UNIVERSITY OF CAMBRIDGE

The Emergence of Network Dynamics in Developing Cortical Circuits

by

Jeremi Chabros

An experimental project report submitted in partial fulfilment of the requirements for the Natural Sciences Tripos

in the

School of Biological Sciences,
Department of Physiology, Development and Neuroscience,
University of Cambridge

April 2021

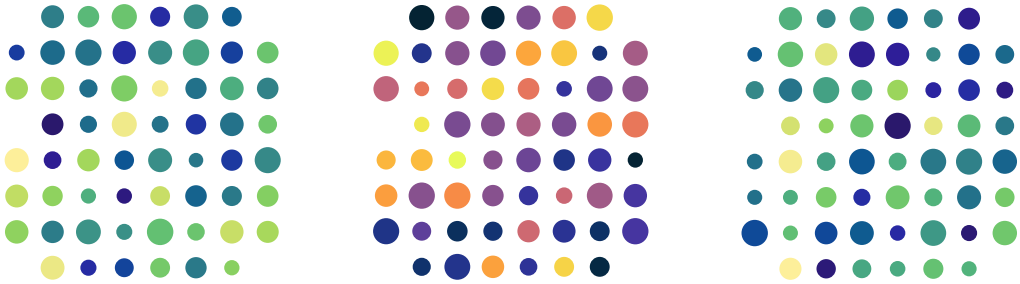
Declaration of Authorship

I, Jeremi Chabros, declare that this thesis, titled "The Emergence of Network Dynamics in Developing Cortical Circuits", and the work presented in it are my own. I confirm that:

- The material in this dissertation is not copied from any published material, unless it is clearly identified as such and a full source reference given.
- The material in this dissertation is not a paraphrase or abstract of any published material, unless it is clearly identified as such and a full source reference given.
- I understand that the examiners may use any means available to test for plagiarism within my dissertation.
- Other than where indicated, this dissertation is my own work.
- The word count is: 8,211 (calculated using TEXCount).

Signed:

Date:



Abstract

Rett syndrome (RTT) is a neurodevelopmental disorder caused by a mutation in the X-linked *MECP2* gene. It is characterised by cortical circuit dysfunction that affects nearly every aspect of a patient's life. Understanding how the disease shapes network structure and dynamics at early developmental stages may help us better understand and, ultimately, modulate the progression of this disease. This study aimed to determine the range and functional coordination of neuronal firing patterns and the extent to which network dynamics can be controlled.

Using microelectrode arrays, I probed the electrical activity of cultured cortical networks over multiple weeks. Using a novel spike detection method I developed to capture spontaneous multiunit activity, I first characterised the network-wide coordination of firing patterns using a measure of dynamical dimensionality called the effective rank. Then, I quantified the extent to which networks realise their propensity for forming functional subcommunities using relative effective rank. To determine the range of circuit dynamics and their controllability, I applied a control theory framework to functional network architecture inferred from parallel spike trains. I quantified the network's dynamical repertoire using average and modal controllability measures. Finally, I determined the extent to which network dynamics could be controlled using the volume of the controllability ellipsoid.

I found that during healthy neurodevelopment, cortical circuits become increasingly structured in a manner optimised to support a wide range of dynamics, and this process is disrupted in Mecp2-deficient networks. Wild type cultures showed low-dimensional, interpretable neural code, and knockout cultures exhibited more desynchronised dynamics. The state-space analysis indicated that healthy circuits were also able to traverse larger surfaces of the energy landscape. Cultures heterozygous for Mecp2 deletion displayed intermediate phenotypes. However, all genotypes were controllable to a similar extent, suggesting that therapeutic modulation of network function in RTT may be feasible.

Acknowledgements

First, I would like to thank my supervisor, Dr Susanna Mierau, for guiding my working knowledge of neuronal networks and the unwavering support both within and outside of my project. Susanna, I owe a particular debt to your generosity in welcoming me to your group. I suppose that this thesis is the start and not merely the culmination of our collaborative efforts to understand network dynamics in Rett syndrome.

I want to thank Alexander Dunn for the tireless work in initiating, maintaining and recording the cell cultures. Alex, thank you for providing me with the invaluable data you have collected over the years. It was my honour to analyse it, and I was fortunate to benefit from your interest and insightful conversations.

My wholehearted gratitude goes to Mr. (soon to be Dr.) Timothy Sit. Tim, thank you for laying the groundwork for this project all these years ago and providing invaluable insights and suggestions all along. It was the greatest inspiration working with you and your expertly written code. The project would not progress without your patience and expertise in troubleshooting my rookie mistakes.

I have had the privilege to work with Dr Bianca Dumitrascu on the computational aspects of this project. Bianca, thank you for your enthusiasm in explaining all the complex mathematical concepts and your understanding in dealing with my algebraic ineptitude.

I want to thank Dr Rachael Feord for her unlimited patience in trying out and troubleshooting my code. Rachael, it was a joy working with you as one of the most talented neuroscientists I know. Without your thoroughness, my project would suffer from many more pitfalls than it has.

Finally, I wish to recognise the invaluable contributions of other members of our group. Rich, thank you for your feedback and understanding in my early days working on this project. Asma, thank you for challenging me with difficult questions that fostered my project. Leo, thank you for allowing me to learn from your work on object detection. Hugo, thank you for developing and explaining the probabilistic thresholding and for the discussions of community detection methods. Lance, thank you for your humour and for showing me your collection of hats in various Zoom meetings. Lucy, thank you for sharing your knowledge of graphics software and consulting colour choices.

Contents

1

Introduction

1.1 Rescuing decline in cognitive function in Rett syndrome

Rett syndrome (RTT) is a severe neurodevelopmental disorder and the most common monogenic cause of intellectual disability in females (Heckman et al., 2014). The hallmark of this condition is an apparent asymptomatic development until 6–18 months old, followed by the regression of motor and language skills resulting in missed developmental milestones (Einspieler & Marschik, 2019). As the disease progresses, it affects nearly every aspect of a patient's life, necessitating lifelong, round-the-clock care. Currently, RTT remains without a cure. Modulating network function at cellular scale is a promising avenue for improving cognitive function in RTT patients (Guy et al., 2007; Castro et al., 2014). Hopefully, with new mechanism-based treatment strategies, we will empower individuals with Rett syndrome to interact with their families and improve their quality of life. Now, the discovery of such treatments requires reliable disease models and developing new methods for studying them.

1.1.1 Disease models of Rett syndrome

The vast majority of RTT cases are caused by a loss-of-function mutation in the X-linked methyl-CpG-binding protein 2 gene (*MECP2*, human; *Mecp2*, mouse) coding for the transcriptional regulator MECP2 (Mecp2 in mouse). Since this discovery (Amir

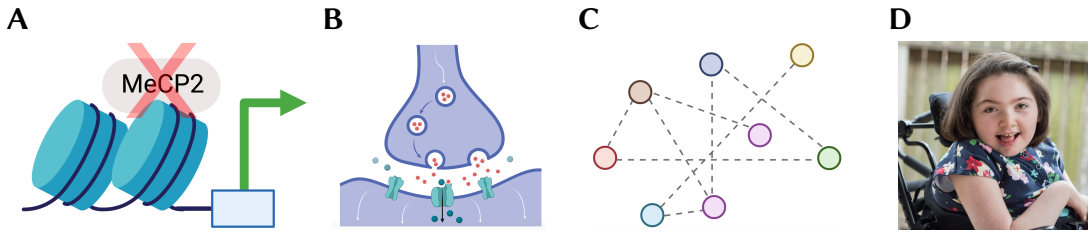


Figure 1.1: Mechanisms of pathogenesis of Rett syndrome. Loss of MeCP2 changes gene expression **(A)**, disrupts synaptic function **(B)**, alters network development **(C)**, and leads to decline in cortical function in human disease **(D)**.

et al., 1999), multiple mouse models of the disease have been generated (Guy et al., 2001; Chen et al., 2001) which recapitulate physiological and behavioural phenotypes as well as the developmental trajectories observed in human disease (Moretti, 2006; Vashi & Justice, 2019). Additionally, cellular mechanisms underlying RTT pathology have been observed in vitro in both the acute brain slices (Dani et al., 2005; Dani & Nelson, 2009) and cell cultures (Baj et al., 2014).

Studying murine in vitro primary dissociated cortical cultures allows us to observe the earliest deficits in network function and how the disease alters the fundamental mechanisms driving network development at the cellular scale, and how these changes impact the developmental trajectories. Moreover, such studies confer us with more experimental freedom than when working with live mice and are cheaper, more ethical, and faster than culturing human stem cells (Marchetto et al., 2010; Walsh & Hochedlinger, 2010).

1.1.2 Rett syndrome as a model circuit disorder

RTT is a disorder of neuronal circuits (Shepherd & Katz, 2011; Goffin & Zhou, 2012; Kron et al., 2012). Early electrophysiological recordings from acute mouse cortical slices suggested that RTT involves synapses and microcircuits primarily with relative sparing of intrinsic neuronal firing properties (Dani et al., 2005). Indeed, excitatory-inhibitory (E-I) imbalances and cortical hypoconnectivity have been later found both in vitro and in vivo (Dani & Nelson, 2009; Wood & Shepherd, 2010; Durand et al., 2012; Sceniak et al., 2016). The cellular mechanisms responsible for the pathogenesis of RTT are shown in Figure 1.1.

However, besides the findings mentioned above, the deficits in functional connectivity and neuronal network dynamics in RTT are poorly understood (Sit, 2018). Further understanding of this disease will require the characterisation of these aspects of network development. This thesis will quantify the effective dimensionality of developing cortical circuits and explore the links between network structure and function using approaches from network control theory.

1.2 Studying network dynamics at cellular scale

Determining the relationship between structure and function is a key challenge in modern neuroscience. The last two decades have seen a rapid increase in conceptual frameworks aiming to address it (Ju & Bassett, 2020). Among these, particular attention has been given to the study of neural networks via connectomic approaches. Using concepts borrowed from graph theory, one can explore how different elements (*nodes*) of the network interact with each other through functional or anatomical connections (*edges*), and how these interactions give rise to neural function at whole-brain (Schröter et al., 2015) and microcircuit scale, both in vivo and in vitro (Poli et al., 2015; Schröter et al., 2017). Despite the widespread use and utility of these approaches, they only capture snapshots of the network function without the temporal dimension in which it exists. Attempts were made to circumvent this obstacle but were largely limited to sliding-window approaches (Sizemore & Bassett, 2018).

Network control theory offers a new perspective on functional brain networks, enabling the modelling of the temporal evolution of network dynamics as a function of its connectivity (Gu et al., 2015). In this study, I aim to characterise the behaviour of neuronal networks (*system*) through the observations of spontaneous spiking activity (*states*). I further define the *state space* as a collection of all states permissible by the biophysical properties of neurons and the connections between them. I apply this framework to neuronal networks in vitro.

During development, cultured neurons wire together, forming networks (Downes et al., 2012; Chiappalone et al., 2019) which can be probed using electrophysiological recordings from microelectrode arrays (MEA). The recorded spontaneous electrical activity can be used to infer functional connectivity (Cutts & Eglen, 2014; Poli et al., 2015) and to characterise the dimensionality of network dynamics (Recanatesi et al., 2019).

1.2.1 Effective dimensionality of network dynamics

The observed repertoire of network activity patterns might span as many dimensions as there are neurons in it. However, the dynamics of the whole network are different from the summed activity of individual components. Previous analyses of neural activity using dimensionality reduction methods revealed that the underlying dynamics can be described by a much lower number of latent (unobserved) dimensions than the original data (Cunningham & Yu, 2014; Gao et al., 2017). These findings are especially relevant for our understanding of brain function because neural networks also perform this kind of computations. During population coding, activities from thousands of individual neurons are combined to robustly decode sensory infor-

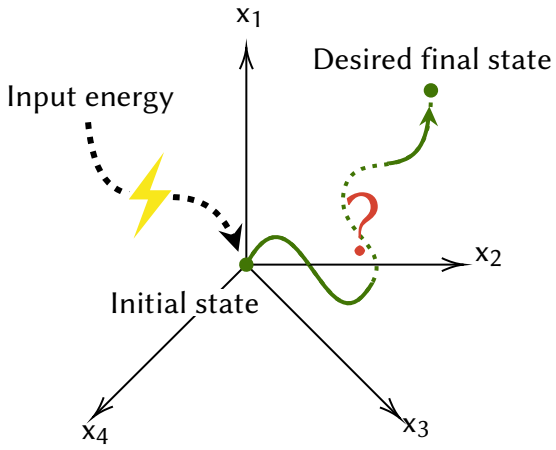


Figure 1.2: Control theory is concerned with finding optimal perturbation strategies. In this study, I investigate whether one can steer a system from an arbitrary initial state to a desired final state with input energy in finite time. In this diagram, the considered system exists in a $n = 4$ dimensional state space (dimensions x_1–x_4). This could represent a network of four neurons, with each dimension in state space specifying the spiking activity of the corresponding neuron. Figure created by me in Ti*k*Z.

mation and produce appropriate functional output (Rossi-Pool & Romo, 2019).

Dimensionality provides a measure of how coordinated the network-wide activity is in terms of the number of dimensions in the state space necessary to explain the observed firing patterns. Because the complexity of network dynamics is strongly regulated by local connectivity motifs (Recanatesi et al., 2019), this framework allows us to quantify the dynamical portrait of circuit computations in developing cortical networks.

1.2.2 Exploring relationships between network structure and dynamics using network control theory

The ultimate proof of our understanding of complex dynamical systems lies in our ability to control them. Over the last century, a rich body of scientific literature has grown to address this challenge, giving birth to the discipline of control engineering. However, little of it has percolated to neuroscience (Kao & Hennequin, 2019).

The notion of control, as studied by control theory, conceptually agrees with our everyday understanding of it. To *control* means to *exert influence in order to achieve the desired goal.* In the scope of control engineering, this means steering a system from one state to another with energy delivered through a set of inputs (Fig. 1.2). Applied to neural networks, we can modulate a selected set of nodes using pharmacological, optogenetic or electrical interventions (either excitatory or inhibitory).

Analysis of neuronal networks in the context of control theory may be crucial

to understanding how neural dynamics relate to brain function in health and disease. Targeting network dynamics in therapeutic interventions can maximise beneficial outcomes for patients. Examples include automatic maintenance of anaesthesia (Ching et al., 2013), optimizing deep brain stimulation in Parkinson's disease (Johnson et al., 2016) and seizure suppression in epilepsy (Berenyi et al., 2012; Jobst et al., 2017). Control engineering also lies at the heart of the rapidly developing field of neuroprosthetics and brain-machine interfaces (Tsu et al., 2015). However, this framework can also be applied to further our understanding of the cellular basis of cognitive disorders, including Rett syndrome and Autism Spectrum Disorders (Sit, 2018).

1.3 Aims

In this study, I aim to:

1. Characterise the relationship between network function and structure inferred from the observed neuronal spike trains across development using dimensionality analysis.

2. Investigate mechanisms underlying the emergence of cortical network dynamics using network control theory.

3. Elucidate the effects of developmental age and Mecp2 deficiency on the dynamical properties of in vitro neuronal networks.

4. Determine whether network dynamics framework is suitable for studying in vitro model of neurodevelopment in RTT.

2

Methods

2.1 Animals used in the study

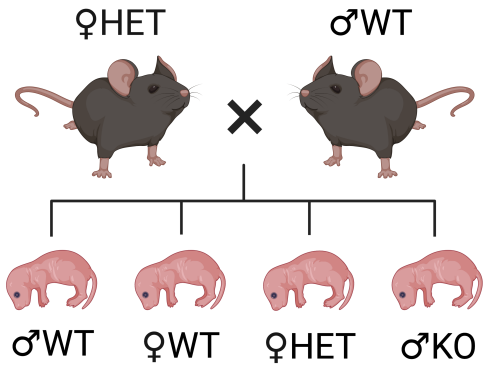


Figure 2.1: Female C57BL mice heterozygous for Mecp2 deletion (*Mecp2*$^{+/-}$) were bred with wild type male mice (*Mecp2*$^{+/y}$), yielding equal genotype distribution in the offspring. WT: wild type, HET: heterozygous, KO: Mecp2 knockout. Figure created in BioRender by me.

All animals were maintained and bred by the personnel of the on-site animal facility. Female C57BL mice heterozygous (HET) for *Mecp2* deletion (*Mecp2*$^{+/-}$) were bred with wild type (WT) male mice (*Mecp2*$^{+/y}$) (Fig. 2.1). This yielded in an approximate 1:1:1:1 offspring genotype spread as shown in Table 2.1. Genotyping was performed by Transnetyx on postmortem tissue.

Genotype	Phenotype	Sex	Abundance
Mecp2 $^{+/y}$	WT	Male	25%
Mecp2 $^{+/+}$	WT	Female	25%
Mecp2 $^{+/-}$	HET	Female	25%
Mecp2 $^{-/y}$	KO	Male	25%

Table 2.1: Approximate distribution of offspring genotypes.

2.2 Microelectrode array preparations

Microelectrode array (MEA) chips (60MEA200/30iR-ITO-gr, Multi Channel Systems, Reutlingen, Germany) were prepared using a published protocol (Charlesworth et al., 2015). MEAs were treated with 0.5 mg$\times$ml^{-1} proteinase K (Sigma number: SAE0009-25MG), incubated for 24–48 hours and then rinsed with sterile distilled water (LifeTechnologies number: 15230071) to remove remaining cellular debris. MEAs were next sterilised in an autoclave. Sterilised MEA grids were treated with 0.01% poly-L-lysine (Sigma number: P4832) and subsequently washed with phosphate-buffered saline (PBS, LifeTechnologies number: 14190094) three times before coating them with laminin (Sigma number: L2020-1MG) 24–48 hours before the initiation of cell culture. Alexander Dunn performed the entire procedure.

