\rightarrow} PP$	0.27	0.53	0.40	-	-
	$V5 \rightarrow PP$	0.15	-0.2	-	-	0.29
	$PP \rightarrow V1$	-	-	-	-	-
	$PP \rightarrow V5$	-	-	0.12	-0.38	-
	$V5 \rightarrow V1$	-	0.12	-	-	-0.10

Table 11.3: Change of the effective connectivity at the frequency f_0 via the MAR model. subject 1 subject 2 subject 3 subject 4 subject 5



Figure 11.1: Time series from three regions and their innovations by MAR model fitting in the task condition. In the top row, the time series extracted from three regions (V1, V5, PP) are shown. Time series, histogram and the autocorrelation function of the innovations (1-step prediction error) obtained by MAR model fitting are shown from the second to the bottom row, respectively.



Figure 11.2: Time series from three regions and their innovations by MAR model fitting in the continuous control condition. In the top row, the time series extracted from three regions (V1, V5, PP) are shown. Time series, histogram and the autocorrelation function of the innovations (1-step prediction error) obtained by MAR model fitting are shown from the second to the bottom row, respectively.



Figure 11.3: Power spectra of three regions of interest (V1, V5, PP) and ANCR among those regions on the task condition (thick line) and the continuous control condition (thin line). The parametric power spectra of V1, V5 and PP are plotted from the left to right in the top row, respectively. The graphs in three another rows show ANCR for all combinations of two regions. The horizontal axis of all the graphs represents the frequency of unit HZ.



Figure 11.4: Bar graph representation of ANCR for two certain periodic components and of ANCR in time domain. In the top three panels, from the left to the right panel, ANCR on the task condition are shown for the periodic components of the frequency 0.018 HZ (around 55 sec./cycle), 0.031 HZ (around 32 sec./cycle) and ANCR in time domain. In the bottom three panels, ANCR on the continuous control condition are shown for the same components as the top panels.



Figure 11.5: Time series from three regions and their innovations by MEMAR model fitting. In the top row, the time series extracted from three regions (V1, V5, PP) are shown. Time series, histogram and the autocorrelation function of the innovations (1-step prediction error) obtained by MAR model fitting are shown from the second to the bottom row, respectively.



Figure 11.6: Power spectra of three regions of interest (V1, V5, PP) and ANCR among those regions on the task condition (thick line) and the control condition (thin line). The parametric power spectra of V1, V5 and PP are plotted from the left to right in the top row, respectively. The graphs in three another rows show ANCR for all combinations of two regions. The horizontal axis of all the graphs represents the frequency of unit HZ.



Figure 11.7: Bar graph representation of ANCR for two certain periodic components and of ANCR in time domain. In the top three panels, from the left to the right panel, ANCR on the task condition are shown for the periodic components of the frequency 0.017 HZ (around 59 sec./cycle), 0.033 HZ (around 30 sec./cycle) and ANCR in time domain. In the bottom three panels, ANCR on the control condition are shown for the same components as the top panels.

Chapter 12

Discussion

In the Part II, for the purpose of evaluating the change of the effective connectivity, two strategies [A1], [A2] have been investigated. The key tools of this research are the multivariate AR model (MAR) and Akaike's noise contribution ratio (ANCR), which measures the contributions of another time series on one time series for every periodic component.

Taking advantage of ANCR for the component of the same period as visual stimulus, in the analysis [A1] the change of the effective connectivity can be evaluated as the difference of ANCR from the experiments with periodic task and with continuous control. The strategy [A2], in which two different AR models are fitted to time series on the task and the control condition, also provides the similar result as that of [A1]. The difficulty, however, occurs in the interpretation of the peak frequency of the power spectrum in the analysis [A2], whereas in the analysis [A1], the period of the stimulus is fully used for the evaluation of the connectivity. Therefore the result of the strategy [A1] has been taken as the final conclusion.

From the result of [A1], we can observe that the bottom-up connectivity tends to increase on the task condition, compared with the continuous control condition. However, we have to be too careful to the interpretation of the result because temporal resolution of the phenomenon in the fMRI data is much slower than that of neuronal activities. Therefore this result can not directly reflect connectivity on neuronal signals among the regions. Hence, what we can say from the result is that, for instance, the local cerebral blood flow in V1 better explains the future behavior of the local cerebral blood flow in V5 on the task condition than that on the control condition.

As the methodology to evaluate the effective connectivity, the multivariate time series approach demonstrated in this thesis is contrast to the structural equation modeling as to the freedom how to specify the direction of connectivity. In the SEM, the direction has to be specified a priori by the analyst, whereas in the multivariate time series approach, this direction can be automatically determined based on the temporal order of events (i.e. Granger causality). There is a possibility to misspecify the connectivity which does not exist in anatomy by the multivariate time series approach. Therefore in the presence of the anatomical knowledge, the method to incorporate these knowledge into the time series approach is desirable. The multivariate time series approach combined with the Bayesian estimation would be a very promising approach and it is expected to provide more meaningful knowledge about connectivity.

The following points could be very important and challenging in the future,

- The Bayesian estimation should be applied instead of the maximum likelihood estimator in order to involve anatomical prior knowledge about the connectivity.
- In order to confirm the validity of the method proposed here, we should apply to the various experimental data.
- For the knowledge discovery the inter-subject analysis is very important, thus the statistical method for this purpose is highly desired.
- The difference of ANCR between two conditions may not be appropriate for quantifying the change of effective connectivity. Another causality measures should be applied or developed.

