UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI FISICA E ASTRONOMIA "GALILEO GALILEI"

Master Degree in PHYSICS OF DATA

Final dissertation

EFFECTIVE CONNECTIVITY AND CONTROLLABILITY IN BRAIN DYNAMICS

Thesis supervisor
DR. SAMIR SUWEIS

Thesis co-supervisor
DR. MICHELE ALLEGRA

Candidate
KARAN KABBUR HANUMANTHAPPA MANJUNATHA
Student ID 1236383

ACADEMIC YEAR 2020/2021

THIS THESIS IS DEDICATED TO MY SUPERVISORS, PROF. SAMIR SUWEIS AND PROF. MICHELE ALLEGRA WHO HAVE GUIDED ME DURING MY THESIS WORK AND FROM WHOM I HAVE DEEPENED MY ABILITIES ON CRITICAL THINKING AND FUNDAMENTAL RESEARCH DURING HOURS OF DISCUSSION.

Abstract

Classical control theory is purported to drive a system towards a desired state by means of appropriate external inputs. Recently, the problem of controllability has been explored in the context of neuroscience, with the goal of inducing desired brain states through electrical (or even behavioral) stimulation. In principle, control theory allows deriving the optimal perturbation of the brains spontaneous ('resting') dynamics needed to reach a specific, arbitrary brain state. However, this requires a precise model of the spontaneous neural dynamics. So far, authors have considered a linear dynamics with inter-areal coupling given by anatomical connections ('structural connectivity', SC), which can be measured with diffusion magnetic resonance imaging. In fact, one can improve over this scheme by directly fitting a linear model to the neural activity data (e.g., from functional magnetic resonance imaging), and thus inferring the effective inter-areal coupling ('effective connectivity', EC). Modelling neural activity through EC gives a much more precise account of the actual dynamics in brain areas than using SC in the linear model. In this work, we address the controllability problem for brain dynamics using functional MRI data from a large public database (LEMON database) and testing different linear models proposed in the literature. The EC from each linear model is used to analyze the controllability of the dynamics (according to different criteria such as Kalman's controllability and the so-called Structural controllability). As a result, we show that controllability properties are robust to the detailed choice of linear model, provided that the models yield an EC structure with a comparable degree of sparsity. Sparsity emerges as the key topological property determining the controllability of the system. However, we also highlight the gap between the theoretical controllability and the controllability in practice as the energy to control the system is very high.

Contents

1

Linear Controllability

In this Chapter, we discuss about well established topics of *Control theory*. In the first section, we directly jump into the main crux of control theory called *Complete State controllability* where knowing the initial state we can reach any desired final state by providing appropriate input to certain nodes of the network. In the next section, we discuss about Kalman's criterion for any kind of Linear system(*LTI*) or network to be controllable. Next we focus on Structural controllability which is a further improvement relatively to Kalman's controllability. We finally discuss the measure called *minimum energy* required to steer the system with *minimum number of driver nodes* found via *Maximum matching*. In the end, we discuss the practical implementation of Kalman and structural controllability and using the result of the latter to obtain complete state controllability.

The Discrete linear time-invariant(LTI) system with time variable $t \in \mathbb{Z}$ has a state equation of the form:

$$\mathbf{x}(t+1) = A\mathbf{x}(t) + B\mathbf{u}(t) \tag{1.1}$$

and the corresponding system in its continuous form is given by:

$$\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = A\mathbf{x}(t) + B\mathbf{u}(t) \tag{1.2}$$

where the notations is as follows:

- $\mathbf{x}$(t) is a $N \times 1$ *state vector* of the form $(x_1(t), ..., x_N(t))^T$ with N being the number of nodes in the system.
- A is $N \times N$ *state matrix* which describes systems wiring diagram and the interaction strength between the components[1], including the possibility of self-loops A_{ii} representing the self-regulation of node i.
- B is $N \times r$ *input matrix* with $r \leq N$ which identifies the nodes controlled by an outside controller with $B_{ij} = 1$ if control input $u_j(t)$ is imposed on node i.

- **u**(t) is time dependent $r \times 1$ *input or control vector* of the form **u**(t) = $(u_1(t), ..., u_r(t))^T$ with r external inputs.

1.1. Complete State controllability or System controllability

Consider the discrete LTI system whose dynamics is given by Eq. 1.1

$$\mathbf{x}(t+1) = A\mathbf{x}(t) + B\mathbf{u}(t), \qquad \mathbf{x}(t=0) = \mathbf{x}_0 = \mathbf{x}_{initial} \tag{1.3}$$

Complete State controllability / System controllability is defined as the ability to transfer a given system from *any initial state* $\mathbf{x}(t = 0) = \mathbf{x}_0$ to *any desired final state* $\mathbf{x}(t = T - 1) = \mathbf{x}_{final}$ in a finite time $T < \infty$ with T steps. Thus, the problem is to find input control signal matrix which has this sequence of control signals $\mathbf{u}(0), \mathbf{u}(1), \mathbf{u}(2), ..., \mathbf{u}((T-1) = (n-1))$ so as to obtain final state $\mathbf{x}(t = T - 1) = \mathbf{x}_{final}$.[2]

Note that the number of time steps considered T should have to be strictly equal to N which is the dimension of A or number of nodes in the system if we are going to apply Structural controllability theory. In order to solve linear equation Eq. 1.3 and obtain the input control signal matrix **u**, it is crucial that the dimensions of the quantities in that equation are to be consistent and so T cannot be any natural number but must be equal to N.

Considering $t = 0, 1, 2, ...N$ for Eq. 1.3, we have:

$$\begin{aligned} \mathbf{x}(1) &= A\mathbf{x}(0)^T + B\mathbf{u}(0)^T \\ \mathbf{x}(2) &= A\mathbf{x}(1)^T + B\mathbf{u}(1)^T = A^2\mathbf{x}(0)^T + AB\mathbf{u}(0)^T + B\mathbf{u}(1)^T \\ &\vdots \\ \mathbf{x}(N) &= A^N\mathbf{x}(0)^T + A^{N-1}B\mathbf{u}(0)^T + ... + B\mathbf{u}(N-1)^T \end{aligned} \tag{1.4}$$

Note: Notation superscript T in the above set of equations represents vector transpose since we require column vector for matrix vector transpose.

The final equation of Eq. 1.4 is rewritten as:

$$\mathbf{x}(N) - A^N \mathbf{x}(0)^T = [B, AB, A^2B, ..., A^{N-1}B] \begin{bmatrix} \mathbf{u}(N-1) \\ \mathbf{u}(N-2) \\ \vdots \\ \mathbf{u}(1) \\ \mathbf{u}(0) \end{bmatrix} \tag{1.5}$$

where $C = [B, AB, A^2B, ..., A^{N-1}B]$ is *Controllability matrix* of dimension $(N, N.r)$ and finally the above equation turns out to be:

$$C^{N \times N.r} \begin{bmatrix} \mathbf{u}(N-1) \\ \mathbf{u}(N-2) \\ \vdots \\ \mathbf{u}(1) \\ \mathbf{u}(0) \end{bmatrix}^{(N.r) \times 1} = \mathbf{x}(N) - A^N \mathbf{x}(0)^T = \mathbf{x}_{final} - A^N \mathbf{x}_{initial}^T \tag{1.6}$$

As a practical implementation of complete state controllability, the procedure is as follows:

1. Fix initial state $\mathbf{x}(0) = \mathbf{x}_{initial}$ and final state $\mathbf{x}(N) = \mathbf{x}_{final}$.

2. Matrix A is already known and input matrix $B = B_{optimal}$ can be computed using algorithm as presented below in Section 1.3 (*Structural controllability via minimization of cost function*) .

3. With this, we solve linear equation 1.6 so as to compute $\mathbf{u}$ after obatining the result we reshape it from $(N.r) \times 1$ to $(N \times r)$.

4. A final cross check is required to see whether we are getting the exact same final state $\mathbf{x}_{final}$ using $\mathbf{u}$ in final equation of Eq. 1.4.

Note that even though the rationale for Kalman's criterion was discussed only for discrete LTI above, a similar reasoning holds also for the continuous LTI systems, yielding exactly the same condition for controllability.

1.2. Kalman Controllability

The system defined by Discrete LTI Eq. 1.1 or its continuous version Eq. 1.2 is said to be controllable if for a suitable choice of input signal vector $\mathbf{u}$(t), it can be driven from any initial state $x_{initial} = x(t = 0)$ to any final state $x_{final} = x(t = T - 1)$ in a finite number of time steps T. This is possible if and only if the controllability matrix

C of dimension $N \times (N.r)$

$$C = [B, AB, A^2B, ..., A^{N-1}B] \tag{1.7}$$

has full rank i.e.

$$rank(C) = N \tag{1.8}$$

Eq. 1.8 is called *Kalman's controllability rank condition*[3]. It represents mathematical condition for controllability, which says that *"The rank of controllability matrix should be equal to the dimension N of the state."*

Discrete time systems - If all the eigenvalues of A have magnitude less than 1 then Eq. 1.1 is stable and so the *discrete Lyapunov equation*:

$$W - AWA^T = BB^T \tag{1.9}$$

has unique solution given by

$$W = \sum_{m=0}^{\infty} A^m BB^T (A^T)^m \tag{1.10}$$

where W is called *Discrete controllability Gramian*. If it is positive definite or equivalently the minimum eigenvalue of W is strictly greater than zero then it is said to be controllable.

Continuous time systems - If all the eigenvalues of A have negative real parts then Eq. 1.2 is stable and so the *Continuous Lyapunov equation*

$$AW + WA^T = -BB^T \tag{1.11}$$

has unique solution

$$W = \int_0^{\infty} e^{At} BB^T e^{A^T t} \tag{1.12}$$

where W is called *continuous controllability Gramian* here. If it is positive definite or equivalently the minimum eigenvalue of W is strictly greater than zero then it is said to be controllable.

1.3. Structural Controllability

Let us begin by introducing terms and notations which will be used often. A system following LTI dynamics denoted as (A, B) with A and B being state matrix and input

matrix respectively, described in the beginning of this chapter, can be represented as directed graph $G(A, B) = (V, E)$. The vertex set defined as $V = V_A \cup V_B$ is a union of state vertices $V_A = \{v_1, ..., v_N\} = \{x_1, ..., x_N\}$ and input vertices/nodes $V_B = \{v_{N+1}, ..., v_{N+r}\} = \{u_1, ..., u_r\}$. In a similar way, edge set is defined as $E = E_A \cup E_B$ with $E_A = \{(x_j, x_i)|a_{ij} \neq 0\}$ which is set of all directed edges from j to i corresponding to non-zero entries of A while $E_B = \{(u_j, x_i)|b_{ij} \neq 0\}$ being set of directed edges from input vertices known as "origins" of $G(A, B)$ to state vertices. These state vertices having incoming edge from input vertices are called *control nodes*.

Example:

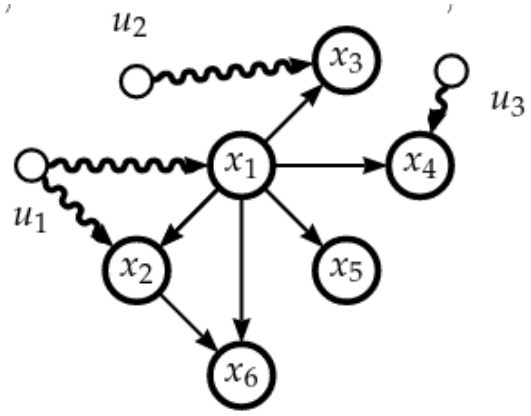


Figure 1.1

In Fig. 1.1, nodes x_1, x_2, x_3, x_4, which are receiving inputs from the input signal vertices or origins u_1, u_2, u_3, are called *control nodes*. However, only x_1, x_3, x_4, which are independent control nodes because they do not share input vertices, are called *driver nodes*.

Note that multiple control nodes can have same input node/origin and so those control nodes which do not share input vertices are called as *driver nodes*[1]. In other words, it is a subset of control nodes not sharing any input vertex among each other. If we denote the number of control nodes as r and driver nodes as r', it is observed that $r \geq r'$.

Before going further into the topic of Structural controllability, we provide an overview of what it means in general and how similar/different it is from Kalman controllability by highlighting the latter's limitations. Kalman controllability refers to a specific pair (A, B). It does not give a simple way to determine the minimum number of driver nodes that must be independently controlled so as to guarantee controllability. If one is free to choose B, Kalman controllability should be in principle tested for all possible $B's$. Moreover, second problem is that it is not fully

robust to noise in matrix A. If A is not fully specified, for instance because its entries are known only with some approximation, then Kalman controllability should be in principle be tested for all possible $A's$. This problem can be circumvented by applying the concept of structural controllability which is more powerful and encompassing approach in comparison to Kalman controllability. Instead of focusing on a single (A_0, B_0) pair, one focuses on the large class of structurally equivalent matrices A, B, and finds graph-theoretical criteria which specifiy how B should be chosen to achieve Kalman controllability of almost all structurally equivalent systems. To be more precise mathematically, it does not consider a single matrix A, but a manifold of matrices $M \subset R^{N\times N}$ with specific entries non-zero. Let us select entries of A that should be strictly non-zero and call it *mask of non-zero entries A* or A_{mask} for short:

$$M = \{A \in R^{N\times N} | a_{ij} \in A_{mask} \neq 0\} \tag{1.13}$$

The manifold M_A is said to be structurally controllable if there exists atleast one $A \in M$ such Eq. 1.1, 1.2 is controllable Eq. 1.8. Note that if the system is structurally controllable, then it is Kalman controllable for all of $A \in M$ except for those in some *proper algebraic variety* which has *Lesbesgue measure zero*.[1], [4], [5].

As defined in [4], given that (A_0, B_0) is known, the system defined by (A_0, B_0) is said to be *Structurally controllable* if and only if there exists a completely controllable pair (A, B) which has the same structure as (A_0, B_0). This means that (A, B) has the same topological structure as (A_0, B_0), in terms of existing links: all zero entries in A_0, B_0 are fixed to zeros in A, B, while the values of non zero entries take arbitrary values. In other words, (A, B) is a structured network model which is a topological model where the presence of edges with non-zero weights is known but not the values of the edge weights.

For a system to be structurally controllable, the following 3 statements are proven to be equivalent according to Lin[4]:

1. $G(A, B)$ follows Kalman's controllability rank condition as given in Eq. 1.8
2. (a) *Graph of a pair (A, B) is such that there exists no non-accessible node.*
   (b) Graph does not contain any *dilation*.
3. $G(A, B)$ is spanned by *cacti*.

