2019

Estimation and Clustering in Statistical Ill-posed Linear Inverse Problems

Rasika Rajapakshage
University of Central Florida

ESTIMATION AND CLUSTERING IN STATISTICAL ILL-POSED LINEAR INVERSE PROBLEMS

by

RASIKA RAJAPAKSHAGE
MS, University of Central Florida, 2016

A dissertation submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy
in the Department of Mathematics
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Summer Term
2019

Major Professor: Marianna Pensky

ABSTRACT

The main focus of the dissertation is estimation and clustering in statistical ill-posed linear inverse problems. The dissertation deals with a problem of simultaneously estimating a collection of solutions of ill-posed linear inverse problems from their noisy images under an operator that does not have a bounded inverse, when the solutions are related in a certain way. The dissertation defense consists of three parts. In the first part, the collection consists of measurements of temporal functions at various spatial locations. In particular, we study the problem of estimating a three-dimensional function based on observations of its noisy Laplace convolution. In the second part, we recover classes of similar curves when the class memberships are unknown. Problems of this kind appear in many areas of application where clustering is carried out at the pre-processing step and then the inverse problem is solved for each of the cluster averages separately. As a result, the errors of the procedures are usually examined for the estimation step only. In both parts, we construct the estimators, study their minimax optimality and evaluate their performance via a limited simulation study. In the third part, we propose a new computational platform to better understand the patterns of R-fMRI by taking into account the challenge of inevitable signal fluctuations and interpret the success of dynamic functional connectivity approaches. Towards this, we revisit an auto-regressive and vector auto-regressive signal modeling approach for estimating temporal changes of the signal in brain regions. We then generate inverse covariance matrices from the generated windows and use a non-parametric statistical approach to select significant features. Finally, we use Lasso to perform classification of the data. The effectiveness of the proposed method is evidenced in the classification of R-fMRI scans.

To my mother and grandmother

Thank you all your support, love and guidance.

ACKNOWLEDGMENTS

I would like to thank my doctoral adviser, Dr. Marianna Pensky, for her guidance, attitude and care during my ups and downs. My appreciation is extended to the members of my dissertation committee, Dr. Jason Sawanson, Dr. Ulas Bagci, Teng Zhang and Dr. Hassan Foroosh. I owe many thanks to Julian Pollifrone for his hours of proofreading. Finally, my gratitude is for my mother and grandmother.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

In many statistical inverse problems, one is often interested in recovering a set of solutions of general ill-posed linear inverse problems $Af_m = q_m$, $m = 1, \cdots, M$, where A is a bounded linear operator that does not have a bounded inverse and the right-hand sides q_m are measured with error. In particular, we assume that some of the curves f_m and hence, q_m are very similar to each other, so that they can be averaged and recovered together. This similarity can be of various kinds. One of the possibilities is that the functions f_m are spatially related. In this case, m is a vector index, e.g., $m = (m_1, m_2, m_3)$ if we have a collection of time dependent functions at various spatial locations, or $m = (m_1, m_2)$ if we consider points in some region on the plane.

Another possibility is that some of the curves f_m and hence, q_m are very similar to each other, so that they can be averaged and recovered together. As a result, one supposedly obtains estimators of f_j with smaller errors. The grouping is usually unknown (as well as the number of groups) and is carried out at a pre-processing step by applying one of the standard clustering techniques with the number of clusters determined by trial and error. Subsequently, the curves in the same cluster are averaged and the errors of those aggregated curves are used as true errors in the analysis.

Problems of this kind appear in many areas of application such as astronomy (blurred images), econometrics (instrumental variables), medical imaging (tomography, dynamic contrast enhanced Computerized Tomography and Magnetic Resonance Imaging), finance (model calibration of volatility) and many others where similar curves are measured and can be recovered together.

Our study is motivated by the analysis of Dynamic Contrast Enhanced (DCE) imaging data.

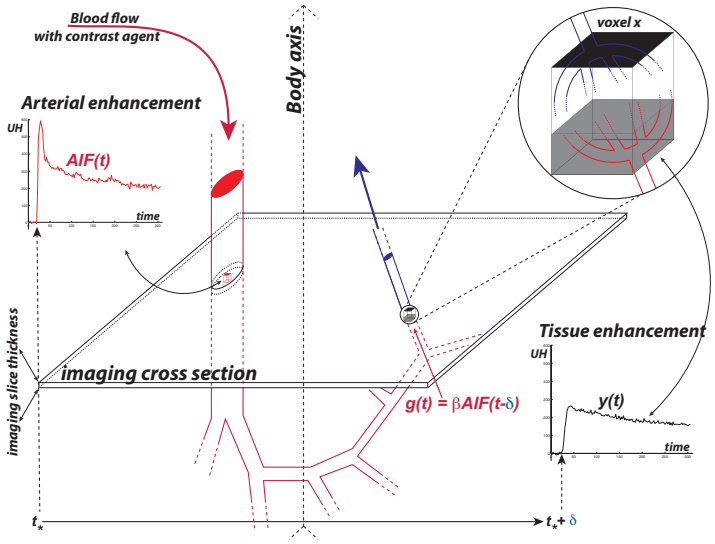

Figure 1.1: DCE imaging experiment and contrast agent circulation. Figure shows a sub-tree of the vascular system going from the artery which receives oxygenated blood (red arrow) to the vein which returns the de-oxygenated blood (blue arrow) after exchanges within the tissue. After passing through the heart, the bolus of the contrast agent, injected into a vein, is distributed, throughout the body along the arterial network to the tissue and later back to the venous system. In the imaging cross-section, the contrast agent induces enhancements first in the artery, providing the AIF, and later in the tissue of interest providing observations $y(t_i),\ i = 1, ..., n$. Enhancements are measured in the voxels of the imaging cross-section

DCE imaging provides a non-invasive measure of tumor angiogenesis and has great potential for cancer detection and characterization, as well as for monitoring. The common feature of DCE imaging techniques is that each of them uses the rapid injection of a single dose of a bolus of a contrast agent and monitors its progression in the vascular network by sequential imaging at times t_i, $i = 1, \cdots, n$. This is accomplished by measuring the pixels' grey levels that are proportional to the concentration of the contrast agent in the corresponding voxels.

At each time instant t_i, one obtains an image of an artery as well as a collection $Y(t_i, \mathbf{x})$ of grey levels for each voxel $\mathbf{x}$. The images of the artery allow to estimate the so called Arterial Input Function, AIF(t_i), that is proportional to the total amount of the contrast

agent entering the tissue area. Comte *et al.* (2017) described the DCE imaging experiment in great detail and showed that the cumulative distribution function $F(z, \mathbf{x})$ of the sojourn times for the particles of the contrast agent entering a tissue voxel $\mathbf{x}$ satisfies the following equation

$$Y(t, \mathbf{x}) = \int_0^{t-\delta} g(t-z)\, \beta(\mathbf{x})(1 - F(z, \mathbf{x}))dz + \epsilon\xi(t, \mathbf{x}). \tag{1.1}$$

Here the errors $\xi(t, \mathbf{x})$ are independent for different t and $\mathbf{x} = (x_1, x_2)$, $g(t) = \mathrm{AIF}(t)$, a positive coefficient $\beta(\mathbf{x})$ is related to a fraction of the contrast agent entering the voxel $\mathbf{x}$ and δ is the time delay that can be easily estimated from data. The function of interest is $f(z, \mathbf{x}) = \beta(\mathbf{x})(1 - F(z, \mathbf{x}))$ where the distribution function $F(z, \mathbf{x})$ characterizes the properties of the tissue voxel $\mathbf{x}$ and can be used as the foundation for medical conclusions. Since the Arterial Input Function can be estimated by denoising and averaging the observations over all voxels of the aorta, its estimators incur much lower errors than those of the left hand side of equation (1.1). For this reason, in our theoretical investigations, we shall treat function g in (1.1) as known. If one is interested in taking the uncertainty about g into account, this can be accomplished using methodology of Vareschi (2015).

Since for each spatial location $\mathbf{x}$, the curves $Y(t, \mathbf{x})$ in (1.1) are very noisy and, in addition, one can combine all tissues in varios spatial locations into three groups: healthy, cancerous and pre-cancerous. For this reason, Lieury *et al.* (2012) clustered all measured spatial curves into those three groups, averaged the curves in each group, so that Comte *et al.* (2017) studies recovery of those averaged curves. Clustering has been done at the pre-processing step and the clustering error is not accounted for in the subsequent error evaluation.

One of the goals of the present paper is to combine the curves that correspond to the neighboring spatial locations or to similar type of tissues in order to improve the accuracy of the reconstructed curves while explicitly accounting for the errors that can occur in this process.

In particular, in the first part of the paper we consider functional deconvolution approach where all curves are combined using combination of wavelet and denoising techniques. In the second part of the paper, we consider a more general problem where, for every $\mathbf{x}$, the curves $f(z, \mathbf{x})$ can be equal to one of the unknown curves h_k, $k = 1, \cdots, M$, but the correspondence is unknown.

In addition, we are studying Functional Brain Connectivity through Auto-Regressive (AR) Models. Functional magnetic resonance imaging (fMRI) and difusion tensor imaging (DTI) are most commonly used imaging modalities to explore functional and structural connectivity patterns of brain regions. The fMRI infers brain activity by indirectly measuring changes in blood flow using magnetic resonance imaging (MRI) and the sub-modality, resting-state fMRI (rs-fMRI) is a powerful method for evaluating regional interactions that occur when a subject is not at rest (i.e., performing an explicit task).

One major challenge in resting-state fMRI(rs-fMRI) is the definition of resting-state itself because it may differ substantially across different subjects, and even for the same subjects, who may never truly be in "resting-state" during scanning. This is mostly due to self-generating thoughts, causing inevitable fluctuations in fMRI signals. Dynamic functional connectivity (dFC) identifies the functional interplay between regions of the brain and the changes that occur over a short time. The most widely used approach to identifying dFC is the sliding window approach where correlations between windowed time-courses of brain regions are estimated.

In this study, our goal is to identify the change points in the signals of rs-fMRI by learning the signal model and use this knowledge to better represent the inter-region interactions. Towards this, we use univariate and multivariate stochastic process models, Auto-Regression (AR) and Vector AR (VAR) respectively, to identify the signal change points. We propose to

use this new model to learn discriminating features/connections in the functional connectivity networks generated using the partial correlations i.e. inverse covariance matrices, using statistical methods and thereby classify functional brain networks into distinct classes.

We propose a new computational platform to better understand the patterns of R-fMRI by taking into account the challenge of inevitable signal fluctuations and interpret the success of dynamic functional connectivity approaches. We revisit an auto-regressive and vector auto-regressive signal modeling approach for estimating temporal changes of the signal in brain regions.

We generate inverse covariance matrices from the generated windows and use statistical approachs (Wilcoxon Rank Sum test,T-test and F-test) to select significant features. Finally, we use Lasso to perform classification of the data. The effectiveness of the proposed method is evidenced in the classification of R-fMRI scans of abnormal and normal control subjects. Patterns of FC networks pertaining to these two groups are more difficult to separate because the patterns of variations are expected to be due to neurocognitive differences, not from diseases.

Proposed framework with AR/VAR signal modeling and statistical feature selection tested along two real data sets and demonstrated the performance. We used Kennedy Krieger Institute (KKI) Autism and control data and Chess Masters and Novice data (CM&N) in order to demonstrate our methodology. In both data sets classification achieved higher accuracy with precision matrix comparing to correlation matrix. KKI data gives higher accuracy with the F-test feature selection while CM&N gives higher accuracy with Wilcoxon Rank Sum feature selection approach.

1.1 Dissertation Organization

Chapter 3 presents the functional Laplace deconvolution studied in [7]. In particular, after introducing some notations in Section 3.1, we present formulation of the problem and estimation methodology in Section 3.2 and 3.3, respectively. We derive the minimax lower bounds and the upper for the risk in Sections 3.4 and 3.5, respectively. In Section 3.6 , we study the finite sample properties of the proposed estimation procedure via a limited simulation study. Section 3.7 provides application of the procedure to the DCE-CT data described above. Proofs for Section 3 can be found in Appendix A.

Chapter 4 is devoted to the issue of clustering in statistical ill-posed linear inverse problems investigated in [58]. Section 4.1 describe the Notation. Section 4.2 introduces problem and its formulation and the assumptions we used in the estimation procedure . Section 4.3 addresses estimation procedure. Section 4.4 is devoted to evaluation of the error of the estimator. In particular, it provides an oracle inequality for the risk and study the minimax lower and the upper bounds for the risk under specific assumptions on the class of underlying functions. We compare the accuracy of estimation with and without clustering, theoretically in Section 4.4.4 and via a limited simulation study in Section 4.5 . Proofs can be found in Appendix B.

Chapter 5 presents the study Functional Brain Connectivity through Auto-Regressive (AR) and Vector Auto-Regressive (VAR) models. In Section 5.1 we review the fMRI imaging technique, imaging modalities and the background of Dynamic functional connectivity analysis. Section 5.2 describe Dynamic functional connectivity and present the Auto-Regressive (AR) and Vector Auto-Regressive (VAR) applicability. Section 5.3 presents the methodology used in the analysis and we explained the Auto-Regressive (AR) and Vector Auto-Regressive (VAR) models. In Section 5.4 explain Dynamic Functional Connectivity with AR/VAR and

discuss the Graphical Lasso. Finally we discuss the results and finding in section 5.5. In here first we introduce the data, then we discuss the Static Functional Connectivity based Classification and its attainment along the real data. At last we introduce Dynamic functional connectivity based classification and discuss its performance with the real data.

CHAPTER 2: LITERATURE REVIEW

2.1 Functional Laplace Deconvolution

The study is motivated by the DCE imaging problem (1.1). Due to the high level of noise in the left hand side of (1.1), a voxel-per-voxel recovery of individual curves is highly inaccurate. For this reason, the common approach is to cluster the curves for each voxel and then to average the curves in the clusters (see, e.g., Rozenholc and Reiß (2012)). As the result, one does not recover individual curves but only their cluster averages. In addition, since it is impossible to assess the clustering errors, the estimators may be unreliable even when estimation errors are small. On the other hand, the functional approaches, in particular, the wavelet-based techniques, allow to denoise a multivariate function of interest while still preserving its significant features.

Our objective is to solve the functional Laplace deconvolution problem (3.1) directly. In the case of the Fourier deconvolution problem, Benhaddou *et al.* (2013) demonstrated that the functional deconvolution solution usually has a much better precision compared to a combination of solutions of separate convolution equations. Below we adopt some of the ideas of Benhaddou *et al.* (2013) and apply them to the solution of the functional Laplace convolution equation.

Specifically, we assume that the unknown function belongs to an anisotropic Laguerre-Sobolev space and recover it using a combination of wavelet and Laguerre functions expansion. Similar to Comte *et al.* (2017), we expand the kernel g over the Laguerre basis and $f(t, \mathbf{x})$, $q(t, \mathbf{x})$ and $Y(t, \mathbf{x})$ over the Laguerre-wavelet basis and carry out denoising by thresholding the coefficients of the expansions, which naturally leads to truncation of the

infinite system of equations that results from the process.

2.2 Clustering in statistical ill-posed linear inverse problems

We consider solution of a set of general ill-posed linear inverse problems $Af_m = q_m$, $m = 1, \cdots, M$, where A is a bounded linear operator that does not have a bounded inverse and the right-hand sides q_m are measured with error. In particular, we assume that some of the curves f_m and hence, q_m are very similar to each other, so that they can be averaged and recovered together. As a result, one supposedly obtains estimators of f_j with smaller errors. The grouping is usually unknown (as well as the number of groups) and is carried out at a pre-processing step by applying one of the standard clustering techniques with the number of clusters determined by trial and error. Subsequently, the curves in the same cluster are averaged and the errors of those aggregated curves are used as true errors in the analysis.

Problems of this kind appear in many areas of application such as astronomy (blurred images), econometrics (instrumental variables), medical imaging (tomography, dynamic contrast enhanced Computerized Tomography and Magnetic Resonance Imaging), finance (model calibration of volatility) and many others where similar curves are measured and can be recovered together. Indeed, clustering has been applied to solution of ill-posed inverse problems for decades in pattern recognition [9], astronomy [62], astrophysics [32], pattern-based time series segmentation [26], medical imaging [21], elastography for computation of the unknown stiffness distribution [5] and for detecting early warning signs on stock market bubbles [44], to name a few. While in some of other settings the main objective is finding group assignments, we are considering only applications where clustering is used merely as a denoising technique. In those applications, routinely, clustering is carried out at the pre-processing step and then the inverse problems are solved for each of the cluster averages separately. As

a result, the errors of the procedures are usually examined for the estimation step only. The objective of this paper is to examine, both theoretically and via simulations, the effect of clustering on the accuracy of the solutions of general ill-posed linear inverse problems.

There exists immense literature on the statistical inverse problems (see, e.g., [2], [10], [20], [28], [52] and monographs [4], [31] and references therein, to name a few). While authors investigated the problem under some special noise scenarios (see, e.g., [70], [42], [51] among others), to the best of our knowledge, the question about the effects of clustering in the statistical inverse problems has never been investigated. Recently, as a part of a more general theory, the effect of clustering on the precision of recovery in multiple regression problems has been studed in [41]. Klopp *et al.* [41] concluded that, even under uncertainty, clustering improves the estimation accuracy.

2.3 Learning Functional Brain Connectivity through Auto-Regressive (AR) Models

A causal system is the system which depends on only the past and present values since it cannot see the future. There are several millions of neurons in the brain and these can be thought of as helping to learn patterns. The rs-fMRI signals of brain scans can also be viewed as causal signals. Similarly, in task-based fMRI scanning, the neurons learn the task being performed. There is no information about the future response, it is determined only by the past signal information. Depending upon the task being performed, different regions of the brain are activated and the response is now dependent upon the past stimulus and present with no information about the future stimulus. So, each region's response is dependent upon the past and could be influenced by the past responses of other regions of the brain. This relationship between two regions was first modeled by the Structural Equation Modeling (SEM) [48] and it was later modeled by the Granger Causality [34],[69]. Different tissues

may cause different delays in the haemodynamic response to activity. This can cause an error in correctly capturing the influence of one region on another as seen through Granger Causality.

In the resting-state, the activations are assumed to be more stable since there is no external stimulus applied. However, there still exist delays in response as well as some uncharacteristic stronger activations which may or may not have been influenced by other regions. These uncharacteristic activations could be due to the fact that the brain is almost never truly at rest ($dreams, etc$) and also to the signal noise. In the preprocessing stage of rs-fMRI analysis, a signal filtering is performed to select responses in a particular frequency range. This step can help to eliminate some of the noise but not all of it. Recently, Khazaee et. al [40] used multivariate Granger Causality to model the overall signal and compute a directed graph to discriminate between Alzheimers and healthy patients. However, the change points were not modeled here.

2.3.1 Functional Connectivity (FC)

Pearson's correlation is the commonly used metric to generate FC matrices from the rs-fMRI signals [35]. More recently, inverse covariance based FC approaches have been shown to be more accurate in classifying signals [50]. In most works, identification of these change points in the signal to determine the window-size has proven challenging with most methods preferring to use a trial and error approach to identifying window size that gives the best modeling performance [45], [72]. Within a window, all time points are given equal weight thereby boosting outlier contribution/effect. To counter this, [3] proposed using tapered windows to discount boundary observations. [22] proposed a regression approach to identify change-points in the signal by iteratively partitioning the signal till no improvement in their

criterion was found. This approach generates different width partitions which can prove challenging when used in the more modern deep-learning approaches. The game of chess requires complex problem solving skills [6] and this cognitive expertise can be studied. [30] hypothesized functional connectivity differences in the intrinsic brain connectivity between chess masters and novice players. Network-based statistics were used to identify differences in the medial temporal lobe and several temporal and parietal areas from the connectivity matrices generated using Pearson's correlation.

CHAPTER 3: FUNCTIONAL LAPLACE DECONVOLUTION

3.1 Notation

In this section, we are going to use the following notations. Given a matrix $\mathbf{A}$, let $\mathbf{A}^T$ be the transpose of $\mathbf{A}$, $\|\mathbf{A}\|_F = \sqrt{\mathrm{Tr}(\mathbf{A}^T\mathbf{A})}$ and $\|\mathbf{A}\| = \lambda_{\max}(\mathbf{A}^T\mathbf{A})$ be, respectively, the Frobenius and the spectral norm of a matrix $\mathbf{A}$, where $\lambda_{\max}(\mathbf{U})$ is the largest, in absolute value, eigenvalue of $\mathbf{U}$. We denote by $[\mathbf{A}]_m$ the upper left $m \times m$ sub-matrix of $\mathbf{A}$. Given a vector $\mathbf{u} \in \mathbb{R}^k$, we denote by $\|\mathbf{u}\|$ its Euclidean norm and, for $p \leq k$, the $p \times 1$ vector with the first p coordinates of $\mathbf{u}$, by $[\mathbf{u}]_p$. For any function $t \in L_2(\mathbb{R}_+)$, we denote by $\|t\|_2$ its L_2 norm on $\mathbb{R}_+$. For vectors, whenever it is necessary, we use the superscripts to indicate dimensions of the vectors and subscripts to denote their components. Also, $a \vee b = \max(a, b)$ and $a \wedge b = \min(a, b)$.

3.2 Formulation of the problem

In this section, we consider the problem of estimating a three-dimensional function f based on observations from its noisy Laplace convolution.

$$Y(t, \mathbf{x}) = q(t, \mathbf{x}) + \epsilon\xi(t, \mathbf{x}) \quad \text{with} \quad q(t, \mathbf{x}) = \int_0^t g(t - z) f(z, \mathbf{x}) dz. \tag{3.1}$$

where $\mathbf{x} = (x_1, x_2)$, $(t, x_1, x_2) \in U = [0, \infty) \times [0, 1] \times [0, 1]$ and $\xi(z, x_1, x_2)$ is the three-dimensional Gaussian white noise such that

$$\mathrm{Cov}\left\{\xi(z_1, x_{11}, x_{12}), \xi(z_2, x_{21}, x_{22})\right\} = \mathbb{I}(z_1 = z_2)\, \mathbb{I}(x_{11} = x_{21})\, \mathbb{I}(x_{12} = x_{22}).$$

Formula (3.1) represents a noisy version of a functional Laplace convolution equation. Indeed, if $\mathbf{x}$ is fixed, then (3.1) reduces to a noisy version of the Laplace convolution equation

$$Y(t) = q(t) + \epsilon\xi(t) \quad \text{with} \quad q(t) = \int_0^t g(t-z)f(z)dz, \tag{3.2}$$

Observe that when g is known, equation (1.1) reduces to the form (3.1) that we study in the present paper. If one is interested in taking the uncertainty about g into account, this can be accomplished using methodology of Vareschi (2015).

We consider the functional version (3.1) of the Laplace convolution equation (3.2). The study is motivated by the DCE imaging problem (1.1). Due to the high level of noise in the left hand side of (1.1), a voxel-per-voxel recovery of individual curves is highly inaccurate. For this reason, the common approach is to cluster the curves for each voxel and then to average the curves in the clusters (see, e.g., Rozenholc and Reiß (2012)). As the result, one does not recover individual curves but only their cluster averages. In addition, since it is impossible to assess the clustering errors, the estimators may be unreliable even when estimation errors are small. On the other hand, the functional approaches, in particular, the wavelet-based techniques, allow to denoise a multivariate function of interest while still preserving its significant features.

The objective is to solve the functional Laplace deconvolution problem (3.1) directly. In the case of the Fourier deconvolution problem, Benhaddou *et al.* (2013) demonstrated that the functional deconvolution solution usually has a much better precision compared to a combination of solutions of separate convolution equations. Below we adopt some of the ideas of Benhaddou *et al.* (2013) and apply them to the solution of the functional Laplace convolution equation.

Specifically, we assume that the unknown function belongs to an anisotropic Laguerre-Sobolev space and recover it using a combination of wavelet and Laguerre functions expansion. Similar to Comte *et al.* (2017), we expand the kernel g over the Laguerre basis and $f(t,\mathbf{x})$, $q(t,\mathbf{x})$ and $Y(t,\mathbf{x})$ over the Laguerre-wavelet basis and carry out denoising by thresholding the coefficients of the expansions, which naturally leads to truncation of the infinite system of equations that results from the process. We derive the minimax lower bounds for the L^2-risk in the model (3.1) and demonstrate that the wavelet-Laguerre estimator is adaptive and asymptotically near-optimal within a logarithmic factor in a wide range of Laguerre-Sobolev balls. We carry out a limited simulation study and then finally apply our technique to recovering of $f(z,\mathbf{x})$ in equation (1.1) on the bases of DCE-CT data. Although, for simplicity, we only consider the white noise model for the functional Laplace convolution equation (3.1), the theoretical results can be easily extended to its sample version by following Comte *et al.* (2017). We carry out a limited simulations study and show that the estimator performs well in a finite sample setting. Finally, we use the technique for the solution of the Laplace deconvolution problem on the basis of DCE Computerized Tomography data.

3.3 Estimation Algorithm

Consider a finitely supported periodized r_0-regular wavelet basis (e.g., Daubechies) $\psi_{j,k}(x)$ on $[0,1]$. Form a product wavelet basis $\Psi_{\boldsymbol{\omega}}(\mathbf{x}) = \psi_{j_1,k_1}(x_1)\psi_{j_2,k_2}(x_2)$ on $[0,1]\times[0,1]$ where $\boldsymbol{\omega}\in\Omega$ with

$$\Omega = \left\{\boldsymbol{\omega} = (j_1,k_1;j_2,k_2) : j_1,j_2 = 0,\cdots,\infty;\ k_1 = 0,\cdots,2^{j_1-1},\ k_2 = 0,\cdots,2^{j_2-1}\right\}. \quad (3.3)$$

Denote functional wavelet coefficients of $f(t,\mathbf{x})$, $q(t,\mathbf{x})$, $Y(t,\mathbf{x})$ and $\xi(t,\mathbf{x})$ by, respectively, $f_{\boldsymbol{\omega}}(t)$, $q_{\boldsymbol{\omega}}(t)$, $Y_{\boldsymbol{\omega}}(t)$ and $\xi_{\boldsymbol{\omega}}(t)$. Then, for any $t \in [0,\infty)$, equation (4.1) yields

$$Y_{\boldsymbol{\omega}}(t) = q_{\boldsymbol{\omega}}(t) + \epsilon\xi_{\boldsymbol{\omega}}(t) \quad \text{with} \quad q_{\boldsymbol{\omega}}(t) = \int_0^t g(t-s) f_{\boldsymbol{\omega}}(s) ds \tag{3.4}$$

and function $f(t,\mathbf{x})$ can be written as

$$f(t,\mathbf{x}) = \sum_{\boldsymbol{\omega}\in\Omega} f_{\boldsymbol{\omega}}(s)\Psi_{\boldsymbol{\omega}}(\mathbf{x}), \quad \mathbf{x} = (x_1, x_2). \tag{3.5}$$

Now, consider the orthonormal basis that consists of a system of Laguerre functions

$$\varphi_l(t) = e^{-t/2} L_l(t), \quad l = 0, 1, 2, \ldots, \tag{3.6}$$

where $L_l(t)$ are Laguerre polynomials (see, e.g., Gradshtein and Ryzhik (1980), Section 8.97)

$$L_l(t) = \sum_{j=0}^{l} (-1)^j \binom{l}{j} \frac{t^j}{j!}, \quad t \geq 0.$$

It is known that functions $\varphi_l(\cdot)$, $l = 0, 1, 2, \ldots$, form an orthonormal basis of the $L^2(0,\infty)$ space and, therefore, functions $f_{\boldsymbol{\omega}}(\cdot)$, $g(\cdot)$, $q_{\boldsymbol{\omega}}(\cdot)$ and $Y_{\boldsymbol{\omega}}(\cdot)$ can be expanded over this basis with coefficients $\theta_{l;\boldsymbol{\omega}}$, g_l, $q_{l;\boldsymbol{\omega}}$ and $Y_{l;\boldsymbol{\omega}}$, $l = 1, \ldots, \infty$, respectively. By plugging these expansions into formula (3.4), we obtain the following equation

$$\sum_{l=0}^{\infty} q_{l;\boldsymbol{\omega}}\, \varphi_l(t) = \sum_{l=0}^{\infty}\sum_{k=0}^{\infty} \theta_{l;\boldsymbol{\omega}}\, g_k \int_0^t \varphi_k(t-s)\varphi_l(s) ds. \tag{3.7}$$

Following Comte *et al.* (2017), for each $\boldsymbol{\omega} \in \Omega$, we represent coefficients of interest $\theta_{l;\boldsymbol{\omega}}$, $l = 0, 1, \ldots$, as a solution of an infinite triangular system of linear equations. Indeed, it is

easy to check that (see, e.g., 7.411.4 in Gradshtein and Ryzhik (1980))

$$\int_0^t \phi_k(x)\phi_j(t-x)dx = e^{-t/2}\int_0^t L_k(x)L_j(t-x)dx = \phi_{k+j}(t) - \phi_{k+j+1}(t).$$

Hence, equation (3.7) can be re-written as

$$\sum_{k=0}^{\infty} q_{k;\boldsymbol{\omega}}\,\varphi_k(t) = \sum_{k=0}^{\infty}\left[\theta_{k;\boldsymbol{\omega}}\;g_0 + \sum_{l=0}^{k-1}(g_{k-l} - g_{k-l-1})\,\theta_{l;\boldsymbol{\omega}}\right]\;\varphi_k(t).$$

Equating coefficients for each basis function, we obtain an infinite triangular system of linear equations. In order to use this system for estimating f, we choose a fairly large M and define the following approximations of f and q based on the first M Laguerre functions

$$f_M(t,\mathbf{x}) = \sum_{\boldsymbol{\omega}\in\Omega}\sum_{l=0}^{M-1}\theta_{l;\boldsymbol{\omega}}\varphi_l(t)\Psi_{\boldsymbol{\omega}}(\mathbf{x}), \quad q_M(t,\mathbf{x}) = \sum_{\boldsymbol{\omega}\in\Omega}\sum_{l=0}^{M-1} q_{l;\boldsymbol{\omega}}\varphi_l(t)\Psi_{\boldsymbol{\omega}}(\mathbf{x}). \tag{3.8}$$

Let $\boldsymbol{\theta}_{\omega}^{(M)}$, $\mathbf{g}^{(M)}$ and $\mathbf{q}_{\omega}^{(M)}$ be M-dimensional vectors with elements $f_{l;\boldsymbol{\omega}}$, g_l and $q_{l;\boldsymbol{\omega}}$, $l = 0,1,\ldots,M-1$, respectively. Then, for any M and any $\boldsymbol{\omega}\in\Omega$, one has $\mathbf{q}_{\omega}^{(M)} = \mathbf{G}^{(M)}\boldsymbol{\theta}_{\omega}^{(M)}$ where $\mathbf{G}^{(M)}$ is the lower triangular Toeplitz matrix with elements $G_{i,j}^{(M)}$, $0\le i,j\le M-1$

$$G_{i,j}^{(M)} = \begin{cases} g_0, & \text{if } \; i=j, \\ (g_{i-j} - g_{i-j-1}), & \text{if } \; j<i, \\ 0, & \text{if } \; j>i. \end{cases} \tag{3.9}$$

In order to recover f in (4.1), we estimate coefficients $q_{l;\boldsymbol{\omega}}$ in (3.8) by

$$\widehat{q}_{l;\boldsymbol{\omega}} = \int_0^{\infty} Y_{\boldsymbol{\omega}}(t)\,\varphi_l(t)\,dt, \quad l=0,2,\ldots, \tag{3.10}$$

and obtain an estimator $\widehat{\boldsymbol{\theta}_{\boldsymbol{\omega}}^{(M)}}$ of vector $\boldsymbol{\theta}_{\boldsymbol{\omega}}^{(M)}$ of the form

$$\widehat{\boldsymbol{\theta}_{\boldsymbol{\omega}}^{(M)}} = (\mathbf{G}^{(M)})^{-1}\widehat{\mathbf{q}_{\boldsymbol{\omega}}^{(M)}}. \tag{3.11}$$

Denote by $\Omega(J_1, J_2)$ a truncation of a set Ω in (3.3):

$$\Omega(J_1, J_2) = \left\{\boldsymbol{\omega} = (j_1, k_1; j_2, k_2) : 0 \le j_i \le J_i - 1,\ k_i = 0, \cdots, 2^{j_i - 1}; i = 1, 2\right\}. \tag{3.12}$$

If we recovered f from all its coefficients $\widehat{\boldsymbol{\theta}_{\boldsymbol{\omega}}^{(M)}}$ with $\boldsymbol{\omega} \in \Omega(J_1, J_2)$, the estimator would have a very high variance. For this reason, we need to remove the coefficients that are not essential for representation of f. This is accomplished by constructing a hard thresholding estimator for the function $f(t, \mathbf{x})$

$$\widehat{f}(t, \mathbf{x}) = \sum_{l=0}^{M-1} \sum_{\boldsymbol{\omega} \in \Omega(J_1, J_2)} \widehat{\theta}_{l;\boldsymbol{\omega}}\, \mathbb{I}\left(|\widehat{\theta}_{l;\boldsymbol{\omega}}| > \lambda_{l,\varepsilon}\right)\, \varphi_l(t)\, \Psi_{\boldsymbol{\omega}}(\mathbf{x}) \tag{3.13}$$

where the values of J_1, J_2, M and $\lambda_{l,\varepsilon}$ will be defined later.