Appendix A

Calculation of ABIC for WMN method

Here calculation of ABIC is shown for the WMN method. The result of this appendix can be easily applied to a class of the WMN solution by adjusting a weight matrix W. For example, setting W = L, ABIC for the LORETA method can be obtained.

The likelihood and prior function for WMN solution are given by

$$p(\boldsymbol{v}|\boldsymbol{j};\sigma^2) \sim N(K\boldsymbol{j},\sigma^2 C_{\boldsymbol{\epsilon}})$$
 (A.1)

$$p(\boldsymbol{v};\tau^2) \sim N(0,\tau^2(W'W)^{-1}).$$
 (A.2)

The detailed calculation of ABIC is as follows. Firstly the posterior distribution can be expressed by,

$$p(\boldsymbol{j}|\boldsymbol{v};\sigma^{2},\tau^{2}) = p(\boldsymbol{v}|\boldsymbol{j};\sigma^{2})p(\boldsymbol{j};\tau^{2})$$

$$= (2\pi)^{-\frac{N_{e}}{2}}(2\pi)^{-\frac{N_{s}}{2}}|\sigma^{2}C_{\boldsymbol{\epsilon}}|^{-\frac{1}{2}}|\tau^{2}(W'W)^{-1}|^{-\frac{1}{2}}$$

$$\exp\left\{-\frac{1}{2\sigma^{2}}||\boldsymbol{v}-\mathbf{K}\boldsymbol{j}||_{C_{\boldsymbol{\epsilon}}}^{2}\right\}\exp\left\{-\frac{1}{2\tau^{2}}||W\boldsymbol{j}||^{2}\right\}$$

$$= (2\pi)^{-\frac{N_{e}+N_{s}}{2}}|\sigma^{2}C_{\boldsymbol{\epsilon}}|^{-\frac{1}{2}}|\tau^{2}(W'W)^{-1}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2\sigma^{2}}E(\boldsymbol{j};\lambda)\right\}$$

$$= (2\pi)^{-\frac{N_{e}+N_{s}}{2}}|\sigma^{2}C_{\boldsymbol{\epsilon}}|^{-\frac{1}{2}}|\tau^{2}(W'W)^{-1}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2\sigma^{2}}E(\boldsymbol{j};\lambda)\right\}$$

$$\exp\left\{-\frac{1}{2\sigma^{2}}(\boldsymbol{j}-\boldsymbol{\hat{j}})'(\mathbf{K}'C_{\boldsymbol{\epsilon}}^{-1}\mathbf{K}+\lambda^{2}W'W)(\boldsymbol{j}-\boldsymbol{\hat{j}})\right\}$$
(A.3)

where $\lambda \equiv \sigma/\tau$, $E(\mathbf{j}; \lambda)$ and $\hat{\mathbf{j}}$ which is the minimizer of $E(\mathbf{j}; \lambda)$ are given by

$$E(\boldsymbol{j}; \lambda) = ||\boldsymbol{v} - \mathbf{K}\boldsymbol{j}||_{C_{\boldsymbol{\epsilon}}^{-1}}^{2} + \lambda^{2}||W\boldsymbol{j}||^{2}.$$
$$\hat{\boldsymbol{j}} = (\mathbf{K}'C_{\boldsymbol{\epsilon}}^{-1}\mathbf{K} + \lambda^{2}W'W)^{-1}\mathbf{K}'C_{\boldsymbol{\epsilon}}^{-1}\boldsymbol{v}.$$

From the third equation to the fourth equation, the next arrangements are employed

$$\exp\left\{-\frac{1}{2\sigma^{2}}E(\boldsymbol{j};\boldsymbol{\lambda})\right\} = \exp\left\{-\frac{1}{2\sigma^{2}}\left|\left|\begin{bmatrix}C_{\boldsymbol{\epsilon}}^{-\frac{1}{2}}\boldsymbol{v}\\0\end{bmatrix} - \begin{bmatrix}C_{\boldsymbol{\epsilon}}^{-\frac{1}{2}}\mathbf{K}\\\boldsymbol{\lambda}W\end{bmatrix}\boldsymbol{j}\right|\right|^{2}\right\}$$
$$= \exp\left\{-\frac{1}{2\sigma^{2}}(||\boldsymbol{\breve{v}} - \mathbf{\breve{K}}\boldsymbol{\hat{j}} + \mathbf{\breve{K}}\boldsymbol{\hat{j}} - \mathbf{\breve{K}}\boldsymbol{j}||^{2})\right\}$$
$$= \exp\left\{-\frac{1}{2\sigma^{2}}(||\boldsymbol{\breve{v}} - \mathbf{\breve{K}}\boldsymbol{\hat{j}}||^{2} + ||\mathbf{\breve{K}}(\boldsymbol{j} - \boldsymbol{\hat{j}})||^{2})\right\},$$
$$e \ \breve{\boldsymbol{v}} = \begin{bmatrix}C_{\boldsymbol{\epsilon}}^{-\frac{1}{2}}\boldsymbol{v}\end{bmatrix} \text{ and } \mathbf{\breve{K}} = \begin{bmatrix}C_{\boldsymbol{\epsilon}}^{-\frac{1}{2}}\mathbf{K}\end{bmatrix}$$

where \breve{v}

ere $\boldsymbol{\breve{v}} = \begin{bmatrix} \mathbf{v} & \mathbf{v} \\ 0 \end{bmatrix}$ and $\mathbf{K} = \begin{bmatrix} \mathbf{v} & \mathbf{v} \\ \lambda W \end{bmatrix}$. Because the second exponential of Eq.(A.3) includes the integrand \boldsymbol{j} , the marginal distribution of \boldsymbol{v} becomes

$$p(\boldsymbol{v};\sigma^{2},\tau^{2}) = \int p(\boldsymbol{v}|\boldsymbol{j};\sigma^{2})p(\boldsymbol{j};\tau^{2})d\boldsymbol{j}$$

