Extensions of Multivariate Dynamical Systems for Simultaneous Explanations of Neural and Behavioral Data

Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Master of Arts in the Graduate School of The Ohio State University

By

Qingfang Liu, B.S.

Graduate Program in Psychology

The Ohio State University

2018

Master’s Examination Committee:

Brandon M. Turner, Advisor

Alexander A. Petrov

Zhong-Lin Lu
© Copyright by

Qingfang Liu

2018
Abstract

A recent approach to study how brain performs cognitive functions and produces overt behaviors is to model behavioral measures and neural measures simultaneously. However, current simultaneous modeling studies fail to provide a whole picture of underlying cognitive mechanisms, because they often do not consider the important temporal dynamics of the neural measures and ignore the brain functional networks for completing cognitive tasks. The purpose of this thesis is to investigate and extend the Multivariate dynamical systems (MDS; Ryali et al., 2011) to model both behavioral and neural measures simultaneously. The MDS framework is constructed in a way such that the temporal dynamics and brain functional connectivities are explicitly contained in the model structures. If such an extension is successful, it could facilitate the development of completely integrative systems of brain-behavior relationships. Here we explore the ability of MDS to account for patterns of neural and behavioral data in perceptual and consumer decision making tasks and examine the ability of current algorithms to estimate the parameters within these MDSs. Two simulation studies with the perceptual decision making MDS model illustrate the ability of the model to capture key characteristics of both behavioral and neural measures of experimental data. Another simulation study with the consumer decision making MDS model shows that the model is able to differentiate between neural theories by generating different predictions of combined behavioral and neural data.
A parameter recovery study suggests the identifiability of the model parameters with a likelihood approximation algorithm. We close the thesis with a discussion about limitations, applications, and future directions.
I dedicate this thesis to my mother and father.
I would like to thank my advisor Dr. Brandon M. Turner. He has been super supportive academically and emotionally through the rough road to finish this thesis and every day of my grad-school life. He consistently allowed this thesis to be my own work, but steered me in the right direction whenever I needed it. I would like to thank my committee members Dr. Alexander A. Petrov for insightful ideas to improve the dynamical system and Dr. Zhong-Lin Lu for practical suggestions to implement this framework. I appreciate the motivation from Abraham Maslow, to always make the choice in my life towards the direction of self-growth.
Vita

2017-present .......................... Graduate Teaching Associate,
                      The Ohio State University.

2016-2017 ............................ Distinguished University Fellowship,
                      The Ohio State University.

2016 ................................. B.S. Psychology,
                      Beijing Normal University, China

Publications

Research Publications

Turner, B. M., Rodriguez, C. A., Liu, Q., Molloy, M. F., Hoogendijk, M., and Mc-
Clure, S. M. “On the Neural and Mechanistic Bases of Self-Control”. Cerebral Cortex,
1–19, 2018.

*Molloy, M. F., *Galdo, M., Bahg, G., Liu, Q., and Turner, B. M. “Whats in a
Response Time?: On the Importance of Response Time Measures in Constraining
Models of Context Effects”. Decision, in press.

Fields of Study

Major Field: Psychology
Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iv</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>v</td>
</tr>
<tr>
<td>Vita</td>
<td>vi</td>
</tr>
<tr>
<td>List of Tables</td>
<td>ix</td>
</tr>
<tr>
<td>List of Figures</td>
<td>x</td>
</tr>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Joint modeling</td>
<td>2</td>
</tr>
<tr>
<td>1.1.1 Integrative approach</td>
<td>5</td>
</tr>
<tr>
<td>1.2 Brain functional integration</td>
<td>7</td>
</tr>
<tr>
<td>1.3 Multivariate Dynamical Systems</td>
<td>9</td>
</tr>
<tr>
<td>1.4 Connection to neural measures</td>
<td>14</td>
</tr>
<tr>
<td>1.4.1 Balloon model</td>
<td>14</td>
</tr>
<tr>
<td>1.4.2 Convolution method</td>
<td>16</td>
</tr>
<tr>
<td>1.5 Explaining behavioral measures</td>
<td>17</td>
</tr>
<tr>
<td>1.5.1 Behavioral DCM</td>
<td>18</td>
</tr>
<tr>
<td>1.5.2 Response time</td>
<td>20</td>
</tr>
<tr>
<td>1.6 Plan of the thesis</td>
<td>20</td>
</tr>
<tr>
<td>2. An MDS Model for Perceptual Decision Making</td>
<td>21</td>
</tr>
<tr>
<td>2.1 Neural correlates underlying perceptual decision making</td>
<td>21</td>
</tr>
<tr>
<td>2.2 Model structure</td>
<td>25</td>
</tr>
<tr>
<td>2.3 Motion detection with unequal coherence</td>
<td>31</td>
</tr>
</tbody>
</table>
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 The $p_L$ condition levels and the corresponding $p_R$ levels and net coherence levels used in the simulation for motion detection with unequal coherence. $p_L$: probability of leftward moving dots. $p_R$: probability of rightward moving dots. The net coherence is the absolute value of difference between $p_L$ and $p_R$.</td>
<td>32</td>
</tr>
<tr>
<td>4.1 Parameter summary: description and value of the parameters in the perceptual decision making MDS model.</td>
<td>71</td>
</tr>
<tr>
<td>4.2 Models with different sets of free parameters.</td>
<td>72</td>
</tr>
</tbody>
</table>
List of Figures

Figure | Page
--- | ---
1.1 An illustration of the three joint modeling approaches for linking neural and behavioral data. N represents the neural data, B represents the behavioral data. $\theta$, $\delta$ and $\Omega$ represent model parameters. Figure extracted from Palestro et al. (2018). | 5

1.2 A brain network example with 3 different regions ($R_1$, $R_2$, $R_3$). Each circle node indicates a brain region. The arrow indicates the direction of the connectivity. An arrow pointed towards the region where it starts from indicates a self-connection within the region. An arrow pointed from one region to another indicates a directional inter-region connection where the direction of the arrow refers to the direction of the connection. Numbers listed along the arrows indicate the strength of the connection. The strength could be positive or negative where positive number means an excitatory effect and negative number means an inhibitory effect. | 11

1.3 The biological assumptions in the Balloon model (Friston et al., 2000; Buxton et al., 1998; Mandeville et al., 1999) used by the DCM. The neuronal model (1) reflects an excitatory - inhibitory balance, which induces the feedforward neurovascular coupling (2) and causes changes in blood inflow. Changes in blood inflow are accompanied by changes in blood outflow, blood volume and deoxyhemoglobin content (3). The rate of blood outflow is influenced by the change in the blood volume. Changes in blood volume and deoxyhemoglobin content are then reflected in the BOLD response (4). Figure extracted from Figure 1 in Havlicek et al. (2015) | 15
2.1 A trial example in a random dot motion task. The monkey decides the net direction of motion, here shown as up versus down. Task difficulty is governed by the fraction of dots that move coherently from one movie frame to the next. The monkey makes a second, voluntary movement to a choice target to indicate his decision. Figure extracted from Gold and Shadlen (2007).

2.2 The cortex-basal ganglia-superior colliculus network in a perceptual decision making task. The caudate nucleus receives inputs from both FEF and LIP. Its output is split along direct and indirect pathways, which are thought to have facilitatory and inhibitory effects, respectively, on behavior. These pathways converge in the substantia nigra, pars reticulata (SNr), which sends the output of the basal ganglia to the SC and, via the thalamus, back up to cortex. The dashed black line represents feedback pathways. Abbreviations: FEF, frontal eye field; GPi and GPe, the internal and external segments of the globus pallidus; LIP, lateral intraparietal area of the parietal cortex; SC, superior colliculus; SNC, substantia nigra pars compacta; SNr, substantia nigra pars reticulata; STN, subthalamic nucleus; VTA, ventral tegmental area. Figure extracted from Ding and Gold (2013).

2.3 A proposed MDS model for explaining neural and behavioral data from a perceptual decision making task. \( U_L \) and \( U_R \) represent the visual inputs for leftward and rightward moving dots, respectively. \( R_1 \) and \( R_2 \) represent the visual cortex neurons that encode the visual inputs of moving dots. \( R_3 \) and \( R_4 \) represent the neurons in regions such as the frontal eye field (FEF) and the lateral intraparietal area (LIP) thought to accumulate evidence for leftward and rightward dots independently. \( R_5 \) is assumed to be the output nuclei of basal ganglia and \( R_6 \) is assumed to be presupplementary motor area (pre SMA). \( U_5 \) provides a constant input for \( R_5 \) and \( R_5 \) inhibits the activation of \( R_6 \). Once the absolute difference of accumulated evidences in \( R_3 \) and \( R_4 \) reaches a threshold, \( R_3 \) and \( R_4 \) jointly inhibit \( R_5 \) so that \( R_6 \) gets disinhibited. \( R_6 \) is assumed to accumulate evidence for the response options, eventually passing an excitatory signal to initiate a movement.

2.4 An example showing the states of activation of 6 hypothetical brain regions in Figure 2.3. Regions are colored corresponding to the nodes in Figure 2.3.
2.5 Predicted choice and response time distribution in the simulation of motion detection with unequal coherence at each $p_L$ condition. At each condition, the response time associated with a left choice is plotted in the negative axis while the response time associated with a right choice is plotted in the positive axis. .......................... 36

2.6 The summary of behavioral data from simulation of motion detection with unequal coherence (left) and simulation of motion detection with equal coherence (right). In the left plot, the mean response time at each $p_L$ level is plotted using the barplot. Error bars are included considering the number of simulation (1000) at each $p_L$ level. Accuracy is defined as correctly choosing the direction with more moving dots and is plotted with the dotline. As a reference, when $p_L$ is .5, choosing right direction is defined as the correct choice. In the right plot, the mean response time at each $p_L$ level is plotted using the barplot. ... 38

2.7 Predicted BOLD activities in brain region $R_3$, $R_4$, and $R_6$, for each of $p_L = .1$ to .9 levels. Activities are averaged across 1000 simulations. 40

2.8 Predicted choice and response time distribution in the simulation of motion detection with equal coherence at each $p_L$ condition. At each condition, the response time associated with a left choice is plotted in the negative axis while the response time associated with a right choice is plotted in the positive axis. .......................... 43

2.9 Predicted BOLD activities in brain region $R_3$, $R_4$, and $R_6$, for each of $p_L = .1$, .2, .3, .4 and .5 levels. Activities are averaged across 1000 simulations. .......................... 44

3.1 Networks implicated in different component processes of temporal discounting: cognitive control (blue), reward valuation (red) and imagery or prospection (green). vmPFC and posterior cingulate cortex are involved in both prospection and valuation. Figure extracted from Peters and Büchel (2011). .......................... 50

3.2 A proposed MDS model scheme for the intertemporal choice task, modified from the perceptual decision making task model in Figure 2.3. The added $R_7$ indicates self-control related brain regions. $R_1$ and $R_2$ stand for the valuation regions. $R_3$ and $R_4$ stand for accumulation regions. $SV_{SS}$ and $SV_{LL}$ represent the subjective values of the smaller sooner option and the lager later option, respectively. .......................... 53
3.3 Predicted behavioral data from different model configurations in the intertemporal choice MDS model. The left panel shows how the proportion of choosing LL choice varies across different $P_{LL}$ levels. The right panel shows how the mean response time varies across different $P_{LL}$ levels. Different colors are for different model configurations, indicated by the legend in the left panel.

3.4 Comparison of predicted behavioral data across different $P_{LL}$ levels (.2, .5 and .8) between M1 and M2, and between M1 and M3. The response time associated with a LL choice is plotted in the negative axis and the response time associated with a SS choice is plotted in the positive axis. The colors used in the histogram plots indicate the corresponding model configurations (M1, M2 and M3) and are consistent with the colors used in the plot title.

3.5 The predicted BOLD activity in the valuation regions and the accumulation regions from the null model M0, at different $P_{LL}$ conditions across .2, .5 and .8.

3.6 The predicted BOLD activity in the valuation regions from model M1, M2, M3 and M4, for each of $P_{LL} = .2$, $P_{LL} = .5$ and $P_{LL} = .8$ conditions.

3.7 The predicted BOLD activity in the accumulation regions from model M1, M2, M3 and M4, for each of $P_{LL} = .2$, $P_{LL} = .5$ and $P_{LL} = .8$ conditions.

4.1 The parameter estimation result of model $V_1$, with parameter $a_0$ and $a_1$ freely estimated. Red vertical lines in the histograms represent the “true” parameter values used for the experimental data whereas the green vertical lines represent the means of the posterior estimates. The blue dash lines in the histograms indicate the prior densities assumed for each parameter. The plot with “log likelihood” against “Iteration” indicates the change of log likelihoods along with the iteration (after burn-in period) for each of 24 chains (shown in different colors). The right bottom panel shows the joint distribution of $a_0$ and $a_1$. The red cross represents the “true” parameter values. The green circle represents the means of the posterior estimates.
4.2 The parameter estimation result of model \( V_2 \), with parameter \( a_0 \) and \( a_2 \) freely estimated. Red vertical lines in the histograms represent the “true” parameter values used for the experimental data whereas the green vertical lines represent the mean of the posterior estimates. The blue dash lines in the histograms indicate the prior densities assumed for each parameter. The plot with “log likelihood” against “Iteration” indicates the change of log likelihoods along with the iteration (after burn-in period) for each of 24 chains (shown in different colors). The right bottom panel shows the joint distribution of \( a_0 \) and \( a_2 \). The red cross represents the “true” parameter values. The green circle represents the averages of the posterior estimates.

4.3 The parameter estimation result of model \( V_3 \), with parameter \( a_0 \), \( a_1 \) and \( a_2 \) freely estimated. Red vertical lines in the histograms represent the “true” parameter values used for the experimental data whereas the green vertical lines represent the means of the posterior estimates. The blue dash lines in the histograms indicate the prior densities assumed for each parameter. The plot with “log likelihood” against “Iteration” indicates the change of log likelihoods along with the iteration (after burn-in period) for each of 24 chains (shown in different colors). The bottom three plots show the pairwise joint distributions of \( a_0 \), \( a_1 \) and \( a_2 \). The red crosses represent the “true” parameter values and the green circles represent the means of the posterior estimates.

4.4 The parameter estimation result of model \( V_4 \), with parameter \( a_0 \), \( a_1 \), \( a_2 \) and \( a_3 \) freely estimated. Red vertical lines in the histograms represent the “true” parameter values used for the experimental dataset whereas the green vertical lines represent the means of the posterior estimates. The blue dash lines in the histograms indicate the prior densities assumed for each parameter. The bottom two rows show the pairwise joint distributions of \( a_0 \), \( a_1 \), \( a_2 \) and \( a_3 \). The red crosses represent the “true” parameter values and the green circles represent the means of the posterior estimates.
Chapter 1: Introduction

The rapid development of brain measurement techniques such as functional magnetic resonance imaging (fMRI), electroencephalography (EEG), magnetoencephalography (MEG), positron-emission tomography (PET) and single-unit recording have contributed substantial insights into the neural correlates of human information processing and cognitive operations in cognitive neuroscience. However, traditional cognitive neuroscience is faced with two major challenges. First, cognitive neuroscientists typically use statistical models to relate the independent variables of an experiment to the measurements of brain activity, as a way of understanding what parts of the brain are systematically related to variables in the experiment. This result can be further used to infer what brain areas, neural processes, and circuits are involved in a particular cognitive process (Forstmann and Wagenmakers, 2015). These analyses provide limited insight to computational mechanisms about how a cognitive function is performed in the brain regions, though these mechanisms are actually of vital importance to understanding cognition. Cognition is about explaining how a cognitive function is operated in a precise and unambiguous manner, so that we are able to predict the brain activity before observing it, or explain the reason of cognitive function deficits. Second, the traditional cognitive neuroscience approach has been devoted to localizing specific brain functions. For example, certain brain regions (e.g.
V5 or middle temporal (MT)) are thought to play major roles in processing visual motion (Maunsell and Van Essen, 1983; Vanduffel et al., 2001), in that the neurons in these regions significantly predict the direction decisions (Gold and Shadlen, 2007). However, the completion of any cognitive function is thought to require the collaboration of a series of functionally segregated brain functions. In the case of visual motion processing, the completion of the function also relies on some basic cortical or subcortical functions such as the basal ganglia to either inhibit the motion impulse or execute the motion command (Hikosaka et al., 2000a; Lo and Wang, 2006). From the late 20th century, researchers started to pay attention to the distributed interactions among functionally segregated regions, coined as functional integration (Friston, 2009). The first challenge gave rise to the cutting-edge field of joint modeling (Turner et al., 2017b; Palestro et al., 2018), and the second challenge brought about the field of brain functional integration, or functional connectivity. While each of these advancements solved specific aspects of the two aforementioned limitations, a better approach would be an integrated solution to solve both limitations simultaneously. The goal of this thesis is to develop an integrated solution to this problem. In this chapter, I first review the joint modeling approach and the functional integration approach. I then propose a new integrated solution that borrows aspects of these two different approaches.