3.4 Minimax lower bounds for the risk.

In order to determine the values of parameters J_1, J_2, M and $\lambda_{l,\varepsilon}$, and to gauge the precision of the estimator $\widehat{f}$, we need to introduce some assumptions on the function g. Let $r \ge 1$ be such that

$$\left.\frac{d^j g(t)}{dt^j}\right|_{t=0} = \begin{cases} 0, & \text{if } \ j = 0, ..., r-2, \\ B_r \neq 0, & \text{if } \ j = r-1, \end{cases} \tag{3.1}$$

with the obvious modification $g(0) = B_1 \neq 0$ for $r = 1$. We assume that function $g(x)$ and its Laplace transform $G(s) = \int_0^\infty e^{-sx} g(x) dx$ satisfy the following conditions:

Assumption A1. $g \in L_1[0,\infty)$ is r times differentiable with $g^{(r)} \in L_1[0,\infty)$.

Assumption A2. Laplace transform $G(s)$ of g has no zeros with nonnegative real parts except for zeros of the form $s = \infty + ib$.

Assumptions 1 and 2 are difficult to check since their verification relies on the exact knowledge of g and the value of r. Therefore, in the present paper, we do not use the value of r in our estimation algorithm and aim at construction of an adaptive estimator that delivers the best convergence rates possible for the true unknown value of r without its knowledge. Hence, we need to derive what is the smallest error that any estimator of f can attain under Assumptions A1 and A2.

For this purpose, we consider the generalized three-dimensional Laguerre-Sobolev ball of radius A, characterized by its wavelet-Laguerre coefficients $\theta_{l;\boldsymbol{\omega}} = \theta_{l;j_1,j_2,k_1,k_2}$ as follows:

$$\mathcal{B}^{s_1,s_2,s_3}_{\gamma,\beta}(A) = \left\{ f : \sum_{l=0}^{\infty}\sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty} 2^{2js_1+2j's_2}(l \vee 1)^{2s_3} \exp\left(2\gamma\, l^{\beta}\right) \sum_{k_1=0}^{2^{j_1}-1}\sum_{k_2=0}^{2^{j_2}-1} \theta^2_{l;\boldsymbol{\omega}} \leq A^2 \right\}, \quad (3.2)$$

where we assume that $\beta = 0$ if $\gamma = 0$ and $\beta > 0$ if $\gamma > 0$. In order to construct minimax lower bounds, we define the maximum L^2-risk over the set V of an estimator $\tilde{f}$ as

$$R_{\varepsilon}(\tilde{f}, V) = \sup_{f \in V} \mathbb{E}\|\tilde{f} - f\|^2. \quad (3.3)$$

Denote

$$\Delta(s_1,s_2,s_3,\gamma,\beta,A)=\begin{cases} A^2\ [A^{-2}\epsilon^2]^{\frac{2s_1}{2s_1+1}}, & \text{if } \ s_1\le \min(s_2,s_3/(2r)),\ \gamma=\beta=0 \\ A^2\ [A^{-2}\epsilon^2]^{\frac{2s_2}{2s_2+1}}, & \text{if } \ s_2\le \min(s_1,s_3/(2r)),\ \gamma=\beta=0 \\ A^2\ [A^{-2}\epsilon^2]^{\frac{2s_3}{2s_3+2r}}, & \text{if } \ s_3\le \min(2rs_1,2rs_2),\ \gamma=\beta=0 \\ A^2\ [A^{-2}\epsilon^2]^{\frac{2s_1}{2s_1+1}}, & \text{if } \ s_1\le s_2,\ \gamma>0,\beta>0 \\ A^2\ [A^{-2}\epsilon^2]^{\frac{2s_2}{2s_2+1}}, & \text{if } \ s_2\le s_1,\ \gamma>0,\beta>0 \end{cases} \tag{3.4}$$

The following theorem gives the minimax lower bounds for the L^2-risk of any estimator $\tilde{f}$ of f.

Theorem 1 *Let* $\min\{s_1,s_2\}\ge 1/2$ *and* $s_3\ge 1/2$ *if* $\gamma=\beta=0$. *Then, if* ϵ, *is small enough, under Assumptions A1 and A2, for some absolute constant* $\underline{C}>0$ *independent of* ϵ, *one has*

$$\inf_{\tilde{f}}\ R_\varepsilon(\tilde{f},\ \mathcal{B}^{s_1,s_2,s_3}_{\gamma,\beta}(A))\ge \underline{C}\,\Delta(s_1,s_2,s_3,\gamma,\beta,A) \tag{3.5}$$

3.5 Upper bounds for the risk.

In order to derive an upper bound for $R_\varepsilon(\widehat{f},\ \mathcal{B}^{s_1,s_2,s_3}_{\gamma,\beta}(A))$, we need some auxiliary statements. Consider $\mathbf{G}^{(m)}$, the lower triangular Toeplitz matrix defined by formula (3.9) with $M=m$. The following results follows directly from Comte *et al.* (2017) and Vareschi (2015).

Lemma 1 (Lemma 4, Comte *et al.* **(2017), Lemma 5.4, Vareschi (2015)).** *Let conditions A1 and A2 hold. Denote the elements of the last row of matrix* $(\mathbf{G}^{(m)})^{-1}$ *by* $v_j,\ j=1,\cdots,m$. *Then, there exist absolute positive constants* C_{G1}, C_{G2}, C_{v1} *and* C_{v2}

independent of m *such that*

$$C_{G1}m^{2r} \leq \|(\mathbf{G}^{(m)})^{-1}\|^2 \leq \|(\mathbf{G}^{(m)})^{-1}\|_F^2 \leq C_{G2}m^{2r}, \tag{3.1}$$

$$C_{\upsilon 1}m^{2r-1} \leq \sum_{j=1}^{m} \upsilon_j^2 \leq C_{\upsilon 2}m^{2r-1}. \tag{3.2}$$

Using Lemma 1, one can obtain the following upper bounds for the errors of estimators $\widehat{\theta}_{l;\boldsymbol{\omega}}$:

Lemma 2 *Let* $\widehat{\theta}_{l;\boldsymbol{\omega}}$ *be the* $l-th$ *element of the vector* $\widehat{\boldsymbol{\theta}_{\boldsymbol{\omega}}^{(M)}}$ *defined in (3.11). Then, under the Assumptions A1 and A2, one has*

$$Var\left[\widehat{\theta}_{l;\boldsymbol{\omega}}\right] \leq C_{\upsilon 2}\ \varepsilon^2\, l^{2r-1} \tag{3.3}$$

$$\mathbb{E}\left[\widehat{\theta}_{l;\boldsymbol{\omega}} - \theta_{l;\boldsymbol{\omega}}\right]^4 \leq 3C_{\upsilon 2}^2\ \varepsilon^4\, l^{4r-2} \tag{3.4}$$

$$\Pr\left(|\widehat{\theta}_{l;\boldsymbol{\omega}} - \theta_{l;\boldsymbol{\omega}}| > \epsilon\sqrt{2\nu\log(\epsilon^{-1})\ l^{-1}}\ \|(\mathbf{G}^{(l)})^{-1}\|\right) \leq \epsilon^{\tau} \tag{3.5}$$

provided $\nu \geq \tau C_{\upsilon 2}/C_{G1}$ *where* C_{G1} *and* $C_{\upsilon 2}$ *are defined in (3.1) and (3.2), respectively.*

Following Lemma 2 we choose J_1, J_2, M and $\lambda_{l,\varepsilon}$ such that

$$2^{J_1} = 2^{J_2} = A^2\epsilon^{-2}, \quad M = \max\left\{m \geq 1 :\ \|(\mathbf{G}^{(m)})^{-1}\| \leq \epsilon^{-2}\right\}, \tag{3.6}$$

and thresholds $\lambda_{l,\varepsilon}$ of the forms

$$\lambda_{l,\varepsilon} = 2\epsilon\ \sqrt{2\,\nu\log(\epsilon^{-1})\ l^{-1}}\ \|(\mathbf{G}^{(l)})^{-1}\|, \tag{3.7}$$

where the value of ν is large enough, so that it satisfies the inequality

$$\nu \geq 12C_{v2}/C_{G1}, \tag{3.8}$$

and C_{v2} and C_{G1} and C_{v2} are defined in (3.1) and (3.2), respectively. Then, the following statement holds.

Theorem 2 *Let* $\min\{s_1, s_2\} \geq 1/2$ *and* $s_3 \geq 1/2$ *if* $\gamma = \beta = 0$. *Let* $\widehat{f}(t, \mathbf{x})$ *be the wavelet-Laguerre estimator defined in (3.13), with* J_1, J_2 *and* M *given by (3.6). Let* $A > 0$, *and let condition (3.2) hold. If* ν *in (3.7) satisfies inequality (3.8), then, under Assumptions A1 and A2, if* ϵ, *is small enough, for some absolute constant* $\overline{C} > 0$ *independent of* ϵ, *one has*

$$R_{\widehat{f},\ \varepsilon}(B^{s_1,s_2,s_3}_{\gamma,\beta}(A)) \leq \overline{C}\, \Delta(s_1, s_2, s_3, \gamma, \beta, A)\ [\log(1/\epsilon)]^d \tag{3.9}$$

where $\Delta = \Delta(s_1, s_2, s_3, \gamma, \beta, A)$ *is defined in (3.4) and*

$$d = \begin{cases} 2s_1/(2s_1+1) + \mathbb{I}(s_1 = s_2) + \mathbb{I}(s_3 = 2rs_1), & \text{if } \ s_1 \leq \min(s_2, \frac{s_3}{2r}),\ \gamma = \beta = 0 \\ 2s_2/(2s_2+1) + \mathbb{I}(s_1 = s_2) + \mathbb{I}(s_3 = 2rs_2), & \text{if } \ s_2 \leq \min(s_1, \frac{s_3}{2r}),\ \gamma = \beta = 0 \\ 2s_3/(2s_3+2r) + \mathbb{I}(s_3 = 2rs_1) + \mathbb{I}(s_3 = 2rs_2), & \text{if } \ s_3 \leq \min(2rs_1, 2rs_2),\ \gamma = \beta = 0 \\ 2s_1/(2s_1+1) + \mathbb{I}(s_1 = s_2), & \text{if } \ s_1 \leq s_2,\ \gamma > 0, \beta > 0 \\ 2s_2/(2s_2+1) + \mathbb{I}(s_1 = s_2), & \text{if } \ s_2 \leq s_1,\ \gamma > 0, \beta > 0 \end{cases}$$

3.6 Simulation Studies

In order to study finite sample properties of the proposed estimation procedure, we carried out a limited simulation study. For each test function $f(t, \mathbf{x})$ and a kernel $g(t)$, we obtained exact values of $q(t, \mathbf{x})$ in the equation (4.1) by integration. We considered n equally spaced

points $t_k = Tk/n$, $k = 1, \cdots, n$, on the time interval $[0; T]$. We created a uniform grid $\{x_{1,i}, x_{2,j}\}$ on $[0,1] \times [0,1]$ with $i = 1, \cdots, n_1$ and $j = 1, \cdots, n_2$, and obtained the three-dimensional array of values $q(x_{1,i}, x_{2,j}, t_k)$. Finally, we obtained a sample $Y_{i,j,k}$ of the left-hand side of the equation (4.1) by adding independent Gaussian $\mathbb{N}(0, \sigma^2)$ noise to each value $q(x_{1,i}, x_{2,j}, t_k)$, $i = 1, \cdots, n_1$, $j = 1, \cdots, n_2$, $k = 1, \cdots, n$.

Next, we constructed a system of M Laguerre functions of the form (3.6). We obtained an estimator $\widehat{f}$ of the form (3.13) with the thresholds $\lambda_{l,\widehat{\epsilon}}$, $l = 0, \cdots, M-1$, given by (3.7) where, by Abramovich *et al.* (2013), $\widehat{\epsilon} = T\widehat{\sigma}/\sqrt{n}$ and $\widehat{\sigma}$ is estimated by the standard deviations of the wavelet coefficients at the highest resolution level.

In our simulations, we used $n_1 = n_2 = n = 32$, $M = 8$ and $T = 5$. We chose $g(x) = \exp(-x/2)$ and carried out simulations with the following test functions

$$\begin{aligned} f_1(t, \mathbf{x}) &= t\, e^{-t}(x_1 - 0.5)^2)\, (x_2 - 0.5)^2, \\ f_2(t, \mathbf{x}) &= e^{-t/2}\, \cos(2\pi x_1 x_2), \\ f_3(t, \mathbf{x}) &= t\, e^{-t}(x_1 - 0.5)^2)\, (x_2 - 0.5)^2 + e^{-t/2}\, \cos(2\pi x_1 x_2), \\ f_4(t, \mathbf{x}) &= e^{-t/2}\, \cos(2\pi x_1 x_2) + (x_1 - 0.5)^2\, (x_2 - 0.5)^2. \end{aligned} \tag{3.1}$$

We also considered three noise scenarios: SNR = 3 (high noise level), SNR = 5 (medium noise level) and SNR = 7 (low noise level). In order for the values of the errors of our estimators to be independent of the norms of the test functions, we evaluated the average relative error as the average L^2-norm of the difference between f and its estimator divided by the norm of f:

$$\Delta(\widehat{f}) = \|\widehat{f} - f\|/\|f\|.$$

Table 3.1 reports the mean values of those errors over 100 simulation runs (with the standard

error of the means presented in parentheses) for the four test functions and the three noise levels.

3.7 Real Data Example

As an application of the proposed technique we studied the recovery of the unknown function $f(t, \mathbf{x}) = \beta(1 - F(t, \mathbf{x}))$ in the equation (1.1) on the basis of the DCE-CT (Computerized Tomography) images of a participant of the REMISCAN cohort study [56] who underwent anti-angiogenic treatment for renal cancer. The data consist of the arterial images and images of the area of interest (AOI) at 37 time points over approximately 4.6 minute interval. The first 15 time points (approximately the first 30 seconds) correspond to the time period before the contrast agent reached the aorta and the AOI (so $\delta = 0$ in equation (1.1)). We used those data points for the evaluation of the base intensity. Since the images of the aorta are extremely noisy, we evaluated the average values of the grey level intensity at each time point and then used Laguerre functions smoothing in order to obtain the values of the Arterial Input Function AIF(t). The images of AOI contain 49×38 pixels. Since our technique is based on periodic wavelets and hence application of the method to a non-periodic function is likely to produce Gibbs effects, we cut the images to the size of 32×32 pixels. In order to achieve periodicity, we obtained symmetric versions of the the images (reflecting the images over the two sides) and applied our methodology to the resulting spatially periodic functions. Consequently, the estimator obtained by the technique is spatially symmetric, so we record only the original part as the estimator $\widehat{f}$. Figure 3.1 shows the averages of the aorta intensities at each time point and its de-noised version that was used as AIF(t). Figure 3.2 presents the values of $\widehat{f}$ at 34 seconds (corresponds to the first time point), 95 seconds (the 12-th time point) and 275 seconds (the last time point).

Table 3.1: The average values of the relative errors $\Delta(\widehat{f})$ (with the standard errors of the means in parentheses) evaluated over 100 simulation runs. The test functions are defined in formula (3.1).

Function	SNR=3	SNR=5	SNR=7
$f_1(t, \mathbf{x})$	0.1107 (0.0110)	0.0694 (0.0066)	0.0511 (0.0049)
$f_2(t, \mathbf{x})$	0.1224 (0.0100)	0.0761 (0.0071)	0.0567 (0.0051)
$f_3(t, \mathbf{x})$	0.1107 (0.0112)	0.0680 (0.0068)	0.0511 (0.0048)
$f_4(t, \mathbf{x})$	0.1080 (0.0117)	0.0690 (0.0058)	0.0519 (0.0046)

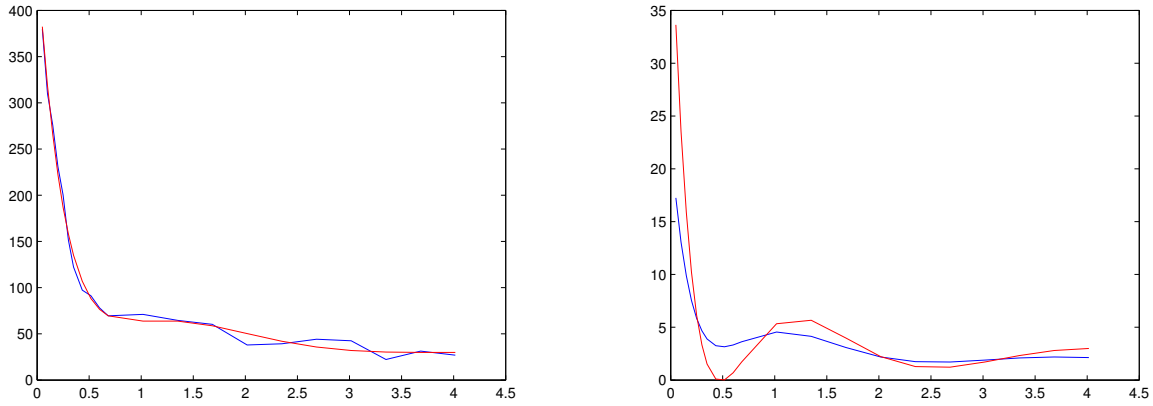

Figure 3.1: Left: the averages of the aorta intensities (blue) and the estimated Arterial Input Function AIF(t) (red). Right: two curves for distinct spatial locations.

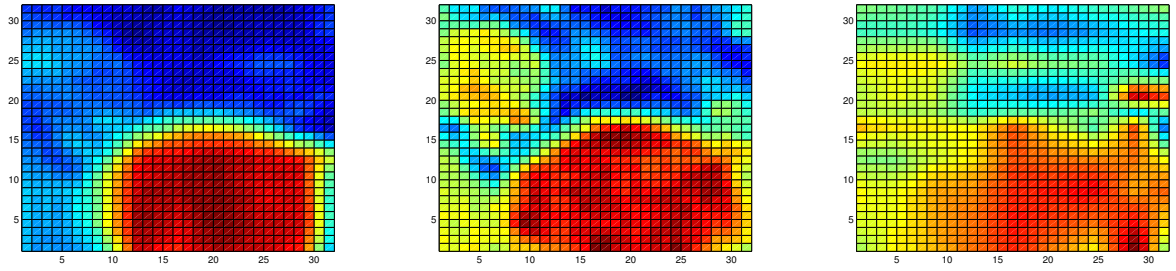

Figure 3.2: The values of $\widehat{f}$ at 34 seconds (corresponds to the first time point), 95 seconds (the 12th time point) and 275 seconds (the last time point).

CHAPTER 4: CLUSTERING IN STATISTICAL ILL-POSED LINEAR INVERSE PROBLEMS

4.1 Notations

We denote $[m] = \{1, \cdots, m\}$. We denote vectors and matrices by bold letters. For any vector $\mathbf{a}$, we denote its l_2- norm by $\|\mathbf{a}\|$ and the l_0 norm, the number of non zero elements by $\|\mathbf{a}\|_0$. For any matrix $\mathbf{A}$, we denote its Frobenius norm by $\|\mathbf{A}\|_F$ and the operator norm by $\|\mathbf{A}\|_{op}$ and the span of the column space of matrix $\mathbf{A}$ by Span($\mathbf{A}$). We denote the Hamming distance between matrices $\mathbf{A}_1$ and $\mathbf{A}_2$, the number of nonzero elemnts in $\mathbf{A}_1 - \mathbf{A}_2$, by $\|\mathbf{A}_1 - \mathbf{A}_2\|_H$. We denote the $k \times k$ identity matrix by $\mathbf{I}_k$ and drop subscript k when there is no uncertainty about the dimension. We denote the inner product and the corresponding norm in a Hilbert space $\mathcal{H}$ by $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and $\| \cdot \|_{\mathcal{H}}$, respectively, and drop subscript $\mathcal{H}$ whenever there is no ambiguity. For any set S, we denote cardinality of S by $|S|$. We denote the set of all clustering matrices for grouping M objects into K classes by $\mathcal{M}(M, K)$. We denote $a_n b_n$ if there exist $c < \infty$ independent of n such that $a_n \leq c b_n$ and $a_n b_n$ if there exist $c > 0$ independent of n such that $a_n \geq c b_n$. Also, $a_n \asymp b_n$ if simultaneously $a_n b_n$ and $a_n b_n$. Finally, we use C for a generic absolute constant independent of n, M and K, which can take different values in different places.

4.2 Formulation of the problem and assumptions

In this section, we consider solution of a set of general ill-posed linear inverse problems $Af_m = q_m,\ m = 1, \cdots, M$, where A is a bounded linear operator that does not have a bounded inverse and the right-hand sides q_m are measured with error. In particular, we

assume that some of the curves f_m and hence, q_m are very similar to each other, so that they can be averaged and recovered together. As a result, one supposedly obtains estimators of f_j with smaller errors. The grouping is usually unknown (as well as the number of groups) and is carried out at a pre-processing step by applying one of the standard clustering techniques with the number of clusters determined by trial and error. Subsequently, the curves in the same cluster are averaged and the errors of those aggregated curves are used as true errors in the analysis.

In particular, we consider the following problem. Let $A : \mathcal{H}_1 \to \mathcal{H}_2$ be a known linear operator where $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces with inner products $\langle \cdot, \cdot \rangle_{\mathcal{H}_1}$ and $\langle \cdot, \cdot \rangle_{\mathcal{H}_2}$, respectively. The objective is to recover functions $f_m \in \mathcal{H}_1$ from

$$X_m(x) = q_m(x) + \sigma n^{-1/2} \epsilon_m(x), \quad q_m = Af_m, \quad m = 1, \cdots, M, \tag{4.1}$$

where $\epsilon_m(x)$ are the independent white noise processes and the goal is to recover the vector function $f = (f_1, \cdots, f_M)$. Assume that observations are taken as functionals of X_m: for any $\psi \in \mathcal{H}_2$

$$\langle X_m, \psi \rangle = \langle Af_m, \psi \rangle + \sigma n^{-1/2} \xi_m(\psi), \tag{4.2}$$

where $\xi_m(\psi)$ are Gaussian random variables with zero means such that

$$\mathbb{E}[\xi_m(\psi_1), \xi_l(\psi_2)] = \begin{cases} \langle \psi_1, \psi_2 \rangle_{\mathcal{H}_2}, & m = l \\ 0, & m \neq l \end{cases} \tag{4.3}$$

In formula (4.2), σ can be viewed as noise level and n as the number of observations.

In what follows we assume that, although M is large, there are only K types of functions $f_m(t)$. In particular, there exists a collection of functions $h_1(t), ..., h_K(t)$ such that $f_m(t) =$

$h_k(t)$ for any m and some $k = z(m)$. In other words, one can define a clustering function $z = z(m)$, $m = 1, \ldots, M$, with values in $\{1, \ldots, K\}$ such that $f_m = h_{z(m)}$. We denote the clustering matrix corresponding to the clustering function $z(m)$ by $\mathbf{Z}$. Note that $\mathbf{Z} \in \{0,1\}^{M \times K}$ and $\mathbf{Z}_{m,k} = 1$ if and only if $z(m) = k$. Hence,

$$\mathbf{Z}^T \mathbf{Z} = \mathbf{D}^2 = \mathrm{diag}(N_1, \cdots, N_K), \tag{4.4}$$

where N_k is the number of functions in cluster k, $k = 1, \cdots, K$.

If we knew the function $z(m)$, we could improve precision of estimating f_m by averaging the signals within clusters,thus, reducing the noise levels, and construct the estimators $\hat{h}_k$ of the common cluster means, subsequently setting $\hat{f}_m = \hat{h}_{z(m)}$. In reality, however, neither the true clustering matrix $\mathbf{Z}_*$, nor the true number of classes K_* are unavailable, so they also need to be estimated.

Note that our objective is accurate estimation of functions f_m, $m = 1, \cdots, M$, rather than recovery of the clustering matrix $\mathbf{Z}$. Moreover, although a true clustering matrix $\mathbf{Z}_*$ always exists (if all functions f_m are different, one can choose $K_* = M$ and $\mathbf{Z}_* = \mathbf{I}_M$), we are not interested in finding $\mathbf{Z}_*$: we would rather incur a small bias resulting from replacement of f_m by $h_k \approx f_m$ than obtain estimators with high variances that are common in inverse problems where each function f_m is estimated separately. On the other hand, while using the clustering procedure, we gather one more type of errors that are due to erroneously pooling together estimators of functions f_m that belong to different classes, i.e., the errors due to mistakes in clustering.

One of the advantages of our estimation procedure is that we do not assume that the number of clusters is known in advance. Instead, we elicit the unknown number of clusters, the

clustering matrix and the estimators of the unknown functions as a solution of a penalized optimization problem. We conclude that clustering does not have an adverse effect on the estimation precision as long as class sizes and the number of observations are large enough. However, significant improvement in accuracy occurs only if the problem is not severely ill-posed.

In order to simplify the problem (4.1), we assume that it allows reduction to the so-called sequence model. In particular, following Donoho (1995), we assume that there exists an orthonormal basis ϕ_j, $j = 1, 2, \cdots$, of $\mathcal{H}_1$ and nearly orthogonal sets of functions $\psi_j, \eta_j \in \mathcal{H}_2$, $j = 1, 2, \cdots$, such that for some constants ν_j, one has

$$A\phi_j = \nu_j^{-1}\eta_j, \quad A^*\psi_j = \nu_j^{-1}\phi_j; \tag{4.5}$$

$$\langle \eta_{j_1}, \boldsymbol{\psi}_{j_2} \rangle_{\mathcal{H}_2} = \delta_{j_1, j_2}; \tag{4.6}$$

$$\| \sum_j a_j \psi_j \|^2 \asymp \sum_j a_j^2, \quad \| \sum_j a_j \eta_j \|^2 \asymp \sum_j a_j^2. \tag{4.7}$$

where $A^* : \mathcal{H}_2 \to \mathcal{H}_1$ is the linear operator conjugate to A. Donoho (1995) showed that conditions (4.5)–(4.7) hold for a variety of linear operators such as convolution, numerical differentiation, Radon transform, $\{\phi_j\}$ is a wavelet basis. Obviously, assumptions (4.5)–(4.7) hold when $\{\phi_j\}$ is the eigenbasis of the operator A.

Expand functions $f_m \in \mathcal{H}_1$ over the basis ϕ_j, $j = 1, \cdots, n$, and denote the matrix of coefficients by $\mathbf{G}$, so that, by assumption (4.5), for $j = 1, \cdots, n$, $m = 1, \cdots, M$, one has

$$\mathbf{G}_{j,m} = \langle f_m, \phi_j \rangle = \nu_j \langle f_m, A^*\psi_j \rangle = \nu_j \langle A f_m, \psi_j \rangle = \nu_j \, \mathbf{Q}_{j,m}. \tag{4.8}$$

Consider matrix of observations $\mathbf{Y}$ and matrix of errors $\mathbf{E}$ with components $\mathbf{Y}_{j,m} = \langle X_m, \psi_j \rangle$ and $\mathbf{E}_{j,m} = \xi_m(\psi)$. Let $\mathbf{G}_*$ and $\mathbf{Q}_*$ be the true matrices of coefficients. Then, it follows from

(4.1), (4.2) and (4.8) that $\mathbf{Y}$ obeys the sequence model $\mathbf{Y} = \mathbf{Q}_* + \sigma n^{-1/2}\mathbf{E} = \mathbf{\Upsilon}^{-1}\mathbf{G}_* + \sigma n^{-1/2}\mathbf{E}$, so that

$$\mathbf{\Upsilon}\mathbf{Y} = \mathbf{G}_* + \frac{\sigma}{\sqrt{n}}\mathbf{\Upsilon}\mathbf{E}, \quad \mathbf{\Upsilon} = \mathrm{diag}(\nu_1, \cdots, \nu_n). \tag{4.9}$$

Here, by (4.3), $\mathbb{E}(\mathbf{E}_{j,m}) = 0$. While $\mathbf{E}_{j,m}$ are independent for different values of m, i.e., $\mathbb{E}(\mathbf{E}_{j_1,m_1}\mathbf{E}_{j_2,m_2}) = 0$ whenever $m_1 \neq m_2$, they are not necessarily independent when $j_1 \neq j_2$. In particular, denote by $\mathbf{\Sigma}$ the matrix with elements $\mathbf{\Sigma}_{i,j} = \langle \psi_i, \psi_j \rangle$ and observe that

$$\mathbb{E}[(\mathbf{E}\mathbf{E}^T)] = M\,\mathbf{\Sigma}, \qquad \mathbb{E}(\mathbf{E}^T\mathbf{E}) = n\,\mathbf{I}_M,$$

so that matrix $\mathbf{E}$ has the matrix-variate normal distribution $\mathbf{E} \sim N(0, \mathbf{\Sigma} \otimes \mathbf{I}_M)$. Consider matrix $\mathbf{S} \in \mathbb{R}^{n\times n}$ such that $\mathbf{\Sigma} = \mathbf{S}\mathbf{S}^T$, so that $\mathbf{\Sigma}^{-1} = \mathbf{S}^{-T}\mathbf{S}^{-1}$ and $\mathbf{S}^{-1}\mathbf{\Sigma}\mathbf{S}^{-T} = \mathbf{I}_n$. Hence, it follows from (4.7) that for some absolute constant C_ψ, one has

$$\|\mathbf{\Sigma}\|_{op} = \|\mathbf{S}\|_{op}^2 \leq C_\psi^2. \tag{4.10}$$

Then, by definition of the matrix-variate normal distribution and Theorem 2.3.1 of Gupta and Nagar (2000), we derive that

$$\boldsymbol{\epsilon} = \mathrm{vec}(\mathbf{E}) \sim N(0, \mathbf{\Sigma} \otimes \mathbf{I}_M) \tag{4.11}$$

Recall that functions f_m belong to K different groups, so that $f_m = h_k$ with $k = z(m)$ where $z = z(m)$ is a clustering function. Denote the matrix of coefficients of functions h_k in the basis ϕ_j by $\mathbf{\Theta}$, so that $\mathbf{\Theta}_{j,k} = \langle h_k, \phi_j \rangle$, $j = 1, \cdots, n$, $k = 1, \cdots, K$. Hence, if $z : \{1, \ldots, M\} \to \{1, \ldots, K\}$ is the clustering function and $\mathbf{Z} \in \{0,1\}^{M\times K}$ is a clustering matrix, then $\mathbf{G}_{i,j} = \mathbf{\Theta}_{z(i),j}$ for $i = 1, \ldots, M$, and $j = 1, \ldots, n$, $\mathbf{\Theta} = \mathbf{G}\mathbf{Z}\mathbf{D}^{-2}$ and $\mathbf{G}$ can be recovered as $\mathbf{G}_{\mathbf{Z},K} = \mathbf{\Theta}\mathbf{Z}^T$.

