COMPUTATIONAL PRINCIPLES OF BRAIN NETWORK DEVELOPMENT



Danyal Akarca

Pembroke College

This thesis is submitted for the degree of Doctor of Philosophy

MRC Cognition and Brain Sciences Unit,

School of Clinical Medicine,

University of Cambridge

November 2022

In loving memory of my mother, Cerys Akarca

– our beacon of kindness and love.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS **8**

PREFACE **9**

SUMMARY **10**

1. CHAPTER 1 General Introduction **12**

1.1. Development is probabilistic not deterministic *13*

1.2. Developmentally-inspired computational modelling *15*

1.3. Modelling brain development as networks undergoing change *19*

1.4. The generative network modelling framework *22*

1.5. Thesis structure *30*

2. CHAPTER 2 A generative network model of neurodevelopmental diversity in structural brain organisation **31**

2.1. Introduction *32*

2.2. Methods *33*

2.2.1. Participants 33

2.2.2. MRI acquisition and pre-processing 35

2.2.3. Connectome construction and cortical morphology 35

2.2.4. Generative network models 36

2.2.5. Cognitive and learning assessments 38

2.2.6. Gene expression data 38

2.2.7. Statistical analyses 39

2.2.8. Gene enrichment analysis and visualisation 41

2.3. Results *42*

2.3.1. Small variations in parameter combinations produce accurate spatially-embedded networks 42

2.3.2. Individual differences in wiring parameters mirror connectome organisation, grey matter morphology and cognitive scores 49

2.3.3. Variability in neurodevelopmental trajectories arises through value-updating over time 53

2.3.4. Genomic patterning of network growth 55

2.4. Discussion *59*

3. CHAPTER 3 Evaluating wiring principles underpinning neuronal network topology *in vitro* across species and time 68

3.1. Introduction *69*

3.2. Methods *73*

3.2.1. High-density microelectrode arrays 73

3.2.2. Rodent primary cortical neuronal cultures 73

3.2.3. Human induced pluripotent stem cell-derived cultures 75

3.2.4. Human cerebral organoid slice cultures 76

3.2.5. Immunohistochemistry 76

3.2.6. Scanning electron microscope imaging 78

3.2.7. Electrophysiological recordings 78

3.2.8. Pharmacological experiments 79

3.2.9. Spike-sorting and post-processing 79

3.2.10. Firing rate and burst statistics 79

3.2.11. Functional connectivity inference 80

3.2.12. Generative network modelling and Voronoi tessellations 81

3.2.13. Model fitting 81

3.2.14. Generative probability distributions 83

3.3. Results *83*

3.3.1. Tracking developing neurons at single-cell resolution 84

3.3.2. Homophilic wiring principles recapitulate the topology of developing rodent neuronal networks *in vitro* 86

3.3.3. Effect of plating density on topology and generative model outcome 91

3.3.4. Topological fingerprints arise from homophilic mechanisms in developing neuronal networks *in vitro* 93

3.3.5. Effect of $GABA_A$ receptor antagonism on generative network wiring 97

3.3.6. Generative wiring principles in human neuronal cultures *in vitro*.....101

3.4. Discussion *105*

4. CHAPTER 4 Spatially-embedded recurrent neural networks exhibit numerous brain-like structural and functional properties.....112

4.1. Introduction..... *113*

4.2. Methods..... *115*

4.2.1. Task paradigm.....115

4.2.2. Recurrent neural network (RNN) modelling 117

4.2.3. Topological analysis 124

4.2.4. Generative network modelling 125

4.2.5. Decoding.....126

4.2.6. Spatial permutation testing of task-relevant information 127

4.3. Results *128*

4.3.1. Modular small-world recurrent networks emerge from Euclidean and topological constraints 128

4.3.2. seRNNs can be approximated via homophily generative models.....131

4.3.3. Functional clustering in space and adaptive coding 135

4.3.4. Mixed selectivity and energetically efficient coding.....139

4.3.5. Regularisation leads to convergent brain-like topology and function 142

4.3.6. Replication across baseline architectures.....146

4.4. Discussion *148*

5. CHAPTER 5 CONCLUSIONS AND FUTURE OUTLOOK 154

5.1. So, what have we learned?..... *155*

5.2. What makes a good computational model of developmental processes? *158*

5.2.1. Simple, but *explainable*, models can recapitulate developmental findings.....158

5.2.2. Models operate across different levels 161

5.2.3. Models span across species 165

5.2.4. Models elucidate developmental processes and time-courses.....168

5.3. Concluding remarks *169*

6. APPENDIX A Supplementary Materials for Chapter 1 171

Graph theory network statistic formalisms..... 172

7. APPENDIX B Supplementary Materials for Chapter 2 175

8. APPENDIX C Supplementary Materials for Chapter 3 195

9. APPENDIX D Supplementary Materials for Chapter 4 218

10. CODE AND DATA AVAILABILITY 231

11. BIBLIOGRAPHY 232

ACKNOWLEDGEMENTS

First, I would like to thank my supervisor, Prof Duncan Astle, for his continual guidance and for allowing me the opportunity to learn from him. In particular, for supporting me through challenging personal times, and creating an incredibly open and fruitful environment for pursuing my work. It really has been a privilege. I am also very thankful to my advisor, Dr Petra Vértes, for her support and insightful guidance.

I want to also acknowledge all members of the Development, Dynamics, Disorders & Data Science (4D) lab who have all been fantastic throughout the duration of my PhD. I am very appreciative of all my direct collaborators, many of whom I hope to work with for a very long time – Jascha Achterberg (Cambridge), Alexander Dunn (Cambridge), Prof Edward Bullmore (Cambridge), Prof John Duncan (Cambridge), Prof Mark Johnson (Cambridge & Birkbeck), Dr Manuel Schröter (ETH Zürich), Prof Stephen Eglen (Cambridge), Prof Ole Paulsen (Cambridge), Dr Susanna Miraeu (Harvard & Cambridge), Sofia Carozza (Cambridge), Prof Matt Botvinick (DeepMind & UCL) and Dr DJ Strouse (DeepMind). I would also like to extend my gratitude to Dr Richard Bethlehem (Cambridge) and Prof Dani Bassett (UPenn & Santa Fe) for examining this thesis.

Very special thanks to the Medical Research Council Doctoral Training Programme (MRC DTP) and Cambridge Trust for their financial support, and to the MRC Cognition and Brain Sciences Unit (CBU) for being an academic home for me.

Most importantly, I would like to thank my close friends and family who have loved and supported me in ways that I simply cannot put into words. I am eternally grateful.

PREFACE

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text. It is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the preface and specified in the text. It does not exceed the prescribed limit of 60,000 words as specified by the Degree Committee of Clinical Medicine and Clinical Veterinary Medicine.

The work outlined within each chapter have been published in the following references:

- **Akarca, D.**, Vértes, P.E., Bullmore, E.T., the CALM team & Astle, D.E. (2021). A generative network model of neurodevelopmental diversity in structural brain organisation. *Nature Communications.*
- **Akarca, D.***, Dunn, A.W.E.*, Hornauer, P.J., Ronchi, S., Fiscella, M., Wang, C. *et al.* (2022). Homophilic generative principles underpin neuronal network topology across scales, species and time. *In review at eLife.*
- Achterberg, J.*, **Akarca D.***, Strouse, DJ., Duncan, J.† & Astle, D.E.† (2022). Spatially-embedded recurrent neural networks reveal widespread links between structural and functional neuroscience findings. *In review at Nature Machine Intelligence.*
- Astle, D., Johnson M.J. & **Akarca, D.** Towards computational neuroconstructivism: a framework for developmental systems neuroscience. (2022). *In review at Trends in Cognitive Sciences.*

*Co-lead authors, †Co-senior authors

The study in **Chapter 3** was done in collaboration with Dr Manuel Schröter, and his team at ETH Zürich, who provided pre-processed neuronal data and (ii) Alexander Dunn, and his team at Cambridge, who helped in interpreting and analysing experimental neuronal data. **Chapter 4** was done in close collaboration with Jascha Achterberg.

SUMMARY

Brain development can be viewed through many lenses and studied at many scales. However, multiple theoretical perspectives have argued that brain organisation develops via competitive interactions between its constituent units, dynamically over time. In this thesis, I focus on modelling these interactions.

In **Chapter 1**, after providing an historical backdrop to the field of developmental systems neuroscience, I introduce generative network models. These relatively new family of models are capable of simulating probabilistic network development. The basis for these models includes simple sets of wiring rules, existing within various imposed biophysical constraints, which steer the developmental trajectory of the network.

In **Chapter 2**, I show how the applications of these models can reveal simple principles that may contribute to our understanding of neurodiversity. In particular, small iterative updates in networks can lead to constrained variability in a child's macroscopic structural brain organisation inferred via *in vivo* diffusion imaging. I highlight how decompositions of networks into the generative components used to construct them in this way can be useful lower-dimensional representations of developmental ingredients. This is particularly relevant when aiming to bridge associations between genomics, cognition and the brain for answering developmental questions.

Generative network models emphasise the evolving economic context of dynamic interactive negotiations between brain regions. These regions can be defined at any scale. In **Chapter 3** I pivot from studying cross-sectional macroscopic connectomes, to modelling the microstructural longitudinal development of *in vitro* neuronal networks at the cellular scale. I show that current instantiations of a homophily generative model are an effective growth model of *in vitro* neuronal network development. This simple model can recapitulate observable local topological organisation of functional networks across species, time, plating densities, cell-types and experimental conditions. Together, **Chapter 2** and **3** can be

considered as a test of whether generative network models can simulate biological brain topologies, in an unsupervised fashion, according to intrinsic wiring rules.

The nervous system has evolved, in part, to sustain and ensure survival of the organism. Therefore, the structural organisation of the brain must be considered with respect to how it directly supports function in order to achieve behavioural goals. However, many current frameworks posit only *associations* between neural structure and function, rather than direct bidirectional influences. In **Chapter 4**, I aim to model how previously aforementioned economic negotiations may facilitate direct structure-function interactions. I introduce an extension of artificial neural networks for which I term spatially-embedded recurrent neural networks (seRNNs). seRNNs add simple biophysical constraints into the model within a regularisation term, that change the nature of how connections change during optimisation. I show that adding local spatial and communication constraints to this neural network points towards a convergent solution whereby optimal functional trade-offs are attained where sparsity, homophily generative mechanisms, small-worldness, functional configuration in space and energetic efficiency together coalesce.

In **Chapter 5** I summarise key take-aways and provide what I believe to be promising future avenues for the applications of computational modelling to developmental systems neuroscience.

1. CHAPTER 1

GENERAL INTRODUCTION

Content included in this chapter is also included in an article currently in review at Trends in Cognitive Sciences: Towards computational neuroconstructivism: A framework for developmental systems neuroscience. Duncan E. Astle, Mark H. Johnson & ***Danyal Akarca****.*

To understand the functional role of neural circuits we must understand the developmental history of those circuits. This is a guiding principle underpinning developmental systems neuroscience. Put simply, we must study trajectories of change rather than end state outcomes alone, because the functional role of any set of integrated neurons is shaped by its interactions with other neural populations, across development. Each part of the brain develops within the context of the rest of the brain – a phenomenon termed *embrainment* (M. H. Johnson & De Haan, 2015; Mareschal et al., 2007). Over the past 30 years, this perspective has been central to how cognitive development and thereby the disorders that can emerge during development have been understood. However, producing robust theoretical models able to incorporate this developmental perspective, in a way that both encompasses biological mechanisms and is falsifiable, represents one of the main conceptual and technical challenges in the field.

1.1. Development is probabilistic not deterministic

One reason that theoretical progress has been challenging is because development is probabilistic, not deterministic. A deterministic view of brain development is simply that genes code for neuroanatomy, neurochemistry and neurophysiology, which in turn determines function, and this consequentially enables particular cognitive processes to come 'online' at particular ages. According to this view, cognitive processes become adult-like when their underpinning neurobiology reaches maturity; under this view, development is considered as a set of incremental and independent processes which occur along a maturational time-course. Likewise, cognitive difficulties reflect specific disruption to the biological pathways upon which they depend. This simple logic implicitly underpins much of cognitive neuroscience and the study of developmental disorders, but this 'deterministic epigenesis' cannot adequately explain a wide range of phenomena (Astle & Fletcher-Watson, 2020; Cicchetti & Rogosch, 1996; Kendler & Greenspan, 2006; Resnik & Vorhaus, 2006; Waddington, 1957; Waggoner & Uller, 2015). Instead, functional activity emerges before brain structures are fully mature, supporting gradually emerging cognition and behaviour. This functional activity, and the nascent

computational processes they support, themselves shape the ongoing development of the neural circuitry in a highly interdependent fashion (Gottlieb, 1998, 2007). In other words, developmental change at any level of analysis is not in fact deterministic, but is instead sensitive to other ongoing developmental processes, both across and within levels. This alternative view, termed 'probabilistic epigenesis' (Gottlieb, 2005, 2007) is a foundational principle of a constructivist approach to developmental cognitive and systems neurosciences (M. H. Johnson, 2000, 2011). However, incorporating this principle into any theory makes it incredibly hard to instantiate formally as a model, because it is difficult to parameterise all of the potentially relevant interactions.

This neuroconstructivist perspective has challenged the wider fields of cognitive psychology and neuroscience because it implies that current cognition or behaviour are themselves an emergent property of dynamic interactions that occurred within that organism's developmental past. That is, the brain progressively becomes constructed via its self-organisation, with individual components of its construction gradually specialising over developmental time. Subsequently the functional role of a brain area is, at least in part, shaped by its interactions with, and inputs from, other areas. This process of functional specification over development has been termed Interactive Specialisation (M. H. Johnson, 2000, 2011). Interactive Specialisation is a domain-general framework that extends the concept of probabilistic epigenesis and posits a number of phenomena we should expect given this principle. Firstly, connectivity between regions will be crucial for their functional development, as it is the connectivity that determines how different areas can influence each other. Over time, reciprocal interactions mean that the resulting functional role of a brain area cannot be considered independently from its connections. Secondly, and relatedly, networks should show some degree of activity-dependent self-organisation. As activity within one brain region is shaped by that occurring elsewhere, over time this will result in more complex emergent functional and structural networks, driven by this coactivity over time. Thirdly, there should be a dynamic correspondence between a cognitive process and neurophysiology. Rather than cognitive processes simply becoming 'operational' when their corresponding neural circuits

reach maturity, there should be a dynamic flexibility in the functional role of neural processes over development. As computational processes change, this may itself drive re-organisation of neural networks.

The field of developmental systems and cognitive neuroscience has produced many findings that can be viewed through this lens. For example, over developmental time functional brain networks gradually integrate (growth of long-range connections connecting neuronal populations over vast anatomical distances) and segregate (pruning of local connections) to produce more specialised systems (Fair et al., 2007; Jones et al., 2022; Sporns, 2013a; Wig, 2017); the neural systems associated with a cognitive process change across development (Durston et al., 2006); many cortical regions involved in perceptual and cognitive functions show increasing specialisation (tuning) to particular stimuli, or stimulus-task relations, and studies consistently show variability in the mapping between the brain and cognition across individuals (Kievit et al., 2014; Siugzdaite et al., 2020). In summary, there is substantial empirical evidence for the principle of probabilistic epigenesis, as implemented in frameworks like Interactive Specialisation. This probabilistic principle is an essential foundational ingredient for any plausible developmental framework.

1.2. Developmentally-inspired computational modelling

Whilst frameworks such as Interactive Specialisation provide a narrative account for the principles of developmental systems and cognitive neuroscience, this is not equivalent to providing formal computational model that can be implemented, tested and falsified. This critical step has proved difficult. From a practical perspective, the data needed to test any specific predictions will necessarily take years to collect. If any predictions are not supported by the data then there could be any number of reasons: the predicted effects could simply have occurred previously in development, or not yet, or the neuroimaging method is not sensitive to the type of change predicted. Empirical progress has been made by using frameworks like Interactive Specialisation to generate specific predictions about targeted systems (Battista et

al., 2018; Joseph et al., 2011). Nonetheless, as a framework, Interactive Specialisation has a limited degree of testability.

Elsewhere, in the domains of cognition, psychiatry or neuroscience, there is a computational revolution underway (Adams et al., 2016; Hassabis et al., 2017; Huys et al., 2016; Kriegeskorte & Douglas, 2018; Murray et al., 2018; Ritter et al., 2013). The mathematical formalisation of theory is having a transformative impact on our ability to integrate different phenomena mechanistically at many different levels of analysis. For example, deep convolutional neural networks have been shown to be excellent models of visual cortex in both humans and monkeys, mirroring the increasingly complex visual representations as we transition from early to deep cortical layers (Khaligh-Razavi & Kriegeskorte, 2014). Dynamic causal models, which aim to infer the causal architecture of distributed dynamical systems, have been used to isolate cellular channel dysfunction by modelling how altered synaptic currents may lead to changes in observed electroencephalography recordings (Symmonds et al., 2018). Recurrent neural network models of prefrontal cortex have been integrated with models of phasic dopaminergic release to form a meta-reinforcement learning system which accounts for otherwise disparate observations (Wang et al., 2018). These examples all showcase the value of computational modelling for unpacking complex interdependent neurobiological processes.

As in other areas of science, a computational framework for developmental systems neuroscience could help to implement theories as more testable accounts. In addition to enabling us to specify theory in mathematical terms, this may precipitate a shift from group-level characterisation of mechanisms towards precision individual-level models. Mathematically formalising predictions made from candidate hypotheses can improve theory building (Blohm et al., 2020; De Schutter, 2001), because it is always relatively easy to interpret neuroscience findings in the context of a hypothesis, simply because it is often *easy to vary*. That is, many hypotheses in developmental systems neuroscience are sufficiently flexible that they can encompass many observations. In contrast, formalised mathematical model can simulate data that would occur if the hypothesis (i.e., the model) were true. This

is a subtle but meaningful difference (Popper, 1959). With this formalism the hypothesis is *hard to vary*, allowing it to be falsified by examining the resultant cross-validated estimate of its accuracy. Importantly, mathematics-based theories are not intrinsically superior as a means to produce knowledge than narrative-based theories *per se*. Both are often complementary and may or may not be better suited to particular questions or fields. Both can be prone to ambiguity if the underlying axioms of the theory are not precisely defined. However, in the study of developmental systems neuroscience where there are numerous ambiguous factors at play, mathematics-based theories – encompassed within computational models – offer an alternative approach to the prevailing narrative-first paradigm. They provide researchers with a set of tools to define their axioms and to falsify competing theoretical accounts.

Computational models for understanding developmental processes are not new. There is a long history of attempts to capture developmental changes in cognitive performance with such models (Davis & Anderson, 1999; Harm & Seidenberg, 1999; Li & Lindenberger, 1999; Mareschal et al., 1999; Mareschal & Shultz, 1996; Nettelbeck & Wilson, 1985; Sternberg, 1983; Thomas & Karmiloff-Smith, 2002, 2003). Connectionist models, in which surface-level performance is achieved by tuning computational parameters, have enabled researchers to assess whether developmental and individual differences in cognition are qualitatively or quantitatively similar (Rumelhart, McClelland, et al., 1986). The formalisation of processes underpinning performance affords a further specificity about processing components that cannot be achieved with a descriptive theory. For example, concepts such as 'processing speed' that are often posited as responsible for neurodevelopmental or individual differences (Nettelbeck & Wilson, 1985; Sternberg, 1983) can be specified mathematically in multiple different ways and compared formally. The connection strengths between nodes can be adjusted systematically, according to different constraints, and the input-output performance of the resulting network compared formally against observed performance. The result is a flexible system that can be trained to perform tasks. Systematically varying the architecture of the network, or the parameters that govern its learning, can reveal candidate mechanisms that explain developmental or individual differences in performance (see Thomas & Karmiloff-

Smith, 2003 for review). This process of formalisation forced scientists to consider the process of development itself. For example, do developmental changes occur as a result of changes in learning parameters (Li & Lindenberger, 1999), learnt weights (Mareschal et al., 1999) or the gradual recruitment of established sub-structures within the model (Mareschal & Shultz, 1996)? Likewise, do disorders that emerge across the course of development stem from different parameters, learning rates, network architectures or some combination therein (Harm & Seidenberg, 1999; Oliver et al., 2000; Thomas & Karmiloff-Smith, 2002)?

While useful for capturing possible factors contributing to developmental differences, these models had several limitations. The models produced were usually specific to a task, in part because the input and output nodes needed to represent the sensory information, and the response parameters, are particular to a paradigm. Thus, whilst these models could accurately model performance on a particular task, or capture important developmental differences in ability on that task, they could not be generalised easily. Secondly, it is not clear how biologically plausible they really are. Whilst connectionist models are designed to mimic neuronal processes in a very general sense, the process by which they learn to perform tasks is not at all constrained by biological processes – it was not clear exactly how nodes of weighted connections map on to brain structure or function. The rapid expansion of neuroimaging applications in paediatric populations meant that developmental systems cognitive neuroscience focussed heavily on observed changes in neurobiology and brain organisation, and computational theory per se became less of a priority.

However, the dramatic growth in the availability and quality of neuroimaging data in paediatric populations is no substitute for mechanistic accounts that can be formally tested. In contrast, elsewhere, developmental neuroscience has produced computational models of specific biological processes, such as synaptogenesis (Baxter & Levy, 2020; Ju et al., 2017), plasticity (Chklovskii et al., 2004; G. Yang et al., 2009), consolidation (Fusi et al., 2005), pruning (Chechik et al., 1998; Navlakha et al., 2015) and axonal guidance (Roccasalvo et al., 2015). These computational models provide insights beyond what could be determined from the observed data alone, such as the role of adaptive synaptogenesis under metabolic

constraints (Ju et al., 2017), the trade-off between cortical developmental pace and performance (Baxter & Levy, 2020) and long-term learning-related changes in dendrites (G. Yang et al., 2009). Biologically inspired computational models like these broadly function by grounding otherwise mathematically-defined processes within a biological context where there are finite resources and/or biologically-reasoned axioms. As such, they retain the advantage of being quantitative, but can be tailored to specific biological problems rather than remaining mathematically abstract. Within developmental systems neuroscience, we need a computational approach that incorporates, and indeed is grounded in, our growing knowledge of brain development.

1.3. Modelling brain development as networks undergoing change

A biophysically grounded computational model for developmental systems neuroscience will need to be broader than the specific examples outlined in the previous section. It will need to consider the complex constellation of connectivity which mediates information transfer across the brain. The challenge is to extend biophysically grounded computational thinking beyond the well-defined scope of specific simple circuits or organisms to address broader neurodevelopmental questions, drawing on rich data from systems neuroscience (e.g., diffusion imaging and genomics).

How can we study this network connectivity? The foundational tool to study networks is *graph theory* – a branch of mathematics that examines networks (also termed *graphs*) which represent an underlying system of components and their interactions (Bassett & Sporns, 2017). In its simplest form, a graph is constituted by nodes and edges, and can be characterised at global or local levels. The nature of the arrangement of the network is termed its network *topology*. Importantly, there are many graph theory measures which compute specific summary statistics at these global and local levels, and I provide the mathematical formalisms for those implemented in this thesis in **Appendix A**.

When computing descriptive graph theory analyses on empirical structural and functional networks, across species, findings tend to converge on the presence of a set of canonical topological attributes. These include a concurrent high clustering and short path lengths, leading to small-worldness (Bassett & Bullmore, 2017; Bullmore & Sporns, 2009), and modules (Heuvel & Sporns, 2013) that are connected via highly integrated hub nodes. These hubs are sometimes termed the 'rich club' (Heuvel & Sporns, 2011; Zamora-López et al., 2010).

Why do these set of attributes tend to consistently emerge together? The most prominent explanation for these findings relates to economics. That is, the brain's organisation is shaped by economic trade-offs, minimising costs whilst allowing the emergence of adaptively valuable topological patterns of anatomical or functional connectivity (Bullmore & Sporns, 2012). This elegant explanation will be elaborated on in depth repeatedly throughout this thesis, as I will argue that it has very important implications for developmental systems neuroscience. However, despite its elegance, it is important to note that the economic account of brain organisation has been argued to be somewhat *hard-to-falsify* as many complex biological phenomena can be explained in terms of competing objectives (for more detail, see Rubinov, 2022). That is, the very inclusion of more than one factor to explain brain organisation (e.g., the observation that cost minimisation (Markov et al., 2013) is not *alone* sufficient to explain network organisation) could be argued to decrease falsifiability because any number of objectives could be pieced together to explain an observation, without it necessarily providing additional explanatory power. How can this be solved? One solution is parsimony. The smaller the number of modelled objectives to approximate empirical observations, the more likely the model is understandable and can be directly linked to plausible mechanisms. A second related solution is to not only outline how models capture the statistics of brain organisation, but to reconcile them directly with biological processes via their parsimonious explanations. In this thesis, both themes of parsimony and biological plausibility are returned to numerous times with relation to the presented models.

As alluded to previously, one aim of computational modelling in this area is to provide explanatory power: moving beyond observations of *what* the findings are, to showing explicitly

why they have occurred. To achieve this, I argue that *generative network models* are a very useful starting point in building a computational framework that incorporates some of the aforementioned neuroconstructivist principles, in order to provide additional explanatory weight to developmental theory.

What are generative network models? Unlike typical neural network models (such as cascade (Fahlman & Lebiere, 1990), recurrent (Rumelhart, Hinton, et al., 1986), convolutional (Lecun et al., 1998), or LSTM networks (Hochreiter & Schmidhuber, 1997)) or network evolutionary algorithms (such as Barabási–Albert (Albert & Barabási, 2002), Erdős–Rényi (Erdős & Rényi, 1959), Watts-Strogatz (Watts & Strogatz, 1998) or Stochastic-block models (Holland et al., 1983)), generative network models typically embed macroscopic regional structures (e.g., parcellated grey matter regions) within a 3D physical space. This physical grounding mirrors the physical neuroanatomy of the brain and allows it to act as simple biophysical constraint within the model. One way of characterising generative network models is to place them somewhere between stochastic blockmodels (which provides a closed form expression for the likelihood of connections depending upon community assignment) and pure developmental models (which directly operate on the timescale of the organism; Betzel & Bassett, 2017b). Another way to delineate generative network models from other model types is the extent to which external data guides their change over time. For example, in contrast to canonical artificial neural networks, generative network models update their topology iteratively according to intrinsic rules. That is, outside of their initial parameterisation, they do not require any externalities (i.e., training data) to shape their topology. This initial parametrisation can be thought of as defining the rule set governing the how nodes negotiate with all other notes. It is important to note that while an advantage of the generative network modelling approach is to embed regional nodes within a 3D physical space, it is possible that artificial neural networks (which make use of external data) could also be embedded in a physical space. As will become clear later in this thesis, these complementary approaches – generative network models and spatially-embedded artificial neural networks – can be used to study both *embrainment* and *embodiment* respectively. The former in so far as generative

network models help us understand how regions of the brain develop their connectivity in tandem with, and with respect to, other regions of the brain. The later in so far as spatially-embedded neural networks help us understand how networks adapt in context of the environment they are situated in.

The very first instantiation of a generative networks model was designed to probe the development of cortical systems networks by minimising costs alone, starting simulations in medio-posterior locations to enable growth in lateral and anterior directions (Kaiser & Hilgetag, 2004). A competing topological value term was then added to the model, mirroring wiring economy principles, to model human macroscopic functional networks (Vértes et al., 2012). As such, most prominently generative network models have been considered in terms of an extension to the economic trade-off principle, which states that the brain's configuration can be accounted for by the economic balancing of wiring cost minimisation and topological efficiency maximisation as explained previously (Bullmore & Sporns, 2012). This is because the brain is inherently a spatially-embedded *topophysical* network (Sperry et al., 2017) and physical constraints resulting from that embedding underlie functionally important network characteristics. The incorporation of economics into computational models, not only within analyses of connectivity but also computation, is explored in more depth in subsequent chapters. In principle, any range of biophysical constraints can be incorporated within the model, such as the empirical lengths of fibres (Betzel et al., 2016) or genomic co-expression patterns between regions (Arnatkevičiūtė et al., 2021; Oldham et al., 2022). More recent developments have extended the extent to which these models can be tuned to formally compare models (Betzel et al., 2016).

1.4. The generative network modelling framework

I now introduce the broad generative network modelling framework that is used, in part, in **Chapters 2-4** with areas of expansion specifically provided in each chapter. The generative network model can be expressed as a simple wiring equation (Kaiser & Hilgetag, 2004; Vértes

et al., 2012) (**Figure 1.1a**). If you imagine a series of locations within a brain network (e.g., at the level of macroscopic connectivity), at each moment in time the wiring equation calculates which two locations will become connected. It calculates this wiring probability by trading-off the cost of a connection forming, against the potential value of the connection being formed. The equation can be expressed as:

$$P_{i,j} \propto (D_{i,j})^{\eta}(K_{i,j})^{\gamma}, \quad (1)$$

where $D_{i,j}$ represents the Euclidean distance between nodes i and j (i.e., 'costs'), and $K_{i,j}$ reflects the value (i.e., 'attractiveness') in forming a connection. $P_{i,j}$ represents the wiring probability as a function of the product of the parameterised costs and value. The end result is a wiring probability matrix which updates over time as new connections are added.

$D_{i,j}$ is parameterised by the scalar η, which changes how node-to-node distances influence their probability of connecting. For example, when η is negative, wiring probabilities decline when distances increase, and this reflects the costliness of forming connections with nodes that are distant. This is traded-off against $K_{i,j}$, which represents some relationship between nodes, which can be thought of as a topological value (or 'rule') driving the intention for node i to connect with node j. $K_{i,j}$ is parameterised by a distinct scalar γ. $K_{i,j}$ can take a range of different forms and can, in principle, be selected from any non-geometric generative rule used to model social and economic networks (Bell et al., 2017; Borgatti et al., 2009; Kumar et al., 2006). One simple example is the 'matching' rule (Vértes et al., 2012): nodes form connections with other nodes on the basis of their normalised overlap in neighbourhood – i.e., whether nodes are connected to similar nodes to themselves. This is a *homophilic* rule, because the measure is defined in terms of similarity (the Greek *homós*, "same") and preference (*philia*, "liking") to the self. This notion of homophily will be revisited numerous times in this work.

To make this more concrete, imagine the following scenario: a network is growing according to the matching rule, preferentially attaching to nodes which are both similarly connected and

spatially proximal. In the wiring equation, this would be represented as η being negative (e.g., η=-1), $K_{i,j}$ represented as normalised neighbourhoods between nodes (i.e., matching) and its parameter γ being positive (e.g., γ=1). In short, a node being far away makes it less likely that a new connection will be formed, but it having a similar a neighbourhood increases the likelihood. Suppose that the right caudal anterior cingulate (Node 2, n^2) is going to wire to one of its six nearest neighbours. Initially, due to an absent network topology, spatial proximity has a great influence in the formation of new connections – it will wire to its nearest neighbour (**Figure 1.1b**). However, gradually over time, the network's developing structural topology means that $K_{i,j}$ (i.e., the relationships between nodes) may now have a greater influence on wiring probabilities. Indeed, the right caudal anterior cingulate may later wire with a node that, although further away than other available nodes, has a greater value (i.e., matching) than the others (**Figure 1.1c**). As the wiring equation separately parameterise costs and value, the presence of a single connection can heavily influence the topology of the network and thus the future updated wiring probabilities. This is because new connections can lead to entirely new overlapping neighbours, which may include distant nodes. As a result, wiring probabilities can change considerably from moment to moment, despite costs remaining fixed (**Figure 1.1d**).

With models such as the generative network model that aims to instantiate developmental principles, it is important to highlight how the notion of *time* is defined. In the generative network model, time is considered only with respect to iterative processes that occur sequentially in a discrete fashion. This is a very different kind of time that one would read from a clock. In these models, rather than time itself being a meaningful entity, it is just a description of the chronology of how processes occur one after each other. Theoretically, so long as two biological processes do not occur at the precisely exact same "clock" time, these models can still capture the essence of developmental time due to its sequential nature. Importantly, using this notion of iterative processes as a proxy for "clock" time may or not be a useful way of thinking about time in context of biological systems, depending on the particular question a researcher has. In the case of generative models (and likely many other

developmental computational models) this definition is a simplification that still allows us to capture the core essence of our questions. One can still include time-related effects (e.g., by allowing parameters to change over the course of the simulation) but this still relates to the same notion of iterative time as described above.

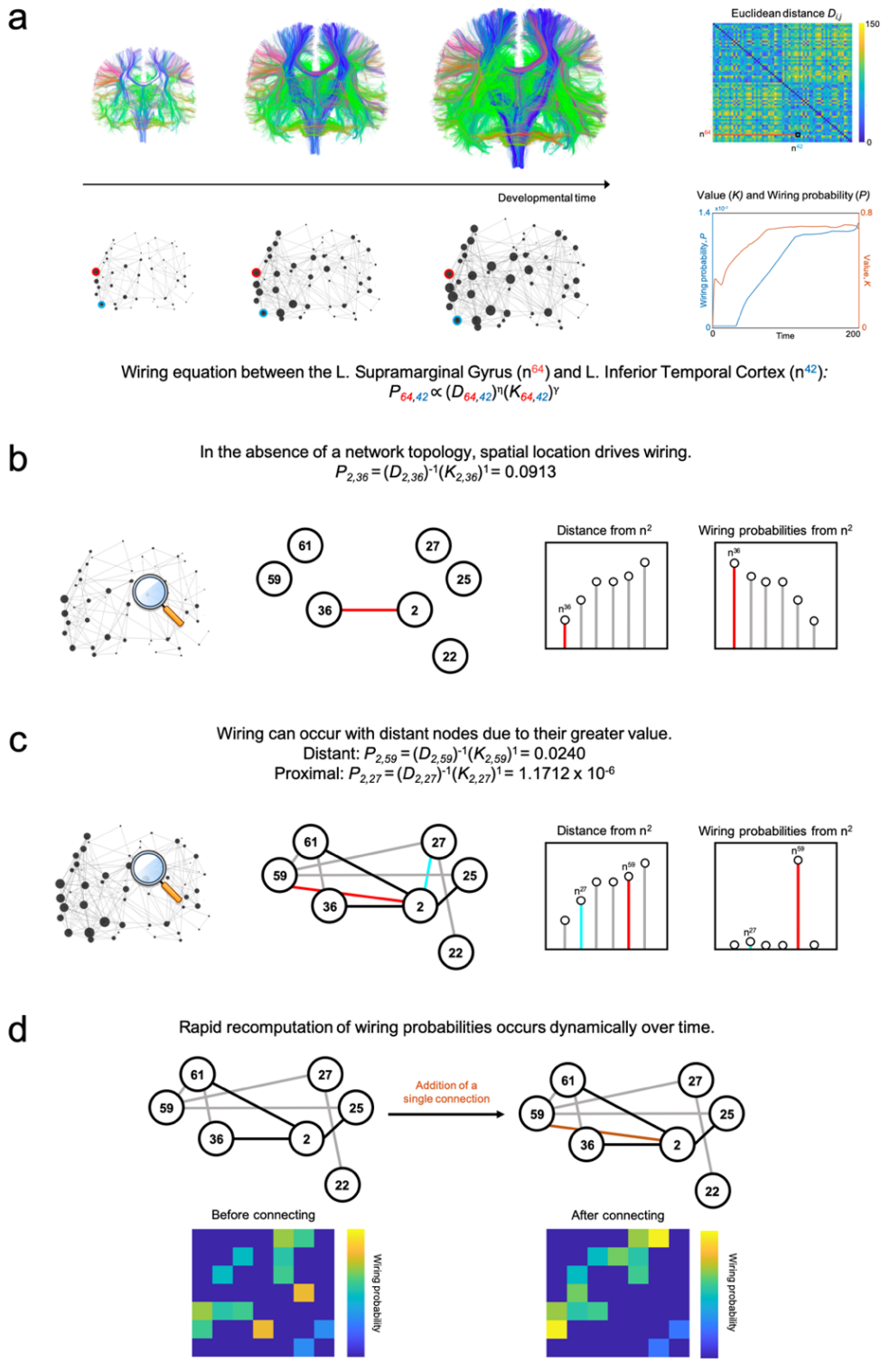


Figure 1.1. Updating wiring probabilities within the generative network model iteratively, based on dynamically changing graphical structures. (a) The brain's structural connectivity is modelled as a generative network which grows over time according to parametrised connection

costs, $(D_{i,j})^{\eta}$ and values, $(K_{i,j})^{\gamma}$. In this illustration, we use subject one's optimal model. (b) Early in network development, the absence of a topology leads to proximal nodes being much more likely to form connections. The displayed distances and probabilities are from the right caudal anterior cingulate (n^2). We display it's six nearest cortical regions. (c) Later, the relative values ($K_{i,j}$) between nodes influence connection probabilities, such that nodes which are more distant (e.g., left rostral anterior cingulate, n^{59} in red) may be preferred to nodes which are closer (e.g., right superior frontal cortex, n^{27} in cyan). (d) As costs and values are decoupled, the wiring probability can be rapidly recomputed when dynamic changes in graphical structure occur over developmental time.

The generative network model is used to simulate network formation under varying wiring rules (**Figure 1.2**), until the overall number of connections matches those found in the observed brain network (of which a range of different observations are compared in this work). Subsequently, to test the accuracy of the simulation, a cost function must be defined which measures the dissimilarity between simulated and observed networks. As alluded to previously, this model fitting step is critical to bridge theory and data, as it is equivalent to evaluating the plausibility of various hypothesises (i.e., the particular wiring equation evaluated).

The first cost function that I introduce here comes in the form of a previously developed energy equation (Betzel et al., 2016), which computes the Kolmogorov-Smirnov (KS) distance between the observed and simulated distributions of a range of network statistics. It then takes the maximum of the four KS statistics considered so that, for any one simulation, no KS statistic is greater than the energy:

$$Energy = \max(KS_k, KS_c, KS_b, KS_e), \tag{2}$$

where KS is the Kolmogorov-Smirnov statistic comparing degree k, clustering coefficient c, betweenness centrality b and edge length e distributions of simulated and observed networks. Minimising energy finds parameters η and γ which generate networks most closely approximating the observed network. Plotting these model fits at varying wiring parameters provides an energy landscape – which tells you how well (and where) the simulation achieves similarity to observations (**Figure 1.2d**).

In principle, any number of statistics can be assessed in a cost function. However, the four measures included in this energy equation are argued to be good candidates for evaluating the plausibility of simulated networks. This is because they are critical statistical properties of realistic networks and have featured within the most well-documented simulated network models (Albert & Barabási, 2002; Barabasi & Albert, 1999; Watts & Strogatz, 1998). Moreover, these statistical properties have been implicated in a number of neuropsychiatric conditions (Gollo et al., 2018; van den Heuvel et al., 2013) in addition to being shown to be heritable (Bohlken et al., 2014). It is important to note, however, that this is just one way to measure the similarity of simulated to observed networks and there are many other possible ways. For example, later in this thesis, an alternative function is introduced which assesses to what extent within-network node-wise statistics are similarly correlated (see Barzel & Biham, 2009 for more detail regarding related approaches). Other algorithmic approaches could in principle be used to assess network similarity, such as using dimensions of statistics that most vary across individual subjects.

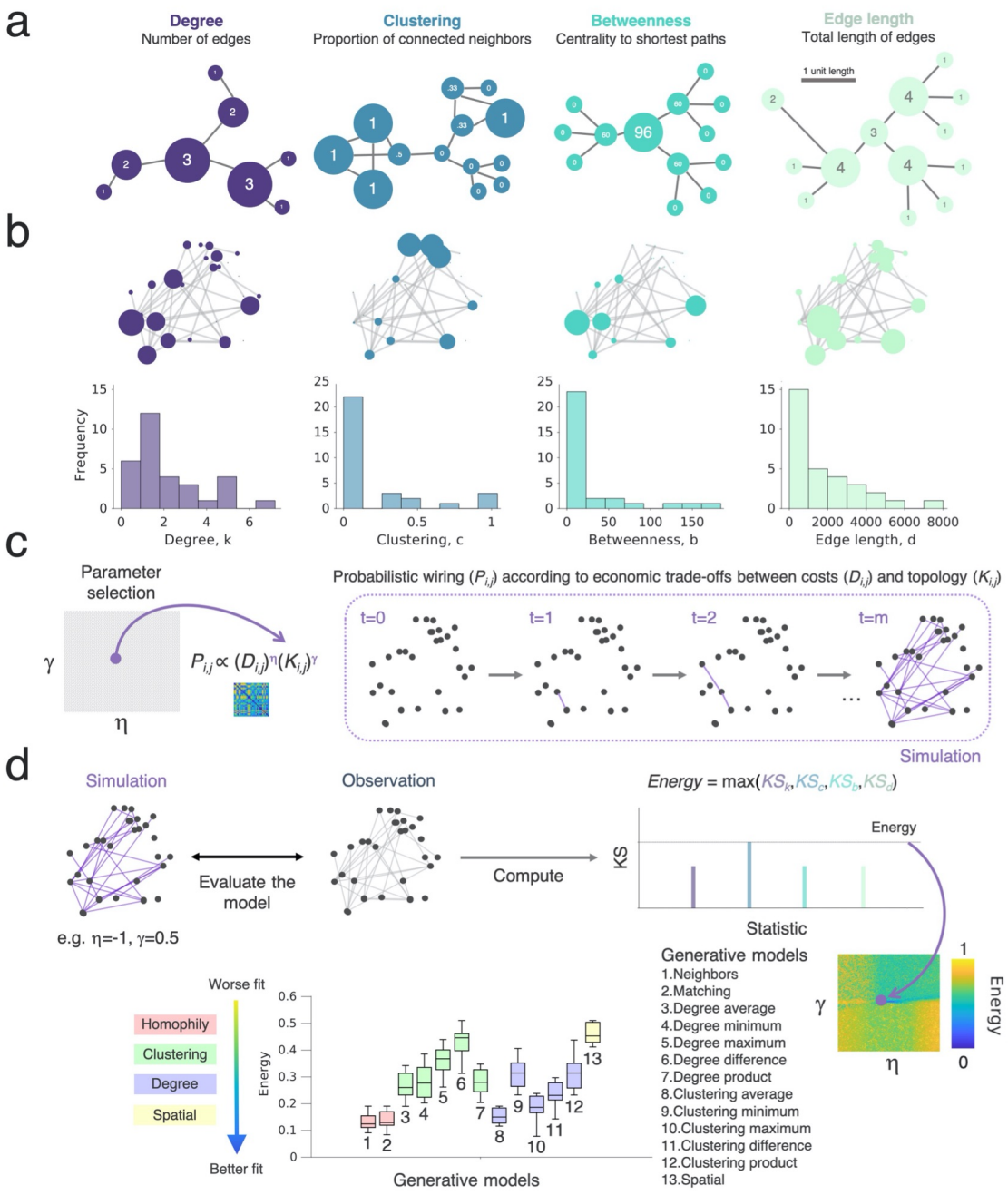


Figure 1.2. Probing wiring principles by simulating topology via generative network models. (a) Example networks highlighting four common graph theoretical metrics. For each schematic, the node size corresponds to the respective graph statistic, and we provide the statistic within the node. The panel on the left shows a schematic for the degree (blue), which relates to the total number of edges that each node has. The left-middle panel shows the clustering coefficient (green), which is a measure of how nodes in a graph tend to cluster together; the right-middle panel shows the betweenness centrality (pink), which relates to the number of shortest paths in the network which pass through the node; the panel on the right shows a schematic for the total edge length (yellow), which relates to the total sum of the edge

lengths among all its connections for each node. (b) Each panel on the top row shows the same network, with the node size corresponding to the respective graph statistic (degree, clustering, betweenness centrality, total edge length). The lower row shows histograms with the distribution of the metric across the network. (c) The generative network model works by simulating network development probabilistically according to the costs incurred in the network (here reflected by the $D_{i,j}$ term, which is modelled as the Euclidean distances) and the topological values. Both terms are parameterised, meaning that for each simulation there is an altered tuning in how costs and values influence the probability with which nodes form connections iteratively. Parameters (η and γ) alter direction and magnitude to which the costs and values respectively influence wiring probabilities. For any combination of wiring parameters, we simulate a network developing over time until it reaches the number of connections, m, equal to the observed empirical network. (d) Comparing the simulated and empirically observed network, we fit the generative network model to achieve a simulation that is the least dissimilar to the observation. This is quantified by the energy equation, and is shown by dark blue in the parameter space. The energy equation is given by the maximum Kolmogorov-Smirnov (KS) statistic across degree, clustering, betweenness centrality and edge length. Each parameter combination is plotted on an energy landscape, which demarcates the model fits across the two-dimensional parameter space. Lower energy values correspond to better fits.

1.5. Thesis structure

Neurocognitive development is not determined, but proceeds probabilistically. As functional processes emerge they themselves start to shape the emergence of complex networks. But how can we capture these processes and test them? **Chapter 2** investigates how variability in structural brain organisation within a large neurodiverse sample of children can be captured by the above described generative modelling framework. **Chapter 3** extends this analysis to investigate how longitudinal *in vitro* neuronal networks can equally be captured by analogous models in rodent primary, human stem cell and cerebral organoid data. **Chapter 4** aims to bridge directly to function by embedding network constraints directly within an artificial neural network model that incorporates structural topological constraints into its optimisation. **Chapter 5** closes by providing key avenues for that may yield fruit in the future.

