COMPARISON OF PARTIAL DIRECTED COHERENCE AND DYNAMIC BAYESIAN NETWORK APPROACH FOR BRAIN EFFECTIVE CONNECTIVITY MODELING USING FMRI

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

OĞUZHAN CAN ÖĞE

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

DECEMBER 2019

Approval of the thesis:

COMPARISON OF PARTIAL DIRECTED COHERENCE AND DYNAMIC BAYESIAN NETWORK APPROACH FOR BRAIN EFFECTIVE CONNECTIVITY MODELING USING FMRI

submitted by **OĞUZHAN CAN ÖĞE** in partial fulfillment of the requirements for the degree of **Master of Science in Electrical and Electronics Engineering Department, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. İlkay Ulusoy
Head of Department, **Electrical and Electronics Eng.** _____

Prof. Dr. İlkay Ulusoy
Supervisor, **Electrical and Electronics Eng., METU** _____

Examining Committee Members:

Prof. Dr. Nevzat Güneri Gençer
Electrical and Electronics Eng., METU _____

Prof. Dr. İlkay Ulusoy
Electrical and Electronics Eng., METU _____

Assoc. Prof. Dr. Yeşim Serinağaoğlu
Electrical and Electronics Eng., METU _____

Assoc. Prof. Dr. Cüneyt F. Bazlamaçcı
Electrical and Electronics Eng., METU _____

Prof. Dr. Metehan Çiçek
Faculty of Medicine, Ankara University _____

Date: 03.12.2019

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: Oğuzhan Can Öğe

Signature:

ABSTRACT

COMPARISON OF PARTIAL DIRECTED COHERENCE AND DYNAMIC BAYESIAN NETWORK APPROACH FOR BRAIN EFFECTIVE CONNECTIVITY MODELING USING FMRI

Öğe, Oğuzhan Can
Master of Science, Electrical and Electronics Engineering
Supervisor: Prof. Dr. İlkay Ulusoy

December 2019, 101 pages

Two of the approaches attempting to model brain effective connectivity are compared. These methods are Partial Directed Coherence (PDC) and Dynamic Bayesian Network (DBN). PDC is based on linear and deterministic signal modelling. It is derived from the Granger Causality approach and underpinned by the Multivariate Auto Regressive (MVAR) model. On the other hand, DBN is based on probabilistic signal modelling, which gives DBN the ability of detecting nonlinear interactions between signals unlike all the other estimator methods. In order to compare these two approaches, linear and nonlinear multivariate synthetic fMRI data whose connectivity is known beforehand is generated. In the generation process Hemodynamic Response Function (HRF) is applied after the generation of data by MVAR model. During data generation, the length of the signals, signal-to-noise ratio of the HRF, and complexity of the network (number of channels) are chosen as variables. All in all, these two methods are compared in terms of these parameters. After the comparison, it can be deduced that PDC performs better on linear signals, while it fails on nonlinear signals completely. DBN performs better on nonlinear signals and gives a satisfactory result for linear ones. Since connections in the brain are highly nonlinear and Dynamic Bayesian Network is the only brain effective connectivity estimator method that can differentiate nonlinear signals, it is certain to say that DBN is a more appropriate

approach for connectivity modelling than PDC. This conclusion is supported by applying two methods to real fMRI collections of dyscalculia patients at the end.

Keywords: Brain Effective Connectivity, PDC, DBN, fMRI, HRF

ÖZ

BEYİN ETKİN BAĞLANTISALLIK MODELLEMESİ İÇİN PDC VE DBN YÖNTEMLERİNİN FMRI KULLANARAK KARŞILAŞTIRILMASI

Öğe, Oğuzhan Can
Yüksek Lisans, Elektrik ve Elektronik Mühendisliği
Tez Danışmanı: Prof. Dr. İlkay Ulusoy

Aralık 2019, 101 sayfa

Beyin etkin bağlantısallığı modellemeye çalışan modellerin iki tanesinin karşılaştırması yapılmıştır. Bu modeller Partial Directed Coherence (PDC) ve Dynamic Bayesian Network (DBN) modelleridir. PDC lineer ve deterministik sinyal modellemesini baz alır. Granger Causality yönteminden türetilmiştir ve Multivariate Auto Regressive (MVAR) modelini içinde barındırır. Diğer yandan, DBN olasılıksal sinyal modellemesine dayanır. Bu yöntem DBN'e diğer bağlantısallık tahmin eden modellerin aksine lineer olmayan sinyallari modelleme yeteneği verir. Bu iki yaklaşımı karşılaştırmak için, bağlantısallık bilgisi önceden bilinen, lineer ve lineer olmayan çok değişkenli sentetik fMRI verileri üretilmiştir. Sentetik veri üretim sürecinde, sinyallerin fMRI sinyallerine benzemesi için Hemodynamic Response Function (HRF) üretilen datalar üzerine uygulanmıştır. Sinyalin boyu, HRF için sinyalin gürültüye oranı ve bağlantının karmaşıklığı (kanal sayısı) veri üretim süresince değişken olarak tutulmuştur. Nihayetinde bu iki yöntem belirtilen değişkenlere göre karşılaştırılmıştır. Karşılaştırmadan sonra, PDC yönteminin lineer sinyallerde etkili olduğu çıkarımı yapılırken lineer olmayan sinyallerde başarısız olduğu görülmüştür. DBN ise lineer olmayan sinyallerde daha doğru sonuç verirken, lineer sinyaller içinse yeterli sonuçlar vermiştir. Beyindeki bağlantısallığın çok büyük oranda doğrusal olmaması ve DBN'in bağlantı tahmin eden modeller arasında

doğrusal olmayan bağlantıları bulan tek model olmasından dolayı, DBN yönteminin PDC'ye göre daha uygun bir yöntem olduğu kesin bir şekilde dile getirilebilir. Bu çıkarım, bu iki yöntemi diskalküli hastalarından alınmış fMRI sinyallerine uygulayarak daha da desteklenmiştir.

Anahtar Kelimeler: Beyin Etkin Bağlantısallık, PDC, DBN, fMRI, HRF

To Melis, and my family, who supported me unconditionally through this period.

ACKNOWLEDGEMENTS

I would like to give my thanks to my supervisor Prof. Dr. İlkay Ulusoy for her guidance and feedbacks. She always welcomes me with a benign attitude and a smile and steers me to a right way with sincere conversations.

I would like to give the deepest gratitude to my girlfriend Melis, who stands by me for every achievement and every failure. With her help and support all through this thesis marathon and in life, I have become a better student and better person. She can count this thesis as her own.

I also give my respects and love to my family, especially my mother. They always give me what they can give. They feel when I am down, and even with a phone call they make my day. They are the backstage workers helping me to reach the end of this long road.

I would like to thank my close friends who keep me busy with social activities during the writing process of this thesis. Without them, I would have completed my thesis sooner, and I would have felt empty without all the last-minute stress that enhances my productivity.

I give my special thanks to Salih Geduk, who is also a Master of Science student in METU EEE, for his help and support.

I acknowledge my company ASELSAN to create opportunities for me to continue my Master's.

The real fMRI data used in this thesis, were collected by the researchers of the TÜBİTAK project "Beynin Sayısal İşlevleriyle İlgili Devre Modellerinin Tasarlanması ve Matematik Öğrenme Güçlüğü (Diskalkuli) Hastalık Haritasının Elde Edilmesi" under the management of Prof. Dr. Metehan Çiçek. The project number is 214S069.

Finally, I give the biggest thank to myself for completing this thesis and getting a master’s degree. It has been a very long and challenging journey.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

ABBREVIATIONS

ACC : Anterior Cingulate Cortex

AIC : Akaike Information Criteria

BOLD : Blood Oxygenation Level-Dependent

CPT : Calculation Performance Tests

DBN : Dynamic Bayesian Network

DD : Developmental Dyscalculia

dDBN : discrete Dynamic Bayesian Network

DTF : Directed Transfer Function

ECoG : Electrocorticogram

EEG : Electroencephalogram

fMRI : functional Magnetic Resonance Imaging

GC : Granger Causality

GCI : Granger Causality Index

GCM : Granger Causality Mapping

HIPS : Horizontal Intraparietal Sulcus

HPC : Hippocampus

HRF : Hemodynamic Response Function

IPS : Intraparietal Sulcus

L_IPS : Left Intraparietal Sulcus

MEG : Magnetoencephalogram

MPFC : Middle Prefrontal Cortex

MST : Mathematical Success Tests

MVAR: Multivariate Autoregressive Model

nDTF : normalized Directed Transfer Function

NIRS : Near Infrared Spectroscopy

OCC : Occipital Lobe

PC : Partial Coherence

PD : Parkinson's Disease

PDC : Partial Directed Coherence

PET : Positron Emission Tomography

R_IPS : Right Intraparietal Sulcus

RS : Resting State

SPSS : Statistical Package for the Social Sciences

var : Variance

CHAPTER 1

INTRODUCTION

1.1. Problem Definition, Research Questions and Purpose of the Study

The brain is an organ that controls all functions of the body, senses information from outside world and creates responses, thoughts, language and emotions. It also stores all this information in the memory. The smallest entity of the brain is the neuron. The neuron cells are organized into neural networks. The communication inside the brain can be examined in multiple scales from synaptic connections between single cells to interconnections of brain regions. Sporns (2007) suggests that brain connectivity means a pattern of anatomical links ("anatomical connectivity"), of statistical dependencies ("functional connectivity") or of causal interactions ("effective connectivity") between distinct parts of the brain [1]. Anatomical connectivity studies are focusing on physical white matter paths between brain regions. Functional connectivity studies are interested in linear and nonlinear correlations among regions without concerning if there are physical paths between these regions or not. In both of the connectivity types, there is no comment on the direction of the information flow. They only try to extract the connections and correlations. However, in the brain connectivity analysis, it is very important to differentiate transmitter region from the receiver region in order to comment on interacting systems. Effective connectivity studies are trying to overcome this problem. These studies focus on direct or indirect causal relationships. Mental diseases, such as Alzheimer, dyslexia, dyscalculia, autism, etc., are suggested to associate with the interruption of effective connectivity.

There are several methods that try to model brain effective connectivity. These methods can be listed as Granger Causality, Multivariate Auto Regressive Model (MVAR), Partial Directed Coherence (PDC), Directed Transfer Function (DTF), and

Dynamic Bayesian Networks (DBN). These methods are deeply analyzed in Chapter 2. Among these models, PDC and DBN are selected for comparison. The reason behind this choice can be found in the characteristics of the methods.

Partial Directed Coherence is a deterministic method which is derived from the Granger Causality method and uses the coefficients of the MVAR model. PDC is based on a linear model. Therefore, it hardly detects the nonlinear relations between the brain regions. Among other deterministic methods, PDC is one of the few methods that can differentiate indirect relations. Therefore, in order to represent the deterministic linear models, PDC is a good choice for a comparison.

Dynamic Bayesian Network is a probabilistic method that intersects probability and graph theories. DBN calculations are mostly based on probabilistic approach. Every variable is named as node, and connectivity between regions is expressed as connection between the nodes. The connection between nodes is defined as conditional probability. Conditional probability distribution can be continuous or discrete. In summary, DBN is a probabilistic method which is not linear. In discrete DBN, which is used in this thesis, conditional probabilities can be shown as table and every element of the table is a parameter of the model. Furthermore, it can also detect the linear relations as well as the multivariate relations. It is chosen to compare in order to represent probabilistic and nonlinear nature of the model. In this thesis, discrete DBN (dDBN) is used in simulations. The discrete DBN will be explained in the Background Information Chapter. However, from now on we will call dDBN as DBN.

The capability comparison of PDC and DBN can be seen in Table 1.1.

Table 1.1. *Capabilities of PDC and DBN*

	PDC	*DBN*
Linearity	Yes	Yes
Nonlinearity	No	Yes
Multivariate	Yes	Yes
Deterministic	Yes	No
Probabilistic	No	Yes

One of the main problems in brain connectivity analysis is the lack of knowledge of how human cognitive functions develop from neuronal structures. Furthermore, comparing the results of different estimator methods without a common database is impracticable. Therefore, in order to compare PDC and DBN on a common database, a synthetic data generation should be done. Since the connections of the synthetic signals are known, the performances of the methods can be compared on the same base.

The first step of synthetic data generation is to generate linear and nonlinear signals. These signals can differentiate the performances of PDC and DBN in the aspect of linearity.

The second step is adding Hemodynamic Response Function (HRF, explained in detail in Chapter 3) to generated data, in order to make the signals resembling to fMRI signals.

In the process of data generation including data generation with HRF, the length of the signals, power of white noise, and number of channels, which corresponds to network complexity, are used as variables.

All in all, the performances of the PDC and DBN is compared in respect of the following variables:

- The effect of linearity and nonlinearity of the data
- The effect of HRF
- The effect of power of white noise
- The effect of the network complexity
- The effect of the length of the data

The main aim of the study is comparing the performances of different brain effective connectivity estimator models in a common base. Furthermore, comparing a deterministic linear model with a probabilistic nonlinear model will give us the knowledge to comment on real applications on fMRI signals.

A further comparison is done on synthetic group data after the individual synthetic data. Two distinct group of data is generated, and estimator models are applied to these data. In the end, the performances of the models are analyzed statistically.

After the comparisons, methods are applied to fMRI data of dyscalculia patients. This application is the result of the study. All the comparison is done to investigate which estimator model is best for real signals. Therefore, while applying the best model to real fMRI data, we could be certain about the model. This will allow researchers to comment on the connectivity measures without concerning the effects of estimator model.

1.2. Literature Search

There is a lack of studies on comparison of brain connectivity estimator methods. The reason behind this situation is that there is no accepted common base to compare these methods on natural signals. Most of the studies that perform comparison are using synthetic data as common source. On the other hand, other studies about brain connectivity analysis only analyzes a connectivity method in a selected aspect.

1.2.1. Studies Comparing Connectivity Methods

One of the main studies that compares two connectivity method is comparing Granger Causality method with DBN and studied by Cunlu Zou and Jianfeng Feng in 2009 [2]. In this paper, the main focus is to compare Granger Causality (GC) and DBN in order to decide which one to use when they give contradictory results. They perform the comparison using both synthetic and experimental data. For synthetic data, first, the methods are tested with a fixed coefficient multivariate signal. They find that for a large data length such as 1000 points, both GC and DBN can reveal correct connectivity values. When they decrease the data length to 80, 40, and 20 both methods are starting to fail. However, DBN seems better for smaller data lengths.

Furthermore, in this paper, after comparing the methods for synthetic linear data, they create a nonlinear data set and perform comparison. For longer data lengths, they indicate that there is not much difference and both methods can find results correctly. When the data length is decreased, the methods start to perform poorly. However, for smaller data lengths, DBN performs better than Granger Causality.

Finally, they apply DBN and GC to an experimental case, where they found that DBN outperforms GC for expected connectivity results.

All in all, they compare two methods in the aspect of data length and linearity. They also verify the results with both synthetic and real data.

In another study, Laura Astolfi (2007) and her coworkers examine the performances of DTF, direct DTF, and PDC for EEG recordings [3]. In this paper, length and signal-to-noise ratio of EEG data are studied as variables affecting the reconstruction of the effective connectivity. Error in estimated connectivity patterns and reconstruction quality are evaluated. In addition to that, particular attention is paid to the ability of the different estimators to distinguish between direct and indirect causality flows. It can be concluded that PDC can show the indirect relations while DTF cannot. However, dDTF can show a small amount of indirect relation.

All in all, they compare three methods in the aspect of data length and SNR and direct vs. indirect relations. They also use synthetic and real data which is a high-resolution EEG.

Another study about comparing the brain connectivity estimator methods is held by Katarzyna J. Blinowska in 2011 [4]. In this article, bivariate measures of connectivity, such as correlation and coherence, and multivariate measures of connectivity, such as MVAR, DTF, and PDC are analyzed and compared. They use a real EEG data to compare the models, and they show that multivariate measures are outperform bivariate measures.

Matthias Winterhalder and his coworkers are also researchers who compare brain connectivity estimator models. In their paper "Comparison of linear signal processing techniques to infer directed interactions in multivariate neural systems" (2005), they compare the performances of different multivariate linear signal processing techniques in the frequency and time domain [5]. The partial cross-spectral analysis and three different quantities measuring Granger causality, i.e. a Granger causality index, partial directed coherence, and the directed transfer function are compared on the basis of different model systems. They compare them in five aspects which are direction of influences, direct versus indirect interactions, nonlinearity of data, specificity in absence of influences, and influences varying with time.

Table 1.2. *Summary of comparison of estimator methods in [4].*

Compared Aspects	*PC*	*GCI*	*DTF*	*PDC*
Direction of influence	(-)	+	+	+
Direct versus indirect interactions	+	+	-	+
Nonlinearity of data	+	-	(+)	(-)
Specificity in absence of influence	+	+	(+)	(+)
Influences varying with time	+	+	+	

where minus sign means incapability, and parenthesis means "in some cases".

In another study, “Learning effective brain connectivity with dynamic Bayesian networks”, Jagath C. Rajapakse and Juan Zhou propose to use dynamic Bayesian networks (DBN) to learn the structure of effective brain connectivity from functional MRI data [6]. Their experiments on synthetic fMRI data demonstrate that the performance of DBN is comparable to Granger causality mapping (GCM) in determining the structure of linearly connected networks. Furthermore, they study the effects of hemodynamic noise, scanner noise, inter- scan interval, and the variability of hemodynamic parameters on the derived connectivity.

In this paper, to see the capabilities of DBN they construct synthetic datasets emulating hemodynamic modulation and temporal sampling of BOLD responses in fMRI. Performances of DBN and GC are compared by using the effects of sampling step size, the amount of hemodynamic noise, and the amount of scanner noise.

Table 1.3. *Comparison Studies*

Paper	*Compared Models*	*Comparison Parameters*	*Data Type*	*Superior Model*
Zou et al, 2009	DBN, GC	Data Length, Linearity	Synthetic, real EEG	DBN
Astolfi et al, 2007	PDC, DTF, dDTF	Data length and SNR of EEG, Direct-Indirect Relations	Synthetic, real EEG	PDC
Blinowska et al, 2011	Bivariate (Coherence, Correlation) vs. Multivariate (MVAR,DTF,PDC) models	-	Real EEG	Multivariate Models
Winterhalder et al, 2005	PC, GCI, DTF, PDC	Direction of influence, Direct-Indirect Relations, Linearity, Absence of influence, Varying influence in time	Real EEG and ECG	Table 1.2
Rajapakse et al, 2007	DBN, GC	Sampling step size, the amount of hemodynamic noise, the amount of scanner noise	Synthetic fMRI	DBN

1.2.2. Studies with PDC and DBN

One of the studies about PDC is run by Daniel Yasumasa Takahashi (2007), where he and coworkers test the usability of PDC for EEG data [7]. They suggest that when not zero, PDC is asymptotically normally distributed and therefore provides means of comparing different strengths of connection between observed time series. Zero PDC indicates an absence of a direct connection between time series. They also analyze the EEG data, before and during epileptic seizure episode using PDC, and showed that PDC is a useful connectivity estimator method in a real application.

As another application of PDC, Wang et al (2016) [8], propose a new approach on the basis of PDC to detect the seizure intervals of epilepsy patients. The proposed method has achieved a good performance with the correct rate of 98.3%, the selectivity rate of 67.88%, the sensitivity rate of 91.44%, the specificity rate of 99.34%, and the average detection rate of 95.39%, which demonstrates that this method is suitable for detecting the seizure intervals of epilepsy patients. By comparing with other existing techniques, the proposed method based on PDC analysis achieves significant improvement in terms of seizure detection.

Another study which uses PDC is a Convolutional Neural Network (CNN) study. In this study Chun-Ren Phang (2019) and his coworkers contribute to the literature by proposing a multi-domain connectome CNN using PDC and granger causality estimator methods [9]. This CNN integrates the information from brain connectivity estimator methods and accurately estimates the Schizophrenic patients from healthy control group.

One of the applications of DBN is done by Warnick et al, in 2018 [10]. They propose a principled Bayesian approach to dynamic functional connectivity, which is based on the estimation of time varying networks. Their method utilizes a hidden Markov model for classification of latent cognitive states, achieving estimation of the networks in an integrated framework that borrows strength over the entire time course of the experiment. Furthermore, they assume that the graph structures, which define the

connectivity states at each time point, are related within a super-graph, to encourage the selection of the same edges among related graphs. They apply their method to simulated task-based fMRI data, where they show how their approach allows the decoupling of the task-related activations and the functional connectivity states. They also analyze data from an fMRI sensorimotor task experiment on an individual healthy subject and obtain results that support the role of particular anatomical regions in modulating interaction between executive control and attention networks.

Another application of DBN is done by Plis et al, in 2011, where they use Bayesian networks to estimate connectivity on two different modalities [11]. They analyze structures of estimated effective connectivity networks using aggregate statistics from the field of complex networks. Their study is conducted on functional MRI and magnetoencephalography data collected from the same subjects under identical paradigms. The results of the paper show some similarities but also reveal some striking differences in the conclusions one would make on the fMRI data compared with the MEG data and are strongly supportive of the use of multiple modalities in order to gain a more complete picture of how the brain is organized given the limited information one modality is able to provide.

1.2.3. Studies with Dyscalculia

One of the main studies that focuses on dyscalculia is a paper named “Dyscalculia: Characteristics, Causes, and Treatments” by Gavin R. Price and Daniel Ansari (2013) [12]. In this paper, dyscalculia disease is discussed in several topics, which are behavioral characteristics, non-numerical deficits, neural characteristics, and treatment. From these topics, the part that is important for this thesis is the neuronal characteristics. In this part, Price and Ansari says that Neuroimaging research in typically developing adults and children has identified the intraparietal sulcus as a key brain region involved in the processing of numerical magnitude representation. Thus, if primary dyscalculia is related to a core deficit in “the number sense” evident at the

brain level, then individuals with dyscalculia can be expected to show atypical activation of the IPS when processing numerical magnitude information. While only a handful of studies to date have tested this robustly, this hypothesis is gaining increasing levels of empirical support. At the functional level, Price et al. (2013) reported reduced modulation of the right IPS in Developmental Dyscalculia children during a nonsymbolic numerical comparison task (i.e., comparing which of two sets of squares was the more numerous) [12]. They reported atypical white matter tracts linking the right IPS to the right fusiform gyrus (part of the ventral visual cortex).

Another study about dyscalculia is from Ruxandra Stanescu-Cosson, in 2000 [13]. In their study with colleagues, they perform an fMRI study for people where they apply four different experiments which are digit naming, comparison, multiplication, and subtraction. Three of them excluding the digit naming, has shown most activation on Intraparietal Sulcus. They also concluded that arithmetic operations with small numbers have a greater reliance on left-lateralized regions, presumably encoding numbers in verbal format. Approximation and exact calculation with large numbers, however, put heavier emphasis on the left and right parietal cortices, which may encode numbers in a non- verbal quantity format. Subtypes of dyscalculia can be explained by lesions disproportionately affecting only one of these networks.

From another study of Brian Butterworth and his colleagues, have stated that if parietal areas, especially the IPS, fail to develop normally, there will be an impairment at the cognitive level in numerosity representation and consequential impairments for other relevant cognitive systems revealed in behavioral abnormalities [14]. The link between the occipitotemporal and parietal cortex is required for mapping number symbols (digits and number words) to numerosity representations. Prefrontal cortex supports learning new facts and procedures. The multiple levels of the theory suggest the instructional interventions on which educational scientists should focus.

In another study, A. Doyle, examines the neurological, cognitive and environmental features of dyscalculia (2016) [15]. It concludes that evidence from neuroimaging and

clinical studies in brain injury support the argument that the parietal lobe, and in particular the intraparietal sulcus (IPS) in both hemispheres, plays a dominant role in processing numerical data, particularly related to a sense of the relative size and position of numbers.

In another study about diagnosing the dyscalculia, Kaufmann et al (2012) explains the diagnostic process as in the following table [16].

Table 1.4. *Differential Diagnostic Consideration in [16]*

Differentiation	***Neurocognitive characteristics***	***Diagnostic Focus***
Isolated dyscalculia	Core deficit: concept of quantity and number (parietal regions, including IPS)	Basic numerical and arithmetical skills
Mathematical learning disability	Multiple deficits: e.g., concept of number/arithmetic + attention/working memory + visuospatial skills (frontoparietal regions)	Arithmetical and non-numerical cognitive functions
Dyscalculia with comorbid disorders	Comorbidities with dyslexia: grapheme/phoneme classification (parietal regions, including angular gyrus); with ADHD: executive functions (prefrontal regions)	Arithmetic + written language + attention/executive functions

As it can be seen from the above table, for isolated dyscalculia Intraparietal Sulcus has important roles.

1.3. Organization of the Thesis

The thesis consists of five main chapters which are Introduction, Background Information, Experiments, Results and Discussion, Applying to Real fMRI Data, and Conclusion. These five chapters thoroughly cover all the necessary information about the thesis topic “COMPARISON OF PARTIAL DIRECTED COHERENCE AND DYNAMIC BAYESIAN NETWORK APPROACH FOR BRAIN EFFECTIVE CONNECTIVITY MODELING USING FMRI”.

In Chapter 1, Introduction, the main focus is the problem definition and purpose of the study. It introduces the study and what is expected as a result from the thesis. In addition to that, a literature search is presented. In this search, related works with this thesis are analyzed and the results are discussed.

In Chapter 2, Background Information, necessary information for reader to understand the work done in this thesis is covered. In this section, there are some background information about brain physics, dyscalculia, brain imaging techniques including EEG and fMRI. In addition to that some information about brain connectivity and connectivity estimator methods are presented to the reader.

The main part of the thesis is Chapter 3, where the experimental data generation process, and results are discussed in detail. In this chapter, firstly synthetic data generation part of the thesis is explained. Equations that are used in generation of nonlinear and linear data are also discussed. Secondly, implementing PDC and DBN methods in Matlab is mentioned. Thirdly, application of PDC and DBN codes to synthetic data is deeply analyzed for every type of synthetic data. After that, by comparing the results of application of PDC and DBN, the two method are compared with each other. After that, a synthetic group data are generated and models are applied to these groups and are tested wheter they are able to distinguish two groups or not. In this application, DTF model is also used.