2.3 Initiation and maintenance of primary dissociated cortical cultures

Newborn mouse pups were sacrificed in accordance with Home Office approved procedures at either postnatal day (P) 0 or 1. Cerebral cortices were dissected in ice-cold PBS (Life Technologies number: 14190094) and transferred under sterile conditions to a 1:1 solution of papain (Sigma number: P5306-25 mg) and PBS for chemical dissociation in a 37℃ water bath for 25 minutes. The papain reaction was halted by adding 4% Foetal Bovine Serum in Neurobasal Media (LifeTechnologies number: 21103049) with B27 supplement (NB-B27, Life technologies number 17504044. Next, cells were manually dissociated and the obtained cell suspension was centrifuged at 0.4 rcf for 10 minutes. The supernatant was aspirated, and the cell pellets were resuspended in 350 µL warm (37℃) NB-B27 media and manually dissociated again. Cell concentration was determined using a haemocytometer in order to calculate the volume of cell suspension required to plate 2×10^5 cells onto each MEA grid. Each MEA

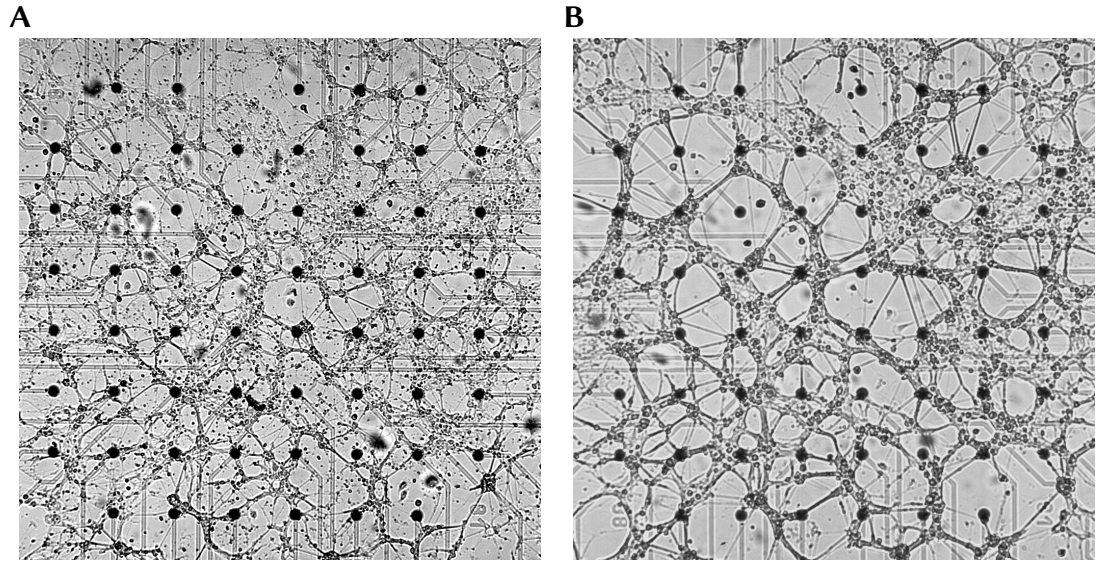


Figure 2.2: Photographs of cortical cultures developing on the MEA grid. **(A)** A sample wild type culture at DIV 7. The grid coverage is satisfactory and some axons are visible. **(B)** The same culture at DIV 21. Note the conspicuous network formation and growth of the axon tracts. Photographs courtesy of Alexander Dunn. Figure created by me in Ti*k*Z.

grid contained 600 µL NB-B27 media containing 1.5 µL (0.25%) L-glutamine (LifeTechnologies number: 25030024) at 37℃. Cells were incubated at 37℃ and 5% CO2/95% air. Three times per week, starting from 3rd day in vitro (DIV3), one third (200 µL) of media was removed and replaced with 0.25% L-glutamine (200 mM) in NB-B27 solution. Each MEA was fitted with a lid to prevent evaporation and to maintain sterile conditions over several weeks (Potter & DeMarse, 2001). Cultures were inspected weekly to ensure no perturbation or death of cells on the grid, and photos of cultures on MEAs (Fig. 2.2) were taken using a microscope setup with the `Wasabi` software (Hamamatsu software). Cultures with fewer than one fifth of electrodes covered by cells were discarded. Cell culture initiation and maintenance have been carried out by Alexander Dunn with assistance from other lab members.

2.4 Electrophysiological recordings using microelectrode arrays

The MEA chip is 2×2 mm and comprises 60 electrodes arranged in an 8×8 grid (without corners). As shown in Figure 2.3, each electrode is 30 µm in diameter, and the electrodes are spaced 200 µm apart. Fifty nine electrodes are used for recordings. One is used as a reference (electrode in column 1, row 5 of the MEA). Spontaneous activity from cell cultures growing on the MEA grid was recorded at 25 kHz sam-

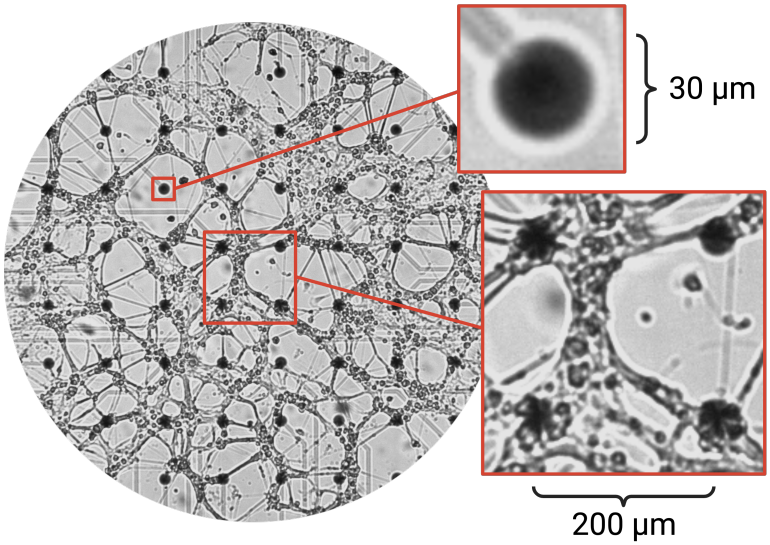


Figure 2.3: Photograph of a cell culture growing on microelectrode array (MEA). MEA is 2 × 2 mm and comprises 60 electrodes arranged in an 8 × 8 grid (without corners). Each electrode (black circle) is 30 µm in diameter, and the electrodes are spaced 200 µm apart. The effective recording range is up to 45 µm from the centre of the electrode. Figure created by me in Ti*k*Z.

pling frequency using MEA2100 amplifier and MC `Rack` software (Multi Channel Systems). Before each recording, the voltage readout was checked for noise levels. Channels with voltage fluctuations exceeding 50 µV were grounded (the signal vector set to that of the reference channel) using MEA `Select` software (Multi Channel Systems). If this occurred over more than 20% of electrodes, the recording was excluded from further analysis. Initial recordings, started at the 7^{th} day in vitro (DIV7) lasted 2 minutes and were followed by weekly 12-minute recordings for 5 weeks (up to DIV35). Temperature of 37℃ was maintained throughout recordings using `TC01` temperature controller and the `TCX-Control` software (Multi Channel Systems). All of these procedures were performed by Alexander Dunn under full blinding to culture genotypes.

File output from MC `Rack` was converted to binary data format (`.raw`) using `MC_DataTool` (Multi Channel Systems) and finally to MATLAB data format (`.mat`) using custom MATLAB scripts. Before pre-processing, voltage traces were plotted and again inspected for possible artifacts that might have originated from damaged electrodes or poor contact between electrodes and amplifier pins. The recording procedure has been summarised in Figure 2.4.

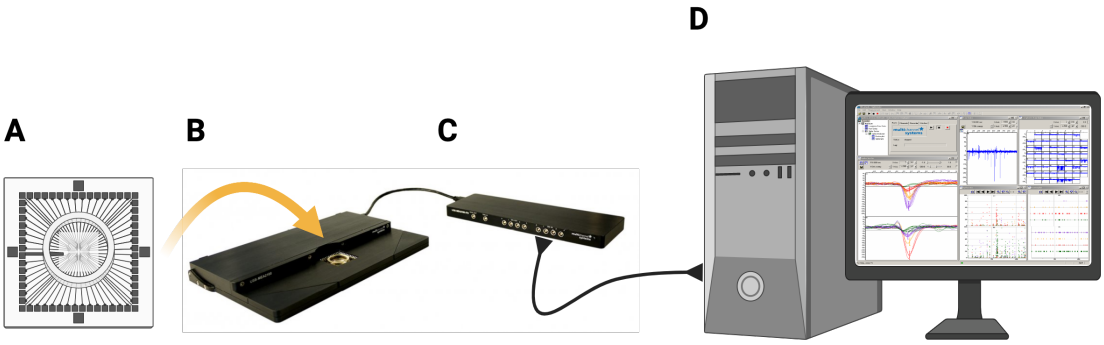


Figure 2.4: MEA recording setup. For recording, MEA chip **(A)** is placed in the headstage **(B)**. Voltage fluctuations originating from neuronal activity or due to noise are detected and sent to the amplifier **(C)**. Amplified analog signal is digitized and sent to the computer where it can be visualised using the MC_Rack software. At this step, channels with excessive baseline voltage fluctuations can be grounded **(D)**. Collected data is then converted to a MATLAB-readable format for analysis. Figure created by me in BioRender.

2.5 Pharmacology

For the validation of detected spikes, three DIV50 cultures have been recorded for three minutes as a baseline. After each recording, half of the media in the MEA well was replaced with 1 µM tetrodotoxin (TTX) in NB-B27. Cells were incubated with TTX for five minutes before a subsequent six-minute recording.

2.6 Data pre-processing

Filtering is a crucial signal processing step that ensures the exclusion of irrelevant background activity and noise. Noise may originate from the cell culture itself (local field potentials and other subthreshold electrical events, heat dissipation), recording electrodes ($1/f$ noise), amplifiers and other electronic instruments, external electrostatic and electromagnetic interference (50-60 Hz mains hum), mechanical vibrations, and the data digitization process (Sherman-Gold, 1993). In this investigation, I applied a 3rd order Butterworth band-pass filter with cutoffs at 600 Hz and 8,000 Hz for the low- and high-pass respectively (Schröter et al., 2015) (Fig. 2.5).

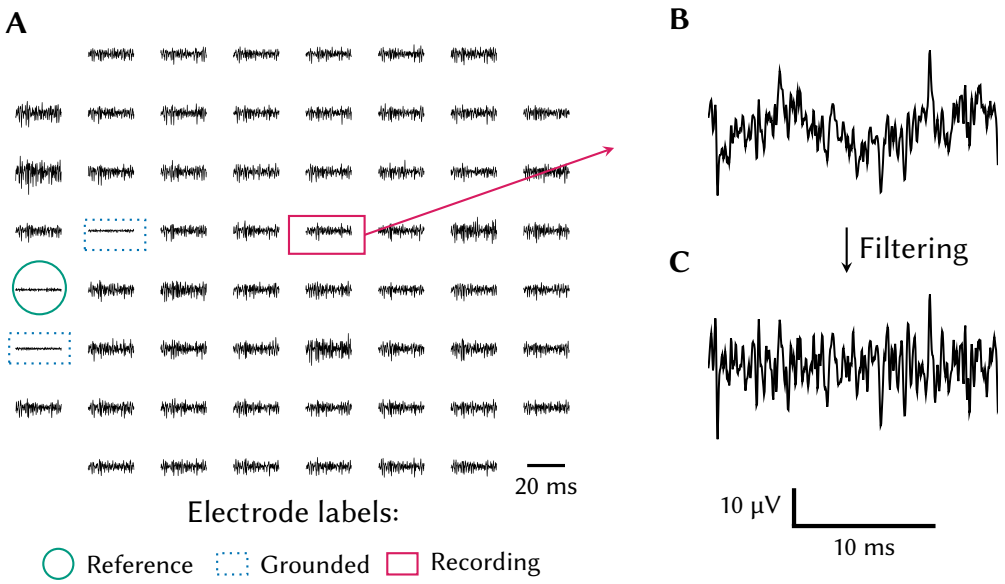


Figure 2.5: Voltage readout from MEA chip. **(A)** Voltage traces for each of the electrodes arranged according to their physical position on the MEA grid. Dotted rectangles represent examples of grounded electrodes. Circle shows the reference electrode at its constant position 15 (column 1, row 5). Unmarked traces correspond to recording electrodes. **(B)** Raw voltage trace. **(C)** Same trace as in (B) after filtering with 3rd order Butterworth bandpass filter with cutoffs at 600 Hz and 8,000 Hz for low- and high-pass respectively. Figure created by me in Ti*k*Z.

2.7 Spike detection

Neural information is encoded and processed as sequences of action potentials called spike trains (Gabbiani & Metzner, 1999). Capturing these signals is paramount for our understanding of neuronal function. Spike detection is a process whereby extracellular action potentials are extracted from electrophysiological recordings also containing irrelevant background neuronal activity and noise. In this investigation, I have developed a custom spike detection method in collaboration with my colleagues. This method has been tailored for the detection of multi-unit activity and relies on continuous wavelet transform (Nenadic & Burdick, 2004) with data-driven template selection. Benchmarking has revealed that it is more accurate than the best published method (Lieb et al., 2017) by at least 8% (Appendix A).

A graphical summary of the experimental methods can be found in Figure 2.6.

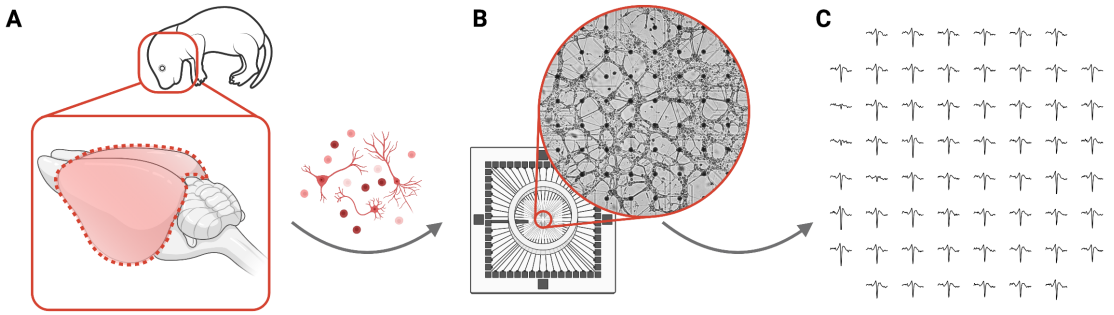


Figure 2.6: Experimental methods summary. **(A)** Entire cerebral cortices were extracted from newborn mouse pups, dissociated, and plated directly onto the MEA grid. **(B)** Cell culture developing on the MEA grid. The aim of this study is to explore the functional characteristics of the networks developing in these cultures. **(C)** Average spike waveforms from a sample MEA recording. Sixty electrodes (59 recording and one reference) were used in the current investigation. Some of the recording electrodes had to be grounded due to excessive voltage fluctuations caused by noise. Neuronal action potentials were extracted in the spike detection step and used to quantify the development of cortical networks.

2.8 Inferring network structure from spontaneous neuronal activity

I have used the spike trains identified from simultaneous recordings of the spontaneous electrical activity across the network to infer the functional connections between neurons. Pairwise correlations in multi-unit activity between electrodes were quantified using spike time tiling coefficient (STTC; Cutts & Eglen, 2014). I have adapted the original code in C provided by the authors, and used the MEX framework to call it from MATLAB, improving computational time by 10-fold compared to a pure MATLAB implementation. I made the code available at `https://github.com/jeremi-chabros/sttc`.

2.9 Statistical analysis

Statistical analysis was performed in MATLAB (R2020b, 9.9.0.1467703, The MathWorks Inc.) and statistical power calculations were done using the G*Power software (Faul et al., 2009). Data were tested for normal distribution quantitatively using the Anderson-Darling test and qualitatively using quantile-quantile plots. Obvious outliers were removed based on the Cook's distance (Altman & Krzywinski, 2016).

Where data was normally distributed, I compared developmental trajectories of different genotypes using repeated measures analysis of variance (rmANOVA) with the assumption of compound symmetry (sphericity). I used Mauchly's test of spheric-

ity to determine if this assumption was violated. If so, I employed the Greenhouse-Geisser Correction, producing correction factor ϵ. Degrees of freedom were then adjusted by multiplying the original values by ϵ. For comparisons between specific ages, I used the Welch's unequal variances t-test. Effect sizes were calculated using Hedges' g for independent samples (Hentschke, 2021) with a small sample bias correction as per Hedges (Hedges, 1981).

Where data was not normally distributed, I used Mann Whitney U test for pairwise comparisons and Friedman's nonparametric repeated measures test for comparing developmental trajectories and genotype effects. Effect sizes were calculated using Mann Whitney U statistic (Conroy, 2012).

$$c = \frac{U}{N_1 \times N_2}$$

where c is the effect size, U is the Mann-Whitney U statistic, and N_1 and N_2 are the corresponding sample sizes.

2.10 Data analysis and visualisation

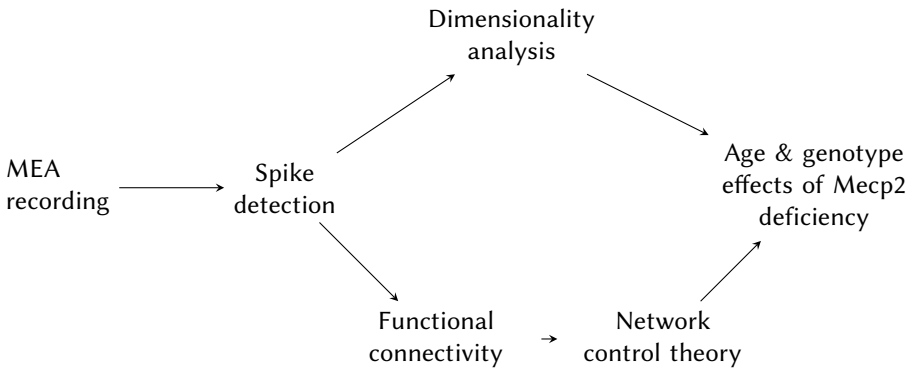


Figure 2.7: A diagram summarizing the data analysis pipeline.

All data analysis was performed using MATLAB and the pipeline is summarised in Figure 2.7. I created all the figures using MATLAB or LaTeXpackage Ti*k*Z. I produced graphics and diagrams in BioRender. I typeset my thesis in LaTeXusing the XeTeXengine. Fonts, colours, and style in this thesis were chosen to be appropriate for people with dyslexia or colourblindness.

To identify the analysis methods I apply in this study, I first evaluated the spike detection methods used by our group. Then, I searched the literature for other published algorithms. To improve the performance of available methods, I developed and benchmarked a novel spike detection algorithm. Next, I used the spike trains to characterise the functional connectivity of developing cortical networks and calculate the

dynamical dimensionality of these networks. Then, I have quantified the control profiles of cortical circuits using energy-related controllability metrics, including average and modal controllability and the controllability ellipsoid. Data analysis and visualisation can be reproduced using scripts publicly available on my GitHub:
`https://github.com/jeremi-chabros`

3

Effective dimensionality of network dynamics

3.1 Discrete vs. continuous dimensionality

Many methods have been employed to characterise the dimensionality of neural network dynamics. Some examples include principal component analysis (PCA), independent component analysis (ICA), linear discriminant analysis (LDA), and factor analysis (FA) (Cunningham & Yu, 2014; Sadiq et al., 2021). However, all these methods quantify the number of dimensions required to explain a specified amount of observed variability based on an arbitrary threshold (e.g. obtained with the elbow method) and in terms of discrete values. Using integers may blur the potential differences between experimental conditions, such as culture age or genotype. In this study, I have chosen the effective rank to address these issues as it is agnostic to any thresholding and provides a continuous measure of the network's dynamical dimensionality.