Non-accessible nodes explanation with a simple example:

$$A = \begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \qquad B = \begin{pmatrix} 0 \\ 0 \\ a_{34} \end{pmatrix}$$

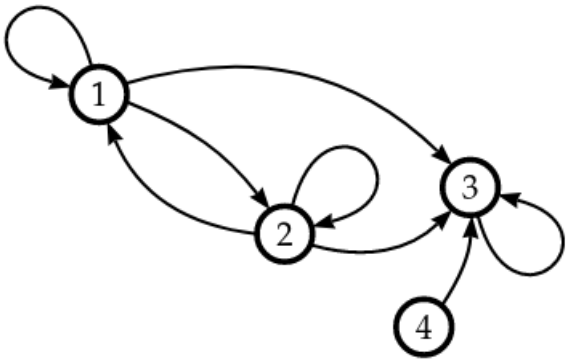


Figure 1.2

The element of the matrix a_{ij} is the weight of the directed edge going from node j to node i. Consider (AB) as a matrix of dimension $N \times (N+1)$ with (N+1)th column given by B. Here in this case $v_{N+1} = 4$ being is called as *origin* of G and so from definition no arrow can point towards vertex $v_{N+1} = 4$ in the example mentioned here.

A node v_i of a graph (AB) is said to be *non-accessible* if and only if there is no possibility of reaching node v_i from the origin node v_{N+1} along path composed of directed edges of graph (AB).

Thus it is possible to see that in the Fig. 1.2, node v_1 and node v_2 are the 2 nodes which are non-accessible from "origin" node v_4 and hence the system given by (A, B) is *not structurally controllable*.

Explanation of dilation with a simple example:

$$A = \begin{pmatrix} 0 & a_{12} & 0 \\ 0 & a_{22} & 0 \\ 0 & a_{32} & 0 \end{pmatrix} \qquad B = \begin{pmatrix} a_{14} \\ a_{24} \\ a_{34} \end{pmatrix}$$

In the Fig. 1.3, according to the graph constructed from (AB), a set S is formed by the nodes $S = \{v_1, v_2, v_3\}$ without the origin v_4. Then, another set $T(S)$ is computed consisting of all the nodes v_j with property that there exists directed edge from v_j to a node in set S. So, in this example, $T(S) = \{v_2, v_4\}$. If the cardinality of $T(S)$ is less than S, or

$$|T(S)| < |S| \tag{1.14}$$

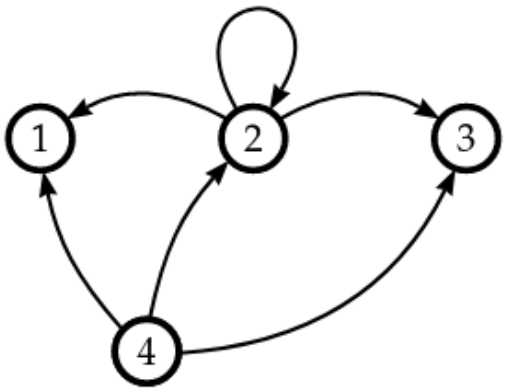


Figure 1.3

the graph (AB) is said to contain a *dilation*.

In general, a graph (AB) is said to contain *a dilation* if the cardinality of set of all nodes S excluding the origins v_{N+1} is k and the cardinality of set of nodes in $T(S)$(set of all source nodes v_j such that there exists a directed edge from v_j to any *target node* in S) is no more than *(k-1)*, then the graph is said to have a dilation. *If the graph has a dilation, the system is not structurally controllable*.[4]

Explanation of Cacti with an example:

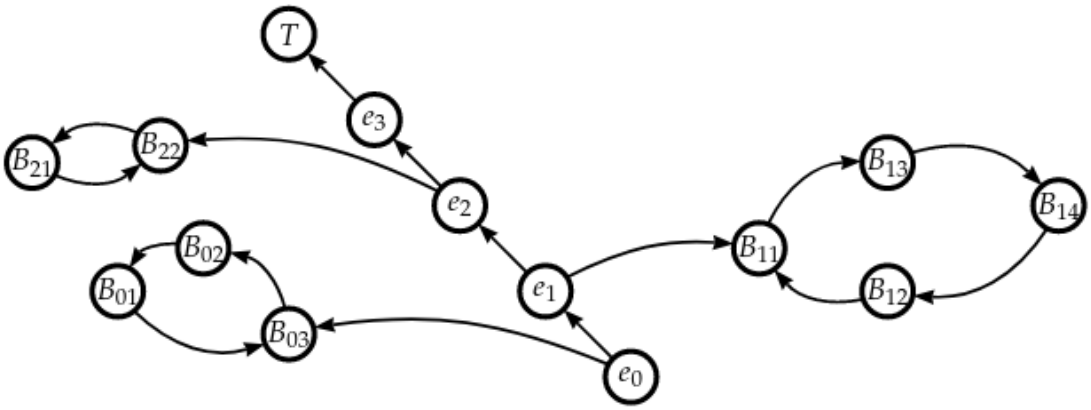


Figure 1.4

A *Stem* is defined as an elementary path originating from the input vertex. In the above figure, $\{e_0, e_1, e_2, e_3, T\}$ forms the stem. A stem consists of initial/input vertex which is *root* of the stem(*here* e_0) and terminal vertex which is *top* of the stem. (*here* T).

A *bud* is defined as an *elementary cycle* with an additional edge that ends at a vertex of the cycle(*but never begins at a vertex of a cycle*). This "additional edge" is termed as *distinguished edge of the cycle*. Here in the above example figure, there are

3 buds: $Bud_0 = (e_0, B_{03}) \cup Cycle_0$, $Bud_1 = (e_1, B_{11}) \cup Cycle_1$ and $Bud_2 = (e_2, B_{22}) \cup Cycle_2$ where $Cycle_0, Cycle_1, Cycle_2$ are made up of nodes set $Cycle_0 : \{B_{01}, B_{02}, B_{03}\}$, $Cycle_1 : \{B_{11}, B_{12}, B_{13}, B_{14}\}$ and $Cycle_2 : \{B_{21}, B_{22}\}$. The distinguished edges of the buds are: $\{e_0 \longrightarrow B_{00}, e_1 \longrightarrow B_{11}, e_2 \longrightarrow B_{22}\}$.

Description of Cacti. A stem is a cactus. With stem S and set of $n+1$ buds $Bud_{set} = \{Bud_0, Bud_1, ...Bud_n\}$, $S \cup Bud_{set}$ is said to be a *cactus* if for every $0 \leq i \leq n$ the *source node i* of distinguished edge of Bud_i is not the top of S and *vertex* $i \in \{Bud_i \cup (S \cup Bud_0 \cup Bud_1 \cup ... \cup Bud_{i-1})\}$). In the above example which is a cactus, e_2 is the initial vertex of distinguished edge $e_2 \longrightarrow B_{22}$ and also not the top of the stem, as well as it belongs to $\{stem\{e_0 \longrightarrow e_1 \longrightarrow e_2 \longrightarrow e_3 \longrightarrow T\} \cup Bud_0 \cup Bud_1\}$
Thus, if a graph $G(A, B)$ is spanned by such *cacti* as described above, then the system is *structurally controllable*.

1.4. Minimum Controllability concepts

1.4.1. Minimum Controllability problem:

Identifying a minimum number of driver nodes(and not control nodes which is $\supseteq$ driver nodes) such that the system is *Kalman Controllable* Eq. 1.8, is known as *Minimum Controllability problem*. By notation, N_D will be henceforth referred as *minimum number of driver nodes* where $N_D = r'_{min}$ with terms/notations as in section 1.3.

1.4.2. Maximum matching:

Finding the minimum number of input vertices or driver nodes so as to gain full control over the network is an *NP-Hard problem*[6]. To avoid such a brute force approach, Liu et al.[1] in their 2011 paper provide an alternative graph based approach to solve the problem of minimum controllability, where it is proved that finding minimum number of driver nodes is indeed related to size of *maximum matching* of the digraph $G(A)$.

Main idea:[1] Given a digraph $G(A)$, there exists a set of edges M called *Matching set* being subset of all the edges of graph(i.e. M $\subset$ E) which consists of edges which do not share common starting or common ending nodes. Each of these edges are called *Matching edge*. If a node is an ending node or target node of these matching edges, then it is called *Matching node*, otherwise it is called unmatched node.

If the cardinality of matching edge set M is maximum, then it is termed as *Maximum Matching*. It is not necessary that there always exists an unique Max matching set M for a digraph, but one can have many and choosing one among the many is

enough to move forward for solving the problem of finding minimum number of driver nodes. To find maximum matching to a directed graph, it is necessary to initially convert it to bipartite graph and then use the *Hopcroft-Karp algorithm*[7], which allows an easy computation of the maximum matching for bipartite graphs. The complexity is of the order $O(\sqrt{V}E)$. Also, note that the maximum matching is said to be *perfect* if all the nodes of the graph are matched. Examples of such graphs directed elementary cycles, graphs with only self loops.

1.4.3. Minimum Input Theorem:

[1] Given a directed graph $G(A)$, the minimum number of inputs or minimum number of driver nodes N_D needed to completely control the system is:

- $N_D = 1$, if there is perfect matching for digraph $G(A)$
- $N_D =$ number of unmatched nodes, otherwise

Mathematically, it can be described as

$$N_D = max\{N - |M^*|, 1\} \tag{1.15}$$

where $|M^*|$ is size or cardinality of maximum matching set M which is also equivalent to number of matched nodes.

Proof of the theorem.

1. N_D = *Number of unmatched nodes* = $N - |M^*|$. If the number of matched nodes are less than total number of nodes in graph $G(A)$ i.e. $|M^*| < N$, then it is not a perfect matching. The matching edges form elementary paths and cycles, henceforth called as *matching path* and *matching cycle*. In the example of cacti(see point 1.3 of section 1.3) discussed earlier, the stem S forms the matching path while $Cycle_0, Cycle_1, Cycle_2$ forms matching cycles.
   $N - |M^*|$ number of stems are formed when we connect an input vertex to each of the unmatched vertex since there are already $N - |M^*|$ unmatched vertices. And all the remaining nodes which are not unmatched nodes are *spanned by matching cycles*.

   There are 2 possible ways to form a bud:

   (a) If for a distinguished edge e_{ij}, if the starting vertex i belongs to stem S and ending vertex j belongs to one of the matching cycles C, then $e_{ij} \cup C$ forms a bud.

   (b) If a matching cycle C cannot form a bud in such a possible way, an alternative way is to connect to one of the input vertices directly and finally form a bud.

We see that the matching cycles do not require extra inputs so as to form buds and hence we finally have a disjoint set of cacti with $N - |M^*|$ inputs. According to Lin's theorem on structural controllability, a graph $G(A, B)$ is structurally controllable if it is spanned by cacti. Since we know that $|M^*|$ is the size of maximum matching in $G(A)$, therefore we have $N_D = N - |M^*|$ as the minimum number of input nodes/driver nodes required to attain structural controllability of a given system.

2. $N_D = 1$. If there is a perfect matching i.e. $|M^*| = N$, then there are no unmatched nodes which means all the nodes are belonging to matching cycles or in other words all the vertices are spanned by one or more matching cycles. A simplest solution to attain controllability in such a case is to introduce a single input vertex(or origin) and connect it to all the cycles of the graph of this type and finally we have buds with all of the buds having the same input vertex of their respective distinguished edges. Finally we have a cacti with a single input and hence we arrive at $N_D = 1$ again by using Lin's equivalent statement on structural controllability.

1.4.4. Maximum matching examples:

Example 1:

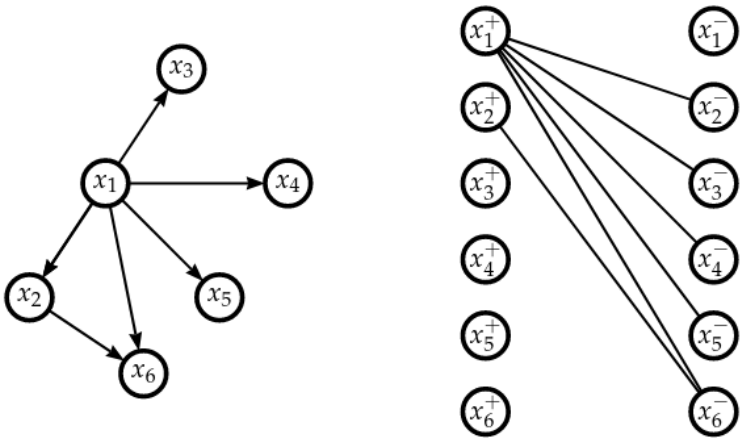


Figure 1.5

In the above example Fig. 1.5, digraph(left) can be transformed to a bipartite graph(right) where a directed edge from x_i to x_j of the directed graph will be represented as an undirected edge from x_i^+ to x_j^- as shown in left of above figure. Applying the concept of maximum matching on the bipartite graph, we see that max. matching set $M^* = \{(x_2^+, x_6^-), (x_1^+, x_5^-)\}$ or $\{(x_2^+, x_6^-), (x_1^+, x_4^-)\}$ or $\{(x_2^+, x_6^-), (x_1^+, x_3^-)\}$ or $\{(x_2^+, x_6^-), (x_1^+, x_2^-)\}$ where the 2 edges in these sets obviously do not have any common starting or ending vertex. And so the matched nodes can be either $\{6, 5\}$ or $\{6, 4\}$ or $\{6, 3\}$ or $\{6, 2\}$ and so unmatched nodes being complement of the matched

nodes among set of all nodes is simply $\{1,2,3,4\}$ or $\{1,2,3,5\}$ or $\{1,2,4,5\}$ or $\{1,3,4,5\}$ respectively. Finally, to fully control the network the minimum number of inputs/driver nodes required are just 2 for this network.

Example 2:

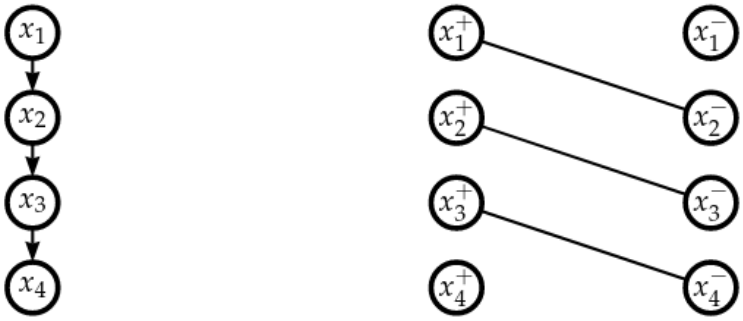


Figure 1.6

In this example Fig. 1.6, after mapping the digraph(left) to bipartite graph, the max. matched edges set $M^* = \{(x_1^+, x_2^-), (x_2^+, x_3^-), (x_3^+, x_4^-)\}$ is unique for this case. And so the matched nodes are $\{2,3,4\}$ and the set of unmatched nodes, being the complement of the set of matched nodes here reduces to $\{1\}$. This shows that the whole network can be controlled by giving input signal to single driver node which is called *Single node controllability*.