1.1 Joint modeling

The joint modeling approach combines cognitive neuroscience with the field of mathematical psychology. Whereas cognitive neuroscience relies on the relationship between experimental stimuli and dynamics of neural data to localize mental processes...
in the brain, mathematical psychology is concerned with understanding underlying mechanisms of how mental processes take place through the analysis of behavioral data. The necessity of incorporating these two fields can be appreciated according to Marr’s three levels of analysis (Marr, 1982): computational, algorithmic and implementational. The computational level is about what goal the system is intended to achieve. The algorithmic level explains how to achieve the goal, such as what representations to use. The implementational level tells physically how the representations could be realized. Under this perspective, mathematical psychologists work on the computational and algorithmic level whereas cognitive neuroscientists work on the implementational level. As many current authors have argued (Love, 2015; Turner et al., 2017b), a better understanding of the complete cognitive system could be realized if all aspects of the system were analyzed together.

Joint modeling is one successful approach to simultaneously connect neural data with behavioral data (Turner et al., 2017b; Palestro et al., 2018). It is coined “joint modeling” due to the theoretical position of applying equal importance to neural and behavioral data. Of course, many other approaches have been developed that place unequal importance on neural and behavioral data, but these approaches are beyond the scope of this thesis (see Turner et al. (2017b) for details).

The gist of the joint modeling is to specify the joint distribution of the neural measures (N) obtained from cognitive neuroscience techniques and behavioral measures (B). Three types of joint models are illustrated in Figure 1.1 via graphical diagrams, where observed variables (N and B) are shown as filled square nodes, and unknown model parameters are shown as empty circles. Paths between the nodes in the graph
indicate dependency among the nodes, where the node being pointed at depends on the node from which the arrow originates.

The left panel of Figure 1.1 shows the “Directed” approach (Palestro et al., 2018; e.g., Cavanagh et al., 2011; Frank et al., 2015; Nunez et al., 2015; Nunez et al., 2017). The Directed approach uses a set of parameters $\delta$ to describe the functional properties of the neural data $N$ through some statistical models, and uses another set of parameters $\theta$ to describe how the behavioral data $B$ are generated from a cognitive model. The word “directed” is illustrated by the modulatory influence that $\delta$ has on $\theta$. The middle panel of Figure 1.1 shows the “Covariance” approach (Palestro et al., 2018; e.g., Turner, 2015; Turner et al., 2013; Turner et al., 2015; Turner et al., 2016; Turner et al., 2017). The Covariance approach represents the functional properties of neural data $N$ and the behavioral data $B$ through “submodels” using parameters $\delta$ and $\theta$, respectively. However, the Covariance approach differs in that it assumes an overarching distribution governed by a set of parameters $\Omega$ that specifies how these parameters interact with one another.

The Covariance approach and the Directed approach are conceptually similar as they both have sets of parameters $\delta$ and $\theta$ to describe functional properties of neural and behavioral data, respectively. Further, the behavioral model parameters $\theta$ and the neural model parameters $\delta$ are connected through a statistical constraint, either through a direct force (Directed approach) or an overarching distribution (Covariance approach).

The most relevant joint modeling approach to be discussed in this thesis is the “Integrative approach”. Different from the previous two approaches, the Integrative approach attempts to develop a single cognitive model capable of predicting both
neural and behavioral measures (Turner et al., 2017b; Palestro et al., 2018). This is shown in the right panel of Figure 1.1: the neural data N and the behavioral data B are explained together through a single set of parameters $\theta$, indicated by the connections from $\theta$ to both N and B. We shall talk in more details about the Integrative approach next, as it is the approach to be applied in this thesis.

1.1.1 Integrative approach

In the Integrative approach, there is only one single model for explaining behavioral and neural data, without separable submodels. As is mentioned earlier, the purpose of combining the cognitive neuroscience and the mathematical psychology is to explain the cognitive processes at all three levels, computational, algorithmic and implementational (Marr, 1982). It means that the single model with a set of parameters $\theta$ should be equipped with sufficient explanatory power to simultaneously explain all variables. It has greater theoretical significance than the other two joint
modeling approaches, since every assumption and every parameter setting should affect the prediction of behavioral and neural data in an explicit way, but in the other two approaches, the existence of submodels might obscure the relation between the submodel specifications and the overall predictive performance. As the linking function relating the two submodels is often purely a statistical one, it can be viewed as lacking theoretical commitment, and so it may have poor generalizability. One could argue that in an absolute sense, the Integrative approach is the most theoretically committed type of joint model.

Despite the strong theoretical benefits of integrative joint models, there are at least three difficulties that merit consideration. First, behavioral and neural data are often on different measuring units (Turner et al., 2017b; Palestro et al., 2018). For example, fMRI neuroimaging data are in the form of time series with temporal resolution at second (depending on the repetition time of pulse sequence in the MRI scanner) and EEG data are usually with millisecond-range temporal resolution. Yet, behavioral data such as the response choice and response time are measured on the trial level. Second, strong commitments are required in this approach to infer both what underlying cognitive processes are involved and where these processes arise in the brain (Palestro et al., 2018). A complete understanding of the underlying cognitive processes at the computational and algorithmic levels requires integrating results from many of previous research studies and reconciling conflicting findings. An implementational level of knowing where these processes arise requires knowledge of how a large brain functional network is related to a cognitive process, which is the topic of Section 1.2. Third, from the model development perspective, if we want to predict both behavioral and neural data from one single model, the single model is expected
to be complex. As the model is complex, fitting the model to data and testing its assumptions is expected to be computationally demanding. For example, because the model is stochastic, specifying the likelihood function is difficult or impossible, and so inferential methods such as maximum likelihood estimation and Bayesian statistics cannot be applied.

1.2 Brain functional integration

Functional integration can be contrasted with functional localization. Whereas functional localization implies that a function can be localized in a cortical area, functional integration speaks to the distributed interactions among functionally segregated regions. The studies of functional localization date back to 19th century when lesion and electrical stimulation paradigms suggested that functions could indeed be localized in the brain (Doty, 1969; Fritsch, 1870; Simonoff, 1866). However, it was not until late 20th century that people started to realize that it was inappropriate to attribute a specific function to a cortical area, given the dependence of cerebral activity on the anatomical connections between distant brain regions (Phillips et al., 1984).

First, the electrical stimulation paradigm in dogs and monkeys supports the idea that there is a relationship between distinct brain regions and specific functions. However, those studies have often overlooked that the elicited movements from electrical stimulation might have originated in related pathways rather than the stimulated regions, because the current from the stimulation may have spread to distant regions (Friston, 2011). Therefore, the electrical stimulation paradigm is not conclusive to
infer functional localization because localizationism discounted interactions or functional integration among different brain areas (Friston, 2011). This problem could be more severe in human fMRI or EEG studies, because the spatial resolutions of fMRI or EEG are worse than that of the electrical stimulation (Dale et al., 2000; Grill-Spector et al., 2006).

Second, lesion studies help identify functional localization through investigating behavioral deficits of patients with focal brain lesions (Godefroy et al., 1998). However, the interpretations can be partially confounded by functional recovery induced by brain plasticity mechanisms (e.g., Duffau et al., 2003). Functional recovery means that different brain areas could perform the same function or yield the same output after some brain regions get damaged (Edelman and Gally, 2001; Friston and Price, 2003; Price and Friston, 2002).

In face of the challenge from brain functional integration, a few methods have been developed to study the brain functional connectivity. Psychophysiological interaction (PPI; Friston et al., 1994, 1997) relies on a regression model to describe modulatory interactions between brain regions. Structural equation modeling (SEM; Büchel and Friston, 1997; Bullmore et al., 2000; Horwitz et al., 1999) combines an anatomical model and the inter-regional covariances of activity to represent the influence of regions on each other through the putative anatomical connections. Granger causal modeling (GCM; Goebel et al., 2003; Valdes-Sosa, 2004) uses vector autoregressive modeling of fMRI time series to look for correlations in the activity of two or more regions. Dynamic causal modeling (DCM; Friston et al., 2003) describes how external stimuli induce changes in hidden neuronal population dynamics that cause variability in the measured neural responses. Each method has its own specific set of
assumptions and benefits and different methods are often framed to address slightly different questions. A full review can be obtained in Smith et al. (2011).

The goal of this thesis is to construct a new framework based on the Multivariate Dynamical Systems (MDS; Ryali et al., 2011). This framework is able to simultaneously predict behavioral data and neural measures of a brain functional integration networks. By constructing this framework, we hope to achieve the following four goals. First, the new framework should provide insights to help promote the understanding of how a large brain network produces an overt behavior. Because the framework is able to explain both neural and behavioral data, it is better situated to explain how the brain relates to behavior. Second, the framework should have a better predictive performance in terms of the combined neural and behavioral data, because it does not rely on the statistical link between two submodels in the other joint modeling approaches. This better predictive performance is an advantage over behavioral-only or neural-only models, and also over the Directed and the Covariance approaches in the joint modeling. Third, this framework provides generative models of brain function, and thus it is invariant to the type of neural measure that is being collected. As such, it might be useful to integrate different modalities such as EEG and fMRI (Turner et al., 2013c, 2016). Fourth, from the model development perspective, the model parameters under this framework should be better constrained, since they are informed from two data streams (Turner et al., 2016).

1.3 Multivariate Dynamical Systems

In general, MDS models are designed to capture unobserved states by specifying a “state equation”. The observed data are modeled as a function of the unobserved
states by specifying an “observation equation”. In the case of neural data, the unobserved states are the neuronal activations in each of M brain regions of interest. We denote the activation for the ith region as \( R_i (i = 1, \cdots, M) \). \( S \) is an \( M \times T \) matrix, and each column of \( M \times 1 \) vector \( S(t) \) is an unobserved state variable representing neuronal activations at time \( t \) in each of \( M \) brain regions of interest. The state equation expresses \( S(t) \) as a function of three components\(^1\) in the following way:

\[
S(t) = CS(t - 1) + DU(t) + \omega(t). \tag{1.1}
\]

First, \( C \) represents the endogenous brain connectivity, independent of any experimental manipulations. \( C \) is an \( M \times M \) matrix showing the strengths of endogenous brain connectivity, so this matrix is compromised with \( M^2 \) connectivity parameters. We might think of a brain network example with three regions \( (R_1, R_2, R_3) \) shown in Figure 1.2, where each circle node indicates a brain region and the arrow indicates the direction of the connectivity. An arrow pointed towards the region where it starts from indicates a self-connection within the region. An arrow pointed from one region to another indicates a directional inter-region connection where the direction of the arrow refers to the direction of the connection. Numbers listed along the arrows indicate the strength of the connection where positive number means an excitatory effect and negative number means an inhibitory effect. This particular brain network has a \( C \) matrix of the form

\(^1\)A bilinear term in the original MDS is intentionally left out because it is not used in the current thesis (see Ryali et al. (2011) for details).
Figure 1.2: A brain network example with 3 different regions \((R_1, R_2, R_3)\). Each circle node indicates a brain region. The arrow indicates the direction of the connectivity. An arrow pointed towards the region where it starts from indicates a self-connection within the region. An arrow pointed from one region to another indicates a directional inter-region connection where the direction of the arrow refers to the direction of the connection. Numbers listed along the arrows indicate the strength of the connection. The strength could be positive or negative where positive number means an excitatory effect and negative number means an inhibitory effect.
\[
C = \begin{pmatrix}
.9 & .7 & 0 \\
.3 & .9 & 0 \\
-.2 & 0 & .9
\end{pmatrix},
\]

where \(C[1, 1] = C[2, 2] = C[3, 3] = .9\) indicates a strong positive connection among all three brain regions. \(C[1, 2] = .7\) indicates a strong positive connectivity from \(R_2\) to \(R_1\). \(C[2, 1] = .3\) indicates a weak positive connectivity from \(R_1\) to \(R_2\). Notice that \(C[1, 2]\) does not necessarily equal to \(C[2, 1]\) since the connectivity strength can differ depending on the direction. \(C[3, 1] = -.2\) indicates a negative connectivity from \(R_1\) to \(R_3\).

Second, the term \(DU(t)\) indicates the direct exogenous effect on \(S(t)\). \(U\) is an \(M \times T\) matrix with (possible) inputs for each brain region (i.e. each row of \(U\)). Each row of \(U\) can be constant across total time points \(T\) but it can also vary to represent temporal fluctuations of the stimuli. \(U(t)\) is an \((M \times 1)\) vector that represents the strength of the external inputs at time point \(t\) and is mainly affected by the experimental stimuli property. \(D\) is an \(M \times M\) diagonal matrix wherein \(D(i, i)\) denotes external stimuli strength to \(R_i\). By specifying a diagonal matrix, each external stimuli (each row of \(U\)) are constrained to affect only on one brain region. Altogether, this term tells nothing about the connectivity within brain regions and is thus conceptually equivalent to the activation coefficients in the traditional fMRI general linear model.

Third, \(\omega(t)\) is an \(M \times 1\) state noise vector following an multivariate normal distribution with \(\omega(t) \sim N_M(0, Q)\). \(Q\), the variance covariance matrix, can be a diagonal matrix if the noises in different brain regions are uncorrelated. Non-diagonal elements in \(Q\) indicate the correlation of noises. Also, \(\omega(1), \omega(2), \cdots, \omega(T)\) are assumed to be independent and identically distributed.
Theoretically, MDS is closely related to other existing models. On the one hand, MDS can be viewed as an multivariate version of the linear dynamical systems (Simon, 2006). Other studies having applied linear dynamical systems include Penny et al. (2005) and Smith et al. (2010). Penny et al. (2005) proposed the bilinear dynamical systems for modeling a single neuron activation. Smith et al. (2010) proposed a switching linear dynamic systems to improve the overall quality and sufficiency of model parameter estimation.

On the other hand, MDS has a lot commonalities with DCM in that they both contain a state equation to model the latent neuronal activations and an observation equation to map the latent neuronal activation to the observed neural signals, such as fMRI blood oxygen level dependent (BOLD) signals. However, there are many differences between MDS and DCM. First, DCM treats the brain as a deterministic dynamic system subject to inputs (Friston et al., 2003), but MDS explicitly includes a stochastic term. Second, DCM and MDS use different observation equations to forward the latent neuronal activation to the BOLD signal. In particular, DCM adopts a “Balloon” model (Friston et al., 2000; Buxton et al., 1998; Mandeville et al., 1999) to describe how latent neuronal activations are transformed into hemodynamic time-series, whereas MDS formulates the relationship as hemodynamic time-series being a convolution of latent neuronal states with a kernel expansion using basis functions. We are going to illustrate the observation equations in more details in Section 1.4, as it is an important consideration when using MDS to predict the fMRI BOLD signal.
1.4 Connection to neural measures

In MDS or DCM, the latent neuronal activations in each region are assumed to generate measurable neural responses through an observation equation. In this thesis, we focus on the fMRI BOLD signal. Many studies have investigated how BOLD signal arises from the functional imaging technique and have agreed on the principle of fMRI BOLD signal. The BOLD signal comes from the hemodynamic response where blood releases oxygen to active neurons at a greater rate than to inactive neurons. The hemodynamic response process further causes a change of the relative levels of oxyhemoglobin and deoxyhemoglobin that can be detected on the basis of their differential magnetic susceptibility.