3.2 Effective rank

Effective rank is a measure of effective dimensionality (Roy & Vetterli, 2007). Since its conception, it has been applied to neuronal networks in a single study in which the effective rank has been calculated from the correlation matrix of neural activity (Sit, 2018). In my current study, I have taken a different approach whereby I

calculated the effective rank directly from the spike trains to achieve a better temporal resolution to quantify network dynamics. Below, I outline the intuition behind this measure.

In linear algebra, a rank of a matrix is an integer value that quantifies the maximal number of linearly independent dimensions (rows or columns). If all of the matrix dimensions are independent (in other words, where no coupling between variables exists), the rank of such matrix will be equal to the lesser of the number of rows and columns. Conversely, if all of the variables are coupled, the rank will be one.

In respect to electrophysiological recordings from neural networks, a column could represent a node's activity across the recording duration. Each row would then summarise the state of all the nodes at a given time bin. In this scope, the rank will equally quantify the number of unique patterns of activity, and the number of dimensions in the state space needed to describe the variability in the observed neural activity. Because each node's activity might participate in different temporal motifs to a variable extent, the whole network's activity will travel through the dimensions of the state space continuously, rather than in saltatory jumps from one dimension to another. Therefore, effective rank is a more accurate measure of dimensionality than the classical interpretation using rank.

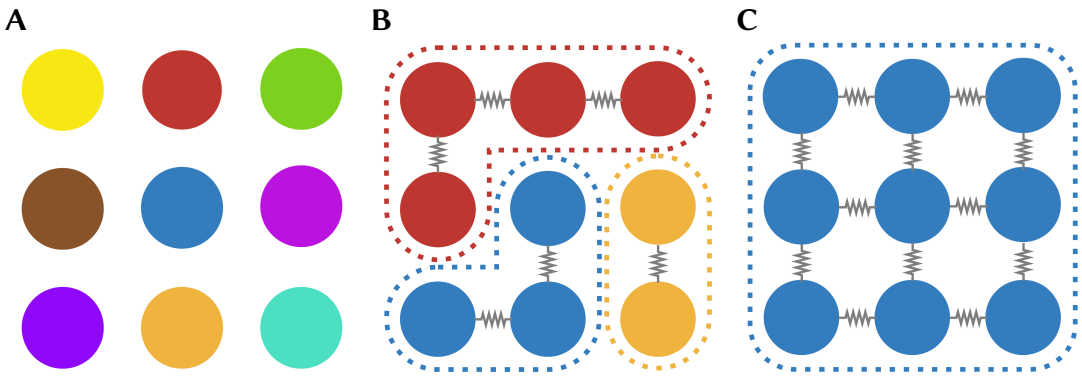


Figure 3.1: Graphical representation of network dimensionality. **(A)** A network with completely decoupled dynamics (complete independence). **(B)** A network with partially coupled dynamics (subgroup independence). **(A)** A network with completely coupled dynamics (complete dependence). Each colour represents a unique pattern of activity and dotted lines represent synchronised subgroups. The connections between nodes represent coupling of their dynamics (i.e. shared patterns of activity).

As a demonstration, let us consider a sample network of nine neurons (Fig. 3.1) whereby we assign a different colour to each unique pattern of activity (dimension in state space). For the purpose of illustration, let us disregard the fact that some colours were generated by a mixture of primary colours, and assume that they are all primary colours (unique and not able to be generated by mixing colours) [1]. In case where the

[1]Notice that we will obtain a more realistic, although much less interpretable, picture of network

dynamics of each node in a network are independent (decoupled) from those of other nodes, the effective rank will be equal to the network size (erank = 9)(Fig. 3.1A). If, due to functional connectivity, some nodes share the same pattern of activity, it can be used to distinguish them from other subgroups with independent activity (Fig. 3.1B). Thus, the effective rank will be equal to the number of subgroups (erank = 3). Finally, if all of the nodes in a network are synchronised, the effective rank will be one (Fig. 3.1C).

3.3 Mecp2 deficiency leads to increased dimensionality of network dynamics

To characterise the development of functional subcommunities sharing patterns of activity in the neuronal networks, I applied effective rank to the MEA recordings from Mecp2-wildtype (WT, n = 17), heterozygous (HET, n = 15) and knock-out (KO, n = 15) cortical cultures. Based on previous results (Perin et al., 2011; Sit, 2018), I hypothesised that during development the effective dimensionality of cortical circuits will decrease following the emergence of synchronised network activity, and that this process will be disrupted in HET and KO cultures.

Compared genotypes	Difference[a]	Confidence interval	p
HET vs KO	−3.43	-6.59 to -0.28	0.03 *
HET vs WT	0.18	-2.97 to 3.34	1.00
KO vs WT	3.62	0.46 to 6.77	0.02 *

Table 3.1: Table summarising statistical differences in effective rank between genotypes (Friedman repeated measures test with multiple comparisons). * $p < 0.05$
[a] Difference in average column rank.

I found that KO cultures had an overall higher effective rank than both WT and HET cultures, which themselves were not found to be different from each other (Tab. 3.1). These results might indicate that over the first five weeks of development, WT and HET cultures had an overall better capacity for network synchronisation than KO cultures. The high dimensionality of Mecp2 deficient cultures suggests developmental delay due to constraints in forming meaningful connections or subcommunities within the larger network. Alternatively, these cultures might compensate for the deficits in subgroup dependency with a more flexible dynamic repertoire, promoting general computations (Recanatesi et al., 2019). The observed genotype differences

dynamics when we lift this assumption. Primary colours will represent the principal directions in the state space (independent temporal motifs). Mixing the colours will mean that each node might participate in each of these patterns to a variable extent. The original whole network activity could then be reconstructed by summing together each primary colour's contributions to the different hues.

might reflect a dose-effect relationship between the expression of Mecp2 and the effective dimensionality of network dynamics.

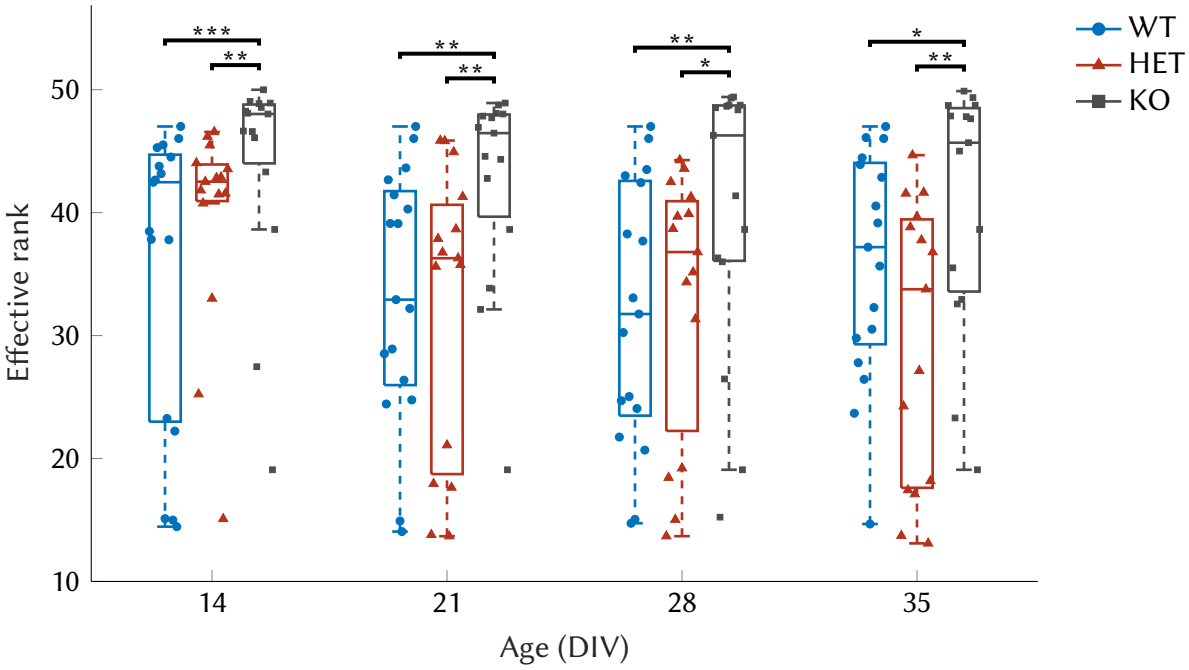


Figure 3.2: Box plot (median, interquartile range, minimum and maximum) of age and genotype effects on effective rank. DIV – days in vitro. Genotypes: WT – wild type, HET – heterozygous, KO – Mecp2 knockout. * $p < 0.05$, ** $p < 0.01$.

Over the whole duration of the study, I observed significantly lower network dimensionality in WT cultures when compared to Mecp2-knockouts (Fig. 3.2). Similarly, the effective dimensionality of HET cultures was lower than that of KO cultures, but I found no statistically significant differences between the WT and HET cultures. All of the pairwise comparisons achieved effect sizes ranging between 1.17 and 1.33, implicating sufficient power (at $1 - \beta = 0.8$ threshold) of the statistically significant results (Tab. 3.2).

At DIV14, WT cultures have already achieved substantial subgroup synchronisation levels that were maintained over the duration of the study. At the same developmental age, the HET cultures had a lower dimensionality than KO cultures, but the values were narrowly distributed. The variability between different HET cultures increased over development, suggesting a developmental trajectory whereby some cultures achieved more meaningful activity patterns than others, which could be explained by the random X-inactivation and therefore variable Mecp2 expression. Despite significant impairments in the formation of communities and local processing (as compared to HET and WT cultures), the effective dimensionality of some Mecp2 deficient networks has decreased during development. This could mean that there exists a combination of function and dysfunction in cortical circuits in RTT, and that their development and functional maturation is delayed. Interestingly, although the

Compared genotypes	Age (DIV)	U	z	p	Effect size [a]
WT vs KO	14	331.00	3.13	0.001 ***	1.30
	21	325.00	2.91	0.002 **	1.27
	28	310.00	2.34	0.009 **	1.22
	35	299.00	1.93	0.027 *	1.17
WT vs HET	14	262.00	0.53	0.299	n/a
	21	237.00	−0.42	0.661	n/a
	28	254.00	0.23	0.410	n/a
	35	207.00	−1.55	0.939	n/a
HET vs KO	14	299.00	2.74	0.003 **	1.33
	21	300.00	2.78	0.003 **	1.33
	28	283.00	2.07	0.019 *	1.26
	35	297.00	2.65	0.004 **	1.32

Table 3.2: Table summarising the statistical results of pairwise comparisons (two-tailed Mann-Whitney U test) between genotypes in respect to effective rank. * $p < 0.05$, ** $p < 0.01$.
[a] Effect sizes reported only for statistically significant differences.

theoretic maximum value of effective rank in current investigation was 59, even KO circuits at earliest developmental stages (DIV14) have shown some levels of network synchronization, indicating that functional circuits emerge early in development, even in RTT disease model.

However, these findings might be potentially biased by variable culture growth during early development and differences in functional network sizes between genotypes. Therefore, a modification to this method was required.

3.4 Relative effective rank reveals changes in dynamical dimensionality due to network size

Relative effective rank is the effective rank normalised towards network size. I define network size as the number of active nodes in the network (those with a spiking frequency exceeding 0.5 Hz). This measure of dimensionality quantifies the extent to which the propensity for subcommunity pattern synchronisation, as quantified by the effective rank, is realised by the network. This measure is bounded by $\left(0, 1\right]$, meaning complete synchronisation or complete independence of firing patterns, respectively.

Compared genotypes	Difference[a]	Confidence interval	p
HET vs KO	–3.35	-6.51 to -0.19	0.03 *
HET vs WT	0.00	-3.15 to 3.15	1.00
KO vs WT	3.35	0.20 to 6.52	0.03 *

Table 3.3: Table summarising statistical differences in relative effective rank between genotypes (Friedman repeated measures test with multiple comparisons). * $p < 0.05$
[a] Difference in average column rank.

I have noted similar trends with relative effective rank as with its non-normalized counterpart (Fig. 3.3, Tab. 3.3, Tab. 3.4). However, I have observed a notable change in respect to the dimensionality of WT and KO cultures at DIV35, where the previously significant difference between cultures has disappeared after accounting for network size. Additionally, the finding that the highest values of relative effective rank approached 0.9 and not 1 suggests that even relatively dysfunctional cultures (compared to WT) exhibited some levels of synchronisation already at early developmental stages.

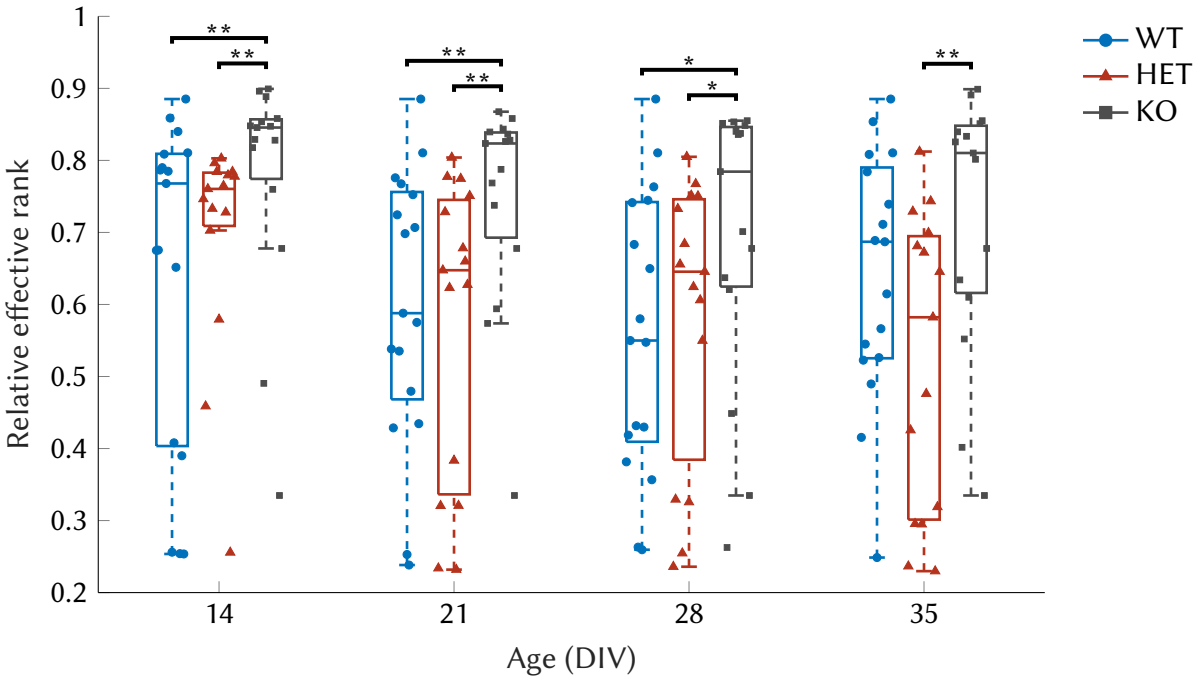


Figure 3.3: Box plot (median, interquartile range, minimum and maximum) of age and genotype effects on relative effective rank. DIV – days in vitro. Genotypes: WT – wild type, HET – heterozygous, KO – Mecp2 knockout. * $p < 0.05$, ** $p < 0.01$.

Compared genotypes	Age (DIV)	U	z	p	Effect size [a]
WT vs KO	14	312.00	2.42	0.008 **	1.22
	21	316.00	2.57	0.005 **	1.24
	28	300.00	1.96	0.025 *	1.22
	35	288.00	1.51	0.065	n/a
WT vs HET	14	240.00	−0.30	0.619	n/a
	21	239.00	−0.34	0.633	n/a
	28	255.00	0.26	0.396	n/a
	35	206.00	−1.59	0.944	n/a
HET vs KO	14	298.00	2.70	0.003 **	1.32
	21	300.00	2.78	0.003 **	1.33
	28	281.00	1.99	0.023 *	1.25
	35	296.00	2.61	0.004 **	1.32

Table 3.4: Table summarising the statistical results of pairwise comparisons (two-tailed Mann-Whitney U test) between genotypes in respect to relative effective rank. * $p < 0.05$, ** $p < 0.01$.
[a] Effect sizes reported only for statistically significant differences.

4

Developing cortical networks: control theory perspective

4.1 Understanding network controllability in RTT

The wiring of neural networks during healthy brain development increasingly supports coordinated control of neural activity (Tang et al., 2017). Circuit hypoconnectivity patterns observed in RTT suggest that the disruption of this process might lead to deficits in cognitive control manifesting in intellectual disability. However, the mechanisms and extent to which these changes affect developing cortical circuits' dynamics and control profiles remain unknown (Sit, 2018).

How do control profiles of cortical networks change over development? Are healthy and diseased networks different from each other? How easily can we steer the network from one arbitrary state to another? In other words, to what extent is the network *controllable*? What types of states, as defined by the *controllability metrics*, can networks reach? Which nodes are the most influential in driving network's dynamics? What is the cost of control?

Answering these questions may help us better understand circuit pathology in RTT and rescue the decline in cognitive function through developing focused, mechanism-based control strategies for effective pharmacological modulation (Scaramuzza et al., 2021) or deep brain stimulation (Hao et al., 2015).

4.2 Modelling network dynamics using control theory

To investigate the developmental trajectories of cortical networks, I searched the neuroscience, mathematics, engineering, and computer science literature for energy-related controllability metrics. This class of metrics deals with the amount of input energy required to steer a system from an initial state to a desired final state in finite time, based on network's topology and dynamics. I approximated the functional architecture of the networks using the spike time tiling coefficient (STTC; Cutts & Eglen, 2014), excluding self-loops. To model network dynamics, I assume that

- We are sampling the continuous behaviour of a network in discrete time steps (here, at 25 kHz frequency),
- The architecture of the network is constant throughout the electrophysiological recording (here, 12 minutes),
- The network exhibits linear dynamics (or its dynamics can be linearised around fixed point),
- We are only able to control a single node at a time.

In control engineering, this set of assumptions is called the discrete-time linear time-invariant (LTI) system approach (for mathematical formulations, see Appendix). In this report, I first apply this framework to determine the role of each node in controlling the network dynamics using average and modal controllability measures. Then, I identify the extent to which the whole network is controllable using a controllability ellipsoid.

4.3 Nodal controllability metrics

Nodal controllability metrics describe which nodes in the network are most influential in constraining or facilitating state trajectories. Each of these diagnostics captures a different goal (Karrer et al., 2020). Together, they answer two questions: 1) which nodes are the most influential in driving the network's dynamics and 2) what types of states can be reached by controlling these nodes. In this study, I quantify the control profile of the entire network by averaging the controllability values across all nodes.