= $(2\pi)^{-\frac{N_{\epsilon}}{2}}|\sigma^{2}C_{\epsilon}|^{-\frac{1}{2}}|\tau^{2}(W'W)^{-1}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2\sigma^{2}}E(\boldsymbol{\hat{j}};\lambda)\right\}$
 $|\sigma^{2}(\mathbf{K}'C_{\epsilon}^{-1}\mathbf{K}+\lambda^{2}W'W)^{-1}|^{\frac{1}{2}}.$ (A.4)

Then we obtain (-2) times type-II log-likelihood as follows,

$$\mathcal{M}(\sigma,\tau) = N_{e} \log 2\pi + \log |\sigma^{2}C_{\epsilon}| + \log |\tau^{2}(W'W)^{-1}| + \frac{1}{\sigma^{2}}E(\hat{\boldsymbol{j}};\lambda) - \log |\sigma^{2}(K'C_{\epsilon}^{-1}K + \lambda^{2}W'W)^{-1}| = N_{e} \log 2\pi + \log |\sigma^{2}C_{\epsilon}| + \log |\tau^{2}(W'W)^{-1}| + \frac{1}{\sigma^{2}}E(\hat{\boldsymbol{j}};\lambda) - \log |\sigma^{2}(K'C_{\epsilon}^{-1}K + \lambda^{2}W'W)^{-1}| = N_{e} \log 2\pi + \log |\sigma^{2}C_{\epsilon}| + \frac{1}{\sigma^{2}}E(\hat{\boldsymbol{j}};\lambda) + \log \left|\frac{(W'W)^{-1}}{\lambda^{2}(K'C_{\epsilon}^{-1}K + \lambda^{2}W'W)^{-1}}\right|.$$
(A.5)

Using the singular value decomposition of $\bar{\mathbf{K}} \equiv C_{\epsilon}^{-1/2} \mathbf{K} W^{-1}$

$$\bar{\mathbf{K}} = USV',$$

The third and fourth term in (A.5) can be further simplified as,

$$E(\hat{\boldsymbol{j}};\lambda) = ||C_{\boldsymbol{\epsilon}}^{-\frac{1}{2}}(\boldsymbol{v} - \mathbf{K}\hat{\boldsymbol{j}})||^{2} + \lambda^{2}||W\hat{\boldsymbol{j}}||^{2}$$

$$= ||\{\mathbf{I} - \bar{\mathbf{K}}(\bar{\mathbf{K}}'\bar{\mathbf{K}} + \lambda^{2}\mathbf{I})^{-1}\bar{\mathbf{K}}'\}C_{\boldsymbol{\epsilon}}^{-\frac{1}{2}}\boldsymbol{v}||^{2} + \lambda^{2}||(\bar{\mathbf{K}}'\bar{\mathbf{K}} + \lambda^{2}\mathbf{I})^{-1}\bar{\mathbf{K}}'C_{\boldsymbol{\epsilon}}^{-\frac{1}{2}}\boldsymbol{v}||^{2}$$

$$= ||U\{\mathbf{I} - S(S'S + \lambda^{2}\mathbf{I})^{-1}S'\}\tilde{\boldsymbol{v}}||^{2} + \lambda^{2}||V(S'S + \lambda^{2}\mathbf{I})^{-1}S'\tilde{\boldsymbol{v}}||^{2}$$

$$= \tilde{\boldsymbol{v}}'\mathrm{diag}\Big(\frac{\lambda^{4}}{(s_{i}^{2} + \lambda^{2})^{2}}\Big)\tilde{\boldsymbol{v}} + \tilde{\boldsymbol{v}}'\mathrm{diag}\Big(\frac{s_{i}^{2}\lambda^{2}}{(s_{i}^{2} + \lambda^{2})^{2}}\Big)\tilde{\boldsymbol{v}}$$

$$= \tilde{\boldsymbol{v}}'\mathrm{diag}\Big(\frac{\lambda^{2}}{s_{i}^{2} + \lambda^{2}}\Big)\tilde{\boldsymbol{v}} \qquad (A.6)$$

and

$$\log \left| \frac{(W'W)^{-1}}{\lambda^{2} (K'C_{\epsilon}^{-1}K + \lambda^{2}W'W)^{-1}} \right| = -\log |\lambda^{2} (\bar{K}'\bar{K} + \lambda^{2}I)^{-1}|$$

$$= -\log |\lambda^{2} (VS'SV' + \lambda^{2}I)^{-1}|$$

$$= -\log |\lambda^{2} (S'S + \lambda^{2}I)^{-1}|$$

$$= -\log \prod_{i=1}^{N_{s}} \frac{\lambda^{2}}{s_{i}^{2} + \lambda^{2}}$$

$$= \sum_{i=1}^{N_{e}} \log \frac{s_{i}^{2} + \lambda^{2}}{\lambda^{2}}.$$
(A.7)

where $\tilde{\boldsymbol{v}} = U'C_{\boldsymbol{\epsilon}}^{-1/2}\boldsymbol{v}$ and s_i is the *i*th diagonal component of the matrix S.