2. CHAPTER 2

A GENERATIVE NETWORK MODEL OF NEURODEVELOPMENTAL DIVERSITY IN STRUCTURAL BRAIN ORGANISATION

Content included in this chapter was published in Nature Communications, 2021: A generative network model of neurodevelopmental diversity in structural brain organization. ***Danyal Akarca****, Petra E. Vértes, Edward T. Bullmore., the CALM team & Duncan E. Astle.*

The formation of large-scale brain networks, and their continual refinement, represent crucial developmental processes that can drive individual differences in cognition and which are associated with multiple neurodevelopmental conditions. But how does this organisation arise, and what mechanisms drive diversity in organisation? In **Chapter 2**, we use generative network modelling to provide a computational framework for understanding neurodevelopmental diversity. Within this framework macroscopic brain organisation, complete with spatial embedding of its organisation, is an emergent property of a generative wiring equation that optimises its connectivity by renegotiating its biological costs and topological values continuously over time. The rules that govern these iterative wiring properties are controlled by a set of tightly framed parameters, with subtle differences in these parameters steering network growth towards different neurodiverse outcomes. Regional expression of genes associated with the simulations converge on biological processes and cellular components predominantly involved in synaptic signalling, neuronal projection, catabolic intracellular processes and protein transport. Together, this provides a unifying computational framework for conceptualising the mechanisms and diversity in neurodevelopment, capable of integrating different levels of analysis – from genes to cognition.

2.1. Introduction

Macroscopic brain organisation is present early in life, even prenatally (Basser et al., 2000; Conturo et al., 1999), and continues to develop through adolescence and into early adulthood (Mori et al., 1999). The emergence and continual refinement of large-scale brain networks, connecting neuronal populations across anatomical distance, allows for increasing functional integration and specialisation (Friston, 2002; Mahon & Cantlon, 2011). This process is thought crucial for the growth of complex cognitive processes such as language (Chai et al., 2016) and executive function (Assem et al., 2020; Astle et al., 2019; Bathelt et al., 2019; Bathelt, Gathercole, et al., 2018; Bathelt, Holmes, et al., 2018; Fiske & Holmboe, 2019). However, there are individual differences in the trajectories of these networks, and these mirror differences in cognitive development. Indeed, differences in macroscopic networks have been implicated

across multiple neurodevelopmental conditions (Siugzdaite et al., 2020), including ADHD (Griffiths et al., 2016), autism (Hong et al., 2019; Roine et al., 2015), and language disorders (Vydrova et al., 2015).

But what mechanisms govern the growth of macroscopic brain networks? And how do these mechanisms give rise to individual differences in developmental outcomes? There are numerous descriptive theories of typical and atypical neurodevelopment (Castellanos et al., 2006; Krishnan et al., 2016; Rogers & Williams, 2006; Vivanti & Rogers, 2014) that speculate about how different levels of analysis (e.g., genes, brain structure and function) interact to guide development. However, as outlined in **Chapter 1**, to better establish mechanistic links between levels of different observations (e.g., genes, biological pathways, system-wide organisation) we need computational models of development.

We used the aforementioned generative network modelling framework (Betzel et al., 2016; Kaiser & Hilgetag, 2004; Vértes et al., 2012) in an intentionally heterogeneous sample of children (n=270), and: (i) tested which topological features should be valued in the growth trade-off to produce highly accurate individual child connectomes; (ii) tested how small changes in these growth parameters alter the organisational properties of the resulting networks; (iii) established relationships between these different growth parameters and cognitive development; (iv) identified genes with expression profiles that were spatially co-located with those topological features; and (v) established the biological pathways that are enriched in these gene lists. Together, this provides a computational framework that mathematically specifies the growth of a network over developmental time, captures individual differences in brain organisation and cognition, and incorporates the genetic and biological pathways that likely constrain this development.

2.2. Methods

2.2.1. Participants

The sample were made up of children referred by practitioners working in specialist educational or clinical services in the East of England (UK) to the Centre for Attention Learning and Memory (CALM), a research clinic at the MRC Cognition and Brain Sciences Unit, University of Cambridge See Holmes, J. *et al.* (Holmes et al., 2019*)* for the full protocol of assessment (see Astle et al., 2019; Bathelt et al., 2019; Bathelt, Gathercole, et al., 2018; Bathelt, Holmes, et al., 2018 for prior work using the same cohort). The composition of this cohort is design to be broadly reflective of children at heightened neurodevelopmental risk for poor developmental outcomes. Consent was obtained from parents and assent was obtained from all youngsters. The study protocol was approved by, and data collection proceeded under the permission of, the local NHS Research Ethics Committee (reference: 13/EE/0157). This cohort of children is intentionally heterogenous. Referrers were asked to identify children with cognitive problems related to learning, with primary referral reasons including difficulties with ongoing problems in "language", "attention", "memory", or "learning / poor school progress". Exclusion criteria were uncorrected problems in vision or hearing, English as a second language, or a causative genetic diagnosis. Children could have single, multiple or no formally diagnosed learning difficulty or neurodevelopmental disorder. Most referrals were made from Special Educational Needs Coordinators (57.0%), followed by Paediatricians (24.1%) and Speech and Language Therapists (4.2%) (**Figure B1a-b**). Subsequently the CALM team supplemented the cohort with a smaller set of unreferred children, recruited from the same schools and neighbourhoods, so that the cohort captures the full ability spectrum. The CALM cohort contains n=967 total children (n=805 referred; n=162 unreferred). Of these, n=299 undertook MRI scanning of which n=279 had usable MRI data. n=270 of these had cognitive data available (see **Figure B1c** for a visualisation of the cognitive variability across the cohort). This sample includes 65.9% boys, mean age 117.8 months, age range was 66-223 months and 78 that came from the non-referred comparison sample. The increased ratio of boys to girls is what we would expect from epidemiology studies of children at neurodevelopmental risk of poor learning or clinical outcomes (Russell et al., 2014).

To validate our modelling, we also included a second cohort of n=140 typically developing children, who had been recruited from local schools (the RED cohort; mean age 9.34 years, SD age 1.41 years, range 6.82–12.8 years, 45.7% boys). These data were collected under the permission of the Cambridge Psychology Research Ethics Committee (references: Pre.2013.34; Pre.2015.11; Pre.2018.53). Parents/legal guardians provided written informed consent and all children provided verbal assent. This second dataset was previously reported by Johnson, A. *et al.* (A. Johnson et al., 2021). In **Table B1** we provide all demographic information of the cohorts used in the study. Race and ethnicity data for the CALM and RED cohorts are not yet available, but are to be published (Holmes et al., 2019).

2.2.2. MRI ACQUISITION AND PRE-PROCESSING

Magnetic resonance imaging data were acquired at the MRC Cognition and Brain Sciences Unit in Cambridge, on the Siemens 3 T Prisma-fit system (Siemens Healthcare) using a 32-channel quadrature head coil. T1-weighted volume scans were acquired using a whole brain coverage 3D Magnetisation Prepared Rapid Acquisition Gradient Echo (MP RAGE) sequence acquired using 1mm isometric image resolution. Echo time was 2.98ms, and repetition time was 2,250ms. Diffusion scans were acquired using echo-planar diffusion-weighted images with an isotropic set of 68 noncollinear directions, using a weighting factor of b=1,000s$\times$mm^{-2}, interleaved with 4 T2-weighted (b=0) volume. Whole brain coverage was obtained with 60 contiguous axial slices and isometric image resolution of 2mm. Echo time was 90ms and repetition time was 8500ms. Both CALM and RED samples underwent the same scanning protocol. n=299 CALM and n=167 RED children underwent MRI scanning. 20 (CALM) and 27 (RED) scans were not useable due to excessive motion (>3 mm movement during the diffusion sequence estimated through FSL eddy), leaving an MRI sample of n=279 CALM and n=140 RED children respectively.

2.2.3. CONNECTOME CONSTRUCTION AND CORTICAL MORPHOLOGY

MRI scans were converted from the native DICOM to compressed NIfTI-1 format. Next, correction for motion, eddy currents, and field inhomogeneities was applied using FSL eddy. Furthermore, we submitted the images to nonlocal means de-noising (Manjón et al., 2009) using DiPy v0.11 (Garyfallidis et al., 2014) to boost signal-to-noise ratio. A constant single angle (CSA) model was fitted to the 60-gradient-direction diffusion-weighted images using a maximum harmonic order of 8 using DiPy. Whole-brain probabilistic tractography was performed with 8 seeds on all voxels. The step size was set to 0.5 and the maximum number of crossing fibres per voxel to 2. For ROI definition, T1-weighted images were submitted to nonlocal means denoising in DiPy, robust brain extraction using ANTs v1.9 (Avants et al., 2011), and reconstruction in FreeSurfer v5.3 (http://surfer.nmr.mgh.harvard.edu). Regions of interest (ROIs) were based on the Desikan-Killiany parcellation of the MNI template (Desikan et al., 2006) with 34 cortical ROIs per hemisphere. FreeSurfer v5.3 was used for tissue classification and anatomical labelling. The technical details of these procedures are described elsewhere (Dale et al., 1999; Fischl et al., 2002; Fischl & Dale, 2000). FreeSurfer morphology statistics were computed for each ROI.

To construct the connectivity matrix, the number of streamlines intersecting both ROIs was estimated and transformed into a density map for each pairwise combination of ROIs. A symmetric intersection was used so that streamlines starting and ending in each ROI were averaged. Self-connections were removed. To produce binarised connectomes from the resulting 68-by-68 streamline matrix, we enforced an average connectome density of 10% to align with prior research (Betzel et al., 2016), resulting in a streamline threshold of 27 streamlines (i.e., a minimum of 27 streamlines must have connected two regions for us to consider the presence of an anatomical connection). In principle, any number of different network thresholds could be used. Later, we discuss the possible effects of this forking-path.

2.2.4. Generative network models

The broad generative network modelling framework was laid out **Chapter 1.** Starting with a sparse seed network (25 bi-directional edges that were common across all n=270 subjects),

edges were added one at a time over a series of steps until a total number of connections were placed that equalled that of the target observed connectome (group level connections, mean=231.4 and SD=19.1; shown in **Figure B2**). The same process was repeated separately for the validation cohort. Each step allows for the possibility that any pair of unconnected nodes will be connected. Connections are formed probabilistically, where the relative probability of connection formation, between nodes i *and* j, is given by **Equation 1.** We used thirteen previously studied non-geometric rules (Betzel et al., 2016; Vértes et al., 2012) to produce energy landscapes. $D_{i,j}$ was defined as the Euclidean distance between node centroids. $(D_{i,j})^{\eta}$ was computed as a power-law, as shown previously to have better performance than exponentials (Vértes et al., 2012). Euclidean distance accounts for 24% and 22% of the variance in fibre length in the CALM and RED samples respectively. These estimates are relatively low by comparison with other cohorts, with previous studies showing Euclidean distance to account for 32% (NKI; Nathan Kline Institute, Rockland, New York), 66% (CHUV; University Hospital Center and University of Lausanne) and 79% (HCP; Human connectome project) of the variance in fibre length (Betzel et al., 2016). Topological parameters were computed using our own internally developed functions adapted from the Brain Connectivity Toolbox (https://sites.google.com/site/bctnet/) (Rubinov & Sporns, 2010).

To evaluate the fitness of synthetic networks and optimise models, we used an energy function that measures how dissimilar a synthetic network is to the observed network (as defined in Betzel et al., 2016). This was given in **Equation 2** in **Chapter 1**. Initially, we ran simulations across a defined a parameter space of 10,000 evenly-spaced combinations ($-7 \leq \eta \leq 7$, $-7 \leq \gamma \leq 7$), for each generative rule. Parameter ranges were based on approximate values that had been evaluated in previous work (Betzel et al., 2016; Vértes et al., 2012; Zhang et al., 2021). This was to capture their respective energy landscapes and to estimate their relative effectiveness at generating plausible networks. Foreshadowing our findings, we then computed a further a set of 50,000 simulations within a much narrower low-energy window ($-3.606 \leq \eta \leq 0.354$, $0.212 \leq \gamma \leq 0.495$) of the matching algorithm for all subsequent analysis. This is because the matching algorithm attained the lowest-energy networks and therefore best

approximated individual-level connectomes. In part of our subsequent analysis, we then computed the standard deviation of the top n=500 performing (lowest energy) wiring parameters to determine how variable the solutions were.

2.2.5. Cognitive and learning assessments

A large battery of cognitive, learning, and behavioural measures was administered in the CALM clinic (Holmes et al., 2019). $^{n=}$9 CALM children did not have available cognitive data (of the n=279 MRI sample) and were therefore excluded, leaving the final sample n=270 children. These children had no missing data. All cognitive scores were age-standardised, controlling for age (for full details of the processing of cognitive data, see Siugzdaite et al., 2020).

The following nine measures of fluid and crystallised reasoning were included: Matrix Reasoning, a measure of fluid intelligence, Wechsler Abbreviated Scale of Intelligence, WASI (Wechsler, 2011); Peabody Picture Vocabulary Test, PPVT (Dunn & Dunn, 2007). Phonological processing was assessed using the Alliteration subtest of the Phonological Awareness Battery, PhAB (Frederickson et al., 1997). Verbal and visuo-spatial short-term and working memory were measured using Digit Recall, Dot Matrix, Backward Digit Recall, and Mr X subtests from the Automated Working Memory Assessment, AWMA (T. Alloway, 2007; T. P. Alloway et al., 2008). Learning measures (literacy and numeracy) were taken from the Wechsler Individual Achievement Test II, WIAT II (Wechsler, 2005) and the Wechsler Objective Numerical Dimensions, WOND (Wechsler, 1996) apart from 78 of controls for which we used multiple subtests from the Woodcock Johnson for Verbal ability (Woodcock et al., 2001).

2.2.6. Gene expression data

Regional microarray expression data were obtained from six post-mortem brains provided by the Allen Human Brain Atlas (http://human.brain-map.org/) (Arnatkevičiūtė et al., 2019; Hawrylycz et al., 2012) which, as far as we are aware, is the only publicly available human 3D

brain map of the transcriptome which covers the full cortex. Of note, other datasets such as the BrainSpan Atlas for the Developing Human Brain (http://brainspan.org) are available, although currently covering 16 regions of interest.

The Allen Human Brain Atlas dataset is based on microarray analysis of post-mortem tissue samples from six human donors aged between 18 and 68 years with no known history of neuropsychiatric or neurological conditions. Data were imported from open-access published work (Arnatkevičiūtė et al., 2019). Since only two of the six brains included samples from the right hemisphere, analyses were conducted on the left hemisphere only. Probes where expression measures do not exceed the background in >50% samples were removed and genes that did not have a corresponding RNA-seq measure were removed. To improve validity of the gene dataset by cross-validating expression measures, probes with a Spearman's correlation < 0.2 with external RNA-seq data were removed, and a representative probe with the highest correlation to RNA-seq data was selected for each gene. Sample assignment was computed by applying a 2 mm distance threshold. In total, a mean of 37.8±22.5 (SD) samples were assigned to each ROI (min=5; max=92).

The fully pre-processed gene data comprised of a 34 by 10,027 matrix of microarray array gene expression data. These data were used for a subsequent PLS analysis and pathway enrichment analysis.

2.2.7. Statistical analyses

2.2.7.1. PREDICTIONS OF SPATIAL EMBEDDING.

To assess the performance of the optimal matching generative network models to produce networks with spatial embedding of topological characteristics, we averaged across each subject's best performing simulation to produce a single 68 (ROIs) by 1 vector for each measure. We did the same for their observed connectomes. We determined the generative error (i.e., mismatch) for the simulations in terms of each statistical network property within the energy equation at the node-level. This was done as a simple subtraction of the observed from

the simulated network. The absolute mean ranked error was then calculated by taking the modulus of the generative error, averaging across the four statistical properties, and taking the rank order. Like the above described analysis, all network measures were calculated using functions from the Brain Connectivity Toolbox (Rubinov & Sporns, 2010).

We presented group level correlational analysis between wiring parameters and observed global graph theory and cortical morphology measures. In each case, the observed connectome and morphological measures were averaged across the whole cortex.

2.2.7.2. PARTIAL-LEAST SQUARES REGRESSION ANALYSIS

We used partial least-squares (PLS) regression to address two distinct aspects of the study. First, we used PLS to determine the latent components of the wiring equation and connectome features which best explain cognitive task performance. The p_{cov} and p_{corr} significance values of each component were determined by permuting the cognitive data 10,000 times and comparing the observed covariance (p_{cov}) and coefficient of determinations (p_{corr}) relative to their null distributions. Second, we used PLS to identify the linear combinations of genes that best predicted average nodal costs and values each subject's optimal simulation (as outlined previously). For this analysis, subject 162 was removed from as they were the only subject to have an optimally performing η that was positive, which biased results due to being an outlier (as parametrised costs are calculated as $(D_{i,j})^{\eta}$), leaving a sample of n=269. On later examination, this subject's scan contained some motion-related artefacts not captured by prior screening. For each of the n=269 subjects, two PLS analyses were performed, providing 538 separate PLS analyses. We performed permutations of the correlations between average scores and costs/values in the same way as previously to determine significance of the PLS modelling across the sample. To assess the significance of each gene in terms of its loading, we ran n=1000 permutations of the response variable for each PLS. This allowed us to compute a gene loading p_{corr} for each component of the PLS which was collapsed across subjects for gene enrichment analysis.

To determine where variability arises in the growth of the networks, we decomposed the wiring equation for each subject. This was achieved by first running the optimal wiring equation for each subject and taking their cost (a static Euclidean distance matrix), matching and wiring probability matrices at each step in the network growth model. For each subject, we took all edges that existed within the simulation and computed their mean and standard deviation and then determined their coefficient of variation. To then explore within-connectome variability, we performed the same analysis but collapsing across subjects to determine how nodes (summed rows of the matrix) and edges (elements of the matrix) vary.

2.2.8. Gene enrichment analysis and visualisation

We next aimed to elucidate the biological processes (BPs) and cellular components (CCs) for which our gene lists converged on. A BP is defined as representing a specific objective that the organism is genetically programmed to achieve. A BP is accomplished by a particular set of molecular functions carried out by specific gene products (or macromolecular complexes), often in a highly regulated manner and in a particular temporal sequence (https://www.ebi.ac.uk/QuickGO/term/GO:0008150) On the other hand, a CC is defined as a location, relative to cellular compartments and structures, occupied by a macromolecular machine when it carries out a molecular function. There are two ways in which the gene ontology describes locations of gene products: (1) relative to cellular structures (e.g., cytoplasmic side of plasma membrane) or compartments (e.g., mitochondrion), and (2) the stable macromolecular complexes of which they are parts (e.g., the ribosome) (https://www.ebi.ac.uk/QuickGO/term/GO:0005575).

To elucidate BPs and CCs across the sample, genes with a $p_{corr}<0.05$ following permutation testing on each component were deemed significant. This provided an individual-level vector of genes that were significant for an individual for each of the nodal costs and nodal values PLS1. To collapse across subjects, genes were then ordered according to their frequency in being significantly associated with connectome growth across subjects for that component. The list stopped when genes were significant for less than 10% of the sample. For each subject,

PLS1 provided an average of 581.5 significant genes (SD 101.4) for nodal costs and 437.6 significant genes (SD 167.4) for nodal values. When collapsed across subjects as described, the nodal costs PLS1 had 1427 genes and the nodal values PLS1 had 1584 genes ordered in terms of importance, which were then submitted to a pathway enrichment analysis (**Figure B3**).

For all information as to the enrichment and visualisation pipeline, please refer to Reimand, J. *et al.* (Reimand et al., 2019). In short, gene ontology (GO) annotations are the most commonly used resource for pathway enrichment analysis. g:Profiler (https://biit.cs.ut.ee/gprofiler/gost) searches a collection of gene sets representing GO terms and, in the ordered test, repeats a modified Fisher's exact test on incrementally larger sub-lists of the input genes and reports the sub-list with the strongest enrichment. Multiple-test correction is applied to produce an adjusted p value (p_{adj}) (Reimand et al., 2007, 2019). To visualise enriched pathways, we used 'EnrichmentMap' within Cytoscape v3.8.0 (http://www.cytoscape.org) (Reimand et al., 2019; Shannon et al., 2003). All default parameters were used. Pathways are shown as nodes (representing enriched biological processes) that are connected by edges if the pathways share genes. Nodes are coloured by their p_{adj} and edges are sized on the basis of the number of genes shared by the connected pathways. To then identify clusters of themes, AutoAnnotate v1.3.3 was used before manually curating the suggested theme names to accurately reflect all pathways within each theme.

2.3. Results

2.3.1. Small variations in parameter combinations produce accurate spatially-embedded networks

From a basic seed network common to all participants (**Figure B2**), we computed the subject-wise optimal generative network model (i.e., network with lowest energy) over a range of 10,000 evenly-spaced parameter combinations ($-7 \leq \eta \leq 7$, $-7 \leq \gamma \leq 7$), using thirteen different generative rules (for rule formulae, see **Table B2**) across our large sample of children (Holmes et al., 2019). In each case, we computed energy landscapes to contextualise how they perform

(**Figure 2.1a-d**). Mirroring findings in adult samples (Betzel et al., 2016; Vértes et al., 2012; Zhang et al., 2021), we found that models driven by geometry and topology outperform the pure geometric spatial model and homophily-based models achieve the lowest energy for our paediatric sample (**Figure 2.1e**). In other words, when one combines the distance penalty with the 'matching rule' we described in **Chapter 1**, it produces the most plausible simulated brain networks. This difference between generative rules is extremely robust. A post-hoc power calculation revealed that the homophily based rules could be distinguished from the next best class with near-perfect statistical power (t=-10.210, p=6.705x10^{-21}; n=270, power>0.99), and that this difference could be detected with around 70 participants. Due to the greater diversity contained within our sample relative to most studied samples, other studies employing similar methods may require more samples to detect this effect.

It is notable that across the matching energy landscape, these plausible networks exist within an extremely narrow window of parameter solutions. That is, as a proportion of the parameter space, the matching rule (and the other homophily-based model 'neighbours') contain the least number of low-energy networks relative to other rules. But as **Figure 2.1e** shows, these networks are the closest to real networks. Thus, varying homophily-based parameters produces the most realistic networks, yet has the lowest variability in the space (**Figure B4**).

While small, variability within this narrow matching window determines inter-individual differences in brain network growth. This is because small changes in parameters (i.e., the magnitude and direction in which costs and values influence wiring probabilities) can lead to networks which are diverse yet include basic structural properties common to all subjects. To derive more precise estimations of optimal generative parameter combinations, we subsequently generated a new set of 50,000 evenly-spaced simulated networks over this narrow low-energy matching window ($-3.606 \leq \eta \leq 0.354$, $0.212 \leq \gamma \leq 0.495$). Focusing on this energy crevasse allows us to detect individual differences in optimal parameter combinations with much greater specificity. In other words, we resampled the parameter combinations focusing

within the low energy window, to make sure we have the most precise estimate of each individual child's optimal parameters.

In **Figure 2.1f**, we show the spatial distribution of these top performing parameter combinations and **Table B3** documents their summary statistics. These finely calibrated networks are even more low energy than in the previous analysis. In **Figure B5** we detail how KS statistics vary across the same space. Importantly, due to the stochastic nature of generative network models, the energy of optimal parameter combinations varies with an average standard deviation (SD) of 0.045 across the sample (1000 independent runs). Therefore, for the rest of this study, we quote our parameter analyses averaged across a variable number of wiring parameters which achieved networks with the lowest energy in the space: n=1 (equating to 0.002% of the space) n=10 (0.02%), n=100 (0.2%) and n=500 (1.0%).

The optimal wiring parameters are significantly negatively correlated with each other, such that subjects with large γ parameters tend to have larger negative η (best n=1 network: r=0.284, p=2.07x10^{-6}; n=10 networks: r=-0.403, p=6.08x10^{-12}; r=-0.460, p=1.58x10^{-15}; n=100 networks: r=-0.460, p=1.58x10^{-15}, n=500 networks: r=-0.497, p=3.21x10^{-18}) (**Figure B5f**). Optimally simulated networks, using this simple wiring equation, are so similar to the actual networks that a support vector machine is unable to distinguish them using the parameters from the energy equation (mean accuracy=50.45%, SD=2.85%).

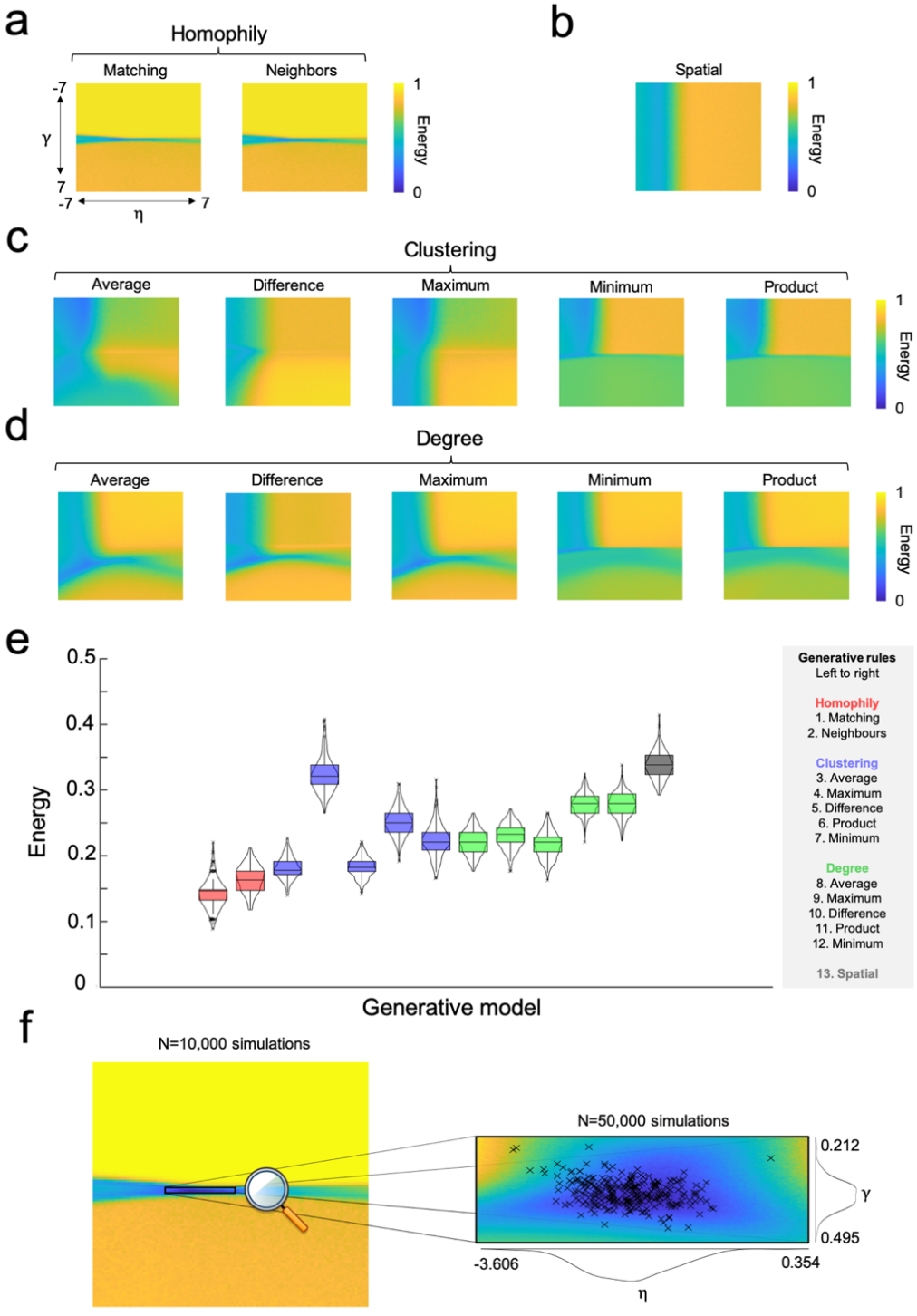


Figure 2.1. Sample-averaged energy landscape visualisation and generative rule comparisons. (a) Homophily based methods. Matching and neighbours algorithms calculate a measure of shared neighbourhood between nodes. (b) The spatial method. This ignores entirely, judging networks only on the basis of their spatial relationship. (c) Clustering based methods. These calculate a summary measure between two nodes in terms of their clustering coefficients. (d)

Degree based methods. These calculate a summary measure between two nodes in terms of their degree. (e) Energy statistics from the best performing simulation across 13 generative rules, showing that matching can achieve the lowest energy networks given the appropriate parameter combination. In total, there are n=270 data points for each of the 13 boxplots. The boxplot presents the median and IQR. Outliers are demarcated as small black crosses, and are those which exceed 1.5 times the interquartile range away from the top or bottom of the box. (f) A further 50,000 simulations were undertaken in the refined matching window, as these defined boundary conditions for which low-energetic networks were consistently achieved. Each cross represents a subject's individually specific wiring parameters that achieved their lowest energy simulated network.

Replicating previous work, we find that our simulated networks, optimised via the statistical properties included in the energy equation via homophily generative mechanisms, accurately capture these properties in observed networks (Betzel et al., 2016; Vértes et al., 2012; Zhang et al., 2021). But do these capture crucial network properties not included in the energy equation, like their spatial embedding? We next examined if the spatial patterning of these network properties arises simply from the generative model.

Averaged across the sample, optimally performing generative models (i.e., those using the 'matching' rule) produce networks which significantly correlate with observed networks in terms of their degree (r=0.522, p=4.96x10^{-5}), edge length (r=0.686, p=1.11x10^{-11}) and betweenness centrality (r=0.304, p=0.012) but not clustering coefficient (r=-0.054, p=0.663) (**Figure 2.2**). That is, the spatial embedding of these network properties seemingly emerges, to mirror those of the observed networks, despite this not being specified in the growth process. We extended this analysis to new measures outside of the energy equation (**Figure B6**). While local efficiency and assortativity cannot be significantly predicted across the sample (r=0.211, p=0.084 and r=-0.096, p=0.116 respectively), optimally performing simulated and observed networks correlate positively in terms of their global number of rich clubs (r=0.316, p=1.11x10^{-7}), maximised modularity (r=0.349, p=3.84x10^{-9}) and transitivity scalar (r=0.411, p=2.11x10^{-12}). In short, despite not being specified in the growth process, the simple homophily rule generates many properties of observed brain networks.

One criticism of our simulations is that their embedding may be an artefact of the seed network (which is on average 10.8% the density of the observed/simulated networks). In short, if by chance the seed network mirrors the final network, it could be inevitable that spatial embedding would emerge, in terms of node degree, betweenness centrality and edge length. If so, one would expect initial local statistical properties of the seed to be significantly associated with regional accuracy of the simulation. To determine if this is the case, we analysed the regional accuracy of our homophily simulations by determining their regional generative error (as a mismatch between observed and simulated outcome, depicted in **Figure B7**). Importantly, the average ranked error (**Figure B7e**) is not correlated with the seed network's connectivity (r=-0.0711, p=0.5643). Furthermore, seed features do not correlate with their own feature's resultant error (degree, r=-0.0408, p=0.7410; betweenness, r=0.1833, p=0.1345; edge length, r=0.1114, p=0.3659). In **Chapter 5** we elaborate on seeding in context of the wider literature.

The large heterogenous sample we chose is ideal for this computational approach to understanding diversity, but it is highly likely that this approach will work for more standard typically developing cohorts of children. In **Figure B8** we replicate our key findings in an independent sample of n=140 children recruited from local primary schools in the same area (A. Johnson et al., 2021).

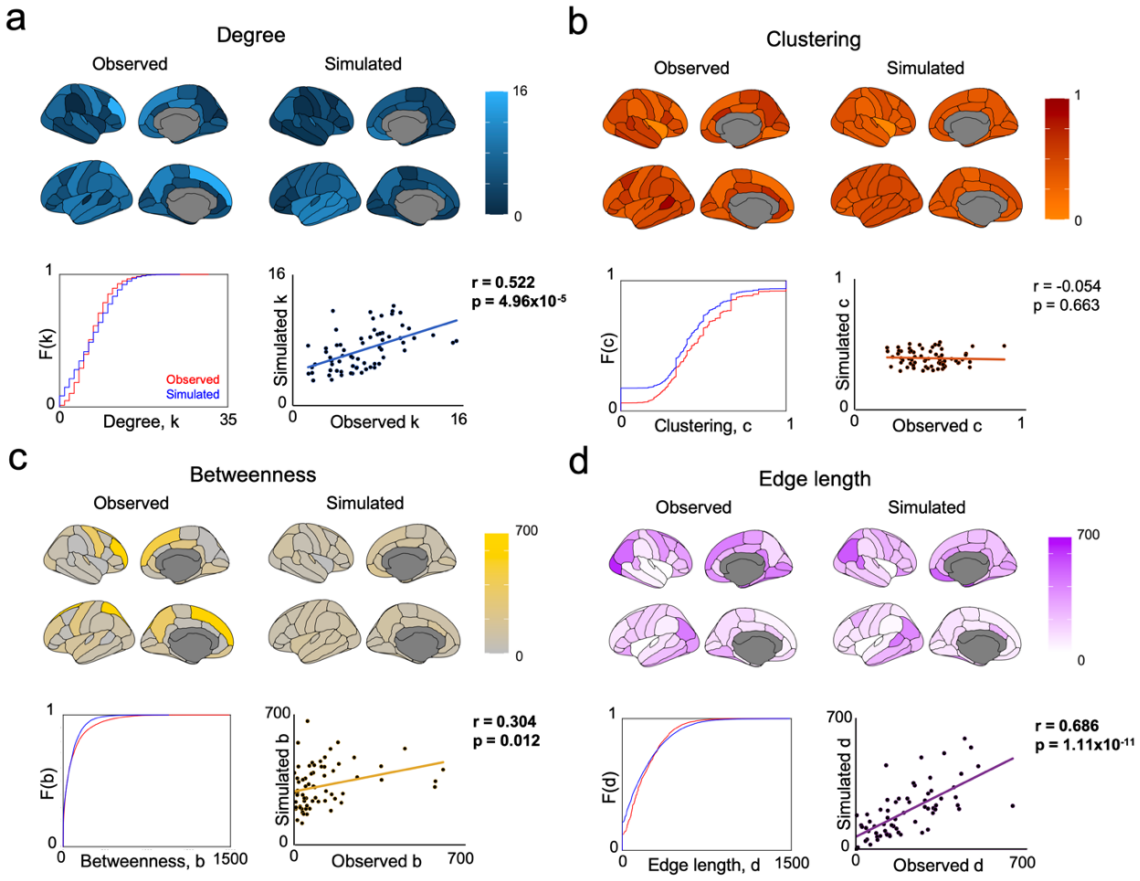


Figure 2.2. Spatial embedding of simulated networks, grown via optimised homophily generative mechanisms. For each network measure, we present the cumulative density functions across all observed versus simulated nodes within each network. Each point in the scatter plot shows one of the 68 across-subject average nodal measures from the observed and optimally simulated networks. We also show a visualisation of these measures. All statistics were computed via two-tailed linear correlations, quoting the Pearson's correlation coefficient. (a) Degree between observed and simulations are significantly positively correlated (r=0.522, p=4.96x10^{-5}). (b) Clustering between observed and simulations at not correlated (r=-0.054, p=0.663). (c) Betweenness centrality between observed and simulations are significantly positively correlated (r=0.304, p=0.012). (d) Edge length (as a summation of all edges from each node) between observed and simulations are significantly positively correlated (r=0.686, p=1.11x10^{-11}). Boldened values are significant correlations at p<0.05. In Figure B6, we present a parallel analysis including local and global measures not included in the energy equation. In

Figure B7, we demarcate for each measure the generative error in spatial embedding, and show the ranked performance for each region.

2.3.2. Individual differences in wiring parameters mirror connectome organisation, grey matter morphology and cognitive scores

A critical benefit of a generative modelling approach is that it allows us to probe the underlying mechanisms occurring over the development of the network (Betzel & Bassett, 2017b). As one explicitly specifies the generative mechanisms involved, the statistical properties which "fall out" of the network can be considered as spandrels (Rubinov, 2016); epiphenomena of the network's development according to the much simpler economical trade-off (in this case, according to the homophily principle). If the generative model is indeed capturing biologically relevant processes, one would expect the wiring equation – at a minimum – to reduce the network's dimensionality simply into the two wiring parameters used to construct it. Under this view, individual differences in wiring parameters should (i) map to wide-ranging statistical properties of the observed network, which may be considered spandrels and (ii) appropriately reduce dimensionality of the connectome such that, for example, one can equivalently predict cognitive scores from parameters as one would be able to from the network properties.

To explore this, we first examined how wiring parameters reflect observed features of brain organisation by quantifying how a subject's and relate to global measures of their observed connectome. Furthermore, for all 270 subjects, cortical morphology data were available. In **Figure 2.3**, we document how global network and morphological measures (most of which are not included in the energy equation) relate to each other, in addition to their reasonably stable association with a varying number of high performing wiring parameters and their associations with age. **Figure 2.3b** shows that η is significantly associated with age – the network parameters needed to form optimal networks over time need to favour longer distance connections for older participants, relative to younger participants (Betzel et al., 2016). To disentangle age-related parameter differences from individual differences, we repeated all of

our correlations across measures whilst partialling out age (**Figure B9**). Associations remain when age has been controlled for, demonstrating that age-related changes in optimal parameters are relatively independent of the individual differences in those parameters. This is not only an important step in demonstrating that these parameters generalise to distinct measures (e.g., morphological observations) not used to train the generative models, but also demonstrates that the generative approach is consistent with the notion that wiring parameters themselves have significant associations with numerous statistical properties in a network.

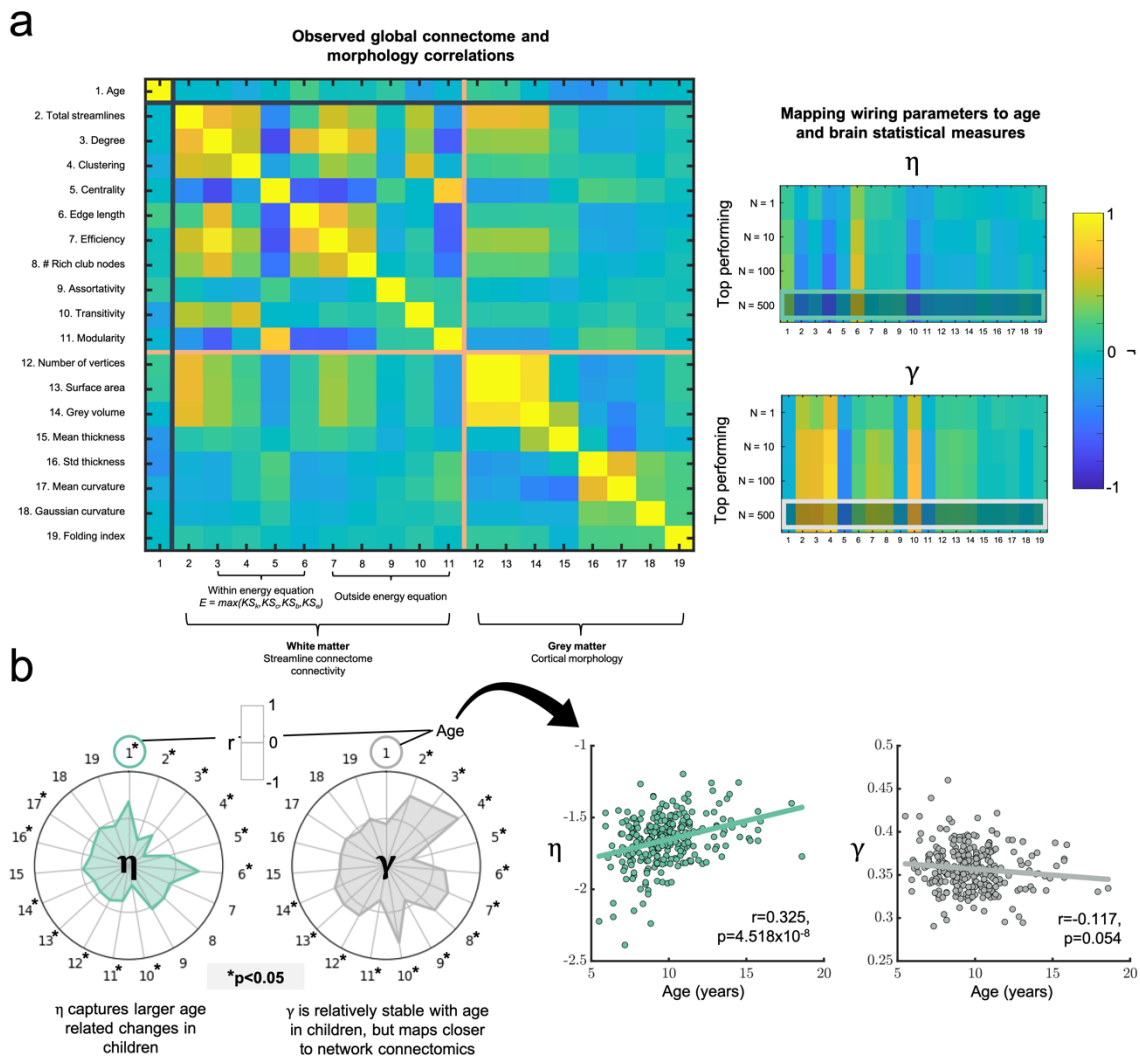


Figure 2.3. Statistical properties of the connectome and cortical morphology, and their relationships with wiring parameters and age. (a) The correlation matrix of connectome and morphological findings show how each measure correlates with every other measure. Measures 3 – 6 were included in the energy equation. Measures 7 – 11 are connectome measures not

included in the energy equation. Measures 12 – 19 are cortical morphological measures. and are each significantly correlated with a range of measures, both inside and outside of the energy equation. Correlation coefficient matrices are shown, the bottom row of which is highlighted and is reflected in the above radar plots (middle), in addition to the significance matrix (bottom), across varying numbers of top performing parameters, for each of the 19 measures investigated. (b) Radar plots depict the correlations between all measures and (left) and (right) averaged across the top n=500 parameters in the parameter space. All statistics were computed via two-tailed linear correlations, quoting the Pearson's correlation coefficient. The asterisk, *, reflects significant correlations at $p<0.05$. Note, the inner edge of the radar plot reflects negative correlations and the outer edge reflects positive correlations. Further scatter plots are provided highlighting the relationship of wiring parameters with age. η has a significantly positive relationship with age (r=0.325, p=4.518x10^{-8}) while γ has a weak non-significant negative relationship with age (r=-0.117, p=0.054).

Next, we tested the ability of the wiring equation to reduce the dimensionality of the connectome. Specifically, if wiring parameters are accurate decompositions of an individual's structural network they should predict cognitive outcomes equivocally to observed features of the connectome. For all 270 subjects we had data from a battery of cognitive tasks, including measures of executive function, phonological awareness, working memory, fluid reasoning and vocabulary.

We tested the relationship between a subject's age-standardised cognition and (i) their optimal wiring parameters and (ii) global measures of their structural connectome (using measures included in the energy equation). This was done using partial least squares (PLS), a multivariate statistical technique which extracts optimally covarying patterns from two data domains (Abdi, 2010). We undertook two separate PLS analyses, which correlated (i) optimal wiring parameter combinations or (ii) global connectome measures across our sample, with cognitive performance in the nine tasks, respectively (**Figure 2.4a**). For both analyses, PLS1 was significant in the amount of explained covariance (p_{cov}=0.009 and p_{cov}=0.049 respectively). PLS1 score predictions, and their cognitive loadings, are extremely similar between wiring parameters and connectome features (r=0.191, p=1.63x10^{-3}, p_{corr}=7x10^{-4} and r=0.210, p=5.29x10^{-4}; each p_{corr}=7x10^{-4} and p_{corr}=4x10^{-4} respectively from 10,000 permutations of scores) (**Figure 2.4b-c**).