MDS and DCM take different modeling approaches to predicting the distributions of these measures. In particular, DCM tries to emulate the real biophysical processes of how BOLD signal arises by making a series of specific assumptions. In contrast, MDS does not put a lot of effort on modeling every step of the biophysical processes and instead utilizes a widely acknowledged convolution method. Both approaches have pros and cons, but from the perspective of simplicity and feasibility, the convolution method used in MDS wins over the biophysical modeling in DCM. In fact, one can view the concrete biophysical assumptions used in DCM as being a generative model of the HRF used in the generic convolution method within MDS.

1.4.1 Balloon model

DCM applies a Balloon model (Friston et al., 2000; Buxton et al., 1998; Mandeville et al., 1999) for the biophysical cascade from neuronal state dynamics to BOLD signal changes. The Balloon model assumes that hemodynamic states are a function of,
Figure 1.3: The biological assumptions in the Balloon model (Friston et al., 2000; Buxton et al., 1998; Mandeville et al., 1999) used by the DCM. The neuronal model (1) reflects an excitatory - inhibitory balance, which induces the feedforward neurovascular coupling (2) and causes changes in blood inflow. Changes in blood inflow are accompanied by changes in blood outflow, blood volume and deoxyhemoglobin content (3). The rate of blood outflow is influenced by the change in the blood volume. Changes in blood volume and deoxyhemoglobin content are then reflected in the BOLD response (4). Figure extracted from Figure 1 in Havlicek et al. (2015) and only of, the neuronal state of each region (Friston et al., 2003). As is shown in Figure 1.3, the neuronal model reflects an excitatory - inhibitory balance, which induces feedforward neurovascular coupling and drives changes in blood inflow. Next, changes in blood inflow are accompanied by changes in blood outflow, blood volume and deoxyhemoglobin content. Then, the rate of blood outflow is influenced by the change in the blood volume. Finally, changes in blood volume and deoxyhemoglobin content alter the paramagnetic properties of local brain tissues and induce variations in the measured BOLD signal.
1.4.2 Convolution method

In MDS, the BOLD signal in each region is modeled as a linear convolution of the hemodynamic response function (HRF) and latent neuronal activations in that region obtained from Equation 1.1. For brain region $R_m$, the neuronal activations come from the $m$-th row of matrix $S$ and are denoted as $S_m$. The first step of the convolution is to create an embedded value $X_m(t)$ containing the past $L$ values of $S_m(t)$, such that

$$X_m(t) = [S_m(t), S_m(t-1), ..., S_m(t-L+1)]',$$

where $X_m(t)$ denotes an $L \times 1$ vector. Then $X_m(t)$ is convolved with an HRF basis, plus a noise term, to obtain the observed BOLD signal at time $t$ in $R_m$, such that

$$Y_m(t) = \phi X_m(t) + e_m(t), \quad (1.2)$$

where $Y_m(t)$ is the observed BOLD signal at time $t$ in $R_m$, $\phi$ is a vector of length $L$ and is the basis for HRF, and $e_m(t)$ is the observation noise with zero mean and variance $\sigma^2_m$, so that

$$e_m(t) \sim N(0, \sigma^2_m),$$

where $N(a, b)$ denotes the Gaussian distribution with the mean parameter $a$ and variance parameter $b$. The noise term $e_m(t)$ is uncorrelated across time points. As for the HRF basis $\phi$, we use the canonical HRF form
\[ \phi = h(x) = A \left( \frac{x^{\alpha_1-1} \beta_1^{\alpha_1} e^{-\beta_1 x}}{\Gamma(\alpha_1)} - c \frac{x^{\alpha_2-1} \beta_2^{\alpha_2} e^{-\beta_2 x}}{\Gamma(\alpha_2)} \right), \]

as it is common in most fMRI studies (Penny et al., 2005; Smith et al., 2010). $x$ references time, $\alpha_1 = 6$, $\alpha_2 = 16$, $\beta_1 = \beta_2 = 1$ and $c = 1/6$. $\Gamma(.)$ indicates the Gamma function\(^2\), which acts as a normalizing parameter. The only unknown parameter in the HRF basis is the amplitude $A$. Note that the convolution used in the MDS is dissimilar from the convolution in traditional fMRI studies. In traditional fMRI studies with general linear model (GLM), this HRF basis is convolved with the stimulus impulse function to obtain a task related regressor and to be used in the design matrix, but in the MDS framework, this HRF basis is instead convolved with the embedded value $X_m$ to obtain the BOLD signal.

### 1.5 Explaining behavioral measures

With connections from neuronal states to neural measures constructed, most studies based on MDS or DCM have exclusively focused on explaining neural data but left behavioral outcomes unexplored (but see Daunizeau et al. (2014); Rigoux and Daunizeau (2015)). To achieve our goal of simultaneously predicting behavioral data and neural measures of a functional integration network, behavioral measures must be incorporated into the MDS framework. In the next section, we discuss the previous relevant work and seek for an improvement based on that.

\(^2\)In the case of positive integer $n$, $\Gamma(n) = (n - 1)!$
1.5.1 Behavioral DCM

Rigoux and Daunizeau (2015) proposes a behavioral DCM to map the hidden neuronal states to behavioral outcomes. The central idea of the behavioral DCM is that the hidden neuronal states are transformed by a logistic regression model to produce a binary behavioral choice. A decision variable calculated from the hidden neuronal states primarily determines one of the two predicted outcomes with some noise added from the bias term, such that

\[
\text{predicted binary outcome} = \text{sign}(\text{decision variable} + \text{bias}),
\]

where \( y = \text{sign}(x) \) denotes a many-to-one mapping from a continuous variable \( x \) to a binary variable \( y \). For example, if \( x \) goes above a pre-defined value, then \( y \) is yes; otherwise \( y \) is no. In the DCM (or MDS) framework, the magnitude of the decision variable is determined by the neuronal activities and external inputs. To be more specific, the decision variable at each time point is affected by the neuronal activities and external inputs at this point and those prior to this point. The decision variable \( r(t) \) at time \( t \) can be expressed as

\[
r(t) = h(S(t),U(t)) \otimes K(t)
\]

\[
= \int_{0}^{\infty} K(\tau)h(S(t-\tau),U(t-\tau))d\tau,
\]

where \( K(t) \) is the finite impulse response of \( r(t) \), and \( h(S(t),U(t)) \) is a many-to-one mapping of neuronal state \( S(t) \) and external inputs \( U(t) \). \( h(S(t),U(t)) \) is convolved with \( K(t) \) to get the decision variable \( r(t) \). This convolution function in practice can be simplified to
\[ \dot{r} = h(S(t), U(t)) - \alpha r, \]

under the assumption that the impact of neuronal activities onto the decision variable decreases exponentially with time. \( \dot{r} \) denotes the rate of change of \( r(t) \) and \( \alpha \) is an unknown rate parameter.

Next, the decision variable \( r(t) \) at each time point is transformed into a variable \( z(t) \) through a sigmoid function,

\[ z(t) = \frac{1}{1 + e^{\rho - r(t)}}, \]

where \( \rho \) is an unknown parameter which serves as a threshold on the decision variable. Then \( z(t) \) is used as a parameter to decide the probability distribution of \( q(t) \). \( q(t) \) indicates a binary behavioral outcome at each time point \( t \) and takes value of 0 or 1. Note that \( q(t) \) has no realistic meaning and the 0 and 1 values are arbitrarily labeled for binary behavioral outcomes. \( q(t) \) has a Bernoulli likelihood with parameter \( z(t) \) indicating probability of being 1, such that

\[ P(q(t)) = z(t)^{q(t)}(1 - z(t))^{1-q(t)}. \]

Finally the likelihoods of \( q(t) \) being 0 or 1 are multiplied across time points to constitute the likelihood of a resulting behavioral outcome \( q \):

\[ P(q) = \prod_{t} P(q(t)). \]
1.5.2 Response time

To the best of our knowledge, the DCM/MDS model has only been applied to binary choice data, and has yet to consider choice response time data. Here we argue that response times provide useful information to infer the latent neuronal states and brain functional connectivity networks for the following reasons. First, as a type of behavioral outcome, response time has been widely shown an important indicator across cognitive tasks (Luce, 1986; Ratcliff and Smith, 2004; Ratcliff et al., 1999; Turner et al., 2016, 2018). Second, the response time data provide temporal information about when the responses are made at each trial, and those terminating points are essentially relevant with the neuronal activation dynamics and thus neural data (This point will be further illustrated in Chapter 2).

1.6 Plan of the thesis

In this thesis, we extend the Multivariate dynamical systems to accommodate behavioral data (choice and response time) with neural data (fMRI BOLD activity), and use this framework to construct two specific models. One is for the perceptual decision making task and another one is for the intertemporal choice task. We present three simulation results showing that the extended MDS models can predict patterns of both behavioral and neural data, and that different model configurations can be differentiated from the combined behavioral and neural data. Next we use approximate Bayesian methods to estimate model parameters from a simulated dataset and suggest that the model parameters of the extended MDS can be successfully recovered. Finally, contributions and limitations of the extended MDS are discussed.
Chapter 2: An MDS Model for Perceptual Decision Making

In this chapter, we develop a model for perceptual decision making intended to capture both neural and behavioral data simultaneously. To make the model more neurally plausible, we construct the model within the MDS framework, where each node is assumed based on extant literature on the neural basis of perceptual decision making. First, we review some important findings about neural mechanisms underlying perceptual decision making, because they provide the theoretical bases on which the model will be constructed. Next, we construct a model within the MDS framework to explain patterns of neural and behavioral data in a perceptual decision making task. We simulate the model under two configurations of perceptual stimuli, with the goal of illustrating the model’s potential for capturing key patterns that are typically observed in experimental data.

2.1 Neural correlates underlying perceptual decision making

The random dot motion task is the most canonical ask used to investigate the neural and cognitive basis of perceptual decision making (Ball and Sekuler, 1982; Britten et al., 1992; Churchland et al., 2008; Forstmann et al., 2010, 2008; Ho et al., 2009; Niwa and Ditterich, 2008; Roitman and Shadlen, 2002; Salzman and Newsome, 1994; Shadlen and Newsome, 2001; van Maanen et al., 2011). In this task, the stimuli
Figure 2.1: A trial example in a random dot motion task. The monkey decides the net direction of motion, here shown as up versus down. Task difficulty is governed by the fraction of dots that move coherently from one movie frame to the next. The monkey makes a second, voluntary movement to a choice target to indicate his decision. Figure extracted from Gold and Shadlen (2007).

consist of an array of moving dots, where some percentage of the dots are moving in a coherent direction. The percentage of dots moving coherently can be varied, and this percentage is often treated as an independent variable to measure the task difficulty (e.g., Britten et al., 1992). Figure 2.1 shows the basic structure of the random dot motion task. The gist of how the brain processes information in the task can be described in three steps. First, sensory visual neurons in the brain areas middle temporal (MT) and medial superior temporal (MST) of extrastriate cortex extract motion information from the visual image and represent the information within the visual cortex (Britten et al., 1992, 1996; Celebrini and Newsome, 1995; Croner and Albright, 1999; Shadlen et al., 1996). Neurons in MT and MST respond selectively to visual stimuli moving in particular directions reflecting the amount of motion energy
to which they are tuned (Albright, 1984; Simoncelli and Heeger, 1998; Zeki, 1974). Second, the map of motion direction in MT and MST is interpreted to guide the response about the net direction of motion. This interpretation function is often considered to be completed in areas of the frontal eye field (FEF) and the lateral intraparietal area (LIP) of the inferior parietal lobe (Andersen et al., 1992; Colby and Goldberg, 1999; Schall et al., 1995; Shadlen and Newsome, 2001). In particular, movement neurons in FEF and LIP initiate a saccade when their spike rate reaches a threshold (Brown et al., 2008; Dorris et al., 1997; Ratcliff et al., 2003, 2007). The cumulative strength of the motion information through time is often taken as evidence of accumulator dynamics in extant decision making models that assume sequential sampling of motion information (Boucher et al., 2007; Carpenter, 1999; Carpenter et al., 2009; Carpenter and Williams, 1995; Gold and Shadlen, 2007; Purcell et al., 2010; Ratcliff et al., 2003, 2007; Shadlen and Newsome, 2001).

The first and second steps only indicate the probability of making a decision choice to a certain direction for a given visual input, but the overt response relies on the ability of downstream neurons to pass the possibility to motor systems (Gold and Shadlen, 2001, 2002). This passing function is thought to be performed by the superior colliculus (SC) and basal ganglia in the third step (Lo and Wang, 2006; Ding and Gold, 2013). The basal ganglia is known to have a critical role in voluntary motor behavior in general (Hikosaka et al., 2000b; Houk et al., 1995; Graybiel, 1995; Wickens, 1997). Neurons in substantia nigra pars reticulate (SNr), an output structure of the basal ganglia, send GABAergic projections to principal cells in the SC, providing a “default” level of tonic inhibition to the SC. This tonic inhibition is released when the SNr receives increased inhibitory inputs from caudate nucleus (CD, part of the
Figure 2.2: The cortex-basal ganglia-superior colliculus network in a perceptual decision making task. The caudate nucleus receives inputs from both FEF and LIP. Its output is split along direct and indirect pathways, which are thought to have facilitatory and inhibitory effects, respectively, on behavior. These pathways converge in the substantia nigra, pars reticulata (SNr), which sends the output of the basal ganglia to the SC and, via the thalamus, back up to cortex. The dashed black line represents feedback pathways. Abbreviations: FEF, frontal eye field; GPi and GPe, the internal and external segments of the globus pallidus; LIP, lateral intraparietal area of the parietal cortex; SC, superior colliculus; SNC, substantia nigra pars compacta; SNr, substantia nigra pars reticulata; STN, subthalamic nucleus; VTA, ventral tegmental area. Figure extracted from Ding and Gold (2013).
striatum), which is driven by excitatory inputs from many cortical areas including the LIP and FEF (Hikosaka et al., 2000b, 2006). The third step of how LIP and FEF affect SC through the mediation of the basal ganglia is shown in Figure 2.2 and is explained as a trade-off mechanism in the striatal hypothesis (Forstmann et al., 2008, 2010). The striatal hypothesis posits that an emphasis on speed promotes excitatory input from cortex to striatum; the increased baseline activation of the striatum acts to decrease the inhibitory control that the output nuclei of the basal ganglia exert over the brain, thereby facilitating faster but possibly premature responses.

2.2 Model structure

Here, we apply the MDS approach to construct a general model of perceptual decision making. Based on the MDS framework, the model is intended to describe how key brain areas systematically activate through time within a trial. Our goal is to first construct a basis set of brain regions, based on the literature reviewed above, and then define the mathematical structure that relates the neural activation among the regions through time. As a reference, Figure 2.3 shows the overarching structure of the model, where we assume a set of six brain regions \((R_1, R_2, R_3, R_4, R_5, \text{ and } R_6)\). By specifying a particular structural relationship between these regions, we can simulate the model’s activity in the context of a random dot motion task.

In Figure 2.3, \(R_1\) and \(R_2\) are visual neuron populations mostly including MT and MST that selectively encode the motion information of the stimulus. \(R_1\) is thought of as a neuron population mainly sensitive to the leftward motion, whereas \(R_2\) is mainly sensitive to the rightward motion. \(U_L\) and \(U_R\) in Figure 2.3 are the leftward moving and rightward moving stimulus input strengths for the nodes \(R_1\) and \(R_2\), respectively.
Figure 2.3: A proposed MDS model for explaining neural and behavioral data from a perceptual decision making task. $U_L$ and $U_R$ represent the visual inputs for leftward and rightward moving dots, respectively. $R_1$ and $R_2$ represent the visual cortex neurons that encode the visual inputs of moving dots. $R_3$ and $R_4$ represent the neurons in regions such as the frontal eye field (FEF) and the lateral intraparietal area (LIP) thought to accumulate evidence for leftward and rightward dots independently. $R_5$ is assumed to be the output nuclei of basal ganglia and $R_6$ is assumed to be presupplementary motor area (pre SMA). $U_5$ provides a constant input for $R_5$ and $R_5$ inhibits the activation of $R_6$. Once the absolute difference of accumulated evidences in $R_3$ and $R_4$ reaches a threshold, $R_3$ and $R_4$ jointly inhibit $R_5$ so that $R_6$ gets disinhibited. $R_6$ is assumed to accumulate evidence for the response options, eventually passing an excitatory signal to initiate a movement.
$R_3$ and $R_4$ are neuron populations of FEF and LIP that further process the visual information from $R_1$ and $R_2$ to guide the responses. Hence, the neuronal activations of $R_1$ and $R_2$ induce the neuronal activations of $R_3$ and $R_4$ respectively, through their endogenous connectivity, and this induction is illustrated as red arrows in Figure 2.3 to represent excitatory effects. Here we conceptualize the neuronal activations of $R_3$ and $R_4$ as two independent decision variables evolving at each time point, which makes $R_3$ and $R_4$ function as two independent accumulators. This independent accumulator assumption has been used in many perceptual decision-making models (Gold and Shadlen, 2007; Shadlen and Newsome, 2001; Kim and Shadlen, 1999; Schall, 2003).