In our model, input energy is delivered as an impulse to a single node at a time. Thence, the control energy is transferred across edges to influence other nodes in the

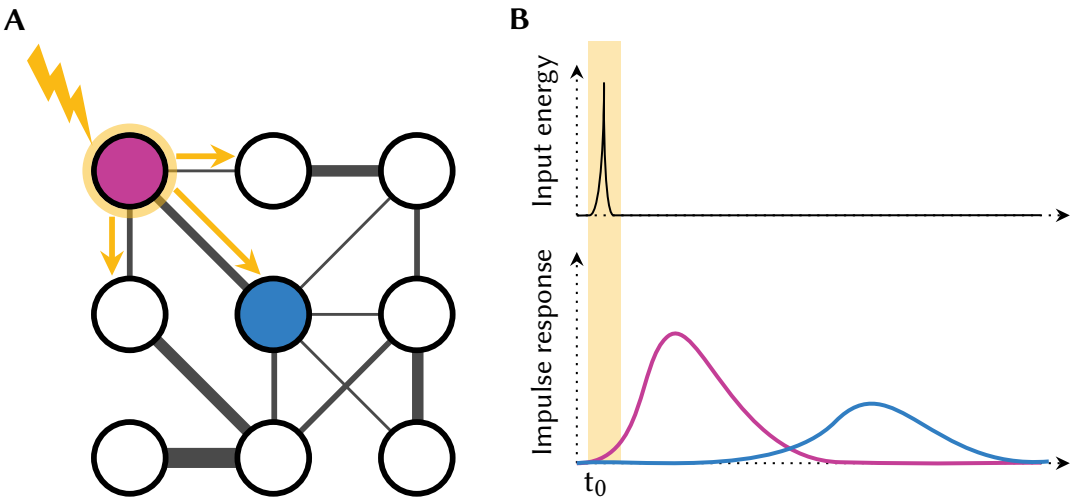


Figure 4.1: Functional architecture determines the network's responses to input energy. **(A)** A sample network of nine neurons with control input (lightning bolt symbol) to a single node (pink circle). The input energy is transferred across connections (yellow arrows) to influence other neurons (e.g., blue neuron) in the network. Edge weights depict the strength of functional connectivity as determined by STTC. **(B)** Input energy (lightning bolt symbol) delivered as an impulse (top) dissipates across connections with a time constant proportional to their strength. The pink curve illustrates the response of the pink neuron in (A), and the blue line shows the response of the blue neuron. Note the time lag and attenuation. Figure created by me in Ti*k*Zfrom synthetic data purely for visualisation purposes.

network (Fig. 4.1A). The input energy passing through connections attenuates with a time constant directly proportional to edge weights. In other words, weaker connections attenuate more (Fig. 4.1B). Because summed responses of each node determine the network's behaviour, the nodal controllability metrics allow us to identify nodes with the most influence over network dynamics.

The impairment of motor and cognitive function in RTT suggests that the volume of the easily reachable subspace might be decreased or that the state space is skewed towards difficult-to-reach states (Fig. 4.2). Such a finding would also support the lower flexibility in network dynamics thought to underlie autistic symptoms (D'Cruz et al., 2013) experienced by approximately half of the patients affected by RTT (Wulffaert et al., 2009). Because mosaic expression of MECP2 has differential effects on different cells and cell populations, identifying the most controllable nodes might help inform further research and target therapeutic modulation, thus reducing potential adverse effects on the rest of the network.

4.3.1 Average controllability

Average controllability identifies nodes that, on average, can steer the system between states with little effort (input energy). The states in consideration are all states

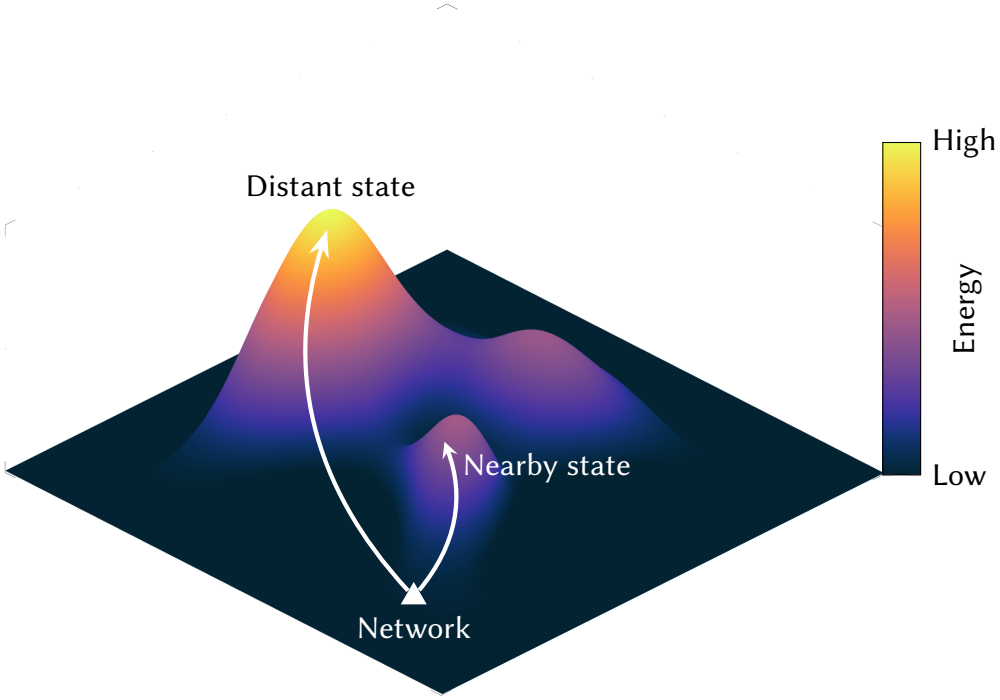


Figure 4.2: Controllability in a three-dimensional energy landscape. Modal controllers are capable of driving the network to distant, difficult-to-reach states, whereas average controllers are better suited to facilitate state transitions towards nearby, easily-reachable states. There are multiple alternative paths to reaching the desired state (not shown). Figure created by me in Ti*k*Zfrom synthetic data purely for visualisation purposes.

permissible by the network architecture. In an energy landscape (Fig. 4.2), low-energy states vastly outnumber high-energy states (Gu et al., 2015). Therefore, we can also use the average controllability to quantify each node's ability to drive the network towards easily reachable states (those nearby the current state in the state space). The higher the average controllability, the more capable a node is in driving network dynamics towards nearby states.

Developing cortical circuits support a wide range of dynamics that can be effectively controlled

In this study, I used average controllability measure to investigate the dynamical repertoire of developing cortical circuits in WT (n = 17), HET (n = 15), and KO (n = 15) cell cultures. An overall increase in average controllability was expected to reflect the increase in functional connectivity over development (Tang et al., 2017) and to be constrained by the edge weight distributions (Pasqualetti & Zampieri, 2014).

I found that WT cultures were on average more controllable than KO, but not HET cultures. Indeed, I also observed no significant overall differences between HET and KO cultures. Together, these results suggest that WT networks are able to traverse

Compared genotypes	Difference[a]	Confidence interval	p
HET vs KO	2.63	-0.52 to 5.79	0.13
HET vs WT	−3.94	-0.78 to 2.37	1.00
KO vs WT	−6.57	-3.42 to -0.26	0.03 *

Table 4.1: Table summarising statistical differences in average controllability between genotypes (Friedman repeated measures test with multiple comparisons). * $p < 0.05$
[a] Difference in average column rank.

vast surfaces of the energy landscape, and that this ability is disrupted to an extent proportional to Mecp2 deficiency (Tab. 4.1).

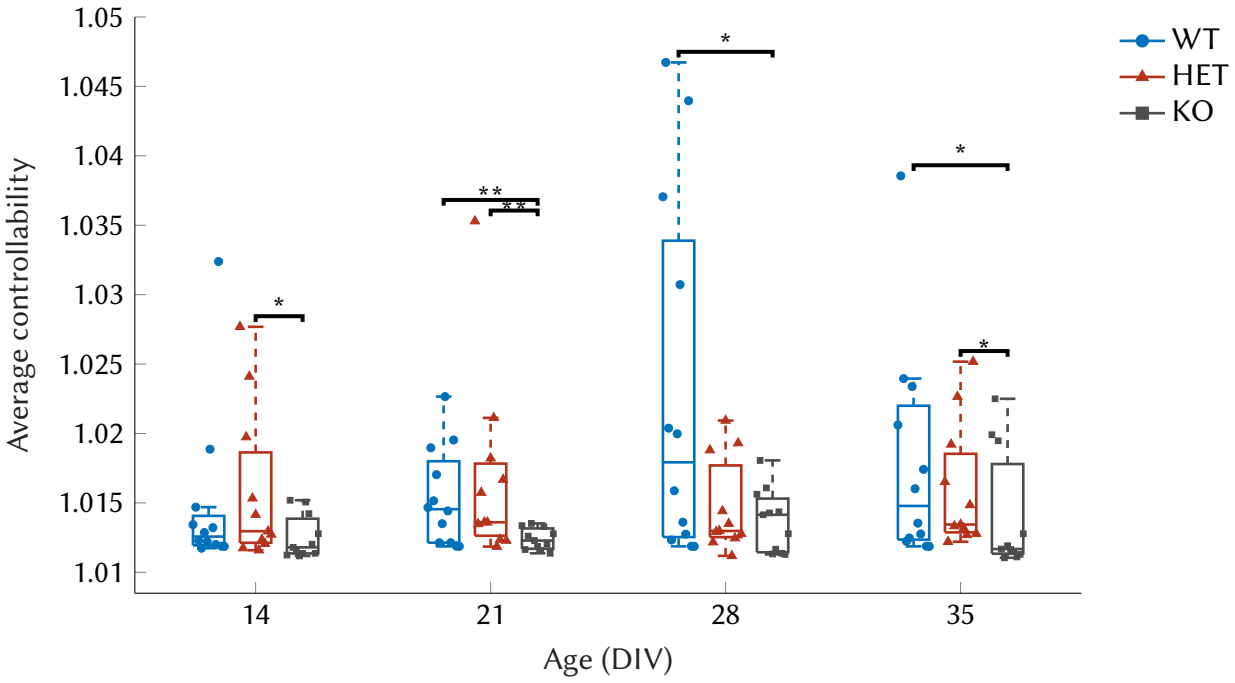


Figure 4.3: Box plot (median, interquartile range, minimum and maximum) of age and genotype effects on average controllability. DIV – days in vitro. Genotypes: WT – wild type, HET – heterozygous, KO – Mecp2 knockout. * $p < 0.05$, ** $p < 0.01$.

At DIV14, all genotypes were controllable to a similar extent, with only significant difference between HET and KO. Starting from DIV21, WT cultures had significantly higher average controllability values than KOs. During development, there were no differences between WT and HET cultures. Interestingly, the controllability of WT and KO cultures increased with time, but HET cultures remained at a steady level throughout the study (Fig. 4.3, Tab. 4.2). All significant results reached large effect sizes (1.30 to 1.39), yielding statistical power above the specified threshold of 0.8.

Compared genotypes	Age (DIV)	U	z	p	Effect size [a]
WT vs KO	14	106.00	−1.57	0.058	n/a
	21	93.00	−2.37	0.009 **	1.39
	28	102.00	−1.82	0.035 *	1.32
	35	98.00	−2.06	0.019 *	1.35
WT vs HET	14	142.00	0.65	0.741	n/a
	21	136.00	0.28	0.609	n/a
	28	112.00	−1.20	0.115	n/a
	35	133.00	0.09	0.537	n/a
HET vs KO	14	96.00	−1.97	0.024 *	1.30
	21	86.00	−2.63	0.004 **	1.38
	28	118.50	−0.49	0.311	n/a
	35	95.50	−2.00	0.023 *	1.30

Table 4.2: Table summarising the statistical results of pairwise comparisons (two-tailed Mann-Whitney U test) between genotypes in respect to average controllability. * $p < 0.05$, ** $p < 0.01$.
[a] Effect sizes reported only for statistically significant differences.

These findings suggest that already at early developmental stages, cortical circuits become increasingly structured to support a wide range of state transitions and that the loss of Mecp2 disrupts this dynamical capacity. However, it is unclear whether these deficits originate in an overall compression of the reachable subspace or in the energy landscape skewed towards difficult-to-reach states. To answer this question, I used modal controllability.

4.3.2 Modal controllability

Modal controllability quantifies the extent to which a node can drive the network towards difficult-to-reach states (those distant in the state space). The energy-based interpretation of this measure defines the *difficult-to-reach* states as those that require substantial input energy (Fig. 4.2). I outline the complementary mechanistic explanation below.

We can describe the network's activity as a weighted sum of specific patterns of activity called *modes* (eigenvectors) that define the directions of the temporal evolution of the system's dynamics in the state space. Corresponding weights (eigenvalues) specify the amplification or attenuation of the system's dynamics in these directions. Provided inputs that control these modes, one can steer the system towards any possible state. An extension of the Popov–Belevitch–Hautus (PBH) test of controllability states that if the weight corresponding to a given mode is small, then this mode is poorly controllable from a given input node. Modal controllability then quantifies the node's ability to control all modes of the network. Generating a precisely structured network activity pattern required for a difficult task requires robust control over multiple modes. Therefore, nodes with high modal controllability are better suited to drive the network towards difficult-to-reach states.

Mecp2-deficient networks are better at reaching difficult-to-reach states

Based on previous reports of negative correlation between average and modal controllability (Gu et al., 2015) and that modal controllability decreases over healthy brain development (Tang et al., 2017), I predicted that this measure in cortical circuits would decrease over time.

Compared genotypes	Difference[a]	Confidence interval	p
HET vs KO	−5.22	-2.07 to 1.08	0.35
HET vs WT	−1.66	1.49 to 4.64	0.77
KO vs WT	0.41	-3.56 to 6.71	0.02 *

Table 4.3: Table summarising statistical differences in modal controllability between genotypes (Friedman repeated measures test with multiple comparisons). * $p < 0.05$
[a] Difference in average column rank.

To test this hypothesis, I have applied modal controllability to functional networks forming in WT ($n = 17$), HET ($n = 15$), and KO ($n = 15$) developing cortical cultures. I found that KO cultures were better suited for transitioning towards

difficult-to-reach states than WT networks. However, I observed no overall difference between KO and HET, or WT and HET cultures (Tab. ??). These results suggest that the dynamical repertoire of Mecp2-null networks is skewed towards difficult-to-reach states.

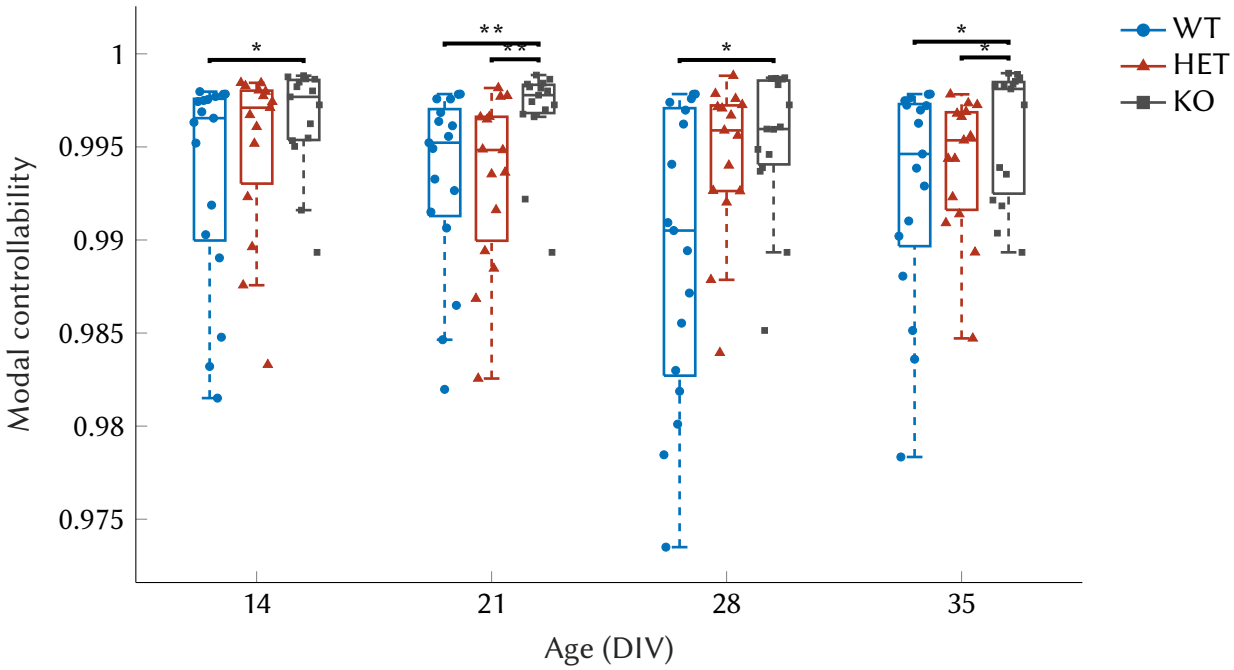


Figure 4.4: Box plot (median, interquartile range, minimum and maximum) of age and genotype effects on modal controllability. DIV – days in vitro. Genotypes: WT – wild type, HET – heterozygous, KO – Mecp2 knockout. * $p < 0.05$, ** $p < 0.01$.

Next, I compared the three genotypes at different developmental stages (Tab. 4.4). I observed that WT cultures were different from KO cultures already at DIV14 and that this continued throughout the study. HET cultures at early developmental age (DIV14) were similar to KO cultures, but during development, they developed lower modal controllability, albeit with some variability (Fig. 4.4). The most notable differences were noted at DIV21, suggesting that significant refinement of functional architecture might happen at this developmental stage. There were no clear developmental trajectories in modal controllability. The statistically significant results have reached the 0.8 statistical power threshold.

Compared genotypes	Age (DIV)	U	z	p	Effect size [a]
WT vs KO	14	230.00	−1.89	0.030 *	1.17
	21	203.00	−2.91	0.002 **	1.27
	28	223.00	−2.15	0.016 *	1.20
	35	223.00	−2.25	0.016 *	1.20
WT vs HET	14	276.00	1.10	0.863	n/a
	21	246.00	−0.04	0.485	n/a
	28	293.00	1.74	0.959	n/a
	35	248.00	0.04	0.515	n/a
HET vs KO	14	205.00	−1.12	0.131	n/a
	21	161.00	−2.94	0.002 **	1.35
	28	214.50	−0.73	0.234	n/a
	35	191.00	−1.70	0.044 *	1.22

Table 4.4: Table summarising the statistical results of pairwise comparisons (two-tailed Mann-Whitney U test) between genotypes in respect to modal controllability. * $p < 0.05$, ** $p < 0.01$.
[a] Effect sizes reported only for statistically significant differences.