Example 3:

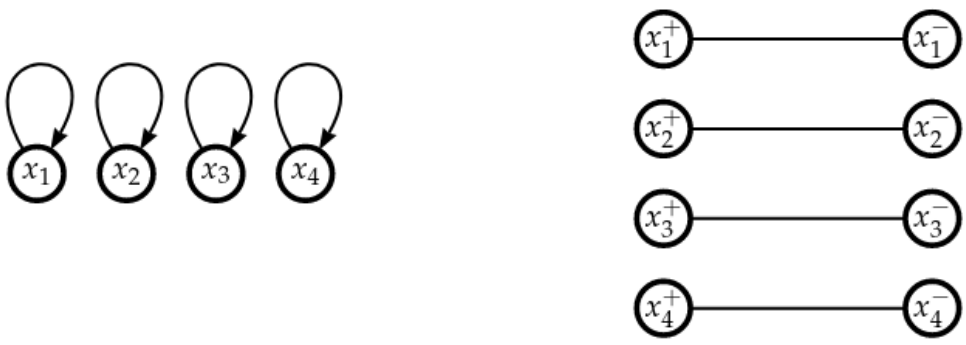


Figure 1.7

This example Fig. 1.7 is a very specific representation of graph whose adjacency matrix contains only diagonal elements, which means the network contains only self-loops and no other edges. So, what about it? In fact, this adjacency matrix is not problematic. The procedure is the same, the first thing is to convert it into a bipartite

representation(right). Then we see that none of the edges have same common or ending vertices and so all the edges are matched i.e.
$M^* = \{(x_1^+, x_1^-), (x_2^+, x_2^-), (x_3^+, x_3^-), (x_4^+, x_4^-)\}$. Thus in the next step we find the matched nodes which are simply the target nodes of max matched edges and in our case all the nodes are matched $\{1, 2, 3, 4\}$. Hence, this is a trivial case of *Perfect matching*. From, the *Minimum input theorem 1.15*, we have that the minimum number of input vertices for this case is simply $N_I = 1$. And so the directed graph in the end can be controlled as shown in the Fig. 1.8.

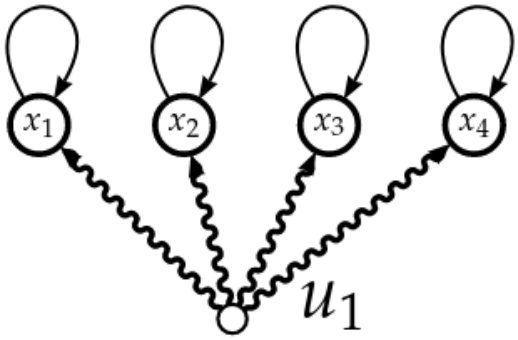


Figure 1.8

1.4.5. Minimum Control energy:

As we have already seen in the section 1.1, it is possible to steer(or control) the system from an initial state to a final state in specified number of time steps using input control matrix $\mathbf{u}$. A metric which captures the amount of energy to steer systems from one state to another, be they general systems (electronic, mechanical, social) or specific biological systems (brain regions in our case), is defined as[8]:

$$E(\mathbf{u}) = \sum_{t=0}^{t_f - 1} ||\mathbf{u}(t)||^2 \tag{1.16}$$

which is energy of discrete time control input $\mathbf{u}$(t), and in its continuous form given by:

$$E(\mathbf{u}) = \int_{t=0}^{t_f} ||\mathbf{u}(t)||^2 \tag{1.17}$$

An optimal input vector $\mathbf{u}^*(t)$ is aimed at minimizing the control energy given by Eq. 1.16,1.17 required to steer the system from any initial to any desired final state. The control energy can be directly obtained from Controllability Gramian W[9]:

$$E_{min} = \mathbf{x}_{final}^T W^{-1} \mathbf{x}_{final} \tag{1.18}$$

Thus Eq. 1.18 indicates that the control energy not only depends on the network A and input matrix B but also on the final desired state to be reached $\mathbf{x}_{final}$.

Minimal control energy: It is the minimum amount of energy required to steer the system or to drive the network along the worst possible eigen-directions of Gramian matrix W[8](which is obtained by solving Lyapunov equation 1.9 or 1.11). It is computed as:

$$E_{min} = 1/\lambda_{min}(W) \tag{1.19}$$

where $\lambda's$ are simply the eigenvalues of controllability gramian W.

1.5. Algorithm implementation for Kalman and Structural Controllability:

1.5.1. Kalman controllability:

- Load the adjacency matrix A from file. If the element a_{ij} refers to directed edge going from node i to j, then transpose it, otherwise leave it as it is. By notation and in general, a_{ij} here should always refer to directed edge from node j to node i.
- Add a check to see if the eigenvalues of A lie in the range ensuring stability of the linear system, i.e., if they lie in the unit circle for discrete LTI or if the maximum of real part of the eigenvalues is negative for continuous LTI. If these conditions are not satisfied, we cannot proceed to test controllability.
- Call a subroutine to construct input matrix B_K which is of dimension (N, r) where N is dimension of A and K is the set of control points or driver nodes such that $K = \{k_1, .., k_r\}$.
$$B_K = [e_{k_1}, ..., e_{k_r}] \tag{1.20}$$
where e_{k_i} is the $i - th$ canonical vector of dimension N.
- Call a subroutine to compute Controllability matrix C as given in Eq. 1.7 by giving input A which is already known and B_K as computed from another function as given in Eq. 1.20
- *Test Single node controllability.* With $B_K = [e_k]$ constructed from a single node input, Kalman's rank condition 1.8 is checked to see if it is satisfied. This is done for each node of the network and list of single nodes is prepared which satisfies rank condition and only for those single nodes *minimum control energy* Eq. 1.19 is computed. If none of them satifies Eq. 1.8, then it is said that *Single node controllability* is not possible for this network.
- *Test Unmatched nodes controllability.* The concept of maximum matching is borrowed from *Structural controllability* topic which is a stronger version of controllability. *Unmatched nodes being driver nodes* of the network are computed

for the digraph as explained in the previous section, B_K being a input matrix of 1's and 0's according to Eq. 1.20 where 1's are present exactly at the row corresponding to node index number of the driver node and the number of columns of B equal to cardinality of set of driver nodes. Again, rank condition is checked and if satisfied then *minimum control energy* is computed. In this procedure, we are controlling a system through a *finite fraction of nodes* where the finite fraction is given by unmatched nodes.

- *Test All nodes controllability.* In this case, we are controlling all the nodes N of the network. So, B_K of Eq. 1.20 simply becomes identity matrix of dimension (N, N) where r = N here. Kalman's controllability rank condition is checked and if satisifed, the *min. control energy* is computed giving us the energy required to steer all the nodes of network to a given state.

1.5.2. Structural controllability via minimization of cost function:

- Similar to the first step of Kalman controllability algorithm, load the adjacency matrix A from file. If the element a_{ij} refers to directed edge going from node i to j, then transpose it, otherwise leave it as it is. By notation and in general, a_{ij} here should always refer to directed edge from node j to node i.

- Add a check to see if the eigenvalues of A lie in the range ensuring stability of the linear system, i.e., if they lie in the unit circle for discrete LTI or if the maximum of real part of the eigenvalues is negative for continuous LTI. If these conditions are not satisfied, we cannot proceed to test controllability.

- Convert the directed graph to a bipartite graph and apply maximum matching's Hopcroft-Karp algorithm to find unmatched nodes. These unmatched nodes are the control nodes or driver nodes which has the possibility to steer the entire network.

- Initialize the input signal matrix $B \equiv B(N, r)$ which is a matrix consisting of independent parameters $\Theta = \{b_1, b_2, ..., b_r\}$ **instead of** 1's and 0's as defined in Eq. 1.20. With a slight abuse of notation let us define input matrix B initialized using Eq. 1.20 as B_{kalman} while instead B as an input matrix of independent parameters as B_{struct}. As an example, let us consider Example 1 of *Maximum matching examples* of subsection 1.4.4. If we consider unmatched nodes(or driver nodes) set $N_D = \{1, 2, 3, 5\}$, then B_{kalman} according to Eq. 1.20 and B_{struct} will be written as:

$$B_{kalman} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad B_{struct} = \begin{pmatrix} b_1 & 0 & 0 & 0 \\ 0 & b_2 & 0 & 0 \\ 0 & 0 & b_3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & b_4 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

where b_1, b_2, b_3, b_4 are free parameters whose values have to be determined.

- The task of determining parameters of B_{struct} is done via minimization of cost function[10] where cost function here refers to the metric *Minimum control energy* given by Eq. 1.19, with the constraints that Kalman's rank condition Eq. 1.8 is satisfied and minimum eigen-value of controllability gramian W is strictly positive. Given (A,B) as the system, $\Theta \equiv \Theta(B)$ collection of parameters of $B \equiv B_{struct}$, the cost function minimization is

$$\min_{\Theta} f_{\Theta}(A, B) = \min_{\Theta} E_{min}(W(A, B)) \tag{1.21}$$

such that the following constraints are satisfied:

$$rank(C) = N \tag{1.22}$$
$$\lambda_{min}(W(A, B)) > 0 \tag{1.23}$$

where $C = [B, AB, A^2B, ..., A^{N-1}B]$ is controllability matrix of dimension $(N, N.r)$.

Finally after obtaining optimized values of the elements of the input signal parameter matrix B_{struct}, using Eq. 1.19 compute minimum control energy using B_{kalman} and $B_{struct} = B_{optimal}$ and compare the energy values to see that $E_{min}^{struct} \leq E_{min}^{kalman}$.

2

Effective connectivity

In this chapter, we begin by discussing about implications of using Structural connectivity matrices for cognitive control by using the concepts of control theory. However, we provide an argument against usage of SC matrices and instead give an alternative approach of using Effective connectivity matrices which embed the dynamics of the neural populations and not only the network topology. In the second section, we henceforth provide three different models to obtain EC matrices from BOLD fMRI time series data, which can be used for brain controllability by applying the concepts that were introduced in the first chapter. Before discussing the models, we briefly introduce the used data (LEMON dataset).

2.1. Control based on structural connectivity

Let us start by precisely defining 3 kinds of connectivity matrices which have to be clarified before we move on.

Functional connectivity. (FC) The correlation of the resting state time series of single brain region with respect to time-series of all the other regions of brain gives rise to a map of functional connectivity(abbreviated as *fcMap*) and it describes the functional connections between the brain regions. It refers to the temporal dependence of neural dynamics of spatially/anatomically seperated brain regions[11].

Structural connectivity. (SC) The structural connections in the brain always refers to *white matter tracts* bundles of millions of long distance axons directly interconnect large groups of spatially seperated neurons. These white matter tracts are the ones responsible for exchange of information and acts as chan-

nels for transport of large amount of functional data between spatially seperated regions.[11]

Effective connectivity. (EC) It refers explicitly to influence that one neural system exerts over another, either at synaptic or at population level.[12]. It "should be understood as experiment and time-dependent, simplest possible circuit diagram that would replicate the observed timing relationships between recorded neurons"[13]. The so called *circuit diagram* referred by the researchers actually means network/graph and *replicate the observed timing relationships between neurons* refers to creating a model and using the adjacency matrix built for such an *Effective connectivity* network, it should be possible to predict the signal data of all the regions considered in the network.

Cognitive control refers to the ability to pursue goal-directed behaviour, in the face of otherwise more habitual or immediately compelling behaviours. This ability is engaged by every faculty that distinguishes human abilities from those of other species, and in virtually every domain of human function from perception to action, decision making to planning, and problem solving to language processing[14]. In the cognitive control, the dynamics of the brain network are altered to enable humans to perform complex tasks. A relationship between the mathematical notions of control theory and the notion of cognitive control of brain systems was first established by *Gu et al.*[15]. They have postulated that controllability of the network using the Control Theory(as mentioned in the previous Chapter) can be a possible mechanism of cognitive control wherein a particular set of nodes(or brain regions) present at critical anatomical positions of brain acts as driver nodes to steer the brain system to a desired state or specific modes of action i.e. cognitive functions.

But there exists a major drawback in the approach of Gu and collaborators[15], who have replaced A in Eq. 1.1 with the SC matrix M directly derived from *Diffusion Tensor Imaging*(DTI) or *Diffusion Spectrum Imaging*(DSI) data. *Tu et al.*[16] have clearly pointed out few critical issues on their approach:

1. Linear dynamics approach used to capture the neural activity of the brain dynamics during its resting state should be given by:

$$A = (1 - \alpha\Delta t)I + cM\Delta t$$

   - α: inverse relaxation time
   - Δt: time step to update the brain state
   - M: Anatomical connectivity matrix of brain
   - I: Identity matrix
   - c: Normalization constant

The relation between the linearized dynamics specified by matrix A and structural connectivity matrix M is model dependent: *Tu et al.*[16] used the above linear model, while *Gu et al.*[15] directly replaced A with M. The criticism for this approach of directly replacing A with M is that it reduces the analysis of dynamical model to a simple analysis of a graph and so we end up assessing dynamical features from a strictly static data.

2. *Gu et al.*[15] has adopted to normalize A

$$A_{norm} = A/(1 + \lambda_{max}(A))$$

where $\lambda_{max}(A)$ absolute of largest eigenvalue of A, so as to ensure that the controllability metrics are well defined. But the criticism for this is that there is no biological meaning attached to this procedure and it is just for mathematical convenience.

3. *Gu et al.*[15] did not perform any comparison with the null models which is a standard approach in network science[17, 18].

Two of the metrics used to measure controllability of brain system as defined by *Gu et al.*[15] are:

Average controllability: It is defined as the average of the input energy from a set of control nodes and over all possible target states. It is mathematically computed as $Trace(W_k)$ where W_k is the controllability gramian of k-th node.

Modal controllability: It is referred as ability of the node to control the dynamics of network towards difficult to reach states i.e. those states that have higher energy with respect to other states. Mathematically it is given as, for an i-th node/brain region:

$$\phi_i = \sum_{j=1}^{n} (1 - \lambda_j(A)^2) v_{ij}^2$$

where v_{ij} is $(ij) - th$ element of the eigenvector matrix of A while $\lambda_i's$ are the eigenvalues of A.

Having defined measures to compute single node controllability, we discuss the points givenforth by *Tu et al.*[16] on why it is improper to use such metrics:

1. Dynamical system in its discrete LTI is of the form we know: $\mathbf{x}(t+1) = A\mathbf{x}(t) + B\mathbf{u}(t)$. *Average controllability* $Trace(W_k)$ assesses the control role of the $k - th$ node (or brain region), in terms of its ability to coordinate with activity of other brain regions when an external input u_k is applied on it. However, the common interpretation that the internal activity of k (without external input) can easily control other regions is incorrect.

2. The formulation of *modal controllability* contains just the eigenvalues and eigenvectors of A which is simply a connectivity matrix from the linearized system dynamics. However, there is no mention of W which is controllability gramian matrix which is a crucial quantity in Control theory framework.

3. As pointed out by *Tu et al.*[16] in their paper, the minimum eigenvalue ($\lambda_{min}(W_k)$) of controllability gramian of k-th region W_k computed for various datasets, are very close to zero. This means that minimum energy E_{min} required to control the system with only that particular single node k is extremely high ($\longrightarrow \infty$) and so it is impossible to control the system in practice. Thus, single node controllability which was said to be possible by *Gu et al.*[15] in their paper is unfeasible. Further evidence for this has also been provided by Tu and collaborators[16] who have used high precision data on C.Elegans connectomes[19] and proved that even small connectomes cannot be controlled by a single node because of this reason.