Observe that, since for the ill-posed inverse problems, the values of ν_j are growing with j, due to equation (4.9), the coefficients $\mathbf{G}_{i,j} = \mathbf{\Theta}_{z(i),j}$ are harder and harder to recover as j is growing. In order to obtain a reasonable solution of the problem (4.9) one needs to ensure that functions h_k, $k = 1, \cdots, K$, allow a sparse representations in the basis ϕ_j. In particular, we assume that h_k belong to a Sobolev ball $h_k \in \mathcal{S}(r, \mathcal{A})$, $k = 1, \ldots, K$, where

$$\mathcal{S}(r, \mathcal{A}) = \left\{ h = \sum_j \theta_j \phi_j : \ \sum_{j=0}^{\infty} |\theta_j|^2 (j+1)^{2r} \leq \mathcal{A}^2, \quad r > 1/2 \right\}. \tag{4.12}$$

The latter implies that

$$\sum_{j=0}^{\infty} |\mathbf{\Theta}_{jk}|^2 (j+1)^{2r} \leq \mathcal{A}^2, \quad r > 1/2. \tag{4.13}$$

4.3 Estimation

Condition (4.13) means that, coefficients $\mathbf{\Theta}_{jk}$ decrease rapidly as j increases and hence, for large n, one does not need to keep all n coefficients for an accurate estimation of functions f_m (and h_k); on the contrary, this will yield an estimator with a huge variance. For this reason, for every function h_k we can choose a set $J_k \subseteq \{1, \ldots, n\}$ and set $\mathbf{\Theta}_{jk} = 0$ if $j \notin J_k$. Note that since conditions (4.13) apply to all $k = 1, \cdots, K$ simultaneously, we can choose $J_k = J$ for every k. Then, one has $\mathbf{G}_{j,m} = 0$ if $j \in J^c$ where the set J^c is complementary to J. In order to express the latter in a matrix form, we introduce matrix

$$\mathbf{W}_J = \mathrm{diag}(\mathbf{w}_1, ..., \mathbf{w}_n) \quad \text{with} \quad \mathbf{w}_j = \mathbb{I}(j \in J), \tag{4.14}$$

and observe that condition $(\mathbf{I}_n - \mathbf{W}_J)\mathbf{G} = \mathbf{0}$ ensures that $\mathbf{G}_{j,m} = 0$, $j \in J^c$. Consider projection matrices

$$\mathbf{\Pi}_{\mathbf{Z},K} = \mathbf{Z}\mathbf{D}^{-2}\mathbf{Z}^T, \quad \mathbf{\Pi}^{\perp}_{\mathbf{Z},K} = \mathbf{I}_M - \mathbf{\Pi}_{\mathbf{Z},K},$$

the projection matrix on the column space of matrix $\mathbf{Z}$ and the projection matrix on the orthogonal subspace, respectively. Here, we use index K to indicate that not only the clustering matrix $\mathbf{Z}$ but also the number of clusters K is unknown. In order to reduce the variances of the estimators of functions f_m, $m = 1, \cdots, M$, we approximate the matrix of coefficients $\mathbf{G}_*$ by $\mathbf{W}_J\mathbf{G}\mathbf{\Pi}_{\mathbf{Z},K}$.

Consider an integer $K \in [M]$ and a set $\mathcal{M}(M, K)$ of clustering matrices that cluster M nodes into K groups. Then, the objective is to find matrices $\mathbf{G}$ and $\mathbf{Z} \in \mathcal{M}(M, K)$, a set J and an integer K such that

$$\|\mathbf{G} - \mathbf{\Upsilon}\mathbf{Y}\mathbf{\Pi}_{\mathbf{Z},K}\|_F^2 + \|\mathbf{\Upsilon}\mathbf{Y}\mathbf{\Pi}^{\perp}_{\mathbf{Z},K}\|_F^2 \Longrightarrow \min \quad \text{subject to} \quad (\mathbf{I}_n - \mathbf{W}_J)\mathbf{G} = \mathbf{0}, \tag{4.15}$$

Since $\|\mathbf{\Upsilon}\mathbf{Y}\mathbf{\Pi}_{\mathbf{Z},K}\|_F^2 + \|\mathbf{\Upsilon}\mathbf{Y}\mathbf{\Pi}^{\perp}_{\mathbf{Z},K}\|_F^2 = \|\mathbf{\Upsilon}\mathbf{Y}\|_F^2$ is independent of $\mathbf{G}$ and $\mathbf{Z}$, the problem can be re-written as

$$\|\mathbf{G}\|_F^2 - 2\mathrm{Tr}(\mathbf{Y}^T\mathbf{\Upsilon}\mathbf{G}\mathbf{\Pi}_{\mathbf{Z},K}) \Longrightarrow \min \quad \text{subject to} \quad (\mathbf{I}_n - \mathbf{W}_J)\mathbf{G} = \mathbf{0}. \tag{4.16}$$

Note though that optimization problem (4.16) has a trivial solution: $K = M$, $J = [n]$ and $\mathbf{Z} = \mathbf{I}_M$. In order to avoid this, we put a penalty on the value of K and the set J.

Then $\mathbf{Z}, \mathbf{G}, J$ and K can be found a solution of the following optimization problem:

$$(\hat{\mathbf{Z}}, \widehat{\mathbf{G}}, \hat{J}, \hat{K}) \in \underset{\mathbf{Z},\mathbf{G},J,K}{\operatorname{argmin}} \left\{ \|\mathbf{G}\|_F^2 - 2\mathrm{Tr}(\mathbf{Y}^T \boldsymbol{\Upsilon} \mathbf{G} \boldsymbol{\Pi}_{\mathbf{Z},K}) + \mathrm{Pen}(J,K) \right\} \tag{4.17}$$
$$\text{subject to } \mathbf{Z} \in \mathcal{M}(M,K), (\mathbf{I}_n - \mathbf{W}_J)\mathbf{G} = \mathbf{0}, J \subseteq [n], K \in [M]$$

Note that if $\hat{\mathbf{Z}}, \hat{J}$ and $\hat{K}$ were known, then it follows from (4.15) that $\widehat{\mathbf{G}}$ would be given by

$$\widehat{\mathbf{G}} = \mathbf{W}_{\hat{J}} \boldsymbol{\Upsilon} \mathbf{Y} \boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \tag{4.18}$$

and problem (5.8) can be presented as

$$(\hat{\mathbf{Z}}, \hat{J}, \hat{K}) \in \underset{\substack{\mathbf{Z}\in\mathcal{M}(M,K) \\ J,K}}{\operatorname{argmin}} \left\{ \|(\mathbf{I} - \mathbf{W}_J)\boldsymbol{\Upsilon}\mathbf{Y}\boldsymbol{\Pi}_{\mathbf{Z},K}\|_F^2 + \|\boldsymbol{\Upsilon}\mathbf{Y}\boldsymbol{\Pi}_{\mathbf{Z},K}^{\perp}\|_F^2 + \mathrm{Pen}(J,K) \right\} \tag{4.19}$$

We choose $\mathrm{Pen}(J,K)$ so that it is of the order of the error of estimating $\mathbf{G}$. In particular, we set

$$\mathrm{Pen}(J,K) = \frac{2C_\psi^2\sigma^2}{n} \left[26K \sum_{j\in J} \nu_j^2 + 39(\max_{j\in J} \nu_j^2) \left\{ M \ln K + |J| \ln\left(\frac{ne}{|J|}\right) + \ln(Mn^{\tau+1}) \right\} \right] \tag{4.20}$$

where C_ψ is defined in (4.10). Penalty (4.20) consists of the four terms. The first term, $26K \sum_{j\in J} \nu_j^2$ represents the error of estimating $|J|$ coefficients for each of the distinct functions h_k, $k = 1, \ldots, K$. The second and the third terms account for the difficulty of clustering M functions into K classes and choosing a set $J \subset \{1, \ldots, n\}$. The last term is of the smaller asymptotic order and offsets the error of the choice of K. Observe that since the data is weighted by the diagonal matrix $\boldsymbol{\Upsilon}$ in (4.9), the last three terms are weighted by $\max_{j\in J} \nu_j^2$.

In practice, we shall solve optimization problem (4.19) separately for each $K \in [M]$ and then

choose the value of K that delivers the smallest value in (4.19). We estimate the matrix of coefficients $\mathbf{G}$ by $\widehat{\mathbf{G}}$ defined in (4.18). After coefficients $\widehat{\mathbf{G}}$ are obtained, we estimate f_m, $m = 1, \ldots, M$, by

$$\hat{f}_m = \sum_{j=1}^{\hat{L}} \widehat{\mathbf{G}}_{j,m} \phi_j, \quad m = 1, \cdots, M. \tag{4.21}$$

4.4 Estimation error

4.4.1 The oracle inequality

The average error of estimating f_m by $\hat{f}_m$, $m = 1, \ldots, M$, is the given by

$$R(\mathbf{f}, \hat{\mathbf{f}}) = M^{-1} \sum_{m=1}^{M} \|\hat{f}_m - f_m\|^2. \tag{4.1}$$

where $\mathbf{f}$ and $\hat{\mathbf{f}}$ are column vector with functional components f_m and $\hat{f}_m$, $m = 1, \ldots, M$, respectively.

It is easy to see that the main portion of the error is due to $M^{-1}\|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2$. The following statement places an upper bound on $\|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2$.

Theorem 3 *Let* $(\hat{\mathbf{Z}}, \widehat{\mathbf{G}}, \hat{J}, \hat{K})$ *be a solution of optimization problem* (5.8) *with the penalty* $Pen(J, K)$ *given by expression* (4.20). *Then, there exists a set* $\Omega = \Omega(\tau)$ *with* $\mathbb{P}(\Omega) \geq 1 - 2n^{-\tau}$ *such that for every* $\omega \in \Omega$ *one has*

$$\|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2 \leq \min_{\mathbf{Z},J,K} \left\{ 3\,\|\mathbf{W}_J \mathbf{G}_* \mathbf{\Pi}_{\mathbf{Z},K} - \mathbf{G}_*\|_F^2 + 4Pen(J,K) \right\} \tag{4.2}$$

Theorem 3 provides an oracle inequality for $\|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2$. The first term in expression (4.2)

is the bias term that quantifies the error of approximation of matrix $\mathbf{G}_*$ when its columns are averaged over K clusters using matrix $\mathbf{Z}$ and one keeps only terms with $j \in J$ in the approximations of each of the K cluster means. This term is decreasing when K and $|J|$ are increasing. The second term, $\mathrm{Pen}(J, K)$, is the variance term that represents the error of estimation for the particular choices of $\mathbf{Z}$, J and K. This term grows when K and $|J|$ are increasing. The error is provided by the best possible bias-variance balance in (4.2).

Note that since the right hand side in (4.2) is minimized over $\mathbf{Z}$ and K, if some of the functions h_k, $k = 1 \cdots, K$, are similar but not exactly identical to each other, it may be advantageous to place those functions in the same cluster, hence, reducing the variance component of the error. Our methodology will automatically take advantage of this opportunity. Theorem 3 however does not provide an explicit expression for the error in the case of a specific collection of functions h_k, $k = 1, \cdots, K_*$ and a clustering matrix $\mathbf{Z}_* \in \mathcal{M}(M, K_*)$. This study is carried out in the next section.

4.4.2 The upper bounds for the risk

In order to study particular scenarios, in what follows, we shall consider the following condition on ν_j:

$$\aleph_1 j^\gamma \exp\left(\alpha j^\beta\right) \leq |\nu_j| \leq \aleph_2 j^\gamma \exp\left(\alpha j^\beta\right) \tag{4.3}$$

for some absolute positive constants $\aleph_1$, $\aleph_2$ and nonnegative γ, α and β where $\beta = 0$ and $\gamma > 0$ whenever $\alpha = 0$. The problem (4.1) is known to be moderately ill-posed if $\alpha = 0$ and severely ill-posed if $\alpha > 0$. Assume that $h_k \in \mathcal{S}(r, \mathcal{A})$, $k = 1, \ldots, K_*$, where $\mathcal{S}(r, \mathcal{A})$ is defined in (4.12). Denote by $\mathbf{h}$ the functional column vector with components h_k, $k = 1, \ldots, K_*$. Consider the maximum risk of our estimator $\hat{\mathbf{f}}$ over all $h_k \in \mathcal{S}(r, \mathcal{A})$, $k = 1, \ldots, K_*$, and all

clustering matrices $\mathbf{Z} \in \mathcal{M}(M, K_*)$

$$R(\hat{\mathbf{f}}, \mathcal{S}(r, \mathcal{A}), M, K_*) = \max_{\mathbf{f}, \mathbf{Z}} R(\mathbf{f}, \hat{\mathbf{f}}) \quad \text{subject to} \tag{4.4}$$

$$\mathbf{f} = \mathbf{Z}\mathbf{h},\ h_k \in \mathcal{S}(r, \mathcal{A}),\ k = 1, \ldots, K_*,\ \mathbf{Z} \in \mathcal{M}(M, K_*),$$

where $\mathcal{S}(r, \mathcal{A})$ is defined in (4.12) and $\mathcal{M}(M, K_*)$ is the set of all clustering matrices that place M objects into K_* classes. In what follows, we assume that M grows as some power of n, so that

$$\ln n \asymp \ln M \asymp \ln(Mn). \tag{4.5}$$

Then, application of the oracle inequality (4.2) with $|J| = L$ and $K = K_*$ provides the following upper bounds for the error.

Theorem 4 *Let assumption* (4.5) *hold and* ν_j, $j = 1, \cdots, n$, *satisfy condition* (4.3) *with* $r > 1/2$. *Let* $(\hat{\mathbf{Z}}, \widehat{\mathbf{G}}, \hat{L}, \hat{K})$ *be a solution of optimization problem* (5.8) *with the penalty given by expression* (4.20). *Then, with probability at least* $1 - 2n^{-\tau}$, *one has* $R(\hat{\mathbf{f}}, \mathcal{S}(r, \mathcal{A}), M, K_*) \leq CR(M, K_*, n)$ *where the constant* C *depends on* $\alpha, \beta, \gamma, r, \tau$ *and* $\mathcal{A}$ *only and*

$$R(M, K_*, n) = \max\left\{ \left(\frac{\sigma^2 \ln K_*}{n} \right)^{\frac{2r}{2r+2\gamma}},\ \left(\frac{\sigma^2 (K_* + \ln n)}{M n} \right)^{\frac{2r}{2r+2\gamma+1}} \right\}, \tag{4.6}$$

if $\alpha = \beta = 0$, *and*

$$R(M, K_*, n) = \max\left\{ \left[\ln\left(\frac{n}{\sigma^2 \ln K_*} \right) \right]^{-\frac{2r}{\beta}},\ \left[\ln\left(\frac{Mn}{\sigma^2 K_*} \right) \right]^{-\frac{2r}{\beta}} \right\}, \tag{4.7}$$

if $\alpha > 0, \beta > 0$.

4.4.3 The minimax lower bounds for the risk

In order to show that the estimator developed in this paper is asymptotically near-optimal, below we derive minimax lower bounds for the risk over all $h_k \in \mathcal{S}(r, \mathcal{A})$, $k = 1, \ldots, K_*$, and all clustering matrices $\mathbf{Z} \in \mathcal{M}(M, K_*)$. For this purpose, we define the minimax risk as

$$R_{\min}(\mathcal{S}(r, \mathcal{A}), M, K_*) = \min_{\tilde{\mathbf{f}}} R(\tilde{\mathbf{f}}, \mathcal{S}(r, \mathcal{A}), M, K_*) \tag{4.8}$$

where $\tilde{\mathbf{f}}$ is any estimator of $\mathbf{f}$ on the basis of matrix of observations $\mathbf{Y}$.

Theorem 5 *Let ν_j, $j = 1, \cdots, n$ satisfy condition (4.3) and $r > 1/2$. Then, with probability at least 0.1, one has $R_{\min}(\mathcal{S}(r, \mathcal{A}), M, K_*) \geq C R_{\min}(M, K_*, n)$ where the constant C depends on α, β, γ, r and $\mathcal{A}$ only and*

$$R_{\min}(M, K_*, n) = \max\left\{ \left(\frac{\sigma^2 \ln K_*}{n} \right)^{\frac{2r}{2r+2\gamma}}, \left(\frac{\sigma^2 K_*}{M n} \right)^{\frac{2r}{2r+2\gamma+1}} \right\}, \tag{4.9}$$

if $\alpha = \beta = 0$, and

$$R_{\min}(M, K_*, n) = \max\left\{ \left[\ln\left(\frac{n}{\sigma^2 \ln K_*} \right) \right]^{-\frac{2r}{\beta}}, \left[\ln\left(\frac{Mn}{\sigma^2 K_*} \right) \right]^{-\frac{2r}{\beta}} \right\}, \tag{4.10}$$

if $\alpha > 0, \beta > 0$.

Observe that expressions (4.7) and (4.10) for the upper and the lower bounds of the risk are identical, so our estimators are asymptotically optimal in the case of $\alpha > 0, \beta > 0$. If $\alpha = \beta = 0$, the first terms in the expressions (4.6) and (4.9) are the same while the second terms differ by a factor $\rho(n, K_*) = (1 + (\ln n)/K_*)^{\frac{2r}{2r+2\gamma+1}}$. Therefore, the estimators are asymptotically optimal unless the second term in (4.6) dominates the first term. In the

latter case, the estimator is asymptotically near-optimal within the factor $\rho(n, K_*)$.

4.4.4 The advantage of clustering

Theorems 4 and 5 allow to answer the question whether clustering in linear ill-posed inverse problems is advantageous or not. Indeed, solving problem (4.1) for each $m = 1, \cdots, M$ separately is equivalent to choosing $K = M = 1$ in the penalty. In this case, one obtains the following corollary.

Corollary 1 *If each of the inverse problems is solved separately, where the penalty is of the form* (4.20) *with* $K = M = 1$ *and* $J = \{1, \cdots, L\}$*, then, with probability at least* $1 - 2n^{-\tau}$*, the average estimation error* $\tilde{R}(n)$ *defined in* (4.1) *is bounded by*

$$\tilde{R}(n) \asymp \begin{cases} \left[\frac{\sigma^2 \ln n}{n}\right]^{\frac{2r}{2\gamma+2r+1}}, & \text{if } \alpha = \beta = 0, \\ \left[\ln\left(\frac{n}{\sigma^2}\right)\right]^{-\frac{2r}{\beta}}, & \text{if } \alpha > 0, \beta > 0. \end{cases} \tag{4.11}$$

If assumption (4.5) *hold, then for* $n \to \infty$*,* $M \to \infty$*, one has*

$$\frac{\tilde{R}(n)}{R(M, K_*, n)} \asymp \begin{cases} M^{\frac{2r}{2\gamma+2r+1}}, & \text{if } \alpha = \beta = 0, K_* = 1 \\ \left[\min\left\{\frac{M \ln n}{K_*}\left(\frac{n(\ln n)^{2r+2\gamma}}{\sigma^2 (\ln K_*)^{2r+2\gamma+1}}\right)^{\frac{1}{2r+2\gamma}}, M\right\}\right]^{\frac{2r}{2r+2\gamma+1}} & \text{if } \alpha = \beta = 0, K_* \geq 2 \\ 1 & \text{if } \alpha > 0, \beta > 0. \end{cases}$$

Therefore, clustering is asymptotically advantageous if $\alpha = \beta = 0$ *and*

$$n \to \infty, \ M \to \infty, \ \frac{n(\ln n)^{2r+2\gamma}}{\sigma^2 (\ln K_*)^{2r+2\gamma+1}} \to \infty.$$

4.5 Simulations

In order to study finite sample properties of the proposed estimation procedure, we carried out limited simulation study. For this purpose, we used $K = 4$ and considered two sets of test functions: smooth functions

$$f_1(x) = \sin(4\pi x), \quad f_2(x) = \cos(4\pi x), \quad f_3(x) = (x - 0.5)^2, \quad f_4(x) = (x - 0.5)^4 \tag{4.1}$$

and non-smooth ones,

$$f_1(x) = f_B(x), \quad f_2(x) = f_W(x), \quad f_3(x) = f_P(x), \quad f_4(x) = |x - 0.5| \tag{4.2}$$

where $f_B(x)$, $f_W(x)$ and $f_P(x)$ are the *blip, wave* and *parabolas* introduced by Donoho and Johnstone [29]. The functions are sampled at n equispaced points j/n, $j = 1, \cdots, n$, on the interval $[0, 1]$. While functions in (4.1) are simpler and easier to recover, functions in (4.2) are more difficult to estimate.

We studied a periodic convolution equation $q = Af = f * g$ with a kernel g that transforms into a product in Fourier domain

$$\tilde{q}_j = \tilde{g}_j \tilde{f}_j, \quad \nu_j = 1/\tilde{g}_j, \ j = 1, \cdots, n, \tag{4.3}$$

where, for any function ψ we denote its j-th Fourier coefficient by $\tilde{\psi}_j$. The periodic Fourier basis ϕ_j serves as the eigenbasis for this operator and, hence, conditions (4.5)–(4.7) hold with ν_j given above.

We carried out simulations with the periodized versions of the following two kernels

$$g_1(x) = \exp(-\lambda|x|), \quad g_2(x) = \exp(-\lambda x^2/2) \tag{4.4}$$

where $g_1(x)$ corresponds to the case of $\alpha = \beta = 0, \gamma = 2$ while $g_2(x)$ corresponds to $\alpha \propto 1/\lambda$, $\beta = 2$ in (4.3). Hence, the problem is moderately ill-posed with g_1 and severely ill-posed with g_2. In addition, recovery of the solution becomes easier as λ grows.

For each of the test functions $f_k, k = 1, \cdots, K$, we evaluated $(Af)_k$ and subsequently scaled them to have equal norms, hence, adjusting $f_k(x)$ accordingly. Furthermore, we generated a clustering function $z : M \to K$ that places M objects into K classes, M/K into each class at random. We obtained the true matrices $\mathbf{F}, \mathbf{Q} \in \mathbb{R}^{n \times M}$ with the sampled versions of the vector functions $f_{z(m)}$ and $(Af)_{z(m)}$, $m = 1, \cdots, M$, respectively. Finally, we generated data $\mathbf{X}$ by adding independent Gaussian noise with the standard deviation σ to every element in $\mathbf{Q}$. We found σ by fixing the Signal-to-Noise Ratio (SNR) and choosing $\sigma = \text{std}(\mathbf{F})/SNR$, where std($\mathbf{F}$) is the standard deviation of the matrix $\mathbf{F}$ reshaped as a vector. In what follows, we considered several noise scenarios: SNR = 3, SNR = 5 and SNR = 7 for g_1 and SNR = 10, SNR = 15, SNR = 20 for g_2.

Since, even for a fixed K, finding $\widehat{\mathbf{Z}}$ that produces a global minimum in (4.19) requires $\binom{M}{K}$ operations, we derived the estimated cluster assignment $\hat{z} : M \to K$, using the `Kmeans` procedure and subsequently averaged observation vectors in each class, thus obtaining K estimated cluster averages $\hat{y}_k$, $k = 1, \cdots, K$. In our simulations, we used the Daubechies 8 wavelet basis as ϕ_j, $j = 1, \cdots, n$. In order to obtain ψ_j in (4.5), we generated wavelet functions ϕ_j using `MakeWavelet` command in WaveLab850 package for Matlab and recovered ψ_j using the second equation in (4.5). We further obtained the estimated wavelet coefficients as a scalar product of ψ_j and $\hat{y}_k$, $j = 1, \cdots, n$, $k = 1, \cdots, K$, and applied hard thresholding

to obtain the set J. Finally, we used the inverse wavelet transform to recover the estimators $\hat{f}_k$ of f_k, $k = 1, \cdots, K$. The equispaced versions of those estimators appear as the estimator $\widehat{\mathbf{F}}$ of the matrix $\mathbf{F}$ with columns $\hat{\mathbf{f}}_{\hat{z}(m)}$ representing functions f_k.

In order to assess the benefits of the clustering, we also obtained estimators without clustering by using the same procedure with the only difference that $K = M$ and $\hat{z}$ is the identity transformation. We measured the accuracy of the estimators by their relative Frobenius error

$$\Delta(\widehat{\mathbf{F}}) = \|\widehat{\mathbf{F}} - \mathbf{F}\|_F / \|\mathbf{F}\|_F. \tag{4.5}$$

Although we carried out simulations with a more diverse sets of parameters, here we report the results for $n = 256$, $M = 60$ and $K = 4$. Tables below report the mean values of $\Delta(\widehat{\mathbf{F}})$ with and without clustering over 100 simulation runs (with the standard error of the means presented in parentheses) for the test functions in (4.1) or (4.2), one of the kernels in (4.4) and various values of λ. In particular, Tables 1 and 2 report results for the set of smooth functions (4.1) with $g_1(x)$ in (4.4) for Table 1 and $g_2(x)$ in (4.4) for Table 2. Tables 3 and 4 report results for the set of non-smooth functions (4.2) with $g_1(x)$ in (4.4) for Table 3 and $g_2(x)$ in (4.4) for Table 4.

Table 4.1: Estimation errors $\Delta(\widehat{\mathbf{F}})$ with and without clustering averaged over 100 simulation runs for the set of smooth functions (4.1) with $g_1(x)$ in (4.4). The standard deviations of the means are in parentheses.

$\lambda = 7$			
	With Clustering	Misclassification Error	Without Clustering
$SNR = 3$	0.0348 (0.0198)	0.0608 (0.1271)	0.1338 (0.0092)
$SNR = 5$	0.0218 (0.0105)	0.0185 (0.0739)	0.0845 (0.0055)
$SNR = 7$	0.0171 (0.0065)	0.0061 (0.0433)	0.0541 (0.0043)
$\lambda = 5$			
$SNR = 3$	0.0602 (0.0144)	0.0584 (0.1262)	0.2688 (0.0118)
$SNR = 5$	0.0397 (0.0103)	0.0711 (0.1321)	0.1637 (0.0069)
$SNR = 7$	0.0283 (0.0068)	0.0231 (0.0791)	0.1148 (0.0050)
$\lambda = 3$			
$SNR = 3$	0.0991 (0.0105)	0.1576 (0.1616)	0.4529 (0.0090)
$SNR = 5$	0.0600 (0.0061)	0.1249 (0.1649)	0.2736 (0.0053)
$SNR = 7$	0.0435 (0.043)	0.1025 (0.1548)	0.1949 (0.0038)

Table 4.2: Estimation errors $\Delta(\widehat{\mathbf{F}})$ with and without clustering averaged over 100 simulation runs for the set of smooth functions (4.1) with $g_2(x)$ in (4.4). The standard deviations of the means are in parentheses.

$\lambda = 20$			
	With Clustering	Clustering Error	Without Clustering
$SNR = 5$	0.0095 (0.0021)	0.1223 (0.1649)	0.0348 (0.0021)
$SNR = 7$	0.0077 (0.0016)	0.0951 (0.1542)	0.0251 (0.0015)
$SNR = 10$	0.0065 (0.0011)	0.0480 (0.1156)	0.0180 (0.0011)
$\lambda = 15$			
$SNR = 5$	0.0142 (0.0016)	0.2423 (0.0959)	0.0245 (0.0015)
$SNR = 7$	0.0138 (0.0011)	0.2115 (0.1402)	0.0199 (0.0011)
$SNR = 10$	0.0137 (0.0008)	0.1718 (0.1595)	0.0169 (0.0008)
$\lambda = 10$			
$SNR = 5$	0.6651 (0.0647)	0.2911 (0.0951)	0.6420 (0.0068)
$SNR = 7$	0.6504 (0.0498)	0.2761 (0.0805)	0.6393(0.0049)
$SNR = 10$	0.6474 (0.0432)	0.2803 (0.0799)	0.6379 (0.0034)

Table 4.3: Estimation errors $\Delta(\widehat{\mathbf{F}})$ with and without clustering averaged over 100 simulation runs for the set of non-smooth functions (4.2) with $g_1(x)$ in (4.4). The standard deviations of the means are in parentheses.

$\lambda = 7$			
	With Clustering	Clustering Error	Without Clustering
$SNR = 3$	0.1474 (0.0070)	0.0000 (0.0000)	0.3782 (0.0047)
$SNR = 5$	0.1027 (0.0057)	0.0000 (0.0000)	0.2577 (0.0061)
$SNR = 7$	0.0864 (0.0044)	0.0000 (0.0000)	0.1897 (0.0028)
$\lambda = 5$			
$SNR = 3$	0.1425 (0.0099)	0.0000 (0.0000)	0.4342 (0.0077)
$SNR = 5$	0.1079 (0.0054)	0.0000 (0.0000)	0.2866 (0.0066)
$SNR = 7$	0.0974 (0.0037)	0.0000 (0.0000)	0.2163 (0.0035)
$\lambda = 3$			
$SNR = 3$	0.1925 (0.0107)	0.0000 (0.0000)	0.5350 (0.0085)
$SNR = 5$	0.1686 (0.0066)	0.0000 (0.0000)	0.3539 (0.0063)
$SNR = 7$	0.1615 (0.0049)	0.0000 (0.0000)	0.2732 (0.0048)

Table 4.4: Estimation errors $\Delta(\widehat{\mathbf{F}})$ with and without clustering averaged over 100 simulation runs for the set of non-smooth functions (4.2) with $g_2(x)$ in (4.4). The standard deviations of the means are in parentheses.

$\lambda = 20$			
	With Clustering	Clustering Error	Without Clustering
$SNR = 5$	0.44704 (0.00013)	0.0000 (0.0000)	0.44830 (0.00022)
$SNR = 7$	0.44701 (0.00009)	0.0000 (0.0000)	0.44765 (0.00013)
$SNR = 10$	0.44699 (0.00006)	0.0000 (0.0000)	0.44731 (0.00008)
$\lambda = 15$			
$SNR = 5$	0.45737 (0.05698)	0.02825 (0.09080)	0.44783 (0.00008)
$SNR = 7$	0.45402 (0.04677)	0.00950 (0.05476)	0.44760 (0.00004)
$SNR = 10$	0.44736 (0.00002)	0.00000 (0.00000)	0.44747 (0.00003)
$\lambda = 10$			
$SNR = 5$	0.77061 (0.07038)	0.16063 (0.15586)	0.73034 (0.00541)
$SNR = 7$	0.75037 (0.05603)	0.11800 (0.15347)	0.72752 (0.00386)
$SNR = 10$	0.73849 (0.04433)	0.08925(0.15463)	0.72593 (0.00270)

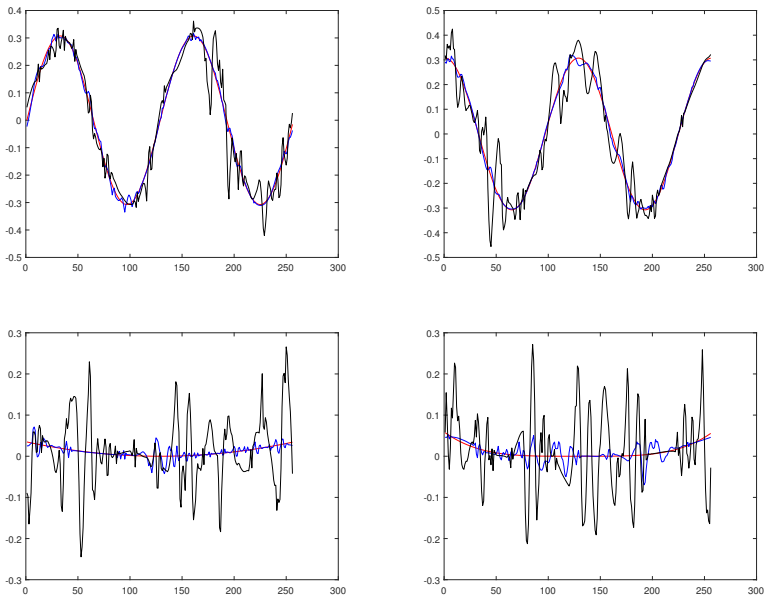

Figure 4.1: True functions (red) and their estimators with clustering (blue) and without clustering (black) for smooth function in (4.1) and kernel g_1 in (4.4) with $\lambda = 3$ and SNR=3. Top row: f_1 (left), f_2 (right). Bottom row: f_3 (left), f_4 (right).

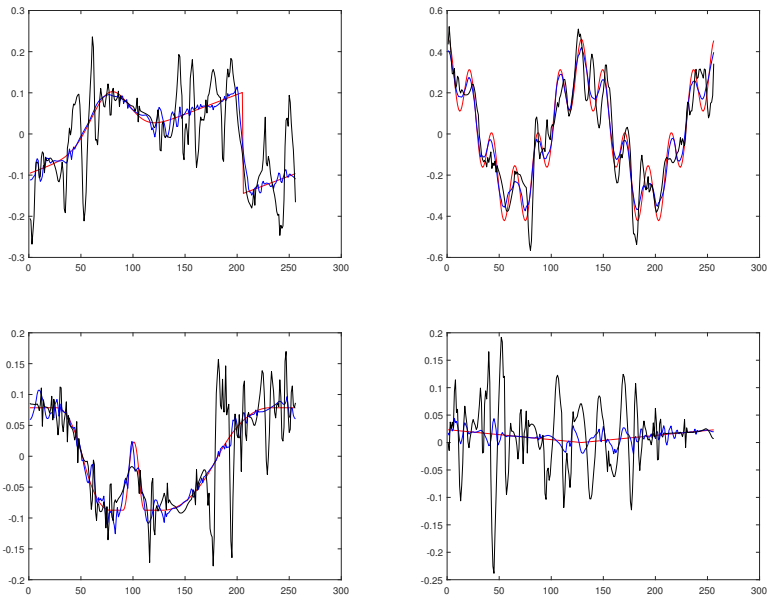

Figure 4.2: True functions (red) and their estimators with clustering (blue) and without clustering (black) for non smooth function in (4.2) and kernel g_1 in (4.4) with $\lambda = 3$ and SNR=3. Top row: f_1 (left), f_2 (right). Bottom row: f_3 (left), f_4 (right).