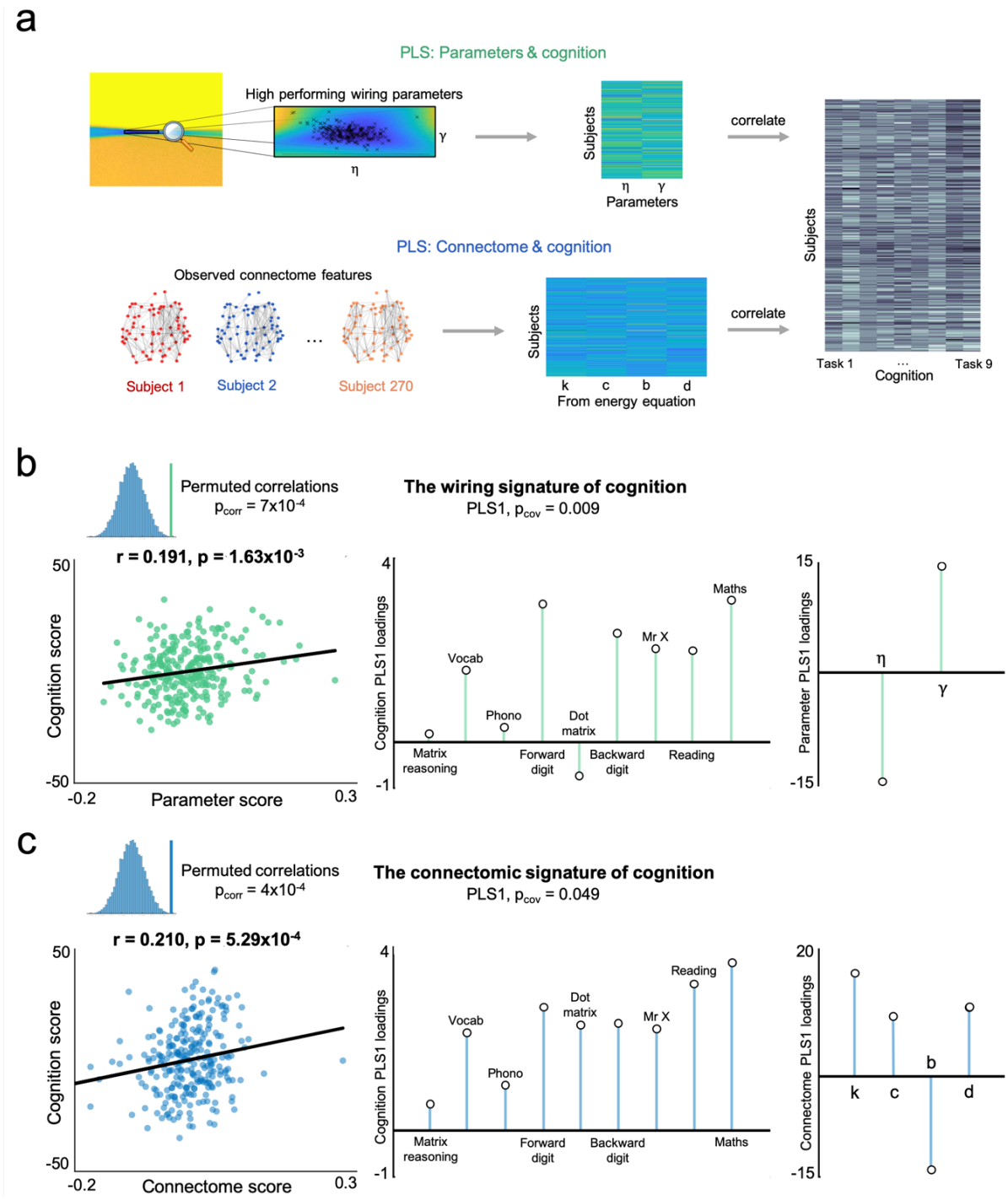


Figure 2.4. Covarying patterns of wiring parameters and connectome features with cognitive performance across nine cognitive tasks. (a) A visual representation of the two PLS analyses undertaken. (b) There is a significant positive correlation (two tailed linear correlation, quoting the Pearson's correlation coefficient) between parameter scores and PLS-derived cognitive scores. PLS1 was statistically significant (p_{corr}=7x10^{-4} and p_{corr}=4x10^{-4} respectively) for both analyses using n=10,000 permutations. Each parameter loads with similar magnitude onto

PLS1. (c) There is an analogous significant positive correlation between connectome scores and PLS-derived cognition scores, using the same statistical procedure.

2.3.3. Variability in neurodevelopmental trajectories arises through value-updating over time

While small generative parameter differences result in differential network properties, we have yet to show how this variability may occur over the development of the networks. That is, how do differences in parameter combinations across subjects manifest themselves when the network is developing? To address this, we examined how between-subject variability in optimal generative network models emerge at the level of cortical nodes and their connections. This is possible by simply decomposing the optimal simulation into its constituent parametrised costs $(D_{i,j})^{\eta}$, values $(K_{i,j})^{\gamma}$ and wiring probabilities $P_{i,j}$ at each time point, for each subject (**Figure 2.5a-b**). This allows us to quantify growth trajectories and thus establish which aspects of network emergence vary most in the sample.

For each subject, we computed the coefficient of variation (CV, $\frac{\sigma}{\mu}$ of their parameterised costs, matching values and wiring probabilities to compare subject-specific variability, as it emerges throughout the simulated growth of connectomes. While subjects exhibit some variability in how parameterised costs influence wiring probabilities (mean CV 2.27), this is dwarfed by their parameterised values over time (mean CV 33.02). This is because the matching value is dynamic, changing at each iteration, unlike relative Euclidean distance between nodes which is static. The result is that significant inter-individual variability arises in the probability of connections forming (mean CV 53.08), leading to the emergence of divergent brain organisation (**Figure 2.5c-f**). Furthermore, the regional patterning of costs and values is not random (**Figure 2.5g**). Nodes and edges with high matching values decline in their variability, suggesting a consistency across subjects in highly 'attractive' nodal structures and their connections. Across the sample, cheaper regions occupy the medial aspects of the cortex while highly valuable regions generally reside in the temporal cortex.

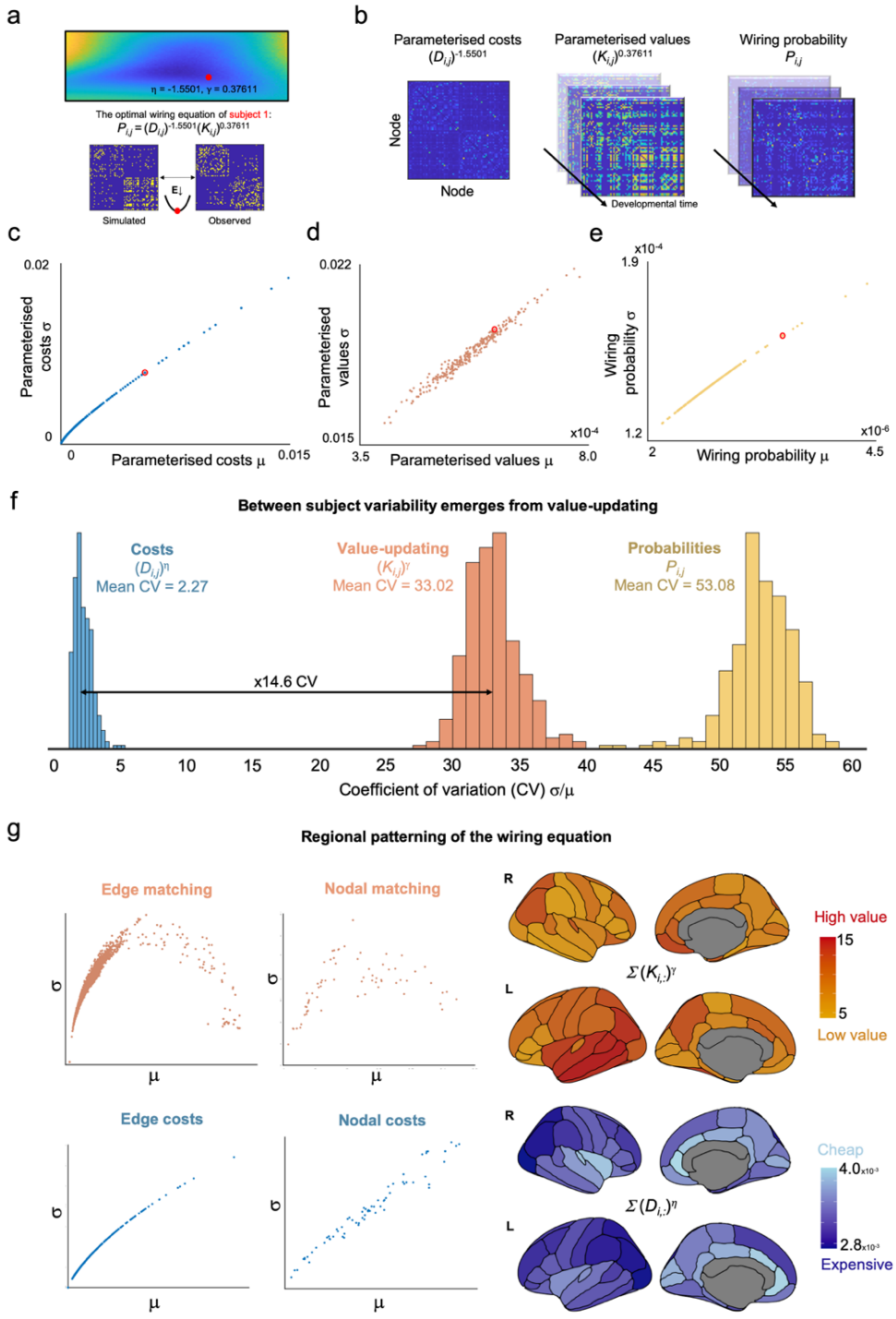


Figure 2.5 Wiring equation decomposition and the subsequent variability across subjects in our heterogeneous sample. (a) For each subject, a simulated network is produced by minimising the energy between the observed and simulated network. Here we present visualisations for subject one (red). (b) Costs $D_{i,j}$ are static, while values $K_{i,j}$ dynamically update according to the matching rule, which enables the computation of wiring probability $P_{i,j}$. (c) The mean and standard deviation for each subject of their edge-wise parameterised

costs, (d) parameterised values and (e) wiring probabilities. (f) Histograms of each subject's coefficient of variation (CV) showing that subjects are more variable in their value-updating compared to costs, leading to large wiring probability variability. (g) Regional patterning of sample-averaged nodal parameterised costs and values, showing highly "valuable" patterning in the left temporal lobe and "cheap" regions generally occupying medial aspects of the cortex. Variability declines as value increases, but increases for costs.

2.3.4. Genomic patterning of network growth

Underlying these macroscopic changes in brain organisation across time are a series of complex molecular mechanisms. These are partly governed by genetically coded processes that vary across individuals. We next tested whether these processes may steer the brain network towards a particular growth trajectory within our generative network models.

Nodal cost and nodal 'matching' value patterning alongside regional gene expression profiles of 10,027 genes using human adult brain microarray data (Arnatkevičiūtė et al., 2019; Hawrylycz et al., 2012) were integrated into two PLS analyses for each subject. For all analyses, gene expression scores at each node were used as the predictor. For each subject's first analysis, their parameterised nodal costs (calculated as the subject's $(D_{i,j})^{\eta}$, as visualised **Figure 2.5g**; fourth row, right) was used as the response variable. For each subject's second analysis, their mean parameterised values (calculated as the subject's $(K_{i,j})^{\gamma}$ averaged over time, as visualised in **Figure 2.5g**; second row, right) was used as the response variable. Each analysis defined PLS components independently which were linear combinations of the weighted gene expression scores at each node (predictor variables) that were most strongly correlated with the subject's nodal costs and nodal values of their simulated growth trajectory. To limit the variability across regions in terms of the samples available, only left hemispheric gene data were analysed (Arnatkevičiūtė et al., 2019).

Across our sample, the first PLS component (PLS1) explained on average 65.0% (SD 1.3%) and 56.9% (SD 9.2%) of the covariance between genetic expression and nodal-costs, and nodal-values respectively. The average nodal-costs PLS1 score significantly correlates with average nodal-costs (r=0.794, p=2.07x10^{-8}, p_{corr}=2x10^{-4}). Similarly, the average nodal-values PLS1 score

significantly correlates with average nodal-values (r=0.718, p=1.71x10^{-6}, p_{corr}=1x10^{-4}) (**Figure 2.6a-b**). To then characterise the genetic profiles associated with each PLS analysis, we permuted the response variable 1000 times to form a null distribution for the loading of each gene, across each subject's PLS1. This provides an estimate of how strong the loading would be by chance, and thus which genes exceed p_{corr}<0.05. Across subjects, PLS1 provided an average of 581.5 significant genes (SD 101.4) for nodal costs and 437.6 significant genes (SD 167.4) for nodal values (**Figure B3a**).

Genes do not act in isolation, but instead converge to govern biological pathways across spatial scales. To move from individual genes to biological processes (BPs) and cellular components (CCs), we performed a pathway enrichment analysis (Reimand et al., 2019). Pathway enrichment analysis summarises large gene sets as a smaller list of more easily interpretable pathways that can be visualised to identify main biological themes. Genes were ordered according to their frequency in being significantly associated with connectome growth across subjects for that component. For example, for nodal-values PLS1, top of the list was the gene associated with connectome growth in the most subjects (*CHI3L1*; significant for 49.4% of our sample), the next was the second most frequent gene (*PRKAB2*; 36.4% of our sample) and so on. Our list stopped when genes were significant for less than 10% of the sample. This left the nodal-costs PLS1 with a list of 1427 genes and the nodal-values PLS1 with a list of 1584 genes ordered in terms of importance, which were submitted to pathway enrichment analysis in g:Profiler (https://biit.cs.ut.ee/gprofiler/gost) (Reimand et al., 2019) (**Figure B3b**). g:Profiler searches a collection of gene sets representing GO terms. In the ordered test, it repeats a modified Fisher's exact test on incrementally larger sub-lists of the input genes and reports the sub-list with the strongest enrichment. Multiple-test correction is applied to produce an adjusted p value (p_{adj}) for each enrichment (Reimand et al., 2007, 2019) as visualised in **Figure B3c-d**, which can be accessed via the links presented in **Table B4**.

The genes identified within the subject-wise PLS are not random, but instead converge on particular biological processes (BP) and cellular components (CC). The nodal costs PLS1 was

most prominently enriched for genes associated with biological processes including catabolic processes and protein localisation (32 BPs; all $p_{adj}<9.58x10^{-3}$), cell projection (14 BPs; all $p_{adj}<4.39x10^{-2}$), immunological processes (34 BPs; all $p_{adj}<4.82x10^{-2}$), regulation of metabolic processes (8 BPs; all $p_{adj}<4.75x10^{-2}$) and regulation of cell development and differentiation (4 BPs; all $p_{adj}<3.87x10^{-2}$). In terms of cellular components, nodal costs PLS1 was enriched for genes associated with the ribosome (14 CCs; all $p_{adj}<2.15x10^{-2}$), vesicular and endoplasmic membranes (19 CCs; all $p_{adj}<4.90x10^{-2}$) and intracellular organelles (8CCs; all $p_{adj}<4.97x10^{-2}$) (**Figure 2.6c**).

The nodal values PLS1 was most prominently enriched for genes associated with biological processes including synaptic signalling (29 BPs; all $p_{adj}<3.96x10^{-2}$), neuronal projection and development (26 BPs; all $p_{adj}<4.21x10^{-2}$) and synapse organisation (2 BPs; all $p_{adj}<2.92x10^{-2}$). In terms of cellular components, nodal values PLS1 was enriched for genes associated with synaptic membranes (60 CCs; all $p_{adj} < 3.15x10^{-2}$) and ion channel complexes (7 CCs; all $p_{adj}<1.18 \times 10^{-2}$) (**Figure 2.6d**).

In **Table B4** we provide links so that readers can run our precise gene ontology queries within a browser and in **Figure B3c-d** we show a visualisation of these enriched gene sets.

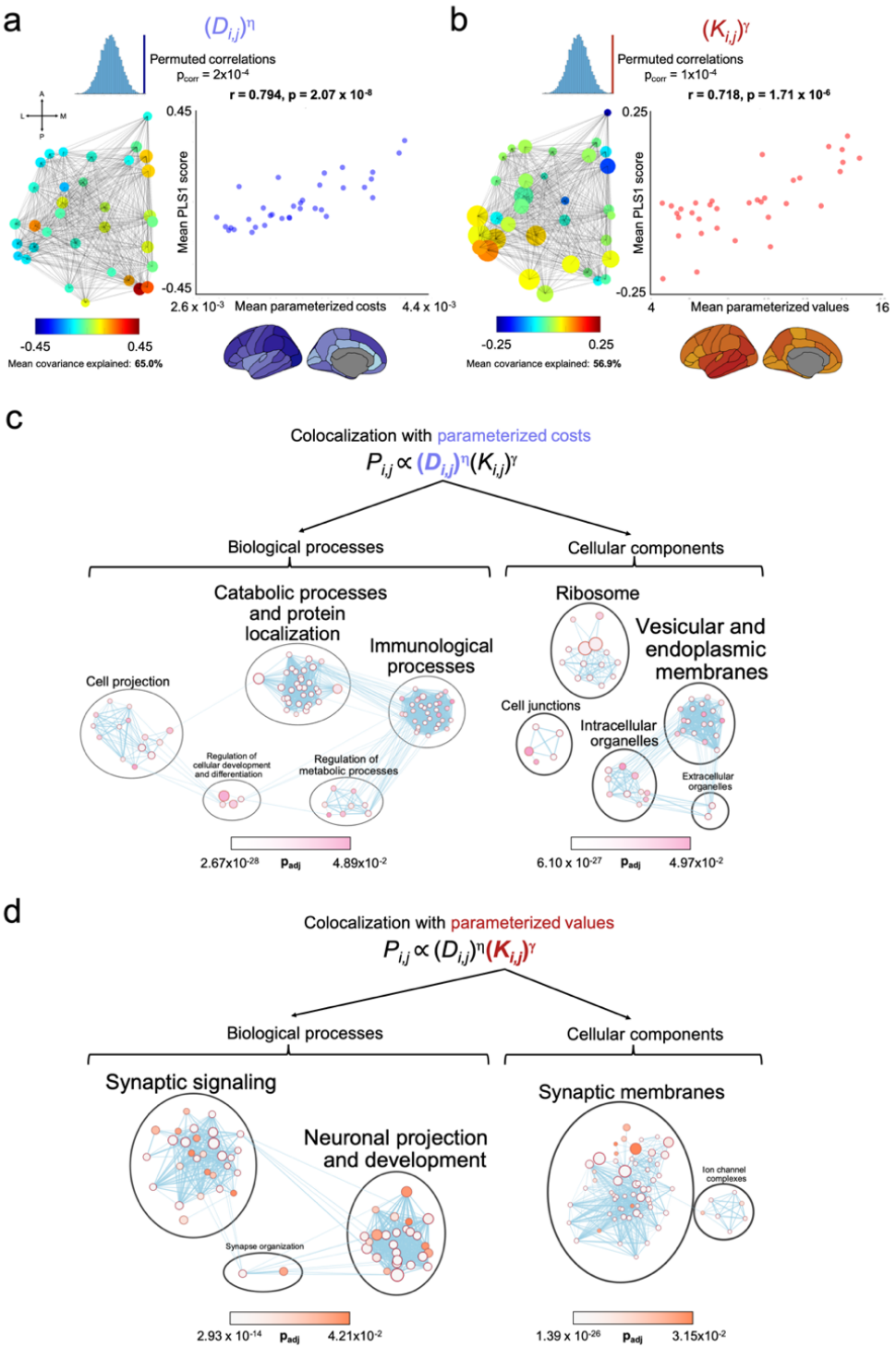


Figure 2.6. Over expressed genes which explain variance in brain wiring across subjects. Both PLS1 components across subjects are enriched for functionally specific biological processes and cellular components. Node size represents the number of genes in the set. The edges relate to gene overlap. (a) Sample averaged parameterised costs significantly correlates with sample averaged PLS1 nodal gene scores, explaining on average 65.0% covariance. Statistics were

computed via two-tailed linear correlations, quoting the Pearson's correlation coefficient, followed by n=10,000 permutations. (b) Sample averaged parameterised values significantly correlates with sample averaged PLS1 nodal gene scores explaining on average 56.9% covariance. Statistics were computed via two-tailed linear correlations, quoting the Pearson's correlation coefficient, followed by n=10,000 permutations. (c) Nodal-costs PLS1 is enriched for genes predominantly associated with protein localisation, catabolic processes and ribosomal/membrane cellular components. (d) Nodal-values PLS1 is enriched for genes predominantly associated with synaptic signalling, neuronal projection and synaptic membranes.

2.4. Discussion

Diversity in macroscopic human brain organisation can be modelled using a generative network. The generative framework does not include time itself as an explicit parameter, but instead models it as a sequence of processes, optimising its connectivity by renegotiating its costs and value (Betzel et al., 2016; Vértes et al., 2012) continuously over iterations. Despite the simplicity of this equation, it results in the dynamic updating of wiring probabilities over time, with multiple network properties, like spatial embedding, being an emergent property of this dynamic updating. Specifically, it is emergent in the sense that these network properties can be explained in terms of the simulated, known, simple mechanisms (i.e., algorithm) defined by the generative model. Importantly, as alluded to in **Chapter 1**, the generative modelling framework here resonates with previously outlined theoretical perspectives implicating dynamic interactions between brain systems over development in progressive, integrative, specialisation (M. H. Johnson, 2011). A key novelty here is that we have drawn direct links to this process in the context of neurodevelopmental diversity; offering a new perspective on the formation of organised macroscopic networks, their possible biological underpinnings, and their association with functional outcomes like cognitive performance. This reflects a theoretical step-change in understanding diversity in neurodevelopment in that it is sufficiently well-specified to generate macroscopic brain networks at the individual level. In turn, this formalisation allows for the unpacking of the computational and/or biological constraints that shape the developing trajectories of networks.

Small changes in wiring parameters of the generative model lead to divergent macroscopic brain networks, with systematically different network properties. Within the model, the key factor that drives individual differences in growth trajectory is the dynamic nature of updating preferences over time. Specifically, as nodes form new connections this dynamically changes their neighbourhoods, and in turn this quickly changes which nodes become 'attractive' for subsequent connections. This is a simple but important observation. The formation of connections, by virtue of existing within the wider network context, will indirectly influence the subsequent statistical relationships between all other neural populations. Mathematically, this is reflected by the formation of a single connection having potentially very large changes to later wiring probability matrices. This means that the course of the developmental trajectory, as modelled, can be altered quite dramatically by early changes in connectivity – leading to differing outcomes. The same is not true for the geometric ($D_{i,j}$) part of the wiring equation, which does not change as a function of updating connectivity. Unlike the values term, geometry acts as a consistent constraint that biases connections towards being shorter. Of note, geometry may be modelled as changing throughout the course of simulation (Oldham et al., 2022). In either case, any changing geometric influences still remain independent of connectivity and, instead, follow a relatively fixed developmental trajectory. Of note, this can also be affected by the, at first seemingly unrelated, way the wiring equation is specified. The current work utilises a multiplicative wiring equation; combining both costs and values into a single wiring probability distribution which determines wiring. As a result, costs and values tend to have coupled influences. In contrast, to parse the separable effects of the two terms, Oldham, *et al.* (Oldham et al., 2022) develop an additive version of the wiring equation. To do this, normalisation of the two wiring terms is required. This has the indirect consequence of actually *reducing* the dynamic influence of changing geometry, precisely because wiring probabilities are defined in relative rather than absolute terms. These otherwise simple methodological alterations therefore change the theoretical lens in which the results can be interpreted. Ultimately, this broadly highlights how changes to generative models must be carefully considered in tandem with theory and this idea will be revisited in **Chapter 5**. Nevertheless, irrespective of these methodological considerations, the present work highlights

that variability will necessarily occur from whatever interactions tend to remain dynamic between neural populations. In **Chapter 3**, this idea will be explored more deeply with relation to the extent to how determined probability distributions are, in addition to how inhibitory mechanisms may be important for sustaining them.

Why do homophily-based generative rules approximate whole-brain networks so well? We propose that the superordinate goal of any developing brain network is to achieve the optimal computational capacity required of it, given finite biological resources. In this light, we suggest that matching produces the lowest-energetic networks precisely because it provides the closest *heuristic* estimate (compared to those tested here and in other works; Betzel et al., 2016; Vértes et al., 2012; Zhang et al., 2021) of the genuine dynamic reappraisals that occur over developmental time. This is because by virtue of preferentially wiring with nodes with shared neighbourhoods, modular architectures emerge (Damicelli et al., 2019), and this reflects the brain's overarching structure. The modular architecture of the brain has been well studied, and has numerous properties enabling effective flexible computations likely important for functional integration (Betzel et al., 2017; Park & Friston, 2013). By virtue of only requiring knowledge of neighbourhood overlap, homophily-based methods may incur less informational costs (Avena-Koenigsberger et al., 2019) relative to other methods which require global information, and therefore may be more biologically plausible. It is of note that homophily is a measure somewhat akin to network communicability – another locally knowable measure containing information that closely relates to the shortest path (Benzi & Klymko, 2013). The link between homophily and communicability will be more closely examined in **Chapter 3** and **4**. Finally, a tentative explanation for homophily at the neuronal level can also be provided in terms of Hebbian-like plasticity (Goulas, Betzel, et al., 2019; Vértes et al., 2012)**.** This opens up the exciting possibility for these generative models to model developing neuronal cultures (Schröter et al., 2015) to explore whether similar wiring principles may operate across scales, which is the subject of the **Chapter 3**.

Our current generative network models operate at a whole-brain level – i.e., a global set of rules governing network formation. But with more biologically realistic information about regional differences it is possible that an alternative *growth* model could be fitted (Nicosia et al., 2013). Unlike a generative model, a growth model captures the graded changes of an established network over time. This would allow for time itself to be incorporated as a parameter within the model, making regionally and temporally sensitive modifications to network growth and therefore encompass multiple longitudinal measurements of the same individual over a biologically meaningful timescale. The only work we know of to have done this is by Nicosia V. *et al.* (Nicosia et al., 2013) in which data from Caenorhabditis elegans (with birth times of ~300 neurons) where integrated into a growth model to reproduce the developing C. elegans network and it's bi-phasic growth rate. With sufficient detailed information it may be possible in future to take a similar growth modelling approach in humans. This idea is elaborated in **Chapter 5**.

Regional variation in nodal costs and values closely mirrored the expression profiles for different sets of genes, which in turn govern different biological processes and cellular components. Since the advent of genome-wide association studies (GWAS), a huge number of genes have been implicated in developmental disorders, including schizophrenia (Ripke et al., 2014) and autism (Grove et al., 2019) but also general cognitive functioning (Davies et al., 2018). It has been challenging to interpret the consequences of these individual implicated genes. The enrichment analysis that accompanied our generative models takes a very different approach. As far as we are aware, this is the first study aiming to bridge models of whole brain organisational emergence and genetics in this way. Nodal costs covaried with genes enriched for highly costly metabolic processes, including catabolic processes, protein transport and cellular components centred around the ribosome and endoplasmic membranes. On the other hand, nodal values covaried with genes enriched for trans-synaptic signalling, neuronal projection and the synaptic membrane. This aligns with recent findings that synaptic genes also colocalise with highly-synergistic regions of the brain, which have been suggested to be crucial for human cognitive evolution (Luppi et al., 2022).

The *omnigenic* model (Boyle et al., 2017) suggests that complex traits are driven by genes that do not have direct effects on the trait per se, but instead propagate through regulatory networks on much smaller numbers of core genes, with more direct effects. This model explains the vast number of GWAS hits for complex traits, as "peripheral" genes necessarily outnumber "core" genes and thus the sum of their small effects exceeds the contribution of core genes. We suggest the omnigenic model may apply to some aspects of gene-development relationships. That is, the many genes that contribute to each PLS1 may not directly contribute to developmental processes themselves, but in the regulation of activity and growth within brain areas that are particularly important for neurodevelopment. Crucially there is variability in enriched genes across subjects (**Figure B3a**).

Our sample is a large mixed cohort of children, the majority of whom were referred from specialists in children's educational and clinical services. This was the ideal testbed for exploring diverse trajectories. The varied referral routes for the cohort makes its composition more reflective of children at heightened neurodevelopmental risk, relative to a more standard case-control design recruited according to strict diagnostic inclusion and exclusion criteria, via a single referral route (Astle & Fletcher-Watson, 2020). But it is important to note that the modelling also works well in a more conventional sample recruited from classrooms in the same area. So, whilst the CALM cohort is ideal for exploring the mechanisms of heterogeneity, these same mechanisms are likely at play in more standard samples. Indeed, this work presents a challenge to the long history of categorising neurodevelopment into discrete groupings based on observed cognitive and/or behavioural traits. Instead, we suggest divergent outcomes may arise via slight trajectory changes that fall out of the continual negotiation of brain connectivity optimisation. While likely that generative preferences are initialised via an individual's genetic preprograming, small changes in wiring preferences over time – possibly via complex interactions of their time course, endocrinological exposure, learning and environment – have profound effects on the emergence of the developmental trajectory. What results is a continual interaction between network growth preferences and the dynamically developing brain, leading to neurodiverse outcomes.

Whilst our sample is designed to capture children at risk, the findings generalised to a more conventional sample, with an even split of boys and girls, recruited in local schools. This suggests that this computational approach could be a powerful tool for developmental scientists more generally. The advent of larger datasets is allowing the study of the developing brain at unprecedented scale and across multiple levels (Casey et al., 2018; Karcher & Barch, 2021). Computational frameworks that allow the integration of different datatypes (e.g., multiple imaging modalities, genetic variability) could provide a valuable tool for building developmental theory that goes beyond correlating different datatypes over time, and fully capitalises on the scale and complexity of those datasets. In the future, this approach could be used to test different theoretical accounts of developmental change, and to make longitudinal predictions where multiple waves of data are available. To realise these opportunities, a crucial next step is to deploy this type of computational modelling to capture the diversity present in larger population-level neuroimaging datasets, with longitudinal data (Bycroft et al., 2018; Casey et al., 2018; Karcher & Barch, 2021) spanning multiple sites.

This computational framework has a number of limitations that provide scope for future improvements. Our generative models are limited to the binary connections which are assumed to be anatomical. This is inevitably a gross simplification of the complex weighted structure of the connectome. Devising ways in which network connections can change in a more graded fashion is a necessary next step to modelling more complete developmental processes. In the future we will need to capture both the strengthening and weakening of connections that has been shown to occur in human brain development (Tymofiyeva et al., 2014; Vértes & Bullmore, 2015). Relatedly, the binarisation of connectomes may introduce some degree of confounding. To our knowledge, there is currently no work formally outlining the effects of density on generative model outcomes. One may stipulate that by increasing the density of the studied networks, nodes increase in connectivity and therefore necessarily all nodes become more correlated (i.e., leading to less node-wise variability). This theoretically could decrease the effectiveness of homophily models to simulate these networks, but as far as we are aware this is yet to be formally demonstrated. Secondly, we currently use one rule, but it is conceivable

that different rules govern growth at different points in the trajectory. It may be possible to accurately approximate the rules governing the remodelling of networks over time, modelled either by changing heuristic estimates (e.g., changing of generative rules over time) or attempting to optimise a superordinate goal (e.g., computational efficiency and/or flexibility). Thirdly, our gene enrichment results are correlational, not causative. There remains an explanatory gap in determining whether and how these specific gene profiles support the sensitivity to connection formation. And crucially, the expression data are derived for a microarray analysis of post-mortem tissue samples from human adults – they are not from the subjects in the study, they are older and there is clear variability within the post-mortem tissue samples (Arnatkevičiūtė et al., 2019; Hawrylycz et al., 2012). Moreover, while RNA-seq data was used to cross-validate gene expression measures by removing probes with a correlation <0.2, it was not possible to externally validate the data with an independent dataset. Caution is therefore required when interpreting our enrichment analysis due to this lack of external validation (Fulcher et al., 2021). The next steps will involve validating these findings in large scale developmental cohorts with available gene data, and forming casual links by applying generative network models to individuals with neurodevelopmental disorders of known genetic origin (Bathelt et al., 2016, 2017; Hawkins et al., 2020). Relatedly, it has been observed that due to the inherent spatial autocorrelations present in transcriptional maps and (often) within neuroimaging phenotypes, that similar gene-ontologies may appear across different phenotypes – limiting the specificity of gene enrichment analyses to particular phenotypes. While our analysis here treats space and topology as separable generative entities as simulated networks develop, it is possible that our findings also reflect some more general gene ontology category. One way to determine the extent to which there is specificity to developmental processes in these enrichments would be to undertake an enrichment analysis over the course of the generative simulation. It would then be possible to assess how significant gene ontologies vary as the network develops in space, rather than at the end of the simulation where autocorrelations may possibly influence interpretations. Fourthly, we have used a parcellation widely used in developmental studies (Alexander et al., 2019; Bathelt et al., 2019; Martin et al., 2019; Tamnes et al., 2017), which aids the comparability of our findings with

the wider literature. Previous work has shown that parcellation choice is not a big determent of optimal wiring parameters (Betzel et al., 2016; although see below for a theoretical caveat to this observation). But within the context of a developmental sample there could be subtle differences in the optimal parcellation across participants, and thus the production of individually optimised parcellations (Eickhoff et al., 2018; Tian et al., 2020) would allow us to test whether and how developmental change in the parcellation itself influence wiring properties. Finally, as in previous studies (Betzel et al., 2016; Vértes et al., 2012; Zhang et al., 2021), our models utilise Euclidean distance as measures of cost in connection formation. While a simplification, this selection removes any *a priori* constraints to the generative model that a more biologically specified cost measure may provide (e.g., fibre lengths, which are sparse and thus limit potential connections). A number of studies have shown relatively inconsistent findings in terms of how much Euclidean distance accounts for fibre length, with findings ranging from 22% (RED, in this work) to 79% (Human Connectome Project) of variance explained. There are very likely to be cohort, parcellation and tractography effects influencing these relationships. Whilst interpolated fibre lengths have been shown to perform equivalently to Euclidean distances within generative network models (Betzel et al., 2016), it is important to consider that in our modelling these two measures only partially overlap. A crucial next step is to test whether microstructural informed tractography (Schiavi et al., 2020), which may provide a more direct measure of biological cost, improves model performance. These ideas will be revisited in the final **Chapter 5**. Another interesting avenue of future research would be to explore the extent to which the distribution of spacing between regions, perhaps deriving from different parcellations, affects the interpretations of generative modelling. For example, it is possible that heterogeneity in spacing (i.e., non-evenly spaced regions) would lead to greater clustering, and *vice versa*. This could feasibly impact wiring parameter estimates. For example, if unevenness in spacing was proportional to greater generated network modularity, η would likely change in the opposite direction. That is, the greater the disparity in spacing distributions, the lower the magnitude of η required to generate the modular network. As described above, the consistency of findings thus far across parcellations may simply be because their spacing distributions are similar.

In conclusion, we provide a unifying computational framework for conceptualising the emergence of structural brain networks and their diversity. The emergence of brain networks can be understood as occurring via continual renegotiations of costs and values, but individuality emergences from their slightly different parameterisation.

3. CHAPTER 3

EVALUATING WIRING PRINCIPLES UNDERPINNING NEURONAL NETWORK TOPOLOGY *IN VITRO* ACROSS SPECIES AND TIME

Content included in this chapter was published in bioRxiv, 2022, and is currently in review at eLife: Homophilic wiring principles underpin neuronal network topology in vitro. ***Danyal Akarca****, Alexander W. E. Dunn*, Philipp J. Hornauer, Silvia Ronchi, Michele Fiscella, Congwei Wang, Marco Terrigno, Ravi Jagasia, Petra E. Vértes, Susanna B. Mierau, Ole Paulsen, Stephen J. Eglen, Andreas Hierlemann, Duncan E. Astle† & Manuel Schröter† (*Co-lead authors, †Co-senior authors).*

The majority of work employing generative models of network topology has focussed on macroscopic connectivity. Yet, economic principles for network organisation extend across levels of analysis within the developing nervous system. Can we apply the same framework to neuronal networks at the cellular scale, to test explanatory consistency across these levels? To do this, we combine high-density microelectrode array (HD-MEA) recordings, which allow for the detailed characterisation of the ongoing extracellular electrical activity of individual neurons *in vitro* with the aforementioned generative modelling approach. As with macroscopic brain organisation, we find the best fitting model also uses a homophilic generative wiring principle in which neurons form connections to other neurons with similar connectivity patterns to themselves. This mechanism for neuronal network emergence accounts for a wide range of observations that are described, but not sufficiently explained, by traditional analyses of network topology. Using rodent and human monolayer and organoid cultures, we show that homophilic generative mechanisms account for the topology of emerging cellular functional connectivity, representing an important wiring principle and determining factor of neuronal network formation *in vitro*.

3.1. Introduction

During mammalian brain development, neuronal networks show remarkable self-organisation that gives rise to complex topological properties, including a greater-than-random clustering and modular community structure (Newman, 2006; Sporns & Betzel, 2016), hierarchies (Hilgetag & Goulas, 2020), heavy-tailed connectivity distributions (Alstott et al., 2014; Hagmann et al., 2007), and richly interconnected hubs (Achard et al., 2006; Bullmore & Sporns, 2009). These distinctive characteristics likely endow neuronal networks with robustness and the capability to support dynamic functional computations (Laughlin & Sejnowski, 2003; E. J. Müller et al., 2020).

Neuronal network development can be characterised across multiple scales (Betzel & Bassett, 2017a; Schröter et al., 2017). At the cellular level, neurons form computational units within circuits. Here, the role of individual neurons can be determined by a combination of factors,

such as their laminar location, connectivity, neurochemical sensitivities and morphology (Kasai et al., 2021). During embryonic development, a series of spatiotemporal-defined genetically-hardwired programs regulate the expression of cell-type specific recognition molecules to initiate axonal and dendritic outgrowth, which ultimately leads to formation of synapses (Jain et al., 2022; Südhof, 2018; Yogev & Shen, 2014). Although there is now a large body of evidence on the mechanisms of specific guidance cues during circuit formation (Sanes & Zipursky, 2020), linking this knowledge to explain the emergence of complex topological features remains challenging.

There is growing evidence that some key organisational properties of nervous systems are conserved across scales, and, in some cases, across species (Allard & Serrano, 2020; Betzel, 2020; Betzel & Bassett, 2017a; Bogdan et al., 2021; Engel et al., 2021; Schröter et al., 2017; Zheng et al., 2020). Nervous systems both at the macro- and micro-scale, for example, have been shown to entail a canonical pattern of small-worldness (Bassett & Bullmore, 2017; Muldoon et al., 2016), a rich-club topology(Harriger et al., 2012; Heuvel & Sporns, 2011; Towlson et al., 2013) and a modular structure (Meunier et al., 2010; Newman, 2006; Sporns & Betzel, 2016). These complex organisational hallmarks allow for functional hierarchies, in which distinct segregated modules perform specialised local computations reflecting basic representational features of incoming signals, while their intermediary nodes integrate those signals to code for a more complex representation of the incoming signal (Lord et al., 2017).

As outlined in **Chapter 1** and **2**, one prominent explanation for these consistent organisational hallmarks is that they reflect the economics of forming and maintaining connections (Bullmore & Sporns, 2012; Fornito et al., 2016; Zhou et al., 2022). Given finite available resources, trade-offs between incurred costs (e.g., material, metabolic) and functional value have to be made continually by all distributed units to ensure optimal network function. In this view, ideas, such as Peters' rule (Braitenberg & Schüz, 1998), which suggest that synaptic contacts simply occur if neurons are close enough in space and if their axons and dendrites overlap, are not sufficient to explain the existence of a connection (Briggman et al., 2011; Takemura et al., 2015). Although spatial embedding clearly has an impact on network architecture (Ercsey-

Ravasz et al., 2013), costly features, such as long-range connectivity hubs, likely exist because they confer some additive functional value beyond the cost of its formation and maintenance (Bullmore & Sporns, 2012). Principles apply not only to the connectivity between brain regions, but also to the cellular and subcellular level (Chen et al., 2006; Laughlin & Sejnowski, 2003; Vértes & Bullmore, 2015).

If an economic trade-off represents an important principle that guides network development, then *what are the specific mechanisms that determine the outcome of that trade-off?* Furthermore, are these mechanisms conserved across spatial scales? In **Chapter 2**, we showed that the homophily principle can generate macroscopic structural networks that are highly statistically similar to observed networks. However, despite this, the extent to which such models actually reflect underlying biological processes remains unclear due to the indirect nature of *in vivo* imaging. An alternative approach is to study neurodevelopmental mechanisms by culturing cellular networks *in vitro*. This allows for the tracking and modelling of the very early events of network formation. HD-MEAs provide a *direct* window into the cellular processes that occur as individual neurons form connections, and in time, complex functional connectivity networks (Schröeter et al., 2015). Today's HD-MEAs contain large numbers of densely packed electrodes, capable of longitudinal measurements of the extracellular electrical activity of developing neuronal networks at very high temporal and spatial resolution (J. Müller et al., 2015; Ronchi et al., 2021). The spontaneous electrical activity, inherent to developing neurons *in vitro,* resembles the patterns of activity that accompany key neurodevelopmental processes occurring during prenatal and early postnatal development *in vivo*, including the refinement of synaptic connectivity (Blankenship & Feller, 2010).

In the current chapter, we translate prior generative modelling research at the level of inter-regional connectivity, including that in the previous chapter (Arnatkevičiūtė et al., 2021; Betzel et al., 2016; Liu et al., 2020; Morgan et al., 2018; Oldham et al., 2022; Vértes et al., 2012; Zhang et al., 2021), to the microscale (functional connectivity between single neurons). In doing so we test whether generative wiring principles are recapitulated across spatial scales. We will also introduce a novel way of considering the local hallmarks of regional connectivity

(coined *topological fingerprints*). This provides an independent method to assess model fits within the generative framework. Moreover, due to the longitudinal nature of the data, we are able to ask more specific questions regarding the capabilities of simple models to recapitulate network statistics over time, examining the capability for developmental trajectories to be resolved. Using HD-MEAs, we acquired electrophysiological network recordings of the spontaneous extracellular neuronal activity from populations of developing primary cells (PCs) derived from dissociated embryonic rodent cortices, three different lines of human induced pluripotent stem cell (iPSC)-derived neurons and more mature sliced human embryonic stem cell (hESC)-derived cerebral organoids (hCOs). Facilitated by the high spatial recording resolution of the used HD-MEAs (centre-to-centre electrode distance of 17.5µm; >3000 electrodes/mm^2), we resolved the activity of several thousands of individual neurons and tracked the development of emerging functional connectivity over several weeks. These data also gave us a unique opportunity to investigate, experimentally, the biological underpinnings that may plausibly arise *in vitro*. For example, how precisely may neurons judge the viability to form connections with each other? A key hypothesis in GABAergic interneurons which regulate the synchronicity of spontaneous neural activity (Bonifazi et al., 2009; Cossart, 2014). By experimentally perturbing $GABA_A$ receptor activity we can ask – in the language of generative network models – how wiring outcomes are influenced.

Using large-scale HD-MEA datasets acquired from developing rodent and human neuronal networks, we comprehensively compared the performance of different generative network models, and probed whether they can account for emerging neuronal network topology over development. We explicitly examine the effect of neuronal plating density on topology and test whether generative network models are capable of recapitulating the local organisational properties of the observed biological neuronal networks via a topological fingerprinting approach. Finally, as briefly alluded to above, by blocking $GABA_A$ receptors we probe whether GABAergic signalling impacts wiring stochasticity in the network. We argue that homophilic wiring mechanisms may reflect a fundamental principle shaping the development of neuronal

networks *in vitro*, and are likely to extend across scales and species in which local structures are refined via activity-dependent interactions over time.

3.2. Methods

3.2.1. High-density microelectrode arrays

Two types of CMOS-based HD-MEA recording systems, produced by MaxWell Biosystems (Zurich, Switzerland), were used in the present study (Ballini et al., 2014; J. Müller et al., 2015). The single-well HD-MEA MaxOne, consisting of 26,400 low-noise electrodes with a centre-to-centre electrode pitch of 17.5µm, arranged in a 120x220 electrode array structure. This HD-MEA can record simultaneously from a total of 1024 (user-selected) readout-channels at 20 kHz; for more technical details see previous studies (Ballini et al., 2014; J. Müller et al., 2015). The second recording system was the multi-well HD-MEA MaxTwo (MaxWell Biosystems, Zurich, Switzerland), comprising the same number of electrodes and readout-channels and electrode specifications as MaxOne for each well. With this system it is possible to simultaneously record from six wells at a time and at a sampling rate of 10kHz. To decrease the impedance and to improve the signal-to-noise ratio (SNR), electrodes were coated with platinum black (J. Müller et al., 2015).

3.2.2. Rodent primary cortical neuronal cultures

Before plating, HD-MEAs were sterilised in 70% ethanol for 30 minutes and rinsed three times with sterile water. To enhance cell adhesion, the electrode area of all HD-MEAs was treated with poly-D-lysine (PDL, 20µL, 0.1mgmL^{-1}; A3890401, Gibco, ThermoFisher Scientific, Waltham, USA) for 1 hour at room temperature and then rinsed three times with sterile water. Next, 10 µL Geltrex (A1569601, Gibco, 0.16 mgmL^{-1}) was pipetted on each array and again left for about one hour at room temperature. For the main analysis of the paper, we used rodent PC neurons prepared as previously described (J. Müller et al., 2015). Briefly, cortices of embryonic day (E) 18/19 Wistar rats were dissociated in trypsin with 0.25% EDTA (Gibco), washed after 20 min of digestion in plating medium (see below), and triturated. Following cell

counting with a haemocytometer, either 50,000 cells (sparse plating condition) or 100,000 cells were seeded on each array, and afterwards placed in a cell culture incubator for 30min at 37°C/5% CO_2. Next, plating medium was added carefully to each well. The plating medium contained 450mL Neurobasal (Invitrogen, Carlsbad, CA, United States), 50mL horse serum (HyClone, Thermo Fisher Scientific), 1.25 mL Glutamax (Invitrogen), and 10mL B-27 (Invitrogen). After two days, half of the plating medium was exchanged with growth medium containing 450mL D-MEM (Invitrogen), 50 mL horse serum (HyClone), 1.25 mL Glutamax (Invitrogen) and 5mL sodium pyruvate (Invitrogen). Across all experiments, the medium was then exchanged twice a week, at least one day before the recording sessions. All animal experiments were approved by the veterinary office of the Kanton Basel-Stadt, and carried out according to Swiss federal laws on animal welfare. A summary of the data used is provided in Table 3.1.