Moving rightward along the diagram in Figure 2.3, $R_5$ is assumed to be the output nuclei of basal ganglia and $R_6$ is assumed to be the presupplementary motor area (pre SMA). The neuronal activation in $R_5$ continuously sends tonic inhibition to $R_6$, preventing $R_6$ from making a response, and this tonic inhibition is illustrated as a purple arrow from $R_5$ to $R_6$ in Figure 2.3. $U_5$ provides a hypothetical constant impulse input for $R_5$ so that when there is no other brain region connected with $R_5$ and $R_6$, $R_5$ remains positively activated and thus $R_6$ remains inhibited. Although there are many other regions (e.g. SC, subcortex) that play an important role in decision making, we have omitted these areas from the MDS model for simplicity and their activities are unlikely to be clearly measured in real experimental data.

$R_3$ and $R_4$ are conditionally connected with $R_5$ through a gated mechanism (denoted by the blue dash line between $R_3$, $R_4$ and $R_5$). In particular, for each accumulator of $R_3$ and $R_4$, we calculate their neuronal activations at each time point,
summed up from the first time point (stimulus onset), and tracked the absolute difference between the two sums. Following the notation of MDS, we can express the cumulative value $E(t)$ as

$$E(t) = | \sum_{t=1}^{t} S_3(t) - \sum_{t=1}^{t} S_4(t) |,$$

where $S_3(t)$ and $S_4(t)$ indicate the neuronal activation of $R_3$ and $R_4$ at time point $t$, respectively. Whenever the cumulative value $E(t)$ reaches a pre-specified threshold value $\theta_1$, the connections from $R_3$ and $R_4$ to $R_5$ are initiated, illustrated as purple arrows from $R_3$ and $R_4$ to $R_5$ to represent inhibitory effects. The relative magnitude between the two sums of $R_3$ and $R_4$ is recorded, determining the response. We denote the time at which the threshold is reached as $t_0$. The joint inhibition from $R_3$ and $R_4$ makes $R_5$ unable to inhibit $R_6$.

Once $R_6$ becomes disinhibited, the neuronal activation of $R_6$ is monitored and accumulated at each moment in time from $t_0$. As soon as it reaches a pre-specified threshold $\theta_2$, $R_6$ sends out a signal to the muscle to initiate a movement, denoted as $t_1$. The blue dashed line between $R_6$ and the “response” stage on the right-hand side indicates another gated connection between the $R_6$ node and the muscle. The response to be made is determined by the relative magnitude of accumulated neuronal activations of $R_3$ and $R_4$ at time $t_0$. As we are not building a model of motor movement, we assume a constant delay parameter $\tau$ to execute the movement. This parameter is often used to model nondecision processes in other decision making models (Brown and Heathcote, 2005, 2008; Ratcliff and Smith, 2004; Smith and Vickers, 1988).
Figure 2.4: An example showing the states of activation of 6 hypothetical brain regions in Figure 2.3. Regions are colored corresponding to the nodes in Figure 2.3.

Following a response, the visual inputs from external stimuli are switched off (i.e. the values of $U_L$ and $U_R$ return to zero). As a result, the mean activations of $R_1$, $R_2$, $R_3$ and $R_4$ return to zero, but fluctuate around this mean due to the noise term in Equation 1.1. The variation of the noise term decreases after making a response because the brain regions are assumed to be in a “resting” stage. Although all of the nodes are still connected, the deactivation back to zero implies that $R_3$ and $R_4$ are no longer able to inhibit the $R_5$ node. $R_5$ becomes disinhibited and $R_6$ reverts back to being inhibited. The system remains at this stasis point until another stimulus is presented.
Figure 2.4 illustrates an example of how this model works by showing a trial of brain activation evolution of the six brain nodes. The three panels show how neuronal activations (y-axis) evolve with time (x-axis) in the six regions (R₁ to R₆). Uᵣ is set to be 4 times larger than Uᵢ and we can observe the neuronal activation oscillations of R₁ and R₂ in the top panel prior to time point t₁. There is a clear pattern that the magnitude of activation within R₂ is higher than that of R₁. The opposite pattern (i.e. R₁ is higher than R₂) mainly arises from the large noise term. We use the same connectivity coefficient from R₁ to R₃ and from R₂ to R₄. As such, prior to time point t₁, most of the time R₄ lies above R₃ (middle panel), similar to the pattern in the top panel (R₂ lies above R₁). The bottom panel of Figure 2.4 illustrates the effect on R₆ from the tonic inhibition from R₅ for the time points before t₀. Here, the activations of R₅ and R₆ are interwoven with each other whereas R₅ looks more stable. The bottom panel within the time window between t₀ and t₁ illustrates the opposite pattern of activations between the R₅ and R₆ nodes. While R₆ rises swiftly and remains highly activated, R₅ declines by a similar rate and remains negatively activated. The time window after t₁ illustrates the neuronal activations of the six brain regions after making the response. The activations of R₁, R₂, R₃ and R₄ fluctuate around zero means. R₅ becomes disinhibited and R₆ is inhibited at the negative value.

The MDS model we have developed can be used to generate predictions about neural and behavioral data through simulations. As a test of the model’s appropriateness, in the following sections, we simulate the model under different configurations of stimuli, and evaluate whether the model generates patterns of data that are reasonable. In Section 2.3, the stimuli consist of relatively common coherencies, whereas in
Section 2.4 the stimuli are somewhat irregular and involve equal evidence for the two options. While in principle equal coherencies for the two options should not present a problem, this particular stimulus configuration presents an interesting challenge to many extant models of decision making.

2.3 Motion detection with unequal coherence

The first simulation involves a standard set of stimuli, where coherence is varied along a single dimension, varying in strength for leftward and rightward response options. Since we constrain the coherence to be either leftward or rightward, we treat those dots moving toward any other directions as irrelevant and do not model them explicitly. For the purposes of our simulation, we assume 1000 dots shown on the screen, with each one moving either leftward or rightward. Then the leftward dots and rightward dots can be subtracted from each other to form a net coherence. For example, if there are 30% leftward dots and 70% rightward dots, then the net coherence level is 40% to the right. The probability of leftward dots is defined as $p_L$ and it is the independent variable in this simulated experiment. Table 2.1 shows that $p_L$ varies from .1 to .9, increasing by .1, implying that the probability of rightward dots $p_R$ varies from .9 to .1, decreasing by .1. When $p_L$ equals to $p_R$, the net coherence is zero. When $p_L$ is less than $p_R$, the net coherence is moving rightward, and the coherence levels are .8, .6, .4 and .2, corresponding to $p_L$ being .1, .2, .3 and .4. Conversely, when $p_L$ is greater than $p_R$, the net coherence is moving leftward, and the coherence levels are .2, .4, .6 and .8, corresponding to $p_L$ being .6, .7, .8 and .9.
Table 2.1: The $p_L$ condition levels and the corresponding $p_R$ levels and net coherence levels used in the simulation for motion detection with unequal coherence. $p_L$: probability of leftward moving dots. $p_R$: probability of rightward moving dots. The net coherence is the absolute value of difference between $p_L$ and $p_R$.

<table>
<thead>
<tr>
<th>$p_L$</th>
<th>.1</th>
<th>.2</th>
<th>.3</th>
<th>.4</th>
<th>.5</th>
<th>.6</th>
<th>.7</th>
<th>.8</th>
<th>.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_R$</td>
<td>.9</td>
<td>.8</td>
<td>.7</td>
<td>.6</td>
<td>.5</td>
<td>.4</td>
<td>.3</td>
<td>.2</td>
<td>.1</td>
</tr>
<tr>
<td>Net coherence</td>
<td>.8</td>
<td>.6</td>
<td>.4</td>
<td>.2</td>
<td>0</td>
<td>.2</td>
<td>.4</td>
<td>.6</td>
<td>.8</td>
</tr>
</tbody>
</table>

With $p_L$ and $p_R$ at hand, we can calculate the number of leftward dots and rightward dots and use them to represent the strengths of the visual stimuli. It is noteworthy that the number of dots reflect the physical properties of the stimuli, but for the simulation purpose, we need to map the physical properties of the stimuli to the perceived visual inputs (i.e., $U_L$ and $U_R$). Here we assume that the perceptual visual input strengths linearly scale with the number of dots. This linear assumption has been applied in many other perceptual decision making models (e.g., Brown and Heathcote, 2008; Smith and Vickers, 1998). In the simulation, we created the impulse function by using the number of dots as strength of input of $U_L$ and $U_R$.

2.3.1 Parameters

We applied the model described in Section 2.2 to a modified MDS based on Equation 1.1:

$$S(t) = C(t)S(t-1) + DU(t) + \omega(t),$$

where $C(t) = C_1$, for $t < t_0$, and $C(t) = C_2$, for $t \geq t_0$. Both $C_1$ and $C_2$ matrices are composed with diagonal elements to indicate the self-connection within
the regions, and nondiagonal elements to indicate the connection paths shown in Figure 2.3. Matrix $C_1$ has the form of

$$C_1 = \begin{pmatrix}
0.7 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.7 & 0 & 0 & 0 & 0 \\
0.7 & 0 & 0.7 & 0 & 0 & 0 \\
0.7 & 0 & 0 & 0.7 & 0 & 0 \\
0 & 0 & 0 & 0 & -0.8 & 0.7 \\
0 & 0 & 0 & 0 & -0.8 & 0.7
\end{pmatrix}.$$  

The diagonal elements in matrix $C_1$ indicate that the within-region connectivity strengths are all 0.7. $C_1[3,1] = C_1[4,2] = 0.7$ indicates that the connectivity strengths from $R_1$ to $R_3$ and from $R_2$ to $R_4$ are both 0.7. $C_1[6,5] = -0.8$ indicates that the connectivity strength from $R_5$ to $R_6$ is -0.8. All the other connectivity strengths are set to zero. Matrix $C_2$ has the form of

$$C_2 = \begin{pmatrix}
0.7 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.7 & 0 & 0 & 0 & 0 \\
0.7 & 0 & 0.7 & 0 & 0 & 0 \\
0.7 & 0 & 0 & 0.7 & 0 & 0 \\
0 & 0 & -0.2 & -0.2 & 0.7 & 0 \\
0 & 0 & 0 & 0 & -0.8 & 0.7
\end{pmatrix}.$$  

Matrix $C_2$ is identical to matrix $C_1$ except that $C_2[5,3] = C_2[5,4] = -0.2$. This indicates that the connectivity strengths from $R_3$ to $R_5$ and from $R_4$ to $R_5$ are both -0.2 after the connection paths are switched to on mode.

For each time point $t$ from time 0 (stimuli-onset) to $t_1$ in a simulation, the number of leftward moving dots ($U_L$) is randomly sampled from a Binomial distribution with a given probability parameter $p_L$:

$$U_L(t) \sim \text{Binomial}(1000, p_L).$$
and the number of rightward moving dots ($U_R$) equals to $U_L$ subtracted from 1000,

$$U_R(t) = 1000 - U_L(t).$$

Hence, as $U_L$ is sampled at each moment in time, the stimulus is stochastic, and the strength of evidence fluctuates through time. The sum of $U_L(t)$ and $U_R(t)$ at each point in time $t$ is always a fixed 1000. Following a response (i.e. $t > t_1$), $U_L(t)$ and $U_R(t)$ are set to zero.

$U_L$ and $U_R$ are put into the Equation 2.1 as the impulse function for $R_1$ and $R_2$, respectively, where $U_L$ and $U_R$ are in fact $U_1$ and $U_2$, constituting the first and second rows of matrix $U$. In practice, $U_L$ and $U_R$ are both divided by 100, which scales the strength of neuronal activation. $U_5$ is fixed to be 1 across the simulation, implying a constant magnitude of impulse function for $R_5$.

The direct input matrix $D$ is composed with diagonal elements for those regions with an external input (i.e. $R_1$, $R_2$ and $R_5$),

$$D = \begin{pmatrix}
.9 & 0 & 0 & 0 & 0 & 0 \\
0 & .9 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & .9 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
\end{pmatrix},$$

so that the direct effect coefficients of $U_1$ ($U_L$), $U_2$ ($U_R$) and $U_5$ on $R_1$, $R_2$ and $R_5$ are all .9, respectively.

The noise term $\omega(t) \sim N_6(0, Q)$, where $N_6(a, b)$ indicates a 6-variate Gaussian distribution with mean parameter $a$ and the variance covariance matrix $b$. Here the variance covariance matrix $Q = \sigma^2 I_6$ ($6 \times 6$ identity matrix). $\sigma^2$ is essentially useful
as a way to manipulate the signal-noise ratio and systematically affect choice accuracy and response times (Ryali et al., 2011). The absence of non-diagonal element in Q means that noises in the six brain regions are uncorrelated. $\sigma^2$ is set to be 1,000 prior to time point $t_1$ to capture the choice and response time distribution shape. $\sigma^2$ decreases to 100 after $t_1$ in order to represent the lower variations of neuronal activations after the response.

The two thresholds $\theta_1$ and $\theta_2$ are both set to be 5000. The non-decision time $\tau$ is set to be 100. Each level of $p_L$ is simulated 1,000 times. The number of total time points allowed in one simulation is 20,000 and none of the simulations reach this limit.

The parameters used for the BOLD signal convolution are set as follows. The length of the kernel L is 16, and the amplitude parameter A is 1. Convolution is done by following the steps in the Section 1.4.2. All negative values are reset to be zero in the convolved result. The magnitudes of the convolved neural activity are further scaled for illustration purpose. The scaling coefficients are $1/10$, $1/10$, $1/18$, $1/18$, $1/8$ and $1/28$ for $R_1$, $R_2$, $R_3$, $R_4$, $R_5$ and $R_6$. The scaling coefficients are the same for $R_1$ and $R_2$, and for $R_3$ and $R_4$, so that the magnitudes of BOLD signals of $R_1$ and $R_2$, and of $R_3$ and $R_4$ are comparable.

2.3.2 Results

Figure 2.5 shows the predicted behavioral data from the simulation. The behavioral data are presented as the choice and response time distribution at each $p_L$ condition. The nine panels are corresponding to the conditions of $p_L$ from .1 to .9. In each panel, the response time associated with a left choice is plotted in the negative axis while the response time associated with a right choice is plotted in the positive
Figure 2.5: Predicted choice and response time distribution in the simulation of motion detection with unequal coherence at each $p_L$ condition. At each condition, the response time associated with a left choice is plotted in the negative axis while the response time associated with a right choice is plotted in the positive axis.
axis. This simulation result indicates that as $p_L$ increases, the proportion of the left choice increases, along with the decrease of the right choice. When $p_L$ equals to .5, approximately same number of choices are made on left and right. This result is sensible according to the settings of the simulation. $p_L$ represents the input strength of the leftward motion relative with the total input strength of leftward and rightward motion. Therefore, when $p_L$ increases, the input strength for the leftward motion increases, along with the decrease of the input strength for the rightward motion, so the proportion of the left choice increases.