Substituting Eqs.(A.6) and (A.7) into (A.5), (-2) times type-II log-likelihood can be obtained as

$$\mathcal{M}(\sigma,\lambda) = N_e \log 2\pi\sigma^2 + \log |C_\epsilon| + \frac{1}{\sigma^2} \sum_{i=1}^{N_e} \frac{\lambda^2}{s_i^2 + \lambda^2} \tilde{v}_i^2 + \sum_{i=1}^{N_e} \log \frac{s_i^2 + \lambda^2}{\lambda^2}.$$
 (A.8)

where \tilde{v}_i is *i*th component of the vector $\tilde{\boldsymbol{v}}$. Here we have replaced the parameter τ by $\lambda = \sigma/\tau$. The hyper-parameter σ, λ can be estimated as so that the function \mathcal{M} is minimized. Differentiating \mathcal{M} by σ^2 , the estimate of σ^2 is provided by

$$\hat{\sigma}^2 = \frac{1}{N_e} \sum_{i=1}^{N_e} \frac{\lambda^2}{s_i^2 + \lambda^2} \tilde{v}_i^2 \tag{A.9}$$

The cost function for estimating the regularization parameter λ is given by

$$\mathcal{M}'(\lambda) = N_e \log 2\pi + N_e \log \hat{\sigma}^2 + \log |C_{\epsilon}| + N_e + \sum_{i=1}^{N_e} \log \frac{s_i^2 + \lambda^2}{\lambda^2}.$$
 (A.10)

When WMN method of a fixed regularized parameter is fitted to time series observations $\boldsymbol{v}_1, \dots, \boldsymbol{v}_T$, the likelihood can be considered to be i.i.d. Therefore the estimates $\hat{\sigma}$ and the cost function \mathcal{M}'_T can be obtained as follows,

$$\hat{\sigma}^2 = \frac{1}{TN_e} \sum_{t=1}^T \sum_{i=1}^{N_e} \frac{\lambda^2}{s_i^2 + \lambda^2} \tilde{v}_{i,t}^2$$
(A.11)

$$\mathcal{M}'_{T}(\lambda) = TN_{e}(\log 2\pi\hat{\sigma}^{2} + 1) + T\log|C_{\epsilon}| + T\sum_{i=1}^{N_{e}}\log\frac{s_{i}^{2} + \lambda^{2}}{\lambda^{2}}.$$
 (A.12)

This ABIC value can be directly used in order to compare instantaneous inverse solutions (WMN solutions) with AIC of dynamic inverse solutions.

Appendix B

Some lemmas on multivariate normal distributions

Two lemmas about the conditional variance and expectation for multivariate normal distribution are introduced and proved here. For detailed issue, see standard textbooks of multivariate analysis such as Mardia et al. 1979; Anderson 2003.

Let x, y, z denote (vector) Gaussian random variables. Furthermore the expectation and covariance matrices of these vectors are denoted by μ_x , Σ_{xy} and so on.

Lemma 1

$$E[x|y] = \mu_x + \Sigma_{xy} \Sigma_{yy}^{-1} (y - \mu_y)$$
 (B.1)

$$\operatorname{Var}[x|y] = \Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx}$$
(B.2)

Lemma 2

$$\mathbf{E}[x|y,z] = E[x|y] + \sum_{xz} \sum_{zz}^{-1} z \tag{B.3}$$

$$\operatorname{Var}[x|y, z] = \operatorname{Var}[x|y] - \Sigma_{xz} \Sigma_{zz}^{-1} \Sigma_{xz}'$$
(B.4)

If $\Sigma_{yz} = 0$ (i.e. $y \perp z$) and $\mu_z = 0$.

Lemma 1 means that when x, y are the Gaussian random variables with mean 0, the conditional expectation is given by

$$\mathbf{E}[x|y] = \Sigma_{xy} \Sigma_{yy}^{-1} y = \mathbf{E}[xy] \mathbf{E}[yy]^{-1} y.$$

In the other words, the conditional expectation corresponds to "the projection" of x onto y (in the sense of L^2 norm). In the original literature of Kalman filtering (Kalman 1960), Lemma 1 and Lemma 2 have been applied to the derivation, where in particular Lemma 2 was referred to as "the orthogonal projection".

Proof of Lemma 1 The joint distribution of x and y is given by

$$p(x,y) = (2\pi)^{-\frac{d}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(x-\mu_x \ y-\mu_y) \begin{bmatrix}\Sigma_{xx} \ \Sigma_{xy}\\\Sigma_{yx} \ \Sigma_{yy}\end{bmatrix}^{-1} \begin{pmatrix}x-\mu_x\\y-\mu_y\end{pmatrix}\right\}$$