4. Tu and collaborators[16] have found that even after randomizing the networks, the relationship between average/modal controllability with respect to weighted degree of the network remains the same as the result obtained from the original anatomical connectivity datasets. See Fig. 2.1.

   This gives evidence that such metrics for controllability cannot be used to determine a *specific or optimized network organization* in the brain. In fact, these quantities are analytically correlated with node degree[16].

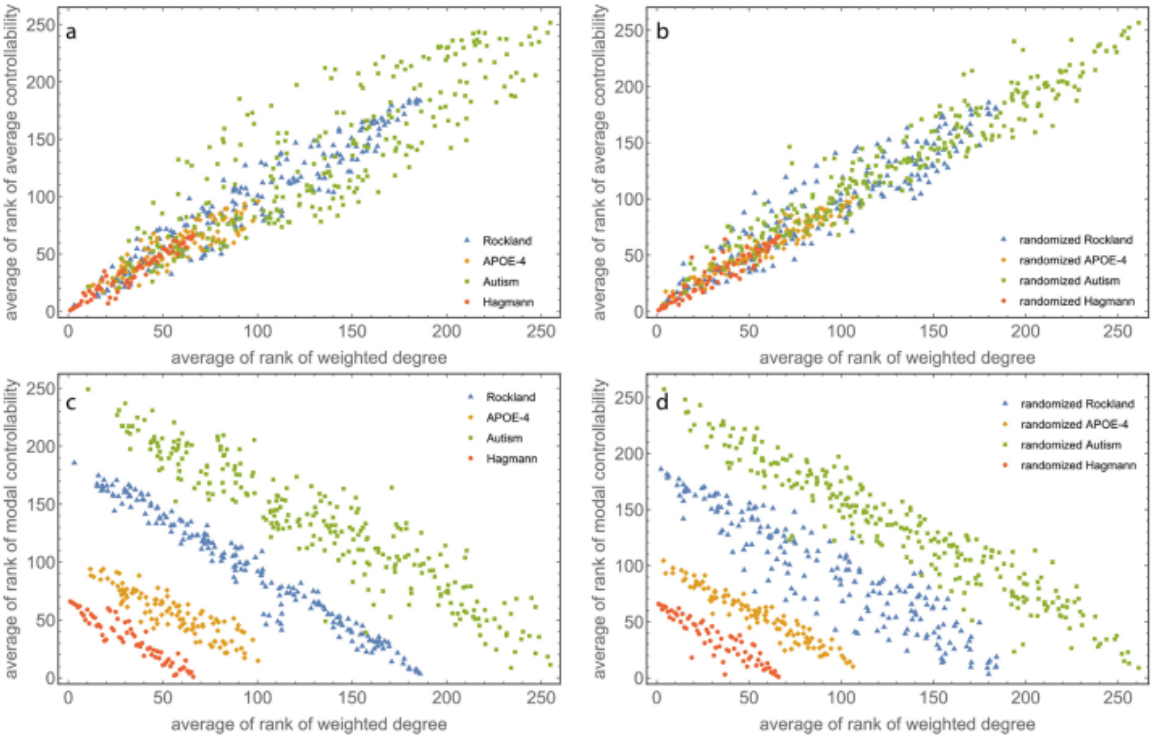


Figure 2.1: Comparison of plots the controllability measures such as average*(top panel)*/modal*(bottom panel)* controllability with respect to ranked weighted degree of 4 original datasets*(top left, bottom left)* and their randomized counterpart*(top right, bottom right)*. *Image reproduced with permission from Tu et al.[16]*

2.2. Linear Brain Dynamics

Despite all the arguments put forth against the methods and framework discussed by *Gu et al.*[15] who were the first to discuss well established network control theory to the context of congnitive systems, their research provides us with necessary room for improvement. An alternative approach to capture the dynamical complexity of the neural systems is to use *Effective Connectivity matrices* instead of *Structural connectivity matrices* which only capture the topology of the brain network. In this chapter, the focus is on obtaining *Effective connectivity matrix A* and implementing the concepts of controllability that was discussed in the previous chapter.

In the initial part we try to see why linear dynamics is simpler, easier and at the same time preferable to use on brain dynamics as compared to non-linear models defined on time series data of neuronal activity of macroscopic regions of brain. With linear dynamics followed by brain, effective connectivity matrix A can be obtained by modelling via *Bassett's Linear model*[20], *Gilson's Multivariate Ornstein-Uhlenbeck(MOU) model* and *Prando's Sparse DCM* . Earlier research consisted of using structural connectivity matrices for controllability [15],[16] but here we are going to use EC matrices instead so as to apply the concepts of controllability discussed in the previous chapter. Before introducing the linear models, we briefly discuss the dataset used (LEMON dataset).

2.2.1. Dataset and Preprocessing

The data of the *Leipzig Study for Mind-Body-Emotion Interactions* (LEMON; Babayan, Anahit, et al. *A mind-brain-body dataset of MRI, EEG, cognition, emotion, and peripheral physiology in young and old adults.* Scientific data 6.1 (2019): 1-21.) comprised 227 subjects in two age groups (younger: 2035, older: 5977).

Brain imaging was performed on a 3T Siemens Magnetom Verio MR scanner (Siemens Medical Systems, Erlangen, Germany) with a standard 32-channel head coil. The participants were instructed to keep their eyes open and not fall asleep while looking at a low-contrast (light grey on dark grey background) fixation cross. The structural image was recorded using an MP2RAGE sequence*Marques et al.* [21] with the following parameters: TI 1 = 700 ms, TI 2 = 2500 ms, TR = 5000 ms, TE = 2.92 ms, FA 1 = 4°, FA 2 = 5°, bandwidth = 240 Hz/pixel, FOV = 256 Œ 240 Œ 176 mm^3, voxel size = $1 \times 1 \times 1$ mm^3. The functional images were acquired using a T2*-weighted multiband EPI sequence with the following parameters: TR = 1400 ms, TE = 30 ms, FA = 69°, FOV = 202 mm, imaging matrix = 88 Œ 88, 64 slices with voxel size = $2.3 \times 2.3 \times 2.3$ mm^3, slice thickness = 2.3 mm, echo spacing = 0.67 ms, bandwidth = 1776 Hz/Px, partial fourier 7/8, no pre-scan normalization, multiband acceleration factor = 4, 657 vol, duration = 15 min 30 s. A gradient echo field map

with the sample geometry was used for distortion correction (TR = 680 ms, TE 1 = 5.19 ms, TE 2 = 7.65 ms).

Data pre-processing was performed using the SPM5 software package (Wellcome Department of Cognitive Neurology, London, UK) with the following steps:

1. correction for slice-timing differences
2. correction of head-motion
3. co-registration of the anatomical image and the mean functional image
4. spatial normalization of all images to the MNI space with a voxel size of $3 \times 3 \times 3\ mm^3$
5. spatial *Independent Component Analysis (ICA)* of the BOLD time-series in MNI space for the removal of artifacts due to blood pulsation, head movement and instrumental spikes.

2.2.2. Linear model approach to extract EC

In this section, among the many computational model we focus on *Bassett's Linear model*[20] and extract EC matrix which can be later used for controllability. The ODE models applied to neuroscience problems are of the form given by:

$$\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = f(\mathbf{x}(t)) + \mathbf{e}_1(t) \tag{2.1a}$$

$$\mathbf{y}(t) = h(\mathbf{x}(t)) + \mathbf{e}_2(t) \tag{2.1b}$$

with initial condition $\mathbf{x}(0) = \mathbf{x}_0$

- $\mathbf{y}$(t) is time series of recorded brain activity obatined either by *rsfMRI (resting state functional Magnetic Resonance Imaging)* whose dimension is *(m, T)* where T is the total number of time steps while m is the number of brain regions or *parcellations*.
- $\mathbf{x}$(t) is time series of *internal* brain states of dimension *(N, T)*. It is not compulsory that $N = m$.
- f and h are non-linear functions
- $\mathbf{e}_1$(t) and $\mathbf{e}_2$(t) are process and measurement noise time series of dimensions *(N, T)* and *(m, T)* respectively.

There is no presence of external input **u**(t) in the Eq. 2.1 because the brain activity is recorded in the *resting state condition* of the subjects. Since the sampling of data of the brain activity of a given subject is a time series **y**(t) which is always discrete, it is necessary to transform Eq. 2.1a in terms of discrete-time dynamics i.e. by approximating $\dot{\mathbf{x}}(t)$ as a first difference $\mathbf{x}(t) - \mathbf{x}(t-1)$. Thus, in discrete time dynamics Eq. 2.1 becomes:

$$\mathbf{x}(t) - \mathbf{x}(t-1) = f(\mathbf{x}(t-1)) + \mathbf{e}_1(t) \tag{2.2a}$$

$$\mathbf{y}(t) = h(\mathbf{x}(t)) + \mathbf{e}_2(t) \tag{2.2b}$$

with $t \in \mathbb{Z}$, and $t = 0, 1, 2, ..., T$ but note that time is measured in units of Δt where Δt is the sampling time of fMRI, which corresponds to roughly 1s.

Since f and h are non-linear functions in Eq. 2.2, to avoid the complications associated with them, a linear class of models can be defined by modeling the dynamics directly at the *BOLD/LFP* level *(BOLD - Blood Oxygen Level Dependent, LFP - Local Field Potential)* by considering $\mathbf{y}(t) = \mathbf{x}(t)$ and combining the noise signals as $\mathbf{e}(t) = \mathbf{e}_1(t) + \mathbf{e}_2(t)$ and considering this $\mathbf{e}(t)$ as a *white noise*. With this, Eq. 2.2 gets simplified to:

$$\mathbf{y}(t) - \mathbf{y}(t-1) = f(\mathbf{y}(t-1)) + \mathbf{e}(t) \tag{2.3}$$

If the non-linear function f is taken to be linear of the form $f(\mathbf{y}(t)) = \mathbf{A}\mathbf{y}(t)$ where **A** is called *Effective connectivity matrix of brain regions* of dimension (N, N), then Eq. 2.3 gets simplified to:

$$\mathbf{y}(t) - \mathbf{y}(t-1) = \mathbf{A}\mathbf{y}(t-1)) + \mathbf{e}(t) \tag{2.4}$$

The fMRI data is *Markovian*. A simple 1-lag linear model as given in Eq. 2.4 would be enough to make a fit since which means that $\mathbf{y}(t-1)$ contains almost all the information to predict the future data $\mathbf{y}(t)$, while $\mathbf{y}(t-2)$ contains little information and previous values of $\mathbf{y}(t)|t < (t-2)$ contain almost no information required for the prediction.

A fit can be performed on the time series data after simply dividing into **training-test set** along the time axis of the recorded data(as in our case fMRi data). There are 2 linear models so as to obtain either *Dense EC*, A_{dense} or *Sparse EC*, A_{sparse}. The Eq. 2.4 can be rewritten as

$$\mathbf{A}\mathbf{y}(t-1) = \mathbf{y}_{diff}(t) \tag{2.5}$$

where $\mathbf{y}_{diff}(t) = \mathbf{y}(t) - \mathbf{y}(t-1)$.
Using only the training dataset, when *Linear least squares minimization method* is used to solve Eq. 2.5, we obtain $\mathbf{A_{fit}} = \mathbf{A_{dense}}$ but when we use *LASSO regularization method* to solve the same, we obtain $\mathbf{A_{fit}} = \mathbf{A_{sparse}}$ and the method improves the

sparsity of the EC matrix. Thus, in the Linear model itself we have *Linear-dense* and *Linear-sparse* models. To check the quality of the fit and also to compare the models, R^2 parameter has been used by Nozari[20]. The error obtained when subtracting the original data from the prediction value of the fitted model is given as:

$$E_{test}(t) = \hat{\mathbf{y}}_{test}(t|t-1) - \mathbf{y}_{test}(t) \tag{2.6}$$

where $\hat{\mathbf{y}}(t|t-1)$ is the estimate of $\mathbf{y}$(t) by the linear model given the previous values $\{\mathbf{y}(\tau)|0 \leq \tau \leq (t-1)\}$. With the error term defined as in Eq. 2.6, the R^2 can be computed by:

$$R_i^2 = 1 - \frac{\sum_t E_i^{test}(t)^2}{\sum_t (y_i^{test} - \overline{y}_i^{test}(t))^2} \tag{2.7}$$

where $i = 1, 2, ...N$ representing N channels(i.e. each channel for each brain region), $\overline{y}_i^{test}(t)$ is the temporal average of time series for $i-th$ region of brain, y_i^{test}.

- $R_i^2 = 1$ indicates perfect model for channel i.
- $R_i^2 = 0$ indicates model which is as good as *constant predictor* i.e. $y_i(t)$ is always equal to its temporal average.

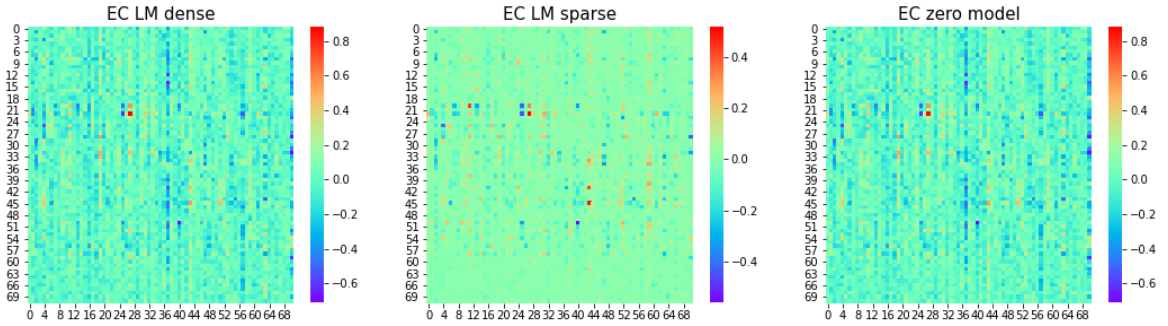


Figure 2.2: EC matrices obatined using Linear model on *sub-010002 of LEMON dataset*

2.2.3. Linearity VS Non-linearity

As mentioned in various literature of neuroscience[22], [23], [24], [25], modelling via nonlinear dynamics is well established at the micro-scale of individual neurons and it is possible also at the meso scale of neuronal populations. This might suggest that nonlinear dynamics can be used also at the macro scale i.e. brain regions. However, Bassett in her paper[20] gives an argument that a very simple yet powerful approach of linear modelling of dynamics of the brain regions is itself enough, instead of non-linear modelling which may further add to the complications. *There are 4 properties of macroscopic neurodynamics and neuroimaging which counteracts or masks the nonlinearities of the neuronal dynamics at the micro level*[20].