4.4 Global controllability metrics

Despite providing intuition for controller placement, nodal controllability metrics do not characterise the control profile of the entire network. In other words, they do not answer how complex (energetically costly) it is to drive the network between states irrespective of the choice of inputs. This issue is relevant in the scope of cortical circuits in RTT because it will answer how feasible therapeutic modulation will be. Global controllability metrics were developed to address this issue and provide a more general description of network controllability. In this study, I have focused on controllability ellipsoid (Bainum & Xing, 1996).

4.4.1 Controllability ellipsoid

Controllability (energy) ellipsoid is a convenient geometrical representation of the system's controllability matrix (Gramian) (Summers et al., 2016). It encloses the controllable subspace and specifies the directions in the state space along which the system can be driven the furthest with a unit of input energy to any node (Fig. 4.5).

Because, to my knowledge, this measure has not yet been applied to neuro-

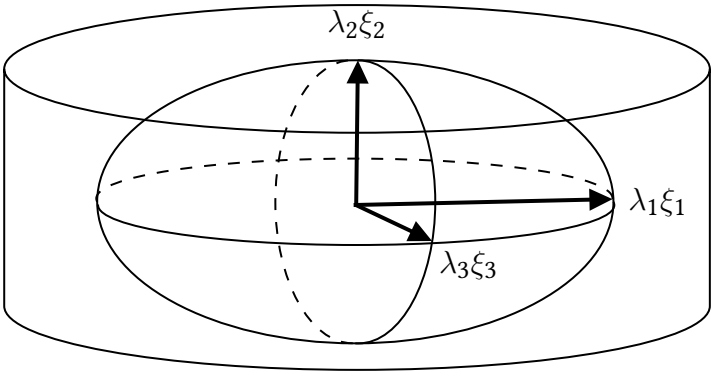


Figure 4.5: Schematic representation of a controllability ellipsoid within a three-dimensional state space (cylinder). The principal axes are given by directions that are hierarchically ordered in terms of controllability. Any state that is reachable with a unit or less input energy (given system's architecture) will lie somewhere within the controllability ellipsoid. In neuronal networks, the state space spans more than three dimensions (Chapter 3).

science before, I hypothesised that the volume of the controllability ellipsoid will mirror the changes in average controllability and that it will increase together with functional connectivity in the developing circuits. If found to be true, this would likely mean that the dysfunction in Mecp2-null networks hinders potential therapeutic modulation within reasonable clinical constraints.

Analysis of controllable subspace reveals feasibility of therapeutically controlling network dynamics in RTT

Compared genotypes	Difference[a]	Confidence interval	p
HET vs KO	−1.07	2.08 to 5.23	0.34
HET vs WT	−3.41	-0.26 to 2.89	0.77
KO vs WT	−5.49	-2.34 to 0.81	0.23

Table 4.5: Table summarising statistical differences in the volume of controllability ellipsoid between genotypes (Friedman repeated measures test with multiple comparisons).
[a] Difference in average column rank.

I used the volume of the controllability ellipsoid $V(\epsilon)$ to quantify the overall controllability of WT (n = 17), HET (n = 15), and KO (n = 15) neuronal networks. I found no overall differences between different genotypes (Tab. 4.5), suggesting similarities in controllable subspaces. This finding could support the notion that RTT is characterised by a mixture of function (breathing, swallowing) and dysfunction (cognitive and motor disability), which might require maintaining sufficiently controllable

neural dynamics. This could also be beneficial for potential therapeutic control.

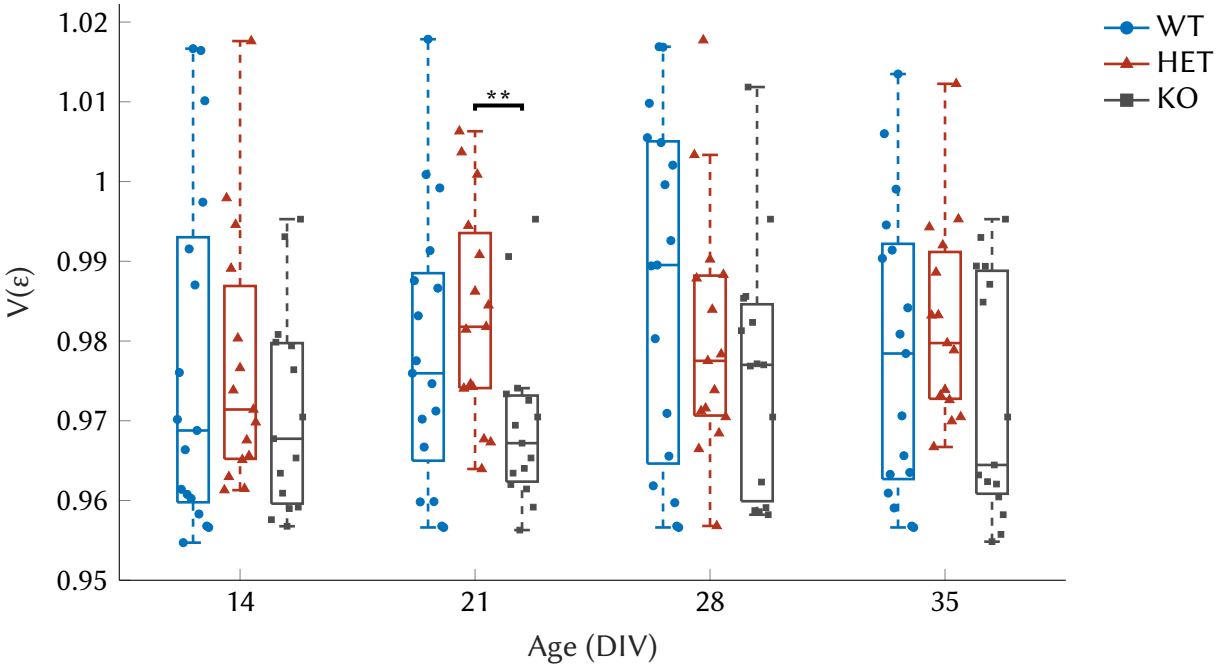


Figure 4.6: Box plot (median, interquartile range, minimum and maximum) of age and genotype effects on the volume of controllability ellipsoid. DIV – days in vitro. Genotypes: WT – wild type, HET – heterozygous, KO – Mecp2 knockout. ** $p < 0.01$.

The only significant difference across genotypes and developmental ages has been observed between KO and HET cultures at DIV21, where the volume of the controllable subspace of Mecp2-null networks was lower (Fig. 4.6, Tab. 4.6). This could be potentially explained by the fact that WT cultures reached possible controllability optimum, HET cultures were slightly developmentally delayed, and connectivity refinements led to apparent overcompensation in respect to KO cultures. This finding also parallels notable changes in other metrics, also occurring at DIV21. However, the statistical power of this finding was 0.59, which was below the power threshold of 0.8. An increase in sample sizes from $n = 15$ to $n = 26$ would be required to confirm this finding.

Compared genotypes	Age (DIV)	U	z	p	Effect size [a]
WT vs KO	14	287.00	0.22	0.821	n/a
	21	322.00	1.55	0.122	n/a
	28	320.00	1.47	0.141	n/a
	35	307.00	0.98	0.326	n/a
WT vs HET	14	269.00	0.79	0.428	n/a
	21	272.00	0.91	0.365	n/a
	28	226.00	−0.79	0.428	n/a
	35	270.00	0.83	0.406	n/a
HET vs KO	14	205.00	−1.12	0.263	n/a
	21	161.00	−2.94	0.003 **	0.72
	28	214.50	−0.73	0.468	n/a
	35	191.00	−1.70	0.081	n/a

Table 4.6: Table summarising the statistical results of pairwise comparisons (two-tailed Mann-Whitney U test) between genotypes in respect to the volume of controllability ellipsoid. * $p < 0.05$, ** $p < 0.01$.
[a] Effect sizes reported only for statistically significant differences.

5

Discussion

5.1 Findings summary

In this investigation, I

1. Found a genotype-dependent reduction in the effective dimensionality of network dynamics in developing cortical circuits,
2. Found that wild type cultures exhibit an increased capability to traverse larger surfaces of the energy landscape than other genotypes,
3. Found that despite the limited state-switching flexibility and diminished dynamical repertoire, Mecp2 deficient cultures had controllable subspaces equal to other genotypes,
4. Shown the potential for the application of the analyses of network dynamics in the study of in vitro cortical networks in health and in a mouse model of Rett syndrome.

5.2 Validity of approach and reliability of results

5.2.1 Cortical cultures for modelling circuit development

In vitro cultures were chosen because they allow control over the experimental conditions and the flexibility of approaches. Using MEAs, it is possible to probe the activity of developing networks over several weeks with a temporal resolution needed to infer functional connectivity and characterise neuronal dynamics. Dissociated neurons offer a simplified model of neurodevelopment lacking factors such as the effects of sensory experience, the role of glia, the mechanical constraints of the skull environment, and the various neuromodulatory signals that in situ circuits receive. Therefore, one must be cautious when extrapolating any results to in vivo conditions. However, by controlling for the factors above, one can infer mechanistic explanations and study the fundamental principles governing network development in health and disease models.

5.2.2 Microelectrode array recordings and spike detection

The large size and spacing of the electrodes on the array means that each of them might record electrical activity from more than a single neuron. Conversely, it is unlikely that more than one electrode records activity from the same neuron. Therefore, this technique does not confer us with single-cell resolution. However, this also means that each electrode can be conveniently abstracted as a node within a network, with relationships between nodes calculated from multi-unit activity (MUA). Indeed, population recordings have revealed that even sparse sampling of network MUA is insightful (Downes et al., 2012; Schröter et al., 2015).

Possible errors in spike detection depend on the choice of spike detection method and parameters. In low signal-to-noise ratio conditions, threshold-based methods perform poorly, necessitating the use of template-based methods. The obvious challenge is the choice of appropriate templates that represent biophysically accurate extracellular spike waveforms. I have mitigated this problem by using a novel spike detection method with data-driven electrode-specific templates. Another problem is the choice of cost parameter. Because there was no ground truth such as paired patch-clamp/juxtacellular and MEA recording, I instead used recordings where TTX was used to validate cost parameters. I performed further validation of the spike detection using synthetic data (Appendix A).

5.2.3 Inferring functional connectivity from correlated spontaneous neuronal activity

We can use spike trains extracted from the recorded voltage traces to infer functional connectivity in developing networks using time-series analysis. Such approaches are helpful in that they reduce the influence of noise through only considering spike times and not correlated noise or measuring quiescent periods as, for example, Pearson's correlation. I chose STTC as a method for inferring functional connectivity because it is independent of firing rates, robust to quiescent periods, mathematically tractable and interpretable (bounded by -1 and 1).

Possible errors in determining the functional connections might arise if inferred relationships are byproducts of simpler phenomena, such as common spatiotemporal tuning (Elsayed & Cunningham, 2017), or correlated noise. This necessitates the normalization to random correlation levels to avoid false positives. To this end, I have adapted and applied probabilistic thresholding (Lancichinetti & Fortunato, 2012, Appendix B).

5.2.4 Limitations of dimensionality and controllability analyses

Calculations of the effective dimensionality rely on the duration of the time bin within which the spiking activity is summed. With different durations, the focus will shift from short to long timescale. In this study, I have chosen the short-timescale correlation time bin of 25 ms (El-Gaby et al., 2021). Because neuronal spiking activity often exhibits Poisson firing rate distributions (Mazzoni et al., 2007), I have applied a variance stabilising transformation by taking the square root of summed activity in each time bin (Yu et al., 2009).

The central assumption in calculating the controllability metrics is that the underlying neuronal dynamics are linear (or linearised around a fixed point). Although this might not always be the case, this assumption is commonly accepted in the study of networked systems, especially where structural interpretations of dynamics are preferred (Gu et al., 2015). Moreover, structural controllability metrics assume a rigid network architecture (Pasqualetti & Zampieri, 2014). Although there exists ample evidence that functional connectivity patterns are highly correlated (Rosenbaum et al., 2017) and constrained (Kordovan, 2019), by the underlying neuroanatomy, one must always exercise caution when extrapolating the findings from functional to structural networks. A recent study investigated the relationships between controllability calculated from functional (fMRI) and structural (tractography) networks, suggesting that

further research into this area is needed (Gu et al., 2021).

Finally, it is important to recognise that the dimensionality, controllability, and functional connectivity analyses were obtained using *spontaneous* neuronal activity. This fact necessitates the question to what extent the captured dynamics represent physiological state space. We could test this experimentally by perturbing the neuronal activity out of equilibrium, thus simulating the potential drive that developing cortical circuits receive from other brain areas (Mazzucato et al., 2016; Das & Fiete, 2020).

With these limitations in mind, I will interpret the obtained results below.

5.3 Interpretation of main findings

5.3.1 Mecp2 deficiency leads to deficits in the emergence of functional subcommunities in cortical networks

I found that already at DIV14, WT cultures show decreased variability in neuronal activity that remains constant over four weeks of early development (DIV14-DIV35). This finding suggests an increased synchronisation of network dynamics achieved through a decreased number of independent spiking patterns. Functionally, network-wide coordination could support low-dimensional, interpretable neural code suitable for global computations.

Starting from DIV14, I observed a decrease in the effective rank in HET cultures, suggesting that they were developmentally delayed in respect to WT cultures but later achieved similar levels of synchronisation. Together with high-dimensional dynamics found in KO networks, these results point to a possible link between Mecp2 expression and capacity for subcommunity synchronisation. Interestingly, these findings were robust to the growth of the networks during development as quantified by the relative effective rank. This could further support the notion that connectivity patterns regulate the complexity of network dynamics more than network size.

Network function requires a balance between dynamical flexibility reflected in a large enough pattern repertoire to support local processing and network synchronisation within and between subcommunities that enables efficient global processing. The disruption of these mechanisms in KO cultures could be justified either by the lack of the overall capacity for network synchronisation or substantial delay in development that could not be observed in the time window of the experiment. To differentiate between these possibilities, we could extend the study duration or focus on longer timescales in the calculations of effective rank.

5.3.2 Developmental changes in controllability metrics reveal network wiring in a manner optimised for control

During network development, I expected average controllability to increase together with increases in functional connectivity. This relationship could be explained by the correlation between average controllability and node strength (the sum of node's weighted edges) (Gu et al., 2015). In the scope of Rett syndrome, the deficits in average controllability could be explained by the cortical hypoconnectivity observed anatomically and functionally in the mouse model of the disease (Shepherd & Katz, 2011; Durand et al., 2012; Sceniak et al., 2016).

I found developmental changes in average and modal controllability to be limited to WT cultures. These cultures had a higher average and lower modal controllability than KO cultures. Heterozygous cultures displayed an intermediate range of values, suggesting a possible dose-effect relationship between controllability profiles and the amount of expressed Mecp2. Additionally, Pasqualetti & Zampieri, 2014 have found that isotropic networks are more challenging to control than anisotropic networks. Isotropic edge weight distribution means that each connection has the same strength, whereas anisotropic networks are heterogeneous. If there is a single strong link from an input node to the controlled node, the energy will be funnelled through this connection, providing effective control. Conversely, if there are many weak connections, the energy will be widely distributed, albeit with the loss of directionality and magnitude.

Biologically, these findings can be explained by the stochastic wiring during early cortical circuit development leading to densely but weakly connected networks. Throughout development, there occurs the activity-dependent refinement of connectivity promoting the maintenance of meaningful, and the pruning of weak connections (Richter & Gjorgjieva, 2017). This process relies, at least in part, on the functional switch of GABA effects from excitatory to inhibitory and the maturation of N-methyl-D-aspartate receptor (NMDAR) subunit composition, both of which are disrupted in RTT (Tang et al., 2016; Mierau et al., 2016). Moreover, differential effects of these alterations on the synapses onto pyramidal and parvalbumin-positive (PV) neurons could lead to excitatory-inhibitory imbalances.

Therefore, both hypoconnectivity and edge weight distributions are altered in RTT, and are plausible explanations of the observed deficits. We could investigate this further by comparing control profiles of pyramidal and PV neurons, for example, using optogenetic approaches or fluorescent markers to distinguish between these cell populations.

In KO cultures, increased modal controllability together with decreased aver-

age controllability suggests that Mecp2-deficient networks have a more limited dynamical repertoire than WT networks and that this repertoire is more restricted to difficult-to-reach states. Altered timecourse of the specialization of these networks could lead to disruptions in flexibility (Harlalka et al., 2019) in switching states and more energy needed to control basic network function. These results are in line with the intellectual disability affecting almost all, and autism spectrum disorder symptoms observed in around 50% of the RTT patients.

Similarities in the volumes of controllable subspaces between KO and WT networks suggest that not all aspects of circuit function are disrupted by Mecp2 deficiency. This could reflect that some aspects of neuronal function are preserved in cortical circuits in RTT, leading to a balance of function and dysfunction observed in human disease.

5.3.3 Unifying dimensionality and control theory approaches in the study of developing cortical networks

So far, I analysed dimensionality and controllability independently. Below, I outline possible ways of combining the snapshots obtained by these methods into a larger picture of network dynamics.

When neuronal dynamics are coupled, the patterns of activity generated by the network can be described by relatively few latent modes (dimensions in state space). Intuitively, exciting a lower number of modes requires fewer driver nodes and less input energy, and therefore such a network will be easier to control on average. Conversely, steering network dynamics towards difficult-to-reach states requires more flexibility conferred by a larger number of degrees of freedom (dimensions in state space) that network activity can independently explore. Therefore, findings of low dimensionality in neural data can be linked to higher average controllability, and findings of many independent patterns of activity support higher modal controllability.

Another link can be found at the level of local connectivity motifs. In this study, the effective rank was operationalized to quantify the capacity of a network to form functional subcommunities with shared patterns of activity. Mathematically, the average controllability of the whole network is equal to the sum of controllability of functional subcommunities weighted by the strength of the edges between them. Therefore, the local subgroup synchronization reflected in the decrease in effective rank will simultaneously increase the average controllability values.

Finally, at the whole-brain scale, population coding relies on combining activities from thousands of neurons into commands that can be decoded using an in-

terpretably low number of latent dimensions. If different functional subcommunities (e.g. brain areas) are synchronized, the computations will require less cognitive effort because the easily reachable subspace will be larger. Disease processes, such as RTT might impair this ability, leading to deficits in network control. The abstract network analyses presented here apply to other biological networks, such as larger-scale brain networks, and could offer a potential avenue for future research of network dynamics in RTT across scales.

5.4 Future directions

5.4.1 Characterising cell type-specific contributions to network dynamics

Circuit pathology in RTT is thought to originate from a disruption in E-I balance stemming from deficits intrinsic to particular neurons or neuronal classes (Tang et al., 2016). The diminished excitatory drive in the network and disruptions in synchronisation brought on principally by parvalbumin-positive (PV) interneuron dysfunction (Mierau et al., 2016) could be rescued by the enhancement of activity. Genetically, it has been achieved in a study whereby re-introduction of Mecp2 at later developmental stages was able to reinstate some behavioural phenotype (Guy et al., 2007). Pharmacologically, a recent report has improved the activity-dependent transcription of some critical genes, neuronal morphology and function in vitro, and rescued motor skills, spatial memory, and delayed the progress of the disease in a mouse model of RTT (Scaramuzza et al., 2021).