Cell line	**No. of units**	**Recording time point**	**Plating density**	**Samples (n)**
Primary rodent cortex (sparse)	129.7 ± 27.9	DIV 7, 10, 12, 14	50,000	6
Primary rodent cortex (dense)	114.7 ± 27.4	DIV 14/15, 21, 28	100,000	12
Human motor neurons	210.7 ± 24.7	DIV 28	100,000	7
Human glutamatergic neurons	191.6 ± 69.1	DIV 28	100,000	8

Human dopaminergic neurons	211.2 ± 55.9	DIV 28	100,000	6
Human cerebral organoids	47.3 ± 27.8	120 days	n/a	6 (slices), 3 organoids

Table 3.1. Overview of all the datasets used in the study. Number of unit results are expressed as the mean±SD of the respective cell line.

3.2.3. Human induced pluripotent stem cell-derived cultures

Three different human iPSC-derived neuronal cell lines were included in the study: iCell DopaNeurons, iCell Motor Neurons and iCell GlutaNeurons, all commercially available from FUJIFILM Cellular Dynamics International (FCDI, Madison, USA). All neural cells were co-cultured with human iCell Astrocytes (FCDI, see above). Cell plating: Cell plating medium consisted of 95mL of BrainPhys Neuronal Medium (STEMCELL Technologies, Vancouver, Canada), 2mL of iCell Neuronal Supplement B (FCDI), 1mL iCell Nervous System Supplement (FCDI), 1mL N-2 Supplement (100X, Gibco), 0.1mL laminin (1mg/mL, Sigma-Aldrich) and 1mL Penicillin-Streptomycin (100X, Gibco). Neurons and astrocytes were thawed in a 37°C water bath for 3 minutes. The cells were then transferred to 50mL centrifuge tubes, and 8 mL plating medium (at room temperature) was carefully added. Cell suspensions were centrifuged at 380xg (1600 RPM) for 5 minutes, and the supernatant was aspirated. Cell pellets were then resuspended in plating medium and combined to achieve a final concentration of 10,000 neurons and 2,000 astrocytes per µL. Finally, 100,000 neurons and 20,000 astrocytes were seeded per HD-MEA by adding 10µL of the prepared solution, after removing the Geltrex droplet. After incubating the cells for one hour at 37°C/5% CO_2, another 0.6mL (small well MaxOne) / 1.2mL (large well MaxOne) of plating medium was added. Half of the medium was changed twice a week.

3.2.4. Human cerebral organoid slice cultures

Human embryonic stem cell (hESC)-derived cerebral organoids (hCOs) were generated from a commercially available hESC stem cell line (Takara Bio, Osaka, Japan), using the STEMdiff cerebral organoid kit (STEMCELL Technologies) following the manufacturer's instructions. Slices were obtained from 120-day old hCOs. Single organoids were first transferred from maturation medium to ice-cold BrainPhys (STEMCELL Technologies) using cut 1000µl pipette tips. Next, cross-sectional 500-µm-thick slices were cut from hCOs using a sterile razor blade and collected in petri dishes filled with BrainPhys medium at room temperature. Before the plating, HD-MEAs were sterilised in 70% ethanol for 30 minutes and rinsed 3 times with distilled water. To improve tissue adhesion, arrays were coated with 0.05% (v/v) poly(ethyleneimine) (Sigma-Aldrich) in borate buffer (pH 8.5, Thermo Fisher Scientific) for 30 minutes at room temperature, rinsed with distilled water, and left to dry. To attach hCOs on HD-MEAs, we applied a thin layer of Matrigel (Corning) to the centre of the HD-MEA and then transferred individual organoid slices to the coated HD-MEAs. After positioning the tissue, we placed a tissue "harp" on top of the organoid slice and applied several drops of recording medium (STEMCELL Technologies, #05793) around the organoid. HD-MEAs were then covered with a lid and placed in a humidified incubator at 37°C, 5% CO_2/95% air for 30 minutes, before adding more medium to a final volume of 2ml per chip. Half of the recording medium was changed every 2-3 days.

3.2.5. Immunohistochemistry

Rodent PC neurons were stained as previously described (D. J. Bakkum et al., 2013). Briefly, PC neurons were fixed using 4% paraformaldehyde solution (ThermoFisher, #FB001). Samples were permeabilised and blocked using a PBS 10X (ThermoFisher, #AM9625) solution containing 10% normal donkey serum (NDS) (Jackson ImmunoResearch, West Grove, USA, #017000001), 1% bovine serum albumin (BSA) (Sigma-Aldrich, 0 5482), 0.02% Na-Az (Sigma-Aldrich, #S2002) and 0.5% Triton X (Sigma-Aldrich, #93443). Permeabilisation facilitated antigen access to the cell, while blocking prevented non-specific binding of antibodies to

neurons. Primary and secondary antibodies were diluted in a PBS solution containing 3% NDS, 1% BSA, 0.02% Na-Az and 0.5% Triton X. The used antibodies are also listed in **Table C1**. Note, immunohistochemistry was performed on control PC cultures prepared as previously outlined (D. J. Bakkum et al., 2013).

Human iPSC-derived neurons were fixed using 8% PFA solution (#15714S, Electron Microscopy Sciences) and blocked for 1 hour at room temperature (RT) in blocking buffer containing 10% normal donkey serum (NDS) (Jackson ImmunoResearch, West Grove, USA, #017-000-001), 1% bovine serum albumin (BSA) (#05482, Sigma-Aldrich), and 0.2% Triton X (Sigma-Aldrich, #93443) in PBS (ThermoFisher Scientific, #AM9625). Primary antibodies were diluted in a blocking buffer and incubated overnight at 4°C. Samples were washed three times with 1% BSA in PBS and incubated with the secondary antibody diluted in blocking buffer for 1 hour at RT. After three additional washes with PBS, DAPI was added for 2 min at RT (1:10000). Images were acquired using the Opera Phenix Plus High-Content Screening System (cat. HH14001000, PerkinElmer, Waltham, MA, USA).

hCOs were fixed using 4% paraformaldehyde (PFA) for 4 hours at room temperature, washed with PBS and immersed in 30% sucrose solution at 4 °C overnight. PFA-fixed organoids were embedded in OCT compound (Sakura Finetek, Alphen aan den Rijn, Netherlands, #4583) and stored at -80 °C. 10 µm sections were cut on a cryostat and collected on Superfrost plus slides (Thermo Scientific, #22-037-246). For immunohistochemistry, sections were permeabilised in 0.1% Triton X-100 and blocked with animal-free blocker (Vector Laboratories, Burlingame, CA, USA, #SP-5030-250). Slides were incubated with primary antibodies for 1 hour at room temperature. Sections were washed in PBS and further incubated with secondary antibodies for 1 hour at room temperature. After washing with PBS, sections were incubated with PureBlu DAPI (Bio-Rad, Hercules, CA, USA, #1351303) for 3 minutes and mounted with ProLong Gold antifade mounting medium (Thermo Scientific, #P36930). Fluorescence images were acquired with a SP8 confocal microscope (Leica, Wetzlar, Germany). The primary and secondary antibodies used for hCO stainings are listed in **Table C1**.

3.2.6. Scanning electron microscope imaging

Fresh tissue samples were fixed in 2.5% glutaraldehyde solution (Sigma-Aldrich, St. Louis, USA) overnight. After fixation, the samples were dehydrated in ascending acetone series (50%, 70%, 80%, 90%, 95%, 100%), and critically point dried (CPD; Quorum Technologies, West Sussex, UK), using CO_2 as the substitution fluid. The procedure is generally suited for SEM preparation and ensures that surface structures of animal tissue samples are preserved in their natural state, i.e., without shrinkage, distortion or dissolution. After CPD, specimens were carefully mounted on aluminium stubs using double sticky carbon-coated tabs as adhesive (Plano, Wetzlar, Germany). Thereafter, they were coated with gold-palladium in a sputter device for 45 seconds (Bio-Rad SC 510, Munich, Germany). SEM analyses were carried out with a Zeiss Digital Scanning Electron Microscope (SUPRA 40 VP, Oberkochen, Germany) in SE2 mode at 5-10kV.

3.2.7. Electrophysiological recordings

In order to track the development of functional connectivity of *in vitro* neuronal networks on HD-MEAs, we performed weekly recordings, starting one week after plating. In order to select a network recording configuration, we performed whole-array activity scans, i.e., series of 1 minute long high-density recordings, covering all 26,400 electrodes of the HD-MEA, using the MaxLab Live software (MaxWell Biosystems). To select recording electrodes, we estimated the multi-unit activity for each electrode using an online sliding window threshold-crossing spike-detection algorithm (window length: 1,024 samples; detection threshold: 4.5 × the root mean squared error (RMSE) of the noise of the 300-3000 Hz band-pass filtered signal). After the activity scan, we selected up to 1024 readout-electrodes, based on the detected average activity and a ranking of the inferred, average amplitude values. Additional high-density network recordings, consisting of 4 x 4 electrode blocks (17.5µm pitch), were acquired for the tracking experiments (see below). The duration of the HD-MEA network recordings was about 30 minutes. The PC neuronal network and the hCO data were acquired by MaxTwo multi-well plates (MaxWell Biosystems); the human iPSC-derived neurons (glutamatergic, motor

and dopaminergic neurons) were recorded on single-well MaxOne HD-MEAs (MaxWell Biosystems).

3.2.8. Pharmacological experiments

Pharmacological experiments with the $GABA_A$ receptor blocker gabazine (SR 95531 hydrobromide, Sigma-Aldrich, #104104509), were performed on sparse (50,000 per well) primary cortical (PC) neuronal cultures. Three cultures were treated with 1µM gabazine one day after plating and tracked until DIV14; media+gabazine exchanges were performed 2-3 times per week.

3.2.9. Spike-sorting and post-processing

All HD-MEA network recordings underwent an initial quality control to assess the overall noise level and signal stability of each recording. Next, we used the software package Kilosort 2 (KS2) (Pachitariu et al., 2016) to spike sort data, applying default parameters. For the developmental tracking analyses, we concatenated all recordings (i.e., DIV7, 10, 12, and 14 for the PC cultures at 50k plating density, and DIV14, 21 and 28 for the PC cultures plated at 100k per well). After spike sorting, we inferred array-wide spike-triggered averages (STAs) for all units labelled as 'good' by KS2. Next, we calculated the spatial similarity between all detected units/STAs to minimise the influence of potential cluster splits that might have occurred during spike sorting of bursty spontaneous activity. The spatial similarity among the inferred array-wide templates was probed by the normalised pairwise maximum cross-correlation: units/STAs that showed a similarity $r>0.75$ and had at least 5 electrodes in common underwent an iterative elimination process using a simple clustering heuristic (Blondel et al., 2008). See **Figure C1** for the number of trackable units for both datasets.

3.2.10. Firing rate and burst statistics

Firing rates across each neuronal unit were calculated as the total number of spikes per unit time (in seconds) in the entire recording. Array values were calculated as the mean across all active units (firing rates>0.01 Hz). Burst rates were calculated using a maximum interspike-

interval (ISI) method (Cotterill et al., 2016) based on the ISI between every Nth spike (ISI_N) (D. Bakkum et al., 2014). The ISI_N threshold for determining the onset/offset of bursting activity was determined by finding the local trough in the bimodal logISI distribution. The two peaks, at short ISIs and long ISIs represent more high frequency bursting and regular activity, respectively. The coefficient of variation (CV) of interburst intervals (IBIs) was calculated as the standard deviation of IBIs relative to the mean IBI in a given neuronal unit; the array value was the mean of this across all neuronal units.

3.2.11. Functional connectivity inference

To detect pairwise correlations in spike trains, here referred to as functional connectivity, we computed the spiketime tiling coefficient (STTC) (Cutts & Eglen, 2014). The STTC aims to mitigate potential confounding in basic correlation indices introduced by different firing rates, by quantifying the proportion of spikes in one train which fall within $\pm\Delta t$ (the synchronicity window) of another. It is given by:

$$STTC = \frac{1}{2}\left(\frac{P_A - T_B}{1 - P_A T_B} + \frac{P_B - T_A}{1 - P_B T_A}\right), \quad (3)$$

where T_A is the proportion of total recording time which lies within $\pm\Delta t$ of any spike from A (T_B is calculated similarly). P_A is the proportion of spikes from A which lies within $\pm\Delta t$ of any spike from B (P_B is calculated similarly). The synchronicity window, Δt, is the only free parameter in the STTC calculation. In the present study, we used a Δt=10ms. STTC was calculated using publicly available Matab code (Giandomenico et al., 2019). We used permutation-based testing to determine the significance of connections. For a given neuronal unit's spike train, spike times were randomly jittered by ±10ms to create a surrogate spike train, using code provided by the Neural Complexity and Criticality Toolbox (Marshall et al., 2016). This was repeated for each neuronal unit for 1000 permutations. To calculate significance of pairwise functional connectivity, experimentally inferred STTC values were compared to the distribution of surrogate SSTC values. A significance value of $p<0.01$ was

used as a cut-off to binarise functional connectivity matrices and calculate network related analysis throughout the manuscript; only units with firing rates >0.01 Hz were considered.

3.2.12. Generative network modelling and Voronoi tessellations

As before, the broad generative network modelling framework was laid out in **Chapter 1.** For each empirical network, we simulated 20,000 networks across a wide parameter space (with η and γ limits -10 to 10) across the same previously studied 13 wiring rules, for each network across all available time points. We used this wide parameter space as there is little prior work guiding our choice of parameters. We also did not select a seed network. A Voronoi tessellation procedure (Betzel et al., 2016) was used as the parameter selection algorithm. This procedure works by first randomly sampling the parameter space and evaluating the model fits of the resulting simulated networks, via the energy equation. Following an initial search of 4000 parameters in this space, we performed a Voronoi tessellation, which establishes two-dimensional cells of the space. We then preferentially sampled from cells with better model fits according to the energy equation (Betzel et al., 2016). Preference was computed with a severity of α=2 which determines the extent to which cell performance led to preferential sampling in the next step. This procedure was repeated a further four times, leading to a total of 20,000 simulations being run for each considered network across all generative rules. Networks were sometimes fragmented during early developmental time points, we restricted calculations of topological metrics to the giant component of the network, i.e., the largest connected part of the network.

3.2.13. Model fitting

In the present study, we make a distinction between simulated networks which, first, mirror the statistical *distributions* of observed networks and, second, those which mirror the topological organisation of those statistics.

The first fitness metric is assessed via the previously used energy equation from **Chapter 2** (Betzel et al., 2016) whereby the model fit is given by the "worst" of the four *KS* distances assessed. As a reminder, the *KS* is the Kolmogorov-Smirnov statistic between the observed and simulated networks at the particular and combination used to construct the simulation, defined by the network degree k, clustering coefficient c, betweenness centrality b and Euclidean edge length d. Note that the *KS* distance between two vectors simply considers their statistical *distributions*. Before moving to local topological organisation, we further assessed the ability of the best performing generative models in each class (foreshadowing results: spatial, matching, clustering average and degree average) to recapitulate network statistics as included in the energy equation, but also two additional measures outside (local efficiency and participation coefficient). We did this via a Monte-Carlo bootstrapping procedure (Beasley & Rodgers, 2012). First, we took the top n=99 performing simulations for each rodent 50k PC culture's model considered, and computed each of the six local statistics as cumulative density plots. For each statistic, we computed a *KS* statistic between the observed local statistics distribution and an average of the statistics of the 99 simulations. We then undertook 99 individual leave-one-out iterations in which we replaced a single simulation of the 99 with the observed distribution. For each of the 99 permutations, we computed the same statistic, forming a null distribution. We then calculated a p_{rank} by ranking how close the original observed statistic was to the mean of this computed null distribution (i.e., how close was the observation to the middle of the null). This was computed for each culture and statistic, for each of the considered generative models. We then quoted the median p_{rank} across cultures. While this measure provides a more direct measure of the ability for each model to recapitulate network statistics, it does so for each network statistic in isolation.

Later in the study, to assess the ability of generative models to capture local statistical relationships in settings with no anatomical reference space (as neurons are randomly distributed on the HD-MEAs), we provide a very simple cost function termed *topological fingerprint dissimilarity* ($TF_{dissimilarity}$). The $TF_{dissimilarity}$ demarcates the ability of *in silico*

network simulations to recapitulate observed local hallmarks of organisation, and is defined as:

$$TF_{dissimilarity_{i,j}} = \sqrt{\sum_i \sum_j (TF_{observed_{i,j}} - TF_{simulated_{i,j}})^2} \quad (4)$$

TF is defined by the n-by-n correlation matrix of n local network statistics for the observed network ($TF_{observed}$) and its corresponding (simulated) network ($TF_{simulated}$). The $TF_{dissimilarity}$ is subsequently equivalent to the Euclidean norm (Lindfield & Penny, 2019) of the difference between observed and simulated topological correlation matrices. Here, we use six common measures of topology to compute the *TF* matrix.

3.2.14. Generative probability distributions

As presently modelled, wiring parameters have a fixed and ongoing effect on network development over the course of a single simulation. Due to the wiring equation, probability updates will subsequently be affected by any parameter differences between conditions. When assessing differences between control versus gabazine cultures in terms of the generative modelling findings, we look directly at the probability ($P_{i,j}$) distributions. This was done by measuring each $P_{i,j}$ element between all node pairs i and j in the network, in 1% intervals (i.e., in a simulation of m=300, every 3 steps were evaluated), before plotting the average distribution of $P_{i,j}$ across these m/100 timesteps. Foreshadowing our findings, it is important to note that the flattening of the probability score distribution means there are more edges have higher probabilities of being connected, leading to decreased specificity of future wiring. This flattening effect is equivalent to the network outcomes being more random, which is a direct result of the γ parameter having a decreased magnitude.

3.3. Results

3.3.1. Tracking developing neurons at single-cell resolution

The datasets of this study comprised recordings of rodent and human neuronal networks that were plated and maintained using established protocols (D. J. Bakkum et al., 2013; Ronchi et al., 2021). Our primary analysis focused on primary rat embryonic day (E) 18/19 cortical cultures (PCs), which were plated at two different plating densities (sparse PC networks: 50,000 cells per well, n=6 cultures; dense PC networks: 100,000 cells per well, n=12 cultures) and used to follow neuronal network development across several weeks *in vitro*. We also analysed human induced pluripotent stem cell (iPSC)-derived neuron/astrocyte co-cultures, containing predominantly glutamatergic, dopaminergic and motor neurons (plated at a density of 100,000 cells per array), and sliced human cerebral organoids (hCOs; n=6 slices).

High-density microelectrode arrays (J. Müller et al., 2015) (HD-MEAs, 3.85 × 2.10 mm^2 sensing area; 26,400 electrodes; 17.5µm centre-to-centre electrode pitch) were used to record the emerging spontaneous neuronal activity of neuronal cultures and to track developing rodent PC neuronal networks at single-cell resolution. We acquired whole-array activity scans to localise the neurons (**Figure 3.1a,b**), and then selected up to 1024 readout electrodes, configured into 4x4 electrode high-density blocks (**Figure 3.1c,d**), at the respective recording start points (i.e., days *in vitro* (DIV) 7 for the sparse PCs and DIV14 for the dense PCs networks). Recordings with the same electrode configuration were acquired also at the consecutive recording time points and concatenated for spike sorting (Pachitariu et al., 2016). Spike-sorted network data allowed us to assign the extracellular electrical activity to individual neurons and to follow single neurons across development (**Figure 3.1e).** The developed approach allowed us to follow single neurons over several weeks *in vitro* (on average 243±37 for the sparse PC networks and 125±29 units for the densely plated PC networks; mean±S.D.; **Figure C1**). In line with previous works, we found that PC neuronal networks developed robust network burst activity (**Figure 3.1c,g**) and that the firing rate of tracked units increased significantly in the first weeks of development (repeated measure analysis of variance (RANOVA): $F(3,12)=7.02$, $p=5.62\text{x}10^{-3}$; n=6 sparse PC networks; **Figure 3.1f**).

To infer functional connectivity among neurons statistically and to characterise neuronal network development of tracked neurons over time, we computed the Spike Time Tiling Coefficient (STTC) (Cutts & Eglen, 2014) – a robust measure of pairwise correlations between spike trains (**Figure 3.1h-j**). Overall STTC increased with development ($F(3,12)=11.82$, $p=6.77\text{x}10^{-4}$, 50k PC cultures; **Figure 3.1i**), as did the network density of inferred functional connectivity graphs ($F(3,12)=11.08$, $p=8.97\text{x}10^{-4}$, $n=6$ sparse PC networks; **Figure C2a**). The probability of inferred STTC connections decayed with inter-neuronal distance (**Figure C2b**).

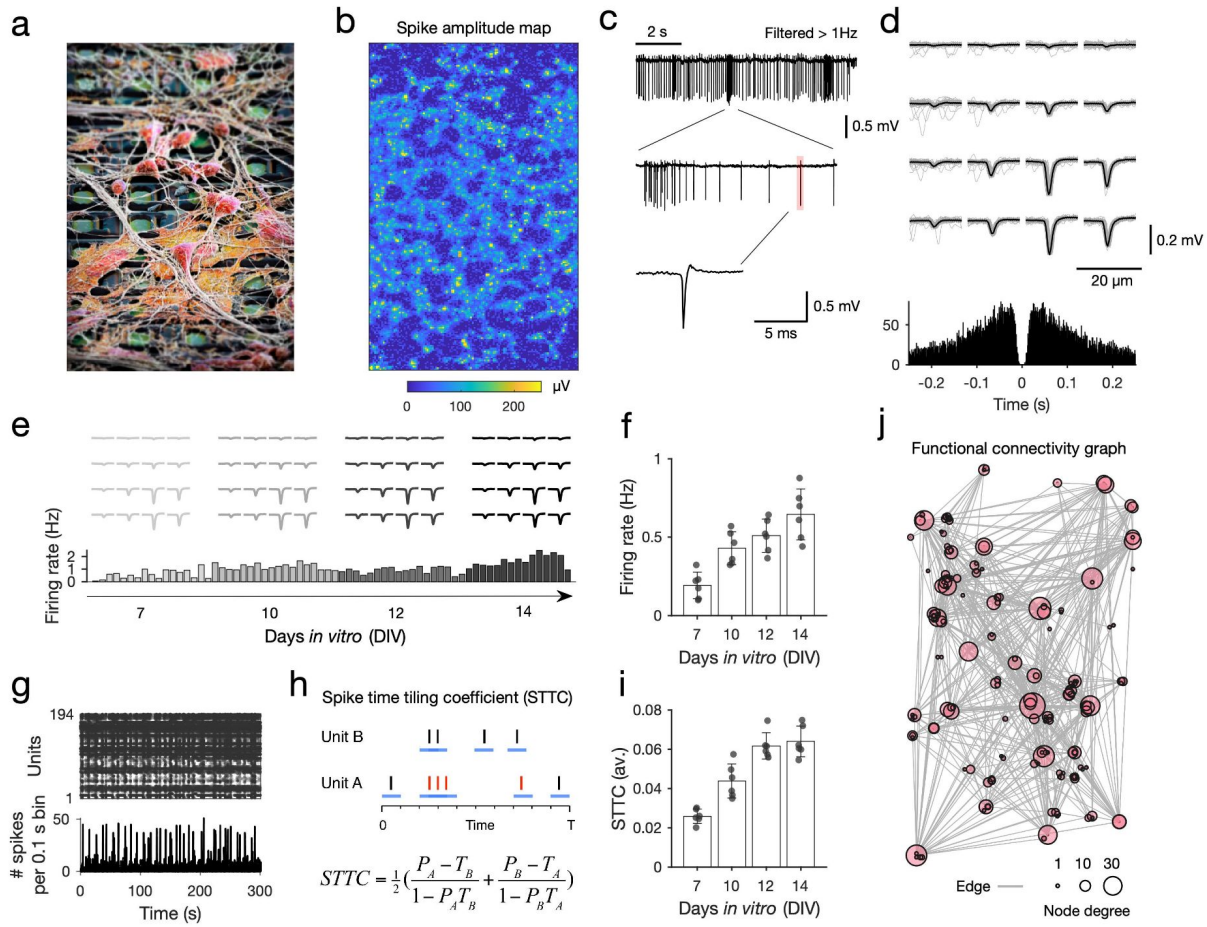


Figure 3.1. High-density microelectrode array extracellular recordings of developing neuronal cultures. (a) Scanning electron microscope (SEM) image of primary neurons plated on high-density microelectrodes (HD-MEA; neurons are coloured in red, electrodes in green; the electrode centre-to-centre distance is 17.5μm). (b) Activity scan across the entire HD-MEA for an example sparse PC neuron culture at days *in vitro* (DIV) 14. Colours depict the average amplitude of online-detected multi-unit activity per electrode (yellow colours indicate the location of potential neurons, i.e., high amplitude deflection, on the array; the array dimensions

are 120 x 220 electrodes; electrode amplitude values are averaged over 1 min recording). (c) Example extracellular trace recorded from one electrode (10s), with high action potential (AP) spiking activity and bursts (middle panel); a single AP is depicted in the lower panel. (d) Electrical footprint (EF) of a single spike-sorted unit on the HD-MEA. The EF is the AP-triggered extracellular potential signature of a single unit on the HD-MEA, here depicted across 16 electrodes of a high-density recording block (in light grey: 100 AP-triggered cut-outs for this EF; in black: the average EF. The lower panel shows a spike autocorrelogram for this unit. (e) Tracking of individual neurons/EFs across development *in vitro*. The upper panel depicts the EFs inferred for the four recording time points (DIV7, 10, 12, 14); the lower panel shows the average activity of the tracked unit (bin size: 100s; grey to black colours correspond to the four recording time points; 30 min per recording day). (f) Spontaneous electrical activity of neuronal cultures, their average firing rate, increased significantly with development (n=6; sparse PC networks). (g) A spike raster plot shows the spike-sorted activity for one culture (300 seconds-long zoom in on a network recording with 194 units); the lower panel depicts the binned activity of the same recording (bin size: 0.1s). All neuronal networks showed a mixture of regular and more synchronised activity periods (network bursts). (h) In order to probe neuronal network development, we used the spike time tiling coefficient (STTC) to infer statistical functional connectivity. The four parameters required to calculate STTC functional connectivity between Unit A and Unit B are P_A, P_B, T_A, and T_B. T_A is calculated as the fraction of the total recording time 'tiled' by $\pm\Delta t$ of any spike from Unit A. P_A is the proportion of spikes from Unit A, which lies within $\pm\Delta t$ of any spike(s) of Unit B (spikes in red). The spike trains for Unit A and B are depicted as black bars; $\pm\Delta t$ is depicted as blue lines for each spike. The significance of each STTC value was estimated against jittered surrogate spike train data. (i) The average STTC increased significantly with development (n=6; sparse PC networks). (j) Example functional connectivity graph inferred from a two-week-old PC network (only the top 5% strongest connections are displayed). The average density of inferred functional connectivity networks was sparse and decayed with distance (Figure C2).

3.3.2. Homophilic wiring principles recapitulate the topology of developing rodent neuronal networks *in vitro*

Following functional connectivity inference, we set out to describe the topology of these networks using graph theory, which provides a mathematical framework for capturing the topological properties of each node within the network, and the network as a whole.

Although graph theoretic measures provide a way to mathematically formalise the topology of networks, they do not provide an explanation as to what topological attachment principles may have shaped network development. To do this, we tested a number of candidate wiring

rules that may best explain the self-organisation of cellular-level functional connectivity graphs over time. This was done using the generative network model procedure previously outlined in **Chapter 1** and **2**. The $D_{i,j}$ term was modelled as the Euclidean distance between tracked neurons i and j at different time points (**Figure C3**).

As in **Chapter 2**, previous studies employing generative models of human macroscopic structural brain organisation have shown that generative rules based on homophilic attachment mechanisms can achieve very good model fits (Betzel et al., 2016; Oldham et al., 2022; Vértes et al., 2012) (although see Arnatkeviciute et al., 2021)). To test what generative models can best simulate microscale connectivity, we applied the generative procedure to inferred cellular functional connectivity graphs.

We first investigate the sparse (50,000 cells per well) PC rodent networks at DIV7, 10, 12 and 14. As previously shown (**Figure 3.1**), PC rodent networks underwent significant developmental changes during this time period. Despite large topological changes (**Figure C2**), we find that over the developmental time course, generative models utilising the homophilic attraction principle as their generative mechanism produce networks with increasingly good model fits (the energy) compared to all other rules tested (**Figure 3.2a**). By DIV14, homophily alone performs best ($p<4.11\text{x}10^{-5}$ for all pairwise comparisons and Cohen's $d>1.46$ reflecting a very large effect size; **Table C2**). The single best performing homophily model, according to the energy equation, was the 'matching model', which – as we outlined from **Chapter 2** – generates network topology according to the overlapping connectivity between nodes (**Figure C4**). We further examined our findings across a range of alternative procedures for formalising the generative models, including a $K_{i,j}$only formalisation, comparison to density- and size-matched random null models, and a comparison of our networks to those constructed using transfer entropy rather than the STTC (**Figure C5**). **Figure 3.2b** shows an immunohistochemistry staining of a PC rodent network at DIV14. **Figure 3.2c** shows the DIV14 energy landscape acquired from the matching generative model.

The matching generative model, beyond providing the lowest energy values (i.e., very good model fits compared to other models), also produced synthetic networks whose aggregate nodal distributions were statistically indistinguishable from the experimentally observed networks (**Figure 3.2d**). We formally demonstrated this using a Monte-Carlo bootstrapping procedure (Beasley & Rodgers, 2012) which directly compares the statistics produced by the well-performing simulations with the observations (degree, p_{rank}=0.0.645; clustering, p_{rank}=0.590; betweenness, p_{rank}=0.815; edge length, p_{rank}=0.585). In **Figure C6** we provide the same bootstrapping analysis, but for each of the best performing generative models of each model class (spatial, clustering average and degree average models). It is of note that this matching generative model was the only model of these able to produce statistically indistinguishable results when compared to the experimental observations across all tested measures; the next best performing non-homophily model, the degree average model, failed to approximate both the empirical edge lengths (p_{rank}=0.025) and participation coefficients (p_{rank}=0.04).

Next, we asked *how well generative models approximate the time-course/trajectories of neuronal network formation.* An advantage of the generative modelling approach is that it allows one to decompose the developmental trajectory. Indeed, if networks are developing according to a homophilic attachment principle then the statistical properties of those simulated trajectories should vary in accordance to our longitudinal observations. To test this, we computed and compared the trajectories of two global network measures of segregation (modularity index Q; Newman, 2006) and integration (global efficiency; Latora & Marchiori, 2001) across the time-course of sparse rodent PC network development. Of note, these measures were both not included in the energy equation. Measures of segregation and integration capture important aspects of how efficient information can be processed across the network (Damicelli et al., 2019; Fair et al., 2007). Next, we selected the best fitting model at DIV14, and decomposed the simulated trajectories up to that point. This allowed us to test whether these simulated trajectories were consistent with the earlier longitudinal observations at DIV7, 10 and 12.

To do this, we compared each of the longitudinal observations (DIV7, DIV10, DIV12 and DIV14) to the simulation at the corresponding time-point of the DIV14 developing simulation (i.e., DIV7, 50%; DIV10, 71%; DIV12, 86% and DIV14, 100%). **Figure 3.2e** and **Figure 3.2f** each show the developmental trajectories for modularity and global efficiency respectively, both unspecified in the energy equation, along with the overlaid observed time-points. Simulations using the homophily generative model clearly captured the same developmental trend for modularity (decreasing over time) and efficiency (increasing over time) and accounted for a substantial amount of variance in both metrics (modularity: R^2=58.5%, r=0.765, p=1.33x10^{-5}; efficiency: R^2=55.8%, r=0.747, p=2.74x10^{-4}).

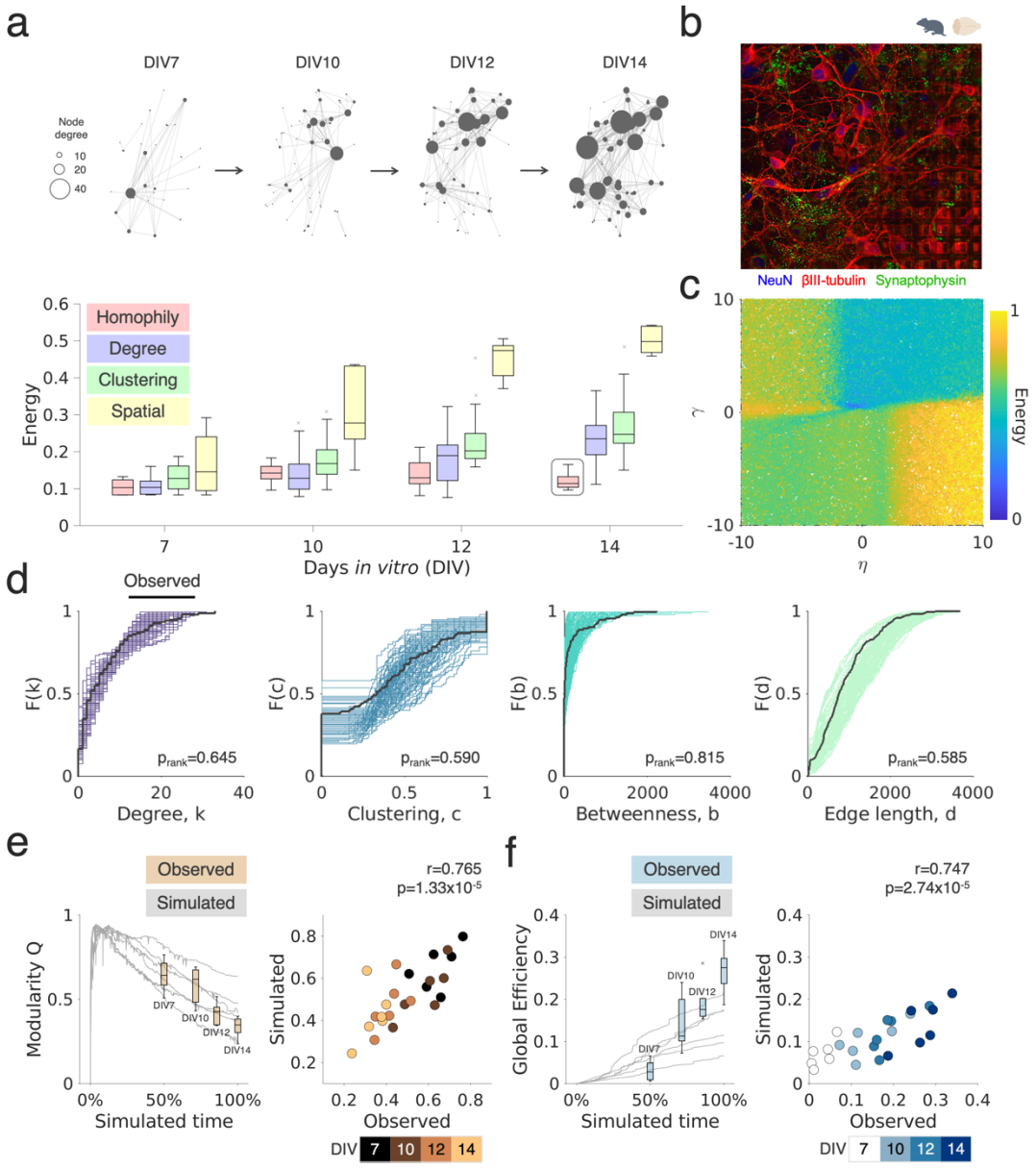


Figure 3.2. Homophilic generative principles underpin rodent primary cortical network development in vitro and explain variance in their maturational trajectories. (a) The top row shows a representative rodent primary cortical (PC) network developing over the first two weeks in vitro on the HD-MEA (DIVs 7, 10, 12 and 14; 50k neuron plating density, where pink nodes correspond to single neurons and the size of each neuron corresponds to its nodal degree). Below we show the energy of the top performing simulated networks, given across the four main categories of studied wiring rules. Boxplots represent the median and IQR. Outliers are demarcated as small black crosses, and are those which exceed 1.5x the interquartile range away from the top or bottom of the box. (b) Immunohistochemical staining of a DIV14 rodent PC culture (control experiment on cover slip), showing staining of NeuN (blue; Neuronal nuclei antibodies), βIII-tubulin (blue) and Synaptophysin (green; Major synaptic vesicle protein p38).

(c) All energy landscapes of the matching generative network model for DIV14 on top of each other. This corresponds to the highlighted box in panel a shown previously. (d) Cumulative density functions (CDFs) of the matching generative model are given, showing that simulated and observed statistics are indistinguishable. CDFs are shown across the four network statistics included in the energy equation of the to$p$=99 simulations for the matching model (best performing generative model) compared to an observed network (black line). Subsequent p values were computed using a Monte-Carlo bootstrapping procedure. (e) Simulated network trajectories were examined to determine if the later generative model's trajectory (DIV14) recapitulated earlier (DIV7, DIV10 and DIV12) observed statistics. Simulated network statistics were derived by computing the modularity Q statistics and global efficiency, at each time-point, within the generative model that was best fit to the DIV14 networks derived from rodent PCs. These were scaled to the developmental time so that observations and simulations could be compared directly. Modularity (left) decreases over time in both simulated trajectories and longitudinal observations. There is a significant positive correlation (right, r=0.765, p=1.33x10^{-5}) between observed and simulated statistics across DIV7 (50% of the DIV14 simulation), DIV10 (71% of the DIV14 simulation), DIV12 (86% of the DIV14 simulation) and DIV14 (100% of the DIV14 simulation). (f) Global efficiency (left) increases over time in both simulated trajectories and longitudinal observations. There is a significant positive correlation (right, r=0.747, p=2.74x10^{-4}) between observed and simulated statistics across DIV7 (50% of the DIV14 simulation), DIV10 (71% of the DIV14 simulation), DIV12 (86% of the DIV14 simulation) and DIV14 (100% of the DIV14 simulation).

3.3.3. Effect of plating density on topology and generative model outcome

So far, we quantified functional connectivity graphs derived from sparse PC neuronal cultures (50,000 cells plated per well). Despite some research on the effect of plating density on the emergence of population activity *in vitro* (Wagenaar et al., 2006), synaptic strength and connectivity (Barral & D Reyes, 2016), there is currently no consensus as to how different plating densities affect neuronal topology. As one critical element of the generative network model is the geometric spacing between neurons, we next probed whether our findings in sparse cultures generalise to networks at higher neuronal plating densities. We, therefore, recorded a second independent dataset of more densely plated rodent PCs (100,000 neurons per well, n=12) in the exact same way as outlined for the first PC rodent dataset and directly compared both densities at DIV14.

We found that only the empirical edge lengths and global clustering differed between sparse versus dense plated PC networks; with dense networks showing relatively shorter connections (p=0.0125) and networks which were more topologically clustered (p=0.0135). All other tested metrics remained indistinguishable (**Figure C7**). Despite these changes across several topological metrics, the global correlational structure of these statistics remained stable (**Figure 3.3a,b**).

Given the topological differences in edge lengths and global clustering, we then asked whether this also translated into significant changes in the energy values among the 13 tested generative network models. In **Figure 3.3c** we show that model energy is unaffected by plating density (p=0.191). In **Figure C8**, we show this broken down by each individual model, in addition to showing that this remains when considering greater numbers of well-performing parameter combinations. All statistical comparisons, for each time-point in the dense PC networks, are presented in **Table C3** showing stability (not only at DIV14, but also at DIV21 and 28). In **Figure 3.3d** we show the energy landscape for both plating densities, which, again, are very similar.

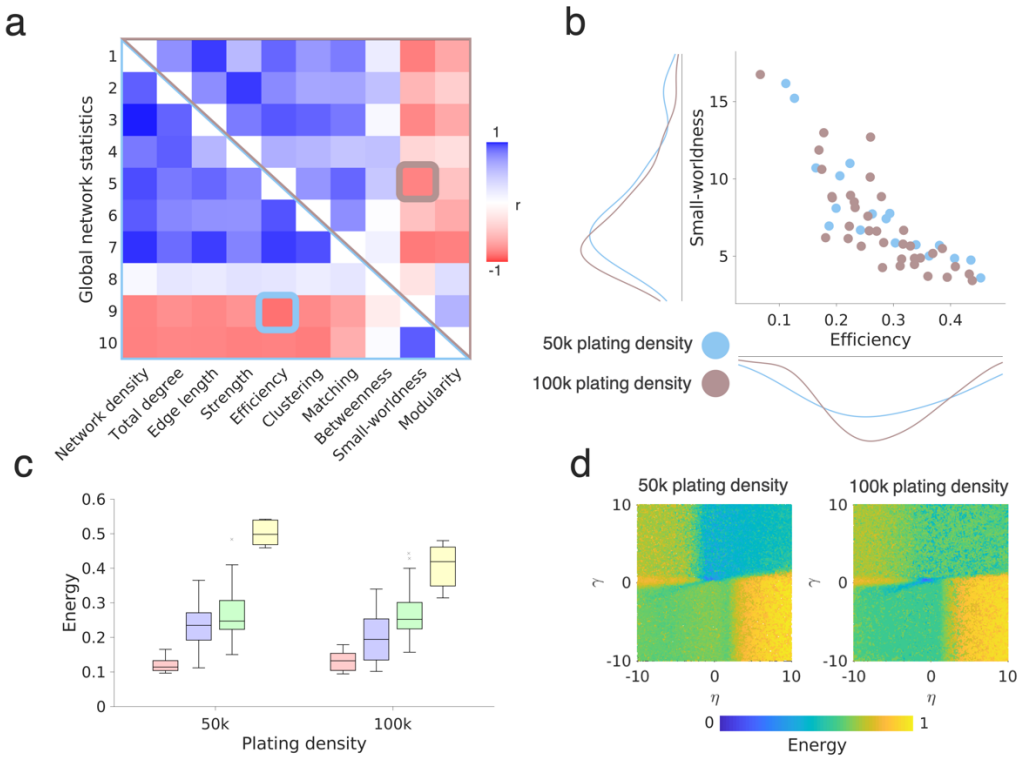


Figure 3.3. Effect of plating density on functional topology and generative network principles. (a) Correlation matrix for global network statistics calculated across 50k (lower triangle) and 100k (upper triangle) PC cultures show a highly similar covariance structure. Colours indicate the Pearson's correlation coefficient. Elements (5,9) and (9,5) of the correlation matrix are further highlighted in the right scatter plot in (b_ showing the correlation between efficiency and small-worldness, for sparse and dense networks. Each point in the scatter plot corresponds to a single network. (c) Both rodent 50k and 100k neuronal cultures can be best simulated by the homophily generative models (DIV14). Note that the leftmost boxplot is the same as that given in Figure 3.2a at DIV14. (d) Energy landscapes for the matching generative network model landscapes for 50 and 100k DIV14. Note that the left landscape is the same as that given in Figure 3.2c.

3.3.4. Topological fingerprints arise from homophilic mechanisms in developing neuronal networks *in vitro*

The results presented so far show that homophily-based generative models produce synthetic networks which are statistically similar to observed functional rodent PCs networks. We have shown this in a way analogous to the macroscopic networks presented in **Chapter 2**. However,

this similarity depends upon the maximum *KS* distance of the four topological statistics as defined in the energy equation. Crucially, this means that while experimentally observed and simulated network statistical distributions mirror each other at the *global* level, the *topological fingerprint* (TF) of these network statistics could differ. That is, nodes within simulated and observed networks could have different *local* relationships to one another, because node-wise local organisational properties are not captured per se by the existing energy equation. Local organisational properties have previously been investigated in terms of how well generative models can recapitulate the locations of organisational features, such as hub-nodes, in the *C. elegans* connectome (Nicosia et al., 2013) or MRI-inferred human brain networks (Arnatkevičiūtė et al., 2021; Oldham et al., 2022).

For example, consider the topological relationship between central and peripheral nodes commonly found in the classical representation of a brain network. Nodes which score highly in centrality measures (e.g., *betweenness centrality* – which determines how many shortest paths pass through – as shown by the red node in **Figure 3.4a**, left) tend *not* to sit within segregated modules, meaning it is common that they concurrently score low in measures of segregation (e.g., clustering coefficients, in which neighbours connect to each other) – and *vice versa* for peripheral nodes (as shown by the green node in **Figure 3.4a**, left). This means that when correlating measures of centrality with measures of clustering across a network, the correlation tends to be negligible or negative (Lord et al., 2017) (**Figure 3.4a**, right).

To assess the ability of generative models to capture these types of local relationships in settings with no anatomical reference space (as neurons are randomly distributed on the HD-MEAs at the time of plating, unlike regions in the macroscopic networks given in **Chapter 2**), we next computed the *topological fingerprint dissimilarity* ($TF_{dissimilarity}$). This demarcates the ability of *in silico* network simulations to recapitulate observe local hallmarks of organisation.