In Chapter 4, DBN, PDC, and DTF are applied to a real life fMRI data of dyscalculic patients and control group, and the results are commented according to the dyscalculia

physiology in the brain. The methods, data collection of fMRI, and statistical analysis of the results are discussed in this chapter.

In Chapter 5, the last part of the thesis, conclusion of all the processes is done and future work is discussed.

CHAPTER 2

BACKGROUND INFORMATION

2.1. Brain Physics

Brain is the main control mechanism of the body. It gets information from the body and sends necessary information to body. To maintain this communication, a wiring system is distributed throughout the body.

The nervous system of a human is built of two cell types which are glial cells and neurons [17].

Brain is based on the cell networks that are created by neurons. Neurons are connected with glial cells, which provide support functions for neural networks [18]. A neuron can be seen in Figure 2.1.

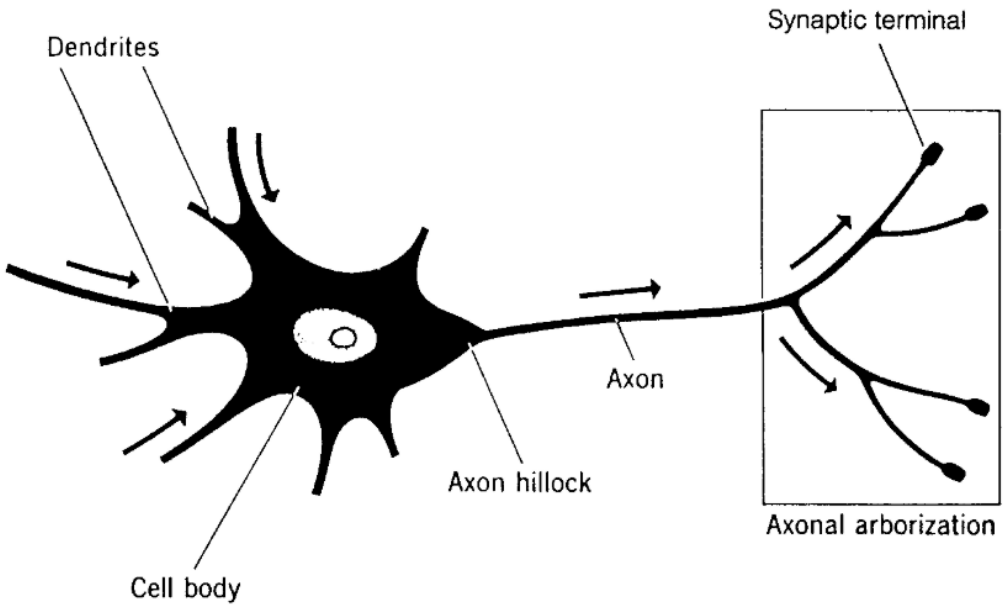

Figure 2.1. A Basic Neuron. *[18]*

The important areas of the brain for this thesis can be listed and explained as follows:

- Hippocampus: The hippocampus plays significant roles in the integration of information from short-term memory to long-term memory.
- Intraparietal Sulcus (IPS): Its main functions are associated with perceptual-motor coordination and visual attention, which allows for visually guided pointing, grasping, and object manipulation that can produce a desired effect. It is thought to play a role in other functions, including processing symbolic numerical information, visuospatial working memory and interpreting the intent of others.
- Anterior Cingulate Cortex: It appears to play a role in a wide variety of autonomic functions, such as regulating blood pressure and heart rate. It is also involved in certain higher-level functions, such as attention allocation, reward anticipation, decision-making, ethics and morality, impulse control (e.g. performance monitoring and error detection), and emotion.
- Occipital Lobe: It is one of the brains main lobes, and it is associated with the visual information.

2.2. Dyscalculia

"Developmental Dyscalculia (DD) is a learning disorder affecting the ability to acquire school-level arithmetic skills, affecting approximately 3-6% of individuals" [12]. The most consistently observed behavioral hallmark of DD is impaired arithmetic fact retrieval. Dyscalculia is a specific learning difficulty that has also been referred to as 'number blindness', in much the same way as dyslexia was once described as 'word blindness' [15].

2.3. Brain Imaging Techniques

The first point of analyzing the brain is to get the electrical and biological activities inside it and convert all this information into some format that can be analyzed. Before we find out and interpret the contextual connectivity of the human brain, we should solve the physical connections between the brain regions. Brain imaging technologies

help scientists to create a non-invasive way to model the physical and neuronal activities inside the brain and provide an opportunity to solve the connectivity between the brain regions. There are lots of brain imaging techniques such as Positron Emission Tomography (PET), Near Infrared Spectroscopy (NIRS), and Magnetoencephalogram (MEG). However, in this thesis, the main focus will be on functional Magnetic Resonance Imaging (fMRI) and Electroencephalogram (EEG) because they are the main techniques that are used in the connectivity estimation analysis.

2.3.1. Functional Magnetic Resonance Imaging (fMRI)

Functional magnetic resonance imaging (fMRI) is an imaging technique that aims to detect the dynamic patterns of activity in the human brain [19].

fMRI uses the blood-oxygen-level-dependent (BOLD) contrast structure in order to measure the neuronal activity inside the brain [20]. When an activity occurs in some part of the brain the blood flow increases on that specific section. Hence, the increase of oxyhemoglobin in blood on this area causes a change on the proton signal from water molecules near the blood vessel for gradient-echo MRI, which creates the BOLD contrast [21].

Figure 2.2 demonstrates the measured reactions in the motor area of the human cerebrum dependent on MR signals which are sensitive to blood stream and blood oxygenation [19].

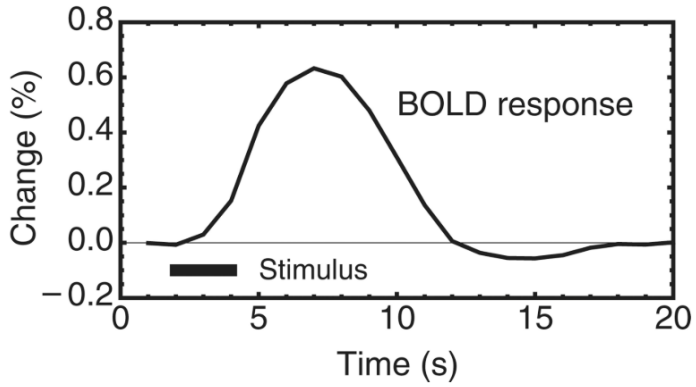

Figure 2.2. BOLD Response. *[19].*

2.3.2. Electroencephalogram (EEG)

Electroencephalography (EEG) is the non-invasive measurement of the brain's electric fields [22]. "EEG is a record of the electric signal generated by the cooperative action of brain cells, or more precisely, the time course of extracellular field potentials generated by their synchronous action. Electroencephalogram derives from the Greek words enkephalo (brain) and graphein (to write). EEG can be measured by means of electrodes placed on the scalp or directly on the cortex. In the latter case, it is sometimes called electrocorticogram (ECoG)" [23]. Due to capability to reflect both the normal and abnormal electrical activity of the brain, EEG has been found to be a very powerful tool in the field of neurology and clinical neurophysiology [24].

The following rhythms have been distinguished in EEG: delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (above 30 Hz). Gamma components are difficult to record by scalp electrodes and their frequency does not exceed 45 Hz; in ECoG components, up to 100 Hz, or even higher, may be registered [23]. In 1958, International Federation in Electroencephalography and Clinical Neurophysiology adopted standardization for electrode placement called 10-20 electrode placement system. This system standardized physical placement and designations of electrodes on the scalp [24].

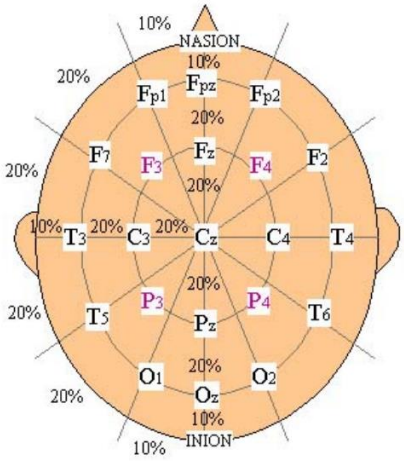

Figure 2.3. 10-20 Placement of EEG electrodes. The figure was taken from *[24]*

2.4. Brain Connectivity

Brain connectivity means interactions among the different neuronal units in cerebrum. Sporns (2007) suggests that brain connectivity means a pattern of anatomical connectivity, functional connectivity, and effective connectivity between different parts of the brain [1]. Brain connectivity researches are attempting to uncover how neurons, brain parts or entire brain communicates with each other when there is presence or absence of a stimulus.

2.4.1. Structural Connectivity

Structural connectivity gives information about which brain part is anatomically associated with another. Some brain imaging techniques, such as DTI, are used to extract the structural connectivity. Voxel based analyses such as diffusion, morphometry, and tractography analyses are the most generally utilized strategies to uncover structural connectivity from imaging data.

Structural networks might be valuable to identify flawed connections between brain parts and may likewise be utilized to help determination of some brain disorders.

2.4.2. Effective Connectivity

Effective connectivity is outlined as the causal relationships that brain units employ over another [25]. Since it focuses on the causal effects that a brain region causes on another part, effective connectivity can be seen in multivariate time dependent systems. Brain executes billions of actions per second. In order to catch the information from brain activity, the data that will be used for effective connectivity analyses should be with high temporal resolution.

2.4.3. Functional Connectivity

Functional connectivity studies are interested in linear and nonlinear correlations among regions without concerning if there are physical paths between these regions or not. In this connectivity type, there is no comment on the direction of the information flow as in the structural connectivity.

2.5. Connectivity Adjacency Matrix

Connectivity adjacency matrix is like a table of connection values across different brain regions. Generally, columns of the connectivity matrix represent the connections as “from” and rows of the connectivity matrix represent the connections as “to” the corresponding nodes. The connection values inside the matrix are usually normalized to 0 and 1. “0” means lack of connection while “1” means there is a connection.

A connectivity matrix can be directed and undirected, or weighted and unweighted. Weighting is about the strength of the connection. An unweighted connection only shows if there is a connection or not. On the other hand, a weighted connection can be assigned numerical values between 0 and 1 which shows the connection strength.

An undirected connectivity matrix doesn’t express the direction information of the connectivity. Therefore, these type of connection matrices are symmetric. On the contrary, a directed connectivity matrix contains the direction information between nodes. Thus, directed matrices generally are not symmetric.

TO

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	0	0	0	0	1	0	0	0	0	0	1	0	1	0
2	0	0	0	0	1	1	1	1	0	0	0	0	1	0
3	0	0	0	0	0	1	0	1	1	0	1	0	1	0
4	0	0	1	0	0	0	0	0	0	0	1	0	0	0
5	0	1	1	0	1	0	0	0	1	0	0	0	1	0
6	0	0	0	1	0	0	0	0	1	0	0	0	1	0
7	0	0	0	0	0	0	0	0	1	1	0	0	0	1
8	0	0	0	0	0	0	1	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	1	0	0	0	0	0	1
10	0	0	0	0	0	0	1	0	0	1	0	1	0	1
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	1	0	1	0	0
13	0	0	1	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	1	0	0	1	0	1

FROM

Figure 2.4. An unweighted connectivity matrix.

2.6. Brain Connectivity Estimator Methods

2.6.1. Granger Causality (GC)

Granger Causality was first made known in 1969 with a purpose of explaining the causal relations between econometric models [26]. GC suggests that "if a signal Y(t) contains information in their past values to predict the behaviour of series X(t), then series Y(t) are said to cause X(t)" [4].

Let X(t) be a signal that can be calculated from its p discrete past values with using a prediction error e1, then X(t):

$$X(t) = \sum_{j=1}^{p} \big(A_{11}(j)X(t-j)\big) + e_1(j) \tag{1}$$

Also using previous values of signal Y with using a prediction error e2, the signal X(t) can also be expressed as:

$$X(t) = \sum_{j=1}^{p} \big(A_{11}(j)X(t-j)\big) + \sum_{j=1}^{p} \big(A_{12}(j)Y(t-j)\big) + e_2(j) \tag{2}$$

Using the GC principles, if the condition of var(e1) > var(e2) is satisfied, then the series Y(t) cause series X(t).

As a bivariate estimator which operates in time domain, Granger causality index (GCI) can be calculated as:

$$GCI_{1\to 2} = \ln\left(\frac{var(e_1)}{var(e_2)}\right) \tag{3}$$

where $GCI_{1\to 2}$ represents the Granger Causality index from first to second signal. Increase in the value of GCI means increase in the causality.

2.6.2. Multivariate Autoregressive Model (MVAR)

MVAR models are generalized form of GC for more than two time series, and first introduced in 1980 [27]. It suggests that a multivariate signal with n number of channels is represented as:

$$\mathbf{X}(t) = (\mathbf{X_1}(t), \mathbf{X_2}(t), \ldots, \mathbf{X_n}(t))^T \quad (4)$$

Then, the construction of the X(t) (combination of all channels) can be expressed as follows:

$$\boldsymbol{X}(t) = \sum_{j=1}^{p} (\mathbf{A}(j)\mathbf{X}(t-j)) + \boldsymbol{e}(t) \quad (5)$$

where p is the model order which represents how many past values of x is effective when generating the x(t), and A(j) is the coefficient matrix containing the j^{th} order autoregressive model coefficients.

In order to find the order of the model "p", several methods are developed. One of the methods that suggests the value p is Akaike Information Criteria (AIC) [28] which is used to calculate p for long multivariate signals [3] [29]. According to AIC, the following formula is used to find model order "p":

$$\mathrm{AIC}(p) = \ln\left|\tilde{\Sigma}(p)\right| + \frac{2}{\hat{T}} pM^2 \quad (6)$$

where $\tilde{\Sigma}$ is the estimated noise covariance of MVAR model for the value p. m is the number of channels and $\hat{T}$ is the number of data samples to fit the model. The model order "p" is selected such as AIC(p) takes minimum value:

$$p_{selcted} = \arg\min_{p} AIC(p) \quad (7)$$

After the selection of p value, the next step is MVAR model fitting where the MVAR coefficients are calculated. However, before the MVAR model fitting is applied, the following inequality;

$$n > m^2 p \tag{8}$$

should be satisfied. In this inequality, n is the number of samples in a channel. Furthermore, like other inequalities in order to be sure about having good results, [30] says that the factor of 10 should be added to the right side of the inequality:

$$n > 10 * (m^2 p) \tag{9}$$

After we ensure about the model is consistent with the MVAR, the MVAR model fitting could be applied. One of the methods that calculates the coefficients of the MVAR model is Yule-Walker Equations which is the model that is used in this study. These Yule-Walker equations estimate the MVAR coefficients Aij(t) in time domain.

The mathematical model of Yule-Walker model can be explained as follows:

Consider the general AR(p)

$$x_{i+1} = \phi_1 x_i + \phi_2 x_{i-1} + \cdots + \phi_p x_{i-p+1} + \xi_{i+1}. \tag{10}$$

for order (p) 1 and k correlation r;

$$\begin{aligned} r_1 &= \sum_{j=1}^{p} \phi_j r_{j-1} \\ r_k &= \sum_{j=1}^{p} \phi_j r_{j-k}. \end{aligned} \tag{11}$$

After that all the equations can be written as follows:

$$\begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_{p-1} \\ r_p \end{pmatrix} = \begin{pmatrix} r_o & r_1 & r_2 & \cdots & r_{p-2} & r_{p-1} \\ r_1 & r_o & r_1 & \cdots & r_{p-3} & r_{p-2} \\ & \vdots & & & \vdots & \\ r_{p-2} & r_{p-3} & r_{p-4} & \cdots & r_o & r_1 \\ r_{p-1} & r_{p-2} & r_{p-3} & \cdots & r_1 & r_o \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_{p-1} \\ \phi_p \end{pmatrix} \tag{12}$$

The autocorrelation r0 is equal to 1. Therefore, equation becomes:

$$\underbrace{\begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_{p-1} \\ r_p \end{pmatrix}}_{\mathbf{r}} = \underbrace{\begin{pmatrix} 1 & r_1 & r_2 & \cdots & r_{p-2} & r_{p-1} \\ r_1 & 1 & r_1 & \cdots & r_{p-3} & r_{p-2} \\ & \vdots & & & \vdots & \\ r_{p-2} & r_{p-3} & r_{p-4} & \cdots & 1 & r_1 \\ r_{p-1} & r_{p-2} & r_{p-3} & \cdots & r_1 & 1 \end{pmatrix}}_{\mathbf{R}} \underbrace{\begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_{p-1} \\ \phi_p \end{pmatrix}}_{\mathbf{\Phi}} \tag{13}$$

From there, coefficients can be found as follows:

$$\hat{\mathbf{\Phi}} = \mathbf{R}^{-1}\mathbf{r}. \tag{14}$$

2.6.3. Partial Directed Coherence (PDC)

PDC is first introduced by Baccala et al. in 2001 [31] and it is based on the MVAR model and uses its coefficients to calculate connectivity values. Astolfi et al. (2007) suggests that PDC reconstruct the connectivity information better than the other estimator methods by separating direct information within the channels from indirect information [3].

Baccala and Sameshima (2001) [31] suggests the frequency domain representation of $A_{ij}(t)$ as $A_{ij}(f)$ where i is the row and j is the column index and computed as follows for each discrete frequency values according to Nyquist Theorem [32]:

$$\mathbf{A_{ij}}(\mathrm{f}) = \begin{cases} 1 - \sum_{r=1}^{p} A_r(i,j)e^{-i2\pi fr}, if\ i = j \\ -\sum_{r=1}^{p} A_r(i,j)e^{-i2\pi fr}, otherwise \end{cases} \tag{15}$$

While some connectivity estimator models like PDC uses frequency domain coefficients $\mathbf{A_{ij}}(\mathrm{f})$ to estimate connections, other estimator models like DTF uses the transfer function of the system, H:

$$\mathrm{H} = \mathbf{A^{-1}}(\mathrm{f}) \tag{16}$$

Baccala (2001) [31] formulate the PDC using MVAR model coefficient matrix $A(f)$ as follows:

$$PDC_{ij}(f) = \frac{A_{ij}(f)}{\sqrt{A_j^*(f)A_j(f)}} \tag{17}$$

The value of $PDC_{ij}(f)$ gives information about the connection from the channel j to i. If the value is close to zero, it means less direct information flow from j to i. On the contrary, if it is close to one, it means a strong direct connection.

As it can be seen from the calculations, PDC operation is not performed in time domain, it is performed in frequency domain. However, it has no direct relation with power spectrum, because no autocorrelation or variance values were used in calculations. PDC only calculates the direct relationships between nodes j and i, without considering other nodes [3].

The main formula of PDC is based on frequency domain, which makes PDC variant to the frequency values.

PDC can also be used for data collected from fMRI and EEG thanks to its multivariate approach. It can analyze and reconstruct the connectivity values of a multichannel dataset [33].

There is a variation of PDC called partial coherence (PC). In order to calculate the PC, the power spectra S(f) should be calculated. S(f) can be formulated as follows:

$$S(f) = H(f)var\big(e(f)\big)H^{*}(f)$$

where,

$$\mathrm{H(f)} = \mathbf{A^{-1}}\mathrm{(f)},$$

$var\big(e(f)\big)$, variance calculated from the FT of noise matrix e(t) (18)

$*$, $complex\ conjugate\ transpose\ operator$

Using the power spectra, coherence between the signals of channel i and j can be calculated:

$$Coh_{ij}(f) = \frac{|S_{ij}(f)|^2}{S_{ii}(f)S_{jj}(f)} \tag{19}$$

Coherence value between channels i and j gives a symmetric coherence matrix. The calculation of PC also gives a symmetric matrix, which means Partial Coherence has no information about the direction.

$$PC_{ij}(f) = \frac{Coh_{ij}(f)}{\sqrt{Coh_{ii}(f)Coh_{jj}(f)}} \tag{20}$$

2.6.4. Directed Transfer Function (DTF)

DTF is introduced by Kaminski and Blinowska in 1991 [34] and the underlying method is Granger Causality.

The unnormalized DTF is simply calculated from the transfer function of the MVAR model which is calculated in Equation (16):

$$\theta_{ij}(f) = H_{ij}(f) \tag{21}$$

In order to get the normalized value of DTF, the following formula is used:

$$DTF_{ij}(f) = \frac{H_{ij}(f)}{\sqrt{\sum_{j=1}^{N} \left|H_{ij}(f)\right|^2}}$$

where,

$$H_{ij}(f), \qquad Unnormalized\ DTF\ of\ a\ channel \tag{22}$$

$$\sqrt{\sum_{j=1}^{N} \left|H_{ij}(f)\right|^2},$$

$$the\ square\ mean\ of\ DTF's\ for\ outgoing\ connectivities$$

In order to state in a different way, it can be said that the nDTF is the ratio of the DTF between two channels to one channel total DTF.

2.6.5. Dynamic Bayesian Networks (DBN)

DBN is first introduced in 1990s. The main purpose of these studies was forecasting. Dagum et al. [35] has used Bayesian Algorithms in these studies (BN). DBN calculations are mostly based on probabilistic approach. Every variable is named as node, and connectivity between regions is expressed as connection between the nodes. The connection between nodes is defined as conditional probability. Conditional probability distribution can be continuous or discrete. In summary, DBN is a probabilistic method which is not linear. In discrete DBN, which is used in this thesis, conditional probabilities can be shown as table and every element of the table is a parameter of the model.

Effective connectivity between two different channels are calculated by assigning a probability value between 0 and 1. This assignment means that there is a probabilistic value of direct causal connection between each channel. Therefore, it can be easily said that DBN algorithms can be used to build probabilistic temporal networks.

The first application of DBN on EEG data is performed by Smith et al. in 2006 [36]. After the first application, many researches were performed on neuroscience using DBN algorithms [37]. DBN maps represent discrete time stochastic processes for each channel time series data X(t):

$$X(t) = (X(t)^1, X(t)^2, \dots, X(t)^n) \tag{23}$$

In DBN, a discrete timestamp is introduced, and the same local model is repeated for each unit of time. That local model is a section of the network called a time slice and represents a snapshot of the underlying evolving temporal process. The nodes within time slice t can be connected to other nodes within the same slice. Also, time slices are interconnected through temporal or transition arcs that specify how variables change from one time point to another. Temporal arcs only flow forward in time, since the state of a variable at one time point is determined by the states of a set of variables at previous time points. A prior BN specifies the initial conditions. In dynamic BNs, the structures of the time slices are identical, and the conditional probabilities are also

identical over time. Therefore, dynamic BNs are time-invariant models, and dynamic only means that they can model dynamic systems. For inference purposes, the structure of a dynamic BN is obtained by unrolling the transition network over all consecutive times [37].

To perform DBN computations, some assumptions should be made:

- First assumption is that the data is stationary which means that causal relations are time invariant.
- Process is assumed to be first order Markovian transition model i.e. [37]:

$$p(X(t)|X(t-1),\dots,X(1)) = p(X(t)|X(t-1)) \quad (24)$$

Higher order and non-stationary Markov models allow more complex temporal processes. However, such complex models pose obvious challenges to structure and parameter estimation.

A discrete dynamic Bayesian network (dDBN) is a specialization of a DBN that models temporal processes [38]. Its graphical topology is divided into columns of nodes such that each column represents a time frame. Each random variable is represented by one node in each of the columns. Links are allowed to connect nodes between columns, provided the link points forward in time. Ideally, there would be one column for every time frame and links could connect nodes separated by arbitrary time steps (including nodes in the same time frame). However, such dDBNs are intractably large and require far more data and computational resources to learn than is likely to be available.

The parameters for the dDBN are commonly computed with maximum likelihood estimates. Given a fully parameterized dDBN, B, the posterior likelihood given a set of data, D, can be computed [38]:

$$\ell(B \mid D) \propto P(B)P(D \mid B) = P(B)\prod_{j=1}^{m}\prod_{t=1}^{T_j}\prod_{i=1}^{n} P\left(X_i = D_{i,t}^{j} \mid \mathrm{Pa}(X_i) = D_{Pa(X_i),t-1}^{j}\right) \quad (25)$$

CHAPTER 3

EXPERIMENTS, RESULTS AND DISCUSSION

3.1. Synthetic Data Generation

Generating the synthetic data is the main part of the thesis. Firstly, since the base connections of the synthetic data is known beforehand, it is easy to test the efficiency of PDC and DBN methods easily. Unlike real fMRI or EEG signals, where the brain connectivity is unknown, applying these methods to synthetic data allows us to compare them objectively. Furthermore, using synthetic data, one can test the performances of PDC and DBN in many aspects since the data can be manipulated easily to see the effects of different variables.

The synthetic data generation can be divided into two different sections. First one is the raw data generation. In this part, the data is generated directly from the connectivity coefficients between channels and past values of the channels based on autoregressive model. Only a white noise is added. The second part is adding BOLD effect to the raw data. This effect can be applied by convolving Hemodynamic Response Function (HRF) with the raw data in order to generate realistic synthetic fMRI data.

The variables that should be controlled are the linearity of the signals, complexity of the network (number of channels), data length, and effect of the power of white noise that used to create raw data.

Order of the signals, i.e., how many past values effect the current signal value, is kept as “1”, since unlike PDC, DBN complexity increases and capability of extracting the connectivity becomes a very hard challenge for connections of a signal that has an order higher than “1”. Therefore, the comparison will be done with the order value as 1 [37].

The communication among neuronal populations, reflected by transient synchronous activity, is the mechanism underlying the information processing in the brain. Although it is widely assumed that the interactions among those populations are highly nonlinear, the amount of nonlinear information transmission and its functional roles are not clear [39]. Therefore, in order to simulate effects of linear and nonlinear signal transmissions, both of them are used to generate synthetic data. This will also help us to distinguish the performances of brain connectivity estimator models for two types of data.