However, the modulation of network function needs to be tightly controlled, as it might have differential effects on excitatory and inhibitory cells. Therefore, functionally dissecting these classes of neurons would be an important future step in network analyses. Experimentally, this could be done either anatomically using fluorescent markers, functionally using optogenetics, or with a combination of both. The putative inhibitory cells could then be characterised based on spike waveforms, spiking frequency and firing regularity.

5.4.2 Expanding control analysis to directed and signed networks

Undirected networks provide a good intuition about network function and are commonly used in network neuroscience (Bullmore & Sporns, 2009). However, the

flow of information in neuronal circuits is often directional. With a single-cell resolution, we could infer putative excitatory and inhibitory connections from spike trains using sophisticated computational methods. Some possibilities include generalized linear models with cross-correlograms (Kobayashi et al., 2019) or approaches from information theory, including Granger causality or transfer entropy (Wibral et al., 2014). Such a framework could offer insights into directional control in terms of distinguishing source, sinks and conduit nodes (Campbell et al., 2015), the minimum number of controllers (Liu et al., 2011) and be used to cross-validate the characterisation of neuronal subtypes (Ren et al., 2020).

In the current investigation, the challenge in inferring directed connectivity was that most of the algorithms outlined above require larger volumes of data than collected during 12-minute recordings. Extending duration of MEA recordings to 30–60 minutes would be needed and would require tighter pH control during experiments. Another issue is that some of these methods are sensitive to multi-unit activity (i.e. unsorted spikes). In this study, I did not attempt spike sorting because it was not necessary for the characterisation of neuronal network dynamics (Trautmann et al., 2019) and because I have observed overlapping spike waveforms (Wagenaar et al., 2006) which could have affected the classification. Spike sorting could be performed in future studies when ground-truth such as paired MEA/juxtacellular or MEA/patch-clamp recordings becomes available.

5.5 Implications of research

In this study, I investigated the mechanisms underlying the emergence of network dynamics in developing cortical circuits using control theory and dimensionality analysis. Using electrophysiological recordings from MEAs, I probed the spontaneous network activity over several weeks of development. I used the detected spike trains to quantify effective dimensionality and to infer functional connectivity. By modelling the relationships between network structure and function using control theory, I characterised dynamical network repertoire and its controllability.

Studying network dynamics offers a framework for the study of neuronal circuit development in health and disease. With functional networks forming in vitro, it is possible to explore the circuit-level effects in disease models. I have shown the utility of these approaches in studying RTT. I found that Mecp-deficiency disrupts the emergence of functional subcommunities in cortical circuits at early developmental stages, distorts network controllability by constraining state transitions and shifting the balance in a dynamical repertoire towards difficult-to-reach states and that it is feasible to modulate network activity based on the volume of controllable subspace.

With this information, it is possible to classify the extent of function and dysfunction in health and RTT, identify the nodes with the most influence over the network function, and design strategies to control these nodes for mechanism-based therapeutic modulation that could be tested experimentally in the future.

Bibliography

Altman, N. & Krzywinski, M. (2016). Analyzing outliers: influential or nuisance? Nat. Methods *13*.4, 281–282.

Amir, R.E., Van den Veyver, I.B., Wan, M., Tran, C.Q., Francke, U. & Zoghbi, H.Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nat Genet *23*.2, 185–188.

Bainum, P.M. & Xing, G.Q. (1996). Digital Control Systems. Implementation and Computational techniques (San Diego, California: Academic Press, Inc.).

Baj, G., Patrizio, A., Montalbano, A., Sciancalepore, M. & Tongiorgi, E. (2014). Developmental and maintenance defects in Rett syndrome neurons identified by a new mouse staging system in vitro. Front. Cell. Neurosci. *8*.

Berenyi, A., Belluscio, M., Mao, D. & Buzsaki, G. (2012). Closed-Loop Control of Epilepsy by Transcranial Electrical Stimulation. Science *337*.6095, 735–737.

Bullmore, E. & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci *10*.3, 186–198.

Campbell, C., Ruths, J., Ruths, D., Shea, K. & Albert, R. (2015). Topological constraints on network control profiles. Sci. Rep. *5*.1, 1–9.

Castro, J., Garcia, R.I., Kwok, S., Banerjee, A., Petravicz, J., Woodson, J., Mellios, N., Tropea, D. & Sur, M. (2014). Functional recovery with recombinant human IGF1 treatment in a mouse model of Rett Syndrome. PNAS USA *111*.27, 9941–9946.

Charlesworth, P., Cotterill, E., Morton, A., Grant, S. & Eglen, S.J. (2015). Quantitative differences in developmental profiles of spontaneous activity in cortical and hippocampal cultures. Neural Dev *10*.1, 1.

Chen, R.Z., Akbarian, S., Tudor, M. & Jaenisch, R. (2001). Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. Nat Genet *27*.3, 327–331.

Chiappalone, M., Pasquale, V. & Frega, M. (2019). In Vitro Neuronal Networks. From Culturing Methods to Neuro-Technological Applications (Cham, Switzerland: Springer International Publishing).

Ching, S., Liberman, M.Y., Chemali, J.J., Westover, M.B., Kenny, J.D., Solt, K., Purdon, P.L. & Brown, E.N. (2013). Real-time Closed-loop Control in a Rodent Model of Medically Induced Coma Using Burst Suppression. Anesthesiology *119*.4, 848–860.

Conroy, R.M. (2012). What Hypotheses do "Nonparametric" Two-Group Tests Actually Test? The Stata Journal *12*.2, 182–190.

Cunningham, J.P. & Yu, B.M. (2014). Dimensionality reduction for large-scale neural recordings. Nat Neurosci *17*.11, 1500–1509.

Cutts, C.S. & Eglen, S.J. (2014). Detecting Pairwise Correlations in Spike Trains: An Objective Comparison of Methods and Application to the Study of Retinal Waves. J. Neurosci. *34*.43, 14288–14303.

D'Cruz, A.-M., Ragozzino, M.E., Mosconi, M.W., Shrestha, S., Cook, E.H. & Sweeney, J.A. (2013). Reduced behavioral flexibility in autism spectrum disorders. Neuropsychology *27*.2, 152–160.

Dani, V.S., Chang, Q., Maffei, A., Turrigiano, G.G., Jaenisch, R. & Nelson, S.B. (2005). Reduced cortical activity due to a shift in the balance between excitation and inhibition in a mouse model of Rett Syndrome. PNAS *102*.35, 12560–12565.

Dani, V.S. & Nelson, S.B. (2009). Intact Long-Term Potentiation but Reduced Connectivity between Neocortical Layer 5 Pyramidal Neurons in a Mouse Model of Rett Syndrome. J Neurosci *29*.36, 11263–11270.

Das, A. & Fiete, I.R. (2020). Systematic errors in connectivity inferred from activity in strongly recurrent networks. Nat Neurosci *23*.10, 1286–1296.

Downes, J.H., Hammond, M.W., Xydas, D., Spencer, M.C., Becerra, V.M., Warwick, K., Whalley, B.J. & Nasuto, S.J. (2012). Emergence of a Small-World Functional Network in Cultured Neurons. PLoS Comput Biol *8*.5. O. Sporns, ed., e1002522.

Durand, S., Patrizi, A., Quast, K.B., Hachigian, L., Pavlyuk, R., Saxena, A., Carninci, P., Hensch, T.K. & Fagiolini, M. (2012). NMDA Receptor Regulation Prevents Regression of Visual Cortical Function in the Absence of Mecp2. Neuron *76*.6, 1078–1090.

Einspieler, C. & Marschik, P.B. (2019). Regression in Rett syndrome: Developmental pathways to its onset. Neurosci Biobehav Rev *98*, 320–332.

Elsayed, G.F. & Cunningham, J.P. (2017). Structure in neural population recordings: an expected byproduct of simpler phenomena? Nat Neurosci *20*.9, 1310–1318.

Faul, F., Erdfelder, E., Buchner, A. & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. Behav. Res. Methods *41*.4, 1149–1160.

Gabbiani, F. & Metzner, W. (1999). Encoding and processing of sensory information. J. Exp. Biol. 202, 1267–1279.

El-Gaby, M., Reeve, H.M., Lopes-dos-Santos, V., Campo-Urriza, N., Perestenko, P.V., Morley, A., Strickland, L.A.M., Lukács, I.P., Paulsen, O. & Dupret, D. (2021). An emergent neural coactivity code for dynamic memory. Nat Neurosci.

Gao, P., Trautmann, E., Yu, B., Santhanam, G., Ryu, S., Shenoy, K. & Ganguli, S. (2017). *A theory of multineuronal dimensionality, dynamics and measurement*. preprint. Neuroscience.

Goffin, D. & Zhou, Z. (2012). The neural circuit basis of Rett syndrome. Front. Biol. *7*.5, 428–435.

Gu, S., Fotiadis, P., Parkes, L., Xia, C.H., Gur, R.C., Gur, R.E., Roalf, D.R., Satterthwaite, T.D. & Bassett, D.S. (2021). Network controllability mediates the relationship between rigid

structure and flexible dynamics. bioRxiv. eprint: https://www.biorxiv.org/content/early/2021/04/26/2021.04.23.441156.full.pdf.

Gu, S. et al. (2015). Controllability of structural brain networks. Nat Commun *6*.1, 8414.

Guy, J., Gan, J., Selfridge, J., Cobb, S. & Bird, A. (2007). Reversal of Neurological Defects in a Mouse Model of Rett Syndrome. Science *315*.5815, 1143–1147.

Guy, J., Hendrich, B., Holmes, M., Martin, J.E. & Bird, A. (2001). A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. Nat Genet *27*.3, 322–326.

Hao, S. et al. (2015). Forniceal deep brain stimulation rescues hippocampal memory in Rett syndrome mice. Nature *526*.7573, 430–434.

Harlalka, V., Bapi, R.S., Vinod, P.K. & Roy, D. (2019). Atypical Flexibility in Dynamic Functional Connectivity Quantifies the Severity in Autism Spectrum Disorder. Front. Hum. Neurosci. *13*, 6.

Heckman, L.D., Chahrour, M.H. & Zoghbi, H.Y. (2014). Rett-causing mutations reveal two domains critical for MeCP2 function and for toxicity in MECP2 duplication syndrome mice. eLife *3*, e02676.

Hedges, L.V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. J. Educ. Stat. *6*.2, 107–128.

Hentschke, H. (2021). hhentschke/measures-of-effect-size-toolbox (GitHub).

Jobst, B.C. et al. (2017). Brain-responsive neurostimulation in patients with medically intractable seizures arising from eloquent and other neocortical areas. Epilepsia *58*.6, 1005–1014.

Johnson, L.A., Nebeck, S.D., Muralidharan, A., Johnson, M.D., Baker, K.B. & Vitek, J.L. (2016). Closed-Loop Deep Brain Stimulation Effects on Parkinsonian Motor Symptoms in a Non-Human Primate – Is Beta Enough? Brain Stimul. *9*.6, 892–896.

Ju, H. & Bassett, D.S. (2020). Dynamic representations in networked neural systems. Nat Neurosci *23*.8, 908–917.

Kao, T.-C. & Hennequin, G. (2019). Neuroscience out of control: control-theoretic perspectives on neural circuit dynamics. Curr. Opin. Neurobiol. *58*, 122–129.

Karrer, T.M., Kim, J.Z., Stiso, J., Kahn, A.E., Pasqualetti, F., Habel, U. & Bassett, D.S. (2020). A practical guide to methodological considerations in the controllability of structural brain networks. J. Neural Eng. *17*.2, 026031.

Kobayashi, R., Kurita, S., Kurth, A., Kitano, K., Mizuseki, K., Diesmann, M., Richmond, B.J. & Shinomoto, S. (2019). Reconstructing neuronal circuitry from parallel spike trains. Nat Commun *10*.1, 4468.

Kordovan, M. (2019). "Constraints on neural activity imposed by the structure of neurons and networks". PhD thesis.

Kron, M., Howell, C.J., Adams, I.T., Ransbottom, M., Christian, D., Ogier, M. & Katz, D.M. (2012). Brain Activity Mapping in Mecp2 Mutant Mice Reveals Functional Deficits in Forebrain Circuits, Including Key Nodes in the Default Mode Network, that are Reversed with Ketamine Treatment. J. Neurosci. *32*.40, 13860–13872.

Lancichinetti, A. & Fortunato, S. (2012). Consensus clustering in complex networks. Sci Rep *2*.1, 336.

Lieb, F., Stark, H.-G. & Thielemann, C. (2017). A stationary wavelet transform and a time-frequency based spike detection algorithm for extracellular recorded data. J. Neural Eng. *14*.3, 036013.

Liu, Y.-Y., Slotine, J.-J. & Barabási, A.-L. (2011). Controllability of complex networks. Nature *473*.7346, 167–173.

Marchetto, M.C., Carromeu, C., Acab, A., Yu, D., Yeo, G.W., Mu, Y., Chen, G., Gage, F.H. & Muotri, A.R. (2010). A Model for Neural Development and Treatment of Rett Syndrome Using Human Induced Pluripotent Stem Cells. Cell *143*.4, 527–539.

Mazzoni, A., Broccard, F.D., Garcia-Perez, E., Bonifazi, P., Ruaro, M.E. & Torre, V. (2007). On the Dynamics of the Spontaneous Activity in Neuronal Networks. PLoS ONE *2*.5. O. Sporns, ed., e439.

Mazzucato, L., Fontanini, A. & La Camera, G. (2016). Stimuli Reduce the Dimensionality of Cortical Activity. Front. Syst. Neurosci. *10*, 11.

Mierau, S.B., Patrizi, A., Hensch, T.K. & Fagiolini, M. (2016). Cell-Specific Regulation of N-Methyl-D-Aspartate Receptor Maturation by Mecp2 in Cortical Circuits. Biol. Psychiatry *79*.9, 746–754.

Moretti, P. (2006). Learning and Memory and Synaptic Plasticity Are Impaired in a Mouse Model of Rett Syndrome. J Neurosci *26*.1, 319–327.

Nenadic, Z. & Burdick, J.W. (2004). Spike detection using the continuous wavelet transform. IEEE Trans. Biomed. Eng. *52*.1, 74–87.

Pasqualetti, F. & Zampieri, S. (2014). On the Controllability of Isotropic and Anisotropic Networks. 53rd IEEE CDC, 6.

Perin, R., Berger, T.K. & Markram, H. (2011). A synaptic organizing principle for cortical neuronal groups. PNAS *108*.13, 5419–5424.

Poli, D., Pastore, V.P. & Massobrio, P. (2015). Functional connectivity in in vitro neuronal assemblies. Front. Neural Circuits *9*.

Potter, S.M. & DeMarse, T.B. (2001). A new approach to neural cell culture for long-term studies. J. Neurosci. Methods *110*.1, 17–24.

Recanatesi, S., Ocker, G.K., Buice, M.A. & Shea-Brown, E. (2019). Dimensionality in recurrent spiking networks: Global trends in activity and local origins in connectivity. PLoS Comput Biol *15*.7. J. Diedrichsen, ed., e1006446.

Ren, N., Ito, S., Hafizi, H., Beggs, J.M. & Stevenson, I.H. (2020). Model-based detection of putative synaptic connections from spike recordings with latency and type constraints. J. Neurophysiol *124*.6, 1588–1604.

Richter, L.M.A. & Gjorgjieva, J. (2017). Understanding neural circuit development through theory and models. bioRxiv. eprint: `https://www.biorxiv.org/content/early/2017/03/29/121574.full.pdf`.

Rosenbaum, R., Smith, M.A., Kohn, A., Rubin, J.E. & Doiron, B. (2017). The spatial structure of correlated neuronal variability. Nat Neurosci *20*.1, 107–114.

Rossi-Pool, R. & Romo, R. (2019). Low Dimensionality, High Robustness in Neural Population Dynamics. Neuron *103*.2, 177–179.

Roy, O. & Vetterli, M. (2007). The Effective Rank: a Measure of Effective Dimensionality. 15th IEEE EUSIPCO, 6.

Sadiq, M.T., Yu, X. & Yuan, Z. (2021). Exploiting dimensionality reduction and neural network techniques for the development of expert brain–computer interfaces. Expert Syst. Appl. *164*, 114031.

Scaramuzza, L. et al. (2021). The enhancement of activity rescues the establishment of *Mecp2* null neuronal phenotypes. EMBO Mol Med *13*.4.

Sceniak, M.P., Lang, M., Enomoto, A.C., James Howell, C., Hermes, D.J. & Katz, D.M. (2016). Mechanisms of Functional Hypoconnectivity in the Medial Prefrontal Cortex of *Mecp2* Null Mice. Cereb. Cortex *26*.5, 1938–1956.

Schröter, M.S., Charlesworth, P., Kitzbichler, M.G., Paulsen, O. & Bullmore, E.T. (2015). Emergence of Rich-Club Topology and Coordinated Dynamics in Development of Hippocampal Functional Networks In Vitro. J Neurosci *35*.14, 5459–5470.

Schröter, M., Paulsen, O. & Bullmore, E.T. (2017). Micro-connectomics: probing the organization of neuronal networks at the cellular scale. Nat Rev Neurosci *18*.3, 131–146.

Shepherd, G.M. & Katz, D.M. (2011). Synaptic microcircuit dysfunction in genetic models of neurodevelopmental disorders: focus on Mecp2 and Met. Curr. Opin. Neurobiol. *21*.6, 827–833.

Sherman-Gold, R. (1993). The Axon Guide for Electrophysiology and Biophysics: Laboratory Techniques (Foster City, CA: Axon Instruments).

Sit, T. (2018). "Dimensionality and Controllability of Network Dynamics in Developing Hippocampal and Cortical Circuits". Undergraduate Thesis. University of Cambridge.

Sizemore, A.E. & Bassett, D.S. (2018). Dynamic graph metrics: Tutorial, toolbox, and tale. NeuroImage *180*, 417–427.

Summers, T.H., Cortesi, F.L. & Lygeros, J. (2016). On Submodularity and Controllability in Complex Dynamical Networks. IEEE Trans. Control Netw. Syst. *3*.1, 91–101.

Tang, E. et al. (2017). Developmental increases in white matter network controllability support a growing diversity of brain dynamics. Nat Commun *8*.1, 1252.

Tang, X. et al. (2016). KCC2 rescues functional deficits in human neurons derived from patients with Rett syndrome. PNAS *113*.3, 751–756. eprint: `https://www.pnas.org/content/113/3/751.full.pdf`.