If homophily is a plausible attachment mechanism by which single neurons together form networks, we should expect homophily-based generative models to produce networks with a

local topological structure resembling the observed data. To probe this (dis)similarity, we calculated the $TF_{dissimilarity}$ between each experimentally inferred functional connectivity graph and the best performing simulated network (according to the energy equation), for each of the 13 generative models, across all recording time points (**Figure 3.4**).

Figure 3.4a provides a schematic of how topological fingerprints (*TFs*) were constructed. Results demonstrate that synthetic networks generated with homophilic attachment rules provide the lowest $TF_{dissimilarity}$ from DIV12 onwards (**Figure 3.4b**). Homophilic rules resulted in the statistically smallest $TF_{dissimilarity}$ (e.g., at DIV14; homophilic rules compared to degree rules: p=1.91x10^{-3}, Cohen's d=1.34; homophilic compared to clustering rules: p=1.37x10^{-7}, Cohen's d=2.59). Of note, homophily and spatial rules could be distinguished significantly at DIV12 (p=9.96x10^{-4}) but not at other time-points (e.g., at DIV14; p=0.0952) possibly indicative of spatial costs driving the topological fingerprint at this level of plating density (all statistical findings are provided in **Table C4)**.

Interestingly, replicate analyses in the dense PC rodent dataset at DIV14, 21 and 28 provided almost identical results but with a stronger distinguishability of homophilic rules relative to spatial models (see **Table C5** and **Figure C9**). The left-most panel in **Figure 3.4c** shows the experimentally observed *TF* matrix (averaged over n=6 sparse PC networks); the adjacent panels show the average *TF* matrices for the matching, clustering-average, degree-average and spatial generative models. Depicted are the best performing models within their generative rule class.

Our results highlight the importance of assessing generative network model simulation performance both in terms of overall global topology (*energy*) and in terms of the local topology generated (*topological fingerprint, TF*). We find that homophily models concurrently outperform the other models on both fronts (**Figure 3.4d**, see **Figure C10** for a replication analysis in the dense rodent PC dataset).

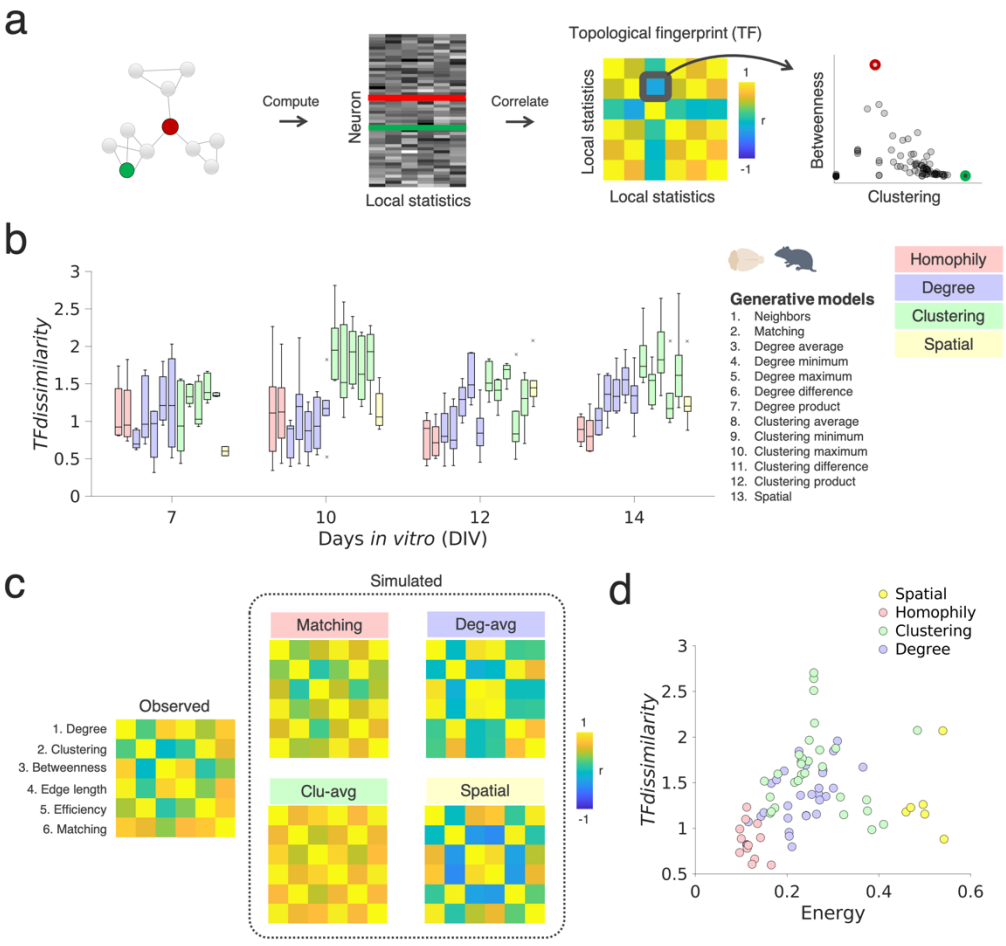


Figure 3.4. Homophilic generative mechanisms best account for local topology in developing neuronal cultures *in vitro*. (a) Schematic illustrating the inference of topological fingerprints. A topological fingerprint matrix is computed as the correlation matrix of several topological statistics at the nodal level (degree, clustering, betweenness, edge length, efficiency and matching). Note that each node here corresponds to a single neuronal unit. The topological correlation matrix (middle left panel) is inferred by correlating each of these local statistical measures. The right panel shows as an example the negative correlation between clustering and betweenness, which captures an aspect of the topological structure of the network shown on the left; the colour bar is clipped to -1 (blue) and 1 (yellow). (b) The $TF_{dissimilarity}$ measures the extent to which simulations capture the topological structure of the observations. Homophily generative models, with spatial models, broadly produce the lowest $TF_{dissimilarity}$, suggesting that it can reconstruct local connectivity patterns in vitro (n=6, 50k plating density; boxplots present the median and IQR; outliers are demarcated as small grey crosses, and are those which exceed 1.5 times the interquartile range away from the top or bottom of the box. (c) Averaged topological fingerprint (TF) matrix for the empirically observed data (left; DIV14), versus the simulated data from generative models with the best fits according to the energy (matching, top left panel), as well as the other best fitting non-homophily

generative models (clustering-average, bottom left panel; degree-average, top right; spatial, bottom right). (d) A scatter plot, with corresponding density plots for each axis, showing the relationship between energy and $TF_{dissimilarity}$ for DIV14 50k rodent cultures broken down by generative rule class (Spatial, yellow. Homophily, red. Clustering, green. Degree, blue). Each scatter plot corresponds to a single simulated network, as measured by the two model fit functions.

3.3.5. Effect of $GABA_A$ receptor antagonism on generative network wiring

GABAergic interneurons act as network hubs (Bonifazi et al., 2009; Cossart, 2014; Hafizi et al., 2021) regulating the synchronisation of spontaneous activity that is critical for formation of connections and plasticity throughout development as well as proper brain function later on (Bonifazi et al., 2009; Cossart, 2014; Le Magueresse & Monyer, 2013; Mann et al., 2009). Here, we examined whether the perturbation of $GABA_A$ receptors would alter formation of functional connectivity, and thereby the outcome of the generative models. To examine this, we cultured sparse rodent PC cultures (50,000 cells per well) under chronic application of gabazine, a selective $GABA_A$ receptor antagonist.

If activity-dependent mechanisms, and spontaneous activity more generally, serve to form functional networks according to wiring principles such as homophily, functional GABAergic interneurons may therefore be key to the implementation of homophily throughout development. By chronically blocking $GABA_A$ receptors, relative to the control cultures, we hypothesised that either homophily would subsequently fail to be implemented, or its implementation would change, leading to differences in network topology. The first hypothesis would be supported by higher energy levels of homophilic rules in gabazine cultures relative to control homophilic rules or compared to other generative rules. The latter hypothesis would be supported by changes in the wiring parameters of a well-fitting homophilic model. Furthermore, after chronic application of gabazine, at DIV14 we washed out gabazine cultures and re-recorded at DIV15 to determine the extent to which network topology recovered in the same cultures.

In **Figure 3.5a** we show representative population activity plots for control and gabazine-treated cultures. We find that homophily generative attachment rules remain the best fitting models for all control, gabazine and washout cultures (**Figure C11a,b**). However, crucially, gabazine cultures are not as well approximated by generative models as controls (p=0.0292; **Figure 3.5b**), which suggests that homophilic generative models cannot approximate the topology of gabazine cultures to the extent that they can control cultures. This finding supports our first hypothesis. Moreover, we find that when gabazine is washed out in the same cultures, the energy obtained from them becomes indistinguishable from controls (p=0.285), demonstrating a very direct specific effect that gabazine is having on network topology due to its previously highlighted mechanism of action. In **Figure C11c** we show how, within the homophily generative models, gabazine cultures are best fit to when the η wiring parameter (which governs the extent to which distance influences wiring probability) increases so that it is closer to zero (i.e., moving from more negative to less negative; p=0.0360) while γ (which varies the extent to which homophily influences wiring probability) remains unaffected (p=0.456). This means that gabazine seemingly weakens the spatial extent of functional connectivity wiring, enabling relatively longer distance connections to contribute towards the network's topology.

Our models may not only reflect differences in wiring parameterisation that are altered by gabazine, but more fundamental differences in neuronal variability that are elicited via changes in the neural dynamics. Within the generative model, connections form via the continual negotiations between the modelled cost and self-similarity that is present between all present neurons simultaneously. If there are clear *relative* winners in this negotiation (e.g., connections that are lower cost than all others, and connections that are more homophilic than all others), these connections are formed with a greater probability ($P_{i,j}$) relative to others. This is because the probability of a connection is proportional to the wiring parameters. Mathematically, the $P_{i,j}$ distribution would look like a canonical lognormal distribution that is found at many anatomical and physiological levels of the brain (Buzsáki & Mizuseki, 2014; Loewenstein et al., 2011) where a small number of possible candidate connections elicit a higher wiring

probability while the majority remain low. However, in the example of gabazine - which increases neuronal synchrony - neurons necessarily entail decreased variability in their dynamics. As a result, we hypothesised that this would lead to a flattening of this lognormal distribution, meaning that the resultant topology became, necessarily, more random, distributed and disordered. Indeed, in **Figure 3.5c** we show this to be the case - gabazine exhibits a flattened wiring $P_{i,j}$ distribution relative to both controls (mean comparisons over development; p=4.49x10^{-94}, Cohen's d=0.550) and when these same cultures are washed-out (mean comparisons over development; p=1.31x10^{-19}, Cohen's d=0.196). In **Figure C11d** we show the entire $P_{i,j}$ distributions over the course of simulated network development. This finding suggests that gabazine alters the network wiring distribution as to become less specific in its wiring, rather than being specific to a smaller number of candidate neurons that are deemed particularly valuable to wire with.

To summarise, we find that homophily generative models are the best fitting model for chronically-treated gabazine cultures but do achieve worse model fits than control cultures. At the level of model parameters, the homophily γ parameter remains stable over all conditions, but it is η - which alters the spatial extent over which wiring is constrained. When gabazine is washed out of DIV14 cultures and recorded at DIV15, numerous topological characteristics are recovered. Within our simulations, this washing out of gabazine is reflected as the network recovering its wiring probability distribution towards a more lognormal distribution.

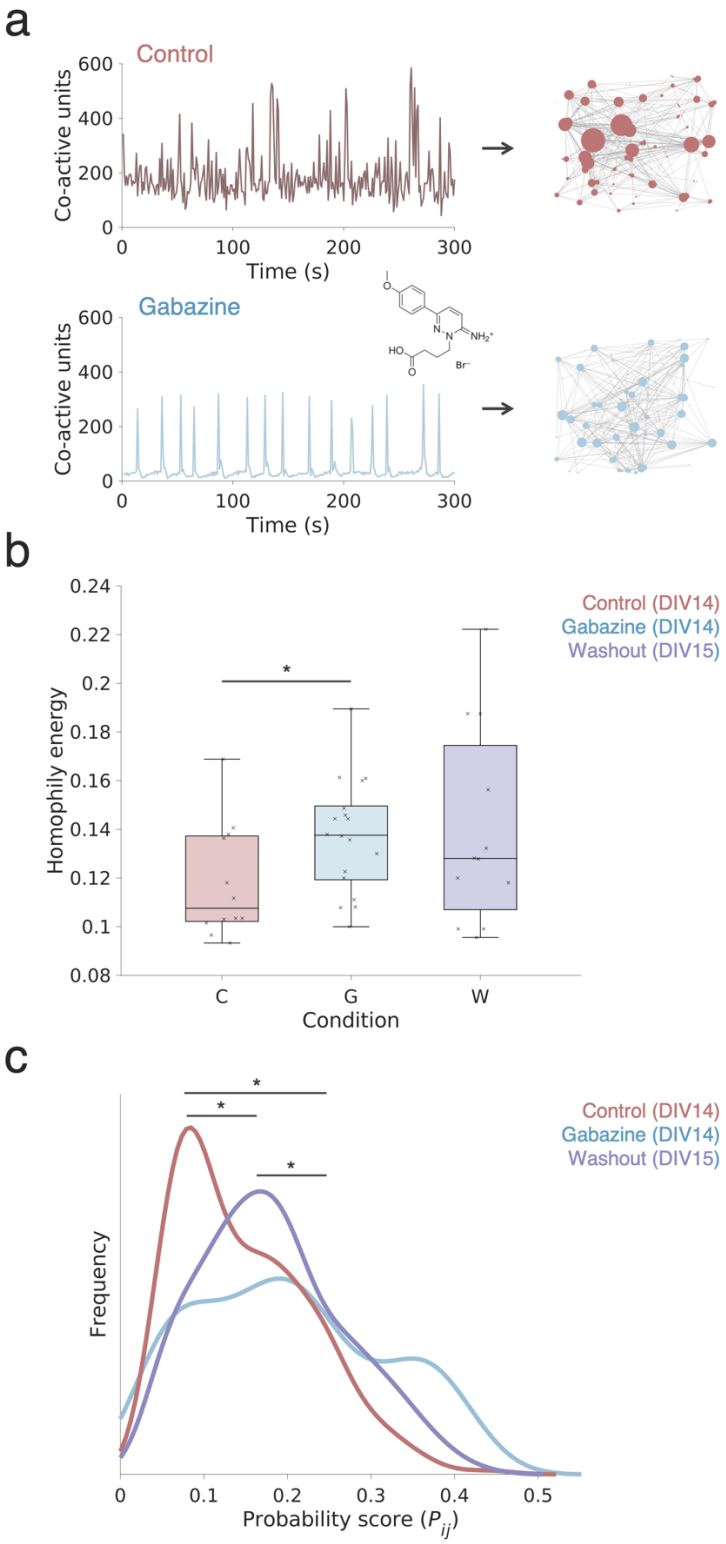


Figure 3.5. Chronic block of $GABA_A$ receptors changes neuronal variability, leading to altered probabilistic wiring. (a) Representative neuronal population activity of a control culture (red, top) and a culture treated chronically with gabazine (bottom, blue) at DIV14; Network activity is synchronised into short bursting activity in the gabazine condition, relative to activity seen

in the control condition. (b) While homophily rules attain the lowest energy compared to all others tested (Figure C11a), the homophily energy for gabazine cultures is greater, suggesting that homophily models cannot approximate the topology of control cultures to the same extent. The energy of the top performing simulated networks are shown. Boxplots represent the median and IQR. (c) The average probability score distributions ($P_{i,j}$) for control (red) and gabazine (blue) sparse PC networks across simulated network development. The average gabazine probability distribution is shifted to the right (i.e., higher scores) and flatter compared to controls. This indicates that wiring becomes less determined across all timepoints, leading to more random outcomes. In Figure C11d,e we show all distributions used to construct these averages.

3.3.6. Generative wiring principles in human neuronal cultures *in vitro*

Finally, we probed whether homophilic wiring principles can be generalised across neuronal networks derived from other species and cultures composed predominantly of specific cell types. In order to test this, we first applied generative network models to human iPSC-derived neuron/astrocyte co-cultures and analysed these networks at DIV28, a time point at which cultures reach a state of relative maturity (Ronchi et al., 2021). Note, this data was not tracked and electrodes selected based on the activity scan at DIV28. The human dataset comprised glutamatergic neurons (GNs, n=8), motor neurons (MNs, n=7), and dopaminergic neuronal cultures (DNs, n=6). In addition, we also studied slice cultures derived from 4-month-old human embryonic stem cell-derived cerebral organoids (hCOs, n=6 slices). Previous studies have indicated that hCOs develop functional networks with increasing complexity as early as 90 days *in vitro* (Giandomenico et al., 2019; Szebényi et al., 2021).

Figure 3.6a provides an overview on the human and rodent neuronal electrophysiological data. Following a t-distributed stochastic neighbour embedding (tSNE) analysis, we find that networks can be clustered according to their overall spike train dynamics (i.e., autocorrelograms derived from the aggregated spike train activity of each network, shown in **Figure C12a**) and grouped according to the respective cell lines. For example, MN dynamics appear closest to rodent primary cortical neurons (PCs) in the tSNE space, and this is reflected by their similarity in underlying spiking dynamics (**Figure C12b**). **Figure 3.6b** shows

immunohistochemical stainings of a DIV21 human iPSC-derived DN culture expressing neuronal and astrocytic markers (MAP2, GFAP, and TH) and **Figure 3.6c** shows stainings for hCOs slices (Tau, NeuN, and GFAP), both derived from control experiments. Representative examples of the observed differences in population activities across different human neuron cultures are depicted in **Figure 3.6d**. As for the rodent PC cultures, we constructed functional connectivity graphs for all human neuron cultures, and assessed key connectivity metrics and topology; **Figure 3.6e** and **Figure 3.6f** show these iPSC-derived network statistics respectively, and **Figure 3.6g** provides an example of a single constructed hCO functional network. It is notable that human iPSC-derived neuronal networks did not differ significantly in average STTC (ANOVA, p=0.0912), however, we found a significant difference in network density (p=1.02x10^{-4}) and topological metrics, such as the small-world index (p=1.51x10^{-4}) whereby MNs show the highest network density, in addition to a small-worldness within the general range of observed rodent PCs, followed by GNs.

Figure 3.6h depicts the generative model findings across all human monolayer and organoid cultures. We find that the findings in generative models approximately mirror the similarity in their underlying spiking dynamics (**Figure 3.6d**), in that human MNs show an energy profile across generative models prominently aligning with empirical rodent PCs, followed by GNs (**Figure 3.6h**, left and middle-left; homophilic rules p<3.70x10^{-3} for all pairwise comparisons and Cohen's d>0.772 reflecting large effect sizes). In contrast, DNs and hCOs, which show differing underlying dynamics, exhibit a very distinct profile with the degree average method achieving the lowest energy, but no non-spatial class of generative rules providing a statistically best model fit (**Figure 3.6h**, middle-right and right; all p>0.054). It is of note that hCOs exhibit significant variability in their spiking dynamics and topology, and provide an energy profile much more consistent with randomly rewired graphs (**Figure B5**, right). This variability in hCOs could be due to a few factors. One reason could be that their relatively nascent activity, compared to rodent PCs and human iPSCs, leads to a greater susceptibility to noise, leading to greater variability. Alternatively, it could be that, irrespective of possible confounds, hCO networks happen to form relatively small networks

meaning it is simply harder to draw confident conclusions about their underlying topology. Both of these reasons may be linked (i.e., the size of the network is smaller due to having less activity). In any case, the results in hCOs are more inconclusive. All statistical findings are provided in full in **Table C6**.

In summary, generative network models based on homophilic wiring mechanisms provide a good explanation for network formation of *in vitro* functional connectivity graphs, depending seemingly upon their underlying spiking dynamics – where some cell-types (tested in isolation) are better approximated than others. Human GNs and, in particular, MNs show the closest resemblance to rodent PCs and as such mirror their underlying homophilic generative mechanisms. Results in human cerebral organoids are as yet more inconclusive, likely due to

their observed variability in functional connectivity.

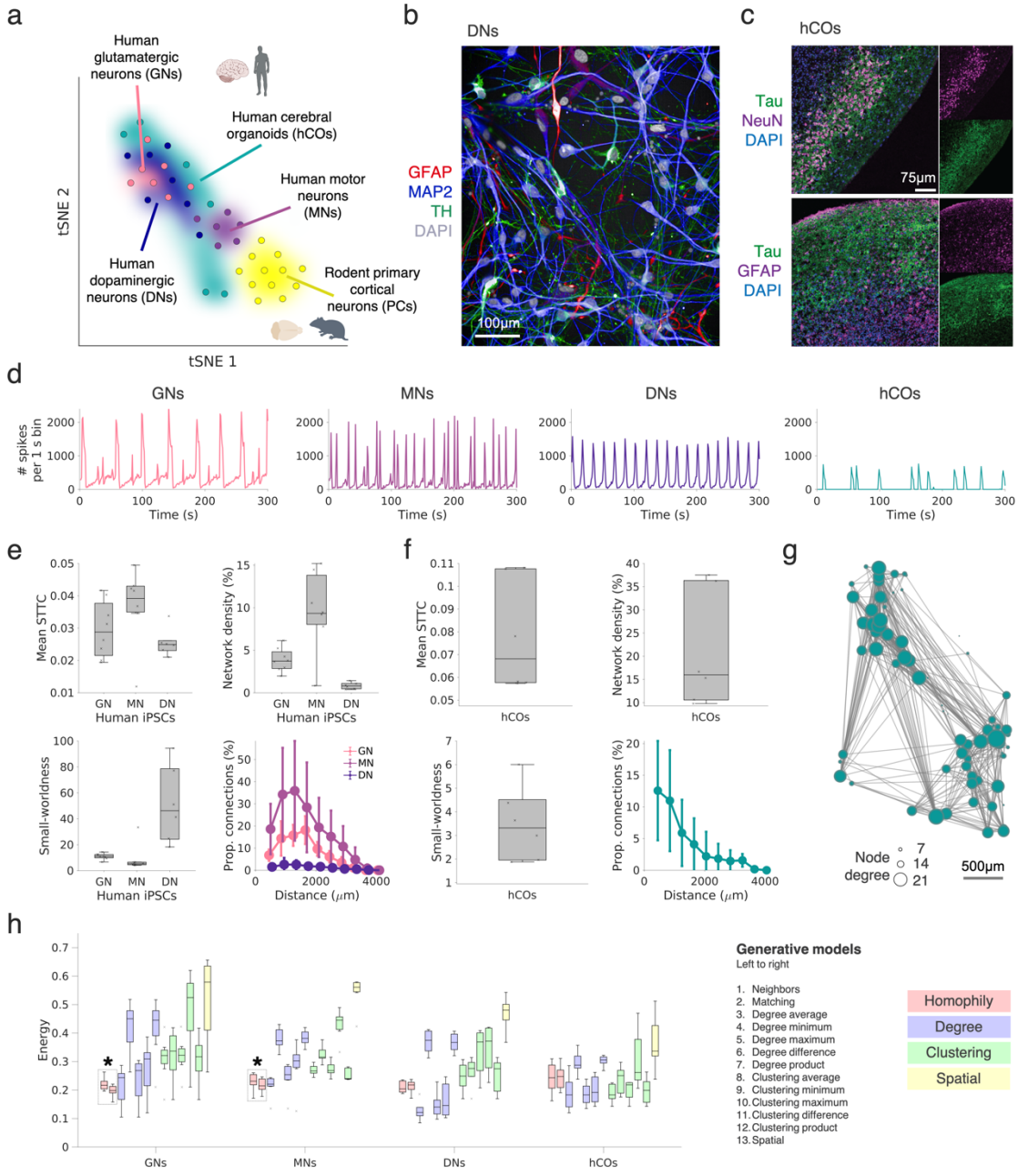


Figure 3.6. Generative network modelling across human neuronal cultures and cerebral organoids. (a) A Clustering of human and rodent neuronal cultures based on T-distributed stochastic neighbourhood embedding (tSNE). Each dot corresponds to one culture; similarity was estimated by correlating the spike train autocorrelograms (activity of all units concatenated) across all datasets and by then applying the tSNE rodent primary cortical neurons (yellow, 100k plating density, n=12), human iPSC-derived dopaminergic (dark-blue, n=6), motor (purple, n=7) and glutamatergic neurons (pink, n=8) form clear clusters in 2D tSNE space; human ESC-derived cerebral organoids (green, n=6) were more scattered. (b) An

immunohistochemical staining of a single dopaminergic culture, which highlights the expression of GFAP (red), MAP2 (blue), TH (green; tyrosine hydroxylase) and DAPI (grey; 4',6-diamidino-2-phenylindole). (c) An immunohistochemical staining of a single hCO slice for Tau (green), NeuN (purple; top), GFAP (purple; bottom) and DAPI (blue; bottom). (d) Representative network population activity for a spike-sorted network of glutamatergic (left), motor (left middle), dopaminergic (right middle) iPSC-derived neuronal cultures and hCO (right). (e) Global topological measures of observed DIV28 human iPSCs in terms of the mean STTC (top left), network density (top right), small-worldness (bottom left) and proportion of extant connection by distance (bottom right). (f) Functional connectivity measures and small-worldness of observed hCO slices including: average STTC (top left), network density (top right), small-worldness (bottom left), and connectivity as a function of distance (bottom right). (g) Functional connectivity graph inferred from a 120-day-old hCO slice. (h) Fits across generative model rules in human iPSC-derived neuronal networks (DIV 28; n=6 for each cell line, left, middle left and middle right) and hCOs (n=6, right). The energy of the top n=1 performing simulated networks are shown. Boxplots represent the median and IQR. Outliers are demarcated as small black crosses, and are those which exceed 1.5x the interquartile range away from the top or bottom of the box. The asterisk reflects a significant comparison at $p<0.05$ for homophily compared to the next best rule (degree). All statistics are provided in Tables C6.

3.4. Discussion

In the current study we applied HD-MEA large-scale electrophysiological recordings to track and characterise single-unit functional connectivity as neural networks develop *in vitro*. Moreover, we systematically tested which candidate topological attachment mechanisms could explain this developing self-organisation, using generative network modelling to create *in silico* models of network formation.

In line with previous work, and extending the previous chapter, we found that functional connectivity increased with development, and that developing neuronal networks *in vitro* exhibited canonical characteristics of complex network architecture (Downes et al., 2012; Schröter et al., 2015). A central aim of our work was to examine which attachment rules would best account for this developing topology. As shown in **Chapter 2**, at the macroscopic level, a consistent result in whole-brain connectivity research has been that homophily-based generative models best recapitulate the topology of human brain connectomes (Betzel et al., 2016; Oldham et al., 2022; Vértes et al., 2012; Zhang et al., 2021). Here, we show that the

same is true for the functional connectivity graphs of developing neuronal networks *in vitro*. Crucially, by combining multiple ways of estimating model fits, we show that homophilic wiring principles can not only account for the distributions of topological statistics at the *global* level, but also for the complex *local* organisation of these statistics, as quantified by their *topological fingerprints*.

Why does homophily offer a dominant account of topology, beyond alternative models? In **Chapter 2**, we suggested that the observation that homophily is somewhat related to the network communicability (Estrada & Hatano, 2008) may be important in interpreting why it achieves good model fits to empirical data. Let us now consider briefly how they relate to each other, and why this may be important. First consider a binary adjacency matrix, A, that defines connections between nodes (where the weight of the edge between i and j, as element $A_{i,j}$) in a network. To compute the distribution of paths on this network after n steps, it is necessary to compute the matrix A^n. An interesting observation is that the element $A^n_{i,j}$ is the number of walks of length n from node i to j. For example, naturally, the matrix A^1 will provide all paths of length 1 between each node – it is simply the original network. But A^2 will show you all paths of length 2 between each node on this network, and so on for any value. To evaluate all possible walks, you therefore need to calculate the weighted sum of all possible walk between i and j. This is network communicability; it is simply the weighted sum all of possible walks between nodes i and j:

$$\sum_{n=1}^{\infty} \frac{A^n}{n!} = e^A = 1 + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots \quad (5)$$

One way to interpret this is to think of communicability as the extent to which, when moving from node to node, a unit of information would flow through a network. In a small network, between two nodes, the shorter path would necessarily be taken more often relative to longer paths. However, as the network gets larger the disparity between shortest paths and communicability widens because there is more scope for alternative routes. Critically, this is a totally local computation – no knowledge of the global structure of the network is required to

undertake this process – highlighted by the fact that communicability can be computed directly from the binary adjacency matrix itself (as highlighted in the above equation). This idea is expanded upon further in the next chapter.

How is this notion of local communication relevant to the present work? Under the hypothesis that neurons update their connection preferences with other neurons as a function of their communicability (or to equivalently reduce its *intuitive* distance with all others (Baram et al., 2018)), a number of outcomes may logically occur indirectly. One outcome is that a preference may emerge for neurons to wire with others that have a topology similar to themselves, because multiple paths to the same node will necessarily improve the likelihood that a walk to that node will occur. This directly resembles a homophilic principle. Indeed, homophily (as operationalised under a matching model) relates to a narrow-version of communicability, in which nodes preferentially wire with other nodes that are commonly two paths away (i.e., A^2) rather than considering the weighted sum of all walks. But, as stated previously, because the shorter paths contribute more than longer paths they may indeed approximate each other very closely. Therefore, homophily likely serves as a heuristic – much like in social networks – (McPherson et al., 2001; Talaga & Nowak, 2020) to facilitate effective communication capable of producing navigable small-world networks (Malkov & Ponomarenko, 2016; Seguin et al., 2018).

The fact that both structural and functional networks can be approximated via the same model seems to indicate that – despite being very different types of data – these networks are under the same fundamental economic constraints. This lends credence to the idea that this model is approximating a general principle that can be applied to developing neural systems, rather than any one particular type of data. That is, this concept of locally communicating units within a network, and its role in shaping topology according to self-similarity, may hold for any interacting developing system (and particularly in the context of large graphs (Son et al., 2021)). This resonates with accounts of Hebbian learning (Hebb, 1949; Vértes et al., 2014) and spike-timing dependent plasticity (STDP) (S. Song et al., 2000) where neurons wire with each other as a function of *similarity* (H. F. Song et al., 2014) to themselves (e.g., concurrent

or temporally precedent neuronal firing, respectively) provided that neurons are sufficiently close in space. It is of note that communication may extend to other means of (local or non-local) cell-to-cell interactions (Armingol et al., 2021), such as via para-, juxta- or endocrinological signalling, that together guide neural circuit construction *in vivo* (de Wit & Ghosh, 2016; Hassan & Hiesinger, 2015). In **Chapter 4**, some of these ideas are explored further by incorporating communicability within an artificial neural network model.

An observation made in this work is that while homophily models tend to perform best, other models – in particular the degree-average (deg-avg) model – tends to consistently follow (e.g., see **Figure C4**). Conceptually, the degree-average generative model will increasingly prefer for the formation of connection if *both* neurons (hence the average) have large numbers of connections, given that $\gamma>0$. In this way, it captures a narrow sense of self-similarity (albeit, only when both neurons have lots of connections) in that high-degree nodes prefer to wire with other high-degree nodes. Most importantly however, the degree-average model (as with all other models apart from homophily) treats the underlying computations for how networks form connections as entirely independent. That is, the computation is made not with direct respect to any node pair – it is made independent on both neurons before performing some other operation (e.g., taking the average, or maximum). In this way, unlike homophily – which is a function of the direct relationship between neurons – these models miss a crucial aspect which how the communication between neurons must in some way constrain the networks self-organisation over time. The extent to which neuronal development is itself intrinsically dependent upon its direct connections to other neurons is unclear, although it is likely to become increasingly important over developmental time – perhaps reflected by the later divergence between homophily and other models (e.g., **Figure 3.2**).

A key advantage of our *in vitro* neuronal network modelling approach is that it can be perturbed under varying experimental conditions, testing how specific cellular mechanisms may change the outcome of generative models in a predictable fashion. In line with previous findings (Mann et al., 2009), the inhibition of $GABA_A$ receptors led to increased neuronal synchronisation (Assenza et al., 2011). Within our wiring equation, the two wiring parameters

influence to what extent costs and topological homophily, respectively, shape wiring probabilities. Importantly, this wiring is probabilistic – the greater in magnitude the wiring parameters, the more that networks are "determined" by their economic negotiations. In the gabazine condition, the homophily generative model provided a subtly worse model fit and exhibited differences in the spatial constraints on wiring. We find that this leads to a shift in the probability distribution, such that more neurons achieved higher probability scores (i.e., a flattening of the distribution). One interpretation for this is observation is that increased synchronicity, caused via the blockage of inhibitory activity via $GABA_A$ receptor antagonism, leads to decreased specificity in neuronal wiring preferences simply because synchronicity lessens the differences between neurons in terms of their activity. That is, the flattening of the distribution shown in **Figure 3.5c** within the gabazine condition, relative to the control's log-normal distribution (Buzsáki & Mizuseki, 2014), is itself a reflection of decreased relative variability in the spiking dynamics within the neuronal network. Almost inevitably, as variability decreases this leads to an increasingly randomised topology. It is an open question precisely how this affects the functioning of the network. Interestingly, as shown in **Chapter 2**, in addition to previous generative network modelling work at the whole-brain scale, lower magnitude wiring parameters are associated with age (**Figure 2.3**), poorer cognitive scores (**Figure 2.4**), and a diagnosis of Schizophrenia (Vértes et al., 2012; Zhang et al., 2021). This may suggest convergent evidence for how developmental randomness, intrinsic to how developing parts interact with each other, may help or hinder outcomes (Hiesinger & Hassan, 2018). A remaining challenge in the field is to be able to directly parse the extent to how stochasticity versus specific economic trade-offs may, together or independently, influence network outcomes under different conditions. Future methodological work should explore how various multiplicative (Betzel et al., 2016; Vértes et al., 2012) or additive (Oldham et al., 2022) generative models may be used to understand these shaping factors.

Importantly, it has to be noted that there is currently no consensus as to how functional connectivity can be inferred from the spontaneous activity of neurons developing *in vitro* (Magrans de Abril et al., 2018). In the present study, we utilised the spike time tiling coefficient

(STTC; Cutts & Eglen, 2014) which was developed to improve on some of the limitations of traditional metrics for the coupling between neurons, such as the correlation index (Cutts & Eglen, 2014). In this work our goal was not to infer synaptic connections, direct structural connections or reconstruct the underlying circuit of synaptic or autaptic connections present. Nevertheless, in **Figure C5** we showed our results were replicated to transfer-entropy based methods (Wong et al., 1993). A further limitation relates to the extension of our modelling to human monolayer and organoid cultures. Crucially, human iPSCs are single cell-types that are purified and artificial. As reflected by the lack of topology exhibited in DNs (which indeed did not show homophilic wiring) there is no surprise they appear different. A more veridical account will likely come from cultures with mixed iPSC lines exhibiting more complex spiking dynamics arising from cell-type heterogeneity – which is highly likely to contribute to energetic efficiency and topological organisation (Perez-Nieves et al., 2021). Relatedly, another limitation relates to the translatability of the ontogeny of *in vitro* systems to empirical *in vivo* timelines of brain development. For example, while recent studies aim to characterise the ontogeny of iPSCs (e.g., Izsak et al., 2022) there is still no clear one-to-one mapping between *in vitro* and *in vivo* observations in terms of their developmental timelines. This mapping is challenging to accurately generate because there remains many biological processes, such as vascularisation and specific developmental cues, that are not captured in *in vitro* systems (see Heide et al., 2018; Vitrac & Cloëz-Tayarani, 2018 for reviews that consider these pitfalls in iPSCs and hCOs respectively). However, this is a fast-moving field, particular with relation to hCOs. Research has advanced from showing that hCOs can recapitulate gene expression programmes existing within *in vivo* foetal neocortical development (Camp et al., 2015) to, more recently, demonstrating that numerous postnatal molecular developmental stages, such as NMDA receptor subunit switches, exist with these *in vitro* systems (Gordon et al., 2021). Nevertheless, a clearer one-to-one ontogeny from *in vitro* systems to *in vivo* observations will enable a significant jump in enabling translation. Finally, it is also possible that the spatial extent of the recording from 3D hCOs onto the 2D HD-MEA system may have limited our network inferences.

In conclusion, we find that the complex topology of developing rodent and numerous human neuronal networks *in vitro* can be best simulated by a simple homophily generative model, where neurons aim to maximise locally-shared connectivity within an economic context. With this, and prior research at the macroscopic level in mind, we suggest that homophily wiring rules provide an adequate isomorphic explanation for any decentralised, locally-computing, developing system.

4. CHAPTER 4

SPATIALLY-EMBEDDED RECURRENT NEURAL NETWORKS EXHIBIT NUMEROUS BRAIN-LIKE STRUCTURAL AND FUNCTIONAL PROPERTIES

Content included in this chapter was published in bioRxiv, 2022, and is currently in review at Nature Machine Intelligence: Spatially-embedded recurrent neural networks reveal widespread links between structural and functional neuroscience findings. Jascha Achterberg, **Danyal Akarca***, DJ Strouse, John Duncan† & Duncan E. Astle† (*Co-lead authors, †Co-senior authors).*

An implicit assumption when studying network neuroscience is that neurobiological networks function in context of the local-constraints to their topology. These constraints likely have direct consequences for how they function. However, in absence of a computational framework able to integrate both structure and function directly, this is hard to test concretely. Moving beyond studies contained in **Chapter 2** and **3,** which used generative modelling of structural and functional networks across scales, in the current chapter we introduce an artificial neural network model bringing together both structural constraints and functional goals within a single objective function, that is learned. We show that by training this neural network, which communicates locally within a 3D Euclidean space, canonical structural and functional findings found in brain networks tend to coalesce. These findings include modularity, small-worldness, homophilic generative mechanisms and functional clustering in space. We argue that this approach can serve as a powerful explanatory tool to better understand how local and global constraints shape the topological and functional outcomes of neural networks that solve tasks.

4.1. Introduction

Neurobiological networks develop and learn to solve problems. As a result, many observed stereotypical structural and functional features of neurobiological networks may be due to having achieved an optimal solution to the common problems (and trade-offs) faced over its evolution and lifetime. At a structural level, one problem brain networks overcome is how to optimally allocate metabolic resources to grow and sustain the network in its physical space. At a functional level, brain networks solve the problem of processing information and appropriately responding to stimuli, given its context and demands (Stokes et al., 2013). While solving these problems, the communication dynamics across the network are simultaneously constrained by the connectivity local to each of its constituent neurons (Batista-García-Ramó & Fernández-Verdecia, 2018). As a result, some of the most basic features of brains, like its sparse (Rosen & Halgren, 2022), small-world (Bassett & Bullmore, 2017) and functionally modular structure (Gallos et al., 2012; Meunier et al., 2010) may result from continual optimisations to problems that they commonly face.

Decades of neuroscientific research has revealed that many of these motifs, like small-worldness and modularity, are common across both human brains (Sporns & Kötter, 2004; Wei et al., 2017) and the animal kingdom (Heuvel et al., 2016). But what key principles shape their functional organisation (Avena-Koenigsberger et al., 2015; Pessoa, 2014)? Computational models have played a key role in addressing this question. For example, as highlighted in prior chapters, modularity in cortical connectivity can in principle arise through the spatial cost of growing a network (Kaiser & Hilgetag, 2004) or via having to regularly task-switch to achieve different goals within an embodied environment (Kashtan & Alon, 2005). Functionally, simple neural network models have shown that orthogonal population dynamics can arise through optimising task performance (G. R. Yang et al., 2019) and predictive coding can occur, without hard-wired constraints, simply via minimising neuronal-unit energy usage in a recurrent neural network (RNN) (Ali et al., 2021).

Is it possible to directly link the problems faced by brain networks, in terms of both their structural and functional optimisations, via analogous computational modelling techniques? Recent interest in reservoir-like networks have made an important step towards actively linking brain structure to functional modelling. This approach works by implementing the structure of connectomes as a fixed hidden layer, with the overall system being trained via a single read-out layer (Goulas et al., 2021; Kawai et al., 2019; Suárez et al., 2021). These models highlight how the macroscopic organisation of networks may have achieved a topology that (within a reservoir framework) can be interpreted as being optimally configured to serve some computational goal such as memory capacity, given its wiring costs (Kawai et al., 2019; Suárez et al., 2021). However, the anatomical substrate of the brain does not explain the whole picture; it likely conditions function, but does not determine it (Batista-García-Ramó & Fernández-Verdecia, 2018). Overall, structure and function likely to influence each other dynamically over time (Mareschal et al., 2007). The question is how do we model this?

We argue that in order to better understand the complex interactions between structure and function in neural networks, it will be necessary to incorporate both within a single coherent model that allows networks to dynamically trade-off between its different objectives in real-

time. To achieve this, we introduce the spatially-embedded recurrent neural network (seRNN). Similar to the trade-offs that many biological brain-networks face, seRNNs are optimised to solve a task in which it makes decisions to achieve functional goals. At the same time, we provide its constituent neurons an economic challenge. The network as a whole must learn to attain a functional goal, whilst each neuron is forced to balance their resources as it updates its connectivity. More specifically, seRNNs are embedded in both 3D Euclidean space and topological space, to mirror the fact that it is metabolically costly for a neuron to sustain connections over long physical distances and that communication between neurons is only possible via direct and indirect connectivity. By allowing seRNNs to dynamically manage both structural and functional objectives simultaneously, we find that a multitude of features commonly observed in neurobiological networks naturally emerge.

4.2. Methods

4.2.1. Task paradigm

We presented RNNs with a simple one-choice inference task. The RNN first observe stimuli A for 20 time-steps, followed by a delay for 10 time-steps, followed by stimuli B for 20 steps. They must then make one choice. This setup can be interpreted as a one-step navigation task, where networks are presented with the goal location (Stimuli A) followed by possible choice directions (Stimuli B). The correctly made choice is in the direction closer to stimuli. **Table 1** outlines all possible trials and defines whether the given trial is included in the baseline version of the task, compared with the harder version – which are regimes tested later in the study. A visualisation of the task set up is shown below in **Figure 4.1a**.

Problem	Stimuli A	Stimuli B	Correct Choice	Baseline Task
1	(Right, Up)	Left, Right	Right	No

2	(Right, Up)	Right, Down	Right	Yes
3	(Right, Up)	Up, Down	Up	No
4	(Right, Up)	Up, Left	Up	Yes
5	(Right, Down)	Left, Right	Right	Yes
6	(Right, Down)	Up, Right	Right	No
7	(Right, Down)	Up, Down	Down	No
8	(Right, Down)	Left, Down	Down	Yes
9	(Left, Up)	Left, Right	Left	Yes
10	(Left, Up)	Left, Down	Left	No
11	(Left, Up)	Up, Down	Up	No
12	(Left, Up)	Up, Right	Up	Yes
13	(Left, Down)	Left, Right	Left	No
14	(Left, Down)	Up, Left	Left	Yes
15	(Left, Down)	Up, Down	Down	No
16	(Left, Down)	Right, Down	Down	Yes

Table 4.1. Stimuli used for the current one-step inference task. Those marked "Yes" for the baseline task are utilised throughout the main work. Later in the study we also make use of the non-baseline combinations to increase the difficulty of the task for a small new-set of RNNs.

All stimuli are one-hot encoded with a vector of eight binary digits. The first four define the goal locations and only one of the four digits would be set to one during the goal presentation. The latter 4 binary digits each stand in for one allowed choice direction and two choice directions would be set to one during the choice options presentation. Gaussian noise with a standard deviation of 0.05 was added to all inputs.

4.2.2. Recurrent neural network (RNN) modelling

We will now introduce how the RNN architecture was constructed, and how we formed the optimisation problem that incorporates both task-solving with structural objectives.

4.2.2.1. THE RNN ARCHITECTURE

We fixed all recurrent neural networks to have 100 units in the hidden layer and defined each unit by the same basic set of equations:

$$h_t = ReLU(W_x x_t + W_h h_{t-1} + b_h) \tag{6}$$

$$\hat{y}_t = \sigma(W_y h_t + b_y), \tag{7}$$

where x_t is a 1x8 binary input vector at time t - the top four elements correspond to the goal location while the bottom four elements correspond to the choices presented. W_x and W_y are 8x100 input and output layer weight matrices (i.e., input and output, to and from, all neurons), initialised using the Xavier method. h_{t-1} and h_t are each 1x100 vectors reflecting the activation of each hidden unit at time t-1 and t respectively, initialised with zeros. W_h is the hidden layer weight matrix (which we later return to), with an orthogonal initialisation. b_h and b_y are 1x100 vectors of hidden and output layer biases which are initialised with zeros.

$ReLU$ refers to the rectified linear unit activation function to compute the activation of each hidden unit, and σ refers to a softmax activation function to compute the network's prediction $\hat{y}_t$. A schematic illustration of a hidden unit is given in **Figure 4.1b**.

4.2.2.2. CANONICAL RNN SETUP

Our work aims to create an optimisation process for a general RNN which allows flexible trade-offs improving task performance within the costs resulting from existing in a biophysical space. We hence seek to incorporate both structural and functional objectives in the network's optimisation process.