We summarize the simulated behavioral results in the left panel of Figure 2.6 by showing how accuracy and mean response time changes with $p_L$. Accuracy is defined as correctly choosing the direction with more moving dots. As such, when $p_L < .5$, the correct choice is to choose right while when $p_L > .5$, the correct choice is to choose left. As a reference, when $p_L$ is .5, choosing right direction is defined as the correct choice. (In our simulation, the proportion of left and right choices when $p_L$ is .5 is .507 and .493.) Figure 2.6 shows that accuracy decreases when $p_L$ increases from .1 to .5 and increases when $p_L$ increases from .5 to .9. Also, accuracy pattern is symmetric around $p_L$ equals to .5. The symmetric pattern is also found on the mean response time but with opposite direction: mean response time increases when $p_L$ increases from .1 to .5 and decreases when $p_L$ increases from .5 to .9. Table 2.1 is helpful to explain the symmetric pattern. In Table 2.1, the net coherence is also symmetric around $p_L = .5$ and the coherence level decreases as $p_L$ increases from .1 to .5, and increases as $p_L$ increases from .5 to .9. Previous studies with the random dot motion paradigm have shown that the net coherence serves as an indicator for the difficulty of the task (Britten et al., 1992; Roitman and Shadlen, 2002; Salzman and Newsome,
Figure 2.6: The summary of behavioral data from simulation of motion detection with unequal coherence (left) and simulation of motion detection with equal coherence (right). In the left plot, the mean response time at each $p_L$ level is plotted using the barplot. Error bars are included considering the number of simulation (1000) at each $p_L$ level. Accuracy is defined as correctly choosing the direction with more moving dots and is plotted with the dotline. As a reference, when $p_L$ is .5, choosing right direction is defined as the correct choice. In the right plot, the mean response time at each $p_L$ level is plotted using the barplot.
Therefore, our simulated random dot motion task is the hardest at $p_L = .5$ condition, and is the easiest at $p_L = .1$ and $p_L = .9$ conditions, and this difficulty pattern is symmetric around $p_L = .5$. This explains the patterns of both accuracy and mean response time.

Figure 2.7 shows the neural results predicted from the simulation. We collapse across simulations and average the predicted BOLD signals of $R_3$, $R_4$, and $R_6$. The shape of the BOLD activity comes from the canonical convolution kernel and it peaks at around time point 200 in our simulation. Comparing different $p_L$ conditions, Figure 2.7 shows in the left panel that as $p_L$ increases, the highest level of BOLD signal in $R_3$ also goes up, which is consistent with our assumption that $R_3$ accumulates evidence for leftward dots. Larger magnitude of leftward dots induces higher level of $R_3$ activity. Correspondingly, the middle panel shows that as $p_L$ increases, the highest level of BOLD signal in $R_4$ goes down, which is consistent with the assumption that $R_4$ accumulates evidence for rightward dots. Larger magnitude of rightward dots induces higher level of $R_4$ activity. The right panel shows that BOLD signal is higher for those boundary conditions (i.e. $p_L = .1$, $p_L = .9$) and lower for middle conditions (i.e. $p_L = .5$), and also the BOLD activity drops faster for boundary conditions compared with middle conditions. This prediction follows immediately with the assumption since $R_6$ stands for the motor cortex which sends out the response command. A response is easier to be made at boundary conditions, leading to a faster rise-up and higher activity level. The response made at boundary conditions takes less time on average, so the BOLD activity drops more quickly after the peak.
2.4 Motion detection with equal coherence

In the simulation of motion detection with unequal coherence, we treat the dots moving towards other directions other than leftward or rightward as the irrelevant “noise” and do not model them. Although we generate some meaningful results from the simulation, it is worth considering if the assumption is tenable. In the next simulation, we aim to simulate the scenario where the leftward dots and rightward dots are of equal amount while varying the ratio of their summed amount compared to the total amount of dots on the screen. This way, we are able to detect if the moving dots towards other directions play a role in the random dot motion paradigm. If these dots are indeed “irrelevant”, the model should predict similar predicted behavioral and neural results with the varied ratio. However, if these dots have an effect on the decision making process, the predicted behavioral and neural data could provide insights to the understanding of the problem.
In fact, this situation of equal coherence is related to the argument between sensitivity to absolute values and to relative values (Teodorescu et al., 2016). Promotors of the relative values postulate that decision making is guided by the relative value difference of the two alternatives, in terms of either the difference or the ratio (Brown and Heathcote, 2008; Ratcliff and Rouder, 1998; Roe et al., 2001; Tversky and Simonson, 1993). However, others argue that task irrelevant absolute values are also important in both humans (Usher and McClelland, 2001) and decentralized biological systems (e.g., bee colonies; Pais et al., 2013). In other words, the absolute value of the alternatives cannot be simply represented by the relative value. An empirical finding is that it takes less time to choose between equally valued alternatives for high value alternatives than for low value alternatives (Pirrone et al., 2014; Teodorescu et al., 2016), which is considered with adaptive advantages. These results indicate that the absolute value of the choice alternatives may not be superfluous. The equal coherence of leftward and rightward dots thus provides an appropriate emulation of this situation.

2.4.1 Parameters

We keep using 1,000 random dots in total, but the 1,000 random dots contain the same probability of leftward and rightward dots, along with dots in other arbitrary moving directions. The effects of dots from other moving directions are still assumed to be offset by summing up. In the 1,000 dots, the probability of leftward dots \( p_L \) (and also rightward dots \( p_R \)) is manipulated across \(.1, .2, .3, .4 \) and \(.5\), so the probability of the dots moving towards other directions \( p_{\text{others}} \) is correspondingly \(.8, .6, .4, .2 \) and
0. For each time point across a simulation, the number of leftward dots, rightward
dots and the others is randomly sampled from a Multinomial distribution, such that

$$(U_L, U_R, U_{\text{others}}) \sim \text{Multinomial}(1000, (p_L, p_R, p_{\text{others}})),$$

where $U_L$, $U_R$ and $U_{\text{others}}$ are the number of dots moving towards left, right and
any other direction, respectively. $p_L$, $p_R$ and $p_{\text{others}}$ are Multinomial distribution
parameters indicating the probabilities for random dots to move towards each of the
directions. Again, $U_L$ and $U_R$ are both divided by 100 and used as $U_1$ and $U_2$ in the
model. The other parameter settings in this simulation are exactly the same as those
in the simulation of motion detection with unequal coherence.

### 2.4.2 Results

Figure 2.8 shows the distribution of response time and choice as $p_L$ (and $p_R$)
increases from .1 to .5. Not surprisingly, equal proportion of left choices and right
choices are made across $p_L$ (and $p_R$) from .1 to .5, because the stimuli inputs provide
equal amount of strength for the leftward and rightward moving dots. However, the
response time shows a decreasing trend as $p_L$ increases. The right panel in Figure 2.6
shows clearly that the mean response time decreases as $p_L$ (and $p_R$) increases. This
result successfully recovers the key response time data findings in previous studies
(Pirrone et al., 2014; Teodorescu et al., 2016).

Figure 2.9 shows the predicted neural results from the simulation. The BOLD
activities are collapsed across simulations and averaged in regions of $R_3$, $R_4$, and
$R_6$. As $p_L$ (and $p_R$) increases, the activity grows faster and arrives at a higher peak
activation level in regions of $R_3$ and $R_4$. This trend is consistent with the result in that
Figure 2.8: Predicted choice and response time distribution in the simulation of motion detection with equal coherence at each $p_L$ condition. At each condition, the response time associated with a left choice is plotted in the negative axis while the response time associated with a right choice is plotted in the positive axis.
Figure 2.9: Predicted BOLD activities in brain region $R_3$, $R_4$, and $R_6$, for each of $p_L = .1, .2, .3, .4$ and .5 levels. Activities are averaged across 1000 simulations.

of simulation of unequal coherence, i.e. larger magnitude of moving dots is associated with higher speed of rise-up and higher peak activation level. In region $R_6$, as $p_L$ (and $p_R$) increases, the activity grows faster and arrives earlier at the peak activation level. However, the peak activity level of $R_6$ does not show obvious difference across $p_L$ conditions.

2.5 Conclusion

In this chapter, we have developed an MDS model for perceptual decision making and simulated the model under two configurations of perceptual stimuli. The two simulation results both show that the model has successfully captured the key characteristics of behavioral and neural data shown in experimental data. In the next chapter, we are going to apply MDS to a different type of decision making task.
Chapter 3: An MDS Model for Intertemporal Choice Decision Making

In this chapter, we develop an MDS model for the intertemporal choice decision making task to capture both behavioral and neural data simultaneously. Although the intertemporal choice shares some key similarities with perceptual decision making, the uniqueness of the intertemporal choice dynamics requires us an adaptation of the perceptual decision making MDS model. First, we review the extant literature on the neural mechanisms underlying the intertemporal choice task, and highlight several similarities between the perceptual decision making and the intertemporal choice decision making. Next, we describe an MDS model for intertemporal choice decision making, and show that different model configurations can be used to connect with different neural theories. Then, we simulate the different model configurations to demonstrate their uniqueness: different model configurations produce different predictions for behavioral and neural data, suggesting that they could be differentiated with experimental data.
3.1 Neural correlates underlying intertemporal choice decision making

Intertemporal choice refers to the tradeoff of choosing between two rewards occurring at different time points. For example, you can either choose to gain $20 tomorrow or to wait for one week to gain $30. In the literature, it is typical to call the first option (i.e., the one with a smaller reward occurring sooner) the smaller sooner (SS) choice, whereas the second option (i.e., the one with a larger reward at a longer delay) the larger later (LL) choice. The intertemporal choice can be viewed as a specific type of the experimental paradigm requiring a choice between two alternatives whose attributes trade off. In the above example, money and delay are the two attributes that trade off. In other words, the structure of the typical experiment guarantees that you cannot gain a larger reward without incurring a time cost (i.e., waiting for longer). Another common example related to the experimental paradigm with the tradeoff feature is about food consumption. For example, choosing between a tasty but unhealthy food and a healthy but not tasty food involves the tradeoff between the health and taste attributes.

A consistent finding is that the subjective values of delayed rewards are often discounted relative to immediate rewards (Frederick et al., 2002; Green and Myerson, 2004; Soman et al., 2005). This process can be characterized for an individual decision maker by measuring a discounting function, which shows how the subjective value of an outcome changes as a function of the delay. As an example, the discounting function can take the form of an exponential function (Samuelson, 1937):

$$SV(r, t) = r\delta^t,$$
where $SV(r, t)$ is the subjective value of reward $r$ to be received after delay $t$, $r$ is the utility of reward $r$ if received immediately, and $\delta$ is a discounting factor that ranges from zero to one, with smaller values of $\delta$ resulting in greater discounting. Another popular form of the discounting function is the hyperbolic discounting function (Ainslie, 1975; Mazur, 1987):

$$SV(r, t) = \frac{r}{1 + kt},$$

where $SV(r, t)$, $r$ and $t$ are with the same meanings as in the exponential discounting function. A positive value $k$ serves as a discounting factor, with larger values of $k$ indicating greater discounting.

Despite the variation in the specific form, the discounting functions show a common tendency that the combination of $r$ and $t$ trade off to produce diminishing “subjective representations” with longer delays (i.e., larger $t$). The subjective representation carries an amount of “strength”, and serves as the main drive for the choice making. Conceptually, the subjective representation is similar to the amount of dots in the random dot motion task. Mapping functions can also be produced to explain how attributes trade off in other domains, such as food consumption. The subjective representation of a food can be determined by the participants’ ratings on how likely they would choose to eat the food (Hare et al., 2009). This perspective shows an important similarity between the perceptual decision making and the intertemporal choice decision making.

Aside from the subjective value of a choice option, the probability that a certain option (e.g. the LL choice) is chosen with the competition of another option can be
manipulated according to a softmax function (Rodriguez et al., 2015; Turner et al., 2016, 2018):

\[ P_{LL} = \frac{1}{1 + e^{-m(SV_{LL} - SV_{SS})}}, \]

where \( m \) accounts for individual sensitivity to changes in discounted value, and \( SV_{SS} \) and \( SV_{LL} \) are calculated through a particular form of discounting function. The value \( P_{LL} \) represents the probability of choosing the LL choice, and is conceptually similar to the “coherence” variable in the perceptual decision making task. In the perceptual decision making task, the coherence level (from 0 to 100%) is used to systematically affect the choice probability for a particular direction. (We used coherence levels 0, .2, .4, .6, and .8 in Chapter 2.) Here \( P_{LL} \) can also be calculated with \( m \) estimated beforehand and \( P_{LL} \) predicts the probability of making a LL choice. Hence, manipulations of choice probability are analogous to manipulations of motion coherence, reflecting a second similarity between the perceptual decision making and the intertemporal choice decision making.

The idea of “subjective value”, though arises from discounting functions fitted into the behavioral data, is also endorsed by the neuroscience studies. The subjective value calculated from the best fitting parameters scales with the neural activities in mainly two brain regions including the ventromedial prefrontal cortex (vmPFC) and ventral striatum across many studies (Kable and Glimcher, 2007; Levy and Glimcher, 2012; Peters and Büchel, 2011; Rangel and Hare, 2010). We refer to this step of obtaining a subjective value from the conflicting attributes as the “valuation process” and its related regions as the “valuation regions”. The location of the valuation regions are shown as red areas in Figure 3.1. The valuation regions, in fact, are
conceptually similar with the visual cortex neurons which encode the input strength, which constitutes the third similarity between perceptual decision making and the intertemporal choice decision making.

Valuation process is important in that it provides the main drive to compare the available options and guide the final decision. However, two options with the same subjective value can still be differentiated as one being superior and another being inferior when considering the realistic situation. To be more specific, considering the economic benefits, the option with more objective reward after a longer delay might be a more rational choice to make than its more immediate competitor. We refer to the ability to forego attractive but ultimately inferior rewards in pursuit of more desirable goals as “self-control”. Studies have shown that executive brain regions such as the dorsolateral prefrontal cortex (dlPFC), anterior cingulate cortex (ACC) and posterior parietal cortex (PPC) are closely associated with self-control in the intertemporal choice behavior (Figner et al., 2010; Hare et al., 2009, 2011, 2014; Kable and Glimcher, 2007; Peters and Büchel, 2011; Turner et al., 2018). We refer to these regions as the “self-control regions” and the locations of them are shown as purple areas in Figure 3.1. Of these regions, dlPFC has been the most extensively studied brain area.

However, the debate continues on how self-control regions interact with valuation regions. One hypothesis suggests that self-control involves dlPFC modulation of valuation process in the vmPFC (Hare et al., 2009, 2011, 2014). Functional connectivity is identified between dlPFC and vmPFC through an intermediate node inferior frontal gyrus (IFG) whenever self-control is exerted. In particular, dlPFC has a negative interaction with IFG, and IFG has a positive interaction with vmPFC. This
connectivity was found to exist in experiments involving both food (Hare et al., 2009) and monetary (Hare et al., 2014) stimuli. This result provides indirect evidence to support that dlPFC exerts self-control through the modulation of vmPFC. Yet, they fail to find a direct connection between dlPFC and vmPFC, and the involvement of a third region introduces more possible explanations.

In contrast, another hypothesis suggests that the dlPFC influences choice behavior without altering valuation process (Figner et al., 2010; Turner et al., 2018). Figner et al. (2010) applied the repetitive transcranial magnetic stimulation (rTMS) to interfere with neural function in the dlPFC. The result shows that disrupting the dlPFC increases impulsive behavior (i.e. increasing the SS choice and decreasing the LL choice), but does not change the valuation judgments of the SS and LL options. This hypothesis is further endorsed by Turner et al. (2018), who tested different model configurations based on different possible mechanisms of self-control. They found that self-control was most likely associated with a lateral inhibition mechanism from
the leaky competing accumulator model (Usher and McClelland, 2001). In particular, the intertemporal choice decision making is considered as a competitive accumulation between the SS and LL options, and lateral inhibition introduces a competitive dynamic among the two alternatives, which is similar to the self-control effect in the behavioral data. Furthermore, they found that the lateral inhibition mechanism significantly correlated with areas in the prefrontal cortex commonly associated with cognitive control (Botvinick et al., 2001; Cavanagh et al., 2009; Ridderinkhof et al., 2004) on a trial-by-trial basis.

The second hypothesis asserts that self-control regions may affect choice preference without directly affecting the valuation process. One possible assumption inspired by Turner et al. (2018) would be that self-control brain regions exert a modulatory effect on the accumulation process. The accumulation process is the accumulation of subjective value for the purpose of action selection (Rodriguez et al., 2015), which is thought to be mostly associated with dorsomedial frontal cortex (dmFC), bilateral posterior parietal cortex, and bilateral lateral prefrontal cortex. The nature of these regions is analogous to accumulator regions such as LIP and FEF in perceptual decision making tasks. Hence, the analogous mechanistic roles played by these regions highlights a fourth similarity between perceptual decision making and intertemporal choice decision making.