The data generation is done in MATLAB.

3.1.1. Raw Data Generation

3.1.1.1. Linear Data Generation

In order to generate the linear synthetic dataset, the following equations are used:

$$\begin{aligned}
x_1(t+1) &= a_{11}x_1(t) + a_{12}x_2(t) + a_{13}x_3(t) + a_{14}x_4(t) + a_{15}x_5(t) + a_{16}x_6(t) + \mathcal{E}_1 \\
x_2(t+1) &= a_{21}x_1(t) + a_{22}x_2(t) + a_{23}x_3(t) + a_{24}x_4(t) + a_{25}x_5(t) + a_{26}x_6(t) + \mathcal{E}_2 \\
x_3(t+1) &= a_{31}x_1(t) + a_{32}x_2(t) + a_{33}x_3(t) + a_{34}x_4(t) + a_{35}x_5(t) + a_{36}x_6(t) + \mathcal{E}_3 \\
x_4(t+1) &= a_{41}x_1(t) + a_{42}x_2(t) + a_{43}x_3(t) + a_{44}x_4(t) + a_{45}x_5(t) + a_{46}x_6(t) + \mathcal{E}_4 \\
x_5(t+1) &= a_{51}x_1(t) + a_{52}x_2(t) + a_{53}x_3(t) + a_{54}x_4(t) + a_{55}x_5(t) + a_{56}x_6(t) + \mathcal{E}_5 \\
x_6(t+1) &= a_{61}x_1(t) + a_{62}x_2(t) + a_{63}x_3(t) + a_{64}x_4(t) + a_{65}x_5(t) + a_{66}x_6(t) + \mathcal{E}_6
\end{aligned} \tag{26}$$

This equation set is for 6-Channel multivariate signal, where $\mathcal{E}$ is the white noise and a_{ij} is the connection coefficient for channel j to channel i. Order of the signal is taken as 1.

While generating the signal, the coefficient matrix A is the main factor that decides the connectivity.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} \end{bmatrix} \quad (27)$$

Figure 3.1 represents the effect of coefficients. This figure shows only the effect of channel one to others in order to show an example of the influence of coefficients.

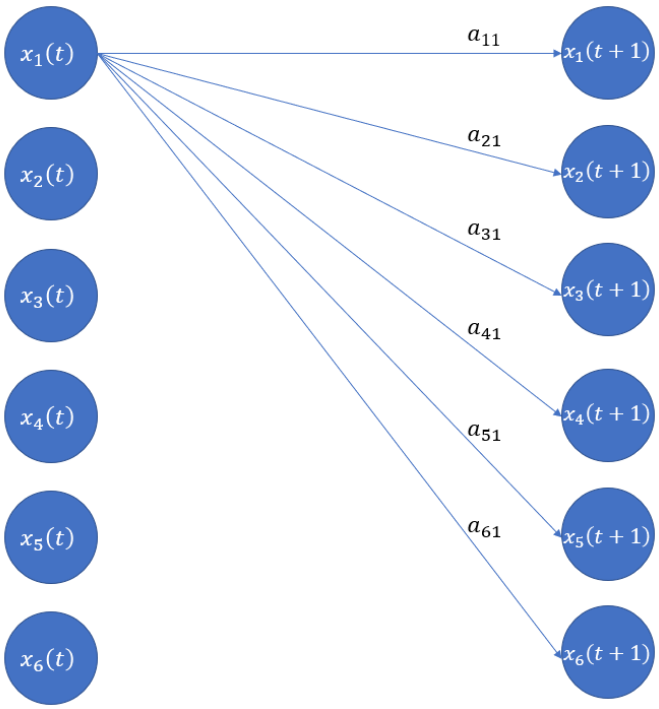

Figure 3.1. Representation of Coefficients on a Graph

In order to distinguish the channels and measure the efficiency of the methods for a general case, the following matrix form is used for coefficients.

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (28)$$

Furthermore, the coefficients are uniformly distributed random variables between the interval of [-1, -0.5] U [0.5, 1]. In this way, every channel is connected to the different number of channels.

An example coefficient matrix that is used to generate a synthetic data can be seen as follows:

$$A = \begin{bmatrix} 0.85 & 0.89 & -0.62 & 0.57 & 0.57 & -0.92 \\ 1.0 & 0.88 & 0.99 & 0.92 & 0.63 & 0 \\ -0.50 & -0.91 & -0.95 & -0.96 & 0 & 0 \\ -0.67 & 0.85 & 0.91 & 0 & 0 & 0 \\ -0.83 & 0.57 & 0 & 0 & 0 & 0 \\ 0.79 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (29)$$

Other than that, the power of white noise that is used in the generation of linear data is kept as variable for only 6 channel data in order to see the effects of the noise. Total of 4 different noise levels are used.

A generated 6 channel linear data can be examined in Figure 3.2.

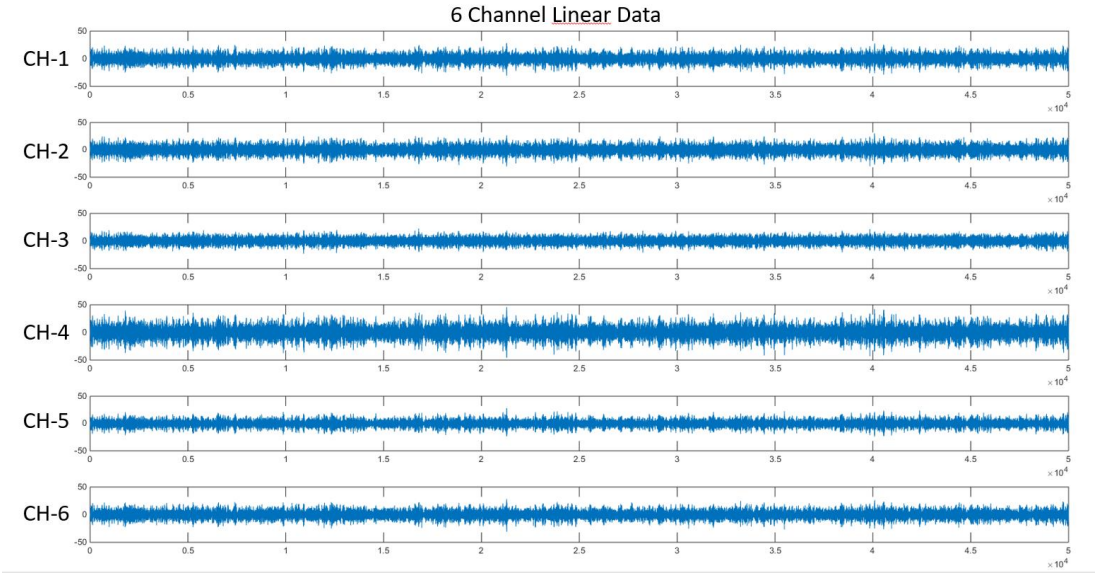

Figure 3.2. 6-Channel Linear Data

The length of the data is 50000 sample. It corresponds to time (signal duration) for real data. While applying the methods, the data will be used as different time interval windows in order to test the effect of data length to retrieve the connectivity information.

All in all, as linear data, 10 different unique synthetic signals are generated for 2, 3, 4, 5, and 6 channel signals which makes 50 datasets for further analyses. For 6 channel data, additionally 6 different noise values were applied, which adds 30 signals to be considered.

Table 3.1. *Variables for Linear Raw Data Generation*

Variables	*Values*
Number of Channels	2,3,4,5,6
Coefficient Matrix	[-1, -0.5] U [0.5, 1] X Equation (28)
White Noise Power	1dB, 2dB, 5dB, 10dB, 20dB, 50dB, 100dB (for 6 Ch only)

3.1.1.2. Nonlinear Data Generation

In order to generate the nonlinear synthetic dataset, the following equations are used:

$$X(t+1) = A \,.\, e^{-\frac{X(t)^2}{2}} + \mathcal{E} \tag{30}$$

where,

$$X(t+1) = \begin{bmatrix} x_1(t+1) \\ x_2(t+1) \\ x_3(t+1) \\ x_4(t+1) \\ x_5(t+1) \\ x_6(t+1) \end{bmatrix} \tag{31}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} \end{bmatrix} \tag{32}$$

This equation set is for 6-Channel nonlinear multivariate signal where $\mathcal{E}$ is the white noise and a_{ij} is the connection coefficient for channel j to channel i. Order of the signal is taken as 1. The basics of nonlinear data generation is taken from [2].

Furthermore, the coefficients are uniformly distributed random variables between the interval of [-1, -0.5] U [0.5, 1] in order to point out the connectivity between the channels.

A generated 6 channel nonlinear data can be examined in Figure 3.3.

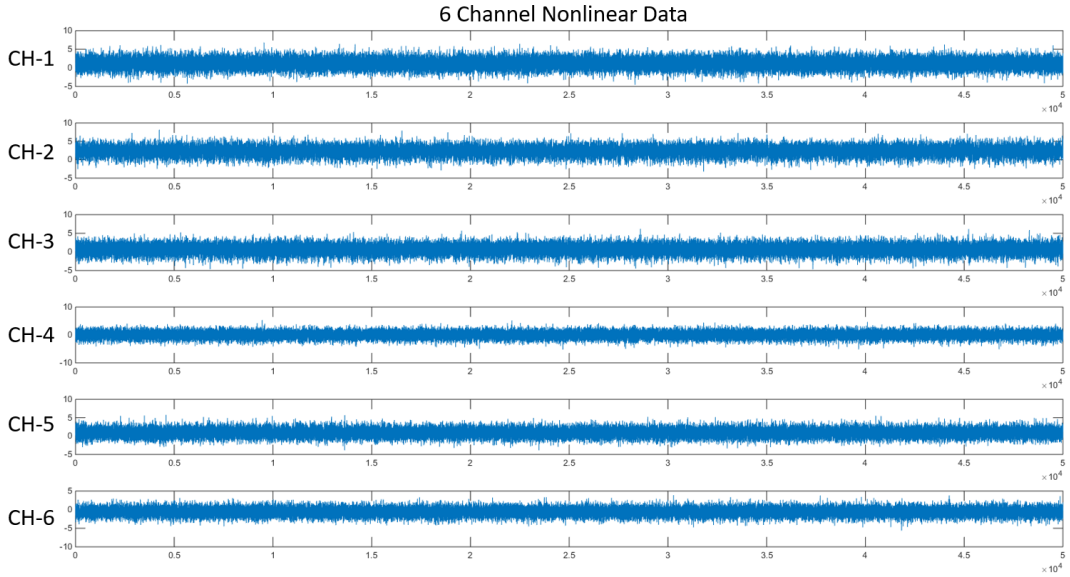

Figure 3.3. 6-Channel Nonlinear Data

As in the linear case the length of the data is 50000 sample.

Table 3.2. *Variables for Nonlinear Raw Data Generation*

Variables	*Values*
Number of Channels	2,3,4,5,6
Coefficient Matrix	[-1, -0.5] U [0.5, 1] X Equation (28)
White Noise Power	1dB, 2dB, 5dB, 10dB, 20dB, 50dB, 100dB (for 6 Ch only)

The Table 3.3 shows number of all the generated raw data.

Table 3.3. *Number of Generated Raw Data*

	Raw Data									
Linearity	Linear					Nonlinear				
Number of Channel	2	3	4	5	6	2	3	4	5	6
# of Data Generated	10	10	10	10	70	10	10	10	10	70
Total # of Data Generated	110					110				
	220									

3.1.2. Linearity Check of the Generated Data

In order the check the linearity or nonlinearity of the synthetic data, the MATLAB code set is used by Habibnia and his co-workers named A Nonlinearity Test for Principal Component Analysis [40]. In this test, results indicate whether the data shows linear characteristics or nonlinear characteristics.

This MATLAB code is developed to test whether the underlying structure within the recorded data is linear or nonlinear. The nonlinearity measure is introduced in Kruger et al (2005) [41]. The measure relies on the division of the recorded range of process operation, or operational range, into disjunct regions. A correlation matrix is then obtained using the data of one of these regions. This is followed by computing thresholds for each matrix element on the basis of the confidence limits for computing the mean and variance of each process variable. Then, the maximum and minimum sum of discarded eigenvalues, or the accuracy bounds, are calculated using the fact that the matrix elements are within known thresholds. Finally, the sum of discarded eigenvalues is obtained for the correlation matrix of each region, noting that the mean and variance of each process variable are obtained for the region for which the accuracy bounds are obtained.

If the sum of discarded eigenvalues for the PCA model of each region is inside the accuracy bounds, the process is said to be linear. Conversely, if at least one of these sums is outside, the process must be assumed to be nonlinear.

The result of the code for linear data can be seen in Figure 3.4.

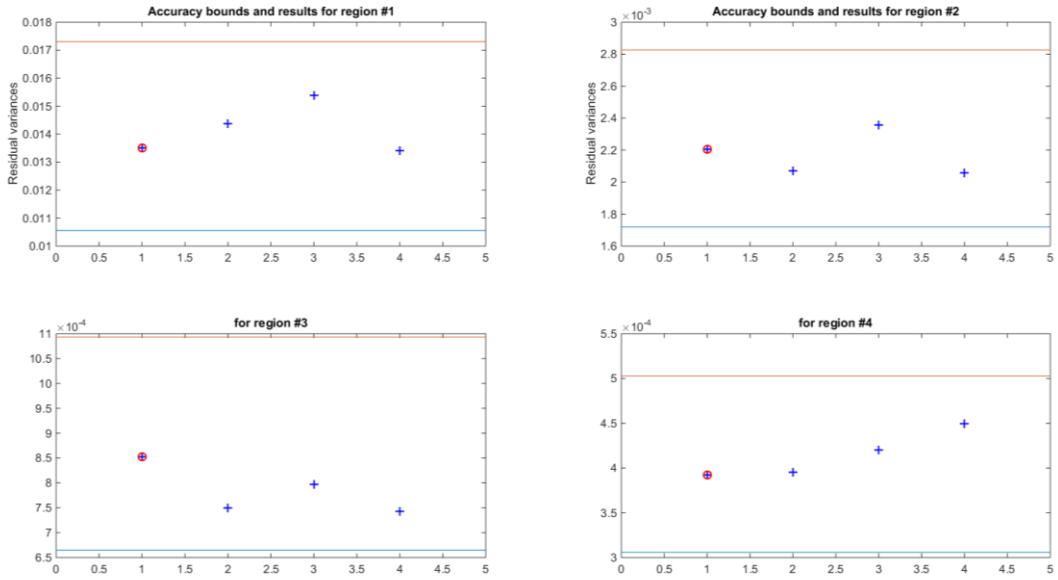

Figure 3.4. Applying Linearity test to the linear synthetic data

This result is shown four plots of benchmarking the residual variances against accuracy bounds of each disjunct region. These plots yield that no violation of the accuracy bounds arise, which leads to the acceptance of the hypothesis that the underlying relationship between the two series is linear.

The result of the code for nonlinear data can be seen in Figure 3.5.

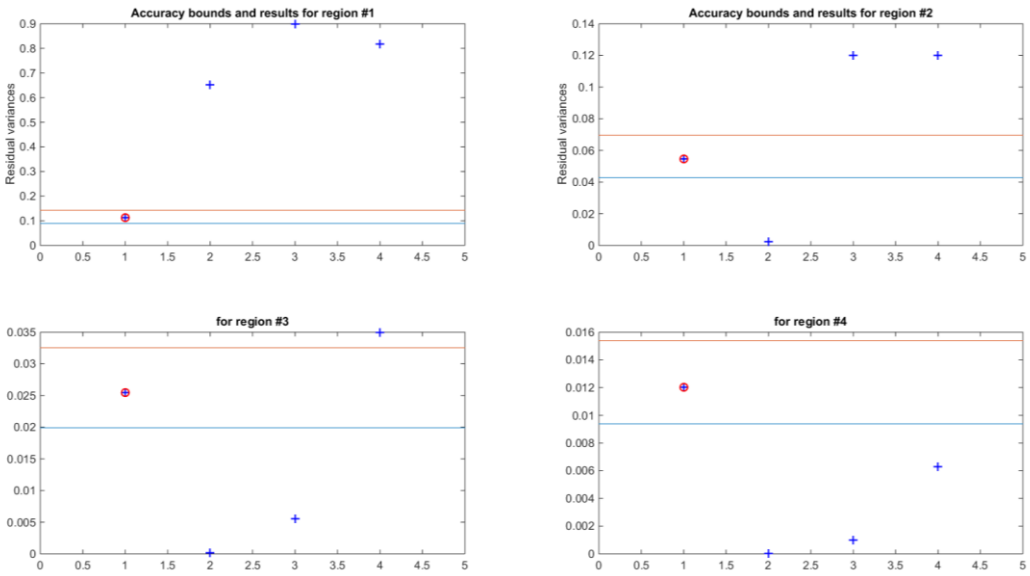

Figure 3.5. Applying Linearity test to the nonlinear synthetic data

The violation here shows that the data is nonlinear.

3.1.3. HRF Application

The BOLD signal time-series is the result of a series of neuronal and vascular events that produce a measurable change in the blood hemoglobin concentration. It is therefore an indirect and noisy observation of the neuronal activity as during neuronal activation local vessels are dilated to increase the blood flow and with it, oxygen and glucose delivery [42].

The simulation of BOLD signals allows experimental control over neuronal and hemodynamic parameters, and this has been achieved mainly by convolving the data with a canonical hemodynamic response function (HRF).

To generate the HRF, following equation is used. This equation is taken from [6].

$$f(t) = \frac{1}{\Gamma\left(\frac{\tau_1}{\tau_3}\right)}\left(\frac{\delta}{\tau_3}\right)^{\frac{\tau_1}{\tau_3}} w(t)^{\frac{\tau_1}{\tau_3}-1}\, e^{-\frac{\delta}{\tau_3}w(t)} - \frac{1}{\tau_5\Gamma\left(\frac{\tau_2}{\tau_4}\right)}\left(\frac{\delta}{\tau_4}\right)^{\frac{\tau_2}{\tau_4}} w(t)^{\frac{\tau_2}{\tau_4}-1}\, e^{-\frac{\delta}{\tau_4}w(t)} \tag{33}$$

The parameters are selected as:

$$\delta = \frac{RT}{16},$$

$$RT = 0.1\ s, \qquad Repeatition\ Time\ of\ Stimulus$$

$$\tau_1 = 6\ s, \qquad delay\ of\ response$$

$$\tau_2 = 16\ s, \qquad delay\ of\ undershoot$$

$$\tau_3 = 1\ s, \qquad dispersion\ of\ response$$

$$\tau_4 = 1\ s, \qquad dispersion\ of\ undershoot$$

$$\tau_5 = 6\ s, \qquad ratio\ of response\ to\ undershoot$$

$$\tau_6 = 32\ s, \qquad length\ of\ the\ kernel$$

$$\Gamma(.), \qquad gamma\ function$$

$$w(t) \epsilon \left\{0, \frac{1}{\delta}, \ldots, \frac{\tau_6}{\delta}\right\}$$

(34)

for generating a general case HRF data as in [6].

The HRF generated for this case can be seen in Figure 3.6.

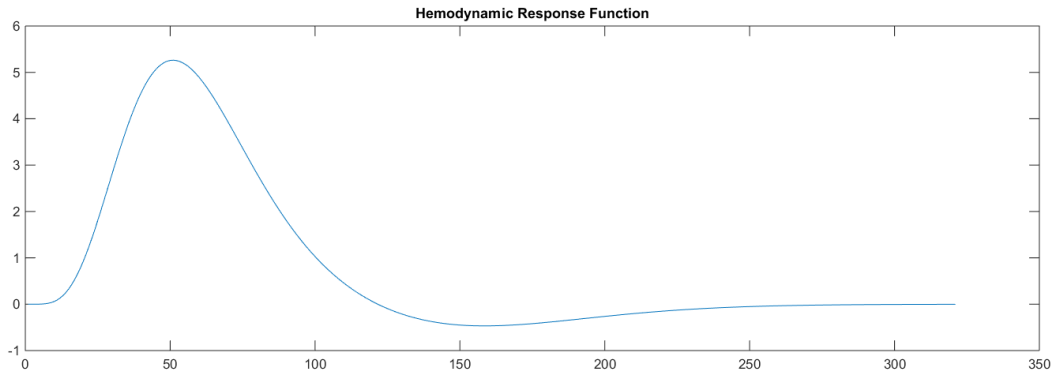

Figure 3.6. Hemodynamic Response Function

After the HRF generation, the raw data is convolved with HRF data and synthetic fMRI data is generated. Only the 6-channel data are convolved with the HRF in order to see the HRF effects. Table 3.4 shows the number of data generated using HRF convolution.

Table 3.4. *Number of Data generated with HRF*

	Data with HRF	
Linearity	Linear	Nonlinear
Number of Channel	6	6
# of Data Generated	10	10
Total # of Data Generated	10	10
	20	

Figure 3.7 shows Channel 1 of the linear signal after HRF application.

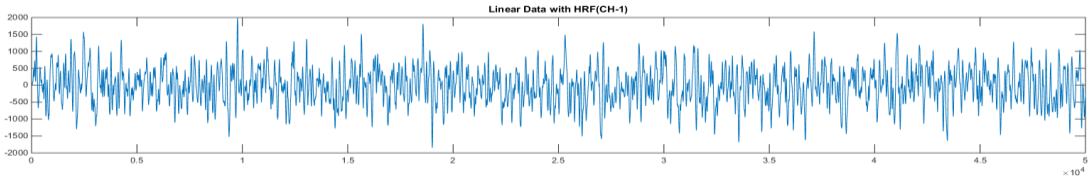

Figure 3.7. Linear signal with HRF

Channel 1 of the nonlinear signal after HRF application is presented in Figure 3.8.

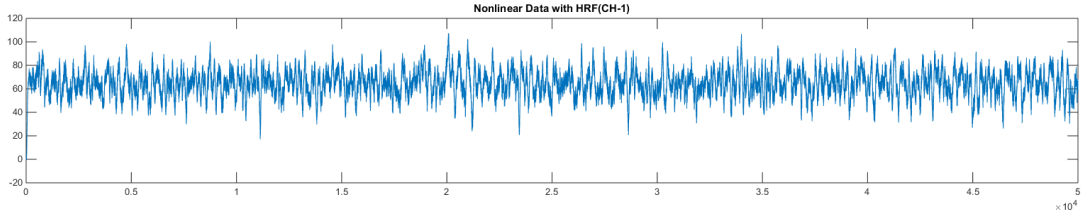

Figure 3.8. Nonlinear signal with HRF

3.2. Implementing PDC and DBN Methods

The PDC and DBN methods are implemented in MATLAB. While implementing the methods, for PDC, the BioSig toolbox from Schlogl, A. is used [43]. For DBN, the Bayesian Network toolbox from Kevin Murphy and MCMC toolbox extension by Husmeier is used. As mentioned earlier, PDC is using the coefficients of the network to construct the connectivity, while DBN directly interprets connectivity from the signals. Therefore, before applying PDC, an MVAR approach is applied inside the mentioned toolbox to extract the coefficients from the signal. Then, PDC is applied using the extracted coefficients.

3.3. Verifying PDC and DBN Methods

The application of PDC method is verified by generating the same data from the paper of Baccala, L. and Sameshima, K. (2001) [31], and getting the same results in the paper. One of the results from the mentioned paper and our trial can be seen in Figure 3.9

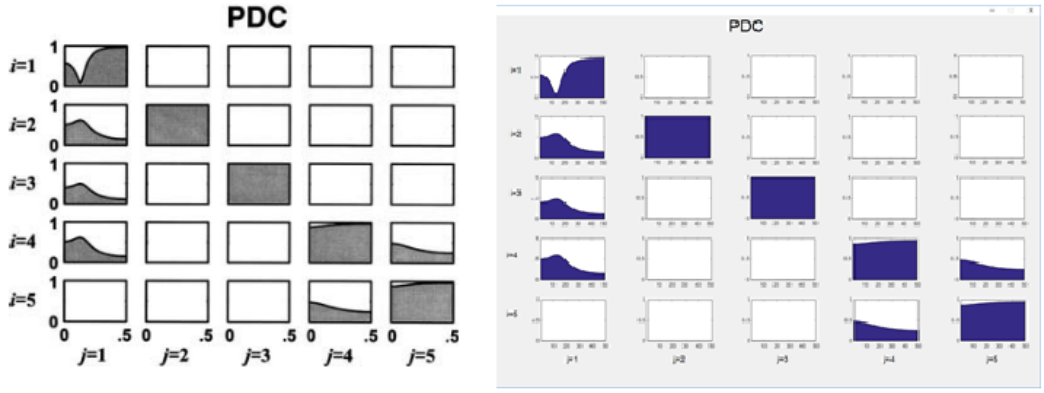

Figure 3.9. PDC Result from [29] (left), PDC result of our application (right)

The application of DBN method is verified by generating a nonlinear data with known coefficients and 500000 samples. The error was zero for that connectivity estimation.

3.4. Applying PDC and DBN to Synthetic Data

Applying PDC and DBN methods to a data gives result as channel-by-channel matrix which indicates the connectivity information. While PDC result matrix gives continuous values between 0 and 1, DBN result matrix gives a binary result indicating whether there is a connection from a channel to another or not as mentioned in Chapter 2.5 Connectivity Adjacency Matrix.

During the application every generated data is given as input to PDC and DBN methods. The output of these methods, the result matrices, are compared with the known connectivity values specified for each data before generation, and error is calculated to analyse the performances.

Since the types of the results are different, different error calculation methods are preferred for PDC and DBN.