CHAPTER 5: LEARNING FUNCTIONAL BRAIN CONNECTIVITY THROUGH AUTO-REGRESSIVE (AR) MODELS

5.1 Introduction

In human connectome projects brain network analysis is an exciting field. The anaysis identifies anatomically and functionally distinct brain areas in health and diseases. Functional magnetic resonance imaging (fMRI) and diffusion tensor imaging (DTI) are most commonly used imaging modalities to explore functional and structural connectivity patterns of brain regions. Diffusion imaging helps to identify structural connections of brain regions, fMRI comprises functional interactions of brain regions, at macroscopic scales.

The fMRI infers brain activity by indirectly measuring changes in blood flow using magnetic resonance imaging (MRI) and the sub-modality, resting-state fMRI (rs-fMRI) is a powerful method for evaluating regional interactions that occur when a subject is not at rest (i.e., performing an explicit task). In various studies that resting-state patterns of correlated brain activity are consistent across individuals, and predictive of disease states when abnormal [49]. Although rs-fMRI is a dynamic imaging modality that measures the neural activities over a period of time, the commonly used methods for analysis of rs-fMRI signals is based on the analysis of entire time series [68]. However, this approach overlooks inter-region interactions and ignores subtle patterns, which may be important for identifying functionally distinct networks.

5.2 Dynamic functional connectivity (dFC)

Dynamic functional connectivity identifies the functional interplay between regions of the brain and the changes that occur over a short time. The most widely used approach to identifying dFC is the sliding window approach where correlations between windowed time-courses of brain regions are estimated [60]. However, identification of these change points in the signal to determine the window-size has proven challenging with most methods preferring to use a trial and error approach to identify window size that gives the best modeling performance [72].

In this work, we aim to identify change-points in the signal and generate windows of equal width using AR & VAR. We further propose a statistical feature selection approach to perform classification on the challenging cognitive expertise data using the previously identified optimal window lengths and stride.

In this study, our goal is to identify the change points in the signals of rs-fMRI by learning the signal model and use this knowledge to better represent the inter-region interactions. Towards this, we use univariate and multivariate stochastic process models, Auto-Regression (AR) and Vector AR (VAR) respectively, to identify the signal change points. We propose to use this new model to then learn discriminating features/connections in the functional connectivity networks generated using the partial correlations i.e. inverse covariance matrices, using statistical methods and thereby classify functional brain networks into distinct classes.

5.3 Methods

5.3.1 Notation

In this section, we are going to use the following notations. We denote vectors and matrices by bold letters. For any matrix $\mathbf{A}$, we denote its Frobenius norm with $\|\mathbf{A}\|_F$ and $\mathbf{A}^T$ by the transpose of $\mathbf{A}$. We denote the entries of the matrix by θ_i^j where index j indicates the row and index i indicates the column. In the set up of a vector the i^{th} entry is denoted as y_t^i.

5.3.2 Auto-Regression (AR)(by collaborator Harish RaviPrakash)

AR models use a weighted sum of the signal's past values to predict the signal's future value under the causal assumption while modeling the residual error as Gaussian. The number of past values that are used to predict the future value of the signal is termed as the *order* of the AR. We propose to use the causality of the fMRI signal and model the problem as an AR problem. Wang et. al [71] proposed the fitting of autoregressive time series with the *LASSO* [63] procedure for a fixed order AR. Consider the time series data $X_1, X_2, \ldots, X_n$ and the auto-regressive process of order k, AR(k), that models this data can be represented as

$$X_t = A_1 X_{t-1} + A_2 X_{t-2} + \ldots + A_k X_{t-k} + \xi_{t+1} \quad \text{for} \quad t = k+1, \ldots, n-1 \tag{5.1}$$

where A_i are the regression coefficients and ξ_{t+1} is the residual error in the prediction. This can be modeled as a Lasso regression problem as

$$\min_A \quad \frac{1}{2}\|\mathbf{Y} - \mathbf{X}\mathbf{A}\|_{\mathbf{2}}^{\mathbf{2}} + \lambda\|\mathbf{A}\|_{\mathbf{1}} \tag{5.2}$$

where $\mathbf{Y} = (\mathbf{X_{k+1}}, \mathbf{X_{k+2}}, \ldots, \mathbf{X_n})^{\mathbf{T}}$, and λ is the regularization term.

For a given rs-fMRI time series signal, we set up the problem with $t = k + 1$ in (5.1). Therefore, the labels, $\mathbf{Y}$, can be represented as $(X_{k+1}, X_{k+2}, \ldots X_n)^T$ i.e. starting from a point where it has a whole order of observations before it. A simple sliding window approach is proposed wherein at each step, the 'k' length block of the signal is taken as the input and the $(k+1)^{th}$ instance as the output. A stride of 's' is used to move the window across untill the output is the last/final observation of the signal. The input blocks are stacked vertically and similarly to the output values so as to have the data in the form of (5.2) the total number of blocks can be calculated as $\frac{T-k}{s} + 1$.

Since the AR is performed for fixed values of the order 'k', different sizes of the window length can be used to test the performance of the learned model. At each prediction step, the errors $(\xi_{t+1}, \ldots, \xi_{t+s})$ are computed from (5.1). If there are P different window sizes tested, the error

$$\mathbf{\Xi} = \begin{bmatrix} \xi_1^1 & \xi_1^2 & \cdots \\ \vdots & \ddots & \\ \xi_s^1 & & \xi_s^P \end{bmatrix}$$

The mean percentage error metric is used to compute a more intuitive error term. The error ξ_j^i, where i is the window length and j is the stride, can be computed as

$$\xi_j^i = \frac{100}{n} \sum_{t=1}^{n} \frac{\xi_{t+j}^i}{x_{t+j}} \tag{5.3}$$

(5.2) and (5.3) are repeated for all the different signals (from the different brain regions) and we have k_i and s_i for each region i. (5.3) is modified to find the average error across all

regions as

$$\xi_j^i = \frac{100}{n * M} \sum_{i=1}^{M} \sum_{t=1}^{n} \frac{\xi_{t+j}^i}{x_{t+j}^i} \tag{5.4}$$

where M is the number of brain regions. To identify the optimal 'k' and 's', a satisfiability criterion is developed. $\mathbf{\Theta}$, a binary matrix indicates the satisfiability of the performance of the learned model.

$$\mathbf{\Theta} = \begin{bmatrix} \theta_1^1 & \theta_1^2 & \dots \\ \vdots & \ddots & \\ \theta_s^1 & & \theta_s^P \end{bmatrix}$$

To determine satisfiability, the following constraint is used.

$$\theta_j^i = \begin{cases} 1, & \text{if } \xi_j^i \leq \tau, \tau = 10 \\ 0, & \text{otherwise} \end{cases} \tag{5.5}$$

The τ in (5.5) is to decide when the learned model is significantly similar to the labels.

5.3.3 Vector Auto-Regression (VAR)

VAR models generalize the univariate AR models by allowing more than one evolving variable. It identifies the linear relationship between multiple time-series and each variable is a linear function of past lag of itself and past lag of other variables.

Consider time series generated from M variables (brain regions) and the process of order k structured as

$$\mathbf{y_t} = \mathbf{A_1 y_{t-1}} + \mathbf{A_2 y_{t-2}} + \ldots + \mathbf{A_k y_{t-k}} + \xi_\mathbf{n} \quad \text{for } \mathbf{t = k + 1}, \ldots, \mathbf{n} \tag{5.6}$$

where $\mathbf{y_t} = (y_t^{(1)}, y_t^{(2)}, ..., y_t^{(M)}) \in R^{M\times 1}$ is the t^{th} sample of M-dimensional time series, $\mathbf{A_i} \in R^{(M\times M)}$ regression coefficient matrices and $\xi_\mathbf{n} \in R^{M\times 1}$ is additive gaussian noise with $E(\xi_\mathbf{t}) = 0$ and $Cov(\xi_\mathbf{t}) = \mathbf{\Sigma}$.

Model (5.6) can be represented in the form of a multivariate linear regression model [39] as follows .

$$\mathbf{y_t} = \mathcal{A}\mathbf{x_t} + \xi_\mathbf{n} \tag{5.7}$$

where $\mathbf{x_t} = (\mathbf{y_{t-1}}; \mathbf{y_{t-2}}; ...; \mathbf{y_{t-k}}) \in R^{(Mk\times 1)}$ and $\mathcal{A} = (\mathbf{A_1}; \mathbf{A_2}; ...; \mathbf{A_k}) \in R^{(M\times Mk)}$. For a given rs-fMRI signal $(M \times n)$, we use a simple sliding window block $(M \times k)$ and at each step the block signal is taken as an input for $\mathbf{x_t}$ and the $(k+1)^{th}$ instance as the output $\mathbf{y_t}$. The block window is moved across the signal by stride s until it reaches the end of the signal and the total number of blocks can be calculated as $r = \frac{T-k}{s} + 1$. We stacked the block signals and created $\mathbf{X} = (\mathbf{x_{k+1}}; \mathbf{x_{k+s+1}}; \mathbf{x_{k+2s+1}}...; \mathbf{x_{k+(r-1)s+1}}) \in R^{(Mk\times r)}$ and $(k+1)^{th}$ stacked into $\mathbf{Y} = (\mathbf{y_{k+1}}; \mathbf{y_{k+s+1}}; \mathbf{y_{k+2s+1}}...; \mathbf{y_{k+(r-1)s+1}}) \in R^{(M\times r)}$. This can be modeled as a group lasso regression problem (5.8) where block windows are groups and λ is the regularization parameter.

$$\hat{\mathcal{A}} \in \underset{\mathcal{A}}{\operatorname{argmin}} \left\{ \|\mathbf{Y} - \mathcal{A}\mathbf{X}\|_F^2 + \lambda \mathrm{Pen}(\mathcal{A}) \right\} \qquad \text{where } \mathrm{Pen}(\mathcal{A}) = \sum_{i=1}^{Mk} \sum_{j=1}^{M} \|(\mathcal{A}_\mathbf{i})_{j:}\|_2 \tag{5.8}$$

The output of (5.8) helps to produce the predicted $(k+1)^{th}$ instances $\hat{\mathbf{Y}} = \mathcal{A}\mathbf{X}$. The prediction error of each block can be computed using (5.7). In order to test the performance of the models we used different window lengths and strides. We recorded relative prediction error (5.9) corresponding to different window lengths k and strides s. We chose the best window

length k and stride s combinations which gives smaller prediction errors.

$$\epsilon = \frac{\|\hat{\mathbf{Y}} - \mathbf{Y}\|_F}{\|\mathbf{Y}\|_F} \tag{5.9}$$

If there are p different window sizes and q different strides, the relative prediction errors can be represented in the matrix form as

$$\mathbf{\Xi} = \begin{bmatrix} \epsilon_1^1 & \epsilon_1^2 & \dots \\ \vdots & \ddots & \\ \epsilon_p^1 & & \epsilon_p^q \end{bmatrix}$$

In order to select the best window, stride combination we setup the threshold $\tau = 0.1$. If $\epsilon_i^j \leq \tau$, and considered that the corresponding window length and stride are significant for further analysis. In order to clearly identify the best window length stride combination, we created a binary matrix $\mathbf{\Gamma}$ with the entries (5.10) .

$$\mathbf{\Gamma} = \begin{bmatrix} \gamma_1^1 & \gamma_1^2 & \dots \\ \vdots & \ddots & \\ \gamma_p^1 & & \gamma_p^q \end{bmatrix}$$

$$\gamma_i^j = \begin{cases} 1, & \text{if } \epsilon_i^j \leq \tau \\ 0, & \text{otherwise} \end{cases} \tag{5.10}$$

5.4 Dynamic Functional Connectivity with AR/VAR

Dynamic Functional Connectivity accounts for the uncharacteristic activations in the brain by treating the signal as non-static. A sliding window approach is used with a connectivity matrix computed for each window. This approach helps to identify the effect of the uncharacteristic activations within smaller windows. The choice of window size k and stride s is determined by using AR (Sec. 5.3.2) or VAR (Sec. 5.3.3) approaches. Connectivity matrices are generated from inverse covariance matrices for each window using [33].

5.4.1 Graphical Lasso

Graph theoretical approaches are considered a powerful framework to study the functional connectivity of fMRI brain networks. The brain is modeled as an undirected graph where the brain regions correspond to the nodes and the connection between the nodes correspond to the edges. The number of edges incident to the node are called the degree of the node. In the network the brain regions, which is our region of interest, the nodes and strength of the edges represent the degree of functional connectivity between a pair of regions.

The correlation method cannot find true physical connections in the brain network. Since two brain regions might show a high correlation even when they are not directly connected, the high correlation could be due to a strong interaction between the two regions with common input from the third region. We can compute the functional connectivity more accurately using partial correlations [18] [54]. Partial correlation provides better characterization of the brain connectivity than the correlation analysis because partial correlation measures the strong direct interactions between the pairs of brain regions while simultaneously removing the influence of the rest of the brain regions. This is done by measuring the correlation between a pair after having regressed out the effects from all other regions in the brain.

The zero value of the partial correlation between two brain regions is interpreted as their conditional independence given the activity of the other brain regions.

The problem with the estimation of functional connectivity using partial correlation is an ill-posed problem when the degree of the node is larger compared to the number of nodes considered. Therefore, the Graphical Lasso method originally proposed by Friedman et.al [33] implements the sparse regularization of the network connections by using the L_1 penalty. It also helps to interpret the estimated networks from the functional view point. Let the empirical covariance matrix of observations be S as the Graphical Lasso estimates the precision matrix Θ by solving the following optimization problem :

$$\underset{\Theta \succeq 0}{max} \quad [\log \det \Theta - tr(S\Theta) - \rho \|\Theta\|_1] \tag{5.11}$$

which maximizes the log likelihood over the non-negative definite matrices Θ.

We implemented Graphical Lasso to find the precision matrix for the implementation of the VAR model. We selected a significant window stride combination and moved across the fMRI signal. At each step (stride), we centered the window entries corresponding to ROI (Region of Interest, n=116) averages. We constructed sample covariance matrices for each window steps and took the average of all constructed covariance matrices. Finally we had one single covariance matrix per subject. We plugged the covariance matrix to (5.11) and obtained the precision matrix.

5.5 Real data Experiments

5.5.1 Data sets

We tested our proposed technique using two different data sets. From the Autism Brain Imaging Data Exchange we obtained the Kennedy Krieger Institute (KKI) data set. KKI data includes Total: 55 (Age: 8.0-12.8), Autism Spectrum Disorders (ASD): 22 (Age: 8.0-12.5)(11 Autism, 11 Asperger's Disorder)and typical controls (TC): 33 (Age: 8.0-12.8). Next, we used Chess Masters and Novice data (CM&N). The studies have demonstrated that non-invasive MRI techniques are valuable for researchers to investigate the underlying neural mechanism of chess. For professional chess players (e.g., chess grand masters and masters CM/Ms), they show what the structural and functional alterations are due to longtime professional practice, and how these alterations, which relate to their behavior, are largely veiled. The data set reported a multi-modal MRI dataset of 29 professional Chinese chess players, most of which are CM/Ms, and 29 well-matched age and gender chess novices. The objective of the data analysis is classify on the basis of f-MRI data. Our aim is to classify KKI testing subjects into Autism Spectrum Disorder or typical control and CM&N testing subjects into professional chess players or novices.

5.5.2 Static Functional Connectivity based Classification

For each subject, we generated the sample correlation matrix and derive the precision matrix by the covariance matrix. We obtained two vectors by taking the upper triangular part of the correlation and the precision matrices, thus obtaining the correlation and precision vectors with 6670 features. For testing our classification procedure, we carried out 10 fold cross validation. For this purpose we divided the samples into the training and testing portions.

We selected only the significant features on these vectors using three different approaches; the Wilcoxon rank sum test, the T-test and F-test. We standardized both training and testing vectors using parameters (mean and standard deviation) of the training set. We included training vectors to **A** and their true classification (binary vector) **y** then used LeastR function in SLAP package to get the weight vector (w) for classification.

$$\min_{w} \quad \frac{1}{2}\|Aw - y\|_2^2 + \lambda\|w\|_1 \tag{5.12}$$

where $A \in R^{m\times n}$, $y \in R^{m\times 1}$ and $x \in R^{n\times 1}$.

We multiplied the weight matrix and the testing precision vector to obtain predictive classification. Since the values are real we convert it into binary values by comparing with zero. Finally, we evaluate the classification accuracy by Sensitivity (SE), Specificity (SP), Positive predictive value (PPV), Negative predictive value (NPV) and Correct rate (Acc).

Table 5.1 summarizes the accuracies obtained for the Chess Masters and Novice (CM&N) data set. We obtained higher accuracy with the precision matrix than with the correlation matrix. According to our study, the Wilcoxon Rank Sum test provides more accurate feature selection for the CM&N data set.

Table 5.2 summarizes results for the Kennedy Krieger Institute (KKI) data set. According to Table 5.2, results are similar to those for the CM&N data set, while we obtained higher accuracy with the precision matrices. Comparatively, the T-test and F-test indicate more accurate feature selection for the KKI data set.

Figure 5.1 represents the behaviour of the accuracies in the Chess master and novice (CM&N) data set against lambda along the different feature selection techniques (Wilcoxon Rank Sum,T-test and F-test). The figure summarizes the behaviour of both correlation and

precision vectors. It is noticeable when lambda increases the accuracy decreases in most of the cases except the Precision+F-test and Precision+T-test. In particular, the Precision+Wilcoxon Rank Sum test reaches higher accuracy in most regularization parameters.

Finally, Figure 5.2 illustrates the accuracy against lambda (regularization parameter) in the KKI data set along the different feature selection techniques (Wilcoxon Rank Sum,T-test and F-test). Similarly, the figure summarizes the behaviour of both correlation and precision vectors. It can observe the decreasing behaviour of accuracy when lambda increases, yet it is not possible to observe with the correlation+F-test and Precision+F-test. The Precision+T-test reaches higher accuracy for most lambdas.

Table 5.1: Classification Accuracy in the context of static functional connectivity for the Chess Masters and Novice dataset obtained by using the Correlation matrices and Precision matrices

	WRST + LASSO	T-Test + LASSO	F-Test+ LASSO
Accuracy (Corr)	59%	52%	57%
Accuracy (Prec)	70%	66%	69%

Table 5.2: Classification Accuracy in the context of static functional connectivity for the the KKI dataset obtained by using the Corelation matrices and Precision matrices

	WRST + LASSO	T-Test + LASSO	F-Test+ LASSO
Accuracy (Corr)	63%	61%	53%
Accuracy (Prec)	73%	76%	76%

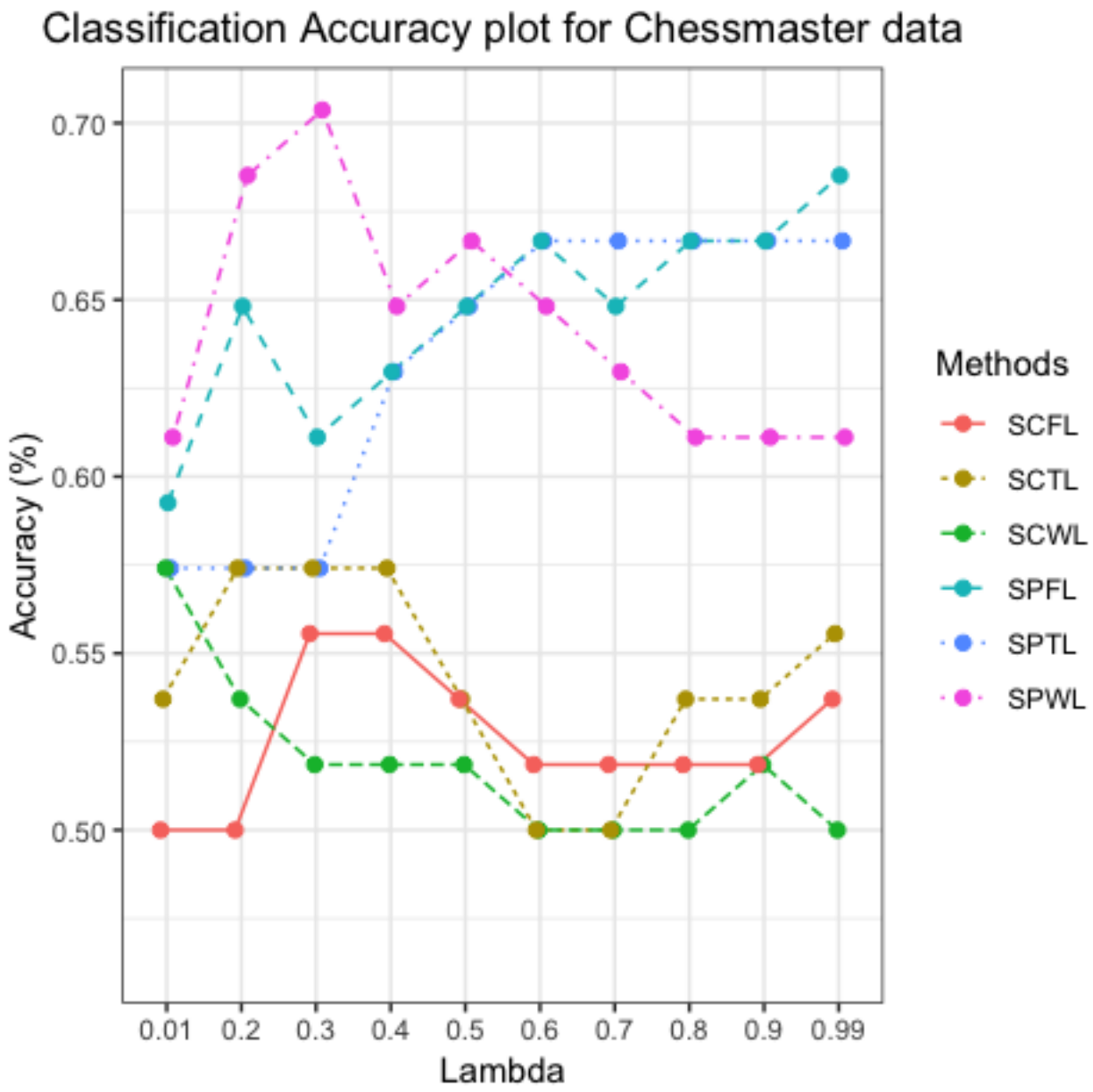

Figure 5.1: Accuracy of Classification of Chess Masters and Novices from various methods for the static functional Connectivity against different λ's in (5.12). Here, S refers to static, C refers to the Corelation matrix, P refers to precision matrix, F refers to $F-$Test, W refers to Wilcoxon rank sum test, T refers to T-test and L refers to Lasso (SCFL stands for Static+ Correlation+F-test+Lasso). The points in the graph represent the accuracy of the corresponding method for the corresponding values of lambda.

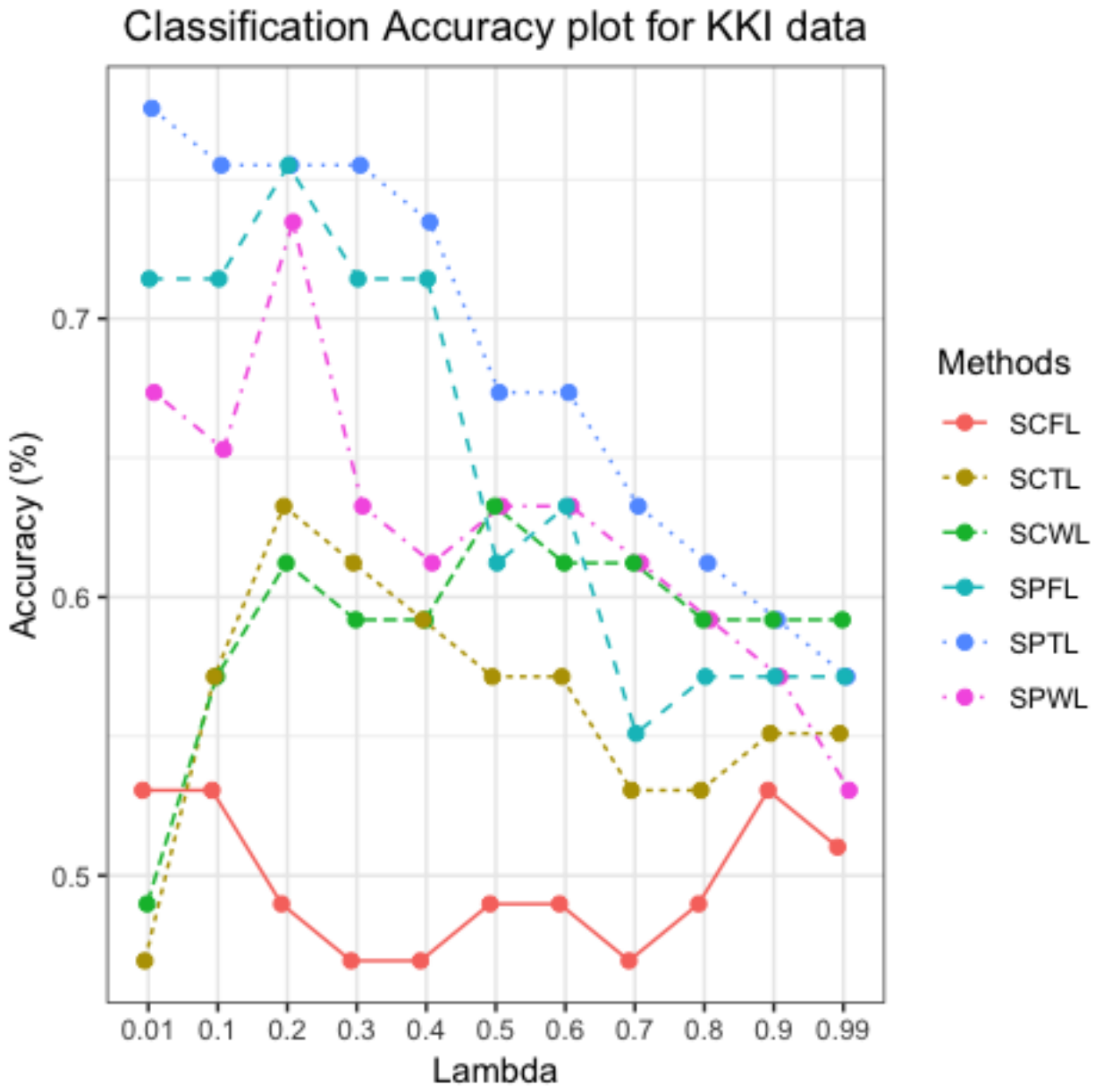

Figure 5.2: Accuracy of Classification for the KKI data set from various methods for the Static functional Connectivity against different λ's in (5.12). Here, S refers to static, C refers to the Corelation matrix, P refers to precision matrix, F refers to F-Test, W refers to Wilcoxon rank sum test and L refers to Lasso. The points in the graph represent the accuracy of the corresponding method for the corresponding values of lambda.

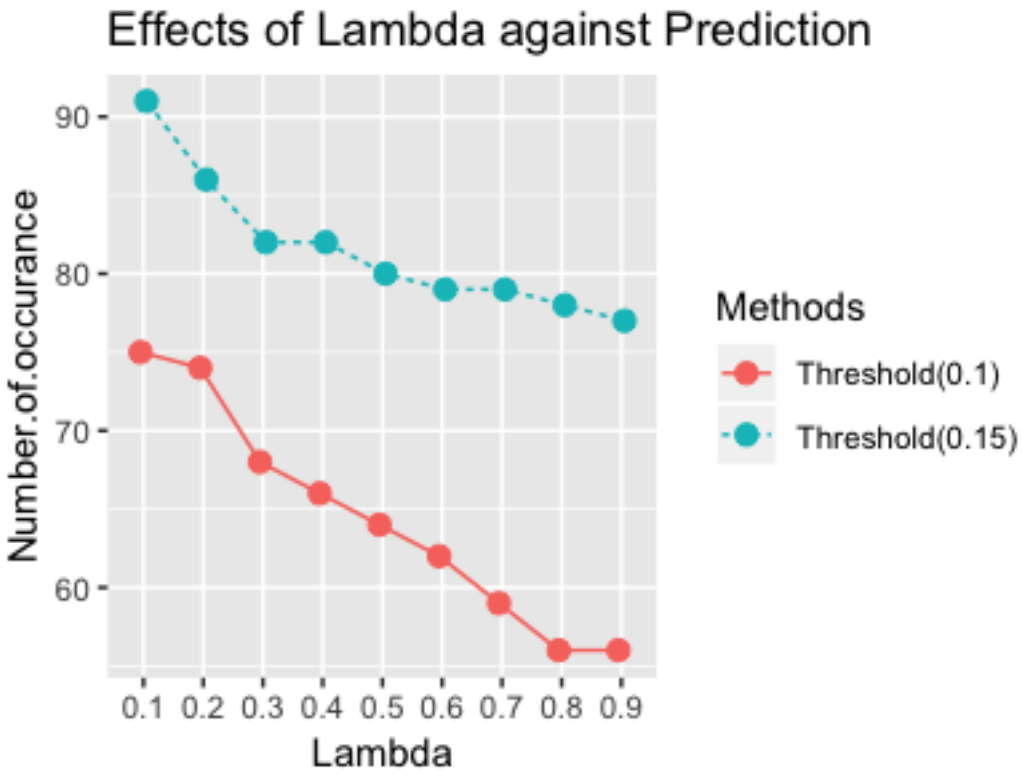

Figure 5.3: In vector auto regression the number of significant window stride combinations occur less than threshold for different λ's in (5.8). Threshold=0.1 (Orange) and Threshold=0.15 (green) for chess masters and novice (CM&N) data.

Table 5.3: In vector auto regression the number of significant window strides combinations occurs less than threshold for different λ's in (5.8). For Threshold=0.1 are results shows in rows 2 and 3. Threshold=0.15 results are shown in rows 5 and 6 for CM&N data.

Thresholed=0.1									
λ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Num of	75	74	68	66	64	62	59	56	56
Thresholed=0.15									
λ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Num of	91	86	82	82	80	79	79	78	77

5.6 Dynamic functional connectivity based classification

In order to determine the order of the VAR processes, 10 subjects were randomly selected from the dataset while ensuring representation from both groups/classes. For the Lasso to Regression order prediction, we fixed λ and then computed the error (5.9) for different window lengths and strides. The process was then repeated across different λ values to study the effect of sparsity on order prediction. These experiments were repeated for determining the order of the VAR process. The observations for VAR can be seen in Table 5.3 and Fig. 5.3 for Chess Masters and Novice data. It can be seen from Fig. 5.3 when lambda increases the number of significant window lengths and strides decreases. Furthermore, at $\lambda = 0.1$ we achieved the highest number of significant window stride combinations, thus we set our regularization parameter as $\lambda = 0.1$ for our further analysis in Chess Masters and Novice data.

Based on the selected λ and their corresponding window lengths k and stride s, we carried out classification in VAR. We selected a significant window stride combination and moved across the fMRI signal. At each step (stride) we centered the window entries' corresponding ROI (Region of Interest, n=116) averages. We constructed a sample covariance matrix for each window step and took the average of all constructed covariance matrices. Finally we had one single covariance matrix per subject. We constructed the Inverse Covariance Matrix (Precision Matrix) using the function "sparse Inverse Covariance" on the SLEP package.

$$\underset{\Theta \succeq 0}{max} \quad [\log |\Theta| - < S, \Theta > -\lambda \|\Theta\|_1] \tag{5.13}$$

where $S \in R^{n \times n}$ is the sample covariance matrix estimated from the data, $\Theta \in R^{n \times n}$ is the (sparse) inverse covariance matrix to be estimated, and $\lambda > 0$ is the regularization parameter.

Then we chose the upper triangular entries form Θ and vectorized it (6670). For testing our classification procedure, we carried out 10 fold cross validation. In each fold we divided the samples into the training and testing portions. Then we divided the training vectors into control and healthy and selected only the significant features using three diferent approaches; the Wilcoxon rank sum test, the T-test and F-test. Finally, we standardized both training and testing vectors using parameters(mean and standard deviation) of the training set. We included training vectors into $\mathbf{A}$ and their true classification $\mathbf{y}$ (binary vector) then used the LeastR function in the SLAP package to get the weight vector (w) for classification.

$$\min_{w} \quad \frac{1}{2}\|Aw - y\|_2^2 + \lambda\|w\|_1 \tag{5.14}$$

where $A \in R^{m\times n}$, $y \in R^{m\times 1}$ and $w \in R^{n\times 1}$.