How do we achieve this? In the canonical RNN setup, all the network's trainable parameters (or in this case, weights; θ) are optimised to minimise the difference between some predicted output ($\hat{y}$) and the observed output (y). So, on an abstract level, the RNN is a function which uses its weights to map an input to a prediction (i.e., we provide an input x_i into the network to receive a prediction $\hat{y}_i$):

$$f(x_i,\theta) = \hat{y}_i \tag{8}$$

We now would want to optimise the weights to minimise the errors of the predictions of the networks, meaning that we aim to minimise the difference between the predicted value $\hat{y}_i$ and the target value y_i. To do this, we define a loss function ($\mathcal{L}$) which defines precisely what should be minimised in order to optimise task performance. Here we utilise the cross-entropy loss ($\mathcal{L}_{CE}$), which calculates the total entropy between the two distributions. This constitutes the task-related component of the objective function. Put simply, minimising $\mathcal{L}_{CE}$ alone will decrease the difference between the neural network's predictions and observations - minimising $\mathcal{L}_{CE}$ will train network performance:

$$\mathcal{L}_{CE} = -\sum_i y_i \log\, \hat{y}_i \tag{9}$$

To help the optimisation process converge on a solution which generalises well to unseen data, a regularisation term is commonly added to this above task-based loss function. This means that the new loss function becomes a combination of *both* the above task loss and some arbitrary regularisation loss. Regularisation is very commonly applied to neural networks in general (Hardt & Recht, 2022) and is also a common ingredient in neuroscientific network models (Ali et al., 2021; Stroud et al., 2021). One common example is the L1 regularisation, which incentivises the network to maximise task performance while minimising the sum of all absolute weights in the neural network. In context of the recurrent weight matrix (W), with dimensions m x m, where m is the number of units in the recurrent layer (as stated above, we utilise 100 units in our simulations), the loss function becomes:

$$\mathcal{L} = \mathcal{L}_{CE} + \lambda\left|\left|W\right|\right| \tag{10}$$

$$\left|\left|W\right|\right| = \sum_i \sum_j w_{i,j}, \tag{11}$$

where $w_{i,j}$ is simply the recurrent weight matrix (previously denoted as W_h). λ determines how much the network is forced to converge on a sparse solution. This parameter is called regularisation strength. Throughout this chapter, we term networks regularised in this way as L1-networks.

4.2.2.3. SPATIALLY-EMBEDDED RNN (SERNN) SETUP

Unlike canonical RNNs, neurobiological networks are embedded in a physical 3D space. To simulate the basic effects caused by existing in this space, we manipulate the regularisation term. We include two constraints to the above objective function that we argue are true for any realistic neural network: Euclidean and communication constraints (see **Chapter 2** and **3**).

The Euclidean embedding makes sure that neurons are not mathematically abstract computational nodes. Put simply, they now exist within a physical space. This bridges to prior

analyses given in **Chapter 2** and **3**. To implement this, we embed units within a 3-dimensional space, such that each unit has a corresponding x, y and z coordinate. Using these coordinates, we generated a Euclidean distance matrix (a schematic example is given in **Figure 4.1c**). The first change we apply to the regulariser is to not simply minimise all weights, as noted above, but instead minimise all weights multiplied by their (simulated) Euclidean distance. This incentivises the network to minimise long distance connections. The element-wise matrix multiplication is denoted with the Hadamard product $\odot$. Adding this to our optimisation term gives us:

$$\mathcal{L} = \mathcal{L}_{CE} + \lambda\left||W \odot D|\right| \tag{12}$$

$$\left||W \odot D|\right| = \sum_i \sum_j w_{i,j} d_{i,j}, \tag{13}$$

where $d_{i,j}$ is the Euclidean distance between hidden units within the recurrent weight matrix. We positioned all units to be equally spaced in a 5x5x4 grid space. In principle, this geometric space can be defined in any way.

Beyond spatial embedding, we further constrained the network such that individual neuronal inputs can only come from other units that are topologically local. This idea was first mentioned in **Chapter 2**, but expanded in the discussion of **Chapter 3**. Put simply, in neurobiological networks, neurons can only communicate with each other as a function of how their propagated signals diffuse through the network to their neighbours (and their neighbours' neighbours, and so on). As a result, if a neuron is topologically distant to another neuron, it is unlikely that their weights will update with respect to each other. In the above equation, despite accounting for the spatial location, connections will update independently of this – meaning that topological communication has no influence on connection updates. In this work, we model these communication constraints as the weighted network communicability:

$$\mathcal{L} = \mathcal{L}_{CE} + \lambda\left||W \odot D \odot C|\right| \tag{14}$$

$$\left||W \odot D \odot C|\right| = \sum_i \sum_j w_{i,j} d_{i,j} c_{i,j} \tag{15}$$

where $c_{i,j}$ is the weighted network communicability between hidden units, within the recurrent weight matrix. The weighted communicability shown in **Figure 4.1d** and is defined as (Crofts & Higham, 2009):

$$c_{i,j} = (e^{S^{-\frac{1}{2}} W S^{-\frac{1}{2}}})_{i,j}, \tag{16}$$

where $S^{-1/2}$ is the diagonal matrix with elements $\frac{1}{\sqrt{s_i}}$ where $s_i$is the strength (sum of all node-wise weights) of node i. Note this equation is similar to **Equation 5** (given in the discussion of **Chapter 3**) which is the binary communicability on an adjacency matrix. But, due to there now being network weights present, this equation additionally controls for nodes with large weights and therefore liable to have an undue influence. **Figure D1-D5** provides a step-by-step explanation of the logic of this measure and **Figure D6** provides intuition as to how communicability itself shapes how weights update during optimisation.

Throughout this work, we term networks regularised in this way as spatially-embedded recurrent neural networks (seRNNs). In this chapter we trained our networks from the starting point of being fully connected with an orthogonal initialisation. This means that when network connections are updated between neuronal units, on the whole, they will tend towards weakening their connection strengths – analogous to pruning – due to the above optimisation process.

4.2.2.4. OPTIMISATION AND REGULARISATION SETTINGS

Throughout our analysis, we compared networks across network classes (e.g., L1 and seRNNs) but also in terms of the regularisation strengths (**Figure 4.1e**). To minimise the loss function

(i.e., for optimisation) we used an adaptive moment (Adam) optimiser with the following parameters: α= 0.001, β_1=0.900, β_2=0.999, epsilo$n$=10^{-7}. We trained networks for 10 epochs. Each epoch consists of 5120 problems, batched in blocks of 128 problems. Later, we also train networks trained on the randomised version of the task to explore how network behaviour and topology changed. These were trained on epochs consisting of 3840 problems to avoid them becoming too sparse too quickly which would prohibit us from analysing the trajectory of metrics.

One of the most critical parameter choices in our analyses is the regularisation strength. While the L1 regularisation and the purely Euclidean regularisation could be matched by average strength of regularisation of the hidden layer, the communicability term of seRNNs makes matching challenging due to it being dependent on the current state of the hidden layer (and hence it dynamically changes throughout training). In absence of a clear analytical method to match regularisation strengths across these classes, we took a pragmatic approach. As performance in the task starts to break down as networks become too sparse to remember past stimuli, we matched regularisation strength using task performance before looking at any of the other structural or functional metrics. In this light, we consider only those networks which converge and between classes these networks have equivalent accuracy on the task (**Figure 4.1f**). Specifically, we set the regularisation spectrum on a linear scale and choose the boundary values so that task performance starts to significantly deteriorate half-way through the set of networks. This was approximately around the 500th network across the set of 1000 networks we tested for each class.

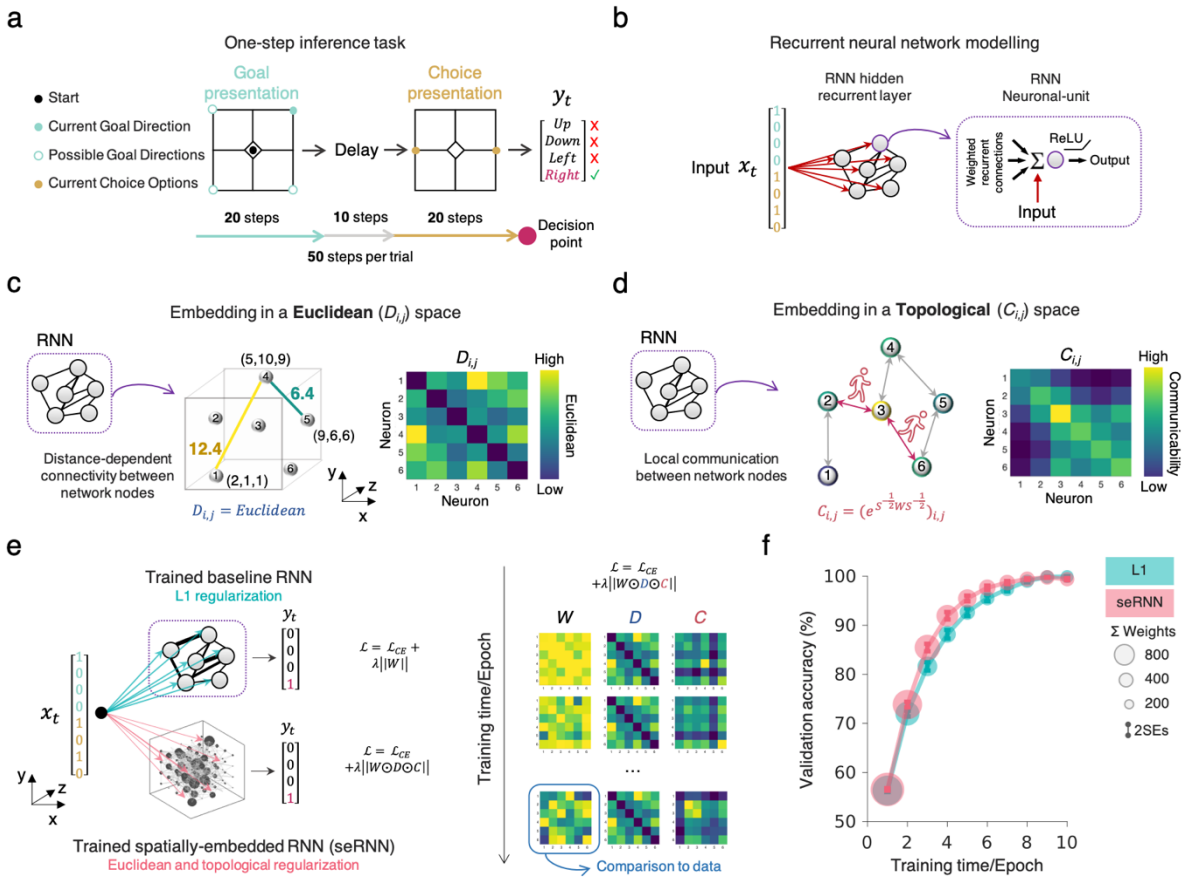


Figure 4.1. Task structure and spatially-embedded recurrent neural networks. (a) The one-step inference task consists of an initial period of 20 steps in which a goal is presented to one of four locations on the grid space: top left, top right, bottom left, or bottom right. This is depicted in light blue in the top right corner of the grid. There is subsequently a 10-step delay period in which this goal location must be kept in memory. Then two choice options are provided for 20 steps. Using the goal location information from before, agents must choose the choice option which is closer to the goal. In this example, as a left and right choice option are provided, the correct decision at the end of this 20-step choice presentation period is to select right. (b) A recurrent neural network (RNN) was trained to complete this task. Each unit of the 100-unit RNN was provided with a vector of input information at the relevant step points. Each neural unit consisted of a ReLU activation function, which also simultaneously took all weighted recurrent connections and provided a non-linear output. (c) We provided Euclidean space constraints on the RNN, by placing the 100-unit RNN into an evenly spaced 5x5x4 grid. We show a schematic demonstration of a small six-node network and show analogous Euclidean distance matrix. (d) We also provided Topological space constraints to the RNN, which means that communication was constrained according to local interactions between neurons given its current topology. This communicability measure is equivalent to an infinite random walk across the network, and the weighted communicability matrix is shown for the same simple six-node network. (e) For the main part of the study, we trained standard L1-

regularised RNNs as a baseline comparison to the spatially-embedded RNNs (seRNNs) which entails both the Euclidean and topological constraints. The training was undertaken such that Euclidean distances and inverse topological communication were minimised. To the right we provide a demonstration on a simple network how the W, D and C matrix all change over training epochs. The distance matrix is static due to fixed locations of units but the communicability matrix changes as the structure of the network (represented by the weight matrix) changes. (f) The validation accuracy of all converging neural networks is shown across L1 RNNs (blue) and seRNNs (pink), showing the equivocal performance is achieved on the shown one-step inference task.

4.2.3. Topological analysis

As the networks learn to perform the task, we also calculated a range of graph theory network statistics are computed. Each statistic was calculated using the Brain Connectivity Toolbox (Rubinov & Sporns, 2010). All mathematical formalisms are provided in **Appendix A**. All network statistics were calculated on the hidden RNN weight matrix (W_h) and, due to there being negative edges, all edges were enforced to be the absolute value of the element to compute the relevant weighted topological measures in the analysis.

The topological measures of modularity (Q) and small-worldness (σ) are the initial focus of our graph theoretical analysis. For modularity (Q), a default resolution parameter γ=1 was used in the algorithm which aims to form a subdivision of the network which maximises the number of within-group edges and minimises between group edges. For small-worldness, a random null model is used in its calculation, and to so we computed n=1000 null random networks. Some measures require binary networks. Therefore, when the measure in question was binary (e.g., small-worldness) a proportional threshold was applied, taking the top 10% of these absolute connections. The decision to use non-signed absolute thresholded weights when considering topology was made because this is the simplest starting point to analyse the networks. But importantly this decision remains relatively arbitrary and may introduce limitations because it fails to highlight any weight-topology interactions that may be functionally relevant. Future work should explore more advanced topological analyses of the seRNNs that incorporate weights, direction and sign.

4.2.4. Generative network modelling

To explore consistency between empirical data and our trained RNN architectures, we employed the same generative network modelling framework as previously outlined in **Chapter 1, 2** and **3.** When simulating networks, used the static Euclidean $(d_{i,j})$ matrix nodes i and j from the above described grid space setup of the RNN. We ran simulations across a defined a parameter space of 1000 evenly spaced combinations ($-3 \leq \eta \leq 3$ and $-3 \leq \gamma \leq 3$), for each of the 13 previously shown generative wiring rules. This produced a total of 13,000 simulations that were compared to each of the RNNs. Of the 1000 RNNs for L1 and seRNNs (sorted by marginally increasing regularisation, as described above), we constrained our analysis between the 100th network (which had been relatively weakly-regularised) to the 600th network (which had been relatively strongly-regularised) that were functional at epoch 9 (defined as having ≥90% validation accuracy), which left 455 and 379 networks for L1 and seRNN networks respectively. This left 834 RNNs for each of the sets, each with 13,000 simulations, leading to a total of 10,842,000 evaluated simulations for the main generative modelling analysis. It is important to note that the generative modelling process is based on binary rather than weighted connections. To therefore binarise our RNN networks to assess models, we enforced a 10% proportional threshold on the absolute weights of the RNN hidden layer. The parameter fitting for each network was done using the same Voronoi tessellation procedure as described in **Chapter 3**, with the same settings (Betzel et al., 2016).

To evaluate the fit of the synthetic networks simulated via our generative modelling, relative to our observed RNNs, we used both the energy equation and topological fingerprint dissimilarity as introduced in **Chapter 1** and **Chapter 3** respectively. When presenting results, we show the model fits for the top n=1 performing parameter combination (i.e., top 0.1% solution of 1000). Previous work in **Chapter 3** showed that over the development of spontaneously firing neuronal networks, homophily does increasingly well in approximating topology relative to other wiring rules. We tested if similar findings occurred in seRNNs and L1-regularised networks by comparing the energy achieved at epoch 3, 5, 7 and 9 in homophily models relative to all other models combined.

4.2.5. Decoding

To analyse the internal function of our trained recurrent neural networks, we record the hidden state activity of every unit while the network solves a set of 640 trials. Each trial is constituted of 50 steps (as shown previously in **Figure 4.1a**). For decoding, the activity is averaged in step-windows of 5, so that there is a total of 10 time-windows. In animal electrophysiology researchers often look at the explained variance per task variable per unit. To allow for comparison of our networks to findings in the literature, we wanted to extract the same metric. Given the nature of our task (which contains the Goal, Choice Options and Correct Choice), variables in our recordings are highly correlated, so that the standard decoding with ANOVA would give biased results. Instead, we constructed a decoding algorithm based on L1 regression, as follows:

1. Apply cross-validated L1 regression with 5 K-Fold to determine the alpha term with the best cross validation performance
2. Split the dataset via repeated K Fold (3 folds, 2 repeats)
3. On each (train, test) dataset:
    1. Train L1 regression with the pre-set alpha term.
    2. Calculate explained variance in test dataset.
    3. Iteratively set all values of a given set of predictors (e.g., all goal predictors) to 0 and recalculate the explained variance and calculate the drop of explained variance per predictor group.
    4. Take the average of drop of explained variance for each group across splits of dataset.

This algorithm results in each unit in every network being assigned an explained variance value for each task variable. This then allows us to… [spell It out!]. Note that the decoding cannot reliably work on networks which make too many errors, so that we only functionally analyse networks with a task performance of $\geq$90%.

4.2.6. Spatial permutation testing of task-relevant information

To examine the spatial clustering of decoded task information of within groups of hidden neural units in the RNNs, we constructed a simple spatial permutation test as follows:

1. Considering a single RNN hidden layer at a particular task time window (note, explained variances change over the course of the task), for each neuron, compute the relative preference for goal versus choice explained variance for each unit. This is calculated as the goal explained variance minus the choice explained variance.
2. Between all n "goal" neurons (i.e., positive difference from step 1), compute the Euclidean distance weighted by the decoding for goal information. This therefore captures the spatial proximity between goal neurons weighted by the magnitude of their "goal" information. Average this matrix to compute a summary statistic. This is the observed statistic.
3. Then repeat this procedure for 1000 times, but for a random set of n neurons taken from the 3D grid space. These 1000 summary statistics constitutes the null distribution.
4. Compute a permuted p-value (p_{perm}) which is simply the location in which the observed statistic (step 2) sits within the null distribution (step 3) normalised to the range [0 1]. This value subsequently corresponds to how clustered or distributed the observed goal decoding information is clustered in space relative to random chance. A small p_{perm} means that information is clustered more than chance and vice versa.
5. Redo step 1-4, but between all "choices" neurons (i.e., negative difference from step 1)
6. Redo step 1-5 for all desired time windows that have been decoded. In the current work, we calculated p_{perm} values for time window 3, time window 6 and time window 9 to reflect different aspects of the task over the sequence of the task.

As in most other areas of our analysis, the above steps were done for all functional RNNs (≥90% validation accuracy) for L1-reguarlised networks and seRNNs. We presented

distributions of these p_{perm} values for goals and choices to highlight how goal and choices information is clustered, distributed or random at key points in the sequence of the task.

4.3. Results

4.3.1. Modular small-world recurrent networks emerge from Euclidean and topological constraints

We first investigated the structural topology of the trained networks. In **Figure 4.2a** we show a representative example of an seRNN network depicted in the 3D Euclidean space it was trained in. In accordance with the training regime, both L1 and seRNN networks implement sparse coding due to the objective function minimising weights and that sparsity increases over the course of the training time (**Figure 4.2b**, left). However, seRNN networks - due to being embedded in a physical costly space - penalise their weights such that connections which are further away tend to become weaker. This is commonly found in empirical brain networks across species and scales (e.g., Betzel & Bassett, 2018) (**Figure 4.2b**, right).

A key goal of neuroscience is to answer how the rich functionality of the brain arises from its structure. One way to address this has been to study the topology of brain networks using systems biology and graph theoretical analyses. As we have seen from previous **Chapters 2** and **3**, empirical brain networks across species and scales have been shown known to exhibit varying degrees of modularity (Bullmore & Sporns, 2009) and small-worldness (Sporns & Zwi, 2004). These are both properties which have been proposed to be important topological characteristics facilitating efficient functioning. On one hand, modularity denotes dense intrinsic connectivity within a module but sparse, weak extrinsic connections between modules. On the other, small-worldness indicates a short average path length between all node pairs, with high local clustering, whereby networks reside somewhere between random and ordered networks. It is of note that modular systems are small-world, but not all small-world networks are modular – brain networks typically exhibit both (Meunier et al., 2010). We aimed to

determine to what extent modularity and small-worldness are properties of L1 and seRNN model systems.

By computing modularity Q statistics and small-worldness, seRNNs consistently exhibit both increased modularity (**Figure 4.2c**) and small-worldness relative to L1 networks (**Figure 4.2d**) over the course of training. Differences are smaller initially, but by later in training, the differences in modularity are very large in terms of effect sizes (at epoch 9, modularity p=2.24x10^{-82}, Cohen's d=1.07; **Figure 4.2c**, right) and are moderate-to-large in terms of small-worldness (p=2.82x10^{-19}, Cohen's d=0.59; **Figure 4.2d**, right). It is of note that, while there is variability across empirical brain networks depending on species, modality and scale, seRNNs achieve modularity Q statistics within ranges commonly found in empirical human networks (Hilger et al., 2017). Both L1 and seRNNs achieve the technical definition of small-worldness of >1 (Watts & Strogatz, 1998), but seRNNs exhibit a higher value more consistent with empirical networks (for review, see Bassett & Sporns, 2017).

In total, we find that by incorporating spatial and topological constraints within a neural network, we produce structural topologies that are much closer to empirical brain network configurations than otherwise possible within more conventional neural network model systems.

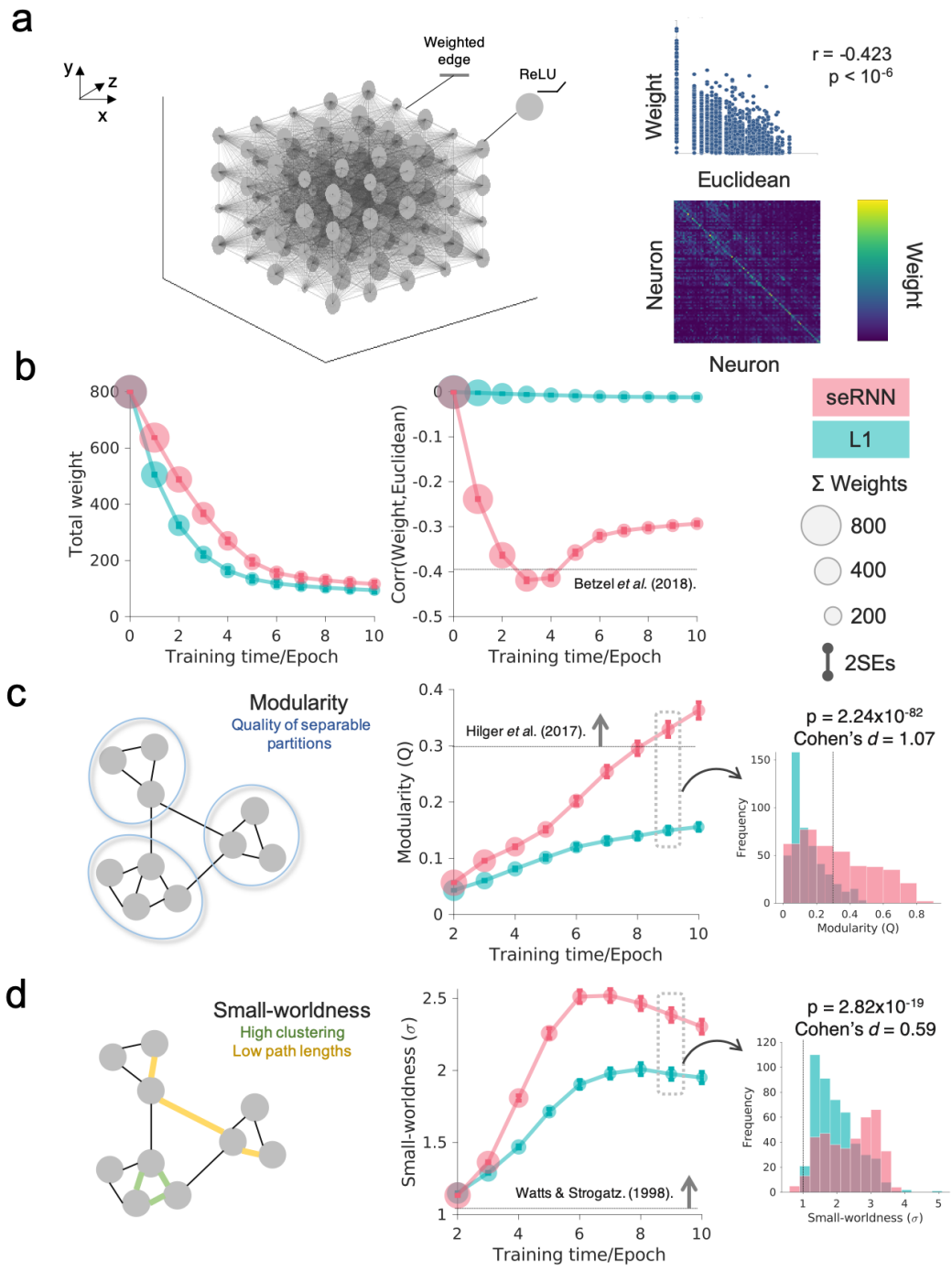


Figure 4.2. Spatially embedded recurrent neural networks (seRNNs) exhibit a brain-like structural topology. (a) Left, an example of a representative seRNN network in the 3D space in which it was trained. This network was taken from epoch 9 at a regularisation of 0.08 and is the network used for visualisations for the rest of this paper. Right, we show its connectivity matrix and weight correlations with Euclidean distance being negative. (b) Left, weights decline for both L1 and seRNN networks over the course of epochs/training. Right, seRNNs have a negative correlation between weight and Euclidean distance over the course of

epochs/training, but in L1 networks there is no relationship between weights and Euclidean distance. These results are across all networks. (c) Left, a schematic illustration of the concept of modularity in networks. While both L1 and seRNN networks exhibit increasing modularity over epochs/training, there is a consistently greater modularity in seRNNs compared to L1 networks. Right, we show very large (Cohen's d=1.07) statistical differences in modularity distributions for functioning (validation accuracy ≥90%) epoch 9 networks in L1 and seRNN networks. (d) Left, a schematic illustration of the concept of small-worldness in networks. While both L1 and seRNN networks exhibit a similar trajectory shape of small-worldness over epochs/training, there is a consistently greater small-worldness in seRNNs compared to L1 networks. Right, we show moderate-to-large (Cohen's d=0.59) statistical differences in small-worldness distributions for functioning (validation accuracy ≥90%) epoch 9 networks in L1 and seRNN networks.

4.3.2. seRNNs can be approximated via homophily generative models

As shown in prior chapters, the topology of empirical structural and functional brain networks, across scales, can be simulated by the topological homophily generative network model (Betzel et al., 2016; Vértes et al., 2012). Using our seRNNs as a benchmark, we next tested the possibility that homophilic mechanisms are themselves simply a consequence of local constrained interactions (e.g., spatial and communication) between neurons within a functioning network. To do this, we examined if seRNNs replicate these empirical benchmarks by simulating their topology in an unsupervised way using generative network models under varying well-established wiring rules (Betzel et al., 2016) analogous to **Chapter 2** and **3** analyses.

We first fit generative network models to functioning seRNNs (≥90% validation accuracy) according to an established energy equation (**Figure 4.3a**, top-right) using a Voronoi tessellation parameter fitting procedure. seRNNs are uniquely best simulated by homophily generative models (One-way ANOVA $F(12,2730) = 814.3$, $p<10^{-90}$; all Tukey HSD comparisons to other rules, $p<3\text{x}10^{-7}$) (**Figure 4.3b**). While able to simulate seRNNs well, homophily models do not uniquely generate the lowest model fits in L1 regularised networks with degree-minimum achieving the lowest energy, followed by degree-product (one-way ANOVA $F(12,3783) = 546.4$, $p<10^{-90}$; both Tukey HSD comparisons with the matching model, $p<3\text{x}10^{-}$

[4]; **Figure D7a).** This is simply because the L1 networks have a more random topology, which all generative rules can approximate (by tuning their wiring parameters closer to zero). When comparing seRNNs and L1 networks directly, seRNNs are approximated by homophily generative models better with a large effect size (p=1.02x10^{-40}, Cohen's d=0.80). In **Figure 4.3c** we show comparison data to macroscopic findings in adolescent structural connectomes (left) and high-density multielectrode arrays of rodent cortical neuronal networks at micrometre scales (right) that were outlined in **Chapter 2** and **3** – both of which are systems under geometric and topological constraints.

In **Chapter 3**, we developed way of computing statistical fits has been to compute the dissimilarity in topological fingerprint (*TF*) (**Figure 3.4a**, bottom-right) which measures the local arrangement of statistics generated via different generative models. To recap, rather than considering *KS* statistics in isolation for each considered statistic, as in the energy equation, the $TF_{dissimilarity}$ computes a measure of how well simulations can generate the *local organisation* of networks statistics. As in observed neuroscience data (Carozza et al., 2022) we also find that the $TF_{dissimilaity}$ generated by homophily and spatial models perform best (one-way ANOVA $F(12,2616) = 12.8$, p=7.63x10^{-43}; all post-hoc Tukey HSD comparisons to other rules p<3.7x10^{-3}) (**Figure 4.3d**). We give visualisations of simulated versus seRNN *TF* matrices (**Figure 4.3e**). Importantly, analogous findings are found in L1 regularised networks (**Figure D7b**) suggesting that the ability for homophily to approximate local arrangements of network statistics may be true of any functioning neural network.

Here we have shown across a range of regularisations that homophily models perform well in generating the topological structure of the seRNNs. But what is the effect of training time and regularisation on generative model performance? In longitudinal recordings of neuronal networks over the course of weeks (given in **Chapter 3**, **Figure 3.2**, 7-14 days *in vitro*), homophilic models do increasingly well at recapitulating topology. Analogously, we find that the energy model fits attained by homophilic models improves relative to all other models with more training, from epoch 3 (where homophily does worse than the mean of all other models)

through to epoch 9 (where homophily performs >2x better than the mean of all other models). This is greater than L1 over time (**Figure 4.3f**) which, as shown above, cannot be uniquely attributed by homophilic rules.

What is the effect of regularisation on model fits? As regularisation increasingly constraints topology, we may expect that homophilic models would better approximate networks the more they are regularised. By correlating the energy achieved by different generative models with the regularisation constraints imposed, we find that there is a unique negative correlation for homophilic rules meaning that as you increasingly constrain seRNNs they can become increasingly approximated by homophily and less by the other generative models (**Figure 4.3g,h**).

Taken together, these results suggest that spatially-embedded regularisation is congruent with the findings that empirical brain networks, which adhere to local spatial and topological constraints. Together, these can be approximated well by homophilic generative network models.

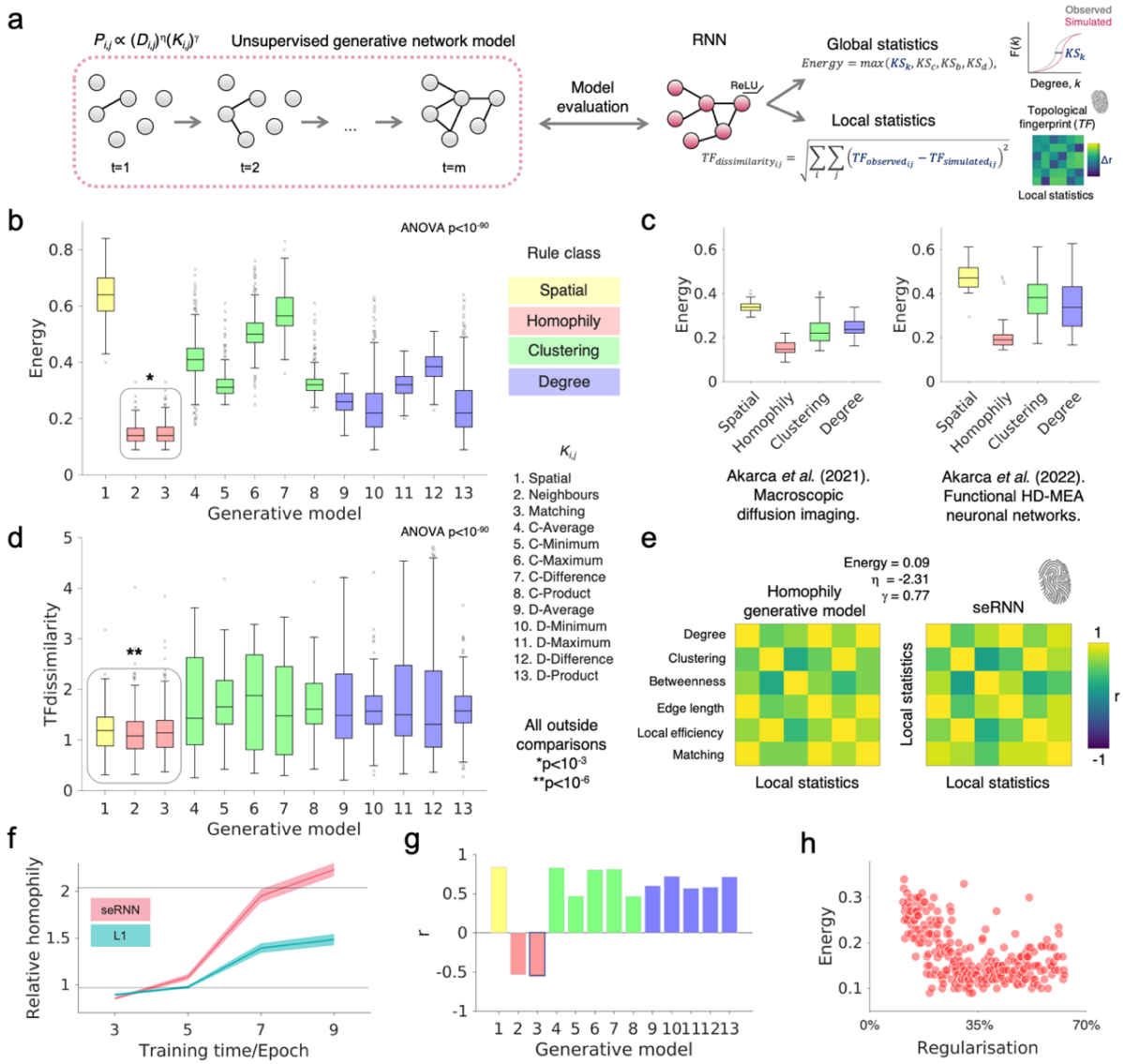


Figure 4.3. Homophilic generative network models can increasingly recapitulate global and local recurrent network topology over training. (a) A schematic of the generative network modelling procedure in which m edges are added to an empty network. The final simulated networks were compared to the seRNNs and assessed via the energy (assessing global statistical distributions) and the topological fingerprint dissimilarity (assessing local correlational structure) to see how well they can recapitulate topology. (b) The energy of seRNN networks show that homophily models achieve the best model fits. The top n=1 solutions are shown across the 379 functioning seRNNs in the evaluated range. (c) These findings are congruent with published data from adolescent whole-brain diffusion connectomes (left) (Akarca et al., 2021) and high-density functional neuronal networks at single-cell resolution (right) (Akarca et al., 2022). (d) Homophily models also can recapitulate the local topological structure of seRNN networks. The top n=1 solutions are shown across the 379 functioning seRNNs in the evaluated range. (e) An example of a best fitting homophily model to a single seRNN network and its corresponding topological fingerprint, showing visual similarity. (f) Homophily models

do increasingly well over epochs/training time over the duration of training. (g) Regularisation uniquely anti-correlates with energy (therefore correlating with model fit). Generative models relate to the key shown to the right of panel d. (h) A scatter plot showing the relationship between regularisation and matching energy.

4.3.3. Functional clustering in space and adaptive coding

So far, we have explored to what extent seRNNs can better approximate empirical structural network findings relative to L1-regularised networks. However, this says nothing about the functional specialisation that neural units have with respect to the overall task they perform. Accordingly, we next examined the spatial patterning of neurons, to test their respective sensitivities to task-relevant stimuli in 3D space (**Figure 4.4a**). Typically, when neuronal units share tuning profiles to a particular stimulus, they tend to cluster in space. This occurs across spatial and functional scales, such as fine-grained cortical surfaces with preferences for detecting colour from orientation (Waskom & Wagner, 2017) (**Figure 4.4b**) to whole-brain functional connectivity which form regular modular network patterns (Ji et al., 2019). However, as typical neural network architectures are not constrained by Euclidean or topological spaces such functional clustering phenomena cannot be explained.

To explore if seRNNs could recapitulate this, we decoded individual unit activity with respect to the goal and choice locations in the task-space, over the course of each trial. This resulted in unit-specific explained variance for goal and choices information throughout the duration of the task. In **Figure 4.4c** we show a visualisation in a representative network and unit-specific preferences over the course of a single trial.

By taking the relative preference for goal versus choice (calculated simply as the goal explained variance minus the choice explained variance) for each unit, we tested to what extent that sensitivity to task-relevant information was concentrated in parts of the Euclidean space that we had embedded the network in. Following 1000 permutations of this random sampling to form a null distribution, we computed a permuted p value (p_{perm}), comparing our observations with the null distribution. As a result, a small p_{perm} value corresponds to the observation that

functionally similar neurons tend to also be significantly clustered in 3D space and a large p_{perm} corresponds to functionally similar neurons being distributed in space (**Figure 4.4d**, top).

We considered three time-windows in the task: (i) an early stage (goal presented, decoded from steps 15-20); (ii) a middle stage (choice-options presented, decoded from steps 30-35) and (iii) a late stage (decision point, decoded from steps 45-50).

During the early stage, as there is only goal information given to the network, neurons preferentially code only for the goal information (seen by wide-spread dark green nodes in **Figure 4.4c**, left). In seRNNs, there is a slight positive skew in p_{perm} values, suggesting clustering of highly goal coding neurons (**Figure 4.4d**, left). Subsequently, in the middle stage, when choice options are then first presented, this goal information clusters within a concentrated area of space, leaving the choice information distributed (seen by clustering of green nodes and distribution of brown nodes in **Figure 4.4c**, middle). This is shown by a large positive skew in p_{perm} values for the goal in seRNN networks (**Figure 4.4d**, middle-top) and correspondingly the opposite for choice information (**Figure 4.4d**, middle-bottom). In the late stage, the clustering of goal information in space dissipates such that by the time a decision must be made the goal information has now spread out slightly more than before – although still retaining some clustering (**Figure 4.4d**, middle-right). At this point, the choice code has now become distributed (**Figure 4.4d**, bottom-right). Note that these findings are unique to seRNNs, shown by the distribution of seRNN p_{perm} values being skewed. L1 p_{perm} values remain uniform, indicative of no functional clustering (no statistical effect). In **Figure D8**, we show that these findings are true when task variables (i.e., goals and choices) are treated independently rather than relatively.

In any given task context, neurons adapt to preserve information relevant to the current behaviour (Duncan, 2021). We next examined the extent to which this *adaptive coding* – the notion that neurons have the capacity to code different types of information – occurs within our modelled networks. In **Figure 4.4e** we show the overall change in explained variances from preferentially coding for goal information (early in the task) to choices information (later

in the task), suggestive that both seRNNs and L1-regularisaed neural networks undertake this adaptive coding. This is also shown in **Figure 4.4f** but across all units of all networks (rather than aggregated across networks.

All together, these findings indicate that rather than having stable codes for task-relevant information that neural units are adaptive in how they code task-relevant stimuli. This is phenomena occurs regardless of the regularisation constraint. However, in seRNNs specifically, there is a specific spatial patterning of information such that early task information (goal) forms a spatially clustered code, but that later information (choices) forms a distributed code. This cannot be recapitulated in L1 RNNs, which are blind to physical constraints.

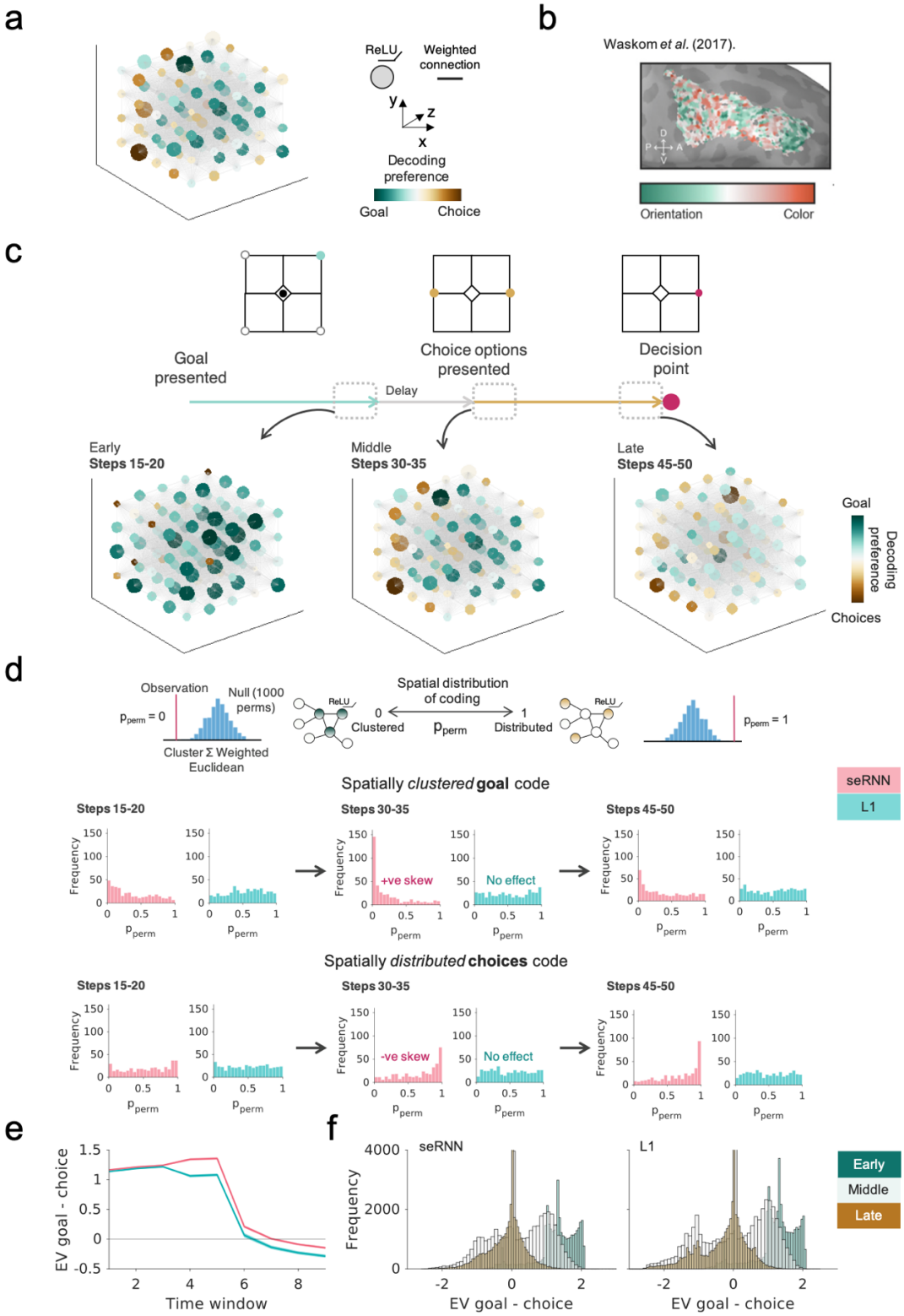


Figure 4.4. Functional clustering and distribution of coding in space. (a) An example of a representative seRNN network. The colour of the nodes relates to the decoding preference of that neuron, where a preference for goal information are represented by green and choices

information by brown. (b) Waldom, *et al.* (2017) highlights the spatial clustering of neuronal ensembles which are preferentially tuned for orientation (green) versus colour (orange) in dorsolateral prefrontal cortex. (c) We show decoding of neuronal units for goal (green) versus choice (brown) information at different points in the trial, within the representative seRNN network. (d) A schematic illustration of the spatial permutation test for determining functional clustering (top-left) or distribution (top-right). The p_{perm} values across RNNs are given for goal information (middle) and choice information (bottom) for seRNNs (pink) and L1 networks (blue). Goal information is shown to be clustered, as given by the positively skewed p_{perm} distributions. Choice information is shown to be distributed, as given by the negatively skewed p_{perm} distributions. (e) The mean and standard error (SE) of the goal–choices explained variance for seRNN versus L1 networks. Positive values correspond to preferential goal information while negative values correspond to preferential choices information. (f) Goal–choices explained variance histograms for early (time window 3), middle (time window 6) and late (time window 9) distributions across all neurons for all functional seRNN (left) and L1 (right) networks.

4.3.4. Mixed selectivity and energetically efficient coding

So far, we have shown that adding spatial constraints to a network as it learns to perform gives rise to patterns of network connectivity which are highly reminiscent of observed biological networks. Nodes become functionally collocated, and the spatial embedding causes differences in how they code task-relevant information. This selectivity profile has been widely studied across species and brain regions. Studies show that neurons tend to show a mixed selectivity profile, meaning that neurons do not only code for a single task variable but instead a mixture of them (Rigotti et al., 2013; Wallach et al., 2022; but see Hirokawa et al., 2019). A mixed selective code is assumed to support a network solving complex tasks by increasing the decodability of information from the network's neurons (Fusi et al., 2016; Johnston et al., 2020). There are many ways to quantify selectivity profiles (e.g., see Bernardi et al., 2020 for recent implementations). One very simple method is to look at the correlation of explained variances per task variable across the population of neurons. These are expected to be uncorrelated, implying a neutrally mixed code where a neuron's coding preference for one variable does not predict its code for another variable. In single unit recordings, correlations

are equal to zero or sometimes slightly positive (Erez et al., 2022). In our analysis we next tested whether seRNNs show a similar uncorrelated code.