In the next section, we will construct an MDS model to simulate behavioral and neural data for the intertemporal choice task. Motivated by the four similarities between the two types of decision making, this MDS model inherits most of the structures intrinsic to the model used above for perceptual decision making. Yet, despite the similarities, the mechanism of self-control within intertemporal choice
tasks are unique. Therefore, a crucial demonstration of the general applicability of the MDS framework as a computational model of brain activity is to use a similar network structure to produces patterns common to both types of decision making.

3.2 Model structure

We construct different configurations of models and map each model configuration to a particular theoretical mechanism of how self-control regions (especially dIPFC) facilitate LL choices in the intertemporal choice process, as is shown in Figure 3.2. Figure 3.2 shows that the model structure for intertemporal choice is similar to the structure for perceptual decision making (i.e., Figure 2.3), and thus we explain the model structure in order of the four differences between the two.

First, the strength of inputs $U_L$ and $U_R$ are changed to $SV_{SS}$ and $SV_{LL}$, where $SV_{SS}$ and $SV_{LL}$ represent the subjective values of the smaller sooner (SS) option and the larger later (LL) option, respectively. $SV_{SS}$ and $SV_{LL}$ may be calculated from the exponential discounting function or the hyperbolic discounting function. The choice of the discounting function is beyond the scope of this thesis, but in practice, we would choose a discounting function by conducting model comparison and then estimate the best-fitting parameters (i.e. $\delta$ in the exponential discounting function or $k$ in the hyperbolic discounting function) to calculate the subjective values.

Second, whereas in the two-choice perceptual decision making model, we manipulated 9 levels of leftward coherence ($p_L$) as the condition, here in the intertemporal choice context, the condition is defined as the expected probability to choose LL choice, denoted as $P_{LL}$. $P_{LL}$ in practice should be calculated from the softmax function, by considering the subjective sensitivity (i.e. parameter $m$ in the softmax function) to
Figure 3.2: A proposed MDS model scheme for the intertemporal choice task, modified from the perceptual decision making task model in Figure 2.3. The added $R_7$ indicates self-control related brain regions. $R_1$ and $R_2$ stand for the valuation regions. $R_3$ and $R_4$ stand for accumulation regions. $SV_{SS}$ and $SV_{LL}$ represent the subjective values of the smaller sooner option and the larger later option, respectively.

the differences between the two subjective values, $SV_{SS}$ and $SV_{LL}$. The typical way to obtain the parameter value $m$ is to use a staircase procedure before the formal experiment. We have three levels of $P_{LL}$ in the intertemporal choice model, .2, .5 and .8.

Third, the nodes of $R_1$, $R_2$, $R_3$ and $R_4$ carry different meanings in the intertemporal choice context. $R_1$ and $R_2$ represent “valuation regions” instead of the visual cortex. $R_3$ and $R_4$ represent “accumulation regions” rather than the LIP and FEF regions. It is noteworthy that although these nodes correspond to different set of brain regions, the nodes still play similar functions as in the perceptual decision making context. Specifically, $R_1$ and $R_2$ encode the strengths of the input, and $R_3$ and $R_4$. 
serve as two independent accumulators. Since the valuation regions and the accumulation regions are both non-separable in the intertemporal choice context, we model the neural activity in the valuation regions by the absolute value of the differences between $R_1$ and $R_2$, and model the neural activity in the accumulation regions by the absolute value of the differences between $R_3$ and $R_4$.

Fourth, an additional node $R_7$ is added to the model structure and this node is crucial since it represents the self-control regions, especially dlPFC. An impulse input function is set as $U_7$ to drive the activation in $R_7$. Node $R_7$ can be connected to each of $R_1$, $R_2$, $R_3$ and $R_4$ through different paths. As an exploration, we set up four different models M1, M2, M3 and M4. Each model contains only one path whereas in Figure 3.2 the model labels are put alongside the corresponding paths. For example, in M1, $R_7$ is only connected to $R_1$, with no connection to any other regions. Besides, we set a “null model” M0 where $R_7$ has no connection with other nodes in the model.

### 3.3 Parameters

The endogenous connectivity matrices $C_1$ and $C_2$ maintain most settings in the perceptual decision making simulation, but they both have an additional 7th row and 7th column indicating the connection of $R_7$ with itself and the other regions, so that

$$C_1 = \begin{pmatrix} .7 & 0 & 0 & 0 & 0 & 0 & c_1 \\ 0 & .7 & 0 & 0 & 0 & 0 & c_2 \\ .7 & 0 & .7 & 0 & 0 & 0 & c_3 \\ 0 & .7 & 0 & .7 & 0 & 0 & c_4 \\ 0 & 0 & 0 & 0 & .7 & 0 & 0 \\ 0 & 0 & 0 & 0 & -.8 & .7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & .7 \end{pmatrix}, \text{ and}$$

$$C_2 = \begin{pmatrix} .7 & 0 & 0 & 0 & 0 & 0 & c_1 \\ 0 & .7 & 0 & 0 & 0 & 0 & c_2 \\ .7 & 0 & .7 & 0 & 0 & 0 & c_3 \\ 0 & .7 & 0 & .7 & 0 & 0 & c_4 \\ 0 & 0 & 0 & 0 & .7 & 0 & 0 \\ 0 & 0 & 0 & 0 & -.8 & .7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & .7 \end{pmatrix}, \text{ and}$$
\[
C_2 = \begin{pmatrix}
.7 & 0 & 0 & 0 & 0 & 0 & c_1 \\
0 & .7 & 0 & 0 & 0 & 0 & c_2 \\
.7 & 0 & .7 & 0 & 0 & 0 & c_3 \\
0 & .7 & 0 & .7 & 0 & 0 & c_4 \\
0 & 0 & -.2 & -.2 & .7 & 0 & 0 \\
0 & 0 & 0 & 0 & -.8 & .7 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & .7
\end{pmatrix}
\]

According to the different model configurations, at most one of \(c_1, c_2, c_3,\) and \(c_4\) can be nonzero. The values of \(c_1, c_2, c_3,\) and \(c_4\) correspond to the connection strengths from \(R_7\) to \(R_1, R_2, R_3\) and \(R_4\) in M1, M2, M3 and M4, respectively. To simulate from model M0, we set all of \(c_1, c_2, c_3,\) and \(c_4\) to be zero. In the simulation, \(c_1 = .5, c_2 = -.5, c_3 = 1.1\) and \(c_4 = -1.1.\) \(C_1[7,7] = C_2[7,7] = .7\) indicates that the self-connection strength within \(R_7\) is \(.7.\)

For each time point \(t\) from time 0 (stimuli-onset) to \(t_1\) in a simulation, the subjective value of the LL choice \((SV_{LL})\) is randomly sampled from a Binomial distribution with a given probability parameter \(p_{LL},\) so that

\[
SV_{LL}(t) \sim \text{Binomial}(1000, p_{LL}),
\]

where \(p_{LL}\) varies across .2, .5 and .8 for different conditions, and the subjective value of the SS choice \((SV_{SS})\) equals to \(SV_{LL}\) subtracted from 1000, so that

\[
SV_{SS}(t) = 1000 - SV_{LL}(t).
\]

\(SV_{LL}\) and \(SV_{SS}\) are put into the Equation 1.1 as the impulse function for \(R_1\) and \(R_2,\) respectively, where \(SV_{LL}\) and \(SV_{SS}\) are in fact \(U_1\) and \(U_2.\) In practice, \(SV_{LL}\) and \(SV_{SS}\) are both divided by 100, which scales the strength of neuronal activation. \(U_5\)
and $U_7$ are fixed to be 1 across the simulation, implying a constant magnitude of impulse function for $R_5$ and $R_7$.

The exogenous connectivity matrix $D$ is also similar with the matrix D in the perceptual decision making MDS model, but it has an additional 7th row and 7th column, where the 7th diagonal element is .9, so that

$$
D = \begin{pmatrix}
.9 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & .9 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & .9 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & .9
\end{pmatrix},
$$

indicating that the effect coefficients of $U_1 (U_L)$, $U_2 (U_R)$, $U_5$ and $U_7$ on $R_1$, $R_2$, $R_5$ and $R_7$ are all .9.

The noise parameter $\omega(t) \sim N_7(0, Q)$, where $N_7$ indicates a 7-variate Gaussian distribution and the variance covariance matrix $Q = \sigma^2 I_7$ ($7 \times 7$ identity matrix). $\sigma^2$ is set to be 500 prior to time point $t_1$, and decreases to 100 after $t_1$. The two thresholds $\theta_1$ and $\theta_2$ are both set to be 5,000. The non-decision time $\tau$ is set to be 100. Each level of $p_{LL}$ is simulated 1,000 times. The number of total time points allowed in one simulation is 20,000 and none of the simulations reach this limit. The parameters used for the BOLD signal convolution are set as follows. The length of the kernel $L$ is 16, and the amplitude parameter $A$ is 1. Convolution is done by following the steps in the Section 1.4.2. All negative values are set to be zero in the convolved result. The magnitudes of the convolved neural activity are further scaled for illustration purpose. The scaling coefficients are $1/10$, $1/10$, $1/18$, $1/18$, $1/8$, $1/28$ and $1/10$ for $R_1$, $R_2$, $R_3$, $R_4$, $R_5$, $R_6$ and $R_7$. The scaling coefficients are the same
between $R_1$ and $R_2$, and between $R_3$ and $R_4$, so that the magnitudes of neural activity between $R_1$ and $R_2$, and between $R_3$ and $R_4$ are comparable.

### 3.4 Results

We summarize the behavioral results from different model configurations in Figure 3.3. The left panel shows how the proportion of choosing LL choice varies across $P_{LL}$ conditions, and the right panel shows how the mean response time varies across $P_{LL}$ conditions. The x-axis denotes the expected probability of choosing LL choice ($P_{LL}$), determined by the relative magnitude of input strengths to the two pathways. The y-axis denotes the proportion of choosing LL choice (left panel) and the mean response time (right panel). First we look at the null model M0 where self-control connection is not added. Figure 3.3 shows that as $P_{LL}$ increases, the model M0 predicts to make more LL choices, and the mean response time is higher at $P_{LL} = .5$, compared with the other two conditions. By adding a self-control node, M1, M2, M3 and M4 all predict a similar change in the pattern of the behavioral data. At condition $P_{LL} = .2$, the proportion of choosing LL increases from a level of almost zero to about .1. This is accompanied by an increase in the response time. The increase of response time is sensible because most of the choices at condition $P_{LL} = .2$ are SS choice and adding the self-control node makes it harder to make a SS choice. At condition $P_{LL} = .5$, the proportion of choosing LL increases from .5 to around .7, with a sharp drop in the mean response time. (Model M0 at condition $P_{LL} = .5$ predicts a really high mean response time value that goes beyond the y-axis limit in the right panel of Figure 3.3.) The drop of the mean response time indicates the difficulty of choice making decreases because the self-control node promotes the pathway of LL choice.
Figure 3.3: Predicted behavioral data from different model configurations in the intertemporal choice MDS model. The left panel shows how the proportion of choosing LL choice varies across different $P_{LL}$ levels. The right panel shows how the mean response time varies across different $P_{LL}$ levels. Different colors are for different model configurations, indicated by the legend in the left panel.

At condition $P_{LL} = .8$, the proportion of choosing LL does not increase. Instead, it has a slight decrease from a high level of almost 1. The mean response time at condition $P_{LL} = .8$ does not show a clear pattern of change, but there is a steady pattern of decreasing response time in M1 and M3. Most choices are made on LL at this condition, so the decrease of response times could indicate that the LL choices are easier to make, which is the effect of adding the self-control node.
Figure 3.4: Comparison of predicted behavioral data across different $P_{LL}$ levels (.2, .5 and .8) between M1 and M2, and between M1 and M3. The response time associated with a LL choice is plotted in the negative axis and the response time associated with a SS choice is plotted in the positive axis. The colors used in the histogram plots indicate the corresponding model configurations (M1, M2 and M3) and are consistent with the colors used in the plot title.
Next we are interested in whether there are differences in the predicted behavioral data by different model configurations. To this end, we compare the predicted behavioral data between M1 and M2, and between M1 and M3. The difference between M1 and M2 could reflect the different effects caused by adding an excitatory connection and an inhibitory connection. The difference between M1 and M3 could reflect the different effects caused by connecting the self-control regions to different nodes (i.e. valuation regions or accumulation regions). Figure 3.4 illustrates the comparison of behavioral predictions across different $P_{LL}$ conditions (.2, .5 and .8) between M1 and M2, and between M1 and M3. There are some slight differences in the predicted behavioral results between different model configurations. For example, M2 predicts slightly longer response times compared with M1, which can also be observed in Figure 3.3. However, response times predicted from different model configurations have similar ranges and averages, and the proportions of choosing LL or SS are also similar. In conclusion, we argue that the behavioral data predicted from different model configurations are similar enough so that we cannot use behavioral data alone to differentiate between these model configurations.

As a possible solution, we investigate predicted neural results in valuation regions and accumulation regions to see if the neural results can help identify between different model configurations. Figure 3.5 shows the predicted neural results from the null model M0 in the valuation regions (left panel) and in the accumulation regions (right panel). The data patterns in these two nodes look quite similar. In both valuation regions and the accumulation regions, BOLD activities are at a similar magnitude level at the conditions $P_{LL} = .2$ and $P_{LL} = .8$, and are higher than that at condition $P_{LL} = .5$. 

60
Figure 3.5: The predicted BOLD activity in the valuation regions and the accumulation regions from the null model M0, at different $P_{LL}$ conditions across .2, .5 and .8.
Then we investigate the BOLD activities in the valuation regions and the accumulation regions for the model configurations with the self-control node ($R_7$), M1, M2, M3 and M4. Since the parameter settings are invariant across M0, M1, M2, M3 and M4 except the self-control connection, we can compare the magnitude of the BOLD activities across different models. Figure 3.6 shows the predicted BOLD activity in the valuation regions. Both M3 and M4 predict similar neural results as the model M0. In particular, the BOLD activities are at similar level at conditions $P_{LL} = .2$ and $P_{LL} = .8$, higher than that at condition $P_{LL} = .5$. However, the BOLD activities predicted by M1 and M2 show different patterns. In particular, M1 predicts that the BOLD activity increases as $P_{LL}$ goes up, whereas M2 predicts that the BOLD activity decreases as $P_{LL}$ goes up. In addition, the highest BOLD activation level predicted by model M1 at $P_{LL} = .8$ condition is higher than that predicted by model M2 at $P_{LL} = .2$ condition.

Figure 3.7 shows the predicted BOLD activity in the accumulation regions. Still, we use the predicted BOLD activity in M0 as the reference. Both M1 and M3 predict that the BOLD activities are highest at $P_{LL} = .8$ and lowest at $P_{LL} = .2$. However, both M2 and M4 predict that the BOLD activities are highest at $P_{LL} = .2$ and lowest at $P_{LL} = .8$. Moreover, the highest activation magnitudes predicted by M1 and M3 are higher than that predicted by M2 and M4. So far, we have obtained the neural predictions at valuation regions and accumulation regions from model configurations M0, M1, M2, M3 and M4.

From the analyses above, predicted neural results from a single brain region may not completely differentiate between different model configurations, but the predicted
Figure 3.6: The predicted BOLD activity in the valuation regions from model M1, M2, M3 and M4, for each of $P_{LL} = .2$, $P_{LL} = .5$ and $P_{LL} = .8$ conditions.
Figure 3.7: The predicted BOLD activity in the accumulation regions from model M1, M2, M3 and M4, for each of $P_{LL} = .2$, $P_{LL} = .5$ and $P_{LL} = .8$ conditions.
neural results from a combination of brain regions are able to tell apart between different model configurations.