We multiplied the weight matrix and the testing precision vector to obtain predictive classification. Since the values are real we convert them to binary values by comparing them with zero. Finally, we evaluate the classification accuracy for Sensitivity (SE), Specificity (SP), Positive predictive value (PPV), Negative predictive value (NPV) and Correct rate (Acc).

5.6.1 Dynamic feature selection and comparison

We obtained higher classification accuracy in the dynamic functional connectivity approach comparative to the static functional connectivity approach . Table 5.4 and Table 5.5 represent the highest accuracy of each case(WRST, T-Test and F-Test). We always achieve more accuracy in precision vectors than with the correlation vector based classification. Table 5.4 shows that the Wilcoxon Rank Sum test obtained the highest classification accuracy with precision vectors while in Table 5.5 the F-test with precision vectors achieve the highest accuracy.

In order to demonstrate accuracy distribution along the different window lengths and strides we create Figure 5.6 for Chess Masters and Novice data. The figure illustrates the accuracy distribution along the different window lengths and strides via box plot against differnt lambda's in (5.14). This box-plot confirms that classification accuracies are correspond higher to $\lambda = 0.1,\ \lambda = 0.2$. Similarly we create Figure 5.7 to present the same distribution setting across the regularization parameter for the KKI data set. Hence the box-plot confirms $\lambda = 0.2$ gives better accuracy along other regularization parameters.

In order to discuss the accuracy against the regularization parameter in (5.14), we create the Figure 5.4 for Chess Masters and Novice data. Similarly Figure 5.5 represents the accuracy against lambda in (5.14) for the KKI data set. In general both graphs indicate when the lambda increases and the accuracy decreases. It is clear from Figure 5.4 the precision vectors with the Wilcoxon Rank Sum test gives the highest accuracy for chess master and novice data. In particular Figure 5.5 indicates that the F-test with the precision matrix gives the highest accuracy for KKI data.

Table 5.4: Highest classification accuracy for Chessmasters/Novice in the context of the Dynamic functional connectivity obtained by using precision and correlation matrices.

	WRST + LASSO	T-Test + LASSO	F-Test+ LASSO
Accuracy (Corr)	61%	72%	65%
Accuracy (Prec)	81%	74%	69%

Table 5.5: Highest classification accuracy for the KKI dataset in the context of Dynamic functional connectivity obtained by using precision and correlation matrices.

	WRST + LASSO	T-Test + LASSO	F-Test+ LASSO
Accuracy (Corr)	63%	61%	69%
Accuracy (Prec)	80%	80%	82%

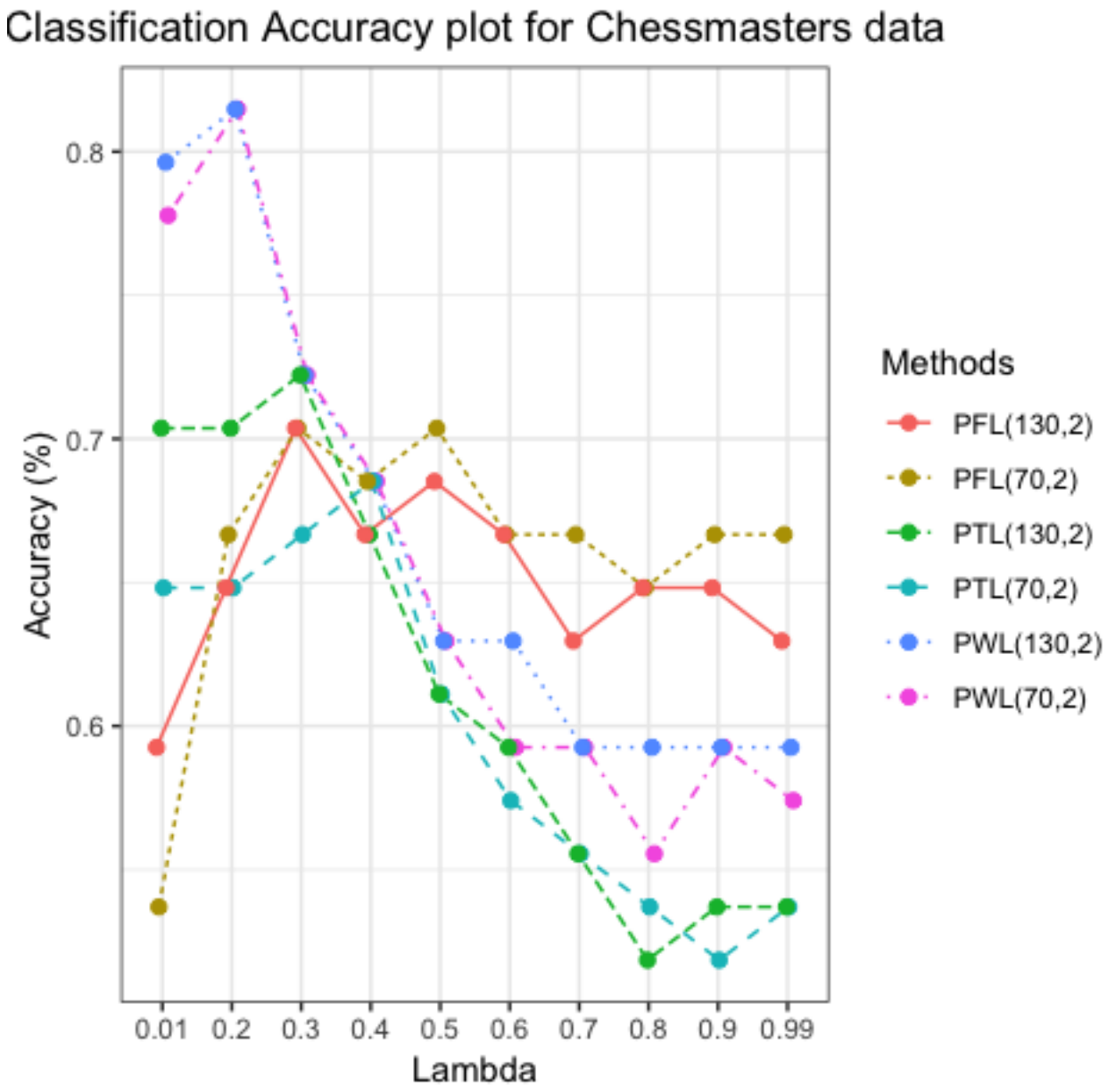

Figure 5.4: Accuracy of Classification of Chess masters/Novice from various methods for dynamic functional connectivity across different λ's in (5.14). Here P refers to the precision matrix, F refers to F-Test, T refers to T-test, W refers to Wilcoxon rank sum test and L refers to Lasso, then (PFL stands for precision+F-test+Lasso /PWL stands for precision+Wilcoxon+Lasso) and $(70, 2)$ refers to the result obtained using the window length 70 and stride 2.

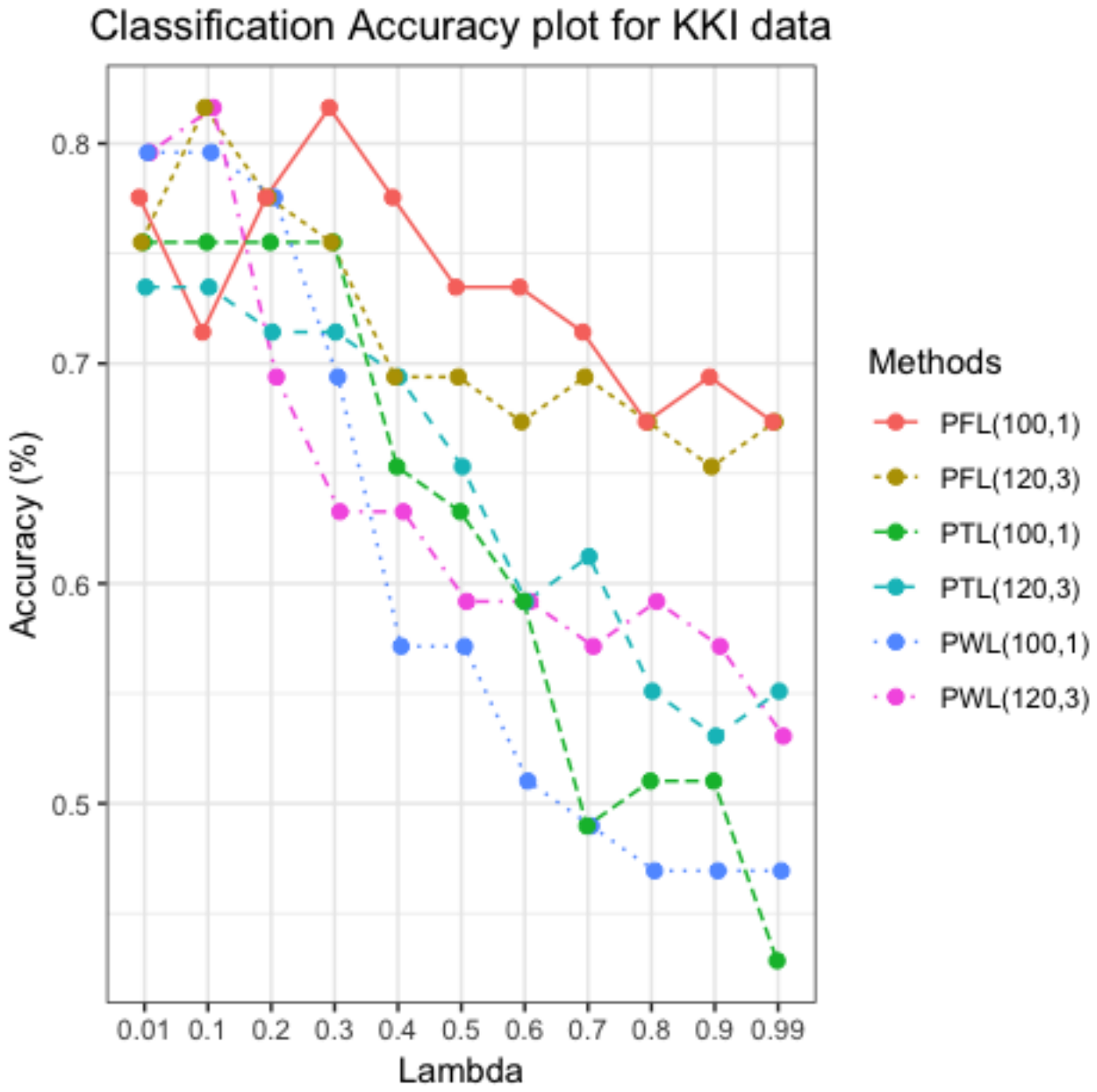

Figure 5.5: Accuracy of Classification of he KKI dataset from various methods for dynamic functional connectivity across different λ's in (5.14). Here P refers to the precision matrix, F refers to F-Test, T refers to T-test and W refers to Wilcoxon rank sum test, L refers to Lasso, then ((PFL stands for precision+F-test+Lasso /PWL stands for precision+Wilcoxon+Lasso) and $(120, 3)$ refers to the result obtained using the window length 120 and stride 3.

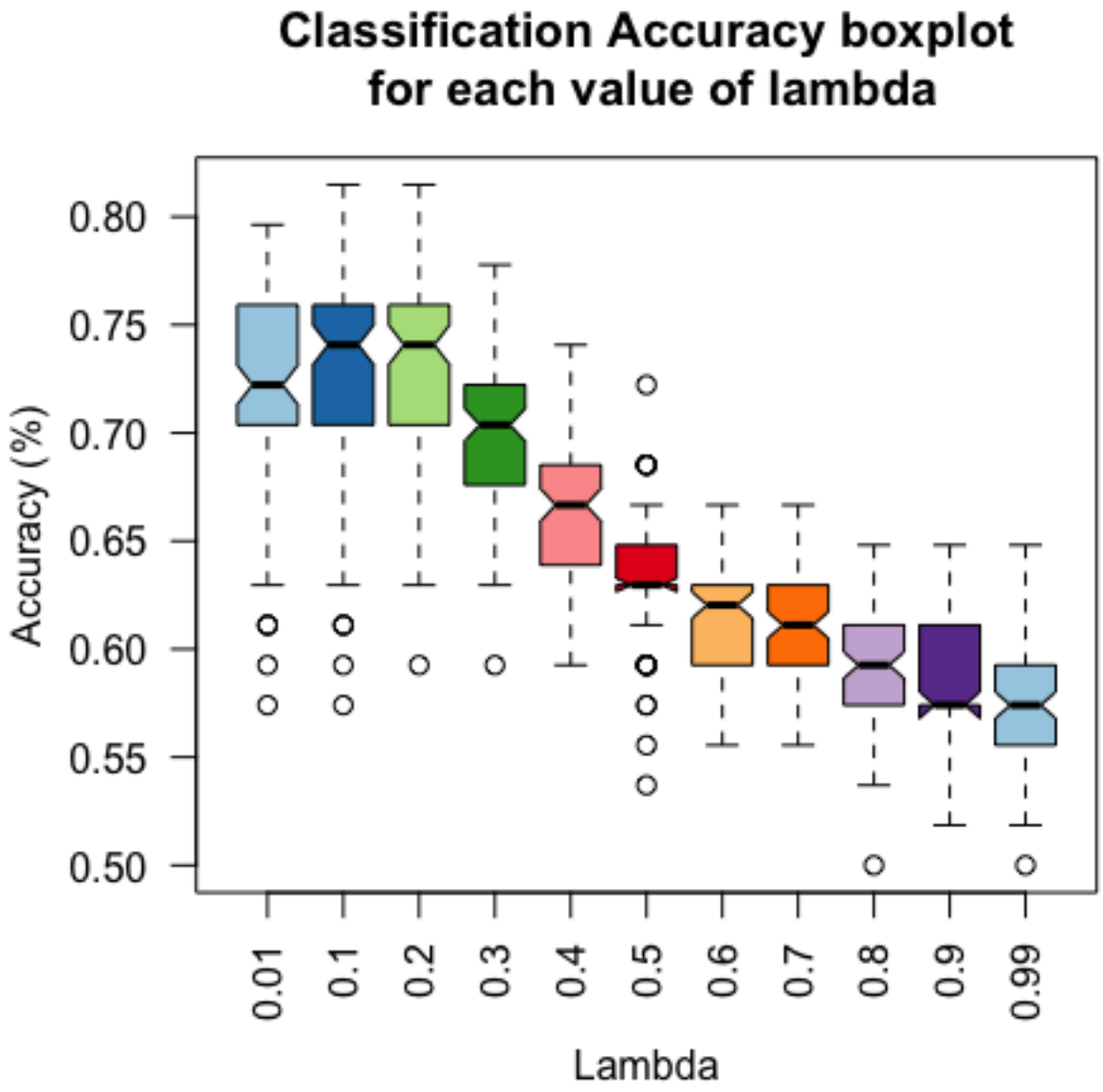

Figure 5.6: Distribution of accuracies of different window lengths and strides against different regularization parameters for classification of Chessmasters and Novice data for the case where precision matrix as connectivity, Wilcoxon rank sum test for feature selection and Lasso for classification for dynamic functional connectivity across different λ's in (5.14). In the Figure, the height of the box represents the inter-quartile range, the horizontal line inside the box represents the median and small circles below and above the whiskers represents the outliers of the accuracy of classification.

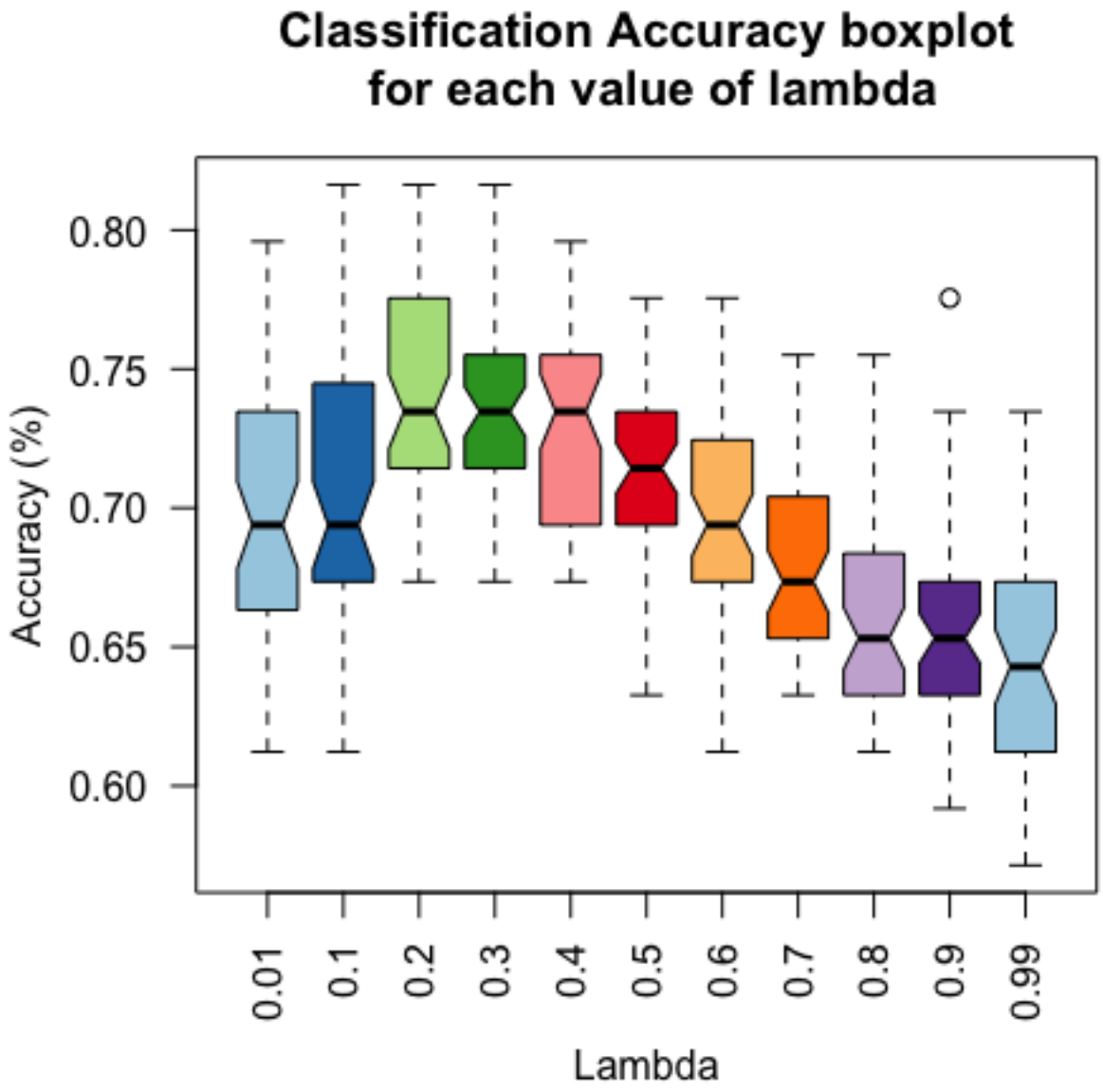

Figure 5.7: Distribution of accuracies of different window lengths and strides against different regularization parameters for of classification of in KKI dataset for the case where precision matrix as connectivity, F- test for feature selection and Lasso for classification for dynamic functional connectivity across different λ's in (5.14). In the Figure, the height of the box represents the inter-quartile range, the horizontal line inside the box represents the median and small circles below and above the whiskers represents the outliers of the accuracy of classification.

CHAPTER 6: CONCLUSION AND FUTURE WORK

6.1 Conclusion

We investigate theoretically and via a limited simulation study, the effect of clustering on the accuracy of recovery in ill-posed linear inverse problems. As we stated earlier, in many applications leading to such problems, clustering is carried out at a pre-processing step and later is totally forgotten when it comes to error evaluation. We conclude that when the sizes of the vectors and the number of functions under investigation are large enough, clustering does not have an adverse effect on the precision. However, as both Corollary 1 and Tables 1–4 show, the improvement due to clustering is more significant when the problem is less ill-posed. It is easy to notice that the difference in precision of estimators with and without clustering is more pronounced for larger values of λ and for $\alpha = \beta = 0$. Indeed, in the case when the problem is not ill-posed ($\alpha = \beta = \gamma = 0$ in (4.3)), as findings of Klopp *et al.* [41] show, clustering always improves estimation precision. On the other hand, when the problem is severely ill-posed ($\alpha > 0, \beta > 0$), the recovery can be very poor even when clustering errors are small or even zero (see, e.g. Table 4 where reconstruction errors are high even when clustering errors are small). This is due to the fact that the reduction in the noise level due to clustering is not sufficient to counteract the ill-posedness of the problem and, thus, and does not lead to a meaningful improvement in estimation accuracy.

6.2 Future Work

At this moment we are working on the extension of our research work on clustering in statistical inverse problems. Currently, we consider a homogeneous set of functions and do not take into account the distances between them. We are planning to extend the methodology to the set of inhomogeneous functions and take the distances between them into account.

We are finishing the rest of the work in Learning Functional Brain Connectivity through Auto-Regressive (AR) Models. We are also planning to extend our analysis with more data sets and identify the top brain regions for classification.

APPENDIX A: DERIVATIONS AND PROOFS PART 1

In order to prove Theorem 1, we use Lemma A1 of Bunea *et al.* (2007), which we will reformulate for the squared risk case.

Lemma 3 *Let Θ be a set of functions of cardinality card$(\Theta) \geq 2$ such that*
(i) $\|f-g\|^2 \geq 4\delta^2,\ for\ f,g \in \Theta,\ f \neq g$,
(ii) the Kullback divergences $K(P_f, P_g)$ between the measures P_f and P_g satisfy the inequality $K(P_f, P_g) \leq \log(card(\Theta))/2,\ for\ f,\ g \in \Theta$.
Then, for some absolute positive constant C_1, one has

$$\inf_{f_n} \sup_{f \in \Theta} \mathbb{E}_f \|f_n - f\|^2 \geq C_1 \delta^2$$

where $\inf_{f_n}$ denotes the infimum over all estimators.

In order to obtain lower bounds, we introduce a triangular Toeplitz matrix associated with Laurent series $(1-z)^{-r}$ and denote by $Q^{(L)} = T_L\left((1-z)^{-r}\right)$ its reduction to the set of indices $0 \leq l \leq L-1$. Following Vareschi (2013), consider function

$$h(t) = \sum_{l=0}^{\infty} h_l \varphi_l(t) \quad \text{with } h_l = \frac{(-1)^l}{\log(l \vee e)} \binom{-1/2}{l} \tag{A.1}$$

Denote $\boldsymbol{\theta}^{(L)} = (\theta_0, \cdots, \theta_{L-1})^T = Q^{(L)} \mathbf{h}^{(L)}$ where $\mathbf{h}_L$ is the vector of the first L coefficients of function h in (A.1). In what follows we shall use Lemma 6.5 of Vareschi (2013) that was in the original version of the paper posted on ArXiv but did not make it to the published version of Vareschi (2015).

Lemma 4 *Let $h(t)$ be a defined in (A.1) and $\boldsymbol{\theta}^{(L)} = Q^{(L)} \mathbf{h}^{(L)}$ where $Q^{(L)} = T_L\left((1-z)^{-r}\right)$ and $\mathbf{h}^{(L)}$ are reductions of the infinite-dimensional Toeplitz matrix $T\left((1-z)^{-r}\right)$ and vector $\mathbf{h}$ of coefficients of $h(t)$ to the set of indices $0 \leq l \leq L-1$. Then, $h(t)$ is square integrable*

and there exist positive constants C_{r1} *and* C_{r2} *that depend on* r *only such that for all* $r \geq 1$ *and any* $l \geq 0$ *one has*

$$C_{r1}\,\frac{(l \vee 1)^{r-1/2}}{\log(l \vee e)} \leq \theta_l \leq C_{r2}\,(l \vee 1)^{r-1/2}. \tag{A.2}$$

Let ϑ be a matrix with components $\vartheta_{k_1,k_2} = \{-1, 1\}$, $k_1 = 0, 1, \cdots, 2^{j_1}-1$, $k_2 = 0, 1, \cdots, 2^{j_2}-1$. Denote the set of all possible values of ϑ by Θ and let functions f_{L,j_1,j_2} be of the form

$$f_{L,j_1,j_2}(t, x_1, x_2) = \rho\; q_L(t)\; p_{j_1,j_2}(x_1, x_2) \tag{A.3}$$

$$q_L(t) = \sum_{l=0}^{L-1} \theta_l\, \varphi_l(t), \quad p_{j_1,j_2}(x_1, x_2) = \sum_{k_1=0}^{2^{j_1}-1} \sum_{k_2=0}^{2^{j_2}-1} \vartheta_{k_1,k_2}\; \psi_{j_1,k_1}(x_1)\, \psi_{j_2,k_2}(x_2), \tag{A.4}$$

where $\boldsymbol{\theta}^{(L)}$ is the vector with components θ_l, $l = 0, \cdots, L-1$ where $\boldsymbol{\theta}^{(L)} = Q^{(L)}\mathbf{h}^{(L)}$ and $Q^{(L)}$ and $\mathbf{h}^{(L)}$ are defined above. Since $f_{L,j_1,j_2} \in \mathcal{B}^{s_1,s_2,s_3}_{\gamma,\beta}(A)$, Lemma 4 implies that one can choose

$$\rho^2 = C_r A^2 2^{-2j_1(s_1+\frac{1}{2})-2j_2(s_2+\frac{1}{2})}\,(L \vee 1)^{-2(r+s_3)}\,\exp\left\{-2\gamma L^{\beta}\right\}, \tag{A.5}$$

where $0 < C_r \leq C_{r2}^2/2r$. If $\tilde{f}_{L,j_1,j_2}$ is of the form (A.3) but with $\tilde{\vartheta}_{k_1,k_2} \in \Theta$ instead of ϑ_{k_1,k_2}, then, by Lemma 4, the L^2-norm of the difference is of the form

$$\|\tilde{f}_{L,j_1,j_2} - f_{L,j_1,j_2}\|_2^2 = \rho^2 \left(\sum_{l=0}^{L-1} \theta_l^2\right) \left(\sum_{k_1=0}^{2^{j_1}-1} \sum_{k_2=0}^{2^{j_2}-1} \mathbb{I}\left(\tilde{\vartheta}_{k_1,k_2} \neq \vartheta_{k_1,k_2}\right)\right) \geq \frac{C_{r1}^2 \rho^2 H\left(\tilde{\vartheta}, \vartheta\right)(L \vee 1)^{2r}}{2r\,[\log(L \vee e)]^2}.$$

Here $H\left(\tilde{\vartheta}, \vartheta\right)$ is the Hamming distance between the binary sequences $\mathrm{vec}(\vartheta)$ and $\mathrm{vec}(\tilde{\vartheta})$ where $\mathrm{vec}(\vartheta)$ is a vectorized version of matrix ϑ.

Observe that matrix ϑ has $\aleph = 2^{j_1+j_2}$ components, and hence, $\mathrm{card}(\Theta) = 2^{\aleph}$. In order to find a lower bound for $H\left(\tilde{\vartheta}, \vartheta\right)$, we apply the Varshamov-Gilbert lemma which states that

one can choose a subset Θ_1 of Θ, of cardinality of at least $2^{\aleph/8}$, and such that $H\left(\tilde{\vartheta},\vartheta\right) \geq \frac{\aleph}{8}$ for any $\vartheta, \tilde{\vartheta} \in \Theta_1$. Hence, for any $\vartheta, \tilde{\vartheta} \in \Theta_1$, one has the following expression for δ^2 defined in Lemma 3:

$$\|\tilde{f}_{L,j_1,j_2} - f_{L,j_1,j_2}\|^2 \geq \frac{C_{r1}^2 \rho^2 2^{j_1+j_2}(L \vee 1)^{2r}}{16r\,[\log(L \vee e)]^2} = 4\delta^2. \tag{A.6}$$

Note that, since $\left|\tilde{\vartheta}_{l,k,k'} - \vartheta_{l,k,k'}\right| \leq 2$, due to (A.3) and (A.4), the Kullback divergence can be written as

$$K(P_f, P_{\tilde{f}}) = \left(2\varepsilon^2\right)^{-1} \|(\tilde{f} - f) * g\|_2^2 \leq 2\varepsilon^{-2}\,\rho^2 4 2^{j_1+j_2} \|q_L * g\|_2^2 = 2\varepsilon^{-2}\,\rho^2 4 2^{j_1+j_2} \|G^{(L)}\boldsymbol{\theta}^{(L)}\|_2^2, \tag{A.7}$$

where matrix $G^{(L)}$ and vector $\boldsymbol{\theta}^{(L)}$ are defined in (3.9) and Lemma 4, respectively. By Lemma **??** in section **??**, and under Assumptions **A1** and **A2**, one obtains that $G^{(L)} = T_L((1-z)^r v(z))$ and $\|T_L(v(z))\|^2 = \lambda_{\max}[T_L^T(v(z))T_L(v(z))] < \|v\|_{circ}^2 < \infty$. Therefore, $G^{(L)}\boldsymbol{\theta}^{(L)} = G^{(L)}Q^{(L)}\mathbf{h}^{(L)} = T_L((1-z)^r v(z))T_L((1-z)^{-r})\mathbf{h}^{(L)}$ and

$$\|G^{(L)}\boldsymbol{\theta}^{(L)}\|_2^2 = \|T_L(v(z))\mathbf{h}^{(L)}\|_2^2 \leq \|T_L(v(z))\|^2 \|\mathbf{h}^{(L)}\|_2^2 \leq \|v\|_{circ}^2 \|h\|_2^2 < \infty \tag{A.8}$$

where $\|h\|_2^2$ is the L^2-norm of the function $h(t)$ and $\|h\|_2 < \infty$ due to Lemma 4. Combination of (A.7) and (A.8) yields $K(P_f, P_{\tilde{f}}) \leq \tilde{C}\epsilon^{-2}\rho^2 2^{j_1+j_2}/16$ where $\tilde{C} = 128\|v\|_{circ}^2\|h\|_2^2$. Application of Lemma 3 requires the constraint

$$K(P_f, P_{\tilde{f}}) \leq \log(\mathrm{card}(\Theta_1)) \leq 2^{j_1+j_2}/16.$$

Therefore, one can choose $\rho^2 = \epsilon^2/\tilde{C}$, so that, by Lemma 3 for some $C_1 > 0$ one has

$$\inf_{f_n} \sup_{f \in \Theta} \mathbb{E}_f \|f_n - f\|^2 \geq C_1 \epsilon^2 2^{j_1+j_2}\,(L \vee 1)^{2r}\,[\log(L \vee e)]^{-2} \tag{A.9}$$

where L, j_1 and j_2 are such that

$$2^{2j_1(s_1+\frac{1}{2})+2j_2(s_2+\frac{1}{2})}\,(L\vee 1)^{2(r+s_3)}\,\exp\left\{2\gamma L^{\beta}\right\}=C_2A^2\epsilon^{-2} \tag{A.10}$$

with $C_2=C_r\tilde{C}$. Thus, one needs to choose j_1, j_2 and L that maximize $2^{j_1+j_2}\,(L\vee 1)^{2r}\,[\log(L\vee e)]^{-2}$ subject to condition (A.10). Denote

$$\tau_\epsilon=\log(A^2\varepsilon^{-2}) \tag{A.11}$$

It is easy to check that the solution of the above linear constraint optimization problem is of the form $\{j_1,j_2,L\}=\left\{0,0,[A^2\varepsilon^{-2}]^{\frac{1}{2s_3+2r}}\right\}$ if $s_3\leq\min\{2rs_1,2rs_2\}$ and $\gamma=\beta=0$, $\{j_1,j_2,L\}=\left\{0,(\log(2))^{-1}(2s_2+1)^{-1}\tau_\varepsilon,e\right\}$ if $s_1\geq s_2$, $s_3\geq 2rs_2-2s_2-1$ and $\gamma=\beta=0$, $\{j_1,j_2,L\}=\left\{(\log(2))^{-1}(2s_1+1)^{-1}\tau_\varepsilon,0,e\right\}$ if $s_1\leq s_2$ and $s_3\geq 2rs_1-2s_1-1$ and $\gamma=\beta=0$. $\{j_1,j_2,L\}=\left\{0,(\log(2))^{-1}(2s_2+1)^{-1}\tau_\varepsilon,e\right\}$ if $s_1\geq s_2$ and $\gamma>0,\beta>0$, and $\{j_1,j_2,L\}=\left\{(\log(2))^{-1}(2s_1+1)^{-1}\tau_\varepsilon,0,e\right\}$ if $s_1\leq s_2$ and $\gamma>0,\beta>0$. By noting that

$$\frac{s_3}{s_3+r}\leq\min\left\{\frac{2s_2}{2s_2+1},\frac{2s_1}{2s_1+1}\right\},\quad if\ s_3\leq\min\{2rs_1,2rs_2\},\quad \gamma=\beta=0, \tag{A.12}$$

$$\frac{2s_1}{2s_1+1}\leq\min\left\{\frac{2s_2}{2s_2+1},\frac{s_3}{s_3+r}\right\},\quad if\ s_1\leq\min\{s_3/2r,s_2\},\quad \gamma=\beta=0, \tag{A.13}$$

$$\frac{2s_2}{2s_2+1}\leq\min\left\{\frac{2s_1}{2s_1+1},\frac{s_3}{s_3+r}\right\},\quad if\ s_2\leq\min\{s_3/2r,s_1\},\quad \gamma=\beta=0, \tag{A.14}$$

and

$$\frac{2s_1}{2s_1+1}\leq\frac{2s_2}{2s_2+1},\quad if\ s_1\leq s_2,\quad \gamma>0,\beta>0. \tag{A.15}$$

we then choose the highest lower bounds in (A.9). This completes the proof of the theorem.