For PDC, Mean Square Error method is used. In this method, error is calculated as follows:

$$Error_{PDC} = \sqrt{\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}(PDCr_{ij} - PDC_{ij})^2}{N^2}} \tag{35}$$

where,

$$\begin{aligned} PDCr, &\quad i\ by\ j\ expected\ connectivity\ result\ matrix \\ PDC, &\quad i\ by\ j\ calculated\ connectivity\ result\ matrix \\ N, &\quad number\ of\ channels \end{aligned} \tag{36}$$

For DBN, using Mean Square Error method is not the best solution to calculate errors, since the result matrix given by the DBN is not continuous. On the contrary, the result matrix is binary. It shows whether there is a considerable connection between channels. The error calculation for DBN is carried out by using the following formula:

$$Error_{DBN} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left|(DBNr_{ij} - DBN_{ij})\right|}{N^2} \tag{37}$$

where,

$$\begin{aligned} DBNr, &\quad i\ by\ j\ expected\ connectivity\ result\ matrix \\ DBN, &\quad i\ by\ j\ calculated\ connectivity\ result\ matrix \\ N, &\quad number\ of\ channels \end{aligned} \tag{38}$$

When different error calculation methods are put into practice, comparing error values of DBN and PDC will not be objective, since comparison will not be on the same frame of reference. The previous calculations express the performances of each

method individually. In order to overcome this problem, two different result matrices should be both continuous or both binary. Converting binary results to continuous values in this situation seems farfetched. The optimized solution is to digitize the PDC result matrix in order to resemble it to the result of DBN. After that, equation (37) can be used to calculate error which is used to compare PDC and DBN results. However, analysing the effects of different parameters on PDC, equation (35) is still the best error calculation method.

To make the error values of both methods resembling, the following calculations are applied to calculate the binary PDC error.

$$E \Rightarrow \ i\ by\ j\ matrix\ consisting\ of\ elements\ e_{ij},\ \ where;$$

$$\boldsymbol{e_{ij}} = \begin{cases} 1\ ,\ if & \left|\frac{(PDCr_{ij} - PDC_{ij})}{1}\right| \leq \frac{standard\ deviation\ of\ PDCr}{2} \\ 0,\ if & \left|\frac{(PDCr_{ij} - PDC_{ij})}{1}\right| > \frac{standard\ deviation\ of\ PDCr}{2} \end{cases} \quad for\ all\ i\ and\ j \qquad (39)$$

$$Error_{PDC_B} = \frac{N^2 - \sum_{i=1}^{N}\sum_{j=1}^{N} \boldsymbol{e_{ij}}}{N^2}$$

In the Equation (39), the standard deviation of the real PDC result is used to decide if the estimated PDC result is close to the real value in order to accept it to be true.

3.4.1. Applying PDC to Synthetic Data

3.4.1.1. Applying PDC to Raw Linear Data

PDC result matrix calculated from one of the 6 channel synthetic linear data and the expected PDC can be examined in Figure 3.10.

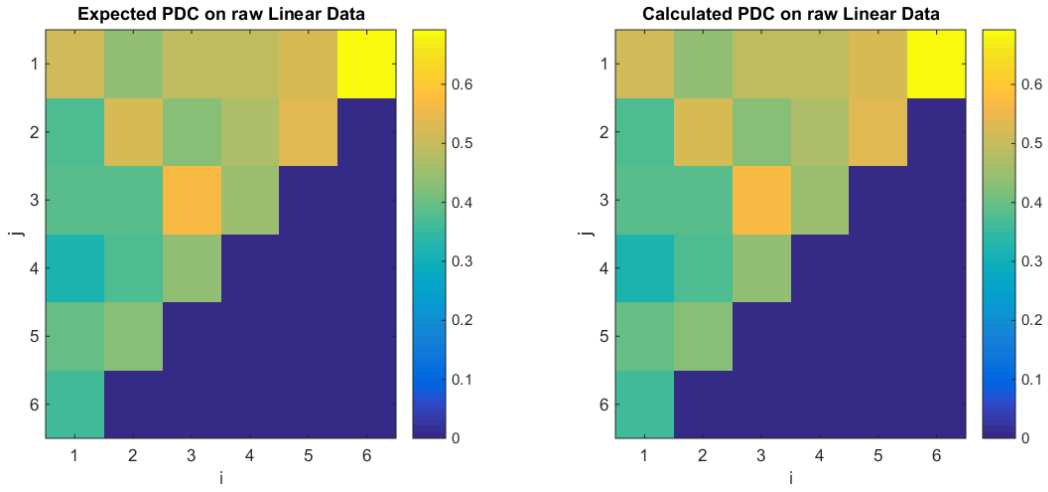

Figure 3.10. PDC results on raw linear data

Error values calculated from the application of PDC to raw linear data are shown on Figure 3.11. In this analysis, the average of error values from ten different data for each channel is used.

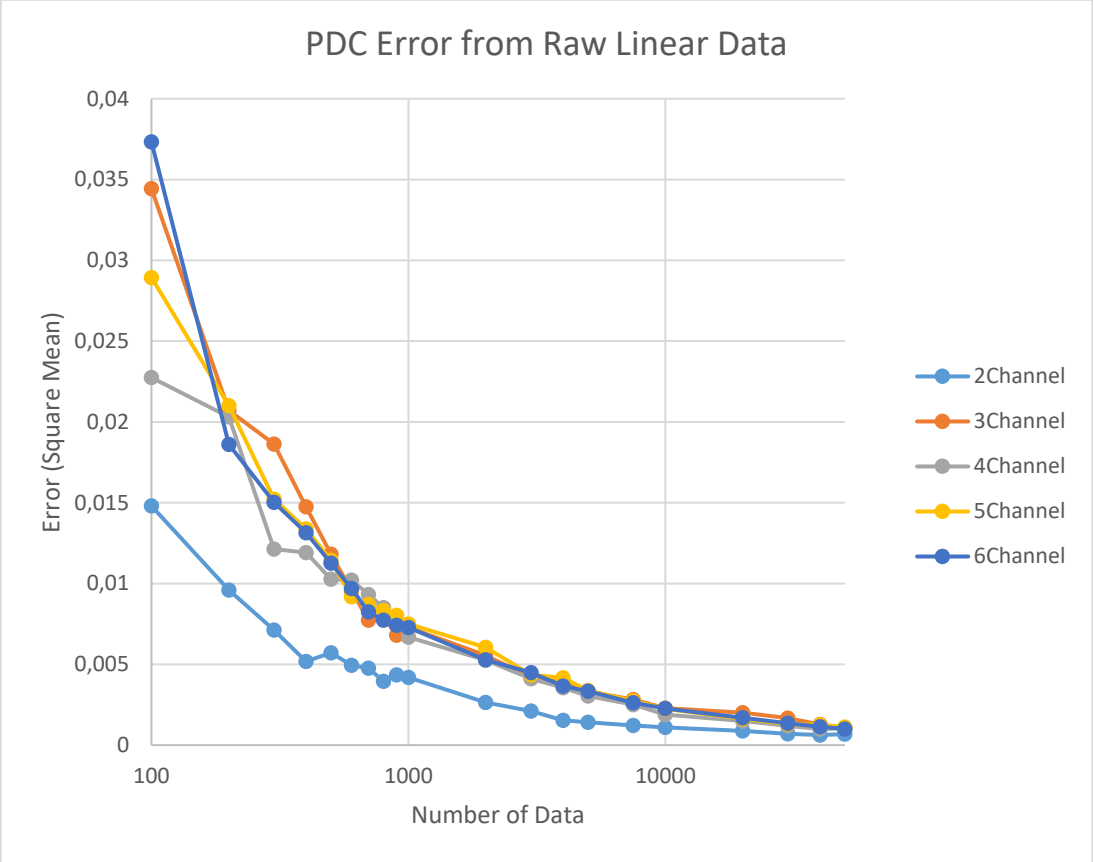

Figure 3.11. PDC Error from Raw Linear Data

From Figure 3.11, the following deductions can be made:

- The performance of PDC to extract connectivity information from given data increases when data length increases. This is something expected as discussed in Chapter 2. However, it can be deduced that even when the data length is small, the error of PDC is satisfactorily low. Since the underlying method in PDC is multivariate auto regressive model, which consists of the linear equations between channels, the effectiveness of PDC on linear data is understandable.

- Increase in number of channels increases the error. However, this increase is not significant for higher data lengths, it shows the effect when the data length is small.

Error values calculated from the application of PDC to raw linear data with different noise power values are shown on Figure 3.12. In this analysis, the average of error values from ten different data for 6-channel data is used.

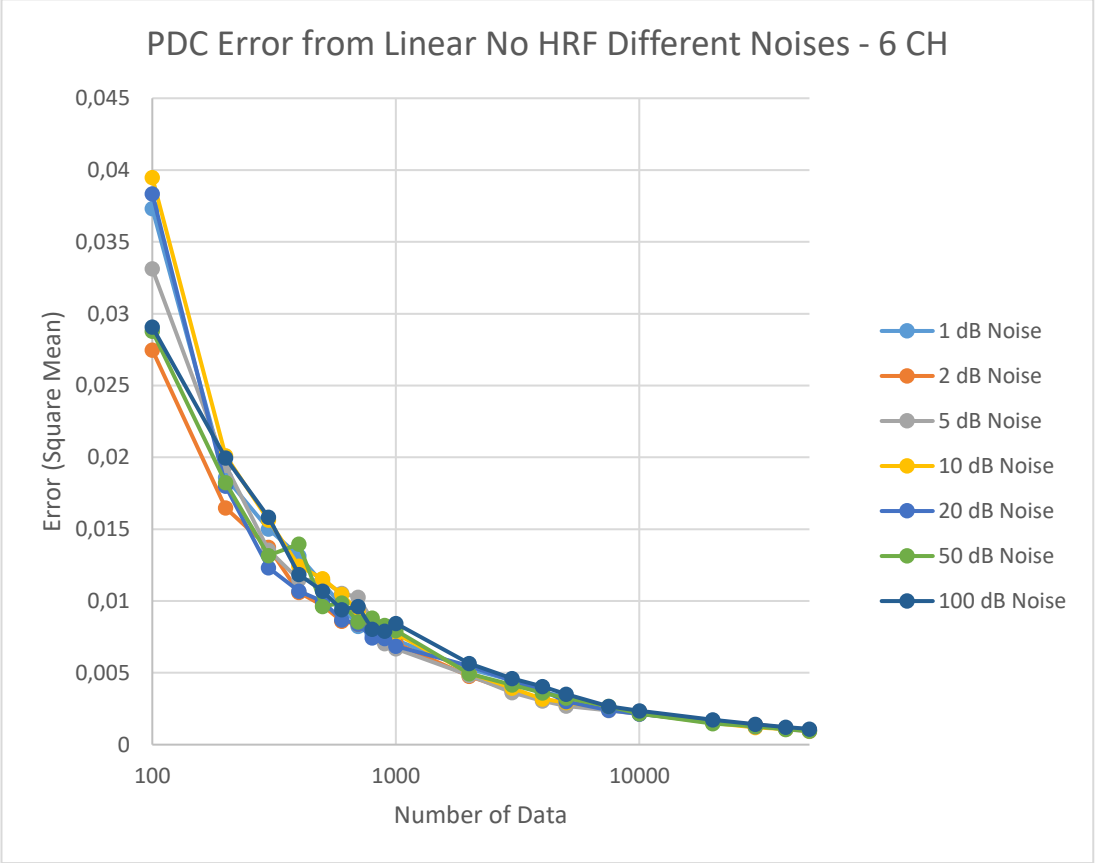

Figure 3.12. PDC Error from Raw Linear Data with Different Noise Powers

The effect of different power values of noise can be seen for smaller values of data length. However, this difference is very small. We can say that for linear data using different noise values in the generation of the data has no effect for PDC to reconstruct the data.

3.4.1.2. Applying PDC to Raw Nonlinear Data

PDC result matrix calculated from one of the 6 channel synthetic nonlinear data can be examined in Figure 3.13.

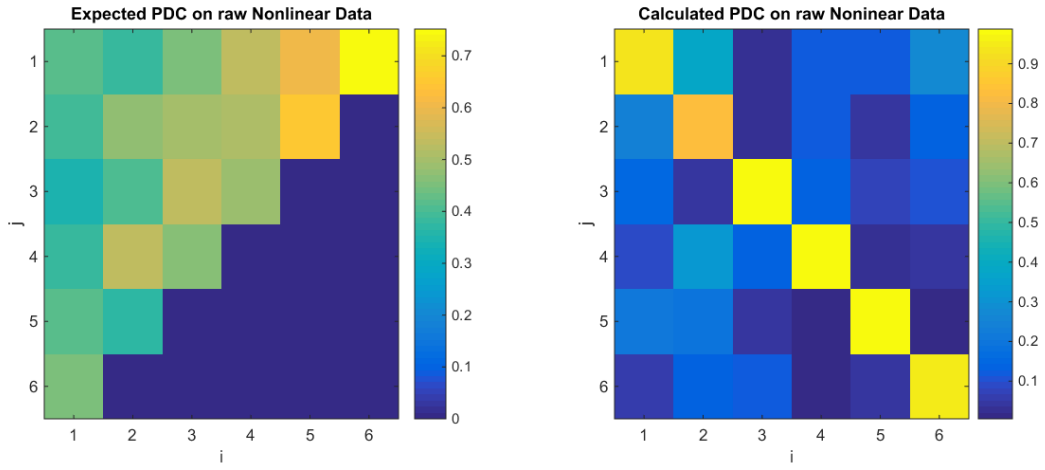

Figure 3.13. PDC results on raw nonlinear data

Error values calculated from the application of PDC to raw nonlinear data are shown on Figure 3.14. In this analysis, the average of error values from ten different data for each channel is used.

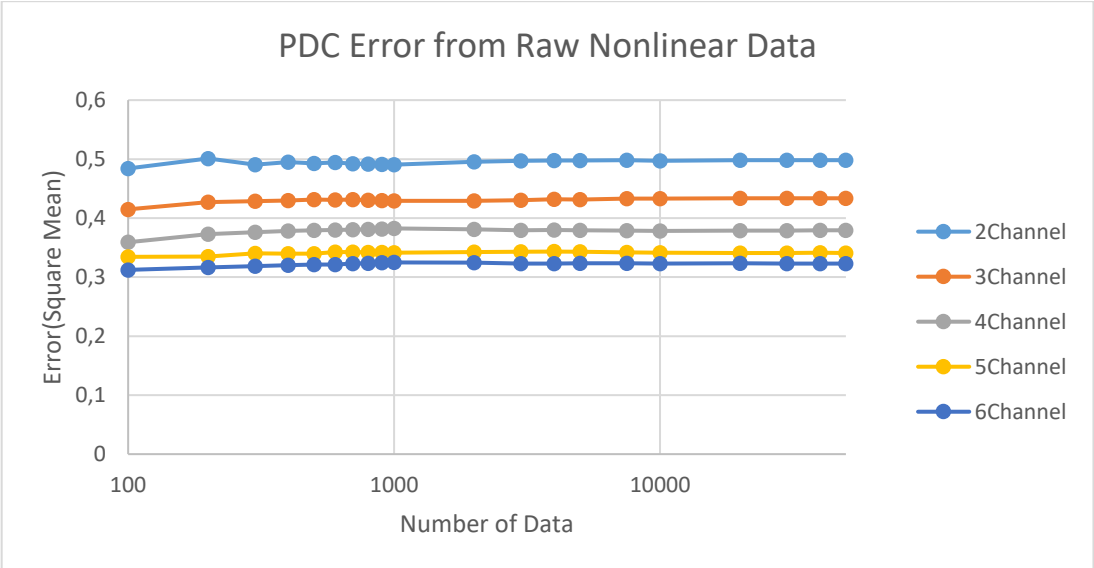

Figure 3.14. PDC Error from Raw Nonlinear Data

From Figure 3.14 , the following deductions can be made:

- The performance of PDC on a nonlinear data is not reliable, it does not extract the connection information. Changes in the length of the data and the number of channels are not worth consideration because their effects are negligible when compared to how big is the error.

Error values calculated from the application of PDC to raw nonlinear data with different noise power values are shown on Figure 3.15. In this analysis, the average of error values from ten different data for 6-channel data is used.

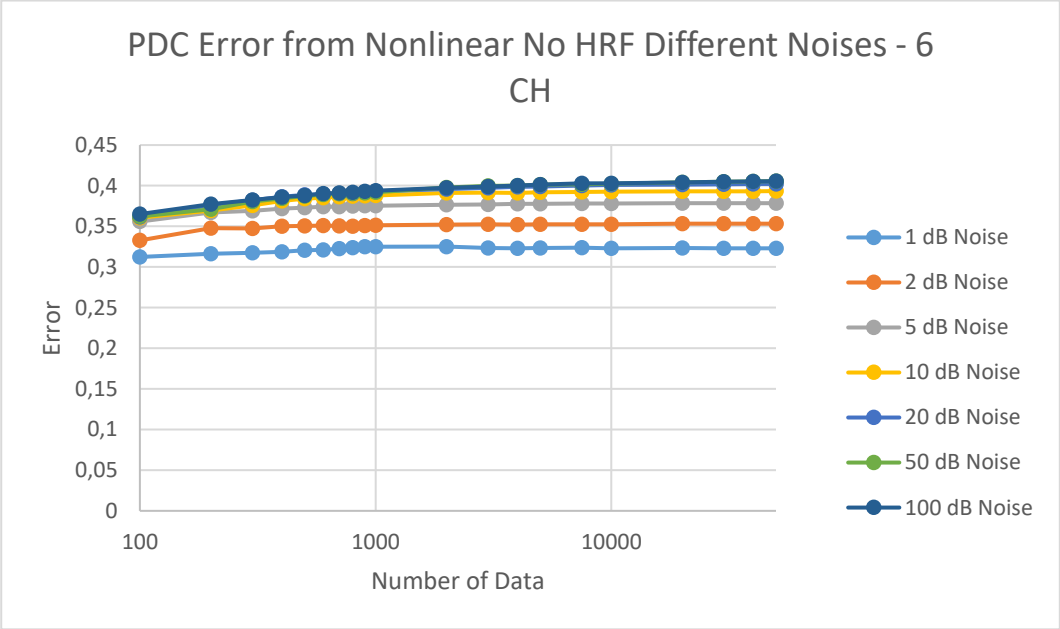

Figure 3.15. PDC Error from Raw Nonlinear Data with Different Noise Powers

- The effect of different power values of noise for PDC on nonlinear data cannot be commented since PDC does not extract the connection information.

After comparing the performance of the PDC in both linear and nonlinear data, it is clear to be said that PDC extracts the connection information much better on a linear

data then nonlinear data. Figure 3.16 concludes the discussion done in Chapters 3.4.1.1 and 3.4.1.2.

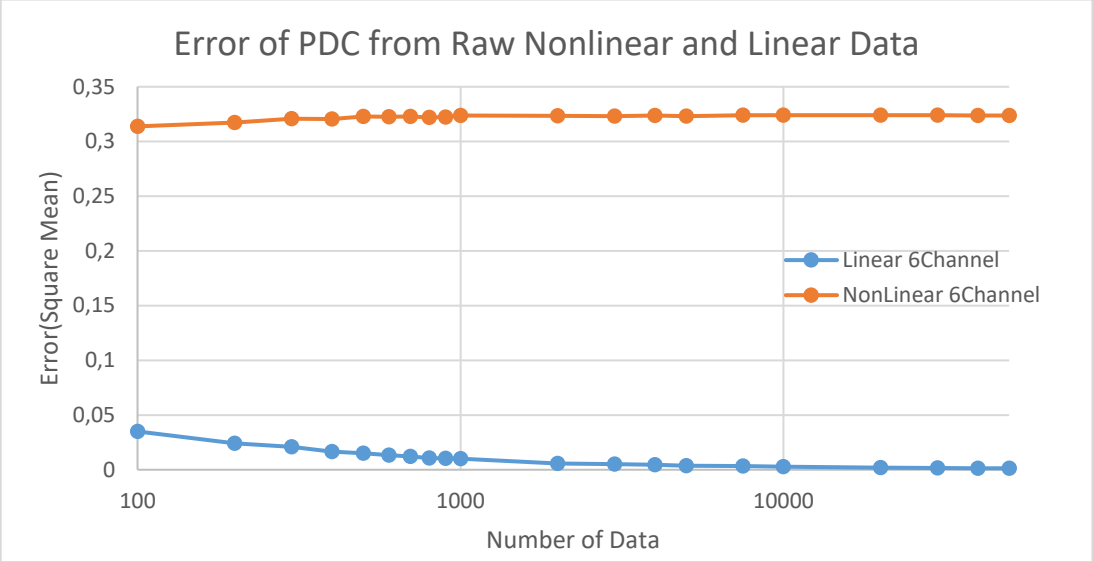

Figure 3.16. Comparison of PDC Error from Raw Nonlinear and Linear Data

3.4.1.3. Applying DBN to Raw Linear Data

DBN result matrix calculated from 6 channel synthetic linear data can be examined in Figure 3.17 together with expected connection values.

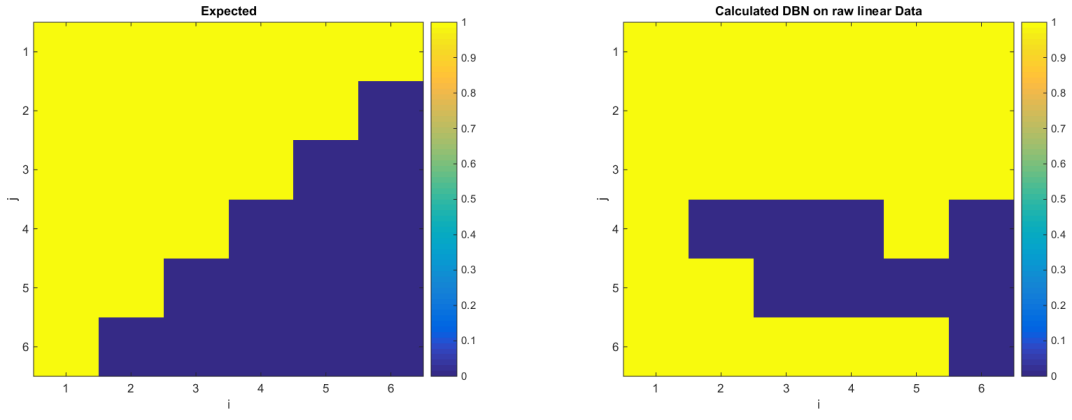

Figure 3.17. DBN results on raw linear Data

Error values calculated from the application of DBN to raw linear data are shown on Figure 3.18. In this analysis, the average of error values from ten different data for each channel is used.

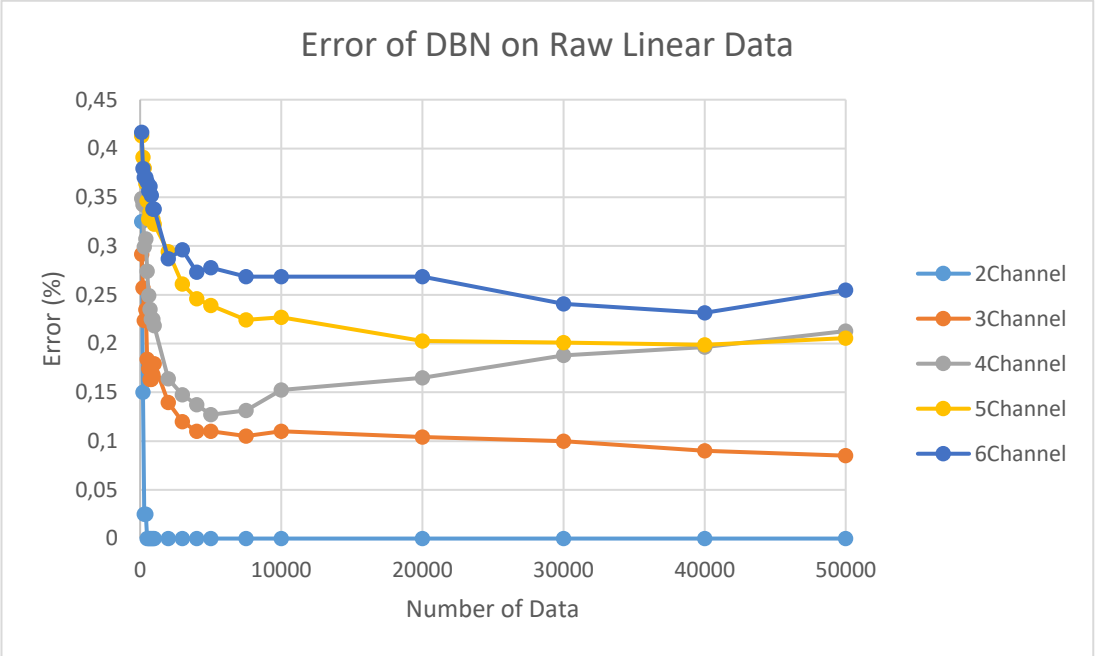

Figure 3.18. DBN Error from Raw Linear Data

From Figure 3.18, the following deductions can be made;

- The performance of DBN to extract connectivity information from given data increases when data length increases. This is something expected as discussed in Chapter 2.
- Increase in number of channels increases the error.

Error values calculated from the application of DBN to raw linear data with different noise power values are shown on Figure 3.19. In this analysis, the average of error values from ten different data for 6-channel data is used and data length is taken as 50000.