3.5 Conclusion

In this chapter, we have develop an MDS model for intertemporal choice decision making and shown that different model configurations are connected with different neural theories. In the simulation, different model configurations produce similar prediction for behavioral data but different predictions for neural data. The modeling results in this chapter provide insights to differentiate different neural theories from experimental data in future studies.
Chapter 4: Fitting MDS to Data

In this chapter, we fit the perceptual decision making MDS model developed in Section 2.2 to a simulated dataset using approximate Bayesian methods. We choose to perform Bayesian inference because one can simultaneously understand both the estimates of the model parameters and the uncertainty about them. We inform the model parameter estimation with both behavioral and neural data to provide more constraints than with only behavioral data or with only neural data (Turner et al., 2016). Our goal of this chapter is to illustrate the identifiability of the MDS model parameters by successfully recovering certain parameter values that have been used to generate the observed dataset.

When using Bayesian statistics, acquiring any posterior distribution depends on efficient evaluation of two functions: (1) the prior distribution for the model parameters, and (2) the likelihood function relating the model parameters to the observed data. The posterior distributions $\pi(\theta \mid D)$ reflect our knowledge about a parameter set $\theta$ after observing a dataset $D$, and it is obtained by combining the prior $\pi(\theta)$ with the likelihood of a parameter set $\theta$:

$$\pi(\theta \mid D) \propto \pi(\theta)L(\theta \mid D).$$

(4.1)
Prior distribution $\pi(\theta)$ reflects our knowledge of the parameter set $\theta$ before estimation and is typically easy to construct. As for the likelihood function $L(\theta \mid D)$, evaluating the function for purely statistical models can be straightforward. However, for simulation-based models that attempt to provide mechanistic explanations for how data manifest, direct evaluation of the likelihood function can be difficult, if not impossible. The MDS perceptual decision making model described in Section 2.2 is one simulation-based model with stochastic characteristics, and thus its likelihood function is hard to specify. As a solution, we use likelihood free method to approximate the likelihood function. This is the first attempt to apply the likelihood free method on the MDS model.

### 4.1 Probability density approximation

We use the probability density approximation (PDA; Miletić et al., 2017; Turner et al., 2013a; Turner and Sederberg, 2012, 2014; Turner et al., 2015a; Turner and Van Zandt, 2012) method to approximate the likelihood function instead of calculating it directly. Essentially, the PDA method relies on numerous simulations of the model for a candidate set of parameters to approximate the likelihood function through a kernel density estimation procedure (KDE; Silverman, 1986). By performing KDE on the simulated data, we obtain a probability density distribution that describes the likelihood of a certain data under that model. Such a likelihood is exactly what is needed to perform the Bayesian estimation. Here, we would like to clarify the usage of “simulated data”. As is stated at the beginning of this chapter, the Bayesian inference is performed on a simulated dataset, but throughout this chapter the simulated dataset is treated as an “experimental data”. To avoid confusion,
in the rest of this chapter, we use “simulated data” to refer to the data generated in the inference procedure and use “experimental data” (or “observed data”) to refer to the data that the inference is based on.

The KDE \( f(D_i \mid X) \) provides an approximate likelihood of experimental data point \( D_i \), given proposal data set \( X \). The likelihood of a set of parameters \( \theta \) given the full data set \( D \ (L(\theta \mid D)) \) is defined as the probability of the model to generate the full data set \( D \) with the given set of parameters \( \theta \). This probability can be approximated with the product of the likelihoods of individual observations \( D_i \) given the proposal data set \( X \):

\[
L(\theta \mid D) = \prod_{i=1}^{N} \text{Model}(D_i \mid \theta) \approx \prod_{i=1}^{N} f(D_i \mid X). \quad (4.2)
\]

The product of all the individual data points is a number very close to 0 (i.e., in the order of \( 10^{-1000} \) or below). Statistics software such as R cannot handle such numbers very well. Instead, a transformation of likelihoods to log-likelihoods is often used:

\[
\log L(\theta \mid D) \approx \sum_{i=1}^{N} \log f(D_i \mid X). \quad (4.3)
\]

Details of how KDE procedure calculates \( f(D_i \mid X) \) are described elsewhere (Turner and Sederberg, 2014; Turner and Van Zandt, 2012) and KDE is a built-in procedure in software packages such as R and Matlab. Hence we do not go through the details of the procedure in this thesis.

The downside of the PDA method is that PDA method approximates the likelihood function from model simulations, which are often very noisy when using MDS.
If likelihood estimates are noisy, chains get stuck: once a chain has found a likelihood that is artificially high due to noise, the probability of accepting a new parameter proposal according to the Metropolis Hastings rule decreases. Hence, any new parameter proposal will only be accepted when its corresponding likelihood estimate is also artificially high to due noise.

4.2 Differential evolution Markov chain Monte Carlo

With a suitable approximation for the likelihood in hand, we use differential evolution with Markov chain Monte Carlo (DE-MCMC; ter Braak, 2006; Turner et al., 2013d, 2015b) to sample from the joint posterior distribution. Conventional MCMC samplers have been shown to generate inefficient proposal sets when the parameters of a model are correlated, but DE-MCMC sampler shows a better performance when handling correlated parameters. DE-MCMC sampler uses multiple interacting chains and the difference between some chains’ current states is used to generate parameter proposals for other chains (Turner et al., 2013d).

In DE-MCMC sampling, a proposal parameter set $\theta^*$ is accepted with Metropolis Hastings probability:

$$\min(1, \frac{\pi(\theta^* \mid D)}{\pi(\theta_t \mid D)}).$$

(4.4)

Hence, if the probability of a proposal parameter set $\theta^*$ is higher than the probability of the current parameter set $\theta_t$, it is accepted. If the probability of $\theta^*$ is lower than the probability of $\theta_t$, the chance of acceptance is equal to the proportion of their probabilities.
4.3 Method

The general procedure for estimating parameters of the MDS model using PDA method is the following:

1. Generate a set of parameters $\theta$.

2. Simulate an experimental data set $D$ under $\theta$ in the way described in Section 2.2, where $D$ contains both behavioral data $B$ and neural data $N$.

3. Generate a proposal set of parameters $\theta^\star$.

4. Simulate a proposal data set $X$ under $\theta^\star$, where $X$ contains both behavioral data $X_B$ and neural data $X_N$.

5. Calculate the log likelihood of the proposal set of parameters $\theta^\star$ given experimental data set $D$ ($\log L(\theta^\star \mid D)$) using Equation 4.2 and Equation 4.3. The Bayes rule (Equation 4.1) and the Metropolis Hastings rule (Equation 4.4) decide whether to accept $\theta^\star$ or not.

6. Repeat steps 3, 4, 5 until a desired number of posterior samples have been obtained.

7. All the accepted parameter proposals $\theta^\star$ constitute the posterior distribution of the parameter set $\theta$.

Wilkinson (2008) has shown that approximate methods can be used to estimate posterior distributions for $\theta$, and Turner and Sederberg (2014) has shown that the more general approach of PDA method also produces accurate posterior estimates. However, in practice some estimation inaccuracy is to be expected. As the method
of acquiring posterior samples is based on two sources of algorithmic approximation: one due to approximating the likelihood from model simulation, and the other due to approximating the posterior distribution itself with DE-MCMC (which is intrinsic to most modern Bayesian applications).

For the purpose of inference, estimating the full matrices $C_1$, $C_2$ and $D$ in Equation 2.1 poses a great computational challenge. As a preliminary investigation, we decompose the matrices into their elements and focus on estimating only a subset of parameters within $C_1$, $C_2$ and $D$. Table 4.1 lists all of the model parameters of the perceptual decision making MDS model.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_0$</td>
<td>Inter-region connection strength, or diagonal elements in $C_1(C_2)$</td>
<td>.7</td>
</tr>
<tr>
<td>$a_1$</td>
<td>Connection strength from $R_1$ to $R_3$, or $C_1<a href="C_2%5B3,1%5D">3,1</a>$</td>
<td>.7</td>
</tr>
<tr>
<td>$a_2$</td>
<td>Connection strength from $R_2$ to $R_4$, or $C_1<a href="C_2%5B4,2%5D">4,2</a>$</td>
<td>.7</td>
</tr>
<tr>
<td>$a_3$</td>
<td>Connection strength from $R_5$ to $R_6$, or $C_1<a href="C_2%5B4,2%5D">6,5</a>$</td>
<td>-.8</td>
</tr>
<tr>
<td>$a_4$</td>
<td>Connection strength from $R_3$ to $R_5$, or $C_2[5,3]$</td>
<td>-.2</td>
</tr>
<tr>
<td>$a_5$</td>
<td>Connection strength from $R_4$ to $R_5$, or $C_2[5,4]$</td>
<td>-.2</td>
</tr>
<tr>
<td>$d_1$</td>
<td>Direct effect from $U_1$ to $R_1$, or $D[1,1]$</td>
<td>.9</td>
</tr>
<tr>
<td>$d_2$</td>
<td>Direct effect from $U_2$ to $R_2$, or $D[2,2]$</td>
<td>.9</td>
</tr>
<tr>
<td>$d_3$</td>
<td>Direct effect from $U_5$ to $R_5$, or $D[5,5]$</td>
<td>.9</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Non-decision time</td>
<td>100</td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>Threshold value for $E(t)$</td>
<td>5,000</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>Threshold value for accumulated movement information in $R_6$</td>
<td>5,000</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>Variance of the error term in Equation 2.1 before $t_1$</td>
<td>1,000</td>
</tr>
<tr>
<td>$A$</td>
<td>Magnitude parameter in the canonical HRF function</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4.1: Parameter summary: description and value of the parameters in the perceptual decision making MDS model.
With more than 10 estimable parameters in Table 4.1, we choose to perform parameter recovery analyses on four sets of model parameters, shown in Table 4.2. Hence, each model in Table 4.2 allows the subset of model parameters to freely vary; so while we knew the true value used to generate our fake data, our goal is to accurately estimate the parameters in each row. Practically, we need to generate a proposal parameter set for these free parameters at each iteration of the model fitting process and derive posterior distributions for the free parameters. For the other parameters listed in Table 4.1 but not contained as free parameters, they are fixed to be the values that produce the experimental data across the iterations. As shown in Table 4.2, in this thesis, we focus on the endogenous connection parameters as the first exploration to fit the MDS model by the PDA method in conjunction with DE-MCMC algorithm.

<table>
<thead>
<tr>
<th>Model</th>
<th>free parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_1$</td>
<td>$a_0, a_1$</td>
</tr>
<tr>
<td>$V_2$</td>
<td>$a_0, a_2$</td>
</tr>
<tr>
<td>$V_3$</td>
<td>$a_0, a_1, a_2$</td>
</tr>
<tr>
<td>$V_4$</td>
<td>$a_0, a_1, a_2, a_3$</td>
</tr>
</tbody>
</table>

Table 4.2: Models with different sets of free parameters.

### 4.3.1 Data generation

We use the same parameter values as in Section 2.2, shown in the Table 4.1, to generate each experimental dataset. We use these parameter values as we have shown in Section 2.2, they produce behavioral and neural results sharing key similarities with the real experimental data. For each simulation study, we constructed a dataset by
simulating the MDS model with 1,000 trials for each of 5 conditions, with $p_L$ varying across .1, .3, .5, .7 and .9. We reduced the total number of conditions from 9 (as in Section 2.2) to 5 to reduce the computational time of the model fitting process. As suggested in Section 2.2, the reduction to 5 levels of $p_L$ condition should maintain the key features in the dynamics of both behavioral and neural data.

4.3.2 Priors

We assumed normal priors for each model parameter in our recovery studies. For model $V_1$, we specified

\[ a_0 \sim N(1,1), \quad \text{and} \]
\[ a_1 \sim N(1,1), \]

where $N(a,b)$ denotes a Normal distribution with mean $a$ and variance $b$. Similarly, for model $V_2$, we specified

\[ a_0 \sim N(1,1), \quad \text{and} \]
\[ a_2 \sim N(1,1). \]

For model $V_3$, we specified

\[ a_0 \sim N(1,1), \]
\[ a_1 \sim N(1,1), \quad \text{and} \]
\[ a_2 \sim N(1,1). \]
For model $V_4$, we specified

\begin{align*}
a_0 & \sim N(1, 1), \\
a_1 & \sim N(1, 1), \\
a_2 & \sim N(1, 1), \text{ and} \\
a_3 & \sim N(-1, 1).
\end{align*}

In summary, across models $V_1$, $V_2$, $V_3$ and $V_4$, we specified the same prior distribution $N(1, 1)$ for parameters $a_0$, $a_1$ and $a_2$, and specified a prior distribution $N(-1, 1)$ for $a_3$. We specified these informative prior distributions because we assume $a_0$, $a_1$ and $a_2$ to be positive and $a_3$ to be negative from the specification of the model structure in Section 2.2. However, we did not constrain them to be positive or negative and instead allow the DE-MCMC algorithm to explore the broad parameter spaces. While the prior distributions are informative, we will see in the Result section that the prior distributions do not have primary effects on the posterior distributions.

4.3.3 Likelihood estimation

For a given candidate parameter value, we simulated the model 1,000 times at each of the 5 $p_L$ conditions (5,000 simulations in total) and obtained the behavioral and neural data. For each $p_L$ condition, the simulated behavioral data are collapsed across 1,000 simulations to form a choice response time distribution, and for the simulated neural data, we focus the inference on BOLD activities in regions of $R_3$, $R_4$ and $R_6$ and time points at 100, 200, 300 and 400, as a way to greatly reduce the computational demands. We choose regions of $R_3$, $R_4$ and $R_6$ because the simulation result in Section 2.2 suggests that the dynamic patterns of $R_3$, $R_4$ and $R_6$ are able to
capture key characteristics of the BOLD activities. We choose these four time points because the BOLD activities around these time points fluctuate significantly so that the activities around these time points are thought to capture the most important features. Therefore, for each $p_L$ condition, we form probability distributions across 1,000 simulations with BOLD activities at each region of $R_3$, $R_4$ and $R_6$, and each time point of 100, 200, 300 and 400. Altogether, we obtained 12 BOLD activity distributions for each $p_L$ condition. For each set of parameter proposal, we based the likelihood estimation on five choice response time distributions and 60 BOLD activity distributions.

The PDA method described above is applied to fit the perceptual decision making MDS model to both behavioral and neural data. Specifically, we estimate the log likelihoods of the response time data with leftward and rightward choices, and also estimate the log likelihoods of the BOLD activity, and then sum them up to form the total log likelihoods. Turner and Sederberg (2014) has shown that to approximate the likelihood function using PDA for data of mixed type (i.e., containing both discrete and continuous measures), we need only normalize each response time distribution by the number of responses made for each choice. Hence, when estimating log likelihoods of response time data with leftward and rightward choices, KDE is done separately for leftward and rightward choices, and multiplied by the proportion of leftward or rightward choices:
\[
\log L(\theta \mid B) = \sum_{j=1}^{J} \log \text{Model}(B_j \mid \theta)
\]
\[
\approx \sum_{j=1}^{J_1} \log(f(B_{j,\text{left}} \mid XB_{\text{left}}) \ast \frac{J_1}{J}) + \sum_{j=1}^{J_2} \log(f(B_{j,\text{right}} \mid XB_{\text{right}}) \ast \frac{J_2}{J}),
\]

where \(B\) is the experimental behavioral data (choice response time distribution) at a \(p_L\) condition and \(\theta\) is a set of parameter proposal. \(XB_{\text{left}}\) and \(XB_{\text{right}}\) are the simulated response time data under \(\theta\) for left and right choice, respectively. \(J_1\) and \(J_2\) are the number of left and right choices out of the total number of simulations \(J\) (i.e., 1,000), respectively. Similarly, for the neural data log likelihoods estimation,

\[
\log L(\theta \mid N) = \sum_{j=1}^{J} \log \text{Model}(N_j \mid \theta)
\]
\[
\approx \sum_{j=1}^{J} \log(f(N_j \mid XN)),
\]

where \(N\) is the experimental neural data (BOLD activity distribution) at a \(p_L\) condition, one of the four time points, and one of the three nodes of regions. \(\theta\) is a parameter proposal, and \(XN\) is the simulated neural data under \(\theta\). Combining the behavioral and neural data, we calculate the total log likelihoods for a proposal set of parameters \(\theta\) in the following:

\[
\log L(\theta \mid B, N) = \sum_{i=1}^{5} \left( \log L(\theta \mid B_i) + \sum_{m=1}^{3} \sum_{t=1}^{4} \log L(\theta \mid N_{i,m,t}) \right).
\]

\(i\) indexes the 5 \(p_L\) conditions (.1, .3, .5, .7 and .9), \(m\) indexes the 3 brain regions \((R_3, R_4 \text{ and } R_6)\), and \(t\) indexes the four time points (100, 200, 300 and 400). \(B_i\)
represents the experimental behavioral data (choice response time distribution) at $p_L$ condition $i$, and $N_{i,m,t}$ represents the experimental neural data (BOLD activity distribution) at region $m$ and time $t$ for $p_L$ condition $i$, both collapsing across 1,000 simulations\(^3\). B and N are the experimental behavioral and neural data, across 3 brain regions, 4 time points and 5 $p_L$ conditions.