The proof of Theorem 2. Denote

$$\mu = \begin{cases} \min\left\{\frac{s_3}{s_3+r}, \frac{2s_2}{2s_2+1}, \frac{2s_1}{2s_1+1}\right\}, & if \quad \gamma = \beta = 0, \\ \min\left\{\frac{2s_2}{2s_2+1}, \frac{2s_1}{2s_1+1}\right\}, & if \qquad \gamma > 0, \beta > 0. \end{cases} \tag{A.16}$$

$$\chi_{\varepsilon,A} = \left[A^{-2}\varepsilon^2 \log(1/\varepsilon)\right] \tag{A.17}$$

$$2^{j_{10}} = \left[\chi_{\varepsilon,A}\right]^{-\frac{\mu}{2s_1}}, \quad 2^{j_{20}} = \left[\chi_{\varepsilon,A}\right]^{-\frac{\mu}{2s_2}} \tag{A.18}$$

and

$$M_0 = \begin{cases} \left[\chi_{\varepsilon,A}\right]^{-\frac{\mu}{2s_3}} & \text{if } \ \gamma = \beta = 0 \\ \left[\frac{\log(1/\varepsilon)}{\gamma}\right]^{1/\beta} & \text{if } \ \gamma > 0, \beta > 0. \end{cases} \tag{A.19}$$

and notice that with the choices of J_1, J_2 and M given by (3.6), the estimation error can be decomposed into the sum of four components as follows

$$\mathbb{E}\|\widehat{f}_n - f\|^2 \leq \sum_{\omega}\sum_{l=0}^{\infty} \mathbb{E}\|\widehat{\theta}_{l:\omega}\mathbb{I}\left(|\widehat{\theta}_{l:\omega}| > \lambda_{l,\varepsilon}\right) - \theta_{l:\omega}\|^2 \leq R_1 + R_2 + R_3 \tag{A.20}$$

where

$$
\begin{aligned}
R_1 &= \sum_{j=m_0}^{J-1}\sum_{j'=m_0'}^{J'-1}\sum_{l=0}^{M-1}\sum_{k,k'}\mathbb{E}\left[\left|\widehat{\theta}_{l:\omega}-\theta_{l:\omega}\right|^2\mathbb{I}\left(\left|\widehat{\theta}_{l:\omega}\right|>\lambda_{l,\varepsilon}\right)\right] \\
R_2 &= \sum_{j=m_0}^{J-1}\sum_{j'=m_0'}^{J'-1}\sum_{l=0}^{M-1}\sum_{k,k'}\left|\widehat{\theta}_{l:\omega}\right|^2\Pr\left(\left|\widehat{\theta}_{l:\omega}\right|<\lambda_{l,\varepsilon}\right) \\
R_3 &= \left(\sum_{j=J}^{\infty}\sum_{j'=J'}^{\infty}\sum_{l=M}^{\infty}+\sum_{j=m_0}^{J-1}\sum_{j'=J'}^{\infty}\sum_{l=M}^{\infty}+\sum_{j=J}^{\infty}\sum_{j'=m_0'}^{J'-1}\sum_{l=M}^{\infty}+\sum_{j=J}^{\infty}\sum_{j'=J'}^{\infty}\sum_{l=0}^{M-1}\cdots\right)\sum_{k,k'}\left|\theta_{l:\omega}\right|^2
\end{aligned}
$$

For R_3, one uses assumption (3.2) to obtain,

$$
\begin{aligned}
R_3 &= O\left(\left(\sum_{j_1=m_0}^{J_1-1}\sum_{j_2=m_0'}^{J_2-1}\sum_{l=M}^{\infty}+\sum_{j_1=m_0}^{J-1}\sum_{j_2=J_2}^{\infty}\sum_{l=1}^{M}+\sum_{j_1=J_1}^{\infty}\sum_{j_2=m_0'}^{J_2-1}\sum_{l=1}^{M}\right)A^2 2^{-2j_1s_1-2j_2s_2}l^{-2s_3}\exp\{-2\gamma l^{\beta}\}\right) \\
&= O\left(A^2 2^{-2J_1s_1}+A^2 2^{-2J_2s_2}+A^2M^{-2s_3}\exp\{-2\gamma M^{\beta}\}\right)
\end{aligned}
\tag{A.21}
$$

If $\gamma=\beta=0$, then since $M\asymp[\varepsilon^2]^{-1/2r}$, R_3 becomes

$$
\begin{aligned}
R_3 &= O\left(A^2\left[A^{-2}\varepsilon^2\right]^{2s_1}+A^2\left[A^{-2}\varepsilon^2\right]^{2s_2}+A^2\left[A^{-2}\varepsilon^2\right]^{\frac{2s_3}{2r}}\right) \\
&= O\left(A^2\left[\chi_{\varepsilon,A}\right]^{\mu}\right)
\end{aligned}
\tag{A.22}
$$

If $\gamma>0,\beta>0$, then

$$
\begin{aligned}
R_3 &= O\left(A^2\left[A^{-2}\varepsilon^2\right]^{2s_1}+A^2\left[A^{-2}\varepsilon^2\right]^{2s_2}\right) \\
&= O\left(A^2\left[\chi_{\varepsilon,A}\right]^{\min\left\{\frac{2s_2}{2s_2+1},\frac{2s_1}{2s_1+1}\right\}}\right)
\end{aligned}
\tag{A.23}
$$

To evaluate the remaining two terms, notice that both R_1 and R_2 can be partitioned into

the sum of two error terms as follows

$$R_1 \leq R_{11} + R_{12}, \quad R_2 \leq R_{21} + R_{22} \tag{A.24}$$

where

$$R_{11} = \sum_{j_1=m_0}^{J_1-1} \sum_{j_2=m_0'}^{J_2-1} \sum_{l=1}^{M} \sum_{k,k'} \mathbb{E}\left[\left|\widehat{\theta}_{l:\omega} - \theta_{l:\omega}\right|^2 \mathbb{I}\left(\left|\widehat{\theta}_{l:\omega} - \theta_{l:\omega}\right| > \frac{1}{2}\lambda_{l;\varepsilon}\right)\right] \tag{A.25}$$

$$R_{12} = \sum_{j_1=m_0}^{J_1-1} \sum_{j_2=m_0'}^{J_2-1} \sum_{l=1}^{M} \sum_{k,k'} \mathbb{E}\left[\left|\widehat{\theta}_{l:\omega} - \theta_{l:\omega}\right|^2 \mathbb{I}\left(|\theta_{l:\omega}| > \frac{1}{2}\lambda_{l;\varepsilon}\right)\right] \tag{A.26}$$

$$R_{21} = \sum_{j_1=m_0}^{J_1-1} \sum_{j_2=m_0'}^{J_2-1} \sum_{l=1}^{M} \sum_{k,k'} |\theta_{l:\omega}|^2 \Pr\left(\left|\widehat{\theta}_{l:\omega} - \theta_{l:\omega}\right| > \frac{1}{2}\lambda_{l;\varepsilon}\right) \tag{A.27}$$

$$R_{22} = \sum_{j_1=m_0}^{J_1-1} \sum_{j_2=m_0'}^{J_2-1} \sum_{l=1}^{M} \sum_{k,k'} |\theta_{l:\omega}|^2 \mathbb{I}\left(|\theta_{l:\omega}| < \frac{3}{2}\lambda_{l;\varepsilon}\right) \tag{A.28}$$

Combining (A.25) and (A.27) and applying Cauchy-Schwarz inequality, Lemma 2 and the fact that $M \asymp [\varepsilon^2]^{-1/2r}$, yields

$$\begin{aligned} R_{11} + R_{21} &= O\left(\sum_{j_1=m_0}^{J_1-1} \sum_{j_2=m_0'}^{J_2-1} \sum_{l=1}^{M} \left(2^{j_1+j_2}\varepsilon^2 l^{2r-1}\varepsilon^{\tau/2} + \varepsilon^{\tau} \sum_{k,k'} |\theta_{l:\omega}|^2\right)\right) \\ &= O\left(\varepsilon^2 2^{J_1+J_2} M^{2r} \left(\varepsilon^2\right)^{\frac{\tau}{4}} + A^2\varepsilon^{\tau}\right) \\ &= O\left(\left(\varepsilon^2\right)^{\frac{\tau}{4}-2} + A^2\varepsilon^{\tau}\right) \end{aligned}$$

Hence, for $\tau \geq 12$ and under condition (3.8), as $\varepsilon \to 0$, one has

$$R_{21} + R_{31} = O\left(\varepsilon^2\right) = O\left(A^2\left[\chi_{\varepsilon,A}\right]^{\mu}\right) \tag{A.29}$$

Now, combining (A.26) and (A.28), and using (3.3) and (3.7), one obtains

$$\begin{aligned}\Delta = R_{12} + R_{22} &= O\left(\sum_{j_1=m_0}^{J_1-1}\sum_{j_2=m_0'}^{J_2-1}\sum_{l=1}^{M}\sum_{k,k'}\min\left\{|\theta_{l:\omega}|^2, \varepsilon^2\log(1/\varepsilon)l^{-1}\|(\mathbf{G}^{(l)})^{-1}\|^2\right\}\right) \\ &= O\left(\sum_{j_1=m_0}^{J_1-1}\sum_{j_2=m_0'}^{J_2-1}\sum_{l=1}^{M}\min\left\{\sum_{k,k'}|\theta_{l:\omega}|^2, 2^{j_1+j_2}\varepsilon^2\log(1/\varepsilon)l^{2r-1}\right\}\right) \quad \text{(A.30)}\end{aligned}$$

Then, Δ can be decomposed into three components, Δ_1, Δ_2 and Δ_3, as follows

$$\Delta_1 = O\left(\left(\sum_{j_1=j_{10}+1}^{J_1-1}\sum_{j_2=m_0'}^{J_2-1}\sum_{l=1}^{M} + \sum_{j_1=m_0}^{J_1-1}\sum_{j_2=j_{20}+1}^{J_2-1}\sum_{l=1}^{M} + \sum_{j_1=m_0}^{J_1-1}\sum_{j_2=m_0'}^{J_2-1}\sum_{l=M_0}^{M}\right)\sum_{k,k'}|\theta_{l:\omega}|^2\right) \quad \text{(A.31)}$$

$$\Delta_2 = O\left(\sum_{j_1=m_0}^{j_{10}}\sum_{j_2=m_0'}^{j_{20}}\sum_{l=1}^{M_0} A^2 2^{j_1+j_2}\left[\chi_{\varepsilon,A}\right] l^{2r-1}\mathbb{I}\left(\eta^c_{l:j_1,j_2}\right)\right) \quad \text{(A.32)}$$

$$\Delta_3 = O\left(\sum_{j_1=m_0}^{j_{10}}\sum_{j_2=m_0'}^{j_{20}}\sum_{l=1}^{M_0}\left[\sum_{k,k'}|\theta_{l:\omega}|^2\right]\mathbb{I}\left(\eta_{l:j_1,j_2}\right)\right) \quad \text{(A.33)}$$

where $\eta_{l:j_1,j_2} = \left\{l, j_1, j_2 : 2^{j_1+j_2}l^{2r} > \left[\chi_{\varepsilon,A}\right]^{\mu-1}\right\}$. For Δ_1, it is easy to see that for j_{10}, j_{20} and M_0 given in (A.18),

$$\Delta_1 = O\left(A^2 2^{-2j_{10}s_1} + A^2 2^{-2j_{20}s_2} + A^2 M_0^{-2s_3}\exp\{-2\gamma M_0^{\beta}\}\right)$$

Consequently, if $\gamma = \beta = 0$, as $\varepsilon \to 0$, one has

$$\Delta_1 = O\left(A^2\left[\chi_{\varepsilon,A}\right]^{\mu}\right) \quad \text{(A.34)}$$

If $\gamma > 0, \beta > 0$, then

$$\begin{aligned}\Delta_1 &= O\left(A^2 2^{-2j_{10}s_1} + A^2 2^{-2j_{20}s_2}\right) \\ &= O\left(A^2 \left[\chi_{\varepsilon,A}\right]^{\min\left\{\frac{2s_2}{2s_2+1}, \frac{2s_1}{2s_1+1}\right\}}\right) \end{aligned} \tag{A.35}$$

For Δ_2 in (A.32), as $\varepsilon \to 0$, one obtains

$$\Delta_2 = O\left(A^2 \left[A^{-2}\varepsilon^2 \log(1/\varepsilon)\right] \left[\chi_{\varepsilon,A}\right]^{\mu-1}\right) = O\left(A^2 \left[\chi_{\varepsilon,A}^{\alpha}\right]^{\mu}\right) \tag{A.36}$$

In order to evaluate (A.33), we need to consider five different cases.

Case 1: $\gamma = \beta = 0$**,** $s_1 \leq \min\{s_2, \frac{s_3}{2r}\}$**.** In this case, $\mu = \frac{2s_1}{2s_1+1}$, (A.33) becomes, as $\varepsilon \to 0$

$$\begin{aligned}\Delta_3 &= O\left(A^2 \sum_{j_1=m_0}^{j_{10}-1} \sum_{l=1}^{M_0} \sum_{j_2=m_0'}^{j_{20}-1} 2^{-2j_1 s_1 - 2j_2 s_2} l^{-2s_3} \mathbb{I}\left(2^{j_1} > 2^{-j_2} \frac{\left[\chi_{\varepsilon,A}\right]^{\mu-1}}{l^{2r}}\right)\right) \\ &= O\left(A^2 \left[\chi_{\varepsilon,A}\right]^{2s_1(1-\mu)} \sum_{l=1}^{M_0} l^{-2(s_3 - s_1 2r)} \sum_{j_2=m_0'}^{j_{20}} 2^{-2j_2(s_2-s_1)}\right) \\ &= O\left(A^2 \left[\chi_{\varepsilon,A}\right]^{\frac{2s_1}{2s_1+1}} \left[\log(\varepsilon^{-1})\right]^{\mathbb{I}(s_1=s_2)+\mathbb{I}(s_1=s_3/2r)}\right) \end{aligned} \tag{A.37}$$

Case 2: $\gamma = \beta = 0$**,** $s_2 \leq \min\{s_1, \frac{s_3}{2r}\}$**.** In this case, $\mu = \frac{2s_2}{2s_2+1}$, (A.33) becomes, as $\varepsilon \to 0$

$$\begin{aligned}\Delta_3 &= O\left(A^2 \sum_{j_1=m_0}^{j_{10}-1} \sum_{l=1}^{M_0} \sum_{j_2=m_0'}^{j_{20}-1} 2^{-2j_1 s_1 - 2j_2 s_2} l^{-2s_3} \mathbb{I}\left(2^{j_2} > 2^{-j_1} \frac{\left[\chi_{\varepsilon,A}\right]^{\mu-1}}{l^{2r}}\right)\right) \\ &= O\left(A^2 \left[\chi_{\varepsilon,A}\right]^{2s_2(1-\mu)} \sum_{l=1}^{M_0} l^{-2(s_3 - s_2 2r)} \sum_{j_1=m_0}^{j_{10}} 2^{-2j_1(s_1-s_2)}\right) \\ &= O\left(A^2 \left[\chi_{\varepsilon,A}\right]^{\frac{2s_2}{2s_2+1}} \left[\log(\varepsilon^{-1})\right]^{\mathbb{I}(s_1=s_2)+\mathbb{I}(s_2=s_3/2r)}\right) \end{aligned} \tag{A.38}$$

Case 3: $\gamma = \beta = 0$**,** $s_3 \leq \min\{2rs_1, 2rs_2\}$**.** In this case, $\mu = \frac{2s_3}{2s_3+2r}$, (A.33) becomes, as

$\varepsilon \to 0$

$$\begin{aligned}
\Delta_3 &= O\left(A^2 \sum_{j_1=m_0}^{j_{10}-1} \sum_{l=1}^{M_0} \sum_{j_2=m_0'}^{j_{20}-1} 2^{-2j_1 s_1 - 2j_2 s_2} l^{-2s_3} \mathbb{I}\left(l^{2r} > 2^{-j_1-j_2} [\chi_{\varepsilon,A}]^{\mu-1}\right)\right) \\
&= O\left(A^2 \left[\chi_{\varepsilon,A}\right]^{-\frac{\mu-1}{2r} 2s_3} \sum_{j_1=m_0}^{j_{10}-1} 2^{-\frac{2j_1}{2r}(2rs_1-s_3)} \sum_{j_2=m_0'}^{j_{20}} 2^{-\frac{2j_2}{2r}(2rs_2-s_3)}\right) \\
&= O\left(A^2 \left[\chi_{\varepsilon,A}\right]^{\frac{s_3}{s_3+r}} \left[\log(\varepsilon^{-1})\right]^{\mathbb{I}(s_2=s_3/2r)+\mathbb{I}(s_1=s_3/2r)}\right) \qquad \text{(A.39)}
\end{aligned}$$

Case 4: $\gamma > 0, \beta > 0$**,** $s_1 \leq s_2$**.** In this case, $\mu = \frac{2s_1}{2s_1+1}$, (A.33) becomes, as $\varepsilon \to 0$

$$\begin{aligned}
\Delta_3 &= O\left(A^2 \sum_{j_1=m_0}^{j_{10}-1} \sum_{l=1}^{M_0} \sum_{j_2=m_0'}^{j_{20}-1} 2^{-2j_1 s_1 - 2j_2 s_2} l^{-2s_3} \exp\{-2\gamma l^{\beta}\} \mathbb{I}\left(2^{j_1} > 2^{-j_2} \frac{[\chi_{\varepsilon,A}]^{\mu-1}}{l^{2r}}\right)\right) \\
&= O\left(A^2 \left[\chi_{\varepsilon,A}\right]^{\frac{2s_1}{2s_1+1}} \sum_{j_2=m_0'}^{j_{20}} 2^{-2j_2(s_2-s_1)}\right) \\
&= O\left(A^2 \left[\chi_{\varepsilon,A}\right]^{\frac{2s_1}{2s_1+1}} \left[\log(\varepsilon^{-1})\right]^{\mathbb{I}(s_1=s_2)}\right) \qquad \text{(A.40)}
\end{aligned}$$

Case 5: $\gamma > 0, \beta > 0$**,** $s_2 \leq s_1$**.** In this case, $\mu = \frac{2s_2}{2s_2+1}$, (A.33) becomes, as $\varepsilon \to 0$

$$\begin{aligned}
\Delta_3 &= O\left(A^2 \sum_{j_1=m_0}^{j_{10}-1} \sum_{l=1}^{M_0} \sum_{j_2=m_0'}^{j_{20}-1} 2^{-2j_1 s_1 - 2j_2 s_2} l^{-2s_3} \exp\{-2\gamma l^{\beta}\} \mathbb{I}\left(2^{j_2} > 2^{-j_1} \frac{[\chi_{\varepsilon,A}]^{\mu-1}}{l^{2r}}\right)\right) \\
&= O\left(A^2 \left[\chi_{\varepsilon,A}\right]^{\frac{2s_2}{2s_2+1}} \sum_{j_1=m_0}^{j_{10}} 2^{-2j_1(s_1-s_2)}\right) \\
&= O\left(A^2 \left[\chi_{\varepsilon,A}\right]^{\frac{2s_2}{2s_2+1}} \left[\log(\varepsilon^{-1})\right]^{\mathbb{I}(s_1=s_2)}\right) \qquad \text{(A.41)}
\end{aligned}$$

Now, to complete the proof, combine formulae (A.22)-(A.41).

APPENDIX B: DERIVATIONS AND PROOFS PART 2

Proof of Theorem 3. Note that it follows from the optimization problem (5.8) that for any fixed $\mathbf{G}, \mathbf{Z}, J$ and K one has

$$\|\widehat{\mathbf{G}}\|_F^2 - 2\mathrm{Tr}(\mathbf{Y}^T\boldsymbol{\Upsilon}\widehat{\mathbf{G}}\boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{K}}) + \mathrm{Pen}(\hat{J},\hat{K}) \leq \|\mathbf{G}\|_F^2 - 2\mathrm{Tr}(\mathbf{Y}^T\boldsymbol{\Upsilon}\mathbf{G}\boldsymbol{\Pi}_{\mathbf{Z},K}) + \mathrm{Pen}(J,K).$$

Therefore, equation (4.9) yields

$$\|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2 \leq \|\mathbf{G} - \mathbf{G}_*\|_F^2 + \frac{2\,\sigma}{\sqrt{n}}\,\mathrm{Tr}[\mathbf{E}^T\boldsymbol{\Upsilon}(\widehat{\mathbf{G}} - \mathbf{G})] + \mathrm{Pen}(J,K) - \mathrm{Pen}(\hat{J},\hat{K}) \tag{B.1}$$

We choose $\mathbf{G} = \mathbf{W}_J\mathbf{G}_*\boldsymbol{\Pi}_{\mathbf{Z},K}$ and in order to analyze the cross term $\mathrm{Tr}[\mathbf{E}^T\boldsymbol{\Upsilon}(\widehat{\mathbf{G}} - \mathbf{G})]$ we use vectorization of the model. For this purpose, we denote

$$\boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{K},\hat{J}} = (\boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{\mathbf{K}}} \otimes \mathbf{W}_{\hat{J}}), \quad \boldsymbol{\Pi}_{\mathbf{Z},K,J} = (\boldsymbol{\Pi}_{\mathbf{Z},K} \otimes \mathbf{W}_J) \tag{B.2}$$

$$\hat{\mathbf{g}} = \mathrm{vec}(\widehat{\mathbf{G}}), \quad \mathbf{g} = \mathrm{vec}(\mathbf{G}), \quad \boldsymbol{\Gamma} = (\mathbf{I}_M \otimes \boldsymbol{\Upsilon}), \quad \boldsymbol{\delta} = (\mathbf{I}_M \otimes \mathbf{S}^{-1})\boldsymbol{\epsilon}, \tag{B.3}$$

here $\mathbb{E}(\boldsymbol{\delta}\boldsymbol{\delta}^T) = \mathbf{I}_{n,M}$,thus $\boldsymbol{\delta} \sim N(0, \mathbf{I}_{n,M})$, where $\boldsymbol{\epsilon}$ is defined in (4.11) and $\boldsymbol{\Sigma} = \mathbf{S}\mathbf{S}^T$. Then, equation (4.9) can be re-written as

$$\boldsymbol{\Gamma}\mathbf{y} = \mathbf{g}_* + \frac{\sigma}{\sqrt{n}}\boldsymbol{\Gamma}(\mathbf{I}_M \otimes \mathbf{S})\boldsymbol{\delta}. \tag{B.4}$$

Observe that by Theorem 1.2.22 of Gupta and Nagar (2000), one has

$$\hat{\mathbf{g}} = \mathrm{vec}(\mathbf{W}_{\hat{J}}\boldsymbol{\Upsilon}\mathbf{Y}\boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{K}}) = \boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{K},\hat{J}}\boldsymbol{\Gamma}\mathbf{y}, \quad \mathbf{g} = \boldsymbol{\Pi}_{\mathbf{Z},K,J}\mathbf{g}_*$$

and $\mathrm{Tr}[\mathbf{E}^T\boldsymbol{\Upsilon}(\widehat{\mathbf{G}} - \mathbf{G})] = \boldsymbol{\delta}^T(\mathbf{I}_M \otimes \mathbf{S}^T\boldsymbol{\Upsilon})(\boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{K},\hat{J}}\boldsymbol{\Gamma}\mathbf{y} - \boldsymbol{\Pi}_{\mathbf{Z},K,J}\mathbf{g}_*)$. Now (B.1) can be rewrite

in a vector form as

$$\|\hat{\mathbf{g}} - \mathbf{g}_*\|^2 \le \|\mathbf{g} - \mathbf{g}_*\|^2 + \Delta + \mathrm{Pen}(J,K) - \mathrm{Pen}(\hat{J},\hat{K}) \tag{B.5}$$

where

$$\Delta = \frac{2\,\sigma}{\sqrt{n}}\boldsymbol{\delta}^T(\mathbf{I}_M \otimes \mathbf{S}^T\boldsymbol{\Upsilon})(\boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{K},\hat{J}}\boldsymbol{\Gamma}\mathbf{y} - \boldsymbol{\Pi}_{\mathbf{Z},K,J}\mathbf{g}_*) = \Delta_1 + \Delta_2 \tag{B.6}$$

with

$$\Delta_1 = \frac{2\,\sigma}{\sqrt{n}}\boldsymbol{\delta}^T(\mathbf{I}_M \otimes \mathbf{S}^T\boldsymbol{\Upsilon})(\boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{K},\hat{J}}(\boldsymbol{\Gamma}\mathbf{y} - \mathbf{g}_*)), \quad \Delta_2 = \frac{2\,\sigma}{\sqrt{n}}\boldsymbol{\delta}^T(\mathbf{I}_M \otimes \mathbf{S}^T\boldsymbol{\Upsilon})(\boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{K},\hat{J}} - \boldsymbol{\Pi}_{\mathbf{Z},K,J})\mathbf{g}_*. \tag{B.7}$$

Derivation of upper bounds for Δ_1 and Δ_2 is based on the following lemma.

Lemma 5 *Let K, J be fixed, $\hat{J}$ be an arbitrary random subset of $\{1,\cdot,n\}$ and $\hat{K}$ be a random integer between 1 and M. Let $\mathbf{Z}$ and $\widehat{\mathbf{Z}}$ be a fixed and a random matrices of ranks K and $\hat{K}$, respectively. Denote the projection matrices on the column spaces of matrices $\mathbf{Z}$ and $\widehat{\mathbf{Z}}$, respectively, by $\boldsymbol{\Pi}_{\mathbf{Z},K}$ and $\boldsymbol{\Pi}_{\widehat{\mathbf{Z}},\hat{K}}$. Let $\mathbf{S}$ be a matrix with $\|\mathbf{S}\|_{op} \le C_\psi$ and $\boldsymbol{\delta} \sim N(0,\mathbf{I}_{nM})$. Then, for any $s > 0$, there exist sets $\Omega_{1\tau}$ and $\Omega_{2\tau}$ with $\mathbb{P}(\Omega_{1\tau}) \ge 1 - n^{-\tau}$ and $\mathbb{P}(\Omega_{2\tau}) \ge 1 - n^{-\tau}$ such that*

$$\|(\boldsymbol{\Pi}_{\mathbf{Z},K} \otimes (\mathbf{W}_J\boldsymbol{\Upsilon}\mathbf{S}))\boldsymbol{\delta}\|^2 \le 2KC_\psi^2(\sum_{j\in J}\nu_j^2) + 3C_\psi^2(\max_{j\in J}\nu_j^2)(\tau\ln n), \quad \forall\omega \in \Omega_{1\tau}; \tag{B.8}$$

$$\begin{aligned}\|(\boldsymbol{\Pi}_{\widehat{\mathbf{Z}},\hat{K}} \otimes (\mathbf{W}_{\hat{J}}\boldsymbol{\Upsilon}\mathbf{S}))\boldsymbol{\delta}\|^2 &\le 2\hat{K}C_\psi^2(\sum_{j\in\hat{J}}\nu_j^2) \\ &+ 3C_\psi^2(\max_{j\in\hat{J}}\nu_j^2)\left\{M\ln\hat{K} + |\hat{J}|\ln\left(\frac{ne}{|\hat{J}|}\right) + \ln M + (\tau+1)\ln n\right\} \quad \forall\omega \in \Omega_{2\tau}.\end{aligned} \tag{B.9}$$

Note that Δ_1 can be re-written as $\Delta_1 = 2\,n^{-1}\sigma^2\,\boldsymbol{\delta}^T(\mathbf{I}_M \otimes \mathbf{S}^T\boldsymbol{\Upsilon})(\boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes \mathbf{W}_{\hat{J}})(\mathbf{I}_M \otimes \boldsymbol{\Upsilon}\mathbf{S})\boldsymbol{\delta}$. Due to $n^{-1/2}\sigma\,\boldsymbol{\Gamma}\boldsymbol{\epsilon} = \boldsymbol{\Gamma}\mathbf{y} - \mathbf{g}_*$ and (B.3), obtain $\Delta_1 = 2\,n^{-1}\sigma^2\,\|(\boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes (\mathbf{W}_{\hat{J}}\Upsilon\mathbf{S}))\boldsymbol{\delta}\|^2$. Therefore, by (B.9), obtain that for $\omega \in \Omega_{2\tau}$

$$|\Delta_1| \leq \frac{2\,\sigma^2\,C_\psi^2}{n}\left[2\hat{K}(\sum_{j\in\hat{J}}\nu_j^2) + 3(\max_{j\in\hat{J}}\nu_j^2)\left\{M\ln\hat{K} + |\hat{J}|\ln\left(\frac{ne}{|\hat{J}|}\right) + \ln M + (\tau+1)\ln n\right\}\right] \tag{B.10}$$

In order to construct an upper bound for Δ_2, consider the following sets

$$\tilde{J} = J \cup \hat{J}, \quad J_1 = J \cap \hat{J}, \quad J_2 = J^c \cap \hat{J}, \quad J_3 = \hat{J}^c \cap J. \tag{B.11}$$

The sets J_1, J_2 and J_3 are non-overlapping and $\tilde{J} = J_1 \cup J_2 \cup J_3$. Furthermore, consider matrix $\tilde{\mathbf{Z}}$ that includes all linearly independent columns in matrices $\mathbf{Z}_K$ and $\hat{\mathbf{Z}}_{\hat{K}}$, so that $\mathrm{Span}\{\tilde{\mathbf{Z}}\} = \mathrm{Span}\{\mathbf{Z}_K, \hat{\mathbf{Z}}_{\hat{K}}\}$. Let $\tilde{K}$ be the number of columns of matrix $\tilde{\mathbf{Z}}$. Then, one has

$$\begin{aligned}
&\boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{K}}\boldsymbol{\Pi}_{\tilde{\mathbf{Z}},\tilde{K}} = \boldsymbol{\Pi}_{\tilde{\mathbf{Z}},\tilde{K}}\boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{K}} = \boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{K}},\\
&\mathbf{W}_J = \mathbf{W}_{J_1} + \mathbf{W}_{J_3},\ \mathbf{W}_{\hat{J}} = \mathbf{W}_{J_1} + \mathbf{W}_{J_2},\ \mathbf{W}_{\tilde{J}} = \mathbf{W}_{J_1} + \mathbf{W}_{J_2} + \mathbf{W}_{J_3}.
\end{aligned}$$

In order to obtain an upper bound for Δ_2 defined in (B.7), note that using notations above, we can rewrite Δ_2 as

$$\begin{aligned}
\Delta_2 &= \frac{2\,\sigma}{\sqrt{n}}\boldsymbol{\delta}^T(\mathbf{I}_M \otimes \mathbf{S}^T\boldsymbol{\Upsilon})[(\boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes \mathbf{W}_{J_2}) + (\boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes \mathbf{W}_{J_1}) - (\boldsymbol{\Pi}_{\mathbf{Z},K} \otimes \mathbf{W}_{J_1}) - (\boldsymbol{\Pi}_{\mathbf{Z},K} \otimes \mathbf{W}_{J_3})]\mathbf{g}_*\\
&= \frac{2\,\sigma}{\sqrt{n}}\boldsymbol{\delta}^T(\mathbf{I}_M \otimes \mathbf{S}^T\boldsymbol{\Upsilon})[(\boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes \mathbf{W}_{J_2}) + (\boldsymbol{\Pi}_{\tilde{\mathbf{Z}},\tilde{K}} \otimes \mathbf{W}_{J_1}) + (\boldsymbol{\Pi}_{\mathbf{Z},K} \otimes \mathbf{W}_{J_3})][(\boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes \mathbf{W}_{J_2})\\
&+ (\boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes \mathbf{W}_{J_1}) - (\boldsymbol{\Pi}_{\mathbf{Z},K} \otimes \mathbf{W}_{J_1}) - (\boldsymbol{\Pi}_{\mathbf{Z},K} \otimes \mathbf{W}_{J_3})]\mathbf{g}_*\\
&= \frac{2\,\sigma}{\sqrt{n}}\boldsymbol{\delta}^T(\mathbf{I}_M \otimes \mathbf{S}^T\boldsymbol{\Upsilon})[(\boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes \mathbf{W}_{J_2}) + (\boldsymbol{\Pi}_{\tilde{\mathbf{Z}},\tilde{K}} \otimes \mathbf{W}_{J_1}) + (\boldsymbol{\Pi}_{\mathbf{Z},K} \otimes \mathbf{W}_{J_3})][(\boldsymbol{\Pi}_{\hat{\mathbf{Z}},\hat{K},\hat{J}} - \boldsymbol{\Pi}_{\mathbf{Z},K,J})]\mathbf{g}_*
\end{aligned}$$