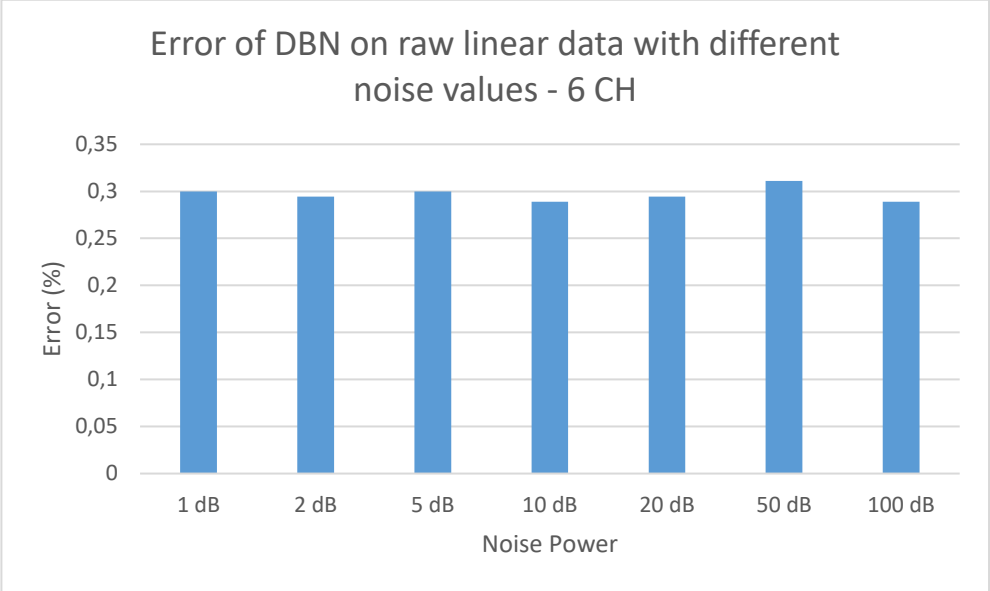

Figure 3.19. DBN Error from Raw Linear Data with Different Noise Powers

The effect of different power values of noise on DBN for raw linear data has shown no effect. For linear signals DBN can return same values for different signals generated with different noise powers.

3.4.1.4. Applying DBN to Raw Nonlinear Data

DBN result matrix calculated from 6 channel synthetic nonlinear data can be examined in Figure 3.20 together with expected connection values.

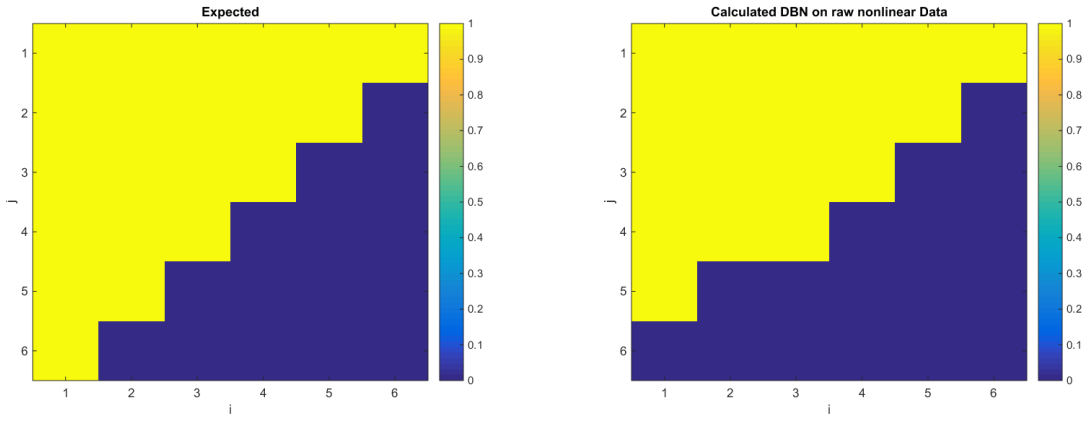

Figure 3.20. DBN results on raw nonlinear Data

Error values calculated from the application of DBN to raw nonlinear data are shown on Figure 3.21. In this analysis, the average of error values from ten different data for each channel is used.

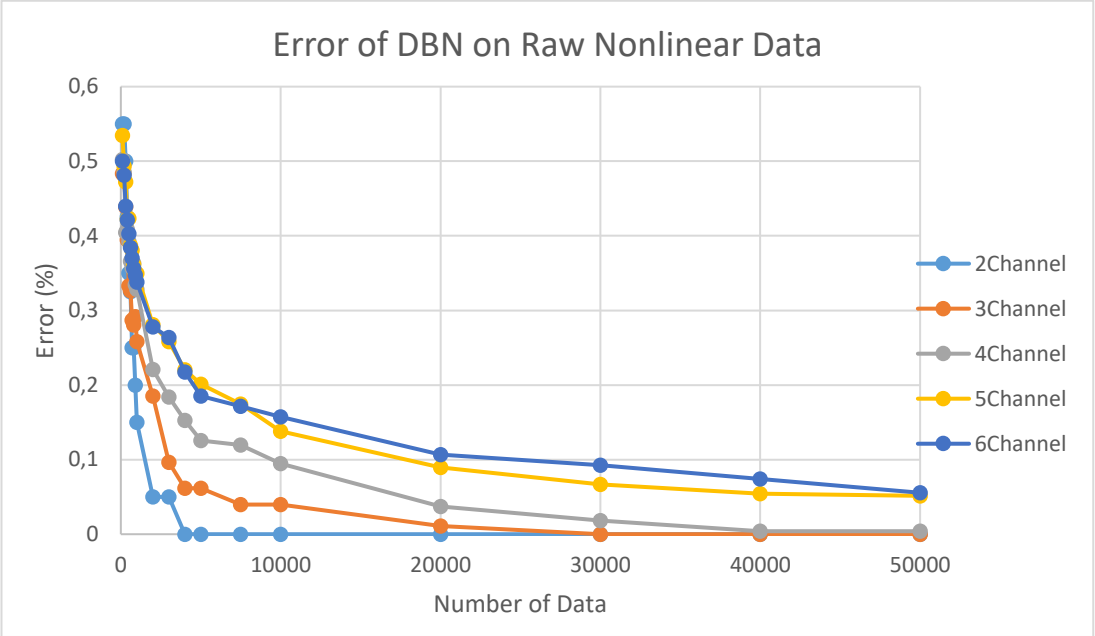

Figure 3.21. DBN Error from Raw Nonlinear Data

From Figure 3.21, the following deductions can be made:

- The performance of DBN to extract connectivity information from nonlinear data increases when data length increases as seen in other case.
- Increase in number of channels increases the error.
- The performance of DBN on nonlinear signals is very satisfactory.

Error values calculated from the application of DBN to raw nonlinear data with different noise power values are shown on Figure 3.22. In this analysis, the average of error values from ten different data for 6-channel data is used and data length is taken as 50000.

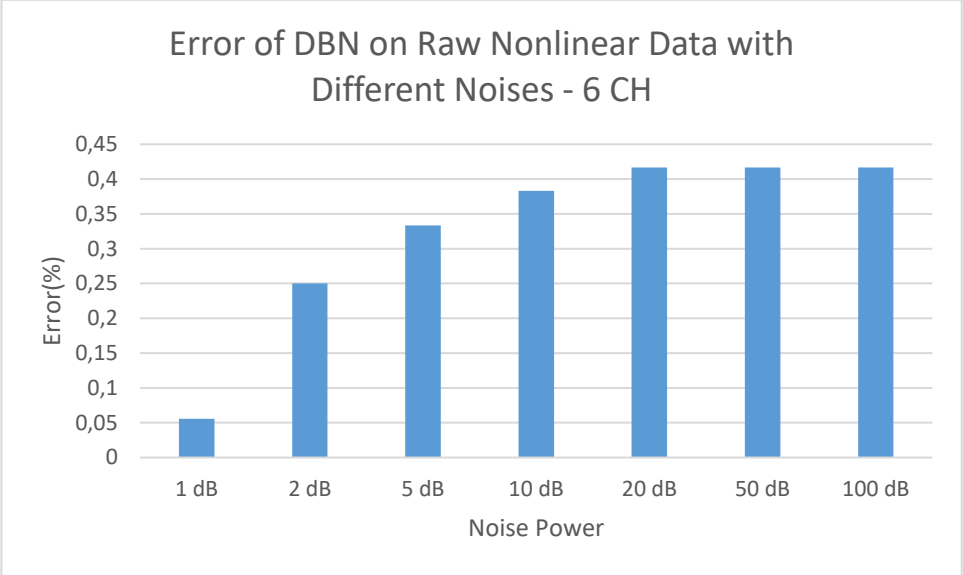

Figure 3.22. DBN Error from Raw Nonlinear Data with Different Noise Powers

As it can be seen, the increase in the power of noise increases the error. This is because the nonlinear data that is generated has the tendency to converge in small values and increase in noise disrupts the data that is generated from the connectivity base, and DBN has some hard time to figure out the connectivity values for bigger noise powers.

After comparing the performance of the DBN in both linear and nonlinear data, it is clear to be said that DBN extracts the connection information much better on a nonlinear data then linear data. Figure 3.23 concludes the discussion done in Chapters 3.4.1.3 and 3.4.1.4.

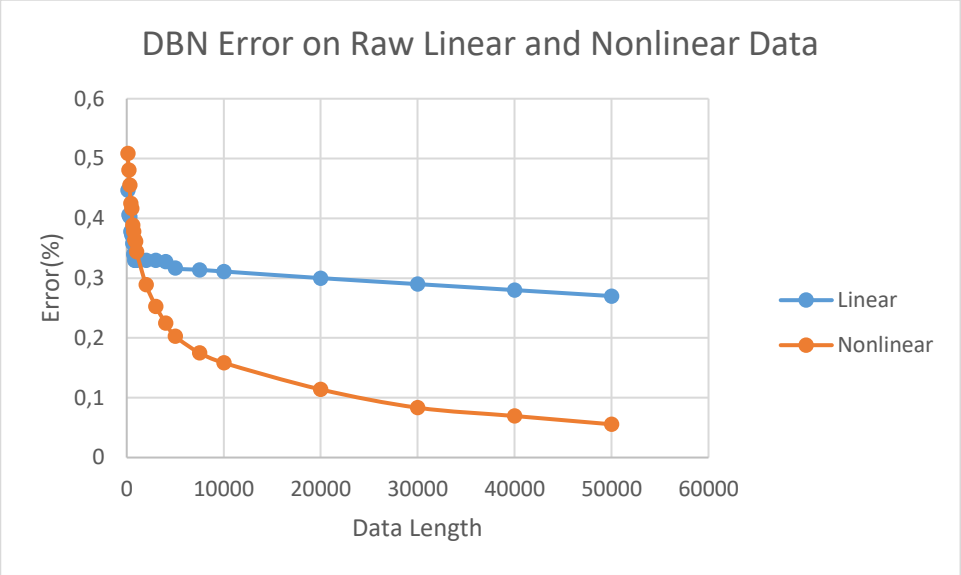

Figure 3.23. Comparison of DBN Error from Raw Nonlinear and Linear Data

3.4.2. Analyzing Data Requirements for PDC

The variable requirements for MVAR model was discussed in Chapter 2.6.2. If those parameters applied to our case, using equation (9) with the following values;

$$\begin{aligned} m &= 6, \text{number channels} \\ p &= 1, \text{order of the signals} \end{aligned} \tag{40}$$

Then,

$$\begin{aligned} n &> 10 * (m^2 p) \\ n &> 10 * (36 * 1) \\ n &> 360 \end{aligned} \tag{41}$$

PDC is based on MVAR model. In order to calculate the data requirements for PDC we can use the MVAR requirements, however there could be slight differences since they are different methods.

From Figure 3.24, it can be seen that the error values of PDC are very small for all of the values; however after the length of the data exceeds 400, the error starts to converge to a smaller value. For a safe assumption, using practice rather than theoretical knowledge, it can be concluded that for a 6-channel data, when the order is 1, data length of 1000 or higher is needed to trust PDC result. Of course, this case is applicable for linear signals.

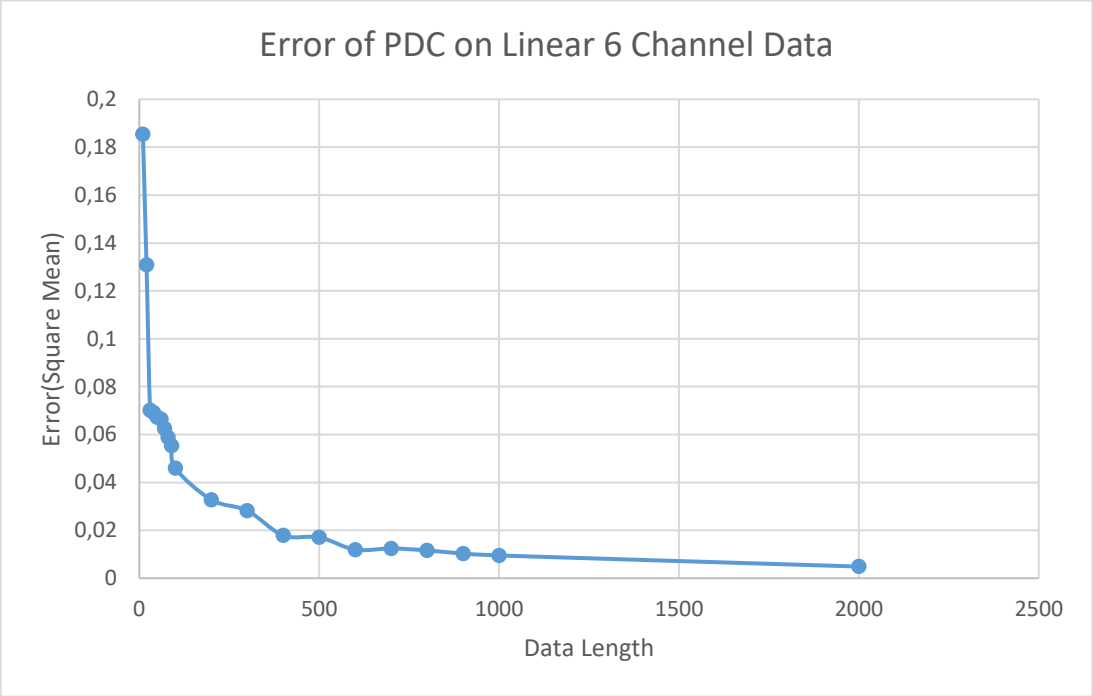

Figure 3.24. Error of PDC on a Raw Linear 6 Channel Data

3.4.3. Analyzing Data Requirements for DBN

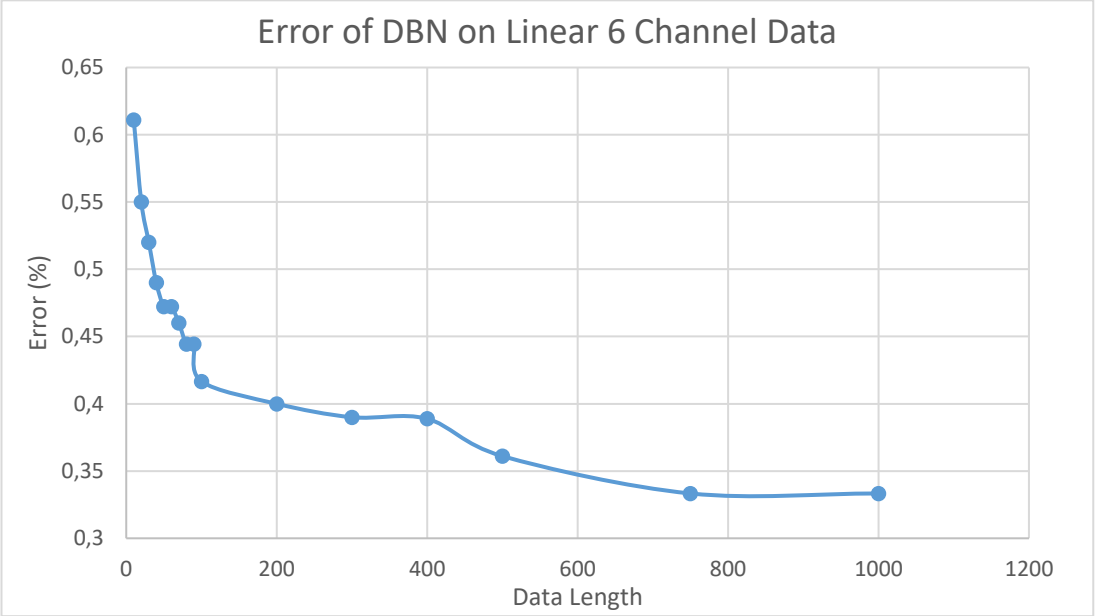

Figure 3.25. Error of DBN on a Nonlinear 6 Channel Data

From Figure 3.25, it can be seen that the error values of DBN starts to converge to a smaller value after the length of the data exceeds 400 to a smaller value. For a safe assumption, it can be concluded that for a 6-channel data, when the order is 1, data length of 1000 or higher is needed to trust DBN result. Of course, this case is applicable for nonlinear signals.

3.4.4. Applying to Data with HRF

3.4.4.1. Applying PDC to Linear Data with HRF

HRF is applied to raw linear 6-channel data. The error values after the application of PDC to linear data with HRF can be seen in Figure 3.26.

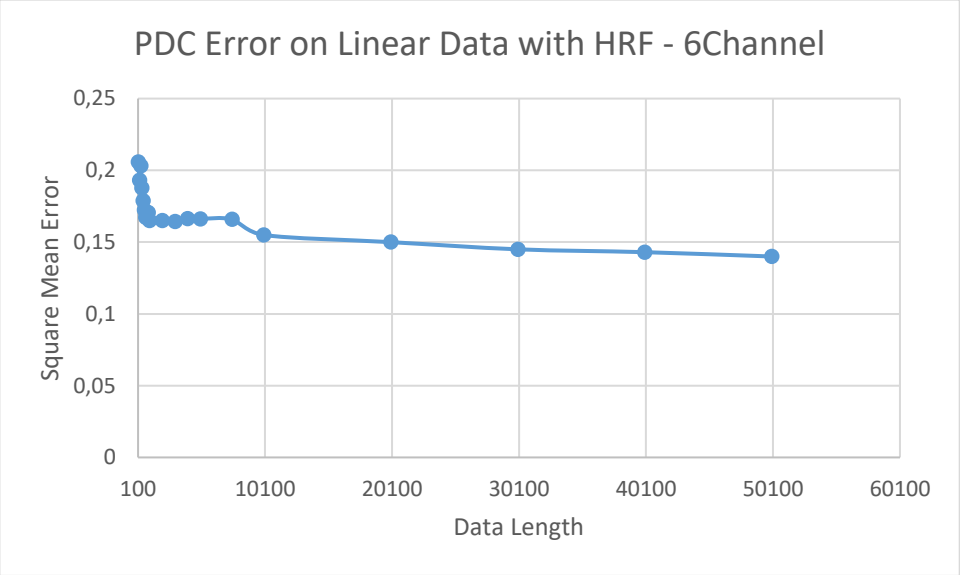

Figure 3.26. PDC Error from 6 Channel Linear Data with HRF

The results will be compared with the error values with no HRF in the following section.

3.4.4.2. Applying PDC to Nonlinear Data with HRF

HRF is applied to raw nonlinear 6-channel data. The error values after the application of PDC to nonlinear data with HRF can be seen in Figure 3.27.

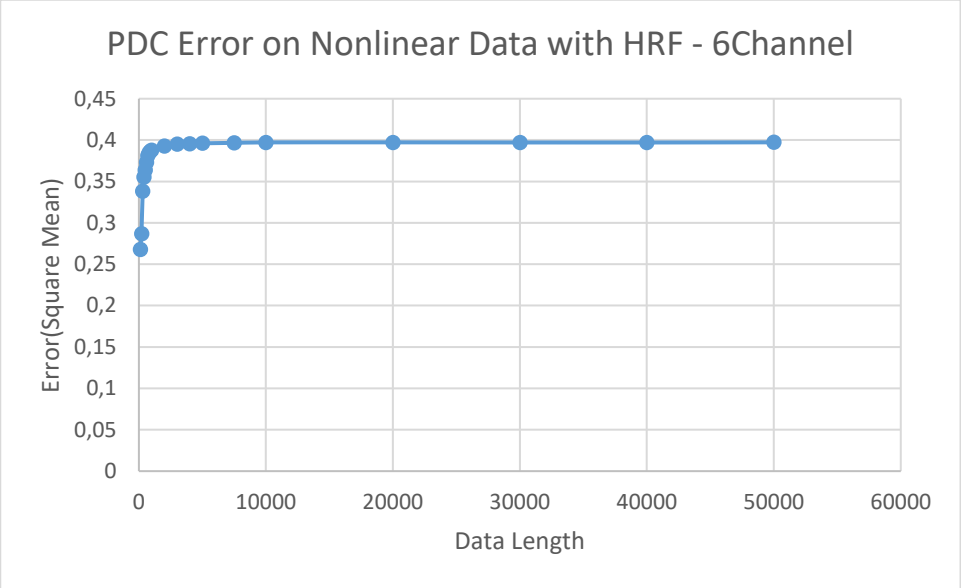

Figure 3.27. PDC Error from 6 Channel Nonlinear Data with HRF

Comparison of error results of PDC for four data types is presented in Figure 3.28.

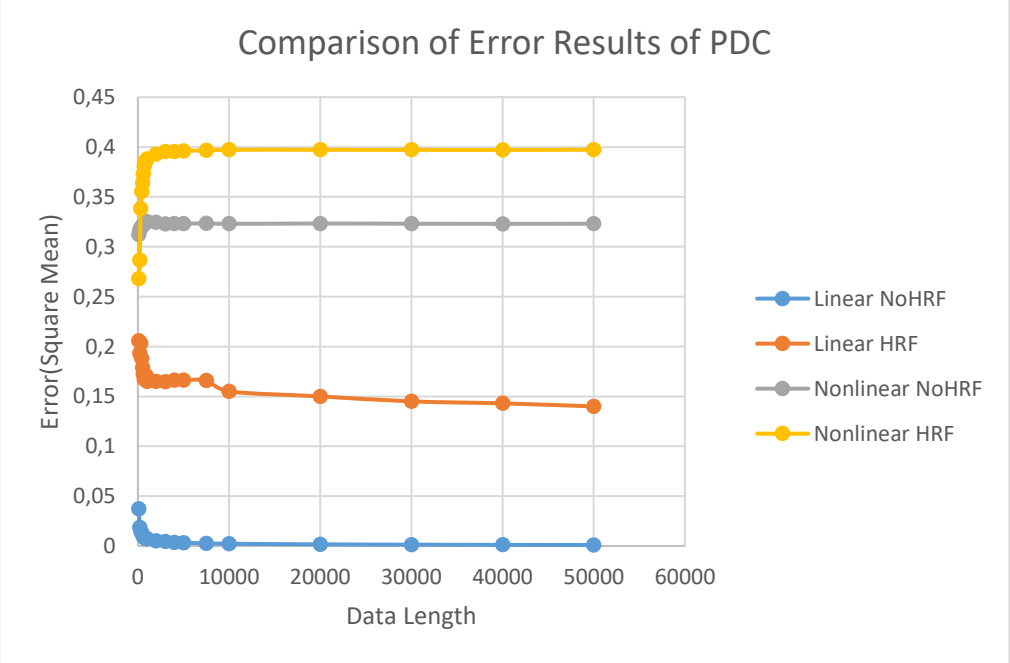

Figure 3.28. Comparison of Error Results of PDC

3.4.4.3. Applying DBN to Linear Data with HRF

HRF is applied to raw linear 6-channel data. The error values after the application of DBN to linear data with HRF can be seen in Figure 3.29.

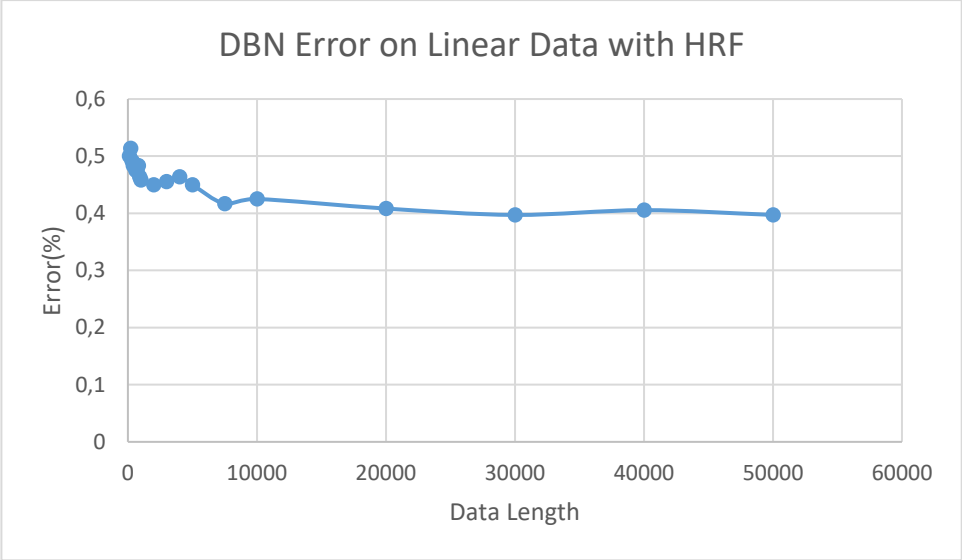

Figure 3.29. DBN Error from 6 Channel Linear Data with HRF

The results will be compared with the error values with no HRF in the following section.

3.4.4.4. Applying DBN to Nonlinear Data with HRF

HRF is applied to raw nonlinear 6-channel data. The error values after the application of DBN to nonlinear data with HRF can be seen in Figure 3.30.