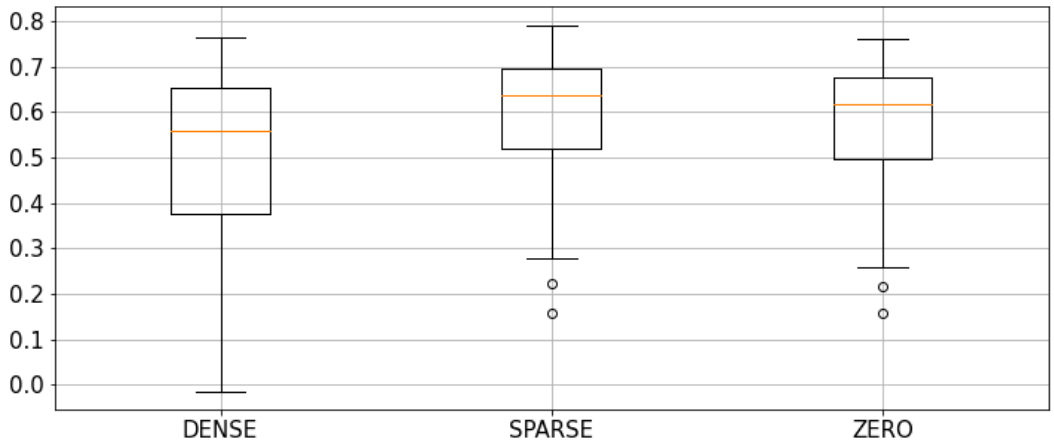


Figure 2.3: R^2 parameter observed for 72 brain regions combined of *sub-010002 of LEMON dataset*. t-test for R^2-Dense and R^2-Zero: t-statistic = -2.389, pvalue=0.018, t-test for R^2-Sparse and R^2-Zero: t-statistic = 0.948, pvalue=0.345. We see that zero model seems to be better than the Linear dense from this result, which contradicts our expectation of Linear model faring better than Null/Zero model which does not consider any previous information to predict.

1. **Averaging over space.** The signals detected by the imaging tools for measuring the macroscopic brain dynamics are an average over the activity of a large number of neurons whose population is of the orders of $\sim 10^2$, $\sim 10^3$ or even $\sim 10^6$. If these neurons or *unit*(neuronal populations in small scale) are either not correlated or their correlations decay with physical distance from one another, then this *spatial averaging* due to use of imaging tools, can weaken the nonlinearities and on the contrary *linearizing effects* could be observed because of this.

2. **Averaging over time.** The BOLD signals captured by the fMRI is always low-pass filtered. The effect of low pass filtering is *temporal averaging* which weakens the nonlinearities observed. Similar to the concept of *spatial correlation* as in *Averaging over space*, here we instead have *autocorrelation* which is, given a signal from a channel(or brain region), it determines how the correlation between the observed signal value and another value at a temporal distance Δt of the time series data decays where $0 \leq \Delta t \leq T$. Stronger low-pass filtering nullifies the non-linear relationships between these signals.

3. **Observation noise.** fMRI/iEEG time series data are affected by observation noise[26], [27]. The nonlinear relationships in the time series gets masked by this. When Signal-to-Noise ratio, $SNR \sim 1$ even in the absence of both *spatial and temporal averaging* it masks completely the nonlinear relationship. It has been shown that in reality $2 \lesssim SNR \lesssim 14$ when the linearizing effect of observation noise is combined with both spatial and temporal averaging for the rsfMRI time series data.

4. **Limited data samples.** The non-linear models are overly dependent on the size of the time series data[20] unlike linear models. There is an inherent loss in the predictive power of the non linear models with increasing dimensionality(which is number of channels or number of regions N). The sample com-

plexity which is number of time series data points required for detecting non-linear relationships is exponentially large for higher dimensions.

In practice, when all the 4 properties described above are combined, they render the non-linearities of the neuronal populations at a macro level ineffective and so linear models are optimal for the time series data of fMRI.

2.2.4. DCM approach to extract EC

Dynamical Causal Modelling(DCM) was first introduced by *Friston et al.*[28]. It is a generative model of measured brain responses that encaptulates both the non-linear and dynamic nature of the brain signals. This is a multiple-input-and- multiple-output(*MIMO*) model where the output *haemodynamic responses* are evoked either by providing known inputs via experimental manipulations, or by neural fluctuations signals arising spontaneously during the resting state. This haemodynamic response is increased flow of blood, oxygen and other nutrients to the nerve cells(or populations of neurons) *in response* to stimulation of neuronal tissues. Since these neuronal populations do not have reservoir of stored energy substrates, there exists a large demand in terms of oxygen, glucose whenever there is any activity among these neurons. This coupling between neuronal activity and blood flow is also referred to as *neurovascular coupling* and this is quantified by Haemodynamic response function(*HRF*).

This DCM consists of 2 parts:

1. *ODE* of coupling among the neuronal populations
2. a mapping from neuronal activity (which is the input) to the output BOLD fmri response signal $y(t)$.

The above can be translated mathematically as[29]:

$$\dot{x}(t) = f(x(t), u(t); \theta_f) + \nu(t) \tag{2.8a}$$

$$y(t) = h(x(t); \theta_h) + e(t) \tag{2.8b}$$

The terms and notations in Eq. 2.8 are:

- $x(t) = [x_1(t)...x_n(t)]^T$ is hidden neural activity of n brain regions at time t.
- $u(t)$ denotes external input such as external stimulations or task demands.
- θ_f denotes parameters that describe the model at neuronal level.
- $\nu(t)$ represents stochastic term taking into account intrinsic brain fluctuations.
- $y(t)$ is the BOLD fmri response or haemodynamic response signal at time t.

- θ_h denotes biophysical parameters that define haemodynamic response.
- $e(t) \sim \mathcal{N}(0, R)$ is observation noise with covariance matrix R.

In the first DCM formulation by *Friston et al.*[28], given Eq. 2.8a, the neural activity is assumed to depend only on external stimuli $u(t)$ and the stochastic term $\nu(t)$ is neglected. The function $f()$ is of the bilinear form. The Eq. 2.8a is modified according to Friston as:

$$\dot{x}(t) = \Big(A + \sum_{j=1}^{m} u_j(t) B_j\Big) x(t) + Cu(t) \tag{2.9}$$

The terms and notations in Eq. 2.9 are:

- $\theta_f = \{A, B_1, ..., B_m, C\}$ encodes the couplings of neural activity and external inputs.
- A represents Effective connectivity matrix(in the absence of external stimuli)
- B_j represents change in neuronal coupling with respect to input j.
- C represents the direct influence of experimental manipulations on neural activity $x(t)$.

Friston in 2014[30] again introduced modifications to Eq. 2.9 to allow for its application to resting state fMRI data where external stimuli is obviously absent since it is a "resting state", $u(t) = 0$. Instead, $\nu(t)$ which was neglected in Eq. 2.9 is reconsidered to be the main driving agent for the neural activity $x(t)$ and so in the end the function $f()$ which was of bilinear form in Eq. 2.9 becomes linear.

$$\dot{x}(t) = Ax(t) + \nu(t) \tag{2.10}$$

The *Haemodynamic Response Function*(*HRF*), given by $h()$ in Eq. 2.8b is modelled by a set of 5 equations which has non-linear and biophysically relevant terms embedded into them so as to map and model the dynamics of experimentally observed haemodynamic response.

For a given i-th region, where $i = 1, ..., n$ in the following set of equations, we have:

$$\dot{r}_i(t) = x_i(t) - \kappa_i r_i(t) - \eta_i(f_i(t) - 1) \tag{2.11a}$$

$$\dot{f}_i(t) = r_i(t) \tag{2.11b}$$

$$\tau \dot{\nu}_i(t) = f_i(t) - \nu_i^{1/\xi_i(t)}(t) \tag{2.11c}$$

$$\tau \dot{q}_i(t) = (f_i(t)/\rho_i)[1 - (1 - \rho_i)^{1/f_i(t)}] - \nu_i^{1/\xi_i(t)-1}(t) q_i(t) \tag{2.11d}$$

$$b_i(t) = V_0 k_1 (1 - q_i(t)) + V_0 k_2 (1 - q_i(t)/\nu_i(t)) + V_0 k_3 (1 - \nu_i(t)) \tag{2.11e}$$

where in the last equation, neural activity $x_i(t)$ is taken as input and BOLD signal $b_i(t)$ is given as an output in this formulation.
$\{r_i, f_i, v_i, q_i\}$ are the 4 haemodynamic states which are biophysical quantities while $\theta_h = \{\kappa_i, \tau_i, \eta_i, \xi_i, \rho_i\}$ are the biophysical parameters defining haemodynamic response.
Explanation of these terms and notations which are direct representation of neurobiological concepts/terms is crucial in understanding the formulation of HRF. As explained by Friston[28]:

- Haemodyamic states:
  - r_i: vasodilatory signal
  - f_i: blood inflow
  - v_i: blood volume
  - q_i; deoxyhaemoglobin content

  In brief, for the i-th region, neuronal activity x_i causes an increase in a vasodilatory signal r_i that is subject to auto-regulatory feedback. Inflow f_i responds in proportion to this signal with concomitant changes in blood volume v_i and deoxyhemoglobin content q_i.[28]

- Biophysical parameters θ_h:
  - κ: rate of signal decay
  - η: rate of flow dependent elimination
  - τ: Haemodynamic transit time
  - ξ: Grubb's exponent
  - ρ: Resting oxygen extraction fraction.

- $V_0 = 0.02$ typically and it represents resting blood volume fraction
- k_1, k_2 and k_3 are constants.

The EC matrix is estimated by inverting the DCM which uses the measured fMRI data. For the resting state fMRI, when the neural activity is modelled by Eq. 2.10 the inversion of DCM becomes more comples to deal with. *Prando et al.*[29] dealt with such a problem by further simplifying the framework of DCM as well as the procedure for its inversion. This simplified model called *Linear model* is more robust and also reduces the computational complexity of the model. This model modifies the original DCM for rsfMRI[30] on two accounts:

1. **Discretization** which is supplemented by low temporal resolution of fMRI scanners where the time interval of measurement of BOLD signals is $T_R \sim 0.5s$ to $3s$. Eq. 2.10 is rewritten as:

$$\begin{aligned} x(kT_R + T_R) &= e^{A(kT_R+T_R-kT_R)}x(kT_R) + \int_{kT_R}^{kT_R+T_R} e^{A(kT_R+T_R-s)}\nu(s)ds \\ &= e^{AT_R}x(kT_R) + \int_0^{T_R} e^{A\tau}\nu(\tau)d\tau \end{aligned} \tag{2.12}$$

Let $x(kT_R) = x(k)$ and $w(k) = \int_0^{T_R} e^{A\tau}\nu(\tau)d\tau$. The above Eq. 2.12 transforms to:

$$x(k+1) = e^{AT_R}x(k) + w(k) \tag{2.13}$$

If we assume that the stochastic term $\nu(t)$ where $t \in R$ is a *white gaussian noise* with variance $\sigma^2 I_n$ where I is identity matrix of size n, then $w(k)$ is also white gaussian and its corresponding variance is given by

$$Q = \sigma^2 \int_0^{T_R} e^{A\tau}e^{A^T\tau} \tag{2.14}$$

2. **Linearization** of original non-linear continuos time model. The set of equations 2.11 which governs the haemodynamic response by considering Finite Impulse response for brain region i, $h_i = [h_{i,0}, ..., h_{i,s-1}]^T$ with s whose value is large enough to maintain temporal dependencies. This FIR model takes neural state $x_i(k)$ as an input and gives the BOLD signal as an output.

$$b_i(k) = \sum_{l=0}^{s-1} h_{i,l}x_i(k-1) \tag{2.15}$$

where $i = 1, ..., n$ being the number of brain regions and $b_i(k) = b_i(kT_R)$ is the simplification considered in above equation. By observing the empirical prior distributions of the *biophysical parameters* θ_h, the FIR h_i is assigned a *Gaussian prior distribution* $h_i \sim \mathcal{N}(\mu_i, \Sigma_i)$. Moving from non-linear stochastic model as given in Eq. 2.8 to linear stochastic model as fomulated by *Prando et al.*[29]:

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{w}(k) \tag{2.16}$$

$$y(k) = \mathbf{H}\mathbf{x}(k) + e(k) \tag{2.17}$$

where the terms are defined as:

(a) $\mathbf{x}(k) = [x^T(k)\ x^T(k-1)\ ...\ x^T(k-s+1)]^T$ where $\mathbf{x}(k) \in R^{ns}$

(b) $\mathbf{w}(k) = [w^T(k)\ 0]^T$ where $\mathbf{w}(k) \in R^{ns}$

(c) $\mathbf{A} = \begin{bmatrix} e^{AT_R} & 0 \\ I_{n(s-1)} & 0 \end{bmatrix}$

(d) Block diagonal matrix $\mathbf{H} = [\mathbf{H}_0\ \mathbf{H}_1\ ...\ \mathbf{H}_{s-1}\ ]$ where

$$\mathbf{H_i} = \begin{bmatrix} h_{1,i} & & & \\ & h_{2,i} & & \\ & & \ddots & \\ & & & h_{n,i} \end{bmatrix} \text{ being diagonal matrix with } i = 0, 1, ..., (s-1)$$

Assumptions were made on stochastic terms:

$$\mathbf{w}(k) \sim \mathcal{N}(0, \mathbf{Q})$$
$$e(k) \sim \mathcal{N}(0, R)$$

where $\mathbf{Q} = \mathbf{blkdiag}(\mathbf{Q}, \zeta I_{n(s-1)})$ is block diagonal matrix and $R = \mathbf{diag}(\lambda_1^2, ..., \lambda_n^2)$ is diagonal matrix. To guarantee that $\mathbf{Q}$ is invertible, $\zeta = 10^{-15}$ is chosen.

The parameters $\theta = \{A, \sigma, h_1, ...h_n, \lambda_1, ...\lambda_n\}$ which specify the linear model Eq. 2.16 is estimated via *Bayesian learning framework* by taking into account the measurements of the BOLD signal $y(k),\ k = 1, ..., n$ being the brain regions. From the most fundamental *Bayes theorem* we know that: *Posterior* $\propto$ *Likelihood* $\times$ *Prior*. Based on the "prior" knowledge of the haemodynamic responses[28] or to support the sparsity of the EC matrix, a prior $p_\gamma(\theta)$ is assigned where γ are the hyperparameters to be determined via this framework.

$$p_\gamma(\theta|\mathbf{Y}) \propto p(\mathbf{Y}|\theta) \times p_\gamma(\theta) \tag{2.18}$$
$$p_\gamma(\theta|\mathbf{Y}) = \int p_\gamma((X), \theta|(Y))d\mathbf{X} \tag{2.19}$$
$$p(\mathbf{Y}|\theta) = \int p(\mathbf{X}, \mathbf{Y}|\theta)d\mathbf{X} \tag{2.20}$$

where the terms and notations are:

1. $\mathbf{Y} = [\mathbf{y}^T(1)....\mathbf{y}^T(n)]^T$ is *measured variable* representing BOLD signal.
2. $\mathbf{X} = [\mathbf{x}^T(1)....\mathbf{x}^T(n)]^T$ is *latent/hidden variable* where the neural activity is unknown.