In the first instance we looked at the selectivity profile of trained networks (epoch 9) for the goal and choices variables At the timepoint in the trial when the network must make a choice, the median selectivity profile falls on r=-0.057 for seRNN but r=-0.303 for L1, showing that L1 networks tend to produce an anti-correlated code while seRNNs have a nearly purely mixed selective code (**Figure 4.5a**). It is possible that this effect is driven by the differential connectome structure of the two groups of networks. While a purely modular and separated network would not automatically also mix codes across variables evenly, we find a well-mixed code in seRNNs. The additional highly communicative connections between modules of the small-worldness characteristic might help seRNNs to organise units in space while retaining a mixed code across the population.

Next, we explored the selectivity profile across epochs, regularisation strengths and time points within the trials. Looking at the selectivity profile across the regularisation spectrum, corresponding to the strength of the spatial embedding, in epoch 9 (**Figure 4.5b**, lower row, green), stronger regularisation led to an anti-correlated code in both L1 and seRNN, but seRNNs' negative slope levelled off more quickly than L1, protecting them from becoming too anti-correlated. With regards to training time, codes became more anti-correlated with higher epoch number, as training duration is related to stronger regularisation (**Figure 4.5b**, upper row with lower row). We also saw the selectivity profile changing within a trial, starting with an anti-correlated code which is becoming more mixed throughout the trial as the point of choice approaches (**Figure 4.5b**, compare blue lines with green lines). This likely is a good strategy for the network, as an initially segregated goal code makes sure that goal information does not vanish too quickly. As the trial goes on, the goal code gradually mixes with the choices code as to arrive at a highly informative mixed selective code when the network needs to infer the correct choice. Like our structural results, we saw that there is significant variation across the population of networks for both L1 and seRNNs, where some networks fall neatly

on r=0 and others might show strongly correlated codes. The following section provides an analysis of which specific networks show mixed selective codes.

The choice of selectivity profile in a population of neurons is strongly linked to the question of energy demand. As the firing of action potentials uses a significant portion of energy (Attwell & Laughlin, 2001), a population of neurons should choose a code with a good trade-off of metabolic cost and information capacity (Johnston et al., 2020). We wanted to investigate whether the structural and functional differences identified so far have an impact on the amount of energy spent on signals sent through the network. To test this, we calculated the mean activation of each unit in a network's recurrent layer (epoch 9 networks) during the period of information integration (after the onset of choices). Then we tested for the difference between seRNNs and L1 networks, controlling for the effect of the average weight strength in the recurrent layer (**Figure 4.5c**). seRNNs spend significantly less energy on unit activations compared to L1 networks (p<0.001). Sustaining a mixed-selective code at the time of choice might help the downstream integration units to decode information more easily so that less unit activations are needed to effectively communicate the correct choice. Note that the effect disappears for the networks with higher average weights as these are also the networks with relatively weak regularisation and hence weaker spatial embedding.

The seRNN, with its brain-like constraints, seems to automatically converge on a mixed selective code at the point of choice (Fusi et al., 2016). The seRNN is not only closer to the brain's code than we observe in L1 networks, but also achieves its functionality spending less energy on unit activations. The result is that seRNNs solve the task with a code that maximises decoding performance while minimising energy consumption. Put simply, it is not just that neurons functionally co-localise, their individual tuning properties are also shaped to minimise costs.

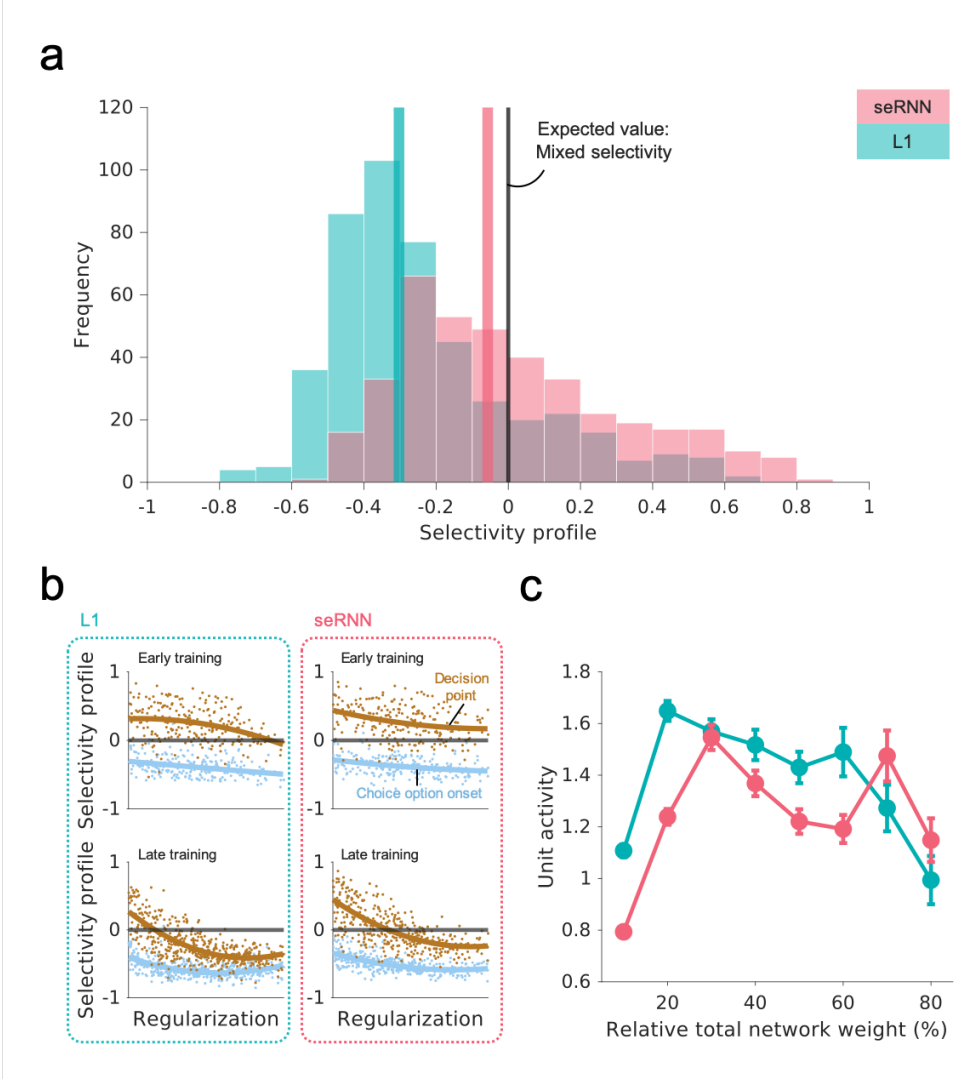


Figure 4.5. Mixed selectivity and energy use. (a) The histogram of selectivity profiles (correlation between explained variance for goal and explained variance for choices) shows how the distribution of seRNNs is more centred around the expected value r=0 than the L1 networks. Coloured lines mark the median of the distribution. The expected value highlights are fully mixed selective code. Coloured lines highlight the median values of the distributions. (b) Longer training and stronger regularisation both cause a more anti-correlated code. At the point of choice, seRNNs naturally centre on a purely mixed selective code as they do not go far below r=0, unlike L1 networks. (c) seRNN networks spend less energy on unit activations than L1 network which are matched for mean weight strength in the recurrent layer.

4.3.5. Regularisation leads to convergent brain-like topology and function

So far, we have seen that seRNNs show a collection of features which are commonly observed in brains but have not been related to each other before. The caveat not addressed so far is that for any feature we observed in seRNNs we also see strong variation across the population of networks (e.g., **Figure 4.2** for modularity or **Figure 4.5** for mixed selectivity). This opens the possibility that these features *do not* actually arise in parallel in seRNNs but instead each feature could emerge in its unique subgroup of networks. This would be unlike biological brains which exist in a critical sweet-spot area (Beggs, 2008) where all the features described in this paper are observed. In this section we tested whether all these seRNN features co-appear in a similar subset of all trained networks, defined by a unique combination of training parameters.

To study the co-occurrence of brain features in seRNNs we looked at the distribution of feature magnitude across the space of training parameters (regularisation strength, number of training epochs passed). **Figure 4.6a** shows matrix plots for accuracy (left), total sum of weights (middle-left), modularity (middle-right) and small-worldness (right) across the entire spectrum of training epochs (x-axis) and regularisation strengths (y-axis). As before, there is substantial variation in the magnitude of features across the population of networks, but now we also see that this variation is structured. Brain-like topology emerges in a sweet-spot of low to medium strength regularisation and during the later training epochs (pink box). The schematic in **Figure 4.6b** highlights this space of sparse, highly accurate, modular small-world networks with an example network exhibiting all properties (**Figure 4.6b**, middle-right). Above this space (i.e., networks with less regularisation, highlighted in orange) networks can solve the task and exhibit small-worldness, but remain very dense and lack the modular organisation found in empirical brain networks. Below this space (i.e., networks with more regularisation, highlighted in light blue) networks exhibit extreme sparsity and modularity, but fail to functionally converge on the task and they lose their small-world topology. Example networks sampled from these spaces are highlighted (**Figure 4.6b**, top-right & bottom-right).

In a next step we wanted to look at the same "sweet spot" in terms of the network's functional properties. The decoding required us to focus this analysis on networks with high task performance. Plots now only show networks with an accuracy $> 90\%$. We focus our analysis

on epoch 9 networks. **Figure 4.6c** shows the functional results across regularisation strengths with the sweet-spot of regularisation from **Figure 4.6a** highlighted with the pink box. In the first two plots from the top we show two structural metrics (sparsity and short connection preference) to highlight that looking at features across training epochs still reveals the same spectrum of ideal regularisation strength. We observed the same distribution when looking at the homophily generative wiring rule across the regularisation spectrum (**Figure 4.3h**). Looking at mixed selectivity (**Figure 4.6c**, 3rd from top) our analyses revealed that networks show a fully mixed-selective code in the sweet-spot window identified before. Units here show a balanced code with information for both goal and choices (**Figure 4.6c**, bottom), whereas very dense or very sparse network show a preference for either goal or choices information. As such the density and related modular small-world structure influences the time horizon of information flowing through the network. **Figure D9** shows a detailed correlation matrix showing all pairwise relationships between features studied in the paper.

Our findings show that there is a critical parameter window in which both structural and functional brain features jointly emerge in seRNNs. Brains are often said to live in a unique but critical niche where all characteristics needed to support their function can exist in parallel (O'Byrne & Jerbi, 2022). seRNNs show the same preference for a critical parameter window but also give us the ability to study networks on their way to converging on brain-like characteristics in the critical parameter window or networks which lie outside the sweet-spot.

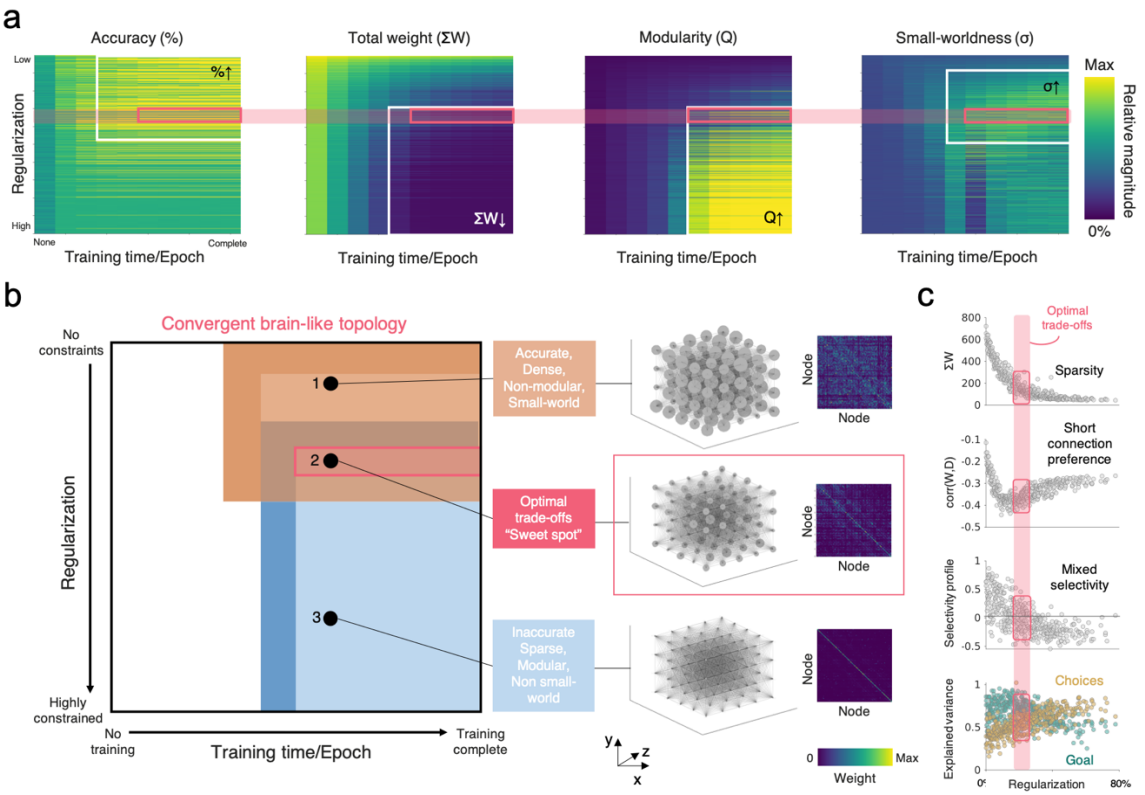


Figure 4.6. The spatially-embedded recurrent neural network (seRNN) parameter space converges on brain-like topology and function. (a) The white borders within the regularisation-training parameter space delineate the conditions where seRNNs achieve robust accuracy (left), sparse connectivity (middle, left), modular networks (middle, right) and small-worldness (right). The pink box shows where all these findings can be found simultaneously. The colour of the matrix corresponds to the relative magnitude of the statistic compared to the maximum. (b) This is further highlighted by a schematic representation which shows the space of possible seRNNs. The pink box shows the overlap of all findings, where accurate, sparse, modular, small-world networks are generated, which we term as being at the optimal trade-off. Network 1, 2 and 3 each represent example networks across the space. The nodes of the representative graph reflect the node's strength, defined as the total sum of the node's in- and out-connection weights. (c) In this pink window, networks are sparse (top), prefer short connections (middle-top), have a selectivity profile centring around zero, consistent with mixed selectivity (middle-bottom) and have equivalent explained variance for both the goal and the choice (bottom).

4.3.6. Replication across baseline architectures

Our analysis has focused on constraints we argue are ubiquitous to any physically instantiated neural system, modelled together and incorporated as a single optimisation within a neural network architecture. But it is possible that space and communicability have distinct roles in outcomes that our analysis has not hitherto uncovered. We next examined the effects of variations in the neural network architecture by constructing several alternative baselines. First, we considered the independent effects that space ($D_{i,j}$ only) and communication ($C_{i,j}$ only) may have on topology (**Figure 4.7a**). Broadly, we find that seRNNs show the largest magnitude in topological measures across training and L1 shows the least, in terms of total weight, weight-distance correlations, modularity and small-worldness – with the independently constrained variations occupying somewhere in between. Interestingly, negative correlations do emerge from communicability-only networks, indicating that local interactions within a physical space may also approximate some limited spatial constraints indirectly. Moreover, while increased modularity is attained by both spatial and communicability networks, small-worldness occurs in absence of spatial constraints.

Beyond incorporating spatial and communication constraints, seRNNs can be conceptualised as combining both functional and structural objectives within a singular unified optimisation process. As such, it is possible to ask: what is the contribution of functional goals versus physical constraints on network outcomes (Kaiser, 2017)? To assess how the task itself shapes topology, we tested two additional variations to the task in seRNN networks: (1) a harder version of the task in which there were greater choices available and (2) one in which the task structure was entirely randomised, detaching goal and choice information from each other. In the comparison of functional requirements, we find that broadly there are subtle changes to the topology depending on task requirements across all considered networks (**Figure 4.7b**). Harder task requirements (shown in blue) retain weights more than seRNNs, have subtly shorter connections, less modularity but retained small-worldness. The absence of a task

requirement (shown in grey) leads to massively sparse networks, subtly less preference for short connections and become less small-world in their topology. In total, this suggests that functional demands and structural constraints together influence topological outcomes.

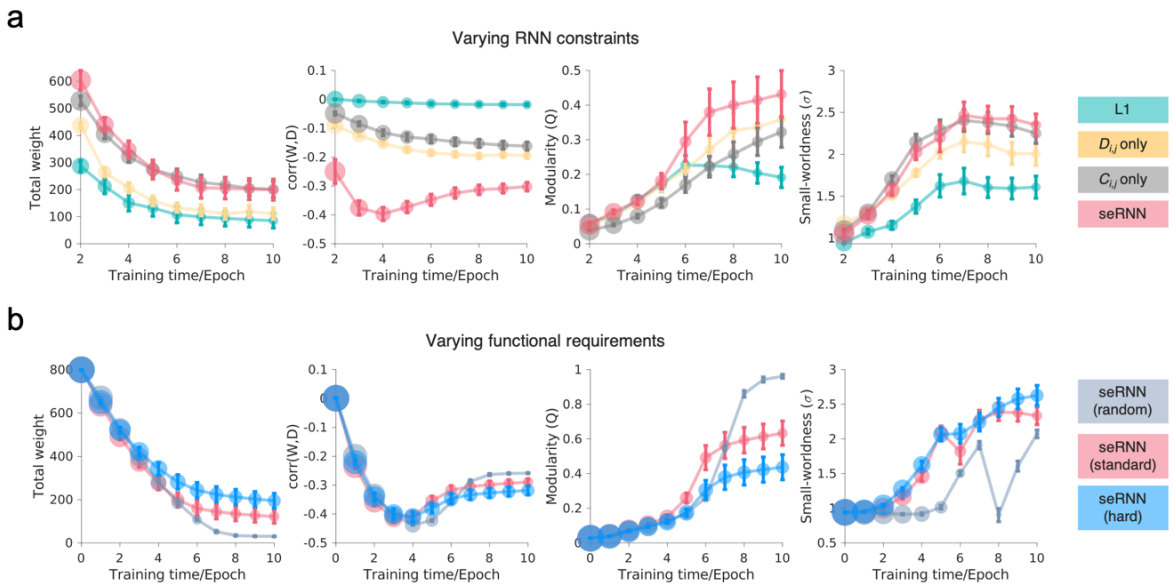


Figure 4.7. Independent replication of key results across baseline regularisations. (a) Recurrent neural networks (RNNs) are shown with varying regularisation constraints to compare their effects on topology. L1 networks are shown in green. $D_{i,j}$ networks were regularised by space but not communicability and are shown in yellow. $C_{i,j}$ networks were regularised by communicability but not space and are shown in grey. SeRNNs are the same networks that have been regularised by both and are shown in pink. These networks are all functioning and have been thresholded >90% validation accuracy. Lines correspond to 2SEs across the 100 networks. The size of the node corresponds to the mean weights. All networks are independent to any that have been shown in the main text, and each network class was trained incrementally over 100 regularisation strengths. (b) The statistical structure of the task was altered to observe the effect of various functional requirements on the topology of the resultant seRNNs. Standard seRNNs are shown in pink. seRNNs trained on a randomised task setup are shown in grey. SeRNNs trained on the harder task setup are shown in blue. Rather than thresholding by function, these are shown across all networks because seRNN (random) networks would otherwise be filtered out as they do not achieve sufficient accuracy as tasks are random. As above, lines correspond to 2SEs across the 100 networks. The size of the node corresponds to the mean weights. All networks are independent to any that have been shown

in the main text, and each network class was trained incrementally over 100 regularisation strengths.

4.4. Discussion

In this chapter we showed that recurrent neural network models, trained in context of local Euclidean and communication constraints, develop a range of canonical hallmarks of brain network topology and function. Our main contribution is to show that rather than requiring the clear dichotomisation of structure and function (or study them with respect to each other), it is possible to understand how concurrent constraints on both structure and function can simultaneously shape one another within the context of a neural network model. We do this by including both types of constraints within a single learning objective. We view spatially-embedded recurrent neural networks (seRNNs) as a class of networks that can be used to probe how various biological or physical factors can influence emergence and functionality within neural networks. This step-change is crucial to build sustained bridges between neuroscience and AI research communities, and to translate these kinds of insights to real life neurodevelopmental contexts.

Why focus on economic principles? Incorporating economics into mathematical models of biological systems is critical for the simple reason that evolution favours solutions which achieve functional goals whilst minimising metabolic costs. Indeed, this principle can be traced back to the dawn of modern neuroscience (Cajal et al., 1995). More recently, this principle has been incorporated across a range of neuroscience disciplines, from understanding cognitive control (Christie & Schrater, 2015) and mental labour (Kool & Botvinick, 2018), to computationally rational decision-making (Gershman et al., 2015). Indeed, as outlined throughout the previous chapters, the predominant explanation for anatomical and functional organisation of the brain has been given in terms of economics: minimising wiring costs while maximising adaptive topological features (Bullmore & Sporns, 2012). But what does it mean to be *adaptive*? Adaptive with respect to what goal or set of problems? In this final empirical chapter, we argue that only in the context of direct, ecologically-valid, functional requirements

can we better understand 'adaptive' more concretely. This chapter suggests a parallel interpretation to the economic argument: biological networks operate within highly locally-constrained settings to achieve global functional goals. Central to this locally-constrained setting is space and communication, which must be resolved alongside wider global functional objectives.

Our findings show that, within a critical window of regularisation, seRNNs recapitulate numerous empirical structural and functional neuroscience observations simultaneously. This suggests that providing artificial neural networks with a *topophysical* structure (Basset, at al. 2010; Sperry et al. 2017) can enhance our ability to directly link computational models of neural structure and function. Notably, recent work placing empirical connectomes within a reservoir framework also partly achieves this goal (Damicelli et al., 2021; Goulas et al., 2021; Suárez et al., 2021). In these works, the functioning network's recurrent structure is fixed – with only the weights within the read-out layer being trained. In contrast, our seRNNs approach allows us to be very specific in defining the networks' constraints and objectives, in both physical structure and function, while allow the network's connectivity to change over time. A related but different approach to ours would be to model the metabolic costs incurred by initiating action potential generation and synaptic transmission (e.g., Sengupta et al., 2010), modelled as minimising the absolute unit pre-activations, rather than post-activations. This subtle methodological change has been shown to lead networks to exhibit some observations consistent with predictive coding (Ali et al., 2021). The key difference between these networks and seRNNs is the level at which the constraints have been implemented; the first is at the neuronal activation-level while seRNNs are constrained in terms of geometry and the local connectivity between neurons. However, in principle it would be possible to implement these all simultaneously. Future work should combine these constraints and examine if they influence outcomes in distinct or shared ways.

Spatial constraints are an intuitively straightforward and principled way to bridge artificial networks with physical networks. A somewhat less intuitive constraint that we include in this work is network communicability. Unlike shortest path lengths, which demarcates a path

between two nodes that entails the shortest number of walks, the communicability instead demarcates the length of time spent on an edge if one were to take an infinitely long random-walk across the network (Estrada & Hatano, 2008). This was introduced in the discussion of **Chapter 3** and in detail in **Figure D1-D6**. However, communicability has a crucial characteristic that is highly relevant in the current context: unlike network properties like shortest paths, communicability requires only local knowledge to calculate. Information can "walk" across the topology of the network without any global knowledge of the network, as a whole (Baram et al., 2018). This is important because it is more akin to polysynaptic communication in which a neuron can only act in the context of its local environment. This is because neurons are decentralised components within the network and do not have global knowledge of the network. In recent work, the incorporation of communication models have shown improved coupling between structure and function (Seguin et al., 2020, 2022); a result consistent with the notion that, to understand neurobiological systems, one must understand its underlying communication dynamics (Srivastava et al., 2020). Models, such as ours, which incorporate local constraints to learning may be more veridical (for a critique on the biological plausibility of common neural network architectures, see Grossberg, 2020).

In the present work, we initialise the seRNN with a fully connected network within the hidden layer. This was motivated by a few practical and biological considerations. First, due to the seRNN architecture, the network requires some initial connections because otherwise it would be unable to propagate information to the output layer, leaving it unable to function. Therefore, at least some connections must exist at initialisation. Second, without sufficient connectivity, it is possible that the network would simply choose to tune all weights down to a point where it failed to converge, simply because this minimises communication. This highlights that a sufficient level of connectivity is required before the network is able to converge on a solution that is able to balance its simultaneous structural and functional objectives. Third, from a biological perspective, while in **Chapter 2** and **Chapter 3** we explicitly studied the *increasing* formation of white-matter structural and *in vitro* microscale functional connections respectively, not all neural connectivity has this trajectory. Indeed, in

human brain development, neural grey matter has an early period of intense proliferation before it undergoes a guided pruning process as the child learns (see S. L. Andersen, 2003 for a broad review on timelines, and Chechik et al., 1998 for a detailed study on what computational capacities this timelines affords the network). This timeline has interesting regional differences across the cortex that may reflect functional specialisation – for example, primary visual cortex peaks at approximately six months (Huttenlocher & de Courten, 1987) and prefrontal cortex at two years (Huttenlocher, 1979). In summary, while the initialisation of full connected networks seems at odds with the prior chapters they are likely modelling different underlying biological processes that are consistent.

There are many areas that we wish to improve upon with future research. Principally, our models did not include a significant amount of biological detail that, while inevitably critical for neuronal functioning, do not speak to the specific observations we aimed to recapitulate in the present study. Nevertheless, implementing such details including specific molecular mechanisms that guide neural circuit development (Moons & De Groef, 2021) or the heterogeneous spiking of neurons (Perez-Nieves et al., 2021) will likely provide a range of new insights. Indeed, the addition of such biological details will help us expand the applicability of our models to explore the effect of developmental time-courses (Baxter & Levy, 2020; Chechik et al., 1998), functional brain specialisation (M. H. Johnson, 2000) and how network variability may underpin individual differences (Siugzdaite et al., 2020).

Beyond biological detail, there are also numerous computational directions that are yet to be examined – for example, currently we do not allow units in the seRNNs to reposition themselves in space. It may be that the distribution of connection lengths is an important computational consideration. For example, in **Chapter 2** and **Chapter 3** empirical spaces were used within which simulations were conducted. Here, we have used an evenly spaced grid as a proof-of-concept that Euclidean distances can be modelled in the system. However, future work may look to more carefully approximate the spatial statistics of particular areas, regions or neurons in the brain at different scales. It may be that greater heterogeneity in the spatial

distribution of nodes may promote the specialisation of task-related information as it further promotes modularity in the guided pruning process. Moving to task-related considerations, to train the current seRNNs we have used a relatively simple task architecture mirroring basic one-step inference in a navigation and working memory context. Future work should expand the scope of functional requirements placed on the network to mirror more closely the types of problems that biological organisms face (e.g., Crosby et al., 2020). It is not clear to what extent different types of networks could be generated by seRNNs if they were forced to solve different tasks. For example, is it possible that if an seRNN solved a different task it could organise into a non-homophilic or even heterophilic network? One plausible way to test this would be to vary the statistical regularity or randomness of the task inputs to the network. For example, with no random perturbations to the task inputs, the seRNN will form a homophilic network as it specialises, where numerous neurons strengthen their connectivity between themselves to undertake shared computations (as has been shown in this work). But this is only possible when the task is somewhat regular and predictable. In contrast, with increasing randomness included within the task, the network would likely develop a more random and non-homophilic topology. For example, if the learnt rule were to change (mirroring counter-conditioning experiements, e.g., Keller et al., 2020). If the randomness were itself somehow predictable (e.g., with a stationary transition probability) it is possible theoretically that a hierarchical network could develop, mirroring the task structure: large modules would reflect the probabilistic transitions over longer time frames but within these modules there may still be non-homophilic connectivity patterns reflecting the shorter time frames. This would naturally make the task intrinsically more challenging, as shown in the extreme case when we randomised all inputs and the seRNNs could then not solve the task at all (**Figure 4.7**). A related fruitful direction to explore would be to examine how the spatial extent of how information is coded in seRNNs links to the temporal sequence of the task inputs. For example, in **Figure 4.4** we find that goal information is clustered but choice information is distributed. We interpret this as mirroring the temporal sequence of the learning process: goal information must be learned first, and thus is encoded in "cheaper" core regions of the seRNN. Then, when choice information is learned, it is done so while goal information must be remembered, and

thus distributes in space. This interpretation could be tested in a multi-step inference task in which multiple different stimuli of different forms are presented in different sequential orders to assess the relationship between stimuli-specific features versus simply the timing of their presentation. We see all these ideas as very exciting future developments.

Our key finding is that modularity, small-worldness, homophily generative mechanisms, spatial configurations of function and mixed selectivity emerge together, within a critical regime of regularised spatially-embedded neural networks. Critically, these findings are a consequence of solving a single optimisation problem that is concurrently structural and functional in nature. In turn, this implies that numerous attributes of brain networks (e.g., Meunier et al., 2010; Sporns, 2013b) may be emergent properties of physically embedded functioning networks (see Rubinov, 2022 for detail on the nature of explainations in systems neuroscience). In **Chapter 5**, we will expand on this concept of incrementally constraining computational models (such as the present one) to increasingly recapitulate biological observations.

In summary, our findings suggest that by incorporating simple spatially-embedding constraints inside a neural network model (termed seRNN), it is possible to recapitulate otherwise disparate structural and functional findings from network neuroscience. These locally-constrained interactions mirror the constraints that neurons face in their communication within biological settings.

5. CHAPTER 5

CONCLUSIONS AND FUTURE OUTLOOK

Content included in this chapter is also included in an article currently in review at Trends in Cognitive Sciences: Towards computational neuroconstructivism: A framework for developmental systems neuroscience. Duncan E. Astle, Mark H. Johnson & ***Danyal Akarca****.*

5.1. So, what have we learned?

The studies given in this thesis have provided a number of insights that I will first summarise briefly, before focusing on their implications and future research questions. **Chapter 2** shows that the simple topological wiring equation outlined in **Chapter 1**, when instantiated with a homophily-based generative rule, can generate networks with good statistical fits to empirical structural DTI connectomes in a large sample of neurodiverse children. When invariable connections across the sample are provided to the model at the beginning of the generative process, a very good level of spatial embedding occurs. The ability to compress the connectome (Betzel & Bassett, 2017b) into generative ingredients, as modelled in the wiring equation, may be very useful in helping bridging across levels. This is because they reduce the dimensionality of the network in a principled manor. Moreover, the wiring parameters correlate with a wide range of white- and grey-matter neuroimaging measures and cognitive performance metrics (corroborating some recent findings, e.g., Zhang et al., 2021). The variability captured by different aspects of a generative model will necessarily depend on their dynamism – components that are static (e.g., distances, which is relatively fixed) will likely have less influence on outcome variability than those that are dynamic (e.g., homophily, which updates). Put simply, when a measure is adjusted continually during the generative process there is greater room for it to influence network stochasticity. Local and regional projections of the wiring equation co-localise with widespread cellular processes and synaptic physiology for costs and homophily, respectively. While there is plenty to still be done, this work showcases a step from studies that test associations to more direct, explanatory (albeit simple), models. This moves us a step closer to the simulation of complex changes that occur over development (M. H. Johnson, 2011).

Numerous studies have speculated that homophily-like generative rules resonate with Hebbian-like mechanisms (Goulas, Betzel, et al., 2019; Vértes et al., 2012, 2014). In **Chapter 3**, we bridged these two concepts, by applying generative network modelling directly to HD-MEA recordings, that can longitudinally detect extracellular neuronal activity at the level of

individual neurons. Homophilic models yield the best model fits for functional connectivity at this microscale. Homophily simulations can also recapitulate of topology such that they are indistinguishable from the observed networks, in addition to the growth trajectories of these networks over the course of weeks, in terms of modularity and global efficiency. Beyond the energy, which measures the dissimilarity to the *distributions* of networks statistics, we introduced the topological fingerprint dissimilarity which evaluates dissimilarity to *local arrangements* of network statistics. Homophilic wiring parameters are affected via experimental perturbation, and these principles extend across numerous data-types including varied plating densities, human stems cells and human cerebral organoids. This work highlights that some principle of self-similarity (as operationalised by homophily) may extend across levels, and draws direct comparisons to underlying neurophysiology.

Topological principles may provide part of the explanation as to what shapes the statistical structure of observable structural or functional networks. However, there remains a significant gap in how these networks balance multiple competing optimisation problems, both in terms of physical constraints, but also solving task-related functional problems simultaneously. Numerous recent studies have embarked on linking structure and function within reservoir frameworks (Goulas et al., 2021; Kawai et al., 2019; Suárez et al., 2021). But such models require that the structural topology is pre-specified, rather than an emergent property of the underlying interactions. As a result, these have limited utility for understanding development, because development centrally entails that topologies change in response to prior influences. In **Chapter 4**, we tackle this problem by training recurrent neural networks to perform a task, whilst simultaneously specifying physical constraints of space and communication. Numerous important observations fall out of this process, including modularity, small-worldness, homophilic generative mechanisms, spatial configuration of function and mixed selectivity. This provides an exciting way to link possibly emergent properties that we observe in brains directly to wider computational optimisations.

This body of work was motivated by the idea, summarised in **Chapter 1**, that to understand the functional role of neural circuits one must understand the developmental history of those

circuits. Past processes will necessarily shape future outcomes – an idea encompassed by the famous Waddington landscape (Waddington, 1957) and which are central to neuroconstructivist principles (Mareschal et al., 2007). So, how well have the analytic approaches presented in the current works captured these developmental processes?

In some ways, they have done well. Generative models have highlighted that the similarity between brain regions (at multiple levels) is likely to play an important role in how the brain develops and organises over time (Betzel et al., 2016; Goulas, Betzel, et al., 2019; Oldham et al., 2022; Vértes et al., 2012; Zhang et al., 2021). This conclusion has both biological and computational interpretations that will likely contribute to new theoretical advances (which in the next section will be elaborated upon). These models also highlight that both variability between brain regions and stochasticity are both likely to be very important factors throughout development. The later introduced seRNN modelling framework highlights that it is possible constrain neural networks in terms of both their structure and function – the lack of which was a crucial limitation in early connectionist literature aiming to answer developmental questions (Thomas & Karmiloff-Smith, 2003). Overall, these analytic methods have done well in showing that to understand developmental trajectories, one must first understand the broad context and constraints in which the system is developing.

In other ways, the current modelling approaches deployed here leave large room for improvement. In particular, it remains unclear to what extent any of the presented models may offer a truly biological veridical account of development and not just purely an account of broad developmental phenomena. Improving these models to draw closer to biology will be challenging because of the difficultly intrinsic to reconciling systems-level computational explanations of development (as have been given often throughout this work) with the messy details of biological systems under change (H. Andersen, 2014). For example, such system-level explanations are prone to circularity or redundancy if not appropriately tested (Rubinov, 2022). A core challenge that must be worked on is balancing parsimony and complexity in these models. For example, parsimony in many of these models is seen as an advantage, but it can equally be seen as a disadvantage – the apparent elegance of the models may lead to

speculative conclusions that are difficult to fully justify in biological terms. This is particularly true when modelling interactions between highly complex developing substrates across different levels and scales. Another example is that seRNNs still retain numerous features that will, inevitably, not map to biology, such its core optimisation algorithm or its node-wise characteristics. Careful reasoning is required to establish how mathematical concepts may map up to biological observations, in context of core principles such as finality of resources, embodiment and evolutionary forces. The very gradual and careful incorporation of new constraints with increasing biological detail will be key in improving computational models to answer questions around development (this will be elaborated on in the next section). Another area for improvement is in how we model behaviour. While in **Chapter 5** we examine how seRNNs learn to perform a simple inference task, the field must better expand the range of problems that model networks are tasked to solve, at key time-points, to better approximate empirical structural and functional observations. Overall, our goal should be to both recapitulate empirical observations while providing plausible mechanistic explanations for how such phenomena are achieved in biological and physical terms.

5.2. What makes a *good* computational model of developmental processes?

In the next section, I take a step back to reflect upon the broader lessons learnt over the empirical chapters of this thesis. In particular, I will reflect upon the ingredients needed for good, and, n time better, computational models of development. As will become apparent, some of these ingredients should be the focus of future studies, and some are encapsulated in the models deployed already in the thesis. In either case, these hallmarks are intended to reflect any good computational framework for developmental systems neuroscience.

5.2.1. Simple, but *explainable*, models can recapitulate developmental findings

Computational models become increasingly compelling when phenomena emerge from the combination of some other principles specified within them, that – on the face of it – seem unrelated. In evolutionary biology, these phenomena are sometimes referred to as 'spandrels', whereby a phenotype is the by-product of another characteristic which was selected for, rather than the consequence of direct adaptive selection.

In the thesis we have shown two clear examples as to how relatively simple models can achieve this. The first example is in the generative network model. Two parameters influence how regional structures weigh up the cost-value trade-off when forming connections probabilistically over time. Resulting networks have complex organisational characteristics mirroring real biological networks, and thus one view is that we can consider the complex topological organisation of the brain in terms of these much simpler local economic negotiations, which occur throughout development. However, more compelling than this is their ability to recapitulate complex local phenomena not at all related to how those networks were simulated, despite the parsimony of the model. For example, not only do the statistics of networks mirror each other, but the local arrangement of these statistics also mirrors those seen in observed networks. We termed this the topological fingerprint (shown in **Chapter 3**). The second example is the seRNN. Here numerous structural and functional findings can be explained via underlying optimisation objectives. They are necessary consequences of learning to solve functional problems in a physical space. That is, the features are not necessarily hard-coded into their design, but occur due to some parallel but quite distinct objective.

These examples of emergent spandrels only make sense when we use parsimonious explanations that do not determine the outcome *a priori*. But how does this marry with another aim, namely that computational models should become increasingly constrained by biology, to narrow-in on biophysical mechanisms? Ultimately, there is an explanatory trade-off, and models necessarily must be able to explain various components of the outcome in terms of different individual or interactive model inputs. For example, towards the end of **Chapter 4**, we find that different constraints have subtly different influences on the outcome. If further constraints were to be incorporated, they too would have to be independently evaluated for

either their novel or redundant contribution to the modelled outcome. Put simply, if models produce elegant features because in some sense those features are already specified within the computational model, this is less compelling. This is because each specified aspect of the model must be somewhat explained on its own terms – reducing the *unique* explanatory power of the model itself (i.e., what the model provides beyond its inputs). To put it concretely, incorporating empirical constraints such as observable fibre lengths (Betzel et al., 2016), morphometric similarity (Seidlitz et al., 2018), cytoarchitecture (Goulas, Majka, et al., 2019), or genetic co-expression (Arnatkevičiūtė et al., 2021; Oldham et al., 2022) will likely provide important insights. But we simultaneously risk constraining the simulation to the point where the outcome becomes an inevitability. Parsimony, and the emergent nature of the complex features arising within the simulations, may itself be a very helpful feature for a developmental computational model. In practice this parsimony will need to be balanced against the complexity of the phenomena being explained (**Figure 5.1**). As a result, I would argue that a good developmental model should gradually and cautiously incorporate constraints to evaluate their shared and independent effects on outcome, with explicit knowledge as to what is and is not "baked-in" to the model. Later in the next section, I will elaborate on one example of a "baked-in" assumption.

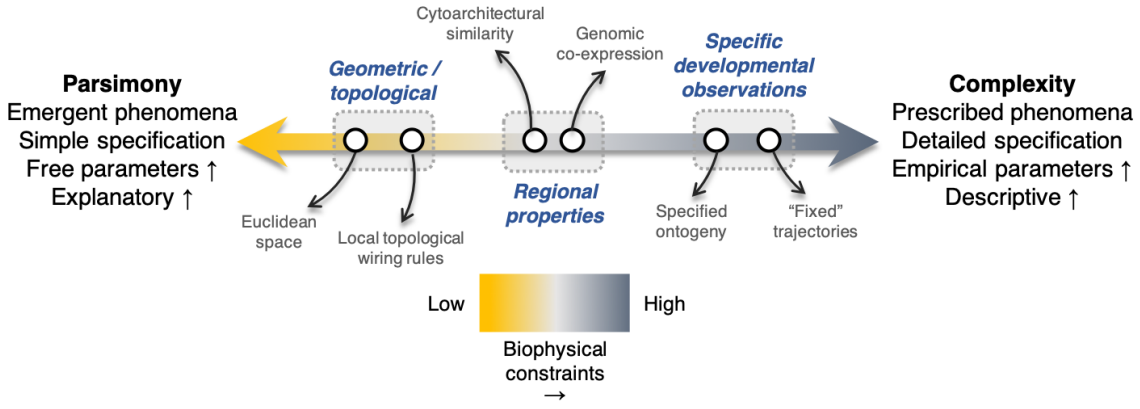


Figure 5.1. Trading-off parsimony and complexity when incorporating biophysical constraints. As one incorporates greater biophysical constraints into the generative model, the explanatory scope narrows and findings become increasingly prescribed and detailed. Example generative model specifications that may make models simpler (e.g., Euclidean-space, as only the physical

geometry is specified) or more complex (e.g., the precise spatiotemporal ontogeny of wiring) are provided.

Importantly, while the current instantiation of the generative network model provides a good account of the emergence of network organisation, it does not account for a number of observations important for developmentalists. For example, they say nothing directly about the response selectivity of particular regions, or how neocortical synapses may be pruned with time. In both cases newer formalisations of generative models will be required that incorporate some level of task-based information-processing into how they work. Notably, some of these questions have been attempted in the seRNN framework. An extended developmental model would evaluate the extent to which regional structures tune and competitively interact with each other over development, and deal with differing developmental time-course dynamics depending on the neural data modelled (e.g., grey-matter pruning (Lebel & Deoni, 2018; Quartz & Sejnowski, 1997), or incremental white-matter increases into adulthood (Lebel & Deoni, 2018)). Such a model would likely require the ability to simulate weighted networks (e.g., reflecting activity-dependent myelination of axonal projections underlying learning (Bonetto et al., 2021), which has been recognised as a key next step in the development of this field (Oldham et al., 2022).

5.2.2. Models operate across different levels

A second characteristic for a good computational model of developmental systems neuroscience is the extent to which it allows us to bridge across levels of biology. One way to form bridges is to colocalise genomic expression with structural organisation directly in developmental populations (Dorfschmidt et al., 2022; Romero-Garcia et al., 2020; Váša et al., 2020; Whitaker et al., 2016). In terms of computational models, as with spandrels, it is compelling when models optimised to explain phenomena at one level also explain emergent phenomena in different biological data – such as genomic expression. As outlined in **Chapter 2**, a benefit of the generative modelling is that developmental ingredients can be extracted in the form of the wiring parameters and projected to the nodal level to evaluate their changes of simulated time

steps. These parameters can be used to leverage alternative independent datasets that capture biological processes occurring at different levels of explanation. For example, they can be integrated with regional variation in gene expression to identify convergent genes with expression profiles that significantly covary with either parameter. In turn, as in **Chapter 2**, these gene lists can then be subjected to a genomic expression analysis, in which we test formally whether the genes code for integrated processes within particular cellular or biological pathways. Likewise, the wiring parameters, the time-courses of network formation, or emergent properties from the model itself can each be associated at an individual level with other phenotypic data, such as cognitive or behavioural assessments (Zhang et al., 2021). The end result is that the computational model has provided an associative bridge between multiple levels of analysis, from gene expression, cellular and biological pathways, the emergence of networks and ultimately cognitive or behavioural outcomes.