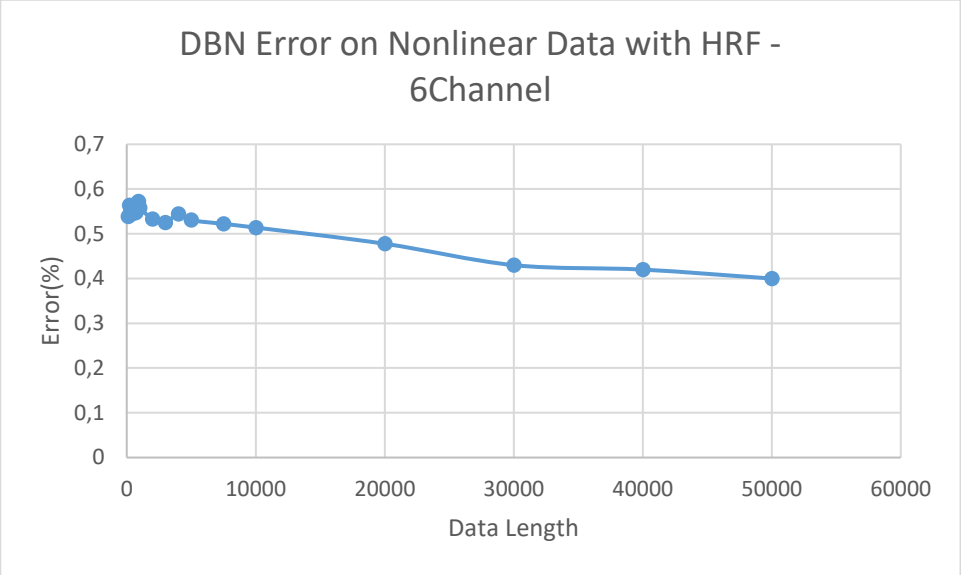

Figure 3.30. DBN Error from 6 Channel Nonlinear Data with HRF

Comparison of error results of PDC for four data types is presented in.

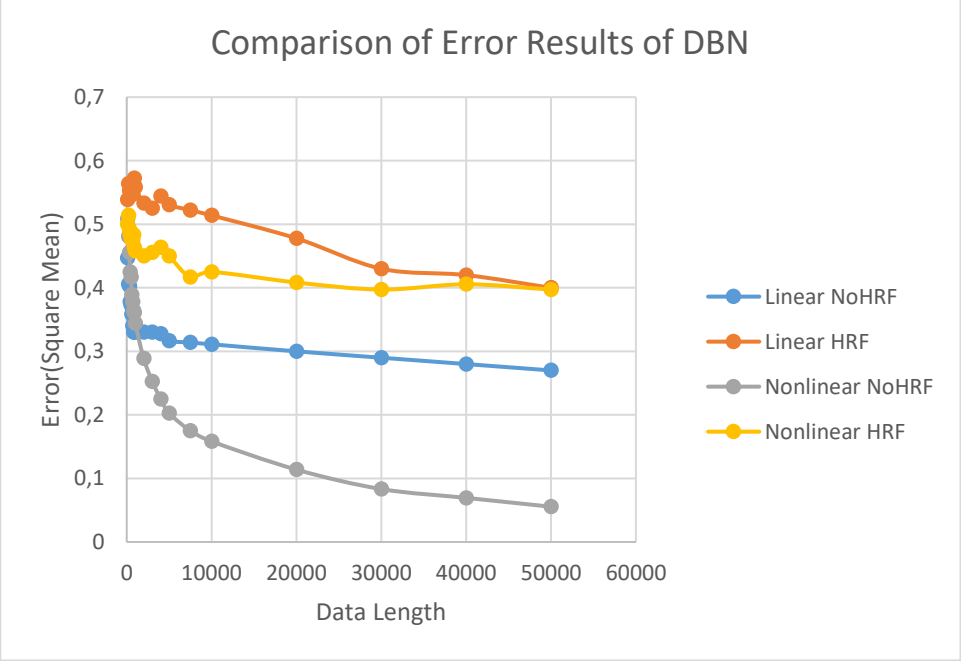

Figure 3.31. Comparison of Error Results of DBN

3.5. Comparison of PDC and DBN Performances on Synthetic Data

As stated earlier, in order to compare the PDC and DBN, the error calculation of PDC is changed. The comparison of PDC and DBN for four types of data (linear, nonlinear, linear with HRF, nonlinear with HRF) will be done using average error of 6 channel data.

Error result values both PDC and DBN for linear data can be analyzed in Figure 3.32.

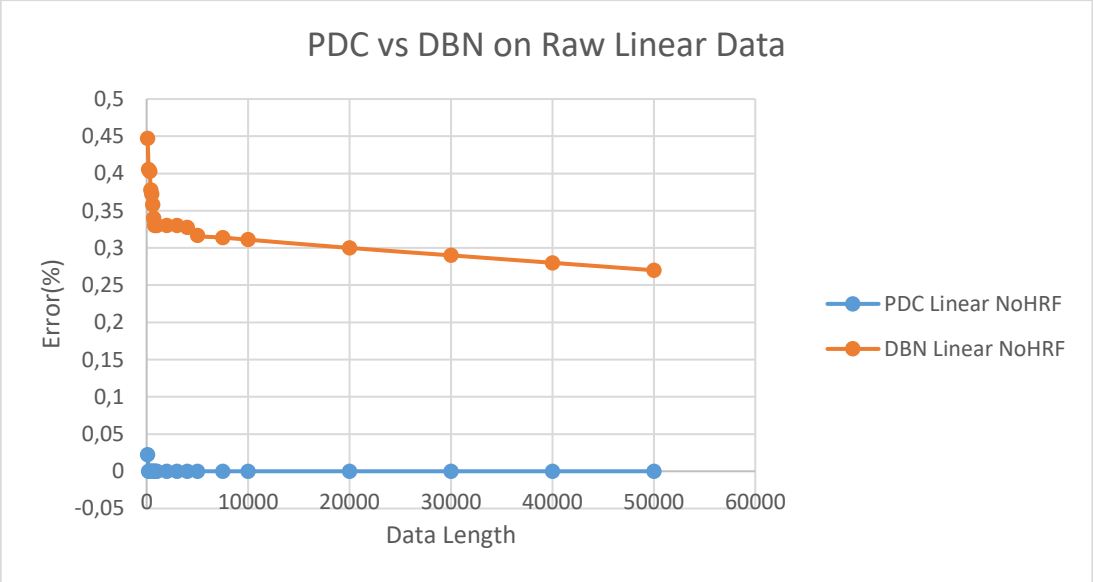

Figure 3.32. Comparison of PDC and DBN on Linear Data

Error result values both PDC and DBN for nonlinear data can be analyzed in Figure 3.33.

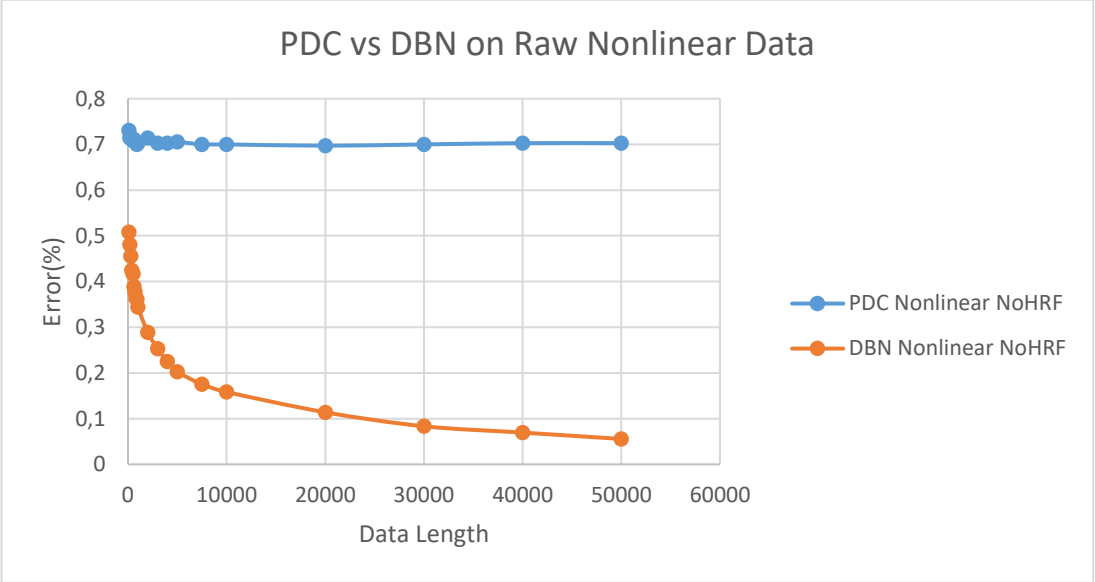

Figure 3.33. Comparison of PDC and DBN on Nonlinear Data

Error result values both PDC and DBN for linear data with HRF can be analyzed in Figure 3.34.

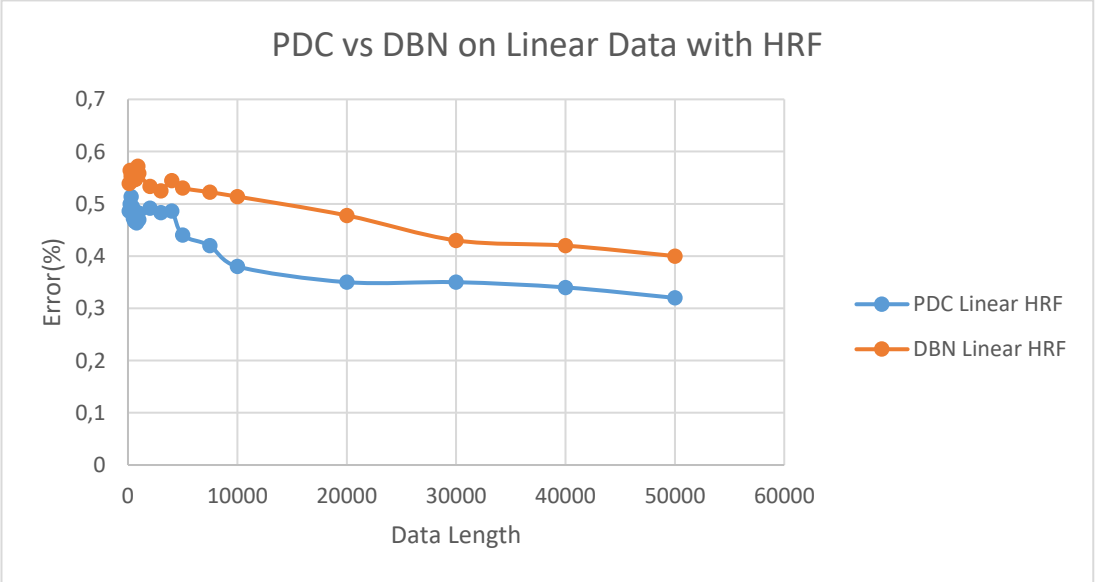

Figure 3.34. Comparison of PDC and DBN on Linear Data with HRF

Error result values both PDC and DBN for nonlinear data with HRF can be analyzed in Figure 3.35.

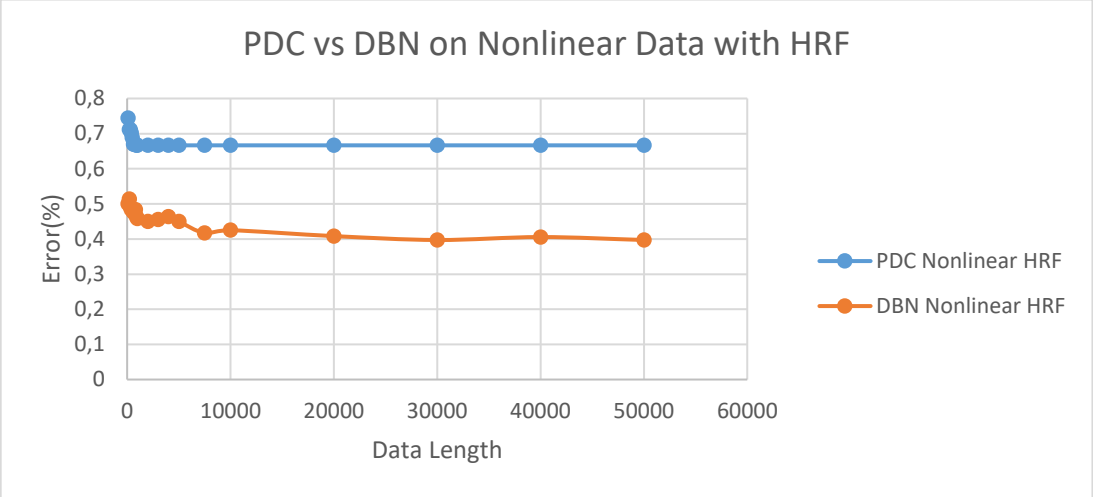

Figure 3.35. Comparison of PDC and DBN on Nonlinear Data with HRF

Figure 3.36 shows the total comparison of PDC and DBN together.

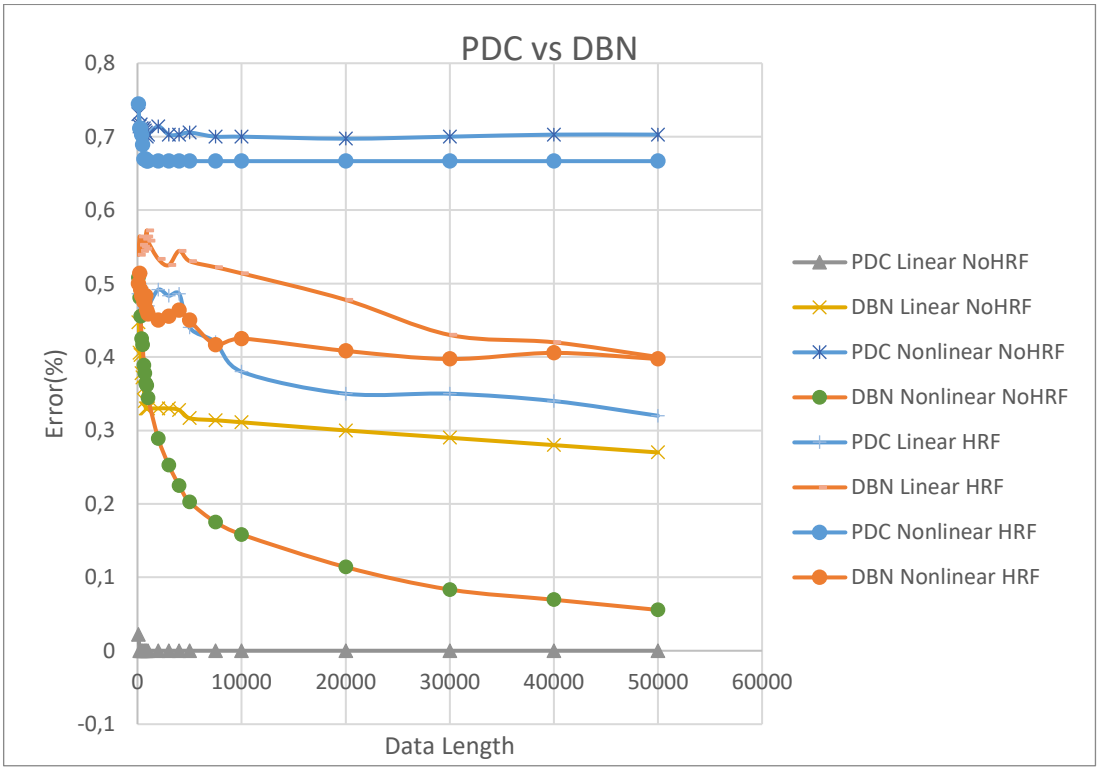

Figure 3.36. Comparison of PDC and DBN on all Data

3.6. Discussion of the Results

In order to summarize the experimental results for constructed synthetic data and discuss them, Table 3.5 can be constructed.

Table 3.5. *Summary of Experimental Results*

	Performance of PDC	*Performance of DBN*
Linear Signals	+	(+)
Nonlinear Signals	-	+
Length of the Signals	Positive Correlation	Positive Correlation
Number of Channels	Negative Correlation	Negative Correlation
Effect of HRF	(+)	(+)
Power of Noise	Neutral	Neutral

Linear signals: When applied to linear signals, both PDC and DBN can construct the network. However, PDC outperforms DBN in linear signals. This is because MVAR model that reconstructs the coefficients to be applied to PDC is based on linear modelling. As a probabilistic method DBN also reconstructs the connectivity values on linear signals, but not as well as PDC.

Nonlinear signals: In our datasets and experimental setup, PDC fails to reconstruct nonlinear signals. Again, it is because its origins are coming from linear systems. On the other hand, DBN shows great impact on nonlinear signals.

Length of the signals: As the sample size increases, the performances of both PDC and DBN also increase. Intuitively, this is the expected result. When the observed part of a signal is large enough, it is easy to interpret the hidden connections.

Number of channels: Complexity of a network is largely dependent on the number of elements inside it. Every channel that added to a network makes the network more complex, since it brings new connection with all the existing channels. In our experiments, when the number of channel increases, i.e. the complexity of the network increases, the performances of PDC and DBN decrease. However, this does not mean

that they fail for larger channel numbers. Our channel number varies between 2 to 6. In the worst-case scenario, when the number of channels is 6, PDC and DBN still recovers the connectivity values. They are only just better with smaller values.

Effect of HRF: Hemodynamic response function simulates the effect of BOLD signals in the brain. When the HRF effect is added to the synthesized signals, it makes harder to interpret the connections of the channels to each other. Therefore, PDC and DBN performs poorly on signals with HRF effect. However, they still manage to extract some of the connectivity information.

Power of White Noise: Changing the power of noise in the linear and nonlinear data generation doesn't really affect the performances of the PDC and DBN. Since the noise is added for every time step, and the amplitude of the next signal in the time domain already has the effect of noise, the power of white noise does not change the type of the signal. However, if the amplitude of the noise starts to reach the level of the signal it distorts the signal and therefore cause models to fail on reconstruction.

These results are the expected results beforehand. Showing the same results with synthetic data and known coefficients add an experimental result to these expectations. Since there is no comparison of PDC and DBN on literature. This work contributes to the understanding of different effective connectivity methods.

3.7. Statistical Comparison of PDC, DTF, DBN Performances on Group Data

In this section, a further comparison of PDC and DBN will be covered. In addition to PDC and DBN, this time the performance of DTF will also be examined.

In order to compare these three estimator methods, we create a group experiment where two synthetic data groups have distinct characteristics different from each other. However, the characteristics within the groups are similar. The main purpose of this experiment is simulating the patient-control group situation. We expect estimator models to differentiate the two groups.

3.7.1. Synthetic Group Data Generation

For both group, 6 channel linear and nonlinear datasets are generated. The groups consist of 1000 subjects. For the first group, the main base of the coefficients is selected as follows:

$$A = \begin{bmatrix} 1 & 1 & -1 & 1 & -1 & 1 \\ -1 & 1 & 1 & -1 & -1 & 0 \\ 1 & 1 & -1 & 1 & 0 & 0 \\ -1 & 1 & 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{42}$$

These coefficients are scaled with variables between the interval of [0.5, 1]. The sign of the coefficients is preserved for every subject within the group.

For the second group, coefficient values from node 1 to 5, and from node 4 to 2 are deleted, and the base of coefficients are selected as follows:

$$A = \begin{bmatrix} 1 & 1 & -1 & 1 & -1 & 1 \\ -1 & 1 & 1 & 0 & -1 & 0 \\ 1 & 1 & -1 & 1 & 0 & 0 \\ -1 & 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{43}$$

These coefficients are scaled with variables between the interval of [0.5, 1]. Again, the sign of the coefficients is preserved for every subject within the group.

After the data generation, the expectation is that the estimator models should point out the differences of groups on the connections from 1 to 5, and from 4 to 2.

3.7.2. Application of the Estimator Methods

DBN, PDC and DTF methods are applied to the generated group data. The group data consist of 1000 patients and 1000 healthy controls. Application of the methods is done on MATLAB as in the Section 3.4. After the application the results are analyzed under the IBM's SPSS (Statistical Package for the Social Sciences) program.

3.7.3. Statistical Analysis and Results for Synthetic Group Data

SPSS (Statistical Package for the Social Sciences), also known as IBM SPSS Statistics, is a software package used for the analysis of statistical data.

Although the name of SPSS reflects its original use in the field of social sciences, its use has since expanded into other data markets. SPSS is commonly used in healthcare, marketing and education research.

From this program the best data test method that applies to our situation is the Mann-Whitney-U Test.

The Mann-Whitney U test is used to compare whether there is a difference in the dependent variable for two independent groups. It compares whether the distribution of the dependent variable is the same for the two groups and therefore from the same population. The test ranks all of the dependent values i.e. lowest value gets a score of one and then uses the sum of the ranks for each group in the calculation of the test statistics.

In order to apply Mann-Whitney U Test in SPSS the following steps are applied.

- From the SPSS Input window, the estimation results of fMRI data were entered, and dyscalculia and control group were selected.

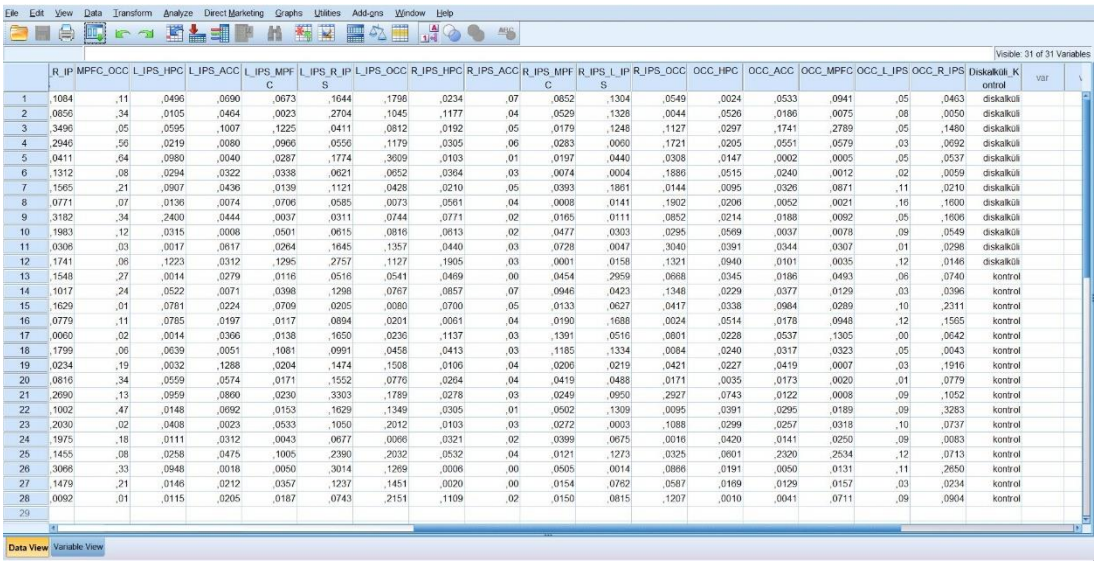

Figure 3.37. Entering the Input to SPSS

- Then, from the Analyze bar, Nonparametric Tests » Legacy Dialogs » 2 Independent Samples is selected.

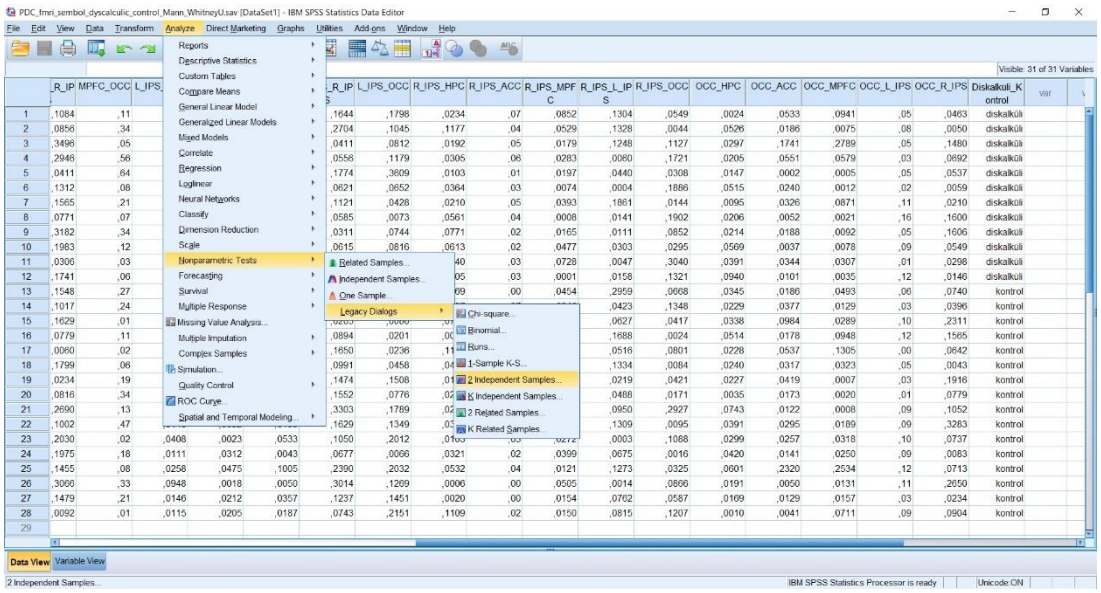

Figure 3.38. 2 Independent Samples Test Selection

- After that, from the opening window, first the connection variables are selected and are moved to the “Test Variable List”. Then, grouping condition is moved

to the “Grouping Variable” part. From the bottom, the desired test is selected, in this case Mann-Whitney U.