Instead of trying to compute high dimensional integral for model posterior as in Eq. 2.19, a workaround is via using Eq. 2.18 with the lower bound of the likelihood as given in Eq. 2.20. In Eq. 2.18, the likelihood $p(\mathbf{Y}|\theta)$ is defined as in Eq. 2.20 and the only missing term to be defined is the *prior* $p_\gamma(\theta)$.

Fundamental concept in sparsifying the EC matrix:

$p_\gamma(\theta)$ for EC matrix A is of the form:

$$p_\gamma(\theta) \propto p_\gamma(A)\ p(\sigma)\ \prod_{i=1}^{n} p(h_i)p(\lambda_i) \tag{2.21}$$

where roles of the terms are:

1. $p(\sigma)$ and $p(\lambda_i)\ i = 1, ..., n$ are uniformative priors.
2. $h_i \sim \mathcal{N}(\mu_h, \Sigma_h)$ is i.i.d gaussian.
3. $p_\gamma(A)$ plays important role in sparsifying the EC and is called *Sparsity inducing prior*.

From the *Sparse Bayesian Learning(SBL)* framework, the elements of EC matrix A are gaussian distributed with zero mean and variance given by γ_i i.e.

$$p_\gamma(\alpha) \sim \mathcal{N}(0, diag(\gamma_1, ..., \gamma_{n^2})) \tag{2.22}$$

where $\alpha = vec(A^T)$ which is the *vectorization of transpose of EC matrix A* whose dimension is $n.n \times 1$.
When γ_i's where $i = 1, ..., n^2$ are estimated by *Marginal Likelihood Maximization*, certain estimates are concentrated around zero and thus the Gaussian posterior distribution of corresponding element α_i of matrix A will also concentrate around zero leading to zero MAP estimate and so in the end we obtain sparse EC matrix[29].

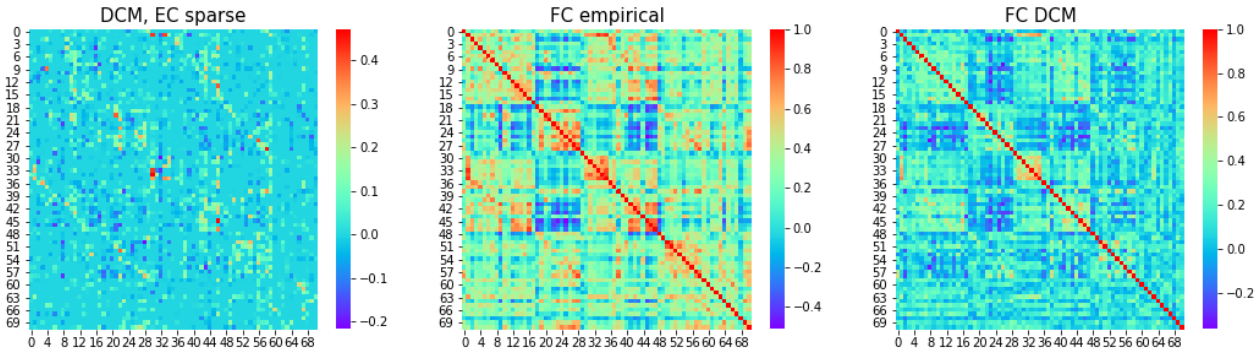


Figure 2.4: EC matrix obtained from sparse DCM(left), Empirical FC matrix from BOLD time series of *sub-010002* from LEMON dataset, Simulated FC matrix from DCM(right). The Pearson Correlation Coefficient of FC simulated via DCM and empirical FC is 0.789

2.2.5. MOU model to extract EC

In this section we discuss the framework formulated by *Gilson et al.*[31] to obtain EC matrix which encapsulates both the synaptic strength(referring to structural connectivity) as well as effective coupling of neural activity of the brain regions. The approach is based on a *Noise Diffusion, (ND)* model which starts from the empirical

FC and computes both an EC matrix and a noise matrix (intrinsic variability experienced by individual brain regions), of which here we focus only on the former which is relevant to our discussion. Also known as *Multivariate Ornstein Uhlenbeck, (MOU)* model where the spatio-temporal covariance pattern is generated by the intrinsic noise which causes fluctuations in neural activity and also shaped by the recurrent connections. This model unlike the DCM, does not explicitly incorporate HRF but instead yields EC matrix which outlines the interactions between brain regions and indirectly relates to the BOLD fMRI signals. Before we discuss the theory introduced by Gilson to extract EC from BOLD time series data, we define the terms and notations: Considering a network of neural populations:

- C refers to EC matrix whose dimension is $n \times n$
- x_i^t refers to neural activity of node/brain region i at time t where $i = 1, ..., n$
- Mean neural activity of node i is $\overline{x}_i = \langle x_i^t \rangle$
- Model time shifted covariances of x_i^t is $Q_{ij}^\tau = \langle (x_i^t - \overline{x}_i)\ (x_j^{t+\tau} - \overline{x}_j) \rangle$ where the average is over the time which is equal to average of random values of neural activity of a node caused by intrinsic noise. If $\tau = 0$ we obtain model zero time shifted covariance Q^0
- Empirical FC with time shift τ and for n_{sim} simulated sessions(if the time series is simulated, else $n_{sim} = 1$) is defined as:

$$\hat{Q}_{ij}^\tau = \frac{1}{n_{sim}\ T} \sum_{0 \leq t < T} (x_i^{n,t} - \hat{x}_i^n)\ (x_j^{n,t+\tau} - \hat{x}_j^n) \tag{2.23}$$

  where $1 \leq n \leq n_{sim}$. Empirical/objective zero time shift covariance is given by $\hat{Q}^0$.

The activity of brain region, x_i is formulated by ODE which is stochastic:

$$dx_i^t = \left[-\frac{x_i^t}{\tau_x} + \sum_{k \neq i} C_{ik} x_i^t \right] dt + dB_i^t \tag{2.24}$$

where the notations are:

- $\tau_x = 1$ typically which is responsible for exponential decay in the x_i^t.
- $C_{ii} = 0$, self-loops are forbidden.
- dB_i^t is white noise with variance σ_i^2 and B_i^t is *Wiener process*.

When assumed stationarity $\frac{dx_i^t}{dt} = 0$ is substituted in Eq. 2.24 and this system has single stable fixed point given by the roots of the equation with J being invertible is necessary condition:

$$J\overline{x} + e = 0 \tag{2.25}$$

where each of the terms in the above equation are matrices and J refers to Jacobian which is of the form:

$$J_{ij} = \frac{\delta_{ij}}{\tau_x} + C_{ij} \tag{2.26}$$

and

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases} \tag{2.27}$$

is Kronecker delta.

From the well established *Ito's formulation*, Lyapunov equation is derived:

$$JQ^0 + Q^0 J^\dagger + \Sigma = 0 \tag{2.28}$$

where the notations are:

- $J^\dagger$ is transpose of J.
- Σ is intrinsic noise matrix with the diagonal elements given by $\sigma_i^2 = \delta_{ij}\langle dB_i^t dB_j^t \rangle$

The model time shifted covariance is given by:

$$Q^\tau = Q^0 \mathbf{expm}[J^\dagger \tau] \tag{2.29}$$

where **expm** is matrix exponential. Thus, J can be calculated easily from Eq. 2.29.

$$J = \frac{1}{\tau} \left\{ \mathbf{logm} \left[(Q^0)^{-1} \, Q^\tau \right] \right\} \dagger \tag{2.30}$$

The EC matrix C is obtained directly from the non-diagonal elements of J.

Having defined all the relevant terms, EC is estimated from the empirical FC *via Lyapunov Optimization*.

1. The cost function to minimize is given by *Lyapunov function*:

$$V(C) = \sum_{m,n} \left[(Q^0_{m,n} - \hat{Q}^0_{m,n})^2 \; + \; (Q^\tau_{m,n} - \hat{Q}^\tau_{m,n})^2 \right] \tag{2.31}$$

   where each term is referred as *matrix distance* and V is positive definite.

2. C is initialized to matrix of dimension $(n \times n)$ whose elements are zeros.
3. Let us assume we know intrinsic noise matrix Σ, the model covariances Q^0 and Q^τ is computed from Eq. 2.28 and Eq. 2.29 respectively.

4. With *optimization rate* ϵ_c being small positive number, the covariance matrix distances is given by:

$$\Delta Q^0_{m,n} = \epsilon_c \left(\hat{Q}^0_{m,n} - Q^0_{m,n}\right) \tag{2.32}$$

$$\Delta Q^\tau_{m,n} = \epsilon_c \left(\hat{Q}^\tau_{m,n} - Q^\tau_{m,n}\right) \tag{2.33}$$

5. The update for the Jacobian can be equivalently written in the form of Eq. 2.30 which depends on the change in model covariances given in Eq. 2.32.

$$\Delta J = \frac{1}{\tau}\left[(Q^0) - 1\,(\Delta Q^0 + \Delta Q^\tau \mathbf{expm}(-J^\dagger \tau))\right]^\dagger \tag{2.34}$$

6. Thus finally, the corresponding update for EC matrix is simply given by:

$$\Delta C_{ij} = \Delta J_{ij} \tag{2.35}$$

with the diagonal elements of C always fixed to zero.

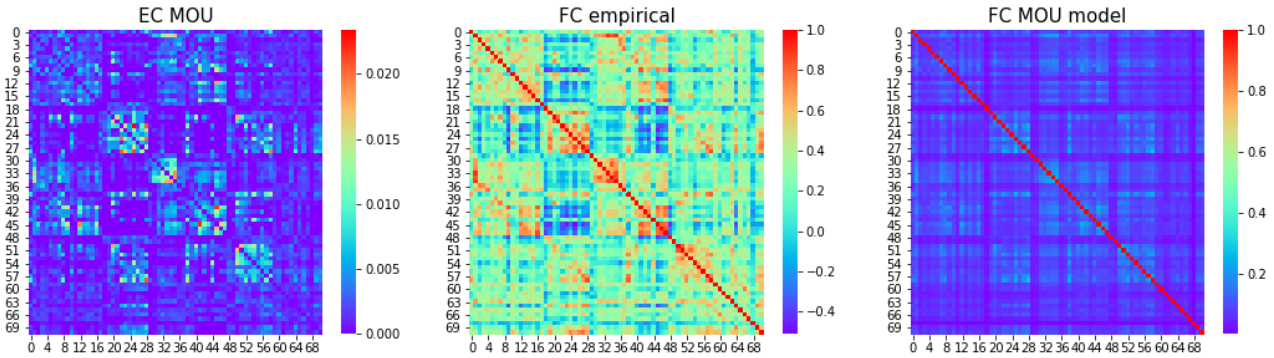


Figure 2.5: EC(left), objective zero time shift FC matrix(middle) and model zero time shift FC matrix(right) obtained using MOU model on *sub-010002 of LEMON dataset*. Pearson correlation of *observed covariance* and *model covariance* is: 0.64588. Note that only in this case FC refers to covariance.

3

Brain controllability with Effective connectivity

In the previous chapter, we have extracted the EC matrices from 3 different frameworks i.e. Linear model, Sparse DCM and MOU model. In this chapter, we are going to see how the concepts of Control theory are applied to study controllability of brain systems.

In the first section, we intend to and to test the algorithms and provide evidence and proof of concepts of well established notions of control theory like Kalman controllability, Structural controllability via cost function minimization and finally the Complete state controllability. To this aim, we consider only first 6 regions of the brain for the sub-010002 of LEMON dataset and results from the DCM model. In the second section, we discuss these notions for the whole brain dynamics (72 regions). We consider only EC matrices from Sparse DCM and MOU model for our study since there exists some inherent problems with Linear model, as discussed in the previous chapter (Fig. 2.3). We will focus on the effects of topology, sparsity and the choice of linear model (DCM vs MOU) on controllability.

3.1. Proof-of-concepts of controllability

3.1.1. Kalman controllability

To demonstrate the methods/code developed, we consider only the first 6 regions of the (72×72) EC matrix from sparse DCM for sub-010002 of LEMON dataset. The diagonal entries are taken from the EC matrix of DCM before it was sparsified.

Case $\mathbf{A} = \mathbf{A_{DCM}}[:\mathbf{6}, :\mathbf{6}]$

$dim(A) = (6, 6)$

$$A = \begin{bmatrix} -0.61502668 & 0 & 0 & 0 & 0 & 0.04969445 \\ 0 & -0.79069292 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0.69419440 & 0 & -0.04564398 & -0.03556111 \\ 0 & 0 & 0.04076849 & -0.53235401 & -0.00303871 & -0.01251824 \\ 0 & 0 & 0.16217933 & 0 & -0.67205625 & 0 \\ 0.02078494 & 0 & 0 & -0.02069000 & 0 & -0.65260954 \end{bmatrix}$$

The above matrix A has nonzero diagonal entries. This is actually necessary in order to have a stable dynamics, as we know that for continuous LTI given by Eq. 1.2 to be stable, the $Real(\ \lambda_{max}(A)\ ) < 0$. However, if nonzero diagonal entries are kept in the identification of driver nodes via maximum matching theorem, we always obtain a control solution which is, at the same time, trivial and unfeasible. This is the control solution discussed in Fig. 1.7,1.8, where all nodes are jointly controlled by a single external input. This trivial solution is not interesting for brain controllability, as there is no way to apply a input jointly over several distant regions: techniques such as transcranial magnetic stimulation only allows to deliver a local input to a single region at the time. Thus, in order to rule out this trivial solution, we address maximum matching by removing the diagonal matrix entries, which leads to a control solution targeting single nodes. Once the driving nodes are found, we reinstate the diagonal entries to compute the energies in Kalman and Structural controllability. By using the above matrix A without the diagonal entries, the unmatched nodes found via max matching are $[2, 5]$ where the 6 nodes are in the order: 1, 2, 3, 4, 5, 6. By *Minimum Input theorem*, number of control nodes/driver nodes = number of unmatched nodes. Thus, the input matrix B in this case is represented as:

$$B = B_{kalman} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

All single nodes violate Kalman's controllability rank condition so *Single node controllability* is not possible for this system given by (A, B) with B being i-th canonical vector of dimension N i.e. $B = e_i$ where i represents the node considered.
Minimum energy required to steer the system using the driver nodes $[2, 5]$ is computed: $E_{min} = 1.73365 \times 10^{12}$.
But when all the nodes are considered as driver nodes, $E_{min} = 1.5813$.

Caveat: A point to be mentioned is that since the sparse DCM is continuous, we obtain E_{min} by solving for controllability Gramian W using continuous Lyapunov Eq. 1.11.