Using Cauchy inequality and $2ab \leq 4a^2 + \frac{b^2}{4}$, obtain

$$|\Delta_2| \leq |\Delta_{2,1}| + |\Delta_{2,2}|, \quad |\Delta_{2,2}| = 0.25 \left\|(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K},\hat{J}}\mathbf{g}_* - \mathbf{\Pi}_{\mathbf{Z},K,J}\mathbf{g}_*)\right\|^2 \tag{B.12}$$
$$|\Delta_{2,1}| = 4\,\frac{\sigma^2}{n}\left\|[(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes \mathbf{W}_{J_2}) + (\mathbf{\Pi}_{\tilde{\mathbf{Z}},\tilde{K}} \otimes \mathbf{W}_{J_1}) + (\mathbf{\Pi}_{\mathbf{Z},K} \otimes \mathbf{W}_{J_3})](\mathbf{I}_M \otimes \mathbf{\Upsilon}\mathbf{S})\boldsymbol{\delta}\right\|^2$$

Applying Cauchy Inequality to the term $\Delta_{2,1}$ and using that $J_2 \subseteq \hat{J}$ and $J_3 \subseteq J$ we rewrite

$$\Delta_{2,1} \leq \frac{12\,\sigma^2}{n}\,\left[\|(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes (\mathbf{W}_{J_2}\mathbf{\Upsilon}\mathbf{S}))\boldsymbol{\delta}\|^2 + \|(\mathbf{\Pi}_{\tilde{\mathbf{Z}},\tilde{K}} \otimes (\mathbf{W}_{J_1}\mathbf{\Upsilon}\mathbf{S}))\boldsymbol{\delta}\|^2 + \|(\mathbf{\Pi}_{\mathbf{Z},K} \otimes (\mathbf{W}_{J_3}\mathbf{\Upsilon}\mathbf{S}))\boldsymbol{\delta}]\|^2\right]$$

$$|\Delta_{2,1}| \leq \frac{12\,\sigma^2}{n}\,\left[\|(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes (\mathbf{W}_{\hat{J}}\mathbf{\Upsilon}\mathbf{S}))\boldsymbol{\delta}\|^2 + \|(\mathbf{\Pi}_{\tilde{\mathbf{Z}},\tilde{K}} \otimes (\mathbf{W}_{J_1}\mathbf{\Upsilon}\mathbf{S}))\boldsymbol{\delta}\|^2 + \|(\mathbf{\Pi}_{\mathbf{Z},K} \otimes (\mathbf{W}_{J}\mathbf{\Upsilon}\mathbf{S}))\boldsymbol{\delta}]\|^2\right]$$

The upper bounds for the first and the third term in the inequality above can be obtained directly from Lemma 5. For the second term, note that since $\tilde{K} \leq K + \hat{K}$ and $J_1 \subseteq J$ and $J_1 \subseteq \hat{J}$ for any $\omega \in \Omega_{1\tau} \cap \Omega_{2\tau}$ one has

$$\|(\mathbf{\Pi}_{\tilde{\mathbf{Z}},\tilde{K}} \otimes (\mathbf{W}_{J_1}\mathbf{\Upsilon}\mathbf{S}))\boldsymbol{\delta}\|^2 \leq C_\psi^2\,\left[2K\sum_{j\in J}\nu_j^2 + 2\hat{K}\sum_{j\in\hat{J}}\nu_j^2\right. \tag{B.13}$$
$$\left. +3\,(\max_{j\in\hat{J}}\nu_j^2)\left\{M\ln\hat{K} + |\hat{J}|\ln\left(\frac{ne}{|\hat{J}|}\right) + \ln M + (\tau+1)\ln n\right\}\right]$$

Combining (B.13) with equations (B.8) and (B.9), obtain

$$|\Delta_{2,1}| \leq 12\frac{\sigma^2}{n}\,C_\psi^2\,\left[4\hat{K}\sum_{j\in\hat{J}}\nu_j^2 + 4K\sum_{j\in J}\nu_j^2 + +3(\max_{j\in J}\nu_j^2)(\tau\ln n)\right. \tag{B.14}$$
$$\left. +6(\max_{j\in\hat{J}}\nu_j^2)\left\{M\ln\hat{K} + |\hat{J}|\ln\left(\frac{ne}{|\hat{J}|}\right) + \ln M + (\tau+1)\ln n\right\}\right]$$

for any $\omega \in \Omega_{1\tau} \cap \Omega_{2\tau}$.

Now consider $|\Delta_{2,2}|$ defined in (B.12). Rewrite $|\Delta_{2,2}|$ as $|\Delta_{2,2}| = 0.25\|(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K},\hat{J}}\mathbf{g}_* - \mathbf{g}_*) - (\mathbf{\Pi}_{\mathbf{Z},K,J}\mathbf{g}_* - \mathbf{g}_*)\|^2$, so that

$$|\Delta_{2,2}| \leq 0.5\,\|(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K},\hat{J}}\mathbf{g}_* - \mathbf{g}_*)\|^2 + 0.5\|(\mathbf{\Pi}_{\mathbf{Z},K,J}\mathbf{g}_* - \mathbf{g}_*)\|^2.$$

Since $\hat{\mathbf{g}} = \mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K},\hat{J}}\mathbf{\Gamma}\mathbf{y}$ and

$$\begin{aligned}\|(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K},\hat{J}}\mathbf{\Gamma}\mathbf{y} - \mathbf{g}_*)\|^2 &= \|(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K},\hat{J}}(\mathbf{g}_* + n^{-1/2}\sigma\,\mathbf{\Gamma}\boldsymbol{\epsilon}) - \mathbf{g}_*)\|^2 \\ &= \|(\mathbf{I} - \mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K},\hat{J}})\mathbf{g}_*\|^2 + n^{-1}\sigma^2\,\|\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K},\hat{J}}\,\mathbf{\Gamma}\boldsymbol{\epsilon}\|^2,\end{aligned}$$

we derive

$$\|\hat{\mathbf{g}} - \mathbf{g}_*\|^2 \geq \|\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K},\hat{J}}\mathbf{g}_* - \mathbf{g}_*\|^2 \tag{B.15}$$

Taking into account that $\mathbf{g} = \mathbf{\Pi}_{\mathbf{Z},K,J}\mathbf{g}_*$, so that $\|\mathbf{g} - \mathbf{g}_*\|^2 = \|\mathbf{\Pi}_{\mathbf{Z},K,J}\mathbf{g}_* - \mathbf{g}_*\|^2$, we obtain

$$|\Delta_{2,2}| \leq 0.5\|\hat{\mathbf{g}} - \mathbf{g}_*\|^2 + 0.5\|\mathbf{g} - \mathbf{g}_*\|^2. \tag{B.16}$$

By combining upper bounds of Δ_1, $\Delta_{2,1}$ and $\Delta_{2,2}$, we derive from (B.10) and (B.14)– (B.16) that for any $\omega \in \Omega_{1\tau} \cap \Omega_{2\tau}$ upper bound for Δ can be written as

$$\begin{aligned}|\Delta| \leq 0.5\|\hat{\mathbf{g}} - \mathbf{g}_*\|^2 + 0.5\|\mathbf{g} - \mathbf{g}_*\|^2 + \frac{2\,\sigma^2\,C_\psi^2}{n}\Bigg\{26\hat{K}\sum_{j\in\hat{J}}\nu_j^2 + 24K\sum_{j\in J}\nu_j^2 \\ +39(\max_{j\in\hat{J}}\nu_j^2)\left[M\ln\hat{K} + |\hat{J}|\ln\left(\frac{ne}{|\hat{J}|}\right) + \ln M + (\tau+1)\ln n\right] + 18(\max_{j\in J}\nu_j^2)\tau\ln n\Bigg\}\end{aligned} \tag{B.17}$$

Since it follows from (B.3) that $\|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2 = \|\hat{\mathbf{g}} - \mathbf{g}_*\|^2$, obtain from (B.5) that for any

$\mathbf{G} = \Pi_{\mathbf{Z},K,J}\mathbf{G}_*$ on the set $\Omega_{1\tau} \cap \Omega_{2\tau}$ one has

$$\|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2 \leq 3\|\mathbf{G} - \mathbf{G}_*\|_F^2 + \frac{2\,\sigma^2\,C_\psi^2}{n}\left\{48K\sum_{j\in J}\nu_j^2 + 36(\max_{j\in J}\nu_j^2)\tau\ln n + 52\hat{K}\sum_{j\in\hat{J}}\nu_j^2\right.$$

$$\left. +78(\max_{j\in\hat{J}}\nu_j^2)\left[M\ln\hat{K} + |\hat{J}|\ln\left(\frac{ne}{|\hat{J}|}\right) + \ln(Mn) + \tau\ln n\right]\right\} + 2[\mathrm{Pen}(J,K) - \mathrm{Pen}(\hat{J},\hat{K})] \tag{B.18}$$

Choose $\mathrm{Pen}(J,K)$ in the form (4.20) and note that all terms containing $\hat{J}$ and $\hat{K}$ in (B.18) cancel. Finally we obtained for any $\mathbf{G} = \mathbf{W}_J\mathbf{G}_*\mathbf{\Pi}_{\mathbf{Z},K}$ that with probability at least $1 - 2n^{-\tau}$

$$\|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2 \leq 3\|\mathbf{G} - \mathbf{G}_*\|_F^2 + \frac{2\,\sigma^2\,C_\psi^2}{n}\left\{48K\sum_{j\in J}\nu_j^2 + 36(\max_{j\in J}\nu_j^2)\tau\ln n\right\} + 2\,\mathrm{Pen}(J,K)$$

which yeilds (4.2).

Proof of Theorem 4. Note that, in this case, the optimal set J is of the form $J = \{1,\cdots,L\}$, so that $|J| = L$, and find $(\hat{\mathbf{Z}},\widehat{\mathbf{G}},\hat{L},\hat{K})$ as a solution of optimization problem (5.8) with the penalty given by expression (4.20).

Note that for the true number of classes K_* with $N_k, k = 1,\ldots,K_*$ elements in each class, $\mathbf{G}$ are coefficients of each f_m and $\mathbf{\Theta}$ is the clustered version of those coefficients. it follows from (4.8) that

$$R(\hat{\mathbf{f}},\mathcal{S}(r,\mathcal{A}),M,K_*) \leq M^{-1}\|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2 + M^{-1}\sum_{k=1}^{K_*} N_k \sum_{j=n+1}^{\infty}\mathbf{\Theta}_{jk}. \tag{B.19}$$

Therefore, application of Theorem 3 with $\widehat{\mathbf{G}} = \mathbf{W}_J\mathbf{G}_*\mathbf{\Pi}_{\mathbf{Z}_*,K_*}$ where $\mathbf{Z}_*$ and K_* are respec-

tively the true clustering matrix and the true number of classes, yields

$$\|\mathbf{W}_J\mathbf{G}_*\mathbf{\Pi}_{\mathbf{Z}_*,K_*}-\mathbf{G}_*\|_F^2=\|(\mathbf{W}_J-\mathbf{I}_n)\mathbf{G}_*\|_F^2=\sum_{k=1}^{K^*}N_k\sum_{j=L+1}^{n}\mathbf{\Theta}_{jk}^2 \tag{B.20}$$

where N_k is the number of functions $f_m=h_k$ in the cluster k, $k=1,\cdots,K^*$, and $\mathbf{\Theta}_{jk}$ are the true coefficients of those functions. Note that it follows from (4.13) that

$$\sum_{j=L+1}^{n}\mathbf{\Theta}_{jk}^2\leq\mathcal{A}^2L^{-2r}. \tag{B.21}$$

Therefore, $\sum_{k=1}^{K^*}N_k=M$, (B.20) and (B.21) yield

$$\|\mathbf{W}_J\mathbf{G}_*\mathbf{\Pi}_{\mathbf{Z}_*,K_*}-\mathbf{G}_*\|_F^2\leq\mathcal{A}^2ML^{-2r} \tag{B.22}$$

Moreover, it follows from (B.22) that

$$\sum_{k=1}^{K_*}N_k\sum_{j=n+1}^{\infty}\mathbf{\Theta}_{jk}\leq\mathcal{A}^2Mn^{-2r},$$

so that the last term in (B.19) is smaller than the first term.

Now, consider the second term in (4.2). Due to the condition (4.3) and $J=\{1,\cdots,L\}$, one obtains

$$\max_{j\in J}\nu_j^2\leq\aleph_2^2\,L^{2\gamma}\exp\left(2\alpha L^{\beta}\right),\quad\sum_{j\in J}\nu_j^2\leq\aleph_2^2\,L^{2\gamma+1}\exp\left(2\alpha L^{\beta}\right).$$

Note also that, due to condition (4.3), in order $M^{-1}\,\|\widehat{\mathbf{G}}-\mathbf{G}_*\|_F^2$ tends to zero as $n\to\infty$, one needs $L\leq Cn^{1/(2\gamma)}$ if $\alpha=0$ and $L\leq[\ln n/(2\alpha)]^{1/\beta}$, so that $\ln(ne/L)\asymp\ln n$. Denote

$$R_1\equiv R_1(K_*,n)=K_*+\ln n,\quad R_2\equiv R_2(M,K_*,n)\asymp M\ln K_*+\ln(Mn). \tag{B.23}$$

Therefore, it follows from (4.2) that

$$M^{-1}\|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2 \leq \tilde{C} \min_L \left\{ L^{-2r} + \frac{\sigma^2 L^{2\gamma}}{M n} \exp\left(2\alpha L^\beta\right) [L R_1(K_*, n) + R_2(M, K_*, n)] \right\} \tag{B.24}$$

where $R_1(K_*, n)$ and $R_2(M, K_*, n)$ are defined in (B.23) and $\tilde{C}$ depends only on μ, $\mathcal{A}$, $\aleph_2$, C_ψ^2 and is independent of M,L, n and K_*.

In order to find the minimum of the right hand side of (B.24), denote

$$R(L, M, K_*, n) = L^{-2r} + \sigma^2 (Mn)^{-1} \exp\left(2\alpha L^\beta\right) \left[L^{2\gamma+1} R_1 + L^{2\gamma} R_2\right] \tag{B.25}$$

and observe that

$$M^{-1} \|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2 \leq \tilde{C} \min_L R(L, M, K_*, n) = \tilde{C}\, R(L_{opt}, M, K_*, n) \tag{B.26}$$

where L_{opt} is the value of L minimizing the right hand side of (B.24). Denote

$$L_{1,opt} = \operatorname*{argmin}_L [L^{-2r} + \sigma^2 (Mn)^{-1} \exp\left(2\alpha L^\beta\right) L^{2\gamma+1} R_1], \tag{B.27}$$

$$L_{2,opt} = \operatorname*{argmin}_L [L^{-2r} + \sigma^2 (Mn)^{-1} \exp\left(2\alpha L^\beta\right) L^{2\gamma} R_2], \tag{B.28}$$

and set $L_{opt} = \min(L_{1,opt}; L_{2,opt})$. It is easy to see that since the first terms in expressions (B.27) and (B.28) are decreasing in L while the second terms are increasing, the values $L_{1,opt}$ and $L_{2,opt}$ are such that those terms are equal to each other up to a multiplicative constant and, therefore, due to (B.25), one has

$$R(L_{opt}, M, K_*, n) \asymp (L_{opt})^{-2r} \tag{B.29}$$

Consider two cases.

Case 1: $\alpha = \beta = 0$.

Direct calculations yield

$$L_{1,opt} \asymp \left(\frac{Mn}{\sigma^2 R_1}\right)^{\frac{1}{2\gamma+2r+1}}, \quad L_{2,opt} \asymp \left(\frac{Mn}{\sigma^2 R_2}\right)^{\frac{1}{2\gamma+2r}},$$

so that

$$L_{1,opt} = \min\left\{\left(\frac{Mn}{\sigma^2 K_*}\right)^{\frac{1}{2\gamma+2r+1}}; \left(\frac{Mn}{\sigma^2 \ln n}\right)^{\frac{1}{2\gamma+2r+1}}\right\},$$
$$L_{2,opt} = \min\left\{\left(\frac{n}{\sigma^2 \ln K_*}\right)^{\frac{1}{2\gamma+2r}}; \left(\frac{Mn}{\sigma^2 \ln (Mn)}\right)^{\frac{1}{2\gamma+2r}}\right\}$$

Then

$$L_{opt} = \min\left\{\left(\frac{n}{\sigma^2 \ln K_*}\right)^{\frac{1}{2\gamma+2r}}; \left(\frac{Mn}{\sigma^2 (K_* + \ln n)}\right)^{\frac{1}{2\gamma+2r+1}}\right\}.$$

which, together with (B.26) and (B.29), yield the expression (4.6).

Case 2: $\alpha > 0$, $\beta > 0$.

Minimizing expressions in (B.27) and (B.28) obtain

$$L_{i,opt} \asymp \left\{\left[\ln\left(\frac{Mn}{\sigma^2 R_i}\right)\right]\right\}^{\frac{1}{\beta}}, \quad i = 1, 2,$$

Taking into account that $R_2 > R_1$ and that, for large M and n, $\ln(Mn \ln n^{-1}) \asymp \ln(Mn)$, obtain

$$L_{1,opt} = \min\left\{\left[\ln\left(\frac{Mn}{\sigma^2 K_*}\right)\right]; \left[\ln\left(\frac{Mn}{\sigma^2 \ln n}\right)\right]\right\}^{\frac{1}{\beta}} = \left[\ln\left(\frac{Mn}{\sigma^2 K_*}\right)\right]^{\frac{1}{\beta}}.$$

Similarly,

$$L_{2,opt} = \min\left\{\left[\ln\left(\frac{n}{\sigma^2 \ln K_*}\right)\right]; \left[\ln\left(\frac{Mn}{\sigma^2 \ln(Mn)}\right)\right]\right\}^{\frac{1}{\beta}} = \left[\ln\left(\frac{n}{\sigma^2 \ln K_*}\right)\right]^{\frac{1}{\beta}}$$

Hence

$$L_{opt} = \min\left\{\left[\ln\left(\frac{n}{\sigma^2 \ln K_*}\right)\right]; \left[\ln\left(\frac{Mn}{\sigma^2 K_*}\right)\right]\right\}^{\frac{1}{\beta}}.$$

which, together with (B.26) and (B.29), yield the expression (4.7).

Proof of Theorem 5. Since the estimation error is comprised of the error due to non-parametric estimation and to clustering, we consider two cases here.

Lower bound for the error due to clustering.

Let K be the fixed number of classes. Consider a subset $\mathcal{Z}(M,K) \subset \mathcal{M}(M,K)$ of the set of all clustering matrices which contain all matrices that cluster $\frac{M}{K}$ vectors into each class. The cardinality of the set $\mathcal{Z}(M,K)$

$$|\mathcal{Z}(M,K)| = \frac{M!}{[(\frac{M}{K})!]^K} \geq \exp\left(\frac{M}{4}\ln K\right) \tag{B.30}$$

by Lemma 5 in Pensky (2018) with $\gamma = 1$. Let set J be of the form $J = \{L_1, ..., L_2\}$ where $1 \leq L_1 < L_2 \leq n$. Choose $\boldsymbol{\Theta}_{jk} = 0$ if $j \notin J$. In what follows, we use the packing lemma (Lemma 4 of Pensky (2018)):

Lemma 6 *(The Packing lemma). Let $\mathcal{Z}(M,K) \subseteq \mathcal{M}(M,K)$ be a collection of clustering matrices. Then, there exists a subset $\mathcal{S}_{M,K}(r) \subset \mathcal{Z}(M,K)$ such that for $\mathbf{Z}_1, \mathbf{Z}_2 \in \mathcal{S}_{M,K}(r)$ one has $\|\mathbf{Z}_1 - \mathbf{Z}_2\|_H = \|\mathbf{Z}_1 - \mathbf{Z}_2\|_F^2 \geq r$ and $\ln|\mathcal{S}_{M,K}(r)| \geq \ln|\mathcal{Z}(M,K)| - r\ln(MKe/r)$.*

Apply this lemma with $r = dM$, $0 < d < 1/4$. Then, by (B.30), derive

$$\ln |\mathcal{S}_{M,K}(dM)| \geq M\ [\ln K - 4d \ln(Ke/d)]\ /4.$$

Use the following statement:

Lemma 7 *If $K \geq 2$ and d is such that*

$$d - d \ln d \leq (\ln 2)/9, \tag{B.31}$$

then $\ln K - 4d \ln(Ke/d) \geq (\ln K)/9$.

It is easy to calculate that, e.g., $d = 0.0147$ satisfies the condition (B.31). Then, for d obeying (B.31), one has

$$\ln |\mathcal{S}_{M,K}(dM)| \geq \frac{M}{36} \ln K, \quad \|\mathbf{Z}_1 - \mathbf{Z}_2\|_H \geq dM \text{ for any } \mathbf{Z}_1, \mathbf{Z}_2 \in \mathcal{S}_{M,K}(dM),\ \mathbf{Z}_1 \neq \mathbf{Z}_2 \tag{B.32}$$

Consider a collection of binary vectors $\boldsymbol{\omega} \in \{0,1\}^{|J|}$. By Varshamov-Gilbert bound lemma, there exists a subset $\mathcal{W}$ of those vectors such that, for any $\boldsymbol{\omega}, \boldsymbol{\omega}' \in \mathcal{W}$ such that $\boldsymbol{\omega} \neq \boldsymbol{\omega}'$ one has $\|\boldsymbol{\omega} - \boldsymbol{\omega}'\|_H \geq |J|/8$ and $\ln |\mathcal{W}| \geq |J| \ln(2)/8$. Choose a subset $\mathcal{W}_K$ of $\mathcal{W}$ such that $|\mathcal{W}_K| = K$. This is possible if $K \leq 2^{|J|/8}$ which equivalent to $|J| \geq 8 \ln K / \ln 2$. Consider a set of vectors $\mathbf{w} \in \{0,1\}^n$ obtained by packing $\boldsymbol{\omega}$ with zeros for components not in J. Then

$$\mathcal{W}_K = \{\mathbf{w}_1, ..., \mathbf{w}_K \in \{0,1\}^n :\ \|\mathbf{w}_i\|_0 \leq |J|,\ \|\mathbf{w}_i - \mathbf{w}_j\|_0 \geq |J|/8,\ i \neq j\} \tag{B.33}$$

Define matrix $\mathbf{W}$ with columns $\mathbf{w}_k$, $k = 1, ..., K$. Finally, form the set $\mathcal{G}_{M,K}$ of matrices $\mathbf{G}$

of the form

$$\mathcal{G}_{M,K} = \left\{ \mathbf{G} \in R^{M \times K} : \mathbf{G} = \theta \mathbf{W} \mathbf{Z}^T, \mathbf{Z} \in \mathcal{S}_{M,K}(dM) \right\}$$

where d satisfies (B.31) and $\theta > 0$ depends on M,n and K. Note that, due to (B.32), one has

$$\ln |\mathcal{G}_{M,K}| \geq (M \ln K)/36 \tag{B.34}$$

Let $\mathbf{Z}_1, \mathbf{Z}_2 \in \mathcal{S}_{M,K}$ be two clustering matrices. Set $\mathbf{G}_1 = \theta \mathbf{W} \mathbf{Z}_1^T$ $\mathbf{G}_2 = \theta \mathbf{W} \mathbf{Z}_2^T$, so that $\mathbf{G}_1, \mathbf{G}_2 \in \mathcal{G}_{M,K}$. Since for any i, i' one has $\|\mathbf{w}_i - \mathbf{w}_{i'}\|_0 = \|\mathbf{w}_i - \mathbf{w}_{i'}\|^2$, derive that

$$\begin{aligned} \|\theta \mathbf{W} \left(\mathbf{Z}_1 - \mathbf{Z}_2\right)^T \|_F^2 = \sum_{m=1}^{M} \sum_{j=1}^{n} \theta^2 \left[\left(\mathbf{w}_{z_1(m)}\right)_j - \left(\mathbf{w}_{z_2(m)}\right)_j \right]^2 = \\ = \theta^2 \sum_{m=1}^{M} \|\mathbf{w}_{z_1(m)} - \mathbf{w}_{z_2(m)}\|^2 \geq \#\{m : z_1(m) \neq z_2(m)\} \, \theta^2 |J|/8. \end{aligned} \tag{B.35}$$

On the other hand, observe that for $\mathbf{Z}_1, \mathbf{Z}_2 \in \mathcal{S}_{M,K}$ one has

$$\#\{m : z_1(m) \neq z_2(m)\} = 0.5 \, \|\mathbf{Z}_1 - \mathbf{Z}_2\|_H \geq dM/2.$$

Therefore, the last two inequalities yield for any $\mathbf{G}_1, \mathbf{G}_2 \in \mathcal{G}_{M,K}$

$$\|\mathbf{G}_1 - \mathbf{G}_2\|_F^2 \geq d \, \theta^2 |J| M/16. \tag{B.36}$$

Now, it is easy to calculate that for any $\mathbf{G}_1, \mathbf{G}_2 \in \mathcal{G}_{M,K}$ and corresponding probability measures $P_{\mathbf{G}_1}$ and $P_{\mathbf{G}_2}$, one has

$$K\left(P_{\mathbf{G}_1}, P_{\mathbf{G}_2}\right) \leq \frac{n}{2\sigma^2 C_\psi^2} \|\mathbf{\Upsilon}^{-1} \left(\mathbf{G}_2 - \mathbf{G}_1\right)\|_F^2 \tag{B.37}$$

Since $\mathbf{G}_1 = \theta \mathbf{W}\mathbf{Z}_1$, $\mathbf{G}_2 = \theta \mathbf{W}\mathbf{Z}_2$, we obtain

$$\| \boldsymbol{\Upsilon}^{-1} (\mathbf{G}_2 - \mathbf{G}_1) \|_F^2 \leq \theta^2 \, \| \mathbf{Z}_2 - \mathbf{Z}_1 \|_{op}^2 \, \| \boldsymbol{\Upsilon}^{-1} \mathbf{W} \|_F^2 \tag{B.38}$$

Note that $\mathcal{S}_{M,K}(dM) \subset \mathcal{Z}(M,K)$, so that for any $\mathbf{Z} \in \mathcal{S}_{M,K}(dM)$ one has $\mathbf{Z}^T\mathbf{Z} = (M/K)\,\mathbf{I}_K$, hence $\|\mathbf{Z}\|_{op} = \sqrt{M/K}$. Then, $\|\mathbf{Z}_1 - \mathbf{Z}_2\|_{op}^2 \leq 4M/K$. Also, due to $J = \{L_1, ..., L_2\}$ and condition (4.3), one has

$$\sum_{j \in J} \nu_j^{-2} \leq \aleph_1^{-2} |J| \, L_1^{-2\gamma} \exp\left(-2\alpha L_1^{\beta}\right). \tag{B.39}$$

Since $\| \boldsymbol{\Upsilon}^{-1} \mathbf{W} \|_F^2 = \sum_{k=1}^{K} \sum_{j \in J} \nu_j^{-2}$, obtain

$$K\left(P_{\mathbf{G}_1}, P_{\mathbf{G}_2}\right) \leq \frac{2}{\aleph_1^2 C_\psi^2} \, \theta^2 \sigma^{-2} n |J| M \, L_1^{-2\gamma} \exp\left(-2\alpha L_1^{\beta}\right). \tag{B.40}$$

Finally, due to condition (4.13), one needs $\theta^2 \sum_{j \in J} (j+1)^{2r} \leq \mathcal{A}^2$, so that we can choose

$$\theta^2 = \mathcal{A}^2 |J|^{-1} L_2^{-2r} \tag{B.41}$$

In order to apply Theorem 2.5 of Tsybakov (2009), we need $K\left(P_{\mathbf{G}_1}, P_{\mathbf{G}_2}\right) \leq \alpha \ln |\mathcal{G}_{M,K}|$ which, due to (B.32), is guaranteed by

$$\frac{\theta^2 n |J|}{\sigma^2 \aleph_1^2 C_\psi^2} L_1^{-2\gamma} \exp\left(-2\alpha L_1^{\beta}\right) \leq \frac{\ln K}{648}. \tag{B.42}$$

If inequality (B.42) holds, then application of Theorem 2.5 of Tsybakov (2009) with $\alpha = 1/9$ yields that, with probability at least 0.1, one has $R_{\min}(\mathcal{S}(r, \mathcal{A}), M, K_*) \geq C R_{\min}(M, K_*, n)$

where, due to (4.1) and (B.36),

$$R_{\min}(M, K_*, n) = \theta^2 |J|. \tag{B.43}$$

Now, we consider two choices of L_1 and L_2: $L_1 = L_2 = L$ and $L_1 = L/2+1$, $L_2 = L$ leading to the following values of θ^2:

$$\theta^2 \asymp \begin{cases} L^{-2r}, & \text{if} \quad L_1 = L_2 = L \\ L^{-(2r+1)}, & \text{if} \quad L_1 = L/2+1, L_2 = L \end{cases} \tag{B.44}$$

We study the cases of $\alpha = \beta = 0$ and $\alpha > 0$, $\beta > 0$ separately.

Case 1: $\alpha = 0$, $\beta = 0$, $L_1 = L_2 = L$, $|J| = 1$.

In this case, by (B.44), inequality (B.42) holds if $L \asymp (\sigma^2 n^{-1} \ln K)^{-\frac{1}{2r+2\gamma}}$. Hence,

$$R_{\min}(M, K_*, n) \left(\sigma^2 n^{-1} \ln K_*\right)^{\frac{2r}{2r+2\gamma}} . \tag{B.45}$$

Case 1 (b) $\alpha = 0$, $\beta = 0$, $L_1 = L/2 + 1$, $L_2 = L$, $|J| = L/2$.

Since $L_1 \asymp L_2 \asymp |J| \asymp L$, inequality (B.42) holds if $L \asymp (\sigma^2 n^{-1} \ln K)^{-\frac{1}{2r+2\gamma}}$ and

$$R_{\min}(M, K_*, n) \left(\sigma^2 n^{-1} \ln K_*\right)^{\frac{2r}{2r+2\gamma}} . \tag{B.46}$$

Case 2: $\alpha > 0$, $\beta > 0$, $L_1 = L_2 = L$, $|J| = 1$.

Plugging the first expression from (B.44) into (B.42), derive that $L^{-(2\gamma+2r)} \exp\left(-2\alpha L^{\beta}\right) \sigma^2 n^{-1} \ln K$,

so that $L \asymp \left[\ln\left(\frac{n}{\sigma^2 \ln K}\right)\right]^{\frac{1}{\beta}}$. Therefore,

$$R_{\min}(M, K_*, n) \left[\ln\left(\frac{n}{\sigma^2 \ln K_*}\right)\right]^{-\frac{2r}{\beta}}. \tag{B.47}$$

Case 2 (b) $\alpha > 0$, $\beta > 0$, $L_1 = L/2 + 1$, $L_2 = L$, $|J| = L/2$.

Plugging the second expression from (B.44) into (B.42), derive that $L^{-(2\gamma+2r)} \exp\left(-2\alpha L^{\beta}\right) \sigma^2 n^{-1} \ln K$. Then, the asymptotic value of L is the same as in the Case 2(a) and

$$R_{\min}(M, K_*, n) \left[\ln\left(\frac{n}{\sigma^2 \ln K_*}\right)\right]^{-\frac{2r}{\beta}}. \tag{B.48}$$

<u>**Lower bound for the error due to estimation**</u>.