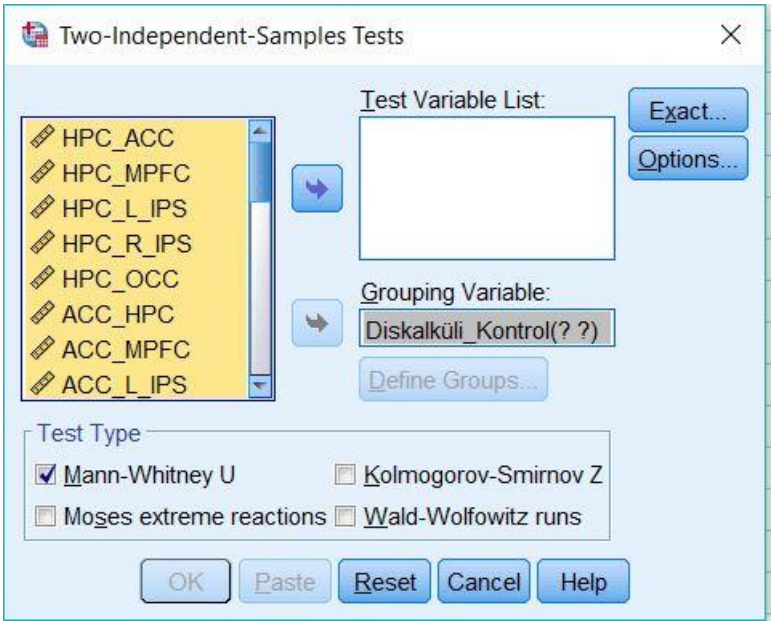

Figure 3.39. Selecting the Test Variables

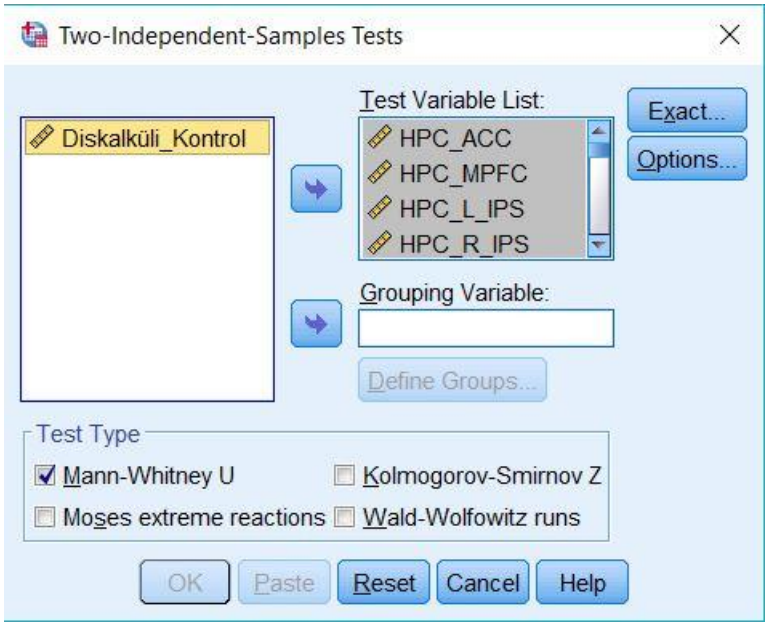

Figure 3.40. Selecting Grouping Variable

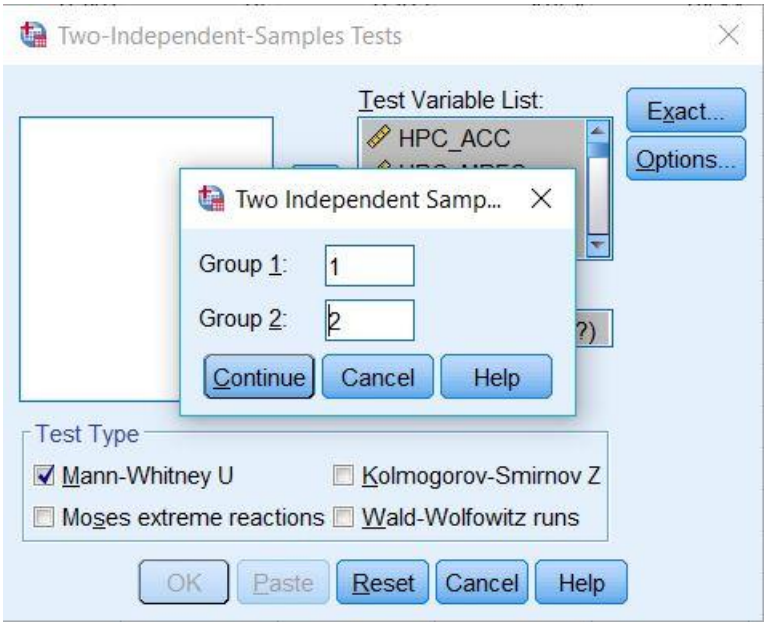

Figure 3.41. Adjusting Grouping Information and Start Test

- Mann Whitney- U Test gives the following output. From this table we are using the z-values in order to differentiate the connection differences between two groups.

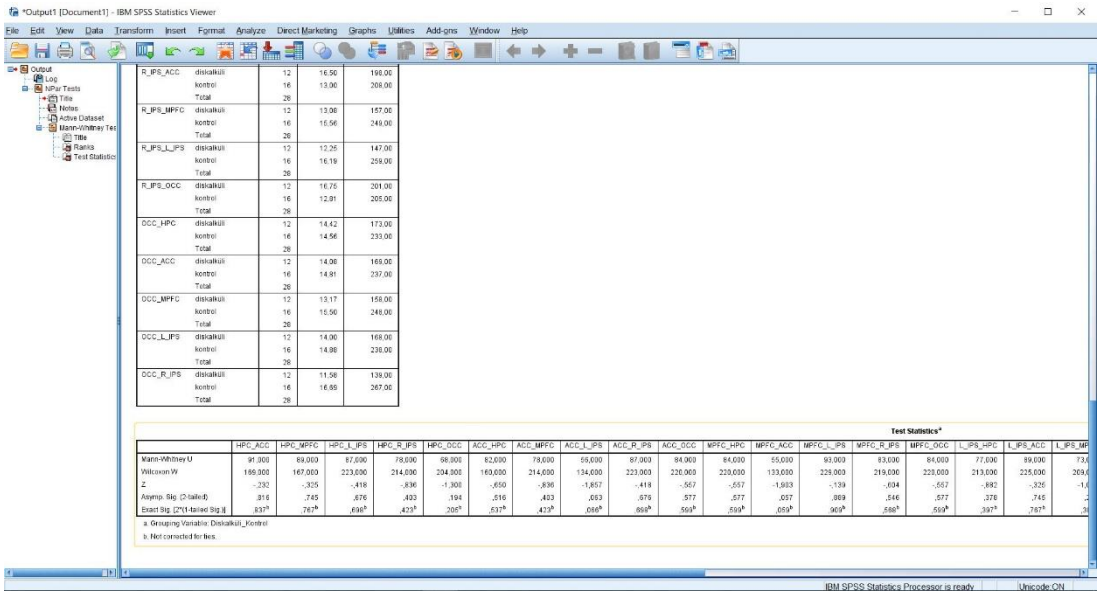

Figure 3.42. Output Window of Mann-Whitney U Test

From z values, the effect size should be calculated in order to decide how big is the difference of variables for two independent groups. An effect size can be calculated by dividing the absolute (positive) Standardized test statistic z by the square root of the number of pairs [44].

$$\text{Effect size} = \sqrt{\frac{z^2}{Total\ Number\ of\ Pairs}} \tag{44}$$

According to Cohen's classification of effect sizes which are 0.1 (small effect), 0.3 (moderate effect) and 0.5 and above (large effect), we have outlined the large effects in the following tables [45]. This value shows us whether there is a considerable difference between two groups or not.

The SPSS results for PDC, DTF, and DBN on synthetic group data can be seen as follows. The large effects are highlighted with red while medium effects are highlighted with yellow.

Table 3.6. *SPSS result for PDC on linear synthetic group data*

PDC Mann-Whitney-U Linear					
Linear Raw			Linear HRF		
Node	Effect Size	z-values	Node	Effect Size	z-values
1 → 2	0,214	-9,559	1 → 2	0,061	-2,717
1 → 3	0,018	-,790	1 → 3	0,065	-2,898
1 → 4	0,190	-8,475	1 → 4	0,078	-3,486
1 → 5	0,855	-38,224	1 → 5	0,498	-22,279
1 → 6	0,076	-3,381	1 → 6	0,066	-2,953
2 → 1	0,027	-1,224	2 → 1	0,041	-1,826
2 → 3	0,005	-,207	2 → 3	0,151	-6,755
2 → 4	0,060	-2,694	2 → 4	0,166	-7,445
2 → 5	0,023	-1,008	2 → 5	0,089	-3,970
2 → 6	0,067	-2,997	2 → 6	0,280	-12,517
3 → 1	0,108	-4,836	3 → 1	0,013	-,587
3 → 2	0,030	-1,328	3 → 2	0,143	-6,410
3 → 4	0,034	-1,524	3 → 4	0,002	-,094
3 → 5	0,173	-7,749	3 → 5	0,051	-2,267
3 → 6	0,042	-1,861	3 → 6	0,067	-3,005
4 → 1	0,268	-11,997	4 → 1	0,019	-,864
4 → 2	0,844	-37,751	4 → 2	0,435	-19,440
4 → 3	0,371	-16,571	4 → 3	0,023	-1,035
4 → 5	0,100	-4,469	4 → 5	0,217	-9,694
4 → 6	0,027	-1,223	4 → 6	0,232	-10,372
5 → 1	0,014	-,627	5 → 1	0,055	-2,472
5 → 2	0,002	-,068	5 → 2	0,133	-5,942
5 → 3	0,094	-4,212	5 → 3	0,091	-4,049
5 → 4	0,024	-1,057	5 → 4	0,110	-4,929
5 → 6	0,012	-,544	5 → 6	0,094	-4,182
6 → 1	0,096	-4,273	6 → 1	0,065	-2,904
6 → 2	0,196	-8,768	6 → 2	0,019	-,849
6 → 3	0,209	-9,367	6 → 3	0,034	-1,521
6 → 4	0,129	-5,786	6 → 4	0,044	-1,966
6 → 5	0,241	-10,776	6 → 5	0,018	-,821

Table 3.7. *SPSS result for PDC on nonlinear synthetic group data*

PDC Mann-Whitney-U Nonlinear					
Nonlinear Raw			Nonlinear HRF		
Node	Effect Size	z-values	Node	Effect Size	z-values
1 → 2	0,428	-19,126	1 → 2	0,681	-30,439
1 → 3	0,232	-10,374	1 → 3	0,222	-9,949
1 → 4	0,758	-33,890	1 → 4	0,515	-23,013
1 → 5	0,589	-26,327	1 → 5	0,238	-10,661
1 → 6	0,792	-35,406	1 → 6	0,583	-26,095
2 → 1	0,837	-37,419	2 → 1	0,532	-23,804
2 → 3	0,118	-5,275	2 → 3	0,428	-19,161
2 → 4	0,277	-12,408	2 → 4	0,237	-10,598
2 → 5	0,866	-38,720	2 → 5	0,846	-37,844
2 → 6	0,722	-32,292	2 → 6	0,218	-9,740
3 → 1	0,276	-12,342	3 → 1	0,220	-9,825
3 → 2	0,126	-5,622	3 → 2	0,051	-2,287
3 → 4	0,200	-8,925	3 → 4	0,257	-11,474
3 → 5	0,716	-32,042	3 → 5	0,120	-5,360
3 → 6	0,232	-10,390	3 → 6	0,066	-2,970
4 → 1	0,054	-2,400	4 → 1	0,147	-6,566
4 → 2	0,817	-36,523	4 → 2	0,624	-27,908
4 → 3	0,384	-17,184	4 → 3	0,509	-22,748
4 → 5	0,768	-34,324	4 → 5	0,305	-13,641
4 → 6	0,244	-10,892	4 → 6	0,020	-,914
5 → 1	0,866	-38,713	5 → 1	0,866	-38,720
5 → 2	0,326	-14,577	5 → 2	0,317	-14,182
5 → 3	0,378	-16,923	5 → 3	0,448	-20,050
5 → 4	0,832	-37,225	5 → 4	0,804	-35,946
5 → 6	0,855	-38,255	5 → 6	0,857	-38,339
6 → 1	0,099	-4,409	6 → 1	0,498	-22,263
6 → 2	0,144	-6,447	6 → 2	0,663	-29,650
6 → 3	0,138	-6,192	6 → 3	0,157	-7,006
6 → 4	0,451	-20,161	6 → 4	0,520	-23,249
6 → 5	0,866	-38,717	6 → 5	0,766	-34,259

Table 3.8. *SPSS result for DTF on linear synthetic group data*

DTF Mann-Whitney-U Linear					
Linear Raw			Linear HRF		
Node	Effect Size	z-values	Node	Effect Size	z-values
1 → 2	0,043	-1,943	1 → 2	0,045	-1,993
1 → 3	0,124	-5,550	1 → 3	0,246	-11,018
1 → 4	0,606	-27,117	1 → 4	0,238	-10,646
1 → 5	0,821	-36,694	1 → 5	0,600	-26,834
1 → 6	0,248	-11,110	1 → 6	0,261	-11,670
2 → 1	0,380	-16,996	2 → 1	0,179	-7,996
2 → 3	0,637	-28,479	2 → 3	0,302	-13,528
2 → 4	0,684	-30,581	2 → 4	0,161	-7,184
2 → 5	0,793	-35,470	2 → 5	0,401	-17,946
2 → 6	0,469	-20,954	2 → 6	0,239	-10,700
3 → 1	0,432	-19,333	3 → 1	0,062	-2,774
3 → 2	0,005	-,218	3 → 2	0,141	-6,325
3 → 4	0,217	-9,687	3 → 4	0,041	-1,821
3 → 5	0,448	-20,015	3 → 5	0,259	-11,592
3 → 6	0,395	-17,648	3 → 6	0,068	-3,046
4 → 1	0,135	-6,016	4 → 1	0,129	-5,786
4 → 2	0,821	-36,710	4 → 2	0,022	-0,993
4 → 3	0,556	-24,870	4 → 3	0,071	-3,196
4 → 5	0,823	-36,806	4 → 5	0,018	-0,819
4 → 6	0,192	-8,586	4 → 6	0,128	-5,707
5 → 1	0,108	-4,852	5 → 1	0,058	-2,610
5 → 2	0,628	-28,066	5 → 2	0,183	-8,183
5 → 3	0,024	-1,070	5 → 3	0,177	-7,923
5 → 4	0,281	-12,546	5 → 4	0,067	-3,001
5 → 6	0,076	-3,412	5 → 6	0,136	-6,085
6 → 1	0,313	-13,982	6 → 1	0,266	-11,908
6 → 2	0,063	-2,839	6 → 2	0,174	-7,775
6 → 3	0,154	-6,903	6 → 3	0,244	-10,917
6 → 4	0,391	-17,488	6 → 4	0,226	-10,118
6 → 5	0,751	-33,573	6 → 5	0,381	-17,029

Table 3.9. *SPSS result for DTF on nonlinear synthetic group data*

DTF Mann-Whitney-U Nonlinear					
Nonlinear Raw			Nonlinear HRF		
Node	Effect Size	z-values	Node	Effect Size	z-values
1 → 2	0,610	-2,797	1 → 2	0,775	-3,553
1 → 3	0,330	-1,512	1 → 3	0,676	-3,099
1 → 4	0,825	-3,780	1 → 4	0,577	-2,646
1 → 5	0,627	-2,873	1 → 5	0,808	-3,704
1 → 6	0,825	-3,780	1 → 6	0,478	-2,192
2 → 1	0,825	-3,780	2 → 1	0,808	-3,704
2 → 3	0,165	-,756	2 → 3	0,165	-,756
2 → 4	0,181	-,832	2 → 4	0,726	-3,326
2 → 5	0,825	-3,780	2 → 5	0,825	-3,780
2 → 6	0,825	-3,780	2 → 6	0,726	-3,326
3 → 1	0,247	-1,134	3 → 1	0,379	-1,739
3 → 2	0,033	-,151	3 → 2	0,297	-1,361
3 → 4	0,049	-,227	3 → 4	0,247	-1,134
3 → 5	0,577	-2,646	3 → 5	0,297	-1,361
3 → 6	0,198	-,907	3 → 6	0,181	-,832
4 → 1	0,379	-1,739	4 → 1	0,231	-1,058
4 → 2	0,825	-3,780	4 → 2	0,627	-2,873
4 → 3	0,445	-2,041	4 → 3	0,297	-1,361
4 → 5	0,610	-2,797	4 → 5	0,049	-,227
4 → 6	0,478	-2,192	4 → 6	0,264	-1,209
5 → 1	0,825	-3,780	5 → 1	0,825	-3,780
5 → 2	0,330	-1,512	5 → 2	0,577	-2,646
5 → 3	0,825	-3,780	5 → 3	0,792	-3,628
5 → 4	0,808	-3,704	5 → 4	0,825	-3,780
5 → 6	0,808	-3,704	5 → 6	0,825	-3,780
6 → 1	0,198	-,907	6 → 1	0,033	-,151
6 → 2	0,231	-1,058	6 → 2	0,561	-2,570
6 → 3	0,561	-2,570	6 → 3	0,264	-1,209
6 → 4	0,330	-1,512	6 → 4	0,528	-2,419
6 → 5	0,726	-3,326	6 → 5	0,759	-3,477

Table 3.10. *SPSS result for DBN on linear synthetic group data*

DBN Mann-Whitney-U Linear					
Linear Raw			Linear HRF		
Node	Effect Size	z-values	Node	Effect Size	z-values
1 → 2	0,249	-3,524	1 → 2	0,116	-1,644
1 → 3	0,024	-,341	1 → 3	0,044	-,628
1 → 4	0,270	-3,821	1 → 4	0,041	-,574
1 → 5	0,492	-6,955	1 → 5	0,674	-9,532
1 → 6	0,100	-1,418	1 → 6	0,020	-,286
2 → 1	0,089	-1,258	2 → 1	0,097	-1,379
2 → 3	0,164	-2,321	2 → 3	0,241	-3,413
2 → 4	0,265	-3,746	2 → 4	0,281	-3,979
2 → 5	0,418	-5,916	2 → 5	0,075	-1,066
2 → 6	0,331	-4,675	2 → 6	0,022	-,311
3 → 1	0,011	-,152	3 → 1	0,011	-,162
3 → 2	0,062	-,882	3 → 2	0,083	-1,169
3 → 4	0,136	-1,929	3 → 4	0,140	-1,977
3 → 5	0,016	-,221	3 → 5	0,234	-3,305
3 → 6	0,066	-,940	3 → 6	0,013	-,182
4 → 1	0,153	-2,164	4 → 1	0,141	-1,992
4 → 2	0,112	-1,587	4 → 2	0,065	-,915
4 → 3	0,054	-,770	4 → 3	0,014	-,195
4 → 5	0,245	-3,464	4 → 5	0,153	-2,164
4 → 6	0,039	-,545	4 → 6	0,024	-,346
5 → 1	0,232	-3,276	5 → 1	0,197	-2,784
5 → 2	0,131	-1,855	5 → 2	0,213	-3,017
5 → 3	0,040	-,564	5 → 3	0,170	-2,398
5 → 4	0,126	-1,780	5 → 4	0,022	-,305
5 → 6	0,223	-3,156	5 → 6	0,000	0,000
6 → 1	0,215	-3,037	6 → 1	0,202	-2,860
6 → 2	0,097	-1,367	6 → 2	0,024	-,335
6 → 3	0,221	-3,123	6 → 3	0,276	-3,898
6 → 4	0,240	-3,397	6 → 4	0,117	-1,652
6 → 5	0,154	-2,171	6 → 5	0,109	-1,548

Table 3.11. *SPSS result for DBN on nonlinear synthetic group data*

DBN Mann-Whitney-U Nonlinear					
Nonlinear Raw			Nonlinear HRF		
Node	Effect Size	z-values	Node	Effect Size	z-values
1 → 2	0,000	0,000	1 → 2	0,230	-3,256
1 → 3	0,000	0,000	1 → 3	0,000	0,000
1 → 4	0,000	0,000	1 → 4	0,072	-1,021
1 → 5	0,997	-14,107	1 → 5	0,676	-9,566
1 → 6	0,000	0,000	1 → 6	0,215	-3,044
2 → 1	0,071	-1,000	2 → 1	0,153	-2,166
2 → 3	0,000	0,000	2 → 3	0,027	-,384
2 → 4	0,000	0,000	2 → 4	0,160	-2,259
2 → 5	0,000	0,000	2 → 5	0,245	-3,458
2 → 6	0,000	0,000	2 → 6	0,218	-3,077
3 → 1	0,099	-1,406	3 → 1	0,000	0,000
3 → 2	0,190	-2,687	3 → 2	0,175	-2,478
3 → 4	0,000	0,000	3 → 4	0,032	-,448
3 → 5	0,000	0,000	3 → 5	0,170	-2,400
3 → 6	0,000	0,000	3 → 6	0,071	-1,008
4 → 1	0,086	-1,220	4 → 1	0,123	-1,741
4 → 2	0,988	-13,966	4 → 2	0,394	-5,573
4 → 3	0,000	0,000	4 → 3	0,111	-1,565
4 → 5	0,000	0,000	4 → 5	0,221	-3,124
4 → 6	0,000	0,000	4 → 6	0,000	0,000
5 → 1	0,040	-,568	5 → 1	0,217	-3,062
5 → 2	0,072	-1,023	5 → 2	0,291	-4,120
5 → 3	0,000	0,000	5 → 3	0,299	-4,230
5 → 4	0,000	0,000	5 → 4	0,362	-5,120
5 → 6	0,000	0,000	5 → 6	0,120	-1,701
6 → 1	0,025	-,360	6 → 1	0,178	-2,522
6 → 2	0,000	0,000	6 → 2	0,217	-3,062
6 → 3	0,000	0,000	6 → 3	0,126	-1,782
6 → 4	0,000	0,000	6 → 4	0,184	-2,600
6 → 5	0,000	0,000	6 → 5	0,183	-2,589

3.7.4. Discussion of Results for Synthetic Group Data

On group analysis, for linear data, PDC has found the desired differences of two group. However, it also shows that there is another medium difference between two groups. For linear data with HRF, it only shows both of the connection differences as medium effect. Other than that, for nonlinear data, PDC gives unreliable results for group analysis.

For linear and nonlinear data, DTF gives unreliable results for group analysis.

DBN has performed perfectly on nonlinear data, while it points out the desired connections, it also recognizes the similarities and gives the smallest effect size values for other connections, which is a sign of a good performance. On nonlinear data with HRF, DBN also finds the desired connections with only one other unrelated medium strength connection. For both of the linear cases, it finds desired connections, however; it also finds one more connection for HRF case.

The results have clearly shown that DBN performs way better than PDC and DTF on differentiating the connections of two independent groups.

-

CHAPTER 4

APPLYING TO REAL FMRI DATA

The real fMRI data used in this thesis, were collected by the researchers of the TÜBİTAK project "Beynin Sayısal İşlevleriyle İlgili Devre Modellerinin Tasarlanması ve Matematik Öğrenme Güçlüğü (Diskalkuli) Hastalık Haritasının Elde Edilmesi" under the management of Prof. Dr. Metehan Çiçek. The project number is 214S069.

Number perception is one of the perceptions that is known to exist in human brain but it is not fully explained by the researchers. Discoveries about the foundation of dyscalculia will make it easier to understand where the numeric perception occurs inside the brain.

In the dyscalculia, the mathematical success of an individual is clearly low according to its education level, age and intelligence level. The difficulty in mathematical skills can be divided into two as calculation difficulties and problem solving difficulties.

In the project, the children diagnosed with dyscalculia are compared with the healthy control group using fMRI and Diffusion Tensor Imaging (DTI). This research aims to compare the physiological differences between the dyscalculic brain and healthy brain and tries to differentiate the pathways of connections inside the brain in order to diagnose the dyscalculia in the future.

The dyscalculia sample used in this project is gathered by scanning of 2000 primay school students. After the scanning, the dyscalculia group is diagnosed by a child psychiatrist and all the other disorders are ignored. In the scanning part before the psychiatric interviews, primary school students were tested using the following three testing methods: Mathematical Success Tests (MST), Calculation Performance Tests (CPT), and Raven Test.

In the MST, the students were tested according to their skills about counting, number patterns, basic arithmetic operations and problems, and fractions. In the CPT, the students were directly asked arithmetic operations including addition, subtraction, multiplication, and division. The Raven Test aims to measure the general ability and spatial ability of the children and consists of diagram puzzles.

According to the test results the children were divided into three groups:

- Dyscalculia Group
    - Among the bottom 25% of the participants in the MST and CPT
    - Among the top 10% of the participants in the Raven Test
- Control Group
    - Among the 35%-75% group of the participants in the MST and CPT
    - Among the top 10% of the participants in the Raven Test
- Not in both
    - The children who are in the top 15% in the SNAP-IV (Swanson Nolan Pelham) scale which is a scale regarding Attention Deficit Hyperactivity Disorder (ADHD).
    - The children repeating a grade level or having a distinct failure in all classes.
    - The children diagnosed with dyslexy or ADHD in psychiatric examination
    - The children below 80 in verbal and performance IQ tests.
    - The children who had a premature birth.

4.1. fMRI Data Collection and Experiments

The fMRI data that is used in this thesis to compare the efficiency of DBN and PDC methods are collected as a two step experiment design from the TÜBİTAK project that is mentioned above. fMRI data was recorded during the Symbolic Number Comparison Test and Dot Comparison Test.

4.1.1. Symbolic Number Comparison Test

In this part, partcipants were exposed to two numbers on the right and left of the screen. The children were asked to click the left or right button on the mouse to select the bigger number. In this test, 36 number pairs were used and by interchanging their position 72 number pair were exposed to children. These pairs have the ratio of 0,5, 0,67, and 0,8.

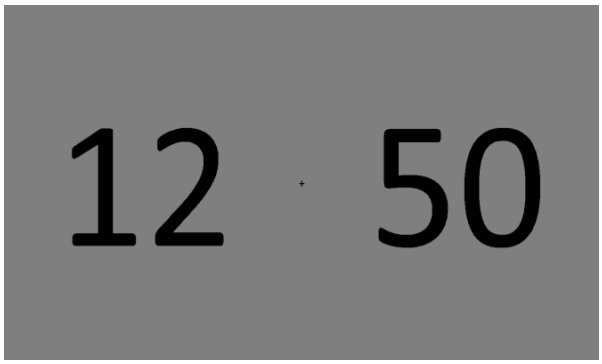

Figure 4.1. Symbolic Number Comparison Test

4.1.2. Dot Comparison Test

In this part, participants were exposed to two different numbers of dot piles on the sides of the screen. While the area that the dot piles cover were the same, number of dots in a pile was different. The children were asked to click the left or right button on the mouse to select the pile which consists bigger number of dots. In this test, 36 number pile pairs were used and by interchanging their position 72 number pile pairs were exposed to children. These pairs have the ratio of 0,5, 0,67, and 0,8.