Trautmann, E.M. et al. (2019). Accurate Estimation of Neural Population Dynamics without Spike Sorting. Neuron *103*.2, 292–308.e4.

Tsu, A.P., Burish, M.J., GodLove, J. & Ganguly, K. (2015). Cortical neuroprosthetics from a clinical perspective. Neurobiol. Dis. *83*, 154–160.

Vashi, N. & Justice, M.J. (2019). Treating Rett syndrome: from mouse models to human therapies. Mamm Genome *30*.5, 90–110.

Wagenaar, D.A., Pine, J. & Potter, S.M. (2006). An extremely rich repertoire of bursting patterns during the development of cortical cultures. BMC neuroscience *7*.1, 1–18.

Walsh, R.M. & Hochedlinger, K. (2010). Modeling Rett Syndrome with Stem Cells. Cell *143*.4, 499–500.

Wibral, M., Vicente, R. & Lizier, J.T. (2014). Directed information measures in neuroscience (Springer).

Wood, L. & Shepherd, G.M. (2010). Synaptic circuit abnormalities of motor-frontal layer 2/3 pyramidal neurons in a mutant mouse model of Rett syndrome. Neurobiol. Dis. *38*.2, 281–287.

Wulffaert, J., Van Berckelaer-Onnes, I.A. & Scholte, E.M. (2009). Autistic disorder symptoms in Rett syndrome. Autism *13*.6, 567–581.

Yu, B.M., Cunningham, J.P., Santhanam, G., Ryu, S.I., Shenoy, K.V. & Sahani, M. (2009). Gaussian-Process Factor Analysis for Low-Dimensional Single-Trial Analysis of Neural Population Activity. J Neurophysiol *102*, 22.

Appendices

A

Spike detection

A.1 Improving and validating spike detection

Initial investigation of microelectrode array (MEA) recordings analysed in this study revealed they had a low signal-to-noise ratio (SNR) and were non-stationary. In such conditions, identifying spikes is challenging. Indeed, I have found that the spike detection methods previously used by our group either over- or under-counted the spikes. The threshold-based spike detection algorithm adapted from Schroeter et al. (2015) was overly sensitive to the selection of a threshold multiplier and proved unsuitable for non-stationary baseline voltage fluctuations. This resulted in a harsh trade-off between sensitivity and precision. Template-based methods focus on the stereotypical waveforms (shapes) of the action potentials and are better suited for the analysis of noisy recordings. The obvious challenge, however, is the selection of appropriate templates. The algorithm proposed by Nenadic & Burdic (2004) relies on continuous wavelet transform (CWT) with templates selected from the built-in wavelet bank from MATLAB Wavelet Toolbox. Because these wavelets poorly resemble biological action potential waveforms, this algorithm lacks the desired sensitivity.

Therefore, I considered searching the neuroscience and biomedical engineering literature for alternative spike detection algorithms.

A.1.1 Threshold-based spike detection

The simplest spike detection methods aim to characterise the baseline voltage magnitude by obtaining a measure of variability or dispersion such as standard deviation, median absolute deviation, or root-mean-square of the signal. The next step is multiplying this value by a user-predefined multiplier (determined empirically) and applying thus obtained threshold to the voltage trace. Any signal exceeding this threshold will be counted as a spike. Sometimes, an artificial refractory period is imposed. Variations of this method are ubiquitous and perform well in high signal-to-noise ratio (SNR) conditions. However, they fail whenever there is a significant amount of multi-unit activity or the SNR is low.

A.1.2 Templated-based correlation

More sophisticated spike detection methods rely on the shape (waveform) of the incoming signal. The aim is to obtain similarity scores by matching portions of the recording to the pre-selected templates. These coefficients are then thresholded to classify the waveforms most closely resembling the original templates as spikes. Several methods have been proposed utilising wavelet transforms (discrete, continuous, stationary) or template matching algorithms. The obvious challenge in this approach is obtaining reasonable templates in the first place.

Having identified a potential avenue for improving one of the existing methods, I have developed, validated, and benchmarked a novel spike detection algorithm.

A.2 Method development

The spike detection method developed in this study aims to combine the continuous wavelet transform (Nenadic & Burdic, 2004) with data-driven templates. It capitalises on both of these approaches. Tailoring the templates to specific recording channels ensures precision, whereas the sensitivity is increased by multiresolution scaling. This idea is best captured by the name: Wavelet Transform with Electrode-Restricted Spike Templates (WATERS). A comparison between generic and custom templates is shown in Fig. A.1.

In the first step, the average spike shapes are extracted from the filtered voltage traces using a simple threshold method. Next, the obtained waveforms are interpolated, smoothed, and projected onto the space of functions orthogonal to constants. If the resultant signal satisfies the requirements (it is zero-mean and square norm one), a mother wavelet is adapted (Fig. A.2).

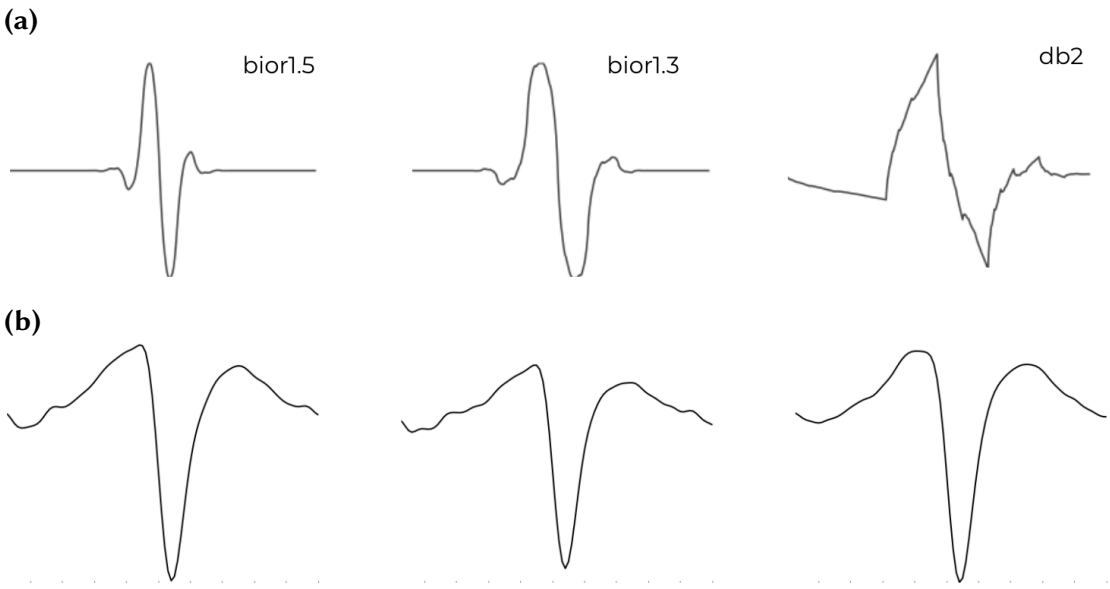


Figure A.1: Differences between wavelets from MATLAB Wavelet Toolbox and custom templates adapted from data. **(a)** Generic wavelets. Notice poor resemblance to physiological extracellular spike waveforms. **(b)** Data-driven templates adapted independently for each electrode (three examples shown).

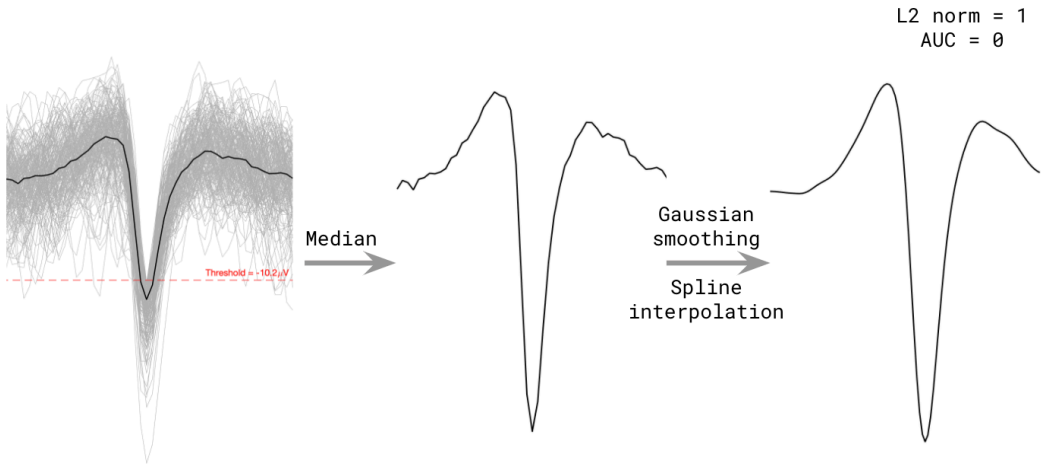


Figure A.2: Procedure of adapting a custom wavelet. First, a pre-specified number of spikes are detected using the threshold method and corresponding waveforms are extracted (left). Then, at each sample the median value is obtained, yielding the median spike waveform (middle). Finally, the template is interpolated and smoothed, and the wavelet is adapted (right).

It is then stretched in time and voltage domains across a specified number of scales, and the degree of similarity between template and signal is quantified using continuous wavelet transform (Fig. A.3). The acquired coefficients are thresholded using Bayesian hypothesis testing with a predefined cost parameter that specifies the trade-off between the cost of omission and the cost of commission of a signal segment to the spike pool. Thus obtained event times are used to extract spike waveforms,

and at this step an optional post hoc artifact removal can be employed to account for signals with biologically implausible amplitudes.

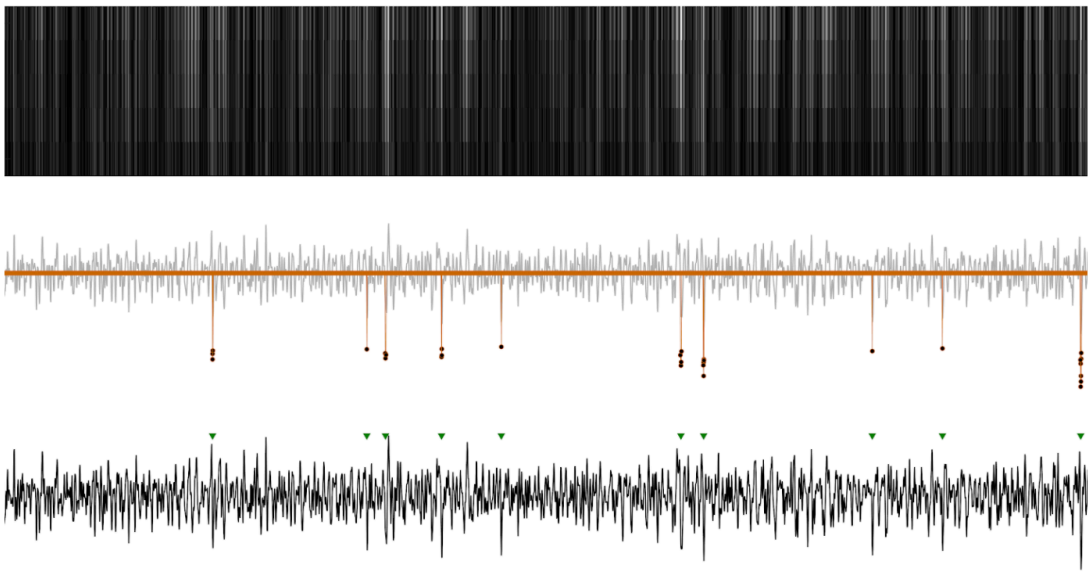


Figure A.3: The inner workings of the spike detection algorithm. **(Top)** The coefficients obtained by scaling the wavelet and correlating it with voltage trace. **(Middle)** Reconstruction of coefficients based on the cost parameter. **(Bottom)** Filtered voltage trace with spike markers over detected spikes.

Validation of spike detection results in this investigation was especially challenging as there was no ground truth data such as simultaneous MEA and patch clamp or juxtacellular recordings. Instead, validation was performed using pharmacology (tetrodotoxin, Fig. A.4) and manual verification of voltage traces with spike time markers. To this end, I have also developed an interactive user interface application available at my GitHub: `https://github.com/jeremi-chabros/WATERS`.

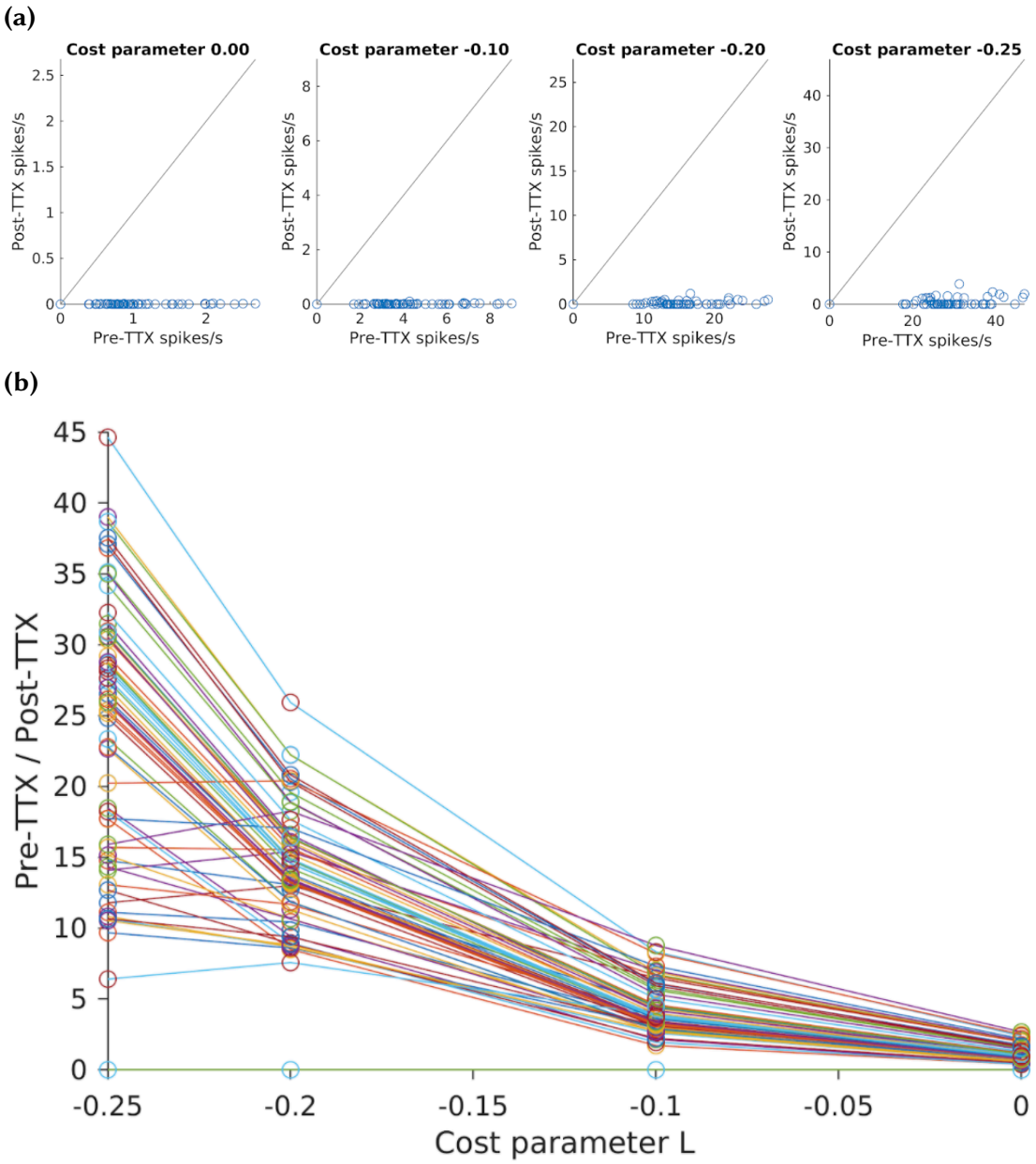


Figure A.4: Optimising cost parameters for spike detection. **(a)** Scatter plots showing spike counts before and after the application of tetrodotoxin (TTX). **(b)** Plot showing the ratio of the spikes detected before and after the application of TTX with respect to the cost parameter.

A.3 Method benchmarking

Development of a novel spike detection algorithm was motivated with unsatisfactory performance of available published methods. Therefore, a thorough benchmarking was required to determine if the method in question is indeed superior. Qualitative benchmarking was performed manually. For quantitative benchmarking, a synthetic data approach modified from Lieb et al. (2017) was used. From a recording, elec-

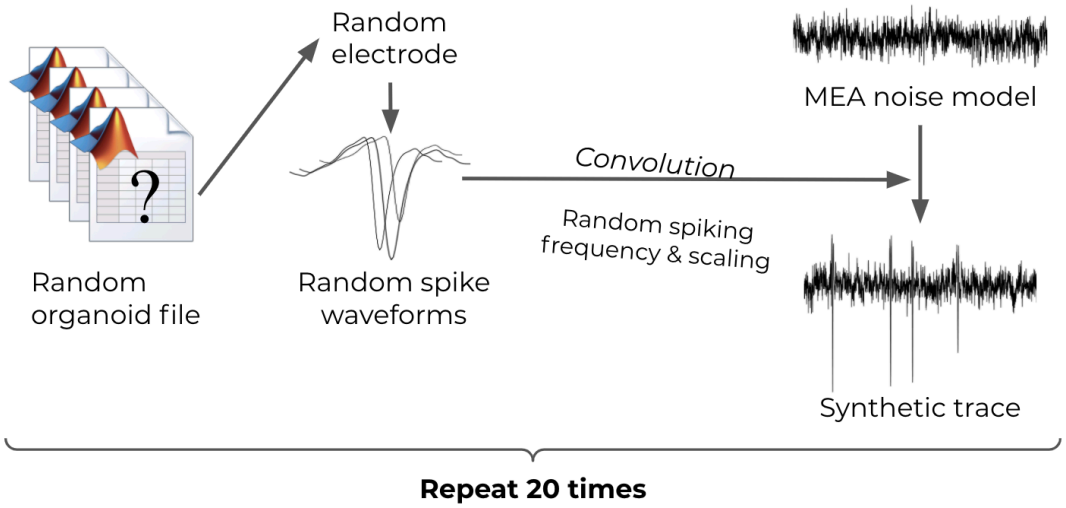


Figure A.5: A graphical summary of the procedure used to create synthetic traces from real MEA data.

trode, and spike detection method selected at random, 20 spike templates have been chosen with full blinding to their sources. Templates were then randomly scaled, reshaped, and convolved with an MEA correlated noise model with magnitude adjusted to the SNR observed in our data (2.5–3.5). Spiking frequencies and recording durations of the synthetic traces have also been matched to experimental data (Fig. A.5).

The results of the benchmarking are shown in Tab. A.1.