Case $\mathbf{A} = \mathbf{A_{DCM}}[:\mathbf{6}, :\mathbf{6}] \times 5$

While the identification of driver nodes is sensitive only to the structure of connections (which links are nonzero) and not to their magnitude, Kalman controllability (and, in particular, control energy), may strongly depend on the magnitude. Therefore, to observe how E_{min} changes with respect to scaling of the values of the EC matrix elements, all the elements are multiplied by 5(scalar product), i.e. $A * 5$.

$$A' = A * 5$$

$$= \begin{bmatrix} -0.61502668 & 0 & 0 & 0 & 0 & 0.24847226 \\ 0 & -0.79069292 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0.69419440 & 0 & -0.22821991 & -0.17780553 \\ 0 & 0 & 0.20384243 & -0.53235401 & -0.01519355 & -0.06259118 \\ 0 & 0 & 0.81089666 & 0 & -0.67205625 & 0 \\ 0.10392470 & 0 & 0 & -0.10345002 & 0 & -0.65260954 \end{bmatrix}$$

The driver nodes are the same as in the previous case of A without diagonal entries, thus input matrix B_{kalman} remains as it is.
In this case $Real(\ \lambda_{max}(A')) = -0.4272$ with A' has the same diagonal entries from dense matrix of DCM while non diagonal elements are multiplied by a factor of 5.

The minimum energy required to steer the system using the driver nodes $[2, 5]$ decreases by 5 orders of magnitude i.e. $E_{min} = 1.24973 \times 10^7$ relative to E_{min} when computed without multiplying by a factor of 5.
Warning: This result would suggest that it is convenient to rescale the EC weights (by multiplying by a factor of 5), since the min. energy decreases by orders of magnitude. However, while this can be easily done on paper, it does not have any practical application since we cannot modify the weights of the dynamical interactions experimentally: the weights are fixed and given by EC matrix derived from BOLD fMRI time series.

3.1.2. Structural controllability via minimization of cost function

With the unmatched nodes being $[2, 5]$ for $A = A_{DCM}[:6, :6]$, the input matrix is given by:

$$B_{struct} = \begin{bmatrix} 0 & 0 \\ b_1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & b_2 \\ 0 & 0 \end{bmatrix}$$

where b_1 and b_2 are the parameters to be found by minimization of cost function.

Case $\mathbf{A = A_{DCM}[:6, :6]}$

Initial guess: $[\ b_1,\ b_2] = [\ 0.54340494,\ 0.27836939\ ]$ were randomly generated from uniform distribution over $[0, 1)$.

$$B_{struct} = B_{fit} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0.543 & 0 \\ 0 & 0 \\ 0 & 1.531 \\ 0 & 0 \end{bmatrix}$$

$E_{min} = 7.39225 \times 10^{11}$ was found. Clearly, when the results are compared with Kalman controllability as presented before, $E_{min}^{struct} < E_{min}^{kalman}$ by an order of magnitude because of the optimization of input matrix B.

Case $\mathbf{A = A_{DCM}[:6, :6] \times 5}$

Initial guess for b_1 and b_2 were the same as in the previous **Case** $\mathbf{A = A_{DCM}[:6, :6]}$.

$$B_{struct} = B_{fit} = \begin{bmatrix} 0 & 0 \\ 1.8860 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 5.30205e+03 \\ 0 & 0 \end{bmatrix}$$

$E_{min} = 0.44456$. This shows that scaling of the elements has greater effect in terms of structural controllability because we see several orders of magnitude difference when compared with EC matrices which are unscaled as in the previous case, $E_{min}(A \times 5, B_{struct}) << E_{min}(A, B_{struct})$.

EC matrix	$Real(\ \lambda_{max}(A)\ )$ $\mathbf{diag}(A) \neq 0$	Kalman E_{min} (unmatched nodes)	Kalman E_{min} (all nodes)	Structural E_{min}
$A[:6,:6]$	-0.5289	1.73365×10^{12}	1.5813	7.39225×10^{11}
$A[:6,:6] \times 5$	-0.4272	1.24973×10^{7}	1.8768	0.44456

3.1.3. Test for Complete state controllability

Case $\mathbf{A} = \mathbf{A_{DCM}}[:\mathbf{6},:\mathbf{6}]$

Given A, B_{struct} as in Case $A = A_{DCM}[:6,:6]$ (ref. 3.1.2) and initializing initial and final state as: $\mathbf{x}_{initial} = [\ 1,\ 1,\ 1,\ 1,\ 1,\ 1\ ]$ and $\mathbf{x}_{final} = [\ 2,\ 2,\ 2,\ 2,\ 2,\ 2\ ]$, the input signal $\mathbf{u}$ is found as mentioned in earlier chapter.

$$\mathbf{u} = \begin{bmatrix} 1.2878 & 6.4182e+04 \\ -1.0183 & 3.37287e+05 \\ 0.8051 & 5.46988e+05 \\ -0.6366 & 8.1585e+04 \\ 0.5034 & -5.21282e+05 \\ -0.3980 & -3.39333e+05 \end{bmatrix}$$

where column 1 and column 2 corresponds to driver nodes 2 and 5 respectively while the rows values correspond to time axis with t = 0 to 5 and Total number of time steps = T = *N(dimension of EC matrix)*. As seen from column 2 of $\mathbf{u}$, the final state is very sensitive to input signal $\mathbf{u}_5$ which is the input signal given to driver node 5.

After computing input signals required to drive the system on the driver nodes, it is possible to cross-check. We know A, B_{struct}, $\mathbf{x_{initial}} = \mathbf{x}(t=0)$ and $\mathbf{u}$ and with this we compute the final state and the result is:

$$\begin{aligned} \mathbf{x}_n &= \mathbf{x}(t = T-1) \\ &= [\ 1.97618996,\ 1.75563045,\ 1.84599689,\ 2.00013076,\ 2.08584186,\ 1.92522838\ ] \end{aligned}$$

The *Mean squared error(MSE)* of the result obtained from cross check above x_n and x_{final} which was initialized to find $\mathbf{u}$ is 0.01615.

Case $\mathbf{A} = \mathbf{A_{DCM}}[:\mathbf{6},:\mathbf{6}] \times \mathbf{5}$

Given A, B_{struct} as in Case $A = A_{DCM}[:6,:6] \times 5$ (ref. 3.1.2) and initializing initial and final state as: $\mathbf{x}_{initial} = [\ 1,\ 1,\ 1,\ 1,\ 1,\ 1\ ]$ and $\mathbf{x}_{final} = [\ 2,\ 2,\ 2,\ 2,\ 2,\ 2\ ]$, the

input signal $\mathbf{u}$ is found as mentioned in earlier chapter.

$$\mathbf{u} = \begin{bmatrix} 0.3710 & 0.0605 \\ -0.2934 & 0.2700 \\ 0.2320 & 0.3541 \\ -0.1834 & 0.0033 \\ 0.1450 & -0.3360 \\ -0.1147 & -0.2003 \end{bmatrix}$$

$\mathbf{x}_n = \mathbf{x}(t = T-1) = [\ 2.0946918,\ 1.75563045,\ 2.34001895,\ 2.01854897,\ 2.41368795,\ 1.93007287\ ]$

Here in $\mathbf{u}$ may not see sensitivity similar to column 2 of $\mathbf{u}$ of the previous case, but instead we see that for B_{fit} of Case $\mathbf{A} = \mathbf{A_{DCM}}[:\mathbf{6}, :\mathbf{6}] \times \mathbf{5}$ the value the parameter takes for after optimization is quite large and so the $\mathbf{x}_{final}$ is sensitive to B_{fit}. The *MSE* of the result obtained from cross check above x_n and x_{final} which was initialized to find $\mathbf{u}$ is 0.06011.

Brief note on Directed Configuration model:

Before proceeding to the next subsection, here we discuss briefly about random model namely Directed Configuration model which will be extensively used. The configuration model generates a random directed pseudo graph (graph with parallel edges and self loops) by randomly assigning edges to match the given degree sequences. [32, 33, 34]

Algorithm:

1. Check: Sum in-degree sequence is equal to Sum of out-degree sequence
2. Create an empty graph of n nodes i.e. $n = |in-degree\ sequence|\ or\ |out-degree\ sequence|$.
3. Build a list of available degree-repeated nodes. For example, for degree sequence [3, 2, 1, 1, 1], the "stub list" is initially [0, 0, 0, 1, 1, 2, 3, 4], that is, node 0 has degree 3 and thus is repeated 3 times, etc.
4. Also, shuffle the stub list in order to get a random sequence of node pairs.
5. Add the edges from out stublist to in stublist with one to one mapping of each elements of the 2 lists.

A non-graphical degree sequence (not realizable by some simple graph) is allowed since this function returns graphs with self loops and parallel edges. This configuration model construction process can lead to duplicate edges and loops. Removing the self-loops and parallel edges(manually) will likely result in a graph that

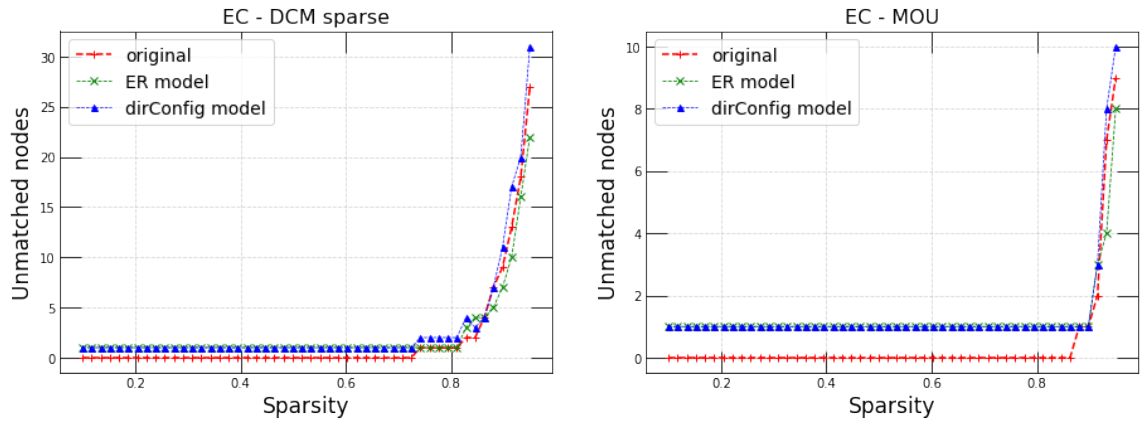
doesn't have the exact degree sequence specified. This *finite-size effect* decreases as the size of the graph increases. The density of self-loops and parallel edges tends to decrease as the number of nodes increases. However, typically the number of self-loops will approach a Poisson distribution with a nonzero mean, and similarly for the number of parallel edges. Consider a node with k stubs(which is degree of node). The probability of being joined to another stub of the same node is basically $(k-1)/N$, where k is the degree and N is the number of nodes. So the probability of a self-loop scales like c/N for some constant c. As N grows, this means we expect c self-loops. Similarly for parallel edges.

3.1.4. Factors influencing controllability: the key role of sparsity

Unlike the previous parts, here we consider the whole brain network i.e. all the 72 regions of EC matrix. We explicitly investigate the possible influence of three factors in determining controllability properties: sparsity, topology, choice of EC. We do so by comparing networks at different degrees of sparsity, with fixed or randomized topology, and with EC obtained by two methods (DCM and MOU).

Given an EC matrix, sparsification is done by killing the entries of the matrix(i.e. setting the entries of the matrix to zero). A method to kill the entries is done we select a fractional value f (which is the fraction of removed links), with $0.1 \leq f \leq 0.95$, and we set to zero the fraction f of links with greater weights in absolute values(i.e., for $f = 0.95$, only the 5% weakest links are kept). Thus, we obtain EC matrix with a fixed sparsity. At each fraction f =0.1 to 0.95, we consider both the sparsified EC given by $A_{sparse}(frac = f)$ and two randomized versions of it. The latter are obtained from $A_{sparse}(frac = f)$, $f = 0.1$ *to* 0.95 by computing the corresponding adjacency matrices from two random models: the *Erdos-Renyi(ER) Model* and the *Directed Configuration(DC) Model*. For ER model, we make use of number of nodes/brain regions N and probability to make a connection between the nodes given by $p = density(A_{sparse}(frac = f))$ computed as density of EC matrix for a given fraction removed f. While, for DC model, we make use of $Out-degree(A_{sparse}(frac = f))$ and $In-degree(A_{sparse}(frac = f))$ for a given fraction f. Both these random models preserves some features of the original matrix: ER model preserves the density and number of nodes while DC model preserves the in and out degree for each of the N nodes. The adjacency matrices obtained from such random models i.e. A_{erdos} and $A_{dirConfig}$ is unweighted, and so we introduce weights by randomly selecting the non-zero entries from the $A_{sparse}(frac = f)$ and placing them in the positions of non-zero entries(here in this case it is 1) of A_{erdos} and $A_{dirConfig}$. In the end, we have weighted adjacency matrices from the random models preserving the weight distribution of the original $A_{sparse}(frac = f)$).

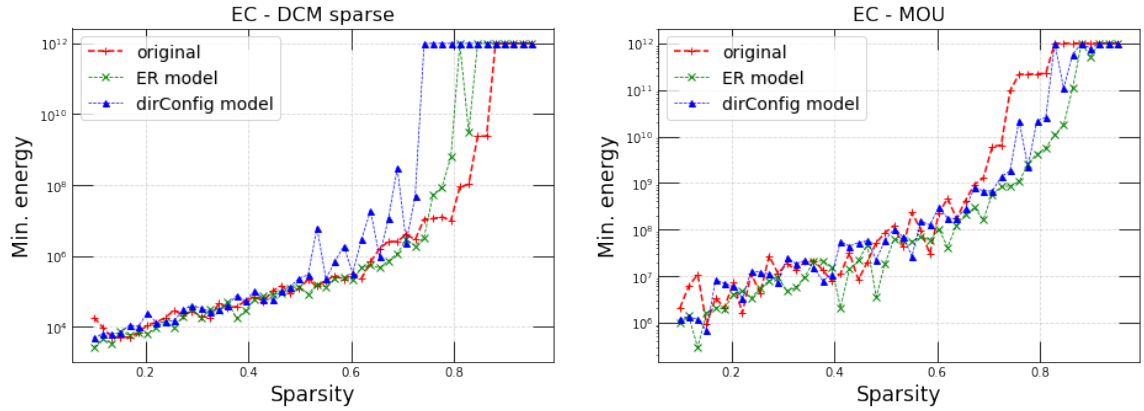
As seen from the Fig. 3.1, we observe that irrespective of the models used for



(a) **(b)**

Figure 3.1: Number of unmatched nodes vs Sparsity/fraction of nodes removed(or set to zero in adjacency matrices) for EC matrix derived from sparse DCM(*left*) and from MOU model(*right*).

extracting EC matrices(either DCM or MOU model) and irrespective of whether the EC matrix is original or random, the number of unmatched nodes computed by applying max. matching concept depends on the extent of sparsity of the adjacency matrix. With the increase in sparsity, we have decrease in density of connections in such networks and we observe that the number of unmatched nodes remains zero(in original but equal to one for random) till a given value of sparsity and then starts to increase exponentially.



(a) **(b)**

Figure 3.2: Min. Energy vs Sparsity/fraction of nodes removed(or set to zero in adjacency matrices) for EC matrix derived from sparse DCM(*left*) and from MOU model(*right*). The input matrix B computed by considering 50% of total 71 nodes(representing 71 brain regions) at random so as compute E_{min} of the system given (A_{sparse}, B) at sparsity or fraction removed ranging from 0.1 to 0.95.