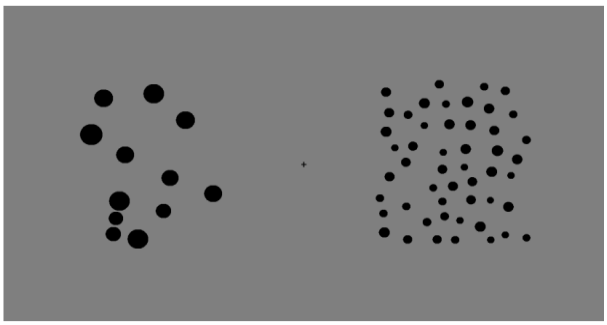

Figure 4.2. Dot Comparison Test

4.2. Application of the Estimator Methods

DBN, PDC and DTF methods are applied the collected fMRI data from the mentioned project. The fMRI data consist of 12 patients and 16 healthy controls. Application of the methods is done on MATLAB as in the Section 3.4. After the application the results are analyzed under the IBM's SPSS (Statistical Package for the Social Sciences) program.

4.3. Statistical Analysis and Results

Statistical analysis of application of three estimator methods on real fMRI data is done on SPSS. The working principle of SPSS is mentioned earlier.

The SPSS results for PDC, DTF, and DBN on real fMRI data of 12 patients and 16 healthy control group can be seen as follows.

According to Cohen's classification of effect sizes which is 0.1 (small effect), 0.3 (moderate effect) and 0.5 and above (large effect), we have outlined the large effects in the following tables [45]. This value shows us whether there is a considerable difference between two groups or not.

Table 4.1. *Mann-Whitney U Results for DBN*

DBN Mann-Whitney-U					
Symbol			Dot		
Connection	Effect Size	z-values	Connection	Effect Size	z-values
HPC→ACC	0,230	-1,219	HPC→ACC	0,070	-0,370
HPC→MPFC	0,058	-0,305	HPC→MPFC	0,066	-0,350
HPC→L_IPS	0,182	-0,964	HPC→L_IPS	0,112	-0,592
HPC→R_IPS	0,176	-0,930	HPC→R_IPS	0,032	-0,169
HPC→OCC	0,231	-1,224	HPC→OCC	0,193	-1,020
ACC→HPC	0,115	-0,609	ACC→HPC	0,226	-1,196
ACC→MPFC	0,115	-0,609	ACC→MPFC	0,030	-0,161
ACC→L_IPS	0,157	-0,832	ACC→L_IPS	0,030	-0,161
ACC→R_IPS	0,115	-0,609	ACC→R_IPS	0,045	-0,238
ACC→OCC	0,030	-0,161	ACC→OCC	0,076	-0,403
MPFC→HPC	0,096	-0,510	MPFC→HPC	0,162	-0,859
MPFC→ACC	0,607	-3,212	MPFC→ACC	0,230	-1,219
MPFC→L_IPS	0,070	-0,370	MPFC→L_IPS	0,016	-0,086
MPFC→R_IPS	0,130	-0,688	MPFC→R_IPS	0,032	-0,169
MPFC→OCC	0,000	0,000	MPFC→OCC	0,066	-0,350
L_IPS→HPC	0,130	-0,688	L_IPS→HPC	0,146	-0,771
L_IPS→ACC	0,384	-2,035	L_IPS→ACC	0,305	-1,612
L_IPS→MPFC	0,193	-1,020	L_IPS→MPFC	0,130	-0,688
L_IPS→R_IPS	0,157	-0,832	L_IPS→R_IPS	0,030	-0,161
L_IPS→OCC	0,070	-0,370	L_IPS→OCC	0,182	-0,964
R_IPS→HPC	0,083	-0,437	R_IPS→HPC	0,152	-0,806
R_IPS→ACC	0,162	-0,859	R_IPS→ACC	0,083	-0,437
R_IPS→MPFC	0,018	-0,096	R_IPS→MPFC	0,130	-0,688
R_IPS→L_IPS	0,305	-1,612	R_IPS→L_IPS	0,211	-1,118
R_IPS→OCC	0,018	-0,096	R_IPS→OCC	0,230	-1,219
OCC→HPC	0,130	-0,688	OCC→HPC	0,276	-1,461
OCC→ACC	0,276	-1,461	OCC→ACC	0,016	-0,086
OCC→MPFC	0,066	-0,350	OCC→MPFC	0,310	-1,638
OCC→L_IPS	0,305	-1,612	OCC→L_IPS	0,070	-0,370
OCC→R_IPS	0,070	-0,370	OCC→R_IPS	0,115	-0,609

Table 4.2. *Mann-Whitney U Results for PDC*

PDC Mann-Whitney-U					
Symbol			Dot		
Connection	Effect Size	z-values	Connection	Effect Size	z-values
HPC→ACC	0,044	-0,232	HPC→ACC	0,316	-1,671
HPC→MPFC	0,061	-0,325	HPC→MPFC	0,158	-0,836
HPC→L_IPS	0,079	-0,418	HPC→L_IPS	0,140	-0,743
HPC→R_IPS	0,158	-0,836	HPC→R_IPS	0,281	-1,486
HPC→OCC	0,246	-1,300	HPC→OCC	0,193	-1,021
ACC→HPC	0,123	-0,650	ACC→HPC	0,097	-0,511
ACC→MPFC	0,158	-0,836	ACC→MPFC	0,079	-0,418
ACC→L_IPS	0,351	-1,857	ACC→L_IPS	0,193	-1,021
ACC→R_IPS	0,079	-0,418	ACC→R_IPS	0,061	-0,325
ACC→OCC	0,105	-0,557	ACC→OCC	0,237	-1,253
MPFC→HPC	0,105	-0,557	MPFC→HPC	0,097	-0,511
MPFC→ACC	0,360	-1,903	MPFC→ACC	0,184	-0,975
MPFC→L_IPS	0,026	-0,139	MPFC→L_IPS	0,053	-0,279
MPFC→R_IPS	0,114	-0,604	MPFC→R_IPS	0,079	-0,418
MPFC→OCC	0,105	-0,557	MPFC→OCC	0,035	-0,186
L_IPS→HPC	0,167	-0,882	L_IPS→HPC	0,026	-0,139
L_IPS→ACC	0,061	-0,325	L_IPS→ACC	0,070	-0,371
L_IPS→MPFC	0,202	-1,068	L_IPS→MPFC	0,219	-1,161
L_IPS→R_IPS	0,123	-0,650	L_IPS→R_IPS	0,026	-0,139
L_IPS→OCC	0,000	0,000	L_IPS→OCC	0,026	-0,139
R_IPS→HPC	0,167	-0,882	R_IPS→HPC	0,044	-0,232
R_IPS→ACC	0,211	-1,114	R_IPS→ACC	0,211	-1,114
R_IPS→MPFC	0,149	-0,789	R_IPS→MPFC	0,053	-0,279
R_IPS→L_IPS	0,237	-1,253	R_IPS→L_IPS	0,149	-0,789
R_IPS→OCC	0,237	-1,253	R_IPS→OCC	0,097	-0,511
OCC→HPC	0,009	-0,046	OCC→HPC	0,158	-0,836
OCC→ACC	0,044	-0,232	OCC→ACC	0,193	-1,021
OCC→MPFC	0,140	-0,743	OCC→MPFC	0,158	-0,836
OCC→L_IPS	0,053	-0,279	OCC→L_IPS	0,009	-0,046
OCC→R_IPS	0,307	-1,625	OCC→R_IPS	0,228	-1,207

Table 4.3. *Mann-Whitney U Results for DTF*

DTF Mann-Whitney-U					
Symbol			Dot		
Connection	Effect Size	z-values	Connection	Effect Size	z-values
HPC→ACC	0,044	-0,232	HPC→ACC	0,263	-1,393
HPC→MPFC	0,061	-0,325	HPC→MPFC	0,132	-0,696
HPC→L_IPS	0,061	-0,325	HPC→L_IPS	0,202	-1,068
HPC→R_IPS	0,175	-0,928	HPC→R_IPS	0,316	-1,671
HPC→OCC	0,158	-0,836	HPC→OCC	0,097	-0,511
ACC→HPC	0,105	-0,557	ACC→HPC	0,061	-0,325
ACC→MPFC	0,167	-0,882	ACC→MPFC	0,035	-0,186
ACC→L_IPS	0,368	-1,950	ACC→L_IPS	0,088	-0,464
ACC→R_IPS	0,088	-0,464	ACC→R_IPS	0,000	0,000
ACC→OCC	0,018	-0,093	ACC→OCC	0,175	-0,928
MPFC→HPC	0,070	-0,371	MPFC→HPC	0,079	-0,418
MPFC→ACC	0,281	-1,486	MPFC→ACC	0,211	-1,114
MPFC→L_IPS	0,053	-0,279	MPFC→L_IPS	0,026	-0,139
MPFC→R_IPS	0,079	-0,418	MPFC→R_IPS	0,018	-0,093
MPFC→OCC	0,035	-0,186	MPFC→OCC	0,079	-0,418
L_IPS→HPC	0,202	-1,068	L_IPS→HPC	0,061	-0,325
L_IPS→ACC	0,123	-0,650	L_IPS→ACC	0,132	-0,696
L_IPS→MPFC	0,193	-1,021	L_IPS→MPFC	0,140	-0,743
L_IPS→R_IPS	0,097	-0,511	L_IPS→R_IPS	0,018	-0,093
L_IPS→OCC	0,009	-0,046	L_IPS→OCC	0,009	-0,046
R_IPS→HPC	0,149	-0,789	R_IPS→HPC	0,061	-0,325
R_IPS→ACC	0,202	-1,068	R_IPS→ACC	0,298	-1,578
R_IPS→MPFC	0,149	-0,789	R_IPS→MPFC	0,018	-0,093
R_IPS→L_IPS	0,237	-1,253	R_IPS→L_IPS	0,097	-0,511
R_IPS→OCC	0,149	-0,789	R_IPS→OCC	0,140	-0,743
OCC→HPC	0,044	-0,232	OCC→HPC	0,097	-0,511
OCC→ACC	0,070	-0,371	OCC→ACC	0,097	-0,511
OCC→MPFC	0,097	-0,511	OCC→MPFC	0,140	-0,743
OCC→L_IPS	0,035	-0,186	OCC→L_IPS	0,132	-0,696
OCC→R_IPS	0,298	-1,578	OCC→R_IPS	0,132	-0,696

ACC: Anterior Cingulate Cortex, (Decision making, error detection)

HPC: Hippocampus, (Memory)

MPFC: Middle Prefrontal Cortex, (Decision making)

OCC: Occipital, (Vison)

L_IPS: Left Intraparietal Sulcus (Processing numerical information)

R_IPS: Right Intraparietal Sulcus (Processing numerical information)

To summarize the Mann-Whitney U result tables, the following statements can be constructed:

- The minus sign of the z values indicates that the lack of connection for the dyscalculia patients.
- Red highlighted parts show that the results indicate great differences for that connection value between dyscalculia and control group.
- Yellow highlighted parts show that the results indicate medium differences for that connection value between dyscalculia and control group.
- As indicated in Chapter 1, it is clearly true that when exposed to a symbolic numerical operation, IPS of dyscalculia patients has shown some deficits. Also, in this experiment, DBN results clearly show that there is a lack of connection between left IPS and ACC of dyscalculic patients. ACC is the decision-making part of the brain. The lack of connection between left IPS to ACC means that there is no information going from left IPS to decide the result. This situation totally fits to Dyscalculic patients, which allows us to think that DBN can show these deficits and discriminate patients and control group. On the other hand, PDC and DTF cannot point out this known lack of connection at the patients.

- Again, in Chapter 1, it is discussed that dyscalculic patients have problems with their IPS when they are exposed to a nonsymbolic mathematical operation. In this experiment, DBN also finds the lack of connection between right IPS and occipital lobe, which also fits to Dyscalculia situation.

Finally, the following can be concluded that when applied to both patient and control group, DBN can extract differences better than PDC. This result supports our claims about DBN is better than PDC to use for real fMRI cases.

CHAPTER 5

CONCLUSION

In this thesis, the performances of Partial Directed Coherence and Dynamic Bayesian Network are compared using synthetic fMRI data. During comparison, the effect of linearity and nonlinearity of the signals was the main parameter that differentiate the two methods. Since PDC is based on linear signal exertion model MVAR, it performed better on solving the connections of linear multivariate signals. On the other hand, unlike PDC, DBN is a probabilistic method, and this gives DBN the ability to differentiate the nonlinear signals. Therefore, this causes a situation that DBN performs better on nonlinear signals.

Other than the linearity and nonlinearity of the signals, the length of the signals, network complexity, and power of white noise were the variables that are controlled during the comparison.

When the length of the signal increases, both PDC and DBN performances increase. This can be explained with the number of information for a method to solve the connections of a network. The larger the data, the bigger the information that methods can work, and hence, they reach an acceptable result. Complexity of the network is caused by the connections between different nodes. The complexity increases when the number of nodes is increased. This elevation causes PDC and DBN to perform worse in the process of reconstructing the network connectivity, because there are a lot of connections to examine.

Changing the power of noise in the linear and nonlinear data generation doesn't really affect the performances of the PDC and DBN. However, if the amplitude of the noise starts to reach the level of the signal, it distorts the signal and therefore cause models to fail on reconstruction.

In the synthetic data analysis, in every case other that linear raw data DBN outperforms PDC. In different studies where Cunlu Zou et al. (2009) and Jagath C. Rajapakse et al. (2007) also show that DBN outperforms Granger Causality which is the base method of Partial Directed Coherence. This results also support our claim. All in all, for a data length of 50000, and channel number of 6, the error values can be listed as follows:

Table 5.1. *All Error Values for 6 Ch Data with 50000 samples*

Analysis Type	*Error (%)*
PDC Linear No HRF	0
DBN Nonlinear No HRF	0,05
DBN Linear No HRF	0,27
PDC Linear HRF	0,32
DBN Nonlinear HRF	0,39
DBN Linear HRF	0,40
PDC Nonlinear HRF	0,66
PDC Nonlinear No HRF	0,70

This statement is also supported by applying both methods and DTF to a synthetic group data. In this analysis, DBN points out the differences of two independent groups way better than PDC and DTF.

At the end of the thesis, the three estimator methods are applied on real fMRI data of dyscalculic patients and control group. The lack of connections or strong connections that DBN found fits perfectly to dyscalculic patient and previous studies with dyscalculia, while PDC cannot differentiate control group and patients.

In the related studies, Gavin R. Price et al. (2013), Ruxandra Stanescu-Cosson et al. (2000), Brian Butterworth et al. (2011), have stated independently that Intraparietal Sulcus is responsible for mathematical calculation. In our study DBN finds that the connections of IPS is weaker for patient group than control group. This situation supports the efficiency of DBN on fMRI data.

To sum up, comparing DBN and PDC has shown that DBN outperforms PDC in many aspects except linearity. When applied to a real fMRI data, DBN can differentiate the patients from the control group, PDC cannot.

5.1. Future Work

In this thesis, the main purpose is to compare the performances of PDC and DBN methods in order to decide which one is useful for fMRI signals. For this purpose, the synthetic data is generated, and methods are applied to these synthetic data. In the generation process, the hemodynamic response function is convolved with the signal to make it more resembling to the real fMRI data. Since real fMRI signals already have the effect of HRF, deconvolving HRF effect from the signals, and applying connectivity estimator methods to these raw signal could also be possible. This process should definitely improve the correctness of estimated connectivity values. As stated in the Chapter 3, when HRF signal is added to the raw data, the performances of PDC and DBN decrease. Therefore, extracting HRF from the real fMRI data correctly would increase the performances, and it would be a really beneficial study for fMRI researches.

In addition to the HRF removal, a further research could be using the DBN in diagnostic purposes. In this study, it is shown that DBN can differentiate the patients from the control group based on their fMRI data. By constructing a neural network using DBN, and by training this network with fMRI data of known patients and control group, it can be possible to diagnose unknown patients. As indicated in Literature Search section, Chun-Ren Phang et al. (2019) designs a convolutional neural network on EEG data using PDC and Granger Causality and can differentiate the control group from schizophrenia patients.

While comparing PDC and DBN, we have seen that PDC performs better on linear signals and DBN performs better on nonlinear signals. Two methods have their own strengths when they are applied to different datasets. In order to create the best method which performs perfect for every type of data, we can use the strengths of these both

methods. By combining their distinct features, the ultimate effective connectivity estimator method can be proposed. However, this will require too much experiment setups, and one should deeply analyse the mathematical background of the two methods. In this thesis, the experimental approaches of the two method are covered and compared. The comparison of mathematical backgrounds would require much more. However, proposing an ultimate estimator method is a motivating challenge for a researcher.

REFERENCES

[1] O. Sporns, "Brain Connectivity," *Scholarpedia,* vol. 2, no. 10, p. 4695, 2007.

[2] C. Zou and J. Feng, "Granger causality vs. dynamic Bayesian network inference: A comparative study," *BMC Bioinformatics,* vol. 10, pp. 1-17, 2009.

[3] L. Astolfi, F. Cincotti, D. Mattia, M. G. Marciani, L. A. Baccala, F. d. V. Fallani, S. Salinari, M. Ursino, M. Zavaglia, L. Ding, J. C. Edgar, G. A. Miller, B. He and F. Babiloni, "Comparison of Different Cortical Connectivity Estimators for High-Resolution EEG Recordings," *Human Brain Mapping,* vol. 28, p. 143–157, 2007.

[4] K. J. Blinowska, "Review of the methods of determination of directed connectivity from multichannel data," *Med Biol Eng Comput,* vol. 49, pp. 521-529, 2011.

[5] M. Winterhalder, B. Schelter, W. Hesse, K. Schwab, L. Leistritz, D. Klan, R. Bauer, J. Timmer and H. Witte, "Comparison of linear signal processing techniques to infer directed interactions in multivariate neural systems," *Signal Processing,* 2005.

[6] J. C. Rajapakse and J. Zhou, "Learning effective brain connectivity with dynamic Bayesian networks," *NeuroImage,* vol. 37, no. 3, pp. 749-760, 2007.

[7] D. Y. Takahashi, L. A. Baccal and K. Sameshima, "Connectivity inference between neural structures via partial directed coherence," *Journal of Applied Statistics,* vol. 34, no. 10, pp. 1259-1273, 2007.

[8] G. Wang, Z. Sun, R. Tao, K. Li, G. Bao and X. Yan, "Epileptic Seizure Detection Based on Partial Directed Coherence Analysis," *IEEE Journal of Biomedical and Health Informatics,* vol. 20, no. 3, pp. 873-879, 2016.

[9] C.-R. Phang, C.-M. Ting, F. Noman and H. Ombao, "Classification of EEG-Based Brain Connectivity Networks in Schizophrenia Using a Multi-Domain Connectome Convolutional Neural Network," pp. 1-15, 2019.

[10] R. Warnick, M. Guindani, E. Erhardt, E. Allen, V. Calhoun and M. Vannucci, "A Bayesian Approach for Estimating Dynamic Functional Network Connectivity in fMRI Data," *Journal of the American Statistical Association,* vol. 113, no. 521, pp. 134-151, 2018.

[11] S. M. Plis, M. P. Weisend, E. Damaraju, T. Eichele, A. Mayer, V. P. Clark, T. Lane and V. D. Calhoun, "Effective connectivity analysis of fMRI and MEG data collected under identical paradigms," *Computers in Biology and Medicine,* vol. 41, no. 12, pp. 1156-1165, 2011.

[12] G. R. Price and D. Ansari, "Dyscalculia: Characteristics, Causes, and Treatments," *Numeracy,* vol. 6, 2013.

[13] R. Stanescu-cosson, P. Pinel, V. D. Moortele, D. L. Bihan, L. Cohen, S. Dehaene, S. H. Fre and S. H. Fre, "Understanding dissociations in dyscalculia: A brain imaging study of the impact of number size on the cerebral networks for exact and approximate calculation," pp. 2240-2255, 2000.

[14] B. Butterworth, S. Varma and D. Laurillard, "Dyscalculia: From brain to education," *Science,* vol. 332, no. 6033, pp. 1049-1053, 2011.

[15] A. Doyle, "Dyscalculia and Mathematical Difficulties: Implications for Transition to Higher Education in the Republic of Ireland," in *Disability Service Symposium,* Dublin, 2010.

[16] L. Kaufmann and M. Von Aster, “The Diagnosis and Management of Dyscalculia,” *Deutsches Arzteblatt International,* vol. 109, no. 45, pp. 767-778, 2012.

[17] B. Dubuc, “The Brain From Top to Bottom,” *McGill University,* pp. 1-2, 2016.

[18] A. M. Andrew, The Handbook of Brain Theory and Neural Networks, vol. 28, 1999, pp. 1084-1094.

[19] R. B. Buxton, “The physics of functional magnetic resonance imaging (fMRI),” 2013.

[20] A. Wang, T. M. Peters, S. de Ribaupierre and S. M. Mirsattari, “Functional magnetic resonance imaging for language mapping in temporal lobe epilepsy.,” *Epilepsy research and treatment,* vol. 2012, p. 198183, 2012.

[21] S. Ogawa, T. M. Lee, A. R. Kay and D. W. Tank, “Brain magnetic resonance imaging with contrast dependent on blood oxygenation (cerebral blood flow/brain metabolism/oxygenation),” 1990.

[22] A. Biasiucci, B. Franceschiello and M. M. Murray, “Electroencephalography,” *Current Biology,* vol. 29, no. 3, pp. R80-R85, 2019.

[23] K. Blinowska and P. Durka, “ELECTROENCEPHALOGRAPHY (EEG)”.

[24] M. Teplan, “Fundamental of EEG Measurement,” *Measurement Science Review,* vol. 2, no. Measurement Science Review, pp. 1-11, 2002.

[25] K. J. Friston, “Functional and Effective Connectivity in Neuroimaging: A Synthesis,” *Human Brain Mapping,* vol. 2, pp. 56-78, 1994.

[26] C. W. J. Granger, “Investigating Causal Relations by Econometric Models and Cross-spectral Methods,” *Econometrica,* vol. 37, no. 3, pp. 424-438, 1969.

[27] C. Granger, “Testing for causality: A personal viewpoint,” *Journal of Economic Dynamics and Control,* vol. 2, no. 1, pp. 329-352, 1980.

[28] H. Akaike, “A new look at the statistical model identification,” *IEEE Transactions on Automatic Control,* vol. 19, no. 6, pp. 716-723, 1974.

[29] H. Lütkepohl, “VAR Analysis in JMulTi,” in *Handbook of Economic Forecasting,* Elsevier, 2006.

[30] A. Delorme, T. Mullen, C. Kothe and e. al, “EEGLAB, SIFT, NFT, BCILAB, and ERICA: New Tools for Advanced EEG Processing,” *Computational Intelligence and Neuroscience,* no. 2011, p. 12, 2011.

[31] L. A. Baccalá and K. Sameshima, “Partial directed coherence: a new concept in neural structure determination,” *Biological Cybernetics,* vol. 84, no. 6, pp. 463-474, 2001.

[32] H. Nyquist, “Certain Topics in Telegraph Transmission Theory,” *Proceedings of the IEEE,* pp. 13-17, 1928.

[33] J. Toppi, F. D. V. Fallani, G. Vecchiato, A. G. Maglione, F. Cincotti, D. Mattia, S. Salinari, F. Babiloni and L. Astolfi, “How the Statistical Validation of Functional Connectivity Patterns Can Prevent Erroneous Definition of Small-World Properties of a Brain Connectivity Network,” *Computational and Mathematical Methods in Medicine,* no. 130985, 2012.

[34] M. J. Kaminski and K. J. Blinowska, “A new method of the description of the information flow in the brain structures,” *Biological Cybernetics,* vol. 65, no. 203, pp. 203-210, 1991.

[35] P. Dagum, A. Galper and E. Horvitz, “Dynamic Network Models for Forecasting,” in *UAI'92 Proceedings of the Eighth international conference on Uncertainty in artificial intelligence,* Stanford, 1992.

[36] E. E. Smith and S. M. Kosslyn, Cognitive psychology: Mind and brain, Upper Saddle River, N.J: Pearson/Prentice Hall, 2007.

[37] C. Bielza and P. Larrañaga, "Bayesian networks in neuroscience: a survey," *Frontiers in Computational Neuroscience,* vol. 8, no. 131, pp. 1-23, 2014.

[38] J. Burge, T. Lane, H. Link, S. Qiu and V. P. Clark, "Discrete dynamic bayesian network analysis of fMRI data," *Human Brain Mapping,* vol. 30, no. 1, pp. 122-137, 1 2009.

[39] D. Marinazzo, W. Liao, H. Chen and S. Stramaglia, "Nonlinear connectivity by Granger causality," *NeuroImage,* vol. 58, pp. 330-338, 2011.

[40] A. Habibnia, E. Rahimikia and H. Mahdikhah, "A Nonlinearity Test for Principal Component Analysis," *MATLAB Central File Exchange,* 2018.

[41] U. Kruger, D. Antory, J. Hahn, G. Irwin and G. McCullough, "Introduction of a nonlinearity measure for principal component models," *Computers & Chemical Engineering,* vol. 29, no. 11-12, pp. 2355-2362, 2005.

[42] J. Rodrigues and A. Andrade, "Synthetic neuronal datasets for benchmarking directed functional connectivity metrics," *PeerJ,* vol. 3, p. e923, 2015.

[43] A. Schlogl and C. Brunner, "BioSig: A Free and Open Source Software Library for BCI Research," *Computer,* vol. 41, no. 10, pp. 44-50, 10 2008.

[44] M. Gilchrist and P. Samuels, "Mann-Whitney U test (Non-parametric equivalent to independent samples t-test)," pp. 1-32, 2015.

[45] H. Chen, P. Cohen and S. Chen, "How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies," *Communications in Statistics: Simulation and Computation,* vol. 39, no. 4, pp. 860-864, 2010.