Now applying the linear transformation

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{bmatrix} I & -\Sigma_{xy}\Sigma_{yy}^{-1} \\ 0 & I \end{bmatrix} \equiv T\begin{pmatrix} x - \mu_x \\ y - \mu_y \end{pmatrix}$$

to the variables x, y gives

$$p(X,Y) = C \exp\left\{-\frac{1}{2}(X \ Y)(T^{-1})' \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}^{-1} T^{-1} \begin{pmatrix} X \\ Y \end{pmatrix}\right\}$$

where C is a constant term. Note that the Jacobian of T is 1. By directly calculating inside of the exponential, we obtain:

$$p(X,Y) = C \exp \left\{ -\frac{1}{2} \begin{pmatrix} X & Y \end{pmatrix} \begin{bmatrix} \Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx} & 0 \\ 0 & \Sigma_{yy} \end{bmatrix}^{-1} \begin{pmatrix} X \\ Y \end{pmatrix} \right\}.$$

Rewriting this probability density using the original variables x, y leads to

$$p(x,y) = C \exp \left[-\frac{1}{2} \left\{ \left(x - \mu_x + \Sigma_{xy} \Sigma_{yy}^{-1} (y - \mu_y) \right)' \Sigma_{x|y}^{-1} \left(x - \mu_x + \Sigma_{xy} \Sigma_{yy}^{-1} (y - \mu_y) \right) + (y - \mu_y)' \Sigma_{yy}^{-1} (y - \mu_y) \right\} \right]$$
(B.5)

where $\Sigma_{x|y} = \Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx}$. Integrating both-hand sides of Eq.(B.5) by x, the marginal distribution of y is obtained as

$$p(y) = C_1 \exp\left\{-\frac{1}{2}(y-\mu_y)'\Sigma_{yy}^{-1}(y-\mu_y)\right\}$$
(B.6)

Since the conditional distribution of x on y is obtained as

$$p(x|y) = \frac{p(x,y)}{p(y)},$$

the conditional expectation and variance are given by

$$E[x|y] = \mu_x + \Sigma_{xy}\Sigma_{yy}^{-1}(y - \mu_y)$$

Var[x|y] = $\Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx}$.

Proof of Lemma 2 Applying (B.1) to the vector $\begin{bmatrix} y \\ z \end{bmatrix}$ in place of y gives

$$\begin{aligned} \mathbf{E}[x|y,z] &= \mu_x + [\Sigma_{xy} \ \Sigma_{xz}] \begin{bmatrix} \Sigma_{yy}^{-1} & 0\\ 0 \ \Sigma_{zz}^{-1} \end{bmatrix} \begin{pmatrix} y - \mu_y \\ z \end{pmatrix} \\ &= \mu_x + \Sigma_{xy} \Sigma_{yy}^{-1} (y - \mu_y) + \Sigma_{xz} \Sigma_{zz}^{-1} z \\ &= \mathbf{E}[x|y] + \Sigma_{xz} \Sigma_{zz}^{-1} z \end{aligned}$$

In the same way, Applying (B.2) to the vector $\begin{bmatrix} y \\ z \end{bmatrix}$ in place of y gives

$$\operatorname{Var}[x|y,z] = \Sigma_{xx} - [\Sigma_{xy} \ \Sigma_{xz}] \begin{bmatrix} \Sigma_{yy}^{-1} & 0\\ 0 & \Sigma_{zz}^{-1} \end{bmatrix} \begin{bmatrix} \Sigma_{yx}\\ \Sigma_{zx} \end{bmatrix}$$
$$= \Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx} - \Sigma_{xz} \Sigma_{zz}^{-1} \Sigma_{zx}$$
$$= \operatorname{Var}(x|y) - \Sigma_{xz} \Sigma_{zz}^{-1} \Sigma_{zx}$$

As can be seen, $\Sigma_{yz} = 0$, that is, the orthogonality between x and y, is essential for the proof.

Acknowledgments

In the course of the past three years, many people have contributed to my work. I would like to mention my gratitude for all of them.

As a supervisor of my Ph. D. course Prof. Tohru Ozaki in the Institute of Statistical Mathematics (ISM) has continued to encourage me and guided me to the goal. In the beginning of my Ph. D. course he motivated me to work for the real data in the neuroscience. Thanks to his instruction I have learned the importance of researching the real problem. He has also introduced many collaborators of his to me and it is of great benefit for me to experience the collaboration with the foreigners.

Dr. Andreas Galka in ISM is the closest collaborator in these three years. We have discussed about many detailed issues of the EEG dynamical inverse problem. The discussion with him is great help for me to write this thesis. He also helped me with the simulation studies in this thesis.

Prof. Roland Biscay in Havana University gave me a lot of helpful comments when we started researching the dynamical inverse problem. His helpful comment enabled us to develop the dynamical inverse solution with the Laplacian matrix incorporated into the system noise covariance.

Prof. Pedro Valdes-Sosa in Cuba Neuroscience Center (CNC) first gave me a lecture of the EEG inverse problem. His enthusiasm for developing the statistical method in neuroscience drove me to work this EEG inverse problem. I should mention that during my first stay in Cuba, I could enjoy myself very much thanks to his consideration.

Prof. Norihiro Sadato in the National Institute for Physiological Science (NIPS) has given me an opportunity to work the fMRI data. It is one of my precious experience to take part in the experiment and to analyze the data acquired in the experiment. This experiment was done with help of many people in NIPS; Prof. Tomohisa Okada, Dr. Daisuke Saito, Satoshi Tanaka and Tokiko Harada. I really appreciate their help.

I would be grateful to my colleagues for their supports and their friendship and also thank Jorge Bosch, Eduardo Martinez-Montes and Nelson Trujillo-Barreto in CNC not only for encouraging me with their interesting joke but also for giving me their serious comments about the research.