After obtaining $A_{sparse}(frac = f)$ for a given fraction f, the driver nodes were identified by applying max. matching and input matrix B was constructed. It

was observed that minimum eigenvalue of controllability Gramian $W(A_{sparse}, B)$ is close to zero. Thus the minimum energy corresponding to a given f(sparsity), for this specific choice of input matrix B, cannot be computed since the energy is extremely high and it leads to numerical issues. For example, for EC matrix of sparse DCM, when the fraction of nodes removed is 0.95, we obtain the highest number of unmatched nodes(which becomes driver nodes in this case) i.e. $N_D = 27$. The $A_{sparse}(frac = 0.95)$ has real part of maximum eigenvalue -0.234 which means continuous LTI equation is stable. With input matrix B constructed from those 27 driver nodes, we obtain minimum eigenvalue of controllabililty Gramian as 9.424×10^{-22} which means $E_{min} \sim 10^{22}$. This mainly arises due to the fact that the number of unmatched nodes computed for a given sparsified A is a relatively low fraction of the total number of nodes. Here, we consider values of $E_{min} > 10^{12}$ as being numerically indistinguishable from ∞, so we approximate all $E_{min} > 10^{12}$ with $E_{min} = 10^{12}$. If we follow this procedure, for all values of f we obtain $E_{min} = 10^{12}$. Thus, controlling the system with unmatched nodes given by the structural controllability is not enough for effective controllability, i.e., controllability with an energy of the order $E_{min} < 10^{12}$.

Thus, we try to explore how the energy scales with f by controlling with a large fraction of nodes. We consider a random 50% of the total number of nodes as driver nodes and then construct input matrix B accordingly. By keeping this $B = B_{random\_50p}$ fixed (where subscript $random\_50p$ stands for randomly selected 50 percent of total nodes), we compute E_{min} for EC matrices extracted from different models (DCM and MOU model) and for random models (ER and DC) at various sparsity as seen in Fig. 3.2. From this figure, we observe that data points for random model coincides with the original: this means that controllability is not dependent on the exact positions of the weighted links of the network, but only on the sparsity of the network and the number of driver nodes, which should amount to a significant fraction of the total nodes. Thus, only if the original EC matrix is very sparse can we hope to achieve pratical controllabilty with driver nodes given by structural controllability. In future work, we will investigate which is the right degree of sparsity that may lead to this situation.

In summary, qualitatively controllability of a system is unaffected by:

1. **Topology of the network**. We observe from plots of E_{min} vs sparsity(Figs. 3.2a, 3.2b) and Unmatched nodes vs sparsity (Figs. 3.1a, 3.1b) that even though the position of entries of A gets rearranged when random models are used, E_{min} seem to approximately coincide with the original curve and number of unmatched nodes follows a similar trend as that of original curve.

2. **Choice of model to extract EC**. By keeping fixed 50 percent of driver nodes taken at random, we again observe from the same plots (Figs. 3.2a, 3.2b) that

initially, when fraction of removed entries (of EC) is low, the minimum energy is less and it increases exponentially with increase in sparsity(or decrease in density) of the EC matrix irrespective of the model we have used to extract EC.

In theory, by applying max. matching, we can identify driver nodes and use those nodes to attain controllability. In practice, However, controllability is not possible when using driver nodes given by structural controllability: one actually needs a larger number of driver nodes to have controllability with a finite energy.

4
Conclusion

We here summarize the main conclusions of this thesis:

- For applying the Controllability on brain networks, using Effective Connectivity matrices that capture the dynamical complexity of the neural systems is desirable instead of Structural connectivity matrices which captures only the topology of the brain network.

- BOLD time series data are affected by 4 properties: averaging over space, averaging over time, observation noise and limited data samples. These properties counteract or mask the non linearities of the neuronal dynamics at the micro level and hence modelling the brain dynamics with most simple Linear model is suitable.

- Among the linear models tested, Sparse DCM proposed by *Prando et al.*[29] is better than MOU model proposed by *Gilson et al.*[31] which neglects the effect of haemodynamic response. An evidence of it is seen when we compare the Pearson correlation coefficient of the observed FC and model FC for the two respective models as: 0.646 and 0.789. The linear model proposed by *Nozari et al.*[20] was found be unsuitable, as linear regression can not correctly reproduce the FC.

- The number of unmatched nodes computed depends on the level of sparsity of any EC matrix. The greater the sparsity of a given EC matrix, the higher the number of unmatched nodes(obtained by applying max. matching). The number of unmatched nodes depends only weakly on the topology of effective connections (which can be seen by randomizing matrices using random models ER and DC) and the method used to infer EC (sparse DCM or MOU).

- It is shown qualitatively that controllability of a system is unaffected by the topology of the network and depends on the sparsity and number of driver nodes. It is further shown that irrespective of the model used(sparse DCM or MOU), controllability metric E_{min} increases exponentially with sparsity (or

decreases with density) when a random subset of nodes (corresponding to a fixed fraction 50%) are used as driver nodes.

Open problems:

- Linear model proposed by *Nozari et al.*[20] has to be revisited starting from its basic formulation and framework. The main problem emerges from the R^2 boxplots(Fig. 2.3) which apparently tells that Linear dense model is worse than the Zero model, which does not use any information contained in past values of the time series to prdict the current value (it only uses a constant predictor). A thorough investigation has to be made to ascertain if the model and code are correct. For the time being, it is advisable to restrict attention to other models such as MOU model and sparse DCM to obtain EC matrices.

- Further investigations have to be done to understand how to deal with diagonal entries in the EC matrix A. The presence of negative entries in the principal diagonal is required to keep the continuous LTI equation stable ($Real(\lambda_{max}(A)) < 0$). However, the presence of diagonal entries in A leads to perfect matching: we cannot identify unmatched nodes > 0 for such case, and a trivial control solution is possible (with a single input delivered jointly to all nodes). To discard this trivial control mode, as a temporary solution we have resorted to not considering the principal diagonal entries for identifying unmatched nodes.

Future perspectives:

- Having demonstrated the proof of concept for *Complete State controllability* where a *small* system can be steered from any desirable initial state to any desirable final state, it should be seen whether complete state controllability is also possible for large brain networks, and whether the input signal $\mathbf{u}$ is feasible in practice to attain full controllability.

- While it is possible in theory to control the system by using the driver nodes given by structural controllability (max. matching), this solution requires extremely high minimum energy to control the system, and so it is not possible in practice to have controllability. We should investigate which is the right degree of sparsity(or density) of EC matrix which ensures that we obtain a large enough number of driver nodes in such a way that minimum energy required to control the system remains finite, and reasonably low (e.g., $< 10^{12}$).

References

[1] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, "Controllability of complex networks," vol. 473, no. 7346, pp. 167–173, May 2011. [Online]. Available: https://doi.org/10.1038/nature10011

[2] K. Ogata, *Modern control engineering*. Upper Saddle River, N.J: Prentice Hall, 1997.

[3] E. D. Sontag, "Kalman's controllability rank condition: From linear to nonlinear." Springer Berlin Heidelberg, 1991, pp. 453–462. [Online]. Available: https://doi.org/10.1007/978-3-662-08546-2_25

[4] C.-T. Lin, "Structural controllability," vol. 19, no. 3, pp. 201–208, Jun. 1974. [Online]. Available: https://doi.org/10.1109/tac.1974.1100557

[5] R. Shields and J. Pearson, "Structural controllability of multiinput linear systems," *IEEE Transactions on Automatic Control*, vol. 21, no. 2, pp. 203–212, 1976.

[6] A. Olshevsky, "Minimal controllability problems," vol. 1, no. 3, pp. 249–258, Sep. 2014. [Online]. Available: https://doi.org/10.1109/tcns.2014.2337974

[7] J. E. Hopcroft and R. M. Karp, "An $n^{5/2} $ algorithm for maximum matchings in bipartite graphs," vol. 2, no. 4, pp. 225–231, Dec. 1973. [Online]. Available: https://doi.org/10.1137/0202019

[8] G. Yan, G. Tsekenis, B. Barzel, J.-J. Slotine, Y.-Y. Liu, and A.-L. Barabási, "Spectrum of controlling and observing complex networks," vol. 11, no. 9, pp. 779–786, Aug. 2015. [Online]. Available: https://doi.org/10.1038/nphys3422

[9] W. J. Rugh, *Linear System Theory (2nd Ed.)*. USA: Prentice-Hall, Inc., 1996.

[10] G. Ramos, A. P. Aguiar, and S. Pequito, "Structural systems theory: an overview of the last 15 years," 2020.

[11] M. P. van den Heuvel and H. E. H. Pol, "Exploring the brain network: A review on resting-state fMRI functional connectivity," vol. 20, no. 8, pp. 519–534, Aug. 2010. [Online]. Available: https://doi.org/10.1016/j.euroneuro.2010.03.008

[12] K. J. Friston, "Functional and effective connectivity: A review," vol. 1, no. 1, pp. 13–36, Jan. 2011. [Online]. Available: https://doi.org/10.1089/brain.2011.0008

[13] A. Aertsen and H. Preissl, "Dynamics of activity and connectivity in physiological neuronal networks," *Nonlinear Dynamics and Neuronal Networks*, 01 1991.

[14] J. D. Cohen, "Cognitive control." John Wiley & Sons, Ltd, Feb. 2017, pp. 1–28. [Online]. Available: https://doi.org/10.1002/9781118920497.ch1

[15] S. Gu, F. Pasqualetti, M. Cieslak, Q. K. Telesford, A. B. Yu, A. E. Kahn, J. D. Medaglia, J. M. Vettel, M. B. Miller, S. T. Grafton, and D. S. Bassett, "Controllability of structural brain networks," vol. 6, no. 1, Oct. 2015. [Online]. Available: https://doi.org/10.1038/ncomms9414

[16] C. Tu, R. P. Rocha, M. Corbetta, S. Zampieri, M. Zorzi, and S. Suweis, "Warnings and caveats in brain controllability," vol. 176, pp. 83–91, Aug. 2018. [Online]. Available: https://doi.org/10.1016/j.neuroimage.2018.04.010

[17] M. Molloy and B. Reed, "A critical point for random graphs with a given degree sequence," vol. 6, no. 2-3, pp. 161–180, Mar. 1995. [Online]. Available: https://doi.org/10.1002/rsa.3240060204

[18] M. Newman, *Networks*. Oxford University Press, Mar. 2010. [Online]. Available: https://doi.org/10.1093/acprof:oso/9780199206650.001.0001

[19] G. Yan, P. E. Vértes, E. K. Towlson, Y. L. Chew, D. S. Walker, W. R. Schafer, and A.-L. Barabási, "Network control principles predict neuron function in the caenorhabditis elegans connectome," vol. 550, no. 7677, pp. 519–523, Oct. 2017. [Online]. Available: https://doi.org/10.1038/nature24056

[20] E. Nozari, M. A. Bertolero, J. Stiso, L. Caciagli, E. J. Cornblath, X. He, A. S. Mahadevan, G. J. Pappas, and D. S. Bassett, "Is the brain macroscopically linear? a system identification of resting state dynamics," Dec. 2020. [Online]. Available: https://doi.org/10.1101/2020.12.21.423856

[21] D. Kumral, F. Şansal, E. Cesnaite, K. Mahjoory, E. Al, M. Gaebler, V. Nikulin, and A. Villringer, "BOLD and EEG signal variability at rest differently relate to aging in the human brain," *NeuroImage*, vol. 207, p. 116373, Feb. 2020. [Online]. Available: https://doi.org/10.1016/j.neuroimage.2019.116373

[22] E. M. Izhikevich, *Dynamical systems in neuroscience*. MIT press, 2007.

[23] V. Booth and J. Rinzel, "A minimal, compartmental model for a dendritic origin of bistability of motoneuron firing patterns," vol. 2, no. 4, pp. 299–312, Dec. 1995. [Online]. Available: https://doi.org/10.1007/bf00961442

[24] W. J. Freeman, "Nonlinear gain mediating cortical stimulus-response relations," vol. 33, no. 4, pp. 237–247, 1979. [Online]. Available: https://doi.org/10.1007/bf00337412

[25] H. R. Wilson and J. D. Cowan, "Excitatory and inhibitory interactions in localized populations of model neurons," vol. 12, no. 1, pp. 1–24, Jan. 1972. [Online]. Available: https://doi.org/10.1016/s0006-3495(72)86068-5

[26] D. N. Greve, , G. G. Brown, B. A. Mueller, G. Glover, and T. T. Liu, "A survey of the sources of noise in fMRI," vol. 78, no. 3, pp. 396–416, Nov. 2012. [Online]. Available: https://doi.org/10.1007/s11336-012-9294-0

[27] Y. Liu, W. G. Coon, A. de Pesters, P. Brunner, and G. Schalk, "The effects of spatial filtering and artifacts on electrocorticographic signals," vol. 12, no. 5, p. 056008, Aug. 2015. [Online]. Available: https://doi.org/10.1088/1741-2560/12/5/056008

[28] K. Friston, L. Harrison, and W. Penny, "Dynamic causal modelling," vol. 19, no. 4, pp. 1273–1302, Aug. 2003. [Online]. Available: https://doi.org/10.1016/s1053-8119(03)00202-7

[29] G. Prando, M. Zorzi, A. Bertoldo, M. Corbetta, M. Zorzi, and A. Chiuso, "Sparse DCM for whole-brain effective connectivity from resting-state fMRI data," vol. 208, p. 116367, Mar. 2020. [Online]. Available: https://doi.org/10.1016/j.neuroimage.2019.116367

[30] K. J. Friston, J. Kahan, B. Biswal, and A. Razi, "A DCM for resting state fMRI," vol. 94, pp. 396–407, Jul. 2014. [Online]. Available: https://doi.org/10.1016/j.neuroimage.2013.12.009

[31] M. Gilson, R. Moreno-Bote, A. Ponce-Alvarez, P. Ritter, and G. Deco, "Estimation of directed effective connectivity from fMRI functional connectivity hints at asymmetries of cortical connectome," vol. 12, no. 3, p. e1004762, Mar. 2016. [Online]. Available: https://doi.org/10.1371/journal.pcbi.1004762

[32] A. Hagberg, P. Swart, and D. S Chult, "Exploring network structure, dynamics, and function using networkx," Los Alamos National Lab.(LANL), Los Alamos, NM (United States), Tech. Rep., 2008.

[33] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, "Random graphs with arbitrary degree distributions and their applications," *Phys. Rev. E*, vol. 64,

p. 026118, Jul 2001. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.64.026118

[34] A.-L. Barabási and M. Pósfai, *Network science*. Cambridge: Cambridge University Press, 2016. [Online]. Available: http://barabasi.com/networksciencebook/

Acknowledgments

I would like to thank my supervisors Prof. Samir Suweis and Prof. Michele Allegra from whom I have learnt extensively on the topics related to my thesis during hours of discussion. With their insightful feedback, constant support and guidance throughout my thesis, it was made possible to bring out the best in me and my skills in research are honed.

I would also like to thank Phd student Giorgia Baron from DEI, University of Padova, for providing me with sparse DCM EC matrices.