Let, as before, $J = \{L_1, ..., L_2\}$ where $1 \leq L_1 < L_2 \leq n$. Consider a set of binary vectors $\boldsymbol{\omega} \in \{0,1\}^{|J|K}$ and set $N = |J|K$. Complete vectors $\boldsymbol{\omega}$ with zeros to obtain vectors $\mathbf{w} \in \{0,1\}^{nK}$. By Varshomov Gilbert lemma, there exists a subset $\mathcal{B}$ of those vectors such that for any $\mathbf{w}, \mathbf{w}' \in \mathcal{B}$ such that $\mathbf{w} \neq \mathbf{w}'$ one has $\|\mathbf{w} - \mathbf{w}'\|_H \geq N/8$ and $\ln|\mathcal{B}| \geq N \ln(2)/8$. Pack vectors $\mathbf{w}$ into matrices $\mathbf{W} \in \{0,1\}^{n \times K}$. Denote the set of those matrices by $\mathcal{W}$ and observe that

$$\|\mathbf{W}_1 - \mathbf{W}_2\|_F \geq N/8 \quad \text{for all} \quad \mathbf{W}_1, \mathbf{W}_2 \in \mathcal{W},\ \mathbf{W}_1 \neq \mathbf{W}_2; \qquad \ln|\mathcal{W}| \geq N/8. \tag{B.49}$$

Let $\mathbf{Z}$ be the clustering matrix that corresponds to uniform sequential clustering, M/K

vectors per class. Finally, form the set $\mathcal{G}_{M,K}$ of matrices $\mathbf{G}$ of the form

$$\mathcal{G}_{M,K} = \left\{ \mathbf{G} \in R^{M \times K} : \mathbf{G} = \theta \mathbf{W} \mathbf{Z}^T, \quad \mathbf{W} \in \mathcal{W} \right\}$$

where $\theta > 0$ depends on M,n and K. Then, for any $\mathbf{G}_1, \mathbf{G}_2 \in \mathcal{G}_{M,K}$, $\mathbf{G}_1 \neq \mathbf{G}_2$, due to $\mathbf{Z}^T\mathbf{Z} = (M/K)\,\mathbf{I}_K$ and (B.49), obtain

$$\|(\mathbf{G}_1 - \mathbf{G}_2)\|_F^2 = \boldsymbol{\theta}^2 \|(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{Z}^T\|_F^2 = \frac{\boldsymbol{\theta}^2 M}{K} \|\mathbf{W}_1 - \mathbf{W}_2\|_F^2 \geq \frac{\boldsymbol{\theta}^2 MN}{8K} \tag{B.50}$$

Now, since $\mathbf{G}_1 = \theta \mathbf{W}_1 \mathbf{Z}$ and $\mathbf{G}_2 = \theta \mathbf{W}_2 \mathbf{Z}$, using formula (B.37), derive that

$$K\left(P_{\mathbf{G}_1}, P_{\mathbf{G}_2}\right) \leq \frac{n\,\theta^2}{2\sigma^2 C_\psi^2} \| \boldsymbol{\Upsilon}^{-1} \left(\mathbf{W}_2 - \mathbf{W}_1\right) \|_F^2 \, \|\mathbf{Z}\|_{op}^2$$

Recalling that $\|\mathbf{Z}\|_{op}^2 = M/K$ and $\| \boldsymbol{\Upsilon}^{-1} \left(\mathbf{W}_2 - \mathbf{W}_1\right) \|_F^2 \leq \sum_{k=1}^K \sum_{j \in J} \nu_j^{-2}$, and using (B.39), arrive at

$$K\left(P_{\mathbf{G}_1}, P_{\mathbf{G}_2}\right) \leq \frac{nM\boldsymbol{\theta}^2}{2\sigma^2\, \aleph_1^2\, C_\psi^2} \, |J| \, L_1^{-2\gamma} \exp\left(-2\alpha L_1^\beta\right).$$

In order to apply Theorem 2.5 of Tsybakov (2009), we need $K\left(P_{\mathbf{G}_1}, P_{\mathbf{G}_2}\right) \leq \alpha \ln |\mathcal{G}_{M,K}|$ which, due to (B.49), is guaranteed by

$$\frac{\theta^2 nM}{\sigma^2 \aleph_1^2 C_\psi^2} L_1^{-2\gamma} \exp\left(-2\alpha L_1^\beta\right) \leq \frac{K}{36}. \tag{B.51}$$

If inequality (B.51) holds, then application of Theorem 2.5 of Tsybakov (2009) with $\alpha = 1/9$ yields that, with probability at least 0.1, one has $R_{\min}(\mathcal{S}(r, \mathcal{A}), M, K_*) \geq C R_{\min}(M, K_*, n)$ where, due to (4.1) and (B.36),

$$R_{\min}(M, K_*, n) = \boldsymbol{\theta}^2 |J| \tag{B.52}$$

Now, as before, we consider two choices of L_1 and L_2: $L_1 = L_2 = L$ and $L_1 = L/2 + 1$, $L_2 = L$ leading to the values of θ^2 given by (B.44). Again, we consider the cases of $\alpha = \beta = 0$ and $\alpha > 0$, $\beta > 0$ separately.

Case 3 (a) $\alpha = 0$, $\beta = 0$, $L_1 = L_2 = L$, $|J| = 1$.

In this case, by (B.44), inequality (B.51) holds if $L \asymp (\sigma^2 n^{-1} M^{-1} K)^{-\frac{1}{2r+2\gamma}}$. Hence,

$$R_{\min}(M, K_*, n) \left(\frac{\sigma^2 K}{M\, n}\right)^{\frac{2r}{2r+2\gamma}} . \tag{B.53}$$

Case 3: $\alpha = 0$, $\beta = 0$, $L_1 = L/2 + 1$, $L_2 = L$, $|J| = L/2$.

Since $L_1 \asymp L_2 \asymp |J| \asymp L$, inequality (B.51) holds if $L \asymp (\sigma^2 n^{-1} M^{-1} K)^{-\frac{1}{2r+2\gamma+1}}$ and

$$R_{\min}(M, K_*, n) \left(\frac{\sigma^2 K}{M\, n}\right)^{\frac{2r}{2r+2\gamma+1}} . \tag{B.54}$$

Case 4: $\alpha > 0$, $\beta > 0$, $L_1 = L_2 = L$, $|J| = 1$.

Plugging the first expression from (B.44) into (B.51), derive that $L^{-(2\gamma+2r)} \exp\left(-2\alpha L^{\beta}\right) \sigma^2 n^{-1} M^{-1} K$, so that $L \asymp \left[\ln\left(\frac{Mn}{\sigma^2 K}\right)\right]^{\frac{1}{\beta}}$. Therefore,

$$R_{\min}(M, K_*, n) \left[\ln\left(\frac{Mn}{\sigma^2 K}\right)\right]^{-\frac{2r}{\beta}} \tag{B.55}$$

Now, in order to obtain the expressions for the lower bounds, we find the maximum of (B.45)

and (B.54) if $\alpha = 0$, $\beta = 0$, and of (B.47) and (B.55) if $\alpha > 0$, $\beta > 0$.

Proof of Corollary 1. First observe that expressions (4.11) are obtained directly from (4.6) and (4.7) by setting $M = K_* = 1$ since all functions belong to the same Sobolev ball (4.12). In order to compare the upper bounds (4.6) and (4.7) obtained with clustering with the upper bound (4.11) derived without clustering, we consider several cases.

Case 1 $\alpha = 0$, $\beta = 0$.

If $K_* = 1$, then $\ln K_* = 0$ and

$$R_{\min}(M, K_*, n) = \left(\frac{\sigma^2 \ln n}{Mn}\right)^{\frac{2r}{2r+2\gamma+1}} < \left(\frac{\sigma^2 \ln n}{n}\right)^{\frac{2r}{2\gamma+2r+1}} .$$

Moreover,

$$\tilde{R}(n)/R_{\min}(M, K_*, n) = M^{\frac{2r}{2\gamma+2r+1}} \to \infty \quad \text{as} \quad M \to \infty \tag{B.56}$$

and clustering is asymptotically advantageous. If $K_* \geq 2$, then clustering is advantageous if

$$\left(\frac{\sigma^2 \ln n}{n}\right)^{\frac{2r}{2r+2\gamma+1}} \geq \max\left\{\left(\frac{\sigma^2 \ln K_*}{n}\right)^{\frac{2r}{2r+2\gamma}}, \left(\frac{\sigma^2 K_*}{M n}\right)^{\frac{2r}{2r+2\gamma+1}}, \left(\frac{\sigma^2 \ln n}{M n}\right)^{\frac{2r}{2r+2\gamma+1}}\right\} \tag{B.57}$$

Compare $(\sigma^2 n^{-1} \ln n)^{\frac{2r}{2r+2\gamma+1}}$ with the first term in the maximum in (B.57)

$$\left(\frac{\sigma^2 \ln n}{n}\right)^{\frac{2r}{2r+2\gamma+1}} \Bigg/ \left(\frac{\sigma^2 \ln K_*}{n}\right)^{\frac{2r}{2r+2\gamma}} = \left[\frac{n (\ln n)^{2r+2\gamma}}{\sigma^2 (\ln K_*)^{2r+2\gamma+1}}\right]^{\frac{2r}{(2r+2\gamma)(2r+2\gamma+1)}}$$

Consider now the ratio between $(\sigma^2 n^{-1} \ln n)^{\frac{2r}{2r+2\gamma+1}}$ and the second term in the maximum in (B.57):

$$\left(\frac{\sigma^2 \ln n}{n}\right)^{\frac{2r}{2r+2\gamma+1}} \Bigg/ \left(\frac{\sigma^2 K_*}{M n}\right)^{\frac{2r}{2r+2\gamma+1}} = \left(\frac{M \ln n}{K_*}\right)^{\frac{2r}{2r+2\gamma+1}}$$

Finally the ratio between $(\sigma^2 \ln n \; n^{-1})^{\frac{2r}{2r+2\gamma+1}}$ and the last term in(B.57) is

$$\left(\frac{\sigma^2 \ln n}{n}\right)^{\frac{2r}{2r+2\gamma+1}} \Bigg/ \left(\frac{\sigma^2 \ln n}{M\, n}\right)^{\frac{2r}{2r+2\gamma+1}} = M^{\frac{2r}{2r+2\gamma+1}}$$

Therefore, clustering asymptotically reduces the estimation error

$$\frac{\tilde{R}(n)}{R_{\min}(M, K_*, n)} \to \infty \quad \text{if} \quad n \to \infty,\ M \to \infty,\ \frac{n\,(\ln n)^{2r+2\gamma}}{\sigma^2\,(\ln K_*)^{2r+2\gamma+1}} \to \infty. \tag{B.58}$$

Case 2 $\alpha > 0$, $\beta > 0$.

If $K_* = 1$, then, due to the condition (4.5),

$$R_{\min}(M, K_*, n) = \left[\ln\left(\frac{Mn}{\sigma^2}\right)\right]^{-\frac{2r}{\beta}} \asymp \left[\ln\left(\frac{n}{\sigma^2}\right)\right]^{-\frac{2r}{\beta}} = \tilde{R}(n). \tag{B.59}$$

If $K_* \geq 2$, then one has $\ln(n/\ln K_*) < 2 \ln n$ and, since $\ln K_* < \ln M$, also

$$\ln(n/\ln K_*) \geq \ln(n/\ln M) \asymp \ln n - \ln\ln M \asymp \ln n - \ln\ln n \asymp \ln n.$$

Hence,

$$R_{\min}(M, K_*, n) = \left[\ln\left(\frac{n}{\sigma^2\ \ln K_*}\right)\right]^{-\frac{2r}{\beta}} \asymp \left[\ln\left(\frac{n}{\sigma^2}\right)\right]^{-\frac{2r}{\beta}} = \tilde{R}(n), \tag{B.60}$$

which completes the proof.

Proof of Lemma 5. Proof of Lemma 5 is based on the following statement provided in Gendre(1999)

Lemma 8 Gendre(1999). *Let* $\mathbf{A} \in R^{p\times p}$ *be a fixed matrix and* $\boldsymbol{\epsilon} \sim N(0, \mathbf{I}_p)$. *Then, for*

any $x > 0$ *one has*

$$\mathbb{P}(\|\mathbf{A}\boldsymbol{\epsilon}\|^2 \geq Tr(\mathbf{A}^T\mathbf{A})) + 2\sqrt{\|\mathbf{A}\|_{op}^2 Tr(\mathbf{A}^T\mathbf{A})x} + 2\|\mathbf{A}\|_{op}^2 x) \leq e^{-x} \tag{B.61}$$

Note that, due to $2ab \leq a^2 + b^2$, probability (B.61) can be re-written as

$$\mathbb{P}(\|\mathbf{A}\boldsymbol{\epsilon}\|^2 \geq 2\|\mathbf{A}\|_F^2 + 3\|\mathbf{A}\|_{op}^2 x) \leq e^{-x} \tag{B.62}$$

Consider $\|(\mathbf{\Pi}_{\mathbf{Z},K} \otimes (\mathbf{W}_J\Upsilon\mathbf{S}))\boldsymbol{\delta}\|^2$ where $\mathbf{Z}, J, K$ fixed. Note that, due to $\|\mathbf{\Pi}_{\mathbf{Z},K}\|_{op}^2 = 1$, $\|\mathbf{S}\|_{op}^2 \leq C_\psi^2$, $\|\mathbf{W}_J\Upsilon\|_{op}^2 = \max_{j\in J}\nu_j^2$ and $\|\mathbf{W}_J\Upsilon\|_F^2 = K\sum_{j\in J}\nu_j^2$, one has

$$\|(\mathbf{\Pi}_{\mathbf{Z},K} \otimes (\mathbf{W}_J\Upsilon\mathbf{S}))\|_{op}^2 \leq \|\mathbf{\Pi}_{\mathbf{Z},K}\|_{op}^2\|\mathbf{W}_J\Upsilon\|_{op}^2\|\mathbf{S}\|_{op}^2 \leq C_\psi^2 \max_{j\in J}\nu_j^2 \tag{B.63}$$

$$\|(\mathbf{\Pi}_{\mathbf{Z},K} \otimes (\mathbf{W}_J\Upsilon\mathbf{S}))\boldsymbol{\delta}\|_F^2 \leq \|\mathbf{\Pi}_{\mathbf{Z},K}\|_{op}^2\|\mathbf{W}_J\Upsilon\|_F^2\|\mathbf{S}\|_{op}^2 \leq KC_\psi^2\sum_{j\in J}\nu_j^2 \tag{B.64}$$

Now applying inequality (B.62) to $\|(\mathbf{\Pi}_{\mathbf{Z},K} \otimes (\mathbf{W}_J\Upsilon\mathbf{S}))\boldsymbol{\delta}\|^2$ where $\boldsymbol{\delta} \sim N(0, \mathbf{I}_{nM})$, obtain for any $x > 0$

$$\begin{aligned}&\mathbb{P}\left(\|(\mathbf{\Pi}_{\mathbf{Z},K} \otimes (\mathbf{W}_J\Upsilon\mathbf{S}))\boldsymbol{\delta}\|^2 \geq 2\|(\mathbf{\Pi}_{\mathbf{Z},K} \otimes (\mathbf{W}_J\Upsilon\mathbf{S}))\|_F^2 + 3\|(\mathbf{\Pi}_{\mathbf{Z},K} \otimes (\mathbf{W}_J\Upsilon\mathbf{S}))\|_{op}^2 x\right) \leq \\ &\mathbb{P}\left(\|(\mathbf{\Pi}_{\mathbf{Z},K} \otimes (\mathbf{W}_J\Upsilon\mathbf{S}))\boldsymbol{\delta}\|^2 - C_\psi^2\left[2K\sum_{j\in J}\nu_j^2 + 3x\max_{j\in J}\nu_j^2\right] \geq 0\right) \leq e^{-x},\end{aligned} \tag{B.65}$$

setting $s = \tau \ln n$ yields (B.8).

In order to prove inequality (B.9), note that for

$$x(M,K,|J|,s) = M\ln K + |J|\ln(\frac{ne}{|J|}) + \ln n + \ln M + s$$

due to $\ln\binom{n}{j} \le j\ln(\frac{ne}{j})$, one has

$$\begin{aligned}\sum_{\mathbf{Z},K,J} e^{-x(M,K,|J|,s)} &\equiv \sum_{K=1}^{M}\sum_{j=1}^{n}\sum_{|J|=j}\sum_{\mathbf{Z}\in\mathcal{M}(M,K)} e^{-x(M,K,j,s)} \\ &= \sum_{K=1}^{M}\sum_{j=1}^{n}\binom{n}{j}K^M e^{-x(M,K,j,s)} \\ &\le \sum_{K=1}^{M}\sum_{j=1}^{n}\left(\frac{ne}{j}\right)^j K^M e^{-x(M,K,j,s)} \le e^{-s} \qquad \text{(B.66)}\end{aligned}$$

Therefore, by (B.65) and (B.66), obtain

$$\mathbb{P}\left(\|(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}}\otimes(\mathbf{W}_{\hat{J}}\Upsilon\mathbf{S}))\boldsymbol{\delta}\|^2 - 2\|(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}}\otimes(\mathbf{W}_{\hat{J}}\Upsilon\mathbf{S}))\|_F^2 - 3\|(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}}\otimes(\mathbf{W}_{\hat{J}}\Upsilon\mathbf{S}))\|_{op}^2\, x(M,\hat{K},|\hat{J}|,s) \ge 0\right) \le$$
$$\sum_{\mathbf{Z},K,J}\mathbb{P}\left(\|(\mathbf{\Pi}_{\mathbf{Z},K}\otimes(\mathbf{W}_J\Upsilon\mathbf{S}))\boldsymbol{\delta}\|^2 - C_\psi^2\left[2K\sum_{j\in J}\nu_j^2 + 3\,x(M,K,|J|,s)\left(\max_{j\in J}\nu_j^2\right)\right] \ge 0\right) \le$$
$$\sum_{\mathbf{Z},K,J} e^{-x(M,K,|J|,s)} \le e^{-s}$$

setting $s = \tau\ln n$ yields (B.9). □

Proof of Lemma 7. By using (B.31), $K \ge 2$ and $0 < d < 1/4$

$$\ln K - 4d\ln(Ke/d) = \ln K - 4[d\ln(K) + d - d\ln d] \ge \ln K - 4d\ln K - \frac{4}{9}\ln K \ge \frac{5}{9}\ln K - \ln K \ge \frac{\ln K}{9}.$$

Starting from (4.19) we can expand it follows

$$\|(\mathbf{I} - \mathbf{W}_J)\boldsymbol{\Upsilon}\mathbf{G}\boldsymbol{\Pi}_{\mathbf{Z},K}\|_F^2 + \|\boldsymbol{\Upsilon}\mathbf{G}\boldsymbol{\Pi}_{\mathbf{Z},K}^{\perp}\|_F^2$$

$$= \|\mathbf{G}\|_F^2 - 2\langle G, \boldsymbol{\Upsilon}\mathbf{Y}\boldsymbol{\Pi}_{\mathbf{Z}}\rangle + \|\boldsymbol{\Upsilon}\mathbf{Y}\boldsymbol{\Pi}_{\mathbf{Z}}\|_F^2 + \|\boldsymbol{\Upsilon}\mathbf{Y}\boldsymbol{\Pi}_{\mathbf{Z}}^{\perp}\|_F^2$$

$$= \|\mathbf{G}\|_F^2 - 2\langle G, \boldsymbol{\Upsilon}\mathbf{Y}\boldsymbol{\Pi}_{\mathbf{Z}}\rangle + \|\boldsymbol{\Upsilon}\mathbf{Y}\|_F^2; \text{ combining last two terms}$$

$$= \|\mathbf{G}\|_F^2 - 2\langle G, \boldsymbol{\Upsilon}\mathbf{Y}\boldsymbol{\Pi}_{\mathbf{Z}}\rangle; \text{ last term independent from } \mathbf{U}, \mathbf{G}, J$$

$$= \|\mathbf{G}\|_F^2 - 2\mathrm{Tr}(\mathbf{Y}^T\boldsymbol{\Upsilon}\mathbf{G}\boldsymbol{\Pi}_{\mathbf{Z}}); \text{ writing inner product as a trace and cyclic rearrangement inside.}$$

LIST OF REFERENCES

[1] Abramovich, F., Pensky, M., Rozenholc, Y. (2013). Laplace deconvolution with noisy observations. *Electron. J. Stat.* **7**, 1094-1128.

[2] Abramovich F and Silverman B W 1998 Wavelet decomposition approaches to statistical inverse problems *Biometrika.* **85** 115–129

[3] Allen, Elena A., et al. "Tracking whole-brain connectivity dynamics in the resting state." Cerebral cortex 24.3 (2014): 663-676.

[4] Alquier P, Gautier E and Gilles Stoltz G 2011 *Inverse Problems and High-Dimensional Estimation* (Berlin: Springer-Verlag)

[5] Arnold A, Reichling S, Bruhns O T and Mosler J 2010 Efficient computation of the elastography inverse problem by combining variational mesh adaption and a clustering technique *Phys Med Biol.* **55** 2035-2056

[6] Atherton, Michael, et al. "A functional MRI study of high-level cognition. I. The game of chess." Cognitive Brain Research 16.1 (2003): 26-31.

[7] Benhaddou, R., Pensky, M., Rajapakshage R., (2018). Anisotropic functional Laplace deconvolution. *Journ. Statist. Plan. Inf.* **199**, 271-285.

[8] Benhaddou, R., Pensky, M., Picard, D. (2013). Anisotropic denoising in functional deconvolution model with dimension-free convergence rates. *Electron. J. Stat.* **7**, 1686-1715.

[9] Bezdek J C and Pal S K 1992 *Fuzzy models for pattern recognition methods that search for structures in data* (New York : IEEE Press)

[10] Bissantz N, Hohage T, Munk A, and Ruymgaart F 2007 Convergence rates of general regularization methods for statistical inverse problems and applications *SIAM J. Numer. Anal.* **45** 2610-2636

[11] Bisdas, S., Konstantinou, G.N., Lee, P.S., Thng, C.H., Wagenblast, J., Baghi, M., Koh, T.S. (2007). Dynamic contrast-enhanced CT of head and neck tumors: perfusion measurements using a distributed-parameter tracer kinetic model. Initial results and comparison with deconvolution- based analysis. *Physics in Medicine and Biology.* **52**, 6181-6196.

[12] Böttcher, A., and Grudsky, S.M. (2000). *Toeplitz Matrices, Asymptotic Linear Algebra, and Functional Analysis.* Birkhauser Verlag, Basel-Boston-Berlin.

[13] Böttcher, A., and Grudsky, S.M. (2005). *Spectral Properties of Banded Toeplitz Matrices*, SIAM, Philadelphia.

[14] Bunea, F., Tsybakov, A. and Wegkamp, M.H. (2007). Aggregation for Gaussian regression. *Ann. Statist.* **35**, 1674–1697.

[15] Cao, Y. (2011). The promise of dynamic contrast-enhanced imaging in radiation therapy. *Semin Radiat Oncol.* **2**, 147–56.

[16] Cao, M., Liang, Y., Shen, C., Miller, K.D., Stantz, K.M. (2010). Developing DCE-CT to quantify Intra-Tumor heterogeneity in breast tumors with differing angiogenic phenotype. *IEEE Trans. on Medical Imaging.* **29**, 1089–1092.

[17] Chao-Gan, Yan, and Zang Yu-Feng. "DPARSF: a MATLAB toolbox for pipeline data analysis of resting-state fMRI." Frontiers in systems neuroscience 4 (2010).

[18] Coloigner, J., Phlypo, R., Coates, T. D., Lepore, N., & Wood, J. C. (2017). Graph Lasso-Based Test for Evaluating Functional Brain Connectivity in Sickle Cell Disease. Brain Connectivity, 7(7), 443453.

[19] Comte, F., Genon-Catalot, V. (2015). Adaptive Laguerre density estimation for a mixed Poisson model. *Electron. J. Stat.* **9**, 1113–1149

[20] Cohen A, Hoffmann M and Reis M 2004 Adaptive wavelet Galerkin methods for linear inverse problems *SIAM Journ. Numer. Anal.* **42** 1479-1501

[21] Comte F, Cuenod C A, Pensky M and Rozenholc Y 2017 Laplace deconvolution on the basis of time domain data and its application to Dynamic Contrast Enhanced imaging *Journ. Royal Stat. Soc., Ser.B.* **79** 69-94

[22] Cribben, Ivor, et al. "Dynamic connectivity regression: determining state-related changes in brain connectivity." Neuroimage 61.4 (2012): 907-920.

[23] Danaher, P., Wang, P., & Witten, D. M. (2014). The joint graphical lasso for inverse covariance estimation across multiple classes. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 76(2), 373-397.

[24] Das, Anup et al. Interpretation of the Precision Matrix and Its Application in Estimating Sparse Brain Connectivity during Sleep Spindles from Human Electrocorticography Recordings. Neural Computation 29 (2017): 603-642.

[25] Delamillieure, Pascal, et al. "The resting state questionnaire: an introspective questionnaire for evaluation of inner experience during the conscious resting state." Brain research bulletin 81.6 (2010): 565-573.

[26] Deng Z, Chung F L and Wang S 2011 Clustering-Inverse: A Generalized Model for Pattern-Based Time Series Segmentation. *Journal of Intelligent Learning Systems and Applications* **3** 26-36

[27] Dey, A.K., Martin, C.F., Ruymgaart, F.H. (1998). Input recovery from noisy output data, using regularized inversion of Laplace transform. *IEEE Trans. Inform. Theory.* **44**, 1125–1130.

[28] Donoho D L 1995 Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition *Applied and Computational Harmonic Analysis* **2** 101–126

[29] Donoho D L and Johnstone I M 1994 Ideal spatial adaptation by wavelet shrinkage *Biometrika* **81** 425-456

[30] Duan, Xujun, et al. "Functional organization of intrinsic connectivity networks in Chinese-chess experts." Brain research 1558 (2014): 33-43.

[31] Engl H W, Hanke M and Neubauer A 2000 *Regularization of Inverse Problems* (Netherlands: Kluwer Academic Publishers)

[32] Fraix-Burnet D and Girard S 2016 *Statistics for Astrophysics Clustering and Classification* (EDP Sciences)

[33] Friedman, J., Hastie, T.& Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. Biostatistics, 9, 432–441.

[34] Goebel, Rainer, et al. "Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping." Magnetic resonance imaging 21.10 (2003): 1251-1261.

[35] Golay, Xavier, et al. "A new correlationbased fuzzy logic clustering algorithm for FMRI." Magnetic Resonance in Medicine 40.2 (1998): 249-260.

[36] Gradshtein, I.S., Ryzhik, I.M. (1980). *Tables of integrals, series, and products.* Academic Press, New York.

[37] Gupta A K and Nagar D K 1999 *Matrix Variate Distributions* (Chapman& Hall/CRC Boca Raton)

[38] Harrison, L., William D. Penny, and Karl Friston. "Multivariate autoregressive modeling of fMRI time series." Neuroimage 19.4 (2003): 1477-1491.

[39] Harrison, L.,Penny, W.D. Friston, K. (2003). Multivariate autoregressive modeling of fMRI time series. NeuroImage, 19, 1477–1491.

[40] Khazaee, Ali, et al. "Classification of patients with MCI and AD from healthy controls using directed graph measures of resting-state fMRI." Behavioural brain research 322 (2017): 339-350.

[41] Klopp O, Lu Y, Tsybakov A B and Zhou H H 2017 Structured matrix estimation and completion `ArXiv:1707.02090`

[42] König C, Werner F and Hohage T 2016 Convergence rates for exponentially ill-posed inverse problems with impulsive noise. *SIAM J. Numer. Anal.* **54** 341–360

[43] Kucyi, Aaron, and Karen D. Davis. "Dynamic functional connectivity of the default mode network tracks daydreaming." Neuroimage 100 (2014): 471-480.

[44] Kürüm E, Weber G W and Iyigun C 2018 Early warning on stock market bubbles via methods of optimization, clustering and inverse problems *Annals of Operations Research* **260** 293-320

[45] Leonardi, Nora, et al. "Principal components of functional connectivity: a new approach to study dynamic brain connectivity during rest." NeuroImage 83 (2013): 937-950.

[46] Lieury, T., Pouzat, C. and Rozenholc, Y. (2012) DynClust: non-parametric denoising and clustering method of noisy images both indexed by time and space, *R-package* available at `http://cran.r-project.org/`.

[47] Mabon, G. (2016). Adaptive deconvolution of linear functionals on the nonnegative real line. *Journ. Statist. Plan. Inf.* **178**, 1-23.

[48] Mclntosh, A. R., and F. GonzalezLima. "Structural equation modeling and its application to network analysis in functional brain imaging." Human brain mapping 2.12 (1994): 2-22.

[49] Mastrovito, Dana. "Interactions between resting-state and task-evoked brain activity suggest a different approach to fMRI analysis." Journal of Neuroscience 33.32 (2013): 12912-12914.

[50] Ng, Bernard, et al. "A novel sparse graphical approach for multimodal brain connectivity inference." International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Berlin, Heidelberg, 2012.

[51] Pensky M 2013 Spatially inhomogeneous linear inverse problems with possible singularities *Annals of Statistics* **41** 2668–2697

[52] Pensky M 2016 Solution of linear ill-posed problems using overcomplete dictionaries *Annals of Statistics* **44** 1739–1764

[53] Pensky M 2018 Dynamic network models and graphon estimation *Annals of Statistics* accepted

[54] Phlypo R, Thirion B, Varoquaux G. 2014. Deriving a multisubject functional-connectivity atlas to inform connectome estimation. In: Golland P, editor; , Hata N,

editor; , Barillot C, editor; , Hornegger J, editor; , Howe J, editor. (eds.) Medical Image Computing and Computer-Assisted Intervention-MICCAI 2014. Cham, Switzerland: Springer International Publishing; pp. 185192 [PubMed]

[55] Preti, Maria Giulia, Thomas AW Bolton, and Dimitri Van De Ville. "The dynamic functional connectome: state-of-the-art and perspectives." Neuroimage 160 (2017): 41-54.

[56] REMISCAN - Project number IDRCB 2007-A00518-45/P060407/STIC 2006; Research Ethics Board (REB) approved- cohort funding by INCa (1M Euros) and promoted by the AP-HP (Assistance Publique Hôpitaux de Paris). Inclusion target: 100 patients. Start in 2007. Closed since July 2015.

[57] Ramsey, Joseph D., et al. "Six problems for causal inference from fMRI." neuroimage 49.2 (2010): 1545-1558.

[58] Rajapakshage, R., Pensky, M., (2018). Clustering in statistical ill-posed linear inverse problems. arXiv:1810.06989, version 1.

[59] Rozenholc, Y., Reiß, M. (2012) Preserving time structures while denoising a dynamical image, *Mathematical Methods for Signal and Image Analysis and Representation (Chapter 12)*, Florack, L. and Duits, R. and Jongbloed, G. and van Lieshout, M.-C. and Davies, L. Ed., Springer-Verlag, Berlin.

[60] Sakolu, nal, et al. "A method for evaluating dynamic functional network connectivity and task-modulation: application to schizophrenia." Magnetic Resonance Materials in Physics, Biology and Medicine 23.5-6 (2010): 351-366.

[61] Smalheiser R., Neil(2017). Data Literacy: Chapter 11-ANOVA, Academic Press ,pp 149 - 155.

[62] Starck J L and Pantin E 2002 Deconvolution in Astronomy : A Review *Publ. Astronom. Soc. of the Pacific* **114** 1051-1069

[63] Tibshirani, Robert. "Regression shrinkage and selection via the lasso." Journal of the Royal Statistical Society. Series B (Methodological) (1996): 267-288.

[64] Tsybakov A B 2009 *Introduction to Nonparametric Estimation* (New York: Springer)

[65] Van Essen, David C., et al. "The WU-Minn human connectome project: an overview." Neuroimage 80 (2013): 62-79.

[66] Vareschi, T. (2013). Noisy Laplace deconvolution with error in the operator. ArXiv 1303.7437, version 2.

[67] Vareschi, T. (2015). Noisy Laplace deconvolution with error in the operator. *Journ. Statist. Plan. Inf.* **157-158**, 16-35.

[68] Wee, ChongYaw, PewThian Yap, and Dinggang Shen. "Diagnosis of autism spectrum disorders using temporally distinct restingstate functional connectivity networks." CNS neuroscience & therapeutics 22.3 (2016): 212-219.

[69] Wen, Xiaotong, Govindan Rangarajan, and Mingzhou Ding. "Is Granger causality a viable technique for analyzing fMRI data?." PloS one 8.7 (2013): e67428.

[70] Werner F and Hohage T 2014 Convergence rates for inverse problems with impulsive noise. *SIAM J. Numer. Anal.* **52** 1203–1221

[71] Wang, Hansheng, Guodong Li, and ChihLing Tsai. "Regression coefficient and autoregressive order shrinkage and selection via the lasso." Journal of the Royal Statistical Society: Series B (Statistical Methodology) 69.1 (2007): 63-78.

[72] Zhou, Yueying, et al. "Simultaneous estimation of low-and high-order functional connectivity for identifying mild cognitive impairment." Frontiers in neuroinformatics 12 (2018): 3.