I would like to thank Prof. Tomoyuki Higuti, Prof. Yoshiyasu Tamura, and Prof. Takashi Tsuchiya in ISM and Dr. Jorge Riera in Tohoku University for reviewing my dissertation of Ph. D. and giving helpful comments for improving this thesis.

And finally I would like to extend my gratitude to my family, who have supported me through the years and continued to encourage me to pursue and achieve my goals .

References

- Akaike, H. (1968). On the use of a linear model for the identification of feedback systems. Annals of the Institute of Statistical Mathematics 20(3), 425–439.
- Akaike, H. (1973). Information theory and an extension of maximum likelihood principle. In 2nd International symposium on information theory, pp. 267–281.
- Akaike, H. (1980a). Likelihood and the bayesian procedure. In J. M. Bernardo, M. H. Degroot, D. V. Lindley, and A. F. M. Smith (Eds.), *Bayesian Statistics*, pp. 141–166. Valencia, Spain: University Press.
- Akaike, H. (1980b). Seasonal adjustment by a Bayesian modeling. *Journal* of Time Series Analysis 1(1), 1–13.
- Akaike, H. and T. Nakagawa (1988). Statistical Analysis and Control of Dynamic Systems. Kluwer Academic Publishers.
- Anderson, T. W. (2003). An Introduction to Multivariate Statistical Analysis (3rd ed.). John Wiley & Sons.
- Ansley, C. F. and R. Kohn (1982). A geometrical derivation of the fixed interval smoothing algorithm. *Biometrika* 69, 486–487.
- Aoki, M. (1987). State Space Modeling of Time Series. Berlin: Springer.
- Ary, J. P., S. A. Klein, and D. H. Fender (1981). Location of sources of evoked scalp potentials: Corrections for skull and scalp thickness. *IEEE Transactions on Biomedical Engineering* 17(2), 141–146.
- Awada, K. A., D. R. Jackson, J. T. Williams, D. R. Wilton, S. B. Baumann, and A. C. Papanicolaous (1997). Computational aspects of finite element modeling in EEG source localization. *IEEE Transactions on Biomedical Engineering* 44(8), 736–752.
- Backus, G. and F. Gilbert (1968). The resolving power of gross earth data. Geophys. J. R. Astr. Soc. 16, 169–205.
- Baillet, S. and J. C. Mosher (2001, November). Electromagnetic brain mapping. *IEEE Signal Processing Magazine*, 14–30.

- Bengtsson, Τ. and J. Ε. Cavanaugh. An improved Akaike information criterion for model selection. state space http://www.cgd.ucar.edu/~tocke/aici_ss.pdf.
- Brockwell, P. J. and R. A. Davis (1996). *Introduction to Time Series and Forecasting*. New York: Springer.
- Brooks, D. H., F. Ghandi, R. S. Macleod, and M. Maratos (1999). Inverse electrocardiography by simultaneous imposition of multiple constraints. *IEEE Transactions on Biomedical Engineering* 46(1), 3–17.
- Büchel, C., R. J. S. Wise, C. J. Mummery, J. B. Poline, and K. J. Friston (1996). Nonlinear regression in parametric activation studies. *NeuroIm*age 4, 60–66.
- de Jong, P. (1989). Smoothing and interpolation with the state-space model. Journal of the American Statistical Association 84 (408), 1085–1088.
- deMunck, J. C. and M. J. Peters (1993). A fast method to compute the potential in the multisphere model. *IEEE Transactions on Biomedical Engineering* 40, 1166–1174.
- Durbin, J. and S. J. Koopman (2001). Time Series Analysis by State Space Models. Oxford Univ. Press.
- Efron, B., T. Hastie, I. Johnstone, and R. Tibshirani (2002, Aug.). Least angle regression. web page.
- Evensen, G. (1994). Sequential data assimilation with a nonlinear quasigeostrophic model using Monte Carlo methods to forecast error statistics. J. geophys. Res. 99(C5), 10,143–10,162.
- Evensen, G. (1997). Advanced data assimilation for strongly nonlinear dynamics. Monthly Weather Review 125, 1342–1354.
- Evensen, G. (2003). The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics* 53, 343–367.
- Friston, K. J., A. P. Holmes, K. P. Worsley, J. B. Poline, C. D. Frith, and R. S. J. Frackowiak (1995). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping 2*, 189–210.
- Friston, K. J., L. G. Ungerleider, P. Jezzard, and R. Turner (1995). Characterizing modulatory interactions between V1 and V2 in human cortex with fMRI. *Human Brain Mapping 2*, 211–224.
- Fukumori, I. (2001). Data assimilation by models. In Satellite Altimetry and Earth Sciences, Chapter 5, pp. 237–265. Academic Pr.
- Fukumori, I. (2002, May). A partitioned kalman filter and smoother. American Meteorological Society 130(5), 1370–1383.