Method	DR (%)	FP (%)
WTEO	83.23	16.77
GABR	88.09	11.91
SWTO	77.80	22.20
MTEO	85.86	14.14
SWTS	84.71	15.29
ABS	71.96	28.04
WATERS	**95.30**	**4.70**
THR	46.62	53.38

Table A.1: The results of the benchmarking of the spike detection methods. Averaged across 20 trials. Duration of signal between 30 and 600 seconds. Spiking frequency between 0.5 Hz and 20 Hz. SNR between 2.5 and 3.5. DR: detection rate, FP: false positive rate. WTEO: wavelet based Teager energy operator, GABR: Gabor transform, SWTO: stationary wavelet transform Teager energy operator, MTEO: multiresolution Teager energy operator, SWTS: stationary wavelet transform, ABS: absolute thresholding algorithm, WATERS: a novel spike detection developed in this investigation, THR: threshold-based method previously used by our lab.

B

Functional connectivity

In current investigation, the functional activity has been inferred from spontaneous multi-unit activity. To this end, I have applied the spike time tiling coefficient (STTC, Cutts & Eglen, 2014). This measure has been chosen because of its mathematical tractability, ease of use and interpretation, and other advantages outlined below.

Necessary properties

- Symmetry. STTC calculated from neuron A in respect to neuron B is the same as from neuron B to neuron A.
- Robustness to variations in firing rates.
- Robustness to recording duration.
- Mathematical tractability and ease of interpretation. This measure is bounded by -1 (perfect anticorrelation) and 1 (perfect correlation).
- Robustness in variations in the time lag parameter.
- Ability to discriminate between lack of correlation and anticorrelation.

Desirable properties

- Ignores periods of quiescence.

- Naive to the spike time distribution (i.e. does not assume a particular distribution in firing rates).
- Minimal number of free parameters. This method relies on a single adjustable parameter.

Mathematical formulations and usage

STTC relies on a single lag parameter that specifies the period (in ms) before and after a given spike. Then, the method searches this time bin for spikes from different channel. Mathematically, it can be defined as

$$\mathrm{STTC} = \frac{1}{2}\left(\frac{P_A - T_B}{1 - P_A T_B} + \frac{P_B - T_A}{1 - P_B T_A}\right)$$

where

T_A is the proportion of the total recording time which lies within $\pm\Delta t$ of any spike from A. T_B is calculated analogically.

P_B is the proportion of spikes in A which lie within $\pm\Delta t$ of any spike from B. P_A is calculated analogically.

Graphical summary of the method can be found in Fig. B.1.

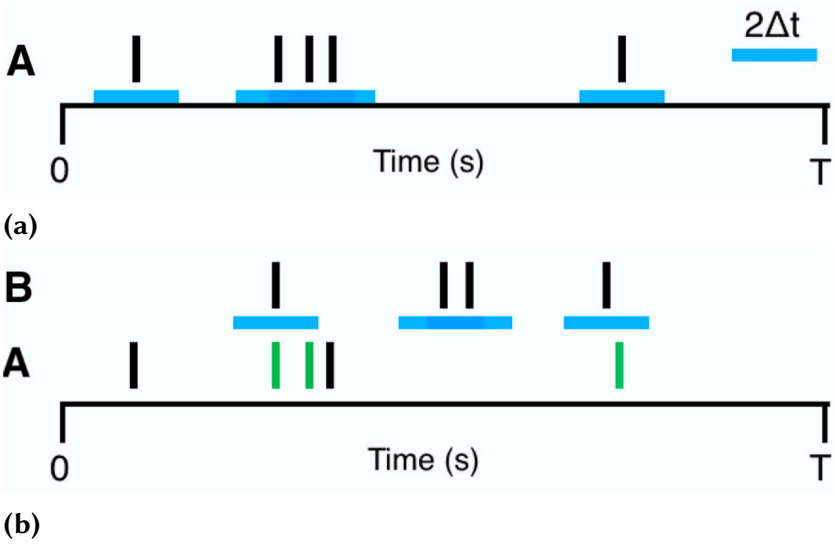


Figure B.1: Graphical representation of the STTC. **(a)** T_A is given by the fraction of the total recording time (black) which is covered (tiled) by blue bars. Here T_A is 1/3. **(b)** P_A is the number of green spikes in A (3) divided by the total number of spikes in A (5). Here P_A is 3/5. Figure adapted from Cutts & Eglen, 2014.

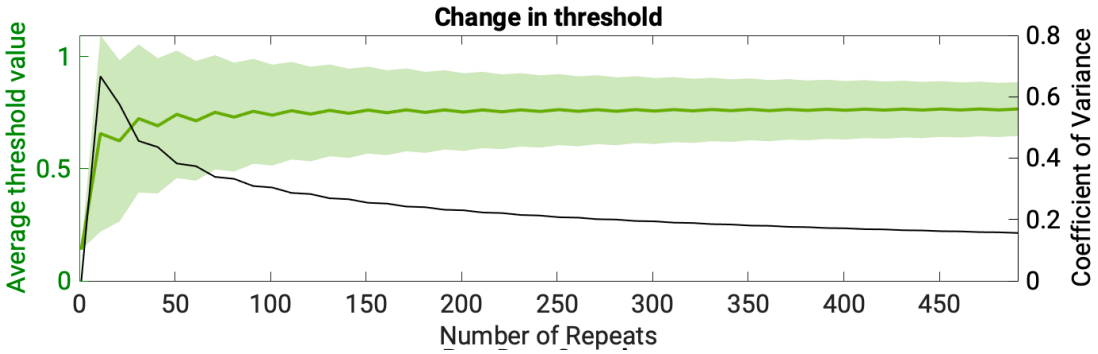


Figure B.2: Probabilistic thresholding. The average threshold value (left) and coefficient of variance (right) across iterations. The aim is to find a value of coefficient of variance lower than the average threshold in the part of the curve where both values have conspicuously stabilised.

Probabilistic thresholding

Despite the accuracy of STTC, it can still return erroneous values. For example, if a single electrode simultaneously records from multiple neurons, the probability that the electrical activity from this electrode will be randomly correlated with another increases. To alleviate this problem, we can generate synthetic data by shuffling the original spike trains, and only accept correlations that occur above chance. In current investigation, to generate synthetic data I have used circular shifting because it preserves the physiological delays between events (i.e. it is more accurate than randomly generated spike trains).

This procedure relies on two pre-specified parameters: the number of iterations used in shuffling the original data, and the probabilistic threshold. I have used the widely accepted threshold of 0.05 meaning that the probability of the connection between a pair of neurons being falsely inferred (false positive) was 5%. To determine the required number of iterations, I have generated a cumulative curve of the number of removed connections and chosen the number of iterations at which it stabilised. I have chosen 250 iterations. A sample plot is shown in Fig. B.2.

C

Effective rank

The neuronal activity recorded at each electrode can be described by a binary vector where ones represent action potentials, and zeros represent periods of quiescence. First, I divided each vector into 25 ms time bins, and calculated the spiking frequency in each time bin. Then, I applied variance stabilising transformation by taking the square root of the obtained values. Finally, I created a matrix of network activity by stacking the activity vectors from all electrodes so as each row described the evolution of nodal dynamics throughout the recording, and each column described the activity of the whole network in a given time bin. I have thus obtained a matrix $M \in \mathbb{R}^{N \times t}$ where N is the number of electrodes, and t is the number of time bins.

Like with PCA, the calculation of effective rank begins with singular value decomposition

$$M = U \Sigma V$$

where U and V are unitary matrices of size $N \times N$ and $t \times t$, respectively, and Σ is an $N \times t$ diagonal matrix containing the (real positive) singular values

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_Q \geq 0,$$

with $Q = \min\{N, t\}$. For notational simplicity, we further define $\sigma = \left(\sigma_1, \sigma_2, \ldots, \sigma_Q\right)^{\mathsf{T}}$ and the singular value distribution

$$p_k = \frac{\sigma_k}{||\sigma||_1} \quad \text{for} \quad k = 1, 2, \ldots, Q,$$

where the superscript $^{\mathsf{T}}$ denotes the transpose, and $\|\cdot\|_1$ the ℓ_1-norm defined as

$$\|\sigma\|_1 = \sum_{k=1}^{Q} |\sigma_k|.$$

In the sequel, all logarithms are to the base e, and we adopt the convention that $0 \log 0 = 0$. The effective rank is defined as follows.

Definition: effective rank

The effective rank of the matrix M, denoted $\mathrm{erank}(M)$ is defined as

$$\mathrm{erank}(M) = \exp\{H(p_1, p_2, \ldots, p_Q)\},$$

where $H(p_1, p_2, \ldots, p_Q)$ is the (Shannon) entropy given by

$$H(p_1, p_2, \ldots, p_Q) = -\sum_{k=1}^{Q} p_k \log p_k.$$

Note that the above entropy is often referred to as spectral entropy.

Relationship between bin duration and effective rank

The choice of the timescale at which neuronal dynamics are studied is arbitrary. Shorter time bins will shift the focus to fine temporal activity patterns (such as mono- or disynaptic), whereas longer time bins will better describe long-term trends in network dynamics (multisynaptic, populational). Moreover, the choice whether to calculate the effective rank from correlation matrix of neural activity, or directly from spike trains, will also depend on the approach. Using correlated activity will blur the temporal resolution of the inferred dynamics but is more robust to stochastic spiking and random correlations, for example between quiescent periods. With time-series data, the effects of noise are reduced, but the choice of the timescale will have a larger impact on the observed dynamics. Quantitative differences between these approaches are summarised in Fig. C.1.

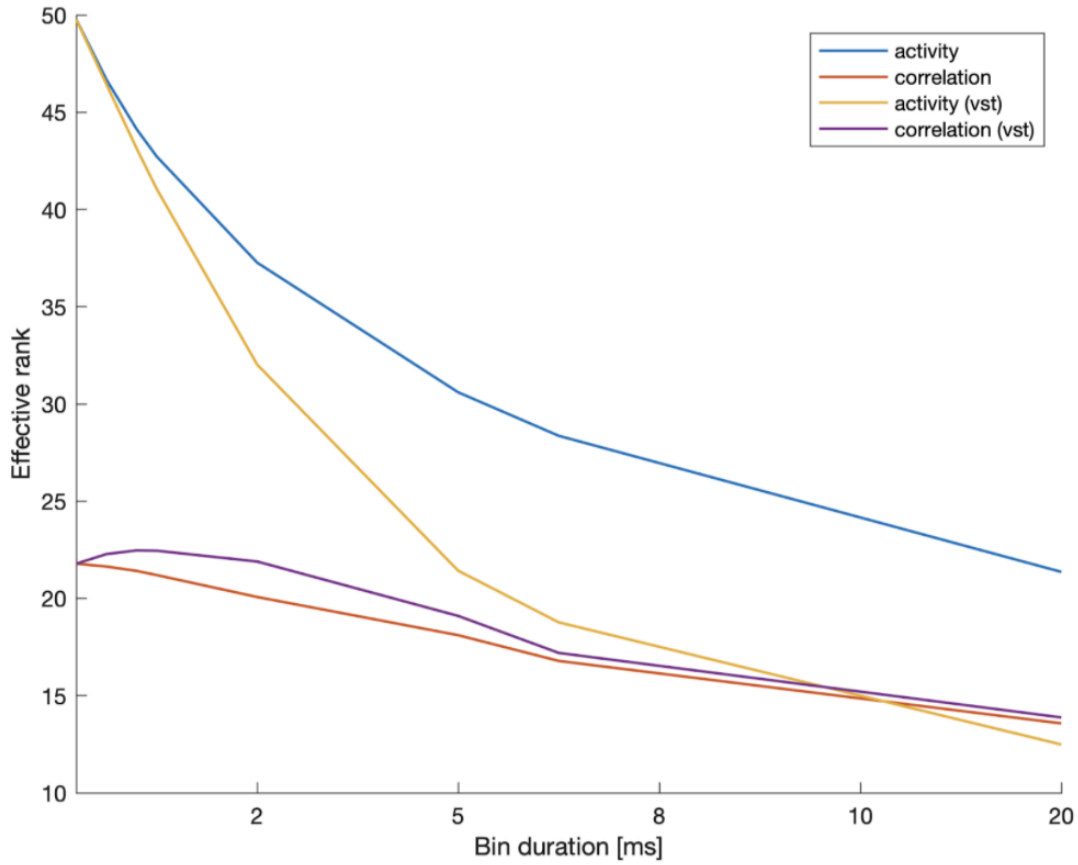


Figure C.1: The effects of different approaches in calculating the effective rank on dynamical dimensionality. *Correlation* means using correlation matrix, and *activity* means using raw spiking data. vst - variance stabilising transformation. Data from a sample recording of a wild type culture at DIV28 (culture code: MPT200209_4A).

D

Network controllability

D.1 Network control theory: background

The core of network control theory (as applied to Neuroscience) is the structural network of neurons and their connections. To model the temporal evolution of network dynamics as a function of its connectivity, we defne a linear time-invariant (LTI) system:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$$

where the column vector $\mathbf{x}(t) = \left[x_1(t), x_2(t), \ldots, x_N(t)\right]^{\mathsf{T}}$ represents the state of the system of N nodes at time t. The adjacency matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$ denotes the relationships between the nodes. In this context, this relation is operationalized as the functional connectivity between each pair of neurons.

In order to impose control over this system, we need to specify inputs. Our model then takes the form

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}_{\kappa}\mathbf{u}_{\kappa}(t)$$

where the binary input matrix $\mathbf{B}_{\kappa} \in \mathbb{R}^{N \times m}$ denotes the set of m control nodes. The term $\mathbf{u}_{\kappa}(t) \in \mathbb{R}^{m \times 1}$ is a vector of m functions denoting the control input, which is the amount of input injected into each of the m control nodes at each time point t. Over time, $\mathbf{u}_{\kappa}(t)$ denotes the injected control input over time. Note that substituting $\mathbf{B}_{\kappa}$ with identity matrix $\mathbf{I} \in \mathbb{R}^{N \times N}$ is equivalent to calculating controllability from each node iteratively, and represents the overall controllability from *any* input.

By applying to this model the electrophysiological data collected with the use of micro-electrode arrays (MEA), we can answer a multitude of questions about how the structure of a network influences its dynamics. In current investigation, we are particularly interested in whether in vitro neural networks are controllable and how the different control theoretical metrics are shaped by neurodevelopment in health and in mouse model of Rett syndrome.

D.2 The effects of time system on controllability metrics

In this investigation, I have used the discrete-time model of network dynamics. However, this analysis can be extended to continuous time, yielding equivalent (to a scalar constant) results.

D.2.1 Discrete time

Natural system

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$$

Controlled system

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}_{\mathcal{K}}\mathbf{u}_{\mathcal{K}}(t)$$

Normalization

For interpretability of results, I scaled the adjacency matrix

$$\mathbf{A}_{\text{norm}} = \frac{\mathbf{A}}{\lambda(\mathbf{A})_{\max} + c} - \mathbf{I}$$

where $\lambda(\mathbf{A})_{\max}$ denotes the largest eigenvalue of $\mathbf{A}$, and c is a scaling constant. By changing c, one can alter the transfer function and enforce the system to dissipate energy (tend to zero) to a variable extent. Note that this normalisation only scales the entries of $\mathbf{A}$ without the loss of generality.

Controllability Gramian

The controllability Gramian is central to calculating almost every controllability metric. For a continuous time system it can be computed as

$$\mathbf{W}_{\mathrm{c}} = \int_0^\infty \exp\left(\mathbf{A}\tau\right) \mathbf{B}_\kappa \mathbf{B}_\kappa^{\mathsf{T}} \exp\left(\mathbf{A}^{\mathsf{T}}\tau\right) \mathrm{d}\tau$$

or, alternatively, by solving the Lyapunov equation

$$\mathbf{A}\mathbf{W}_{\mathrm{c}} + \mathbf{W}_{\mathrm{c}}\mathbf{A}^{\mathsf{T}} + \mathbf{B}_\kappa \mathbf{B}_\kappa^{\mathsf{T}} = 0$$

limitation being that it can only be applied to (Hurtwitz) stable systems.

Almost as important as the Gramian itself is its eigendecomposition. The eigenvalues of the controllability Gramian determine how difficult is to control network dynamics in a direction specified by the corresponding eigenvector. Larger eigenvalues mean less control energy, whereas smaller eigenvalues specify directions in state space where control is difficult (energetically costly). This fact will be useful in understanding the average controllability.

Modal controllability

$$\phi_i = \sum_{j=1}^{N} \left(1 - e^{\lambda_j(\mathbf{A})}\right) \mathbf{v}_{ij}^2$$

where $\lambda_j(\mathbf{A})$ and $\mathbf{v}_j$ are the j-th eigenvalue and eigenvector of $\mathbf{A}$. Note that this measure is a sum of modes that can be excited by a given node, weighted by the corresponding eigenvalues (i.e. the larger the eigenvalue, the more controllable that mode is).

D.2.2 Discrete time

Natural system

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t)$$

Controlled system

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}_\kappa \mathbf{u}_\kappa(t)$$

Normalization

$$\mathbf{A}_{\text{norm}} = \frac{\mathbf{A}}{|\lambda(\mathbf{A})_{\max}| + c}$$

Controllability Gramian

$$\mathbf{W}_{\text{c}} = \sum_{\tau=0}^{T-1} \mathbf{A}^{\tau} \mathbf{B}_{\kappa} \mathbf{B}_{\kappa}^{\mathsf{T}} \left(\mathbf{A}^{\mathsf{T}}\right)^{\tau}$$

Modal controllability

$$\phi_i = \sum_{j=1}^{N} \left(1 - \lambda_j^2(\mathbf{A})\right) \mathbf{v}_{ij}^2$$

D.2.3 Metrics independent of time system

Average controllability

$$\text{ctrb}_{\text{ave}} = \text{Tr}(\mathbf{W}_{\text{c}})$$

Notice that the trace of the Gramian can be calculated as the sum of the elements on its diagonal or, more importantly, as the sum of its eigenvalues. Because, due to the above normalization, the eigenvalues are close to 1, this metric would be strongly biased by the network size. I have controlled for it while calculating the average controllability by dividing obtained scores by the number of active nodes.

Controllability ellipsoid

We can define the volume of the controllability ellipsoid as

$$V(\epsilon) = H_n \sqrt{\det(\mathbf{W}_{\text{c}})},$$

where n is the dimension of the hyperellipsoid (equal to the number of nodes in the network), and H_n is the hypersphere volume coefficient given by

$$H_n = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2} + 1)}$$

For simplicity, in the current report the volume of the controllability ellipsoid has been approximated by excluding the hypersphere volume coefficient from the calculations. However, because in higher dimensions the determinant becomes numerically problematic, and because logarithm is monotone, I have adapted the volume of controllability ellipsoid as proposed by Summers et al., 2016.

$$V(\epsilon) = \log \det \mathbf{W}_c$$