### 4.3.4 Details of the sampling algorithm

Differential evolution algorithm applies principles of evolution to optimization to quickly find a best possible solution and to avoid getting stuck in local minima. We used 24 chains, and ran the algorithm for 200 iterations following a burn-in period of 50 iterations, resulting in 3,600 samples of the joint posterior. A migration step was used (see Turner and Sederberg, 2012; Turner et al., 2013d) with probability 0.1 for all the iterations. We also used a purification step every 10 iterations to ensure that the chains are not stuck in spuriously high regions of the approximate posterior distribution (Holmes, 2015). Convergence was checked by visual inspection.

### 4.4 Results

Figure 4.1 shows the parameter estimation result of model $V_1$. The top two histograms show marginal posterior distributions of parameters $a_0$ and $a_1$. Red vertical lines in the histograms represent the “true” parameter values used for the simulated dataset whereas the green vertical lines represent the mean of the posterior estimates. For both $a_0$ and $a_1$, the posterior means are close to the “true” parameter values, and this closeness indicates that the posterior distributions have successfully recovered

\(^3\)B\(_i\) and $N_{i,m,t}$ are written as B and N in Equations 4.5 and 4.6 to avoid confusion of notation.
the parameters. The blue dash lines in the histograms indicate the prior densities assumed for each parameter (the dash line can hardly be seen in histogram of parameter $a_0$ because the densities are very close to zero). Recall that we specified $N(1, 1)$ as the prior distribution for both $a_0$ and $a_1$, the posterior distributions are clearly different from the prior distributions. The bottom left plot shows how log likelihoods vary with iterations after the burn-in period for each of 24 chains. Though with fluctuations, the average log likelihoods remain in steady levels, and most chains have converged at the end of the 150th iteration. The joint distribution in the bottom right panel shows a strong negative linear relationship between the posterior estimates of $a_0$ and $a_1$. The red cross indicates the “true” parameter values used for the experimental dataset whereas the green circle indicates the means of the posterior estimates. The closeness of the two in the joint distribution further indicates a good recovery of parameters $a_0$ and $a_1$.

Figure 4.2 shows the parameter estimation result of model $V_2$. The top two histograms show marginal posterior distributions of parameters $a_0$ and $a_2$. Red vertical lines in the histograms represent the “true” parameter values whereas the green vertical lines represent the means of the posterior estimates. For both $a_0$ and $a_2$, the posterior means are close to the “true” parameter values, but less accurate than the posterior estimates of $a_0$ and $a_1$ in model $V_1$. The posterior distributions are different from the prior distributions indicated by the blue dash lines. Due to the symmetric characteristic of the leftward and rightward pathways in the perceptual MDS model, we would expect the parameters $a_1$ and $a_2$ to show similar posterior estimates. We attribute the different accurate performances of parameter estimation in model $V_1$ and model $V_2$ to random error occurring in the estimation process. The bottom left
Figure 4.1: The parameter estimation result of model $V_1$, with parameter $a_0$ and $a_1$ freely estimated. Red vertical lines in the histograms represent the “true” parameter values used for the experimental data whereas the green vertical lines represent the means of the posterior estimates. The blue dash lines in the histograms indicate the prior densities assumed for each parameter. The plot with “log likelihood” against “Iteration” indicates the change of log likelihoods along with the iteration (after burn-in period) for each of 24 chains (shown in different colors). The right bottom panel shows the joint distribution of $a_0$ and $a_1$. The red cross represents the “true” parameter values. The green circle represents the means of the posterior estimates.
Figure 4.2: The parameter estimation result of model $V_2$, with parameter $a_0$ and $a_2$ freely estimated. Red vertical lines in the histograms represent the “true” parameter values used for the experimental data whereas the green vertical lines represent the mean of the posterior estimates. The blue dash lines in the histograms indicate the prior densities assumed for each parameter. The plot with “log likelihood” against “Iteration” indicates the change of log likelihoods along with the iteration (after burn-in period) for each of 24 chains (shown in different colors). The right bottom panel shows the joint distribution of $a_0$ and $a_2$. The red cross represents the “true” parameter values. The green circle represents the averages of the posterior estimates.
plot shows that log likelihoods remain in steady levels along the iterations and most chains have converged at the end of the 150th iteration. The joint distribution in the bottom right panel shows a negative linear relationship between the posterior estimates of $a_0$ and $a_2$, with the red cross representing the “true” parameter values and the green circle representing the posterior means. The negative correlation could be a possible reason to cause the misestimation of parameter values.

Figure 4.3 shows the parameter estimation result of model $V_3$. The three histograms show marginal posterior distributions of parameters $a_0$, $a_1$ and $a_2$. The green vertical lines represent the average of the posterior estimates, but the “true” parameter values (.7, .7, .7) are beyond the plotting ranges. This indicates a worse model fitting performance compared with models $V_1$ and $V_2$. The posterior distributions are different from the prior distributions indicated by the blue dash lines. The middle right plot shows that log likelihoods remain in steady levels along the iterations and the chains have all converged at the end of the 150th iteration. From the pairwise joint distribution of parameters $a_0$, $a_1$ and $a_2$ in the bottom panels, there are strong negative linear relationships between the posterior estimates of $a_0$ and $a_1$, and between $a_0$ and $a_2$, and there is a strong positive linear relationship between the posterior estimates of $a_1$ and $a_2$. The true parameter values (indicated by the red crosses) lie outside of the clouds of the posterior estimates and are far from the posterior means (indicated by the green circles). However, the true parameter values lie in the same linear directions formed by the posterior estimates in each joint distribution plot, and it suggests that the linear correlations between parameters might be important causes of the misestimations.
Figure 4.3: The parameter estimation result of model $V_3$, with parameter $a_0$, $a_1$ and $a_2$ freely estimated. Red vertical lines in the histograms represent the “true” parameter values used for the experimental data whereas the green vertical lines represent the means of the posterior estimates. The blue dash lines in the histograms indicate the prior densities assumed for each parameter. The plot with “log likelihood” against “Iteration” indicates the change of log likelihoods along with the iteration (after burn-in period) for each of 24 chains (shown in different colors). The bottom three plots show the pairwise joint distributions of $a_0$, $a_1$ and $a_2$. The red crosses represent the “true” parameter values and the green circles represent the means of the posterior estimates.
Figure 4.4: The parameter estimation result of model $V_4$, with parameter $a_0$, $a_1$, $a_2$ and $a_3$ freely estimated. Red vertical lines in the histograms represent the “true” parameter values used for the experimental dataset whereas the green vertical lines represent the means of the posterior estimates. The blue dash lines in the histograms indicate the prior densities assumed for each parameter. The bottom two rows show the pairwise joint distributions of $a_0$, $a_1$, $a_2$ and $a_3$. The red crosses represent the “true” parameter values and the green circles represent the means of the posterior estimates.
As an exploration, we further added another parameter $a_3$ in model $V_4$. Figure 4.4 shows the parameter estimation result of model $V_4$. The histograms show marginal posterior distributions of parameters $a_0$, $a_1$, $a_2$, and $a_3$. The marginal distributions of $a_0$, $a_1$ and $a_2$ look almost identical to the marginal distributions in model $V_3$. The posterior estimates of $a_3$ are more accurate than the other three parameters. The posterior distributions are different from the prior distributions indicated by the blue dash lines. We skip plotting the log likelihoods with iterations as it looks very similar as in model $V_1$, $V_2$ and $V_3$. We plot the pairwise joint distributions of all the parameters in model $V_4$ in the bottom two rows. The pairwise joint distributions of $a_0$, $a_1$ and $a_2$ are very similar with those in model $V_3$. The correlations between them are little affected by adding another free parameter $a_3$. In contrast, the correlations of $a_3$ with each of $a_1$, $a_0$ and $a_2$ are not as strong as the pairwise correlations between $a_0$, $a_1$ and $a_2$. The weaker correlations of $a_3$ with other parameters provides a reasonable explanation for why the posterior estimates of $a_3$ are the most accurate among the four parameters.

We can conceptually understand the correlational patterns between the parameters from the model structure. The nodes of $R_1$, $R_2$, $R_3$ and $R_4$ are connected throughout the simulation, and the accumulated evidences calculated from nodes of $R_3$ and $R_4$ are directly related with the choice and response time prediction. $a_0$ is related with all nodes, $a_1$ connects $R_1$ and $R_2$, and $a_2$ connects $R_3$ and $R_4$. It is thus not surprising that the posterior estimates between $a_0$, $a_1$ and $a_2$ are pairwise strongly correlated. However, the nodes of $R_5$ and $R_6$ are not connected with the other four nodes at the beginning of a simulation and their activities are related only with the
response time, but not with the choice. As such, the posterior estimates of \(a_3\) only correlate with \(a_0\) to a lesser extent, but not correlate with \(a_1\) and \(a_2\).

4.5 Conclusion

In this chapter, we have performed parameter estimation on four sets of model parameters in the perceptual decision making MDS model. Although the estimations did not successfully recover all parameter in each set, they still suggested the identifiability of the MDS model. The parameter estimation of models \(V_1\) and \(V_2\) are generally satisfactory, and the misestimation in models \(V_3\) and \(V_4\) are mostly due to the correlations between the parameters. Although the DE-MCMC algorithm was expected to tackle the problem of correlation, the noise in the likelihood calculation with the PDA method impeded the DE-MCMC algorithm to play its role.
Chapter 5: Discussion

The framework we provided in this thesis is able to predict behavioral and neural data at the same time by successfully extending the Multivariate Dynamical Systems (Ryali et al., 2011, 2016). The predicted behavioral and neural data share key similarities with the experimental data, shown in both perceptual decision making and intertemporal choice decision making. This extended MDS framework is a valuable exploration of the Integrative approach of joint modeling (Turner et al., 2017b; Palestro et al., 2018). The behavioral and neural data are predicted through a single model, and this single model has a strong theoretical commitment to disclosing underlying cognitive mechanisms in the computational, algorithmic and implementational levels (Marr, 1982). The latent states in MDS models contain temporal dynamics and brain functional connectivities of cognitive processes, which contributes to an integrative brain-behavioral relationship.

5.1 Limitation of MDS

Although the MDS models developed in this thesis are useful in predicting behavioral and neural data, they show limitations in following aspects.
5.1.1 Noise term

In Chapter 2 and 3, we predict different choice and response time distributions across different levels of the independent variable (i.e. $p_L$ or $P_{LL}$) mainly by virtue of a noise term in Equation 1.1. Essentially, this noise term helps to allow the existence of “error” choices. For example, when the input strength of the leftward motion pathway is larger than the input strength of another pathway (e.g. rightward motion), there could still be a portion of rightward behavioral choices due to the noise term. The variance parameter $\theta$ of the noise term has to be large enough in the prediction for both perceptual decision making ($\theta = 1,000$) and intertemporal choice decision making ($\theta = 500$). The downside of employing such large noises in the system is that they make the latent neuronal states highly stochastic, and then the predicted neural BOLD signals through the observation equation carry over the randomness. The stochastic characteristics pose a great challenge to the inference of model parameters: the likelihood function is impossible to calculate and a proposal set of parameters has to be simulated a huge amount of times to approximate the likelihood accurately. The requirement of simulation times increases the time cost and might introduce a bias in the estimation if the simulation times are not sufficient.

5.1.2 Self-control node

For the intertemporal choice MDS model in Chapter 3, we expect to promote the LL choice by adding a self-control node ($R_7$). The self-control node can either excite the LL pathway or inhibit the SS pathway to play its role. The behavioral prediction presented in Figure 3.3 shows that the models with the self-control nodes ($M_1$, $M_2$, $M_3$, and $M_4$) predict higher proportions of choosing LL compared with the
model without a self-control node \((M_0)\) only at \(P_{LL}\) conditions .2 and .5, but models with the self-control nodes produce slightly lower proportions of LL compared with the model without a self-control node at \(P_{LL}\) condition .8. However, a more realistic prediction is that the proportion of choosing LL choice is constantly higher (at least not lower) for models with the self-control nodes than the model without a self-control node, at all conditions. The current predicted pattern in Chapter 3 might be caused by the large noise term. We will explore the reason in more depth.

### 5.1.3 Intertemporal choice MDS model

The intertemporal choice MDS model developed in Chapter 3 is mostly based on the model structure of perceptual decision making. Although we went through the similarities and differences between the two types of decision making, the intertemporal choice is far more realistic and complex, compared with the perceptual decision making usually happening in the laboratory experiment. The hyperbolic and exponential discounting functions can only approximate the subjective value (e.g. \(SV_{SS}\) and \(SV_{LL}\)) and the probability of choosing one option (e.g. \(P_{LL}\)) from a combination of delay and reward, but there are not explicit input such as the amount of dots moving leftward or rightward. Moreover, the discounting parameters are subject to change over time. Altogether, the less accurate calculation of model inputs is likely to bias the model prediction. Also, there has been an interest in how the dynamic interactions between delay and reward information in the choice making process. The subjective value may not fully represent the functions played by delay and reward, and so a next step to enrich the intertemporal choice MDS model is to represent how different attributes (e.g. delay and reward) are processed.
5.2 Parameter recovery

In Chapter 4, we fit the perceptual decision making MDS model to a simulated dataset to examine the identifiability of the model parameters. We did not use the expectation maximization algorithm that is the typical inference procedure in most DCM or MDS studies (Friston et al., 2003; Ryali et al., 2011, 2016). Rather, we tackle the difficulty of likelihood calculation by adopting the likelihood-free algorithm, in particular, the PDA method. The PDA method has been successfully applied in many established cognitive models such as Linear Ballistic Accumulator model (Turner and Sederberg, 2014) and Leaky Competing Accumulator model (Miletić et al., 2017; Turner and Sederberg, 2014). To the best of our knowledge, it is the first attempt to apply the PDA method to the likelihood calculation of a model capable of predicting behavioral and neural data at the same time. Hence, the success of the model fit could broaden the application field of the PDA method.

The posterior estimates in Chapter 4 show linear correlations between parameters $a_0$, $a_1$, $a_2$ and $a_3$. The correlational relationships suggest that freely estimating all parameters may not be possible, but a portion of parameters can still be estimated if we fix some parameters to be particular values when estimating the other parameters.

The problem of likelihood calculation is that when calculating the likelihood functions based on the neural experimental data, we have focus on a limited set of nodes and limited number of time points, and this might introduce a bias to the posterior estimates. In the future, we might consider increasing the number of simulations in the likelihood calculation and calculating likelihood based on the whole set of nodes. Of course, we will try estimating other parameter values listed in Table 4.1.
5.3 Future directions

MDS is a theoretical framework describing how cognitive functions are performed in the brain functional networks. This is done by modeling latent neuronal states in connected brain regions. Because MDS models latent neuronal states instead of observable data, it is not dependent on a specific modality of data, compared with the modeling studies focused on a specific modality of data. As such, the MDS framework enables us to transcend specific modalities of measurement, and to construct a much more integrative model. Besides modeling fMRI data, there have been observation equations to forward the latent neuronal states in MDS/DCM to other modalities such as electroencephalography (EEG; Garrido et al., 2008), magnetoencephalography (MEG; David et al., 2006), and local field potentials (LFPs) obtained from invasive recordings in humans or animals (Moran et al., 2013). Specific observation equations can be developed for other modalities. Ideally, the MDS framework is expected to incorporate observations from these different modalities. If such a goal can be achieved, it would promote the integration of research findings from different modalities and help gain a deeper and more comprehensive understanding of cognition.
Bibliography


Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.


*Commons, ML.; Mazur, JE.; Nevin, JA*, pages 55–73.


Simonoff, L. (1866). Die hemmungsmechanismen der säugethiere experimentell bewiesen.