- Galka, A., T. Ozaki, O. Yamashita, R. Biscay, and P. Valdes-Sosa (2002). A solution to the dynamical inverse problem of EEG generation using spatiotemporal kalman filtering. Research Memorandum 862, The Institute of Statistical Mathematics.
- Geselowitz, D. B. (1967). On bioelectric potentials in an inhomogeneous volume conductor. *Biophysical journal* 7, 1–11.
- Geweke, J. F. (1982). Measurement of linear dependence and feedback between multiple time series. Journal of the American Statistical Association 77, 304–324.
- Geweke, J. F. (1984). Measures of conditional linear dependence and feedback between time series. Journal of the American Statistical Association 79, 907–915.
- Goronidnitsky, I. F. and B. D. Rao (1997). Spatial signal reconstruction from limited data using FOCUS: A reweighted minimum norm algorithm. *IEEE Transactions on Signal Processing* 45, 1–16.
- Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37(3), 424–438.
- Grave de Peralta Menendez, R. and S. L. Gonzalez Andio (1999). Backus and gilbert method for vector fields. *Human Brain Mapping* 7, 161–165.
- Grave de Peralta Menendez, R. and S. L. Gonzalez Andio (2002). Comparison of algorithms for localization of focal source evaluation with simulated data and analysis of experimental data. *International Journal of Bioelectromagnetism* 4(1).
- Grave de Peralta Menendez, R., O. Hauk, S. L. Gonzalez Andio, H. Vogt, and C. Michel (1997). Linear inverse solutions with optimal resolution kernels applied to electromagnetic tomography. *Human Brain Mapping 5*, 454–467.
- Hämäläinen, M. S. and R. J. Ilmoniemi (1984). Interpreting measured magnetic fields of the brain: estimates of current distributions. Technical Report TKK-F-A559, Helsinki University of Technology.
- Hamilton, J. D. (1994). Time Series Analysis. Princeton University Press.
- Hansen, P. C. (1992). Analysis of discrete ill-posed problems by means of the L-curve. SIAM Review 34, 561–580.
- Hansen, P. C. (1994). Regularization tools: A matlab package for analysis and solution of discrete ill-posed problems. Numerical Algorithms 6, 1– 35.
- Hansen, P. C. (2000). The L-curve and its use in the numerical treatment of inverse problems. In P. Johnston (Ed.), *Computational Inverse Problems* in Electrocardiology. WIT Press.

- Harrison, L., W. D. Penny, and K. Friston (2003). Multivariate autoregressive modeling of fMRI time series. *NeuroImage 19*, 1477–1491.
- Hosoya, Y. (1991). The decomposition and measurement of the interdependency between second-order stationary process. *Probability Theory and Related Fileds* 88, 429–444.
- Hurvich, C. and C. Tsai (1989). Regression and time series model selection in small samples. *Biometrika* 76, 297–307.
- Kailath, T. (1980). Linear Systems. Prentice-Hall.
- Kaipio, J. P., P. A. Karjalainen, E. Somersalo, and M. Vauhkonen (1999). State estimation in time-varying electrical impedance tomography. Annals New York Acad Sci 873, 430–439.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. Journal of Basic Engineering 82, 35–45.
- Kaminski, M., M. Ding, W. A. Truccolo, and S. T. Bressler (2001). Evaluating causal relations in neural systems: Granger causality, directed transfer function and statisitical assessment of significance. *Biological Cybernetics* 85, 145–157.
- Karjalainen, P. A., M. Vauhkonen, and J. P. Kaipio (1997). Dynamic reconstruction in SPET. In Proc. 19th Int. Conf. IEEE Eng. Med. Biol. Society.
- Kitagawa, G. and W. Gersh (1996). Smoothness Priors Analysis of Time Series. New York: Springer.
- Klepfer, R. N., C. R. Johnson, and R. S. Macleod (1997). The effect of inhomogenetities and anisotropies on electrocardiographic fields. *IEEE Transactions on Biomedical Engineering* 44(8), 706–719.
- Mardia, K., J. T. Kent, and J. M. Bibby (1979). *Multivariate Analysis*. San Diego: Academic Press.
- Matsuura, K. and Y. Okabe (1995). Selective minimum-norm solution of the biomagnetic inverse problem. *IEEE Transactions on Biomedical En*gineering 42(6), 608–615.
- Matsuura, K. and Y. Okabe (1996). A robust reconstruction of sparse biomagnetic sources. *IEEE Transactions on Biomedical Engineering* 44(8), 720–726.
- Mazziotta, J. C., A. Toga, A. C. Evans, P. Fox, and J. Lancaster (1995). A probabilistic atlas of the human brain: theory and rationale for its development. *NeuroImage 2*, 89–101.
- McGuire, J. J. and P. Segall (2003). Imaging of aseismic fault slip transients recorded by dense geodetic networks. *Geophys. J. Int.* (155), 778–788.

- McIntosh, A. R. (2000). Towards a network theory of cognition. *Neural Networks* 13, 861–870.
- McIntosh, A. R. and F. Gonzalez-Lima (1994). Structural equation modeling and its application to network analysis in functional brain imaging. *Human Brain Mapping 2*, 2–22.
- Meinhold, R. J. and N. D. Singpurwalla (1983). Understanding the Kalman filter. *The American Statistician* 37(2), 123–127.
- Mosher, J. C., R. M. Leahy, and P. S. Lewis (1999). EEG and MEG: Forward solutions for inverse methods. *IEEE Transactions on Biomedical Engineering* 46, 245–259.
- Osborne, M. R., B. Presnell, and B. A. Turlach (2000). A new approach to variable selection in least squares problems. *IMA Journal of Numerical Analysis 20*, 389–404.
- Pascual-Marqui, R. D. (1999). Review of methods for solving the eeg inverse problem. International Journal of Bioelectromagnetism 1(1), 75–86.
- Pascual-Marqui, R. D. (2002). Standardized low resolution brain electromagnetic tomography(sLORETA): technical details. Methods & Findings in Experimental & Clinical Pharmacology 24, 5–12.
- Pascual-Marqui, R. D. and C. M. Michel (1994). LORETA: New authentic 3d functional images of the brain. In W. Skrandies (Ed.), *ISBET newsletter*, Number 5, pp. 4–25.
- Pascual-Marqui, R. D., C. M. Michel, and D. Lehmann (1994). Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int. J. Psychophysiol. 18, 49–65.
- Phillips, C., M. D. Rug, and K. J. Friston (2002). Systematic regularization of linear inverse solutions of the eeg source localization problem. *NeuroImage* 17, 287–301.
- Phillipse, C. L. M. (2001). Source estimation in EEG. Ph. D. thesis, Universite de Liege.
- Riera, J. J. and M. E. Fuentes (1998). Electric lead field for a piecewise homogeneous volume conductor model of the head. *IEEE Transactions* on Biomedical Engineering 45(6), 746–753.
- Rodin, E. A. and M. J. Rodin (1995). Dipole sources of the human alpha rhythm. *Brain Topography* 7, 201–208.
- Rosenfeld, M., R. Tanami, and S. Abboud (1996). Numerical solution of the potential due to dipole sources in volume conductors with arbitrary geometry and conductivity. *IEEE Transactions on Biomedical Engineering* 43(7), 679–689.

- Salheen, H. I. and T. K. Ng (1997). New finite difference formulations for general inhomogeneous anisotropic bioelectric problems. *IEEE Transac*tions on Biomedical Engineering 44 (9), 800–809.
- Sarvas, J. (1987). Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Physics in Medicine and Biology* 32, 11–22.
- Scherg, M. and J. S. Ebersole (1994). Brain source imaging of focal and multifocal epileptiform EEG activity. *Clinical Neurophysiology* 24, 51– 60.
- Schlitt, H. A., L. Heller, R. Aaron, E. Best, and D. M. Ranken (1995). Evaluation of boundary element methods for EEG forward problem: Effect of linear interpolation. *IEEE Transactions on Biomedical Engineering* 42(1), 52–58.
- Schmidt, U. and A. K. Louis (2002a). Efficient algorithms for the regularization of dynamic inverse problems: I. theory. *Inverse Problems* 18, 645–658.
- Schmidt, U. and A. K. Louis (2002b). Efficient algorithms for the regularization of dynamic inverse problems: II. applications. *Inverse Problems 18*, 659–676.
- Schmidt, U., A. K. Louis, F. Darvas, H. Buchner, and M. Fuchs (2001). Numerical aspects of spatio-temporal current density reconstruction from EEG-/ MEG-data. *IEEE Trans. Med. Imaging 20*, 314–324.
- Schwarz, G. (1978). Estimating the dimension of a model. Annals of Statistics 6, 461–464.
- Schweppe, F. (1965). Evaluation of likelihood functions for gaussian signals. IEEE Transactions on Information Theory 11, 61–70.
- Segall, P. and M. Matthews (1997). Time dependent inversion of geodesic data. J. geophys. Res. 22(102), 22,391–22,400.
- Shumway, R. H. (2000). *Time series analysis and its applications*. New York: Springer.
- Tanokura, Y. and G. Kitagawa (2003). Extended power contribution that can be applied without independence assumption. Technical Report 886, The Institute of Statistical Mathematics.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. J. Roy. Stat. Soc. B 58, 267–288.
- Tong, P. and J. N. Rosettos (1976). *Finite-element method basic technique* and implementation. Mit Press.
- Uutela, K., M. Hämäläinen, and E. Somersalo (1999). Visualization of magnetoencephalographic data using minimum current estimates. *NeuroIm*age 10, 173–180.

- Valdés-Sosa, P., J. Bosch, R. Grave de Peralta-Menendez, J. L. Hernández, R. Pascual, and R. Biscay (1992). Frequency domain models for the EEG. Brain Topography 4, 309–319.
- Vauhkonen, M., P. A. Karjalainen, and J. P. Kaipio (1998). A kalman filter approach to track fast impedance changes in electrical impedance tomography. *IEEE Transactions on Biomedical Engineering* 45(4), 486–493.
- Vauhkonen, P. J., M. Vouhkonen, and J. P. Kaipio (2001). Fixed-lag smoothing in electrical impedance tomography. Int. J. Num. Meth. Eng. 50, 2195–2209.
- Wahba, G. (1990). Spline Models for Observational Data. Pennsylvania: Society for Industrial and Applied Mathematics.
- Worsley, K. J., C. H. Liao, J. Aston, V. Petre, G. H. Duncan, F. Morales, and A. C. Evans (2002). A general statistical analysis for fMRI data. *NeuroImage* 15, 1–15.
- Yamashita, O., A. Galka, T. Ozaki, R. Biscay, and P. Valdes-Sosa (2002). Dynamic inverse problem and the practical computation procedure for the solution. Technical Report 861, The Institute of Statistical Mathematics.
- Yamashita, O., A. Galka, T. Ozaki, R. Biscay, and P. Valdes-Sosa (2004). Recursive penalized least squares solution for dynamical inverse problems of EEG generation. *Human Brain Mapping*. in press.
- Yamashita, O. and O. Ozaki (2003). Approximate method of high dimensional kalman filtering for EEG dynamical inverse problem. In Japanese Statistical Conference 2003.