An alternative approach is to integrate information (e.g., gene expression, as above) directly across levels of biology, into a single developmental model (e.g., Arnatkevičiūtė et al., 2021; Oldham et al., 2022). For example, Oldham, *et al.* (Oldham et al., 2022) utilise the analogous gene-expression data that was shown in **Chapter 2** (from the Allen Human Brain Atlas) to generate a node-node correlation index, termed the correlated gene-expression (CGE). The uncorrected CGE (uCGE) itself contains a large spatial-autocorrelation, in that regions that are closer in space also tend to have greater correlated gene expression. When using this model as a single matrix to probabilistically add edges, according to the following equation:

$$P_{i,j} \propto uCGE_{i,j}{}^{\gamma}, \tag{17}$$

the authors found very good model fits to empirical macroscopic brain networks. They showed that, when starting from an empty network, this improves spatial-embedding in space relative to more canonical topological models (we return to this consideration in a moment). Notably, model fits were shown to be comparable to the matching topological model that was emphasised in **Chapter 2** and **3**. How should such a result be considered in context of the

topological findings in this thesis and elsewhere? Numerous authors implicitly suggest that genetic and topological arguments for wiring are somewhat incongruent with each other (Arnatkevičiūtė et al., 2021; Oldham et al., 2022; Rubinov, 2016, 2022). However, I would argue that a better interpretation is to consider that genomics and topology are likely harmonious. Evolution has likely shaped how the genome *compresses* the coding for topological outcomes, in the context of finite-economic resources, via a genomic "bottleneck" (Zador, 2019). Indeed, one pervasive way to think about the early development of neural structures is as them having undergone evolutionary *pre-training*; where the historical training of the organism is provided via its genetic code. This may endow neural networks with an initial structure that induces numerous computational capacities, without requiring experience-dependent plasticity (see Zador, 2019 for an elaboration of this argument). This idea is neatly demonstrated in recent work (Stöckl et al., 2021) which formalises a generative model, termed the probabilistic skeleton, which is a recurrent spiking neural network developed from empirical connectivity probabilities between 17 neuronal types in mouse V1 (Billeh et al., 2020) that is shown to exhibit basic computational functionality. Crucially, the relatively broad computational function of these neural networks is induced through their connectivity structure (rather than through synaptic plasticity). In biological systems, any genetic influence on connectivity will occur in content of the numerous constraints, such as those as outlined in **Chapter 4** which, in contrast, shows how an embedded network may learn to achieve a specific, albeit simple, navigational task. Taken together, neurobiological levels (e.g., genes, neurons, agents, environments) do not operate in isolation, and therefore a good developmental model should also not operate in isolation. I would argue that it is much more likely that levels complement and shape each other – a central early idea introduced in **Chapter 1** (Gottlieb, 1998).

When spanning across-levels, the same problem outlined in the previous section remains pertinent – what information should we pre-specify? Previously, I argued that one should be cautious when adding constraints into a developmental model, as possible explanations may leak across other domains. One concrete example is in the seeding of generative network

models, as outlined in **Chapter 2** and **3**. While at first, this may seem like a simple design choice, the locations that connections initially form has a clear theoretical implication with possible consequences for results and interpretations. This was recognised in the first generative modelling study (Kaiser & Hilgetag, 2004), which positioned the initial two areas if connection formation in medio-posterior positions to facilitate growth into lateral and anterior directions, mirroring human cortical development. In contrast, in **Chapter 2**, we initialised seeding across subjects to be edges that are common to all subjects (i.e., present in 100% of subjects, see **Figure B2**), meaning that simulations begin from a point at which there are no differences between subjects. This facilitated the study of how *variability* arises via the generative network process – the core topic of that study. There are numerous implications regarding such a design choice. These connections are by definition, common to all subjects. Any argument regarding neurodiversity or variability will naturally not extend to these *invariable* edges – and must be explained in their own terms. Indeed, one must think about what drives this disparity in the consistency of edges. One possibility is that highly consistent edges are under greater genetic influence in their formation, but variable edges are under greater environmental influence (Arnatkevičiūtė et al., 2021). Importantly, when setting seeds as common to all participants, it is possible that early genetic effects may be "baked-in" into the analysis – which is perfectly valid so long as the resulting interpretations are justified, given model inputs are considered. Of note, in **Chapter 3**, no seed network was specified and this reflects the spontaneous nature of neuronal network self-organisation over time at this microscale. Finally, when one provides a seed to the generative algorithm, the effect of temporal accumulation (i.e., the old-get-richer; Kaiser, 2017)) can be investigated as a consequence of simply the time-availability for wiring opportunity (see Oldham & Fornito, 2019 for a review on the development of hub connectivity). Indeed, rich club organisation of highly interconnected cortical hubs has been observed to arise by 30 weeks' gestation (Ball et al., 2014). Using this rich club network as the seed would likely be more biologically veridical.

Another point to consider, related to seeding, is to what extent that connections which develop early in the model simulation can cause biologically unrealistic instabilities in the outcome.

For example, it is possible that a low-probability event that occurs early may have a disproportionate effect on the outcome. It is hard to justify if this is a biologically relevant phenomenon or simply an artefact of the model. Nevertheless, this observation underscores the importance of both the initial conditions and constrained stochasticity in wiring parameterisation – both of which determine the stability of the outcome.

The potential role of temporal proximity and its contribution to topology extends the broad idea of self-similarity beyond just topology and space (as emphasised in **Chapter 2** and **3**) to self-similarity in space, topology and *time*. This observation opens the door to the speculation that self-similarity principles may also extend across levels of analysis. For example, one avenue, would be to investigate how well generative models can explain the development of representational spaces with characteristic structures (e.g., temporal lobe connectivity and semantic representational networks). How may such questions be addressed? As outlined previously, a current limitation is that most bridges between biological levels are not direct, but associative. A future possibility is that multiple levels of biology could be formalised and integrated into a multi-layered generative network model (Bazzi et al., 2020; Betzel & Bassett, 2017a) which directly spans levels (Stöckl et al., 2021). It may be that homophily may also approximate these networks. Alternatively, various non-homophilic interactions (even heterophilic) may indeed be very adaptive at certain levels, enabling diverse dynamics between levels (Bazinet et al., 2022; Betzel & Bassett, 2018; Park & Friston, 2013). Future work could elaborate on how concurrent structural and functional optimisation algorithms (as in **Chapter 4**) may be able to capture how both interaction types emerge within singular coherent objective functions, given various computational demands.

5.2.3. Models span across species

Evolution shapes the brains of humans and other animals (Cisek & Hayden, 2022). Thus, just as we can bridge different levels of analysis within a species, we can also bridge across different species. Comparative connectomics aims to examine the products of evolution in the context of brain networks (Heuvel et al., 2016). Specifically, it aims to quantitatively delineate the

similarities and differences between the brain networks of humans and other species with reference to the phylogenetic tree; allowing us to uncover common features across all mammals versus features that are uniquely human.

A good developmental systems model should be applicable across the phylogenetic tree, enabling for insights as to both universal principles across species and specific principles within species. Why is this important? Most developmental theories to date make little reference to other species. This leaves a significant explanatory gap in the neurosciences between macroscopic descriptions of the human brain and cognition, and more detailed descriptions which are possible in non-human animal research (Barron et al., 2021). A good systems developmental model must therefore establish in what ways brains differ in terms of their development, as a function of a common ancestry (Finlay et al., 2001). In **Chapter 3**, some headway was made in this endeavour, by showing commonalities across species in terms of generative wiring findings.

Connectomics may be a particularly useful lens to view how cognitive capacities can rapidly change over evolution, particularly in how sparsity, modularity and laminar differentiation may occur in context of possibly very small regulatory-gene changes (Changeux et al., 2021). However, as evolution describes the processes by which organisms change over time according to natural selection pressures, the various demands placed on these organisms to survive must be central to the explanatory story. That is, it will not be sufficient to simply observe the anatomical and functional organisation of the brain in isolation of the behavioural goals that the organism was selected upon. The modelling framework put forward in **Chapter 4** may be particularly useful in tackling this problem. Depending on the various organism, it may be possible to model directly the range of behavioural requirements placed on them, given the particular physical constraints encountered by its neural architectures, simulating the various problems that the organism faces in the real world. Put simply, rather than positing how differences in network architecture between closely-related species may indicate specific adaptations in cognition and behaviour (Heuvel et al., 2016), this could be examined directly. This could allow for a wider scope of questions to be answered, such as how the availability of

resources to various organisms in their ecosystems may concurrently influence anatomy and function. For example, a critical influence on biodiversity is the stability and resource-richness of the climate, highlighted by the fascinating observation of the latitudinal diversity gradient: the number of species increases as you move from the poles to the equator (Fine, 2015; Saupe et al., 2019). Without directly modelling how neural architectures exist within these complex environments, comparative connectomics is limited but expansions on the framework set out in **Chapter 4** could begin to make advances in this area. In summary, moving modelled neural architectures closer to both the world in which organisms reside, in addition to the structural and metabolic problems they must solve, may yield significant insights into species-specific behaviours as a function of the phylogenetic tree.

There are also practical benefits of spanning across species for understanding developmental processes. Non-human animal research has the additional natural advantage of allowing us to embrace experimental methods to probe causal mechanisms, which is simply not possible in humans. This can allow us to delineate via the study of various organisms the origins of variability. On one hand, some organisms such as the *Caenorhabditis elegans* have remarkable reproducibility of the observed network across individual organisms (i.e., they are largely identical), allowing for the detailed study as to how this neuronal connectivity is encoded directly into the genome (Barabási & Barabási, 2020; Taylor et al., 2021), and matures (Witvliet et al., 2021). On the other hand, "higher" organisms exhibit significant experience and learning-dependent variability in their networks. Non-human animal research in this context, combined with computational models, may be a particularly powerful tool in elucidating how variability may be causally influenced by these factors. One example is how early-life adversity may induce brain structure adaptation, such as in the work by Carozza, S. *et al.* (Carozza et al., 2022), who show how early adversity may directly lead to heightened stochasticity in network development. Stochasticity is a core developmental principle (Hiesinger & Hassan, 2018; Vogt, 2015) in brain network formation, and is argued to potentially make networks more robust to perturbation to facilitate effective responses to

future threats. This highlights how a good developmental model should coherently integrate developmental principles across species.

5.2.4. Models elucidate developmental processes and time-courses

As described in **Chapter 1**, constructivist principles entail the core principle that to understand the functional role of neural circuits we must understand the developmental history of those circuits. Therefore, a good developmental systems model should tackle – at its core – the temporal and stochastic nature of interactions between each part of the system, over time. This is an intrinsically challenging aspect of developmental systems and cognitive neuroscience, because the data necessary to take years to collect. A good computational model of developmental processes needs to go beyond explaining differences in end state, and provide insight into the ongoing processes that produced that end state. For example, one of the classical observations in developmental systems neuroscience is the integration and segregation of networks over time (Fair et al., 2007). Variability to the time-courses of these processes is significantly associated with multiple neurodevelopmental conditions (Jones et al., 2022).

Considering stochasticity is critical to specifying generative models because it shapes the variability of developmental outcomes. On one hand, if a model has no stochasticity then development can be considered purely in terms of determined interactions between regions where there are identical outcomes upon multiple runs. On the other hand, if the model is wholly stochastic then development is equivalent to completely random wiring. Therefore, the goal of a good developmental model must be to balance determined versus stochastic effects within a "constrained developmental manifold" which reflects the expected dimensions along which typical neurodiversity can be expected to exist and be observed. The constrained element of this can be studied via the incorporation of known neurodevelopmental details into a model at biologically meaningful time-courses (Betzel & Bassett, 2017b), such as in the work of Nicosia, V. *et al.* (Nicosia et al., 2013) with the nematode worm *Caenorhabditis elegans* connectome (which was discussed in more detail previously in **Chapter 2**). The stochastic

manifold can be studied via simulations of neurodiversity with generative modelling. In turn, this can be used to evaluate the space of wiring parameters and network outcomes that map to empirical data. A fruitful area of future study would be to explicitly detach any "determined" versus "stochastic" effects within the generative model and study their interaction and/or temporal dynamics. This is particularly important because, as mentioned previously, stochasticity is likely a critical feature of brain network development that endows robustness to the system (Hiesinger & Hassan, 2018) and flexibility to adaptation (Carozza et al., 2022).

A generative model allows us to establish the mechanisms by which this may occur within synthetic networks, revealing insight into the real thing. For example, despite being a simple generative model, we find in **Chapter 3** that segregation and integration (modelled as network modularity and efficiency) happen as an emergent property of economic negotiations in the generative model when fit to neuronal networks. In the future, a computational developmental systems model could be ideal for testing more specific hypotheses about temporal shifts in network statistics and developmental discontinuities in brain development (e.g., so-called sensitive periods), as well as how these timings are influenced by the social and economic environment of children (Tooley et al., 2021). To achieve this, a developmental model may be required to incorporate within-network longitudinal measurements directly into the formalisation. With an appropriate dataset, this could be achieved by fitting generative models directly to the change of developing observations (i.e., a single cost function that encompasses multiple longitudinal observations) rather than fit to each observation individually (as was done in **Chapter 3**) to recapitulate empirical growth curves (e.g., Bethlehem et al., 2022). In conclusion, only with a good developmental systems model that elucidates developmental processes and time-courses can we hope to make progress bridging mechanism with observation.

5.3. Concluding remarks

We make scientific progress when conceptual frameworks, methods and techniques are developed and make it possible to test competing explanations of data. However, due to the sheer complexity of brain development, it is challenging to move from description to explanation because *hard-to-vary* hypotheses – which can be systematically tested – are themselves hard to come by. The present work shows how computational models may be a very useful tool in addressing this. So far, these models highlight how the temporal unfolding of simple economical negotiations between the components of a developing system can lead to the emergence of highly complex outcomes. These trade-offs are both structural and functional in nature and are likely critical to how the brain constructs itself over development across multiple scales of analysis. The present work lays a foundation for numerous new and exciting multidisciplinary research directions with promise to make scientific progress in our understanding of brain development.

6. APPENDIX A

SUPPLEMENTARY MATERIALS FOR CHAPTER 1

Graph theory network statistic formalisms

Throughout the presented works, a range of graph theory network statistics are computed. Each statistic was calculated using the Brain Connectivity Toolbox (Rubinov & Sporns, 2010), and the mathematical formalisms are provided.

Degree. The degree is the number of edges connected to a node. The degree of node i is given by:

$$k_i = \sum_{j \in N} a_{i,j}, \tag{18}$$

where $a_{i,j}$ is the connection status between i and j. $a_{i,j}$=1 when link i, exists (when i and j are neighbours); $a_{i,j}$=0 otherwise ($a_{i,i}$=0 for all i).

Clustering coefficient. The clustering coefficient is the fraction of a node's neighbours that are neighbours of each other. The clustering coefficient for node i is given by:

$$c_i = \frac{1}{n} \sum_{i \in N} \frac{2t_i}{k_i(k_i - 1)}, \tag{19}$$

where c_i is the clustering coefficient of node i (c_i=0 for k_i<2).

Betweenness centrality. The betweenness centrality is the fraction of all shortest paths in the network that contain a given node. Nodes with high values of betweenness centrality therefore participate in a large number of shortest paths. The betweenness centrality for node i is given by:

$$b_i = \frac{1}{(n-1)(n-2)} \sum_{h,j \in N} \frac{\rho_{hj}(i)}{\rho_{hj}}, \tag{20}$$

where ρ_{hj} is the number of shortest paths between h and j, and $\rho_{hj}(i)$ is the number of shortest paths between h and j that pass through i.

Edge length. The edge length is the total edge lengths connected to a node. It is given by:

$$d_i = \sum_{j \in N} a_{i,j} d_{i,j}, \tag{21}$$

where $d_{i,j}$ is the Euclidean distance between i and j.

Global efficiency. The global efficiency is the average of inverse shortest path length. It is given by:

$$E_i = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N, j \neq i} d_{i,j}^{-1}}{n-1} \tag{22}$$

Matching. The matching index computes the proportion of overlap in the connectivity between two nodes. It is given by:

$$M_{i,j} = \frac{|N_{i/j} \cap N_{j/i}|}{|N_{i/j} \cup N_{j/i}|}, \tag{23}$$

where $N_{i/j}$ refers to neighbors of the node i excluding node j. Where global measures of matching have been used, we averaged across the upper triangle of the computed matching matrix.

Small-worldness. Small-worldness refers to a graph property where most nodes are not neighbours of one another, but the neighbours of nodes are likely to be neighbours of each other. This means that most nodes can be reached from every other node in a small number of steps. It is given by:

$$\sigma = \frac{c/c_{rand}}{l/l_{rand}}, \tag{24}$$

where c and c_{rand} are the clustering coefficients, and l and l_{rand} are the characteristic path lengths of the respective tested network and a random network with the same size and density of the empirical network. Networks are generally considered as small-world networks at $\sigma>1$. In our work, we computed the random network as the mean statistic across a distribution of n=1000 random networks. The characteristic path length is given by:

$$L_i = \frac{1}{n}\sum_{i\in N}\frac{\sum_{j\in N, j\neq i} d_{i,j}}{n-1} \tag{25}$$

Modularity. The modularity statistic, Q, quantifies the extent to which the network can be subdivided into clearly delineated groups:

$$Q = \frac{1}{l}\sum_{i,j\in N}\left(a_{i,j} - \frac{k_i k_{,j}}{l}\right)\delta_{m_i m_j} \tag{26}$$

where m_i is the module containing node i, and $\delta_{m_i m_j}$=1 if m_i=m_j, and 0 otherwise.

Participation coefficient. The participation coefficient is a measure of diversity of intermodular connections of individual nodes, where community allocation was determined via a Louvain algorithm, with a resolution parameter γ=1, which aims to form a subdivision of the network which maximises the number of within-group edges and minimises between group edges.

7. APPENDIX B SUPPLEMENTARY MATERIALS FOR CHAPTER 2

Content included in this appendix was published in Nature Communications, 2021: A generative network model of neurodevelopmental diversity in structural brain organisation. Danyal Akarca, Petra E. Vértes, Edward T. Bullmore., the CALM team & Duncan E. Astle.

Figure B1. Referral rates and cognitive variability across the CALM cohort.

Figure B2. Seed network construction and generative model overview.

Figure B3. Gene loading analysis and pathway enrichment.

Figure B4. Wiring parameter variability across the n=500 lowest energy simulations.

Figure B5. Exploration of energy and KS statistics, and the resultant parameter combination relationships.

Figure B6. Out-of-sample network measures, not included in the energy equation, and their relationships between simulated and observed measures.

Figure B7. Computing the generative mismatch between low energy simulated networks and those observed.

Figure B8. Replication of key findings in the external n=140 RED dataset.

Figure B9. Controlling for age has little influence on the relationships between statistical measures, and how they relate to wiring parameters.

Table B1. Demographic information of cohorts used.

Table B2. Generative rules, their formulae and descriptive statistics across the broad parameter space.

Table B3. Descriptive statistics of high performing wiring parameters and the energy of the networks they produce.

Table B4. Gene queries submitted to g:Profiler.

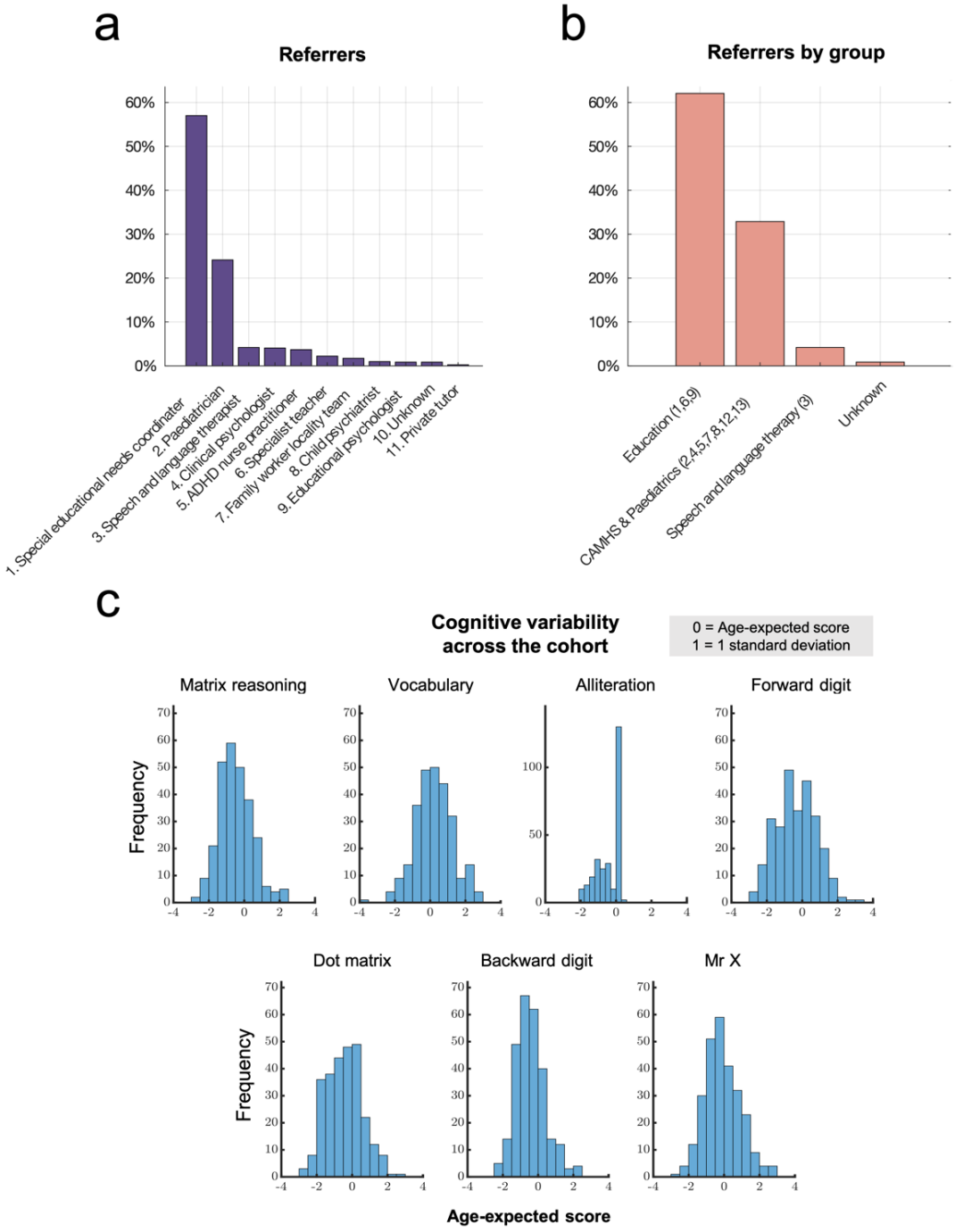


Figure B1. Referral rates and cognitive variability across the CALM cohort. (a) The majority of referrals came from Special Educational Needs Coordinators followed by Paediatricians. (b) Referrers by broad group. CAMHS; Child and Adolescent Mental Health Services. (c) Age-expected scores across the seven cognitive variables.

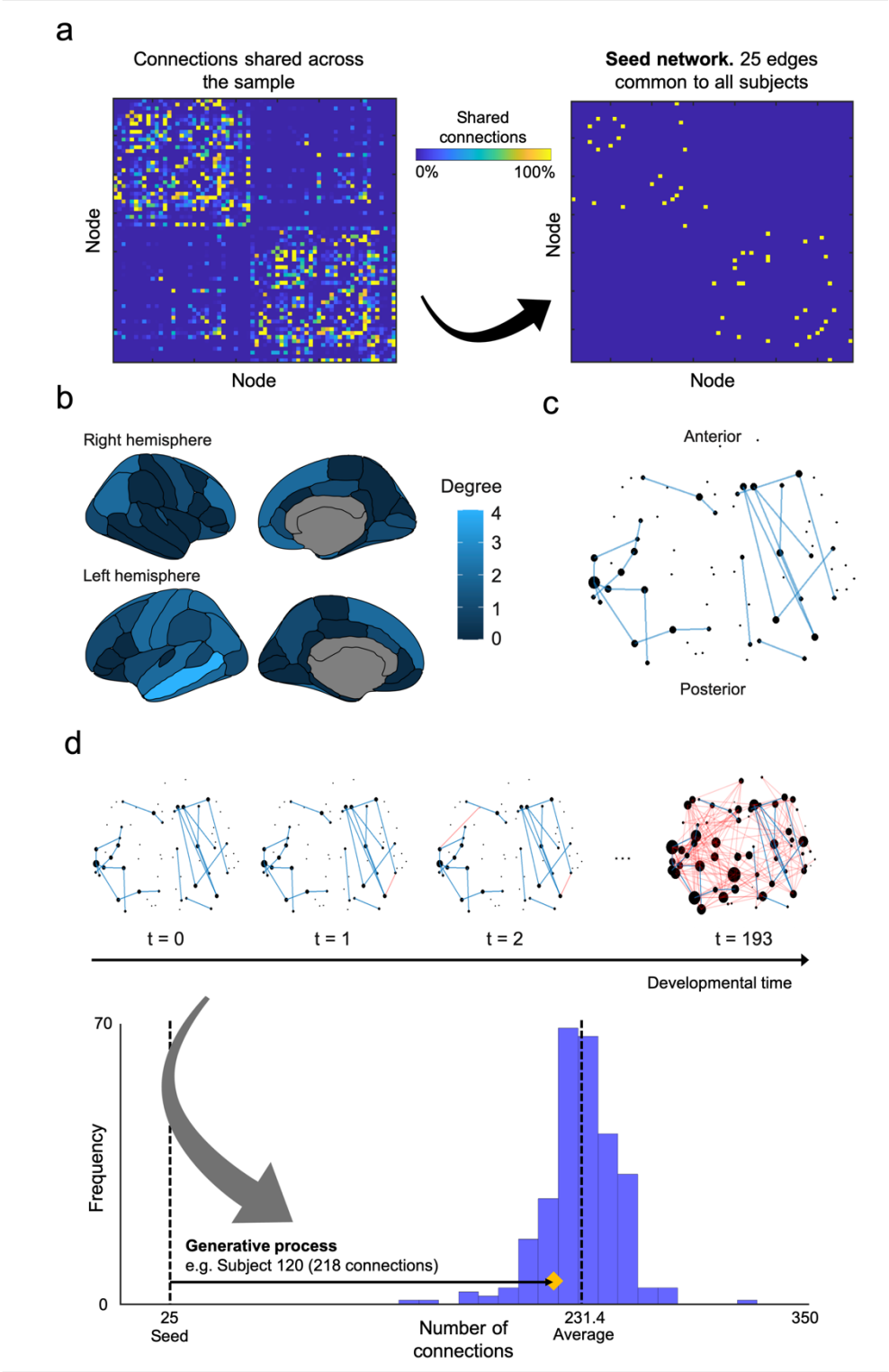
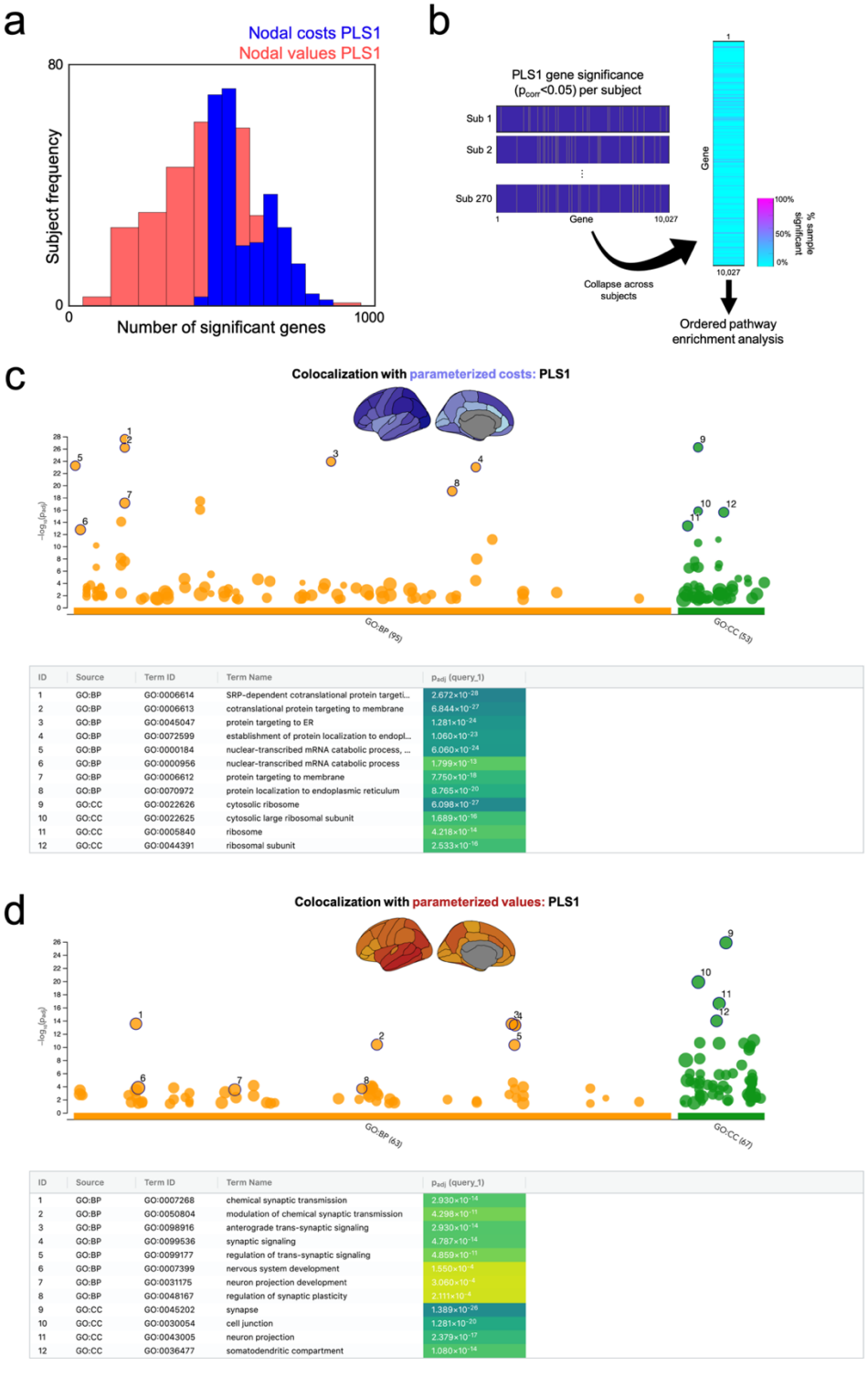


Figure B2. Seed network construction and generative model overview. (a) The seed network was constructed by taking connections that were shared across 100% of the sample. This left

25 bidirectional connections that were common across all subjects. (b) The degree distribution of the seed network. (c) An edge-wise view of the seed network. Node size relates to degree. (d) A schematic illustration of how generative models aim to uncover the wiring rules that give rise to an observed brain network. Starting from the seed network (10.8% density of the average observed network), edges are added to the network according to parameterised costs and value (i.e., rule) terms until it reaches the same number of edges as observed. For this illustration, subject 120 is shown, where 193 edges are added to the seed's 25 edges, leaving 218 edges in the final network. The distribution of number of observed edges are shown, with a mean of 231.4 edges and standard deviation of 19.1.



ID	Source	Term ID	Term Name	p_{adj} (query_1)
1	GO:BP	GO:0006614	SRP-dependent cotranslational protein targeti...	2.672×10^{-28}
2	GO:BP	GO:0006613	cotranslational protein targeting to membrane	6.844×10^{-27}
3	GO:BP	GO:0045047	protein targeting to ER	1.281×10^{-24}
4	GO:BP	GO:0072599	establishment of protein localization to endopl...	1.060×10^{-23}
5	GO:BP	GO:0000184	nuclear-transcribed mRNA catabolic process, ...	6.060×10^{-24}
6	GO:BP	GO:0000956	nuclear-transcribed mRNA catabolic process	1.799×10^{-15}
7	GO:BP	GO:0006612	protein targeting to membrane	7.750×10^{-18}
8	GO:BP	GO:0070972	protein localization to endoplasmic reticulum	8.765×10^{-20}
9	GO:CC	GO:0022626	cytosolic ribosome	6.098×10^{-27}
10	GO:CC	GO:0022625	cytosolic large ribosomal subunit	1.689×10^{-16}
11	GO:CC	GO:0005840	ribosome	4.218×10^{-14}
12	GO:CC	GO:0044391	ribosomal subunit	2.533×10^{-16}

ID	Source	Term ID	Term Name	p_{adj} (query_1)
1	GO:BP	GO:0007268	chemical synaptic transmission	2.930×10^{-14}
2	GO:BP	GO:0050804	modulation of chemical synaptic transmission	4.298×10^{-11}
3	GO:BP	GO:0098916	anterograde trans-synaptic signaling	2.930×10^{-14}
4	GO:BP	GO:0099536	synaptic signaling	4.787×10^{-14}
5	GO:BP	GO:0099177	regulation of trans-synaptic signaling	4.859×10^{-11}
6	GO:BP	GO:0007399	nervous system development	1.550×10^{-4}
7	GO:BP	GO:0031175	neuron projection development	3.060×10^{-4}
8	GO:BP	GO:0048167	regulation of synaptic plasticity	2.111×10^{-4}
9	GO:CC	GO:0045202	synapse	1.389×10^{-26}
10	GO:CC	GO:0030054	cell junction	1.281×10^{-29}
11	GO:CC	GO:0043005	neuron projection	2.379×10^{-17}
12	GO:CC	GO:0036477	somatodendritic compartment	1.080×10^{-14}

Figure B3. Gene loading analysis and pathway enrichment. (a) Distributions of significant genes (p_{corr}<0.05) for the nodal-costs and nodal-values PLS1, which provided an average of 581.5 significant genes and 437.6 significant respectively across the sample. (b) To collapse

across subjects, we ranked genes according to the number of subjects for which each gene was significant (according to n=1000 permutations, p_{corr}<0.05) and submitted genes above 10% to an ordered pathway enrichment analysis, using g:Profiler. Pathway enrichment analysis summarises large gene sets as a smaller list of more easily interpretable pathways that can be visualised to identify main biological themes. (c) Parameterised costs pathway enrichment analysis, for biological processes and cellular components, and (d) Parameterised values.

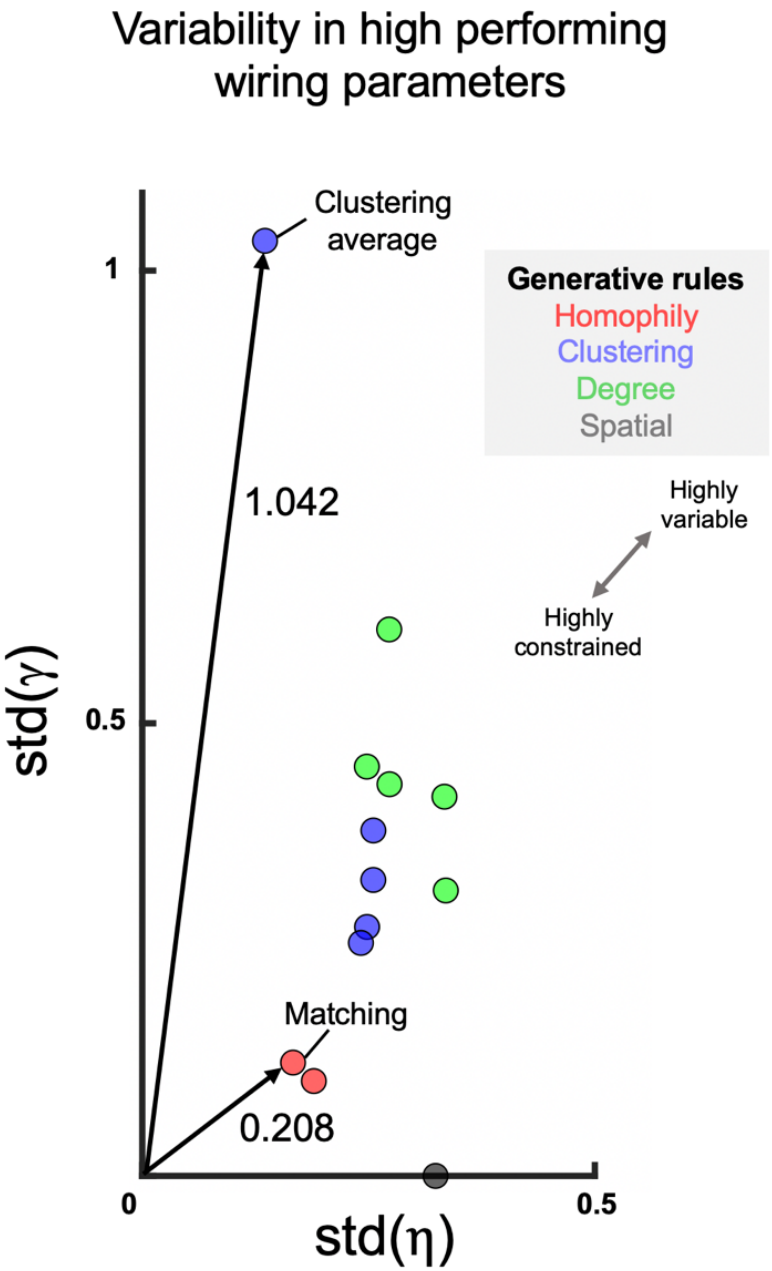


Figure B4. Wiring parameter variability across the n=500 lowest energy simulations. Each point corresponds to a coordinate denoting the standard deviation of the n=500 lowest energy η (x-axis) and γ (y-axis) parameters for each generative rule. Two vectors are plotted for the rule with the lowest variability (Matching) with a Euclidean of 0.208, and the highest variability (Clustering average) with a Euclidean of 1.042.

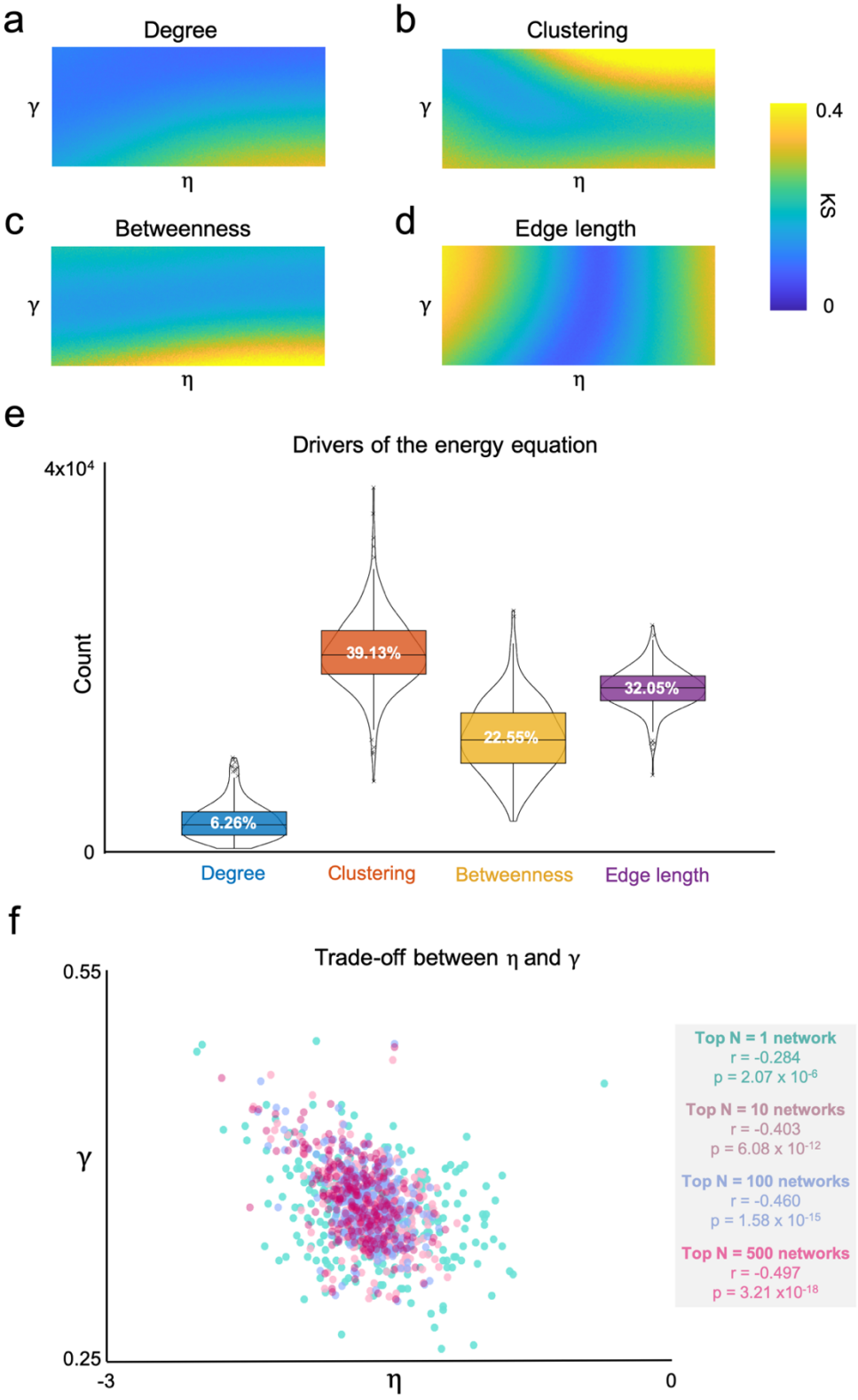


Figure B5. Exploration of energy and KS statistics, and the resultant parameter combination relationships. (a) Visualisation of group-averaged degree KS statistics in the window. (b)

Clustering coefficient KS statistics (c) Betweenness-centrality KS statistics. (d) Edge length statistics. (e) The KS statistics which were "maxed" as to produce the energy statistic, across all subjects and simulations in the narrow parameter window. In total, there are n=270x50,000=13,500,000 simulations. On average, the energy equation was driven by clustering 39.13% (n=5,282,550) of the time, followed by edge length at 32.05% (n=4,326,750), followed by betweenness at 22.55% (n=3,044,250) and degree at 6.26% (n=845,100). The boxplot presents the median and IQR. Outliers are demarcated as small black crosses, and are those which exceed 1.5 times the interquartile range away from the top or bottom of the box. (f) Within the narrow window, η and γ trade-off such that subjects with a larger γ have a greater preponderance to have a larger magnitude negative η. All statistics were computed via two-tailed linear correlations, quoting the Pearson's correlation coefficient.

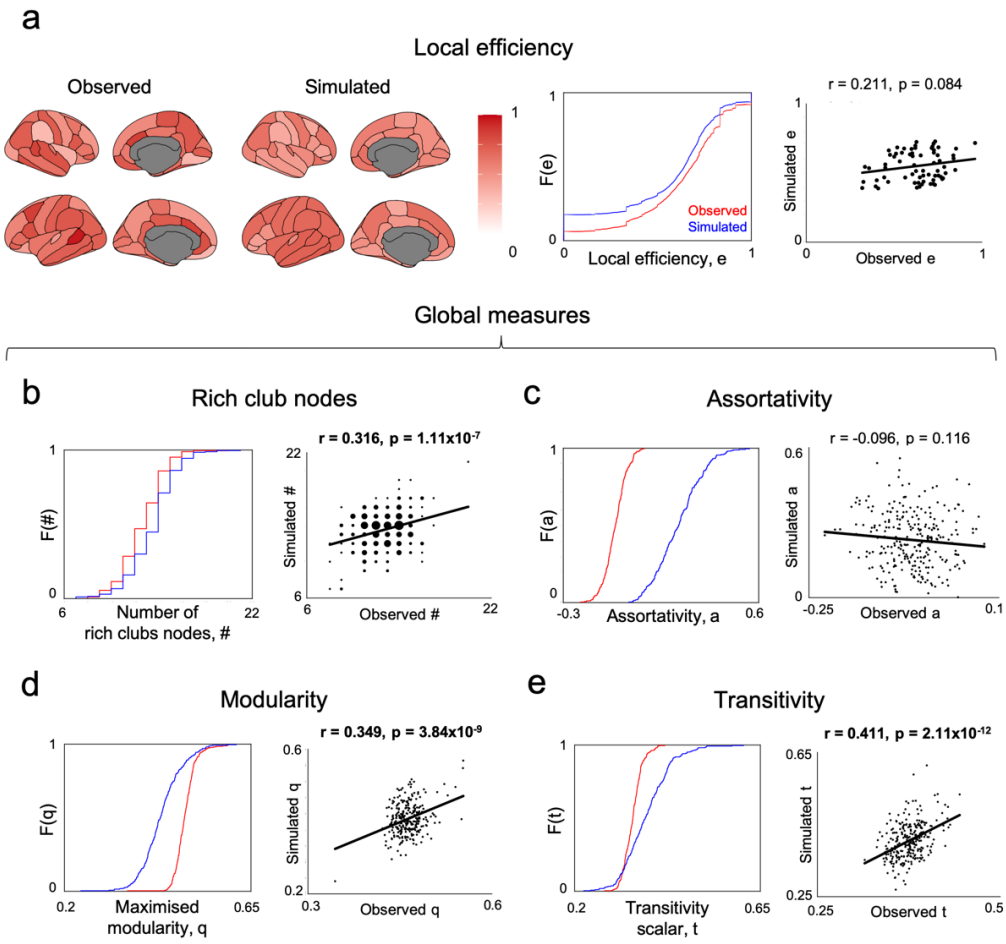


Figure B6. Out-of-sample network measures, not included in the energy equation, and their relationships between simulated and observed measures. Local efficiency is the only local measure, and thus, as in Figure 2.2, each point in the scatter plot reflects the across-subject average nodal measure from observed and optimal simulated networks. All other measures are global and therefore each point in the scatter represents a subject in the sample. All statistics were computed via two-tailed linear correlations, quoting the Pearson's correlation coefficient. (a) Local efficiency between observed and simulated networks are not significantly correlated (r=0.211, p=0.084). (b) The number of rich club nodes between observed and simulated networks are significantly positively correlated (r=0.316, p=1.11x10^{-7}). (c) Assortativity between observed and simulated networks are not significantly correlated (r=-0.096, p=0.116). (d) Modularity between observed and simulated networks are significantly positively correlated (r=0.349, p=3.84x10^{-9}). (e) Transitivity between observed and simulated networks are significantly positively correlated (r=0.411, p=2.11x10^{-12}). Boldened values are significant correlations at p<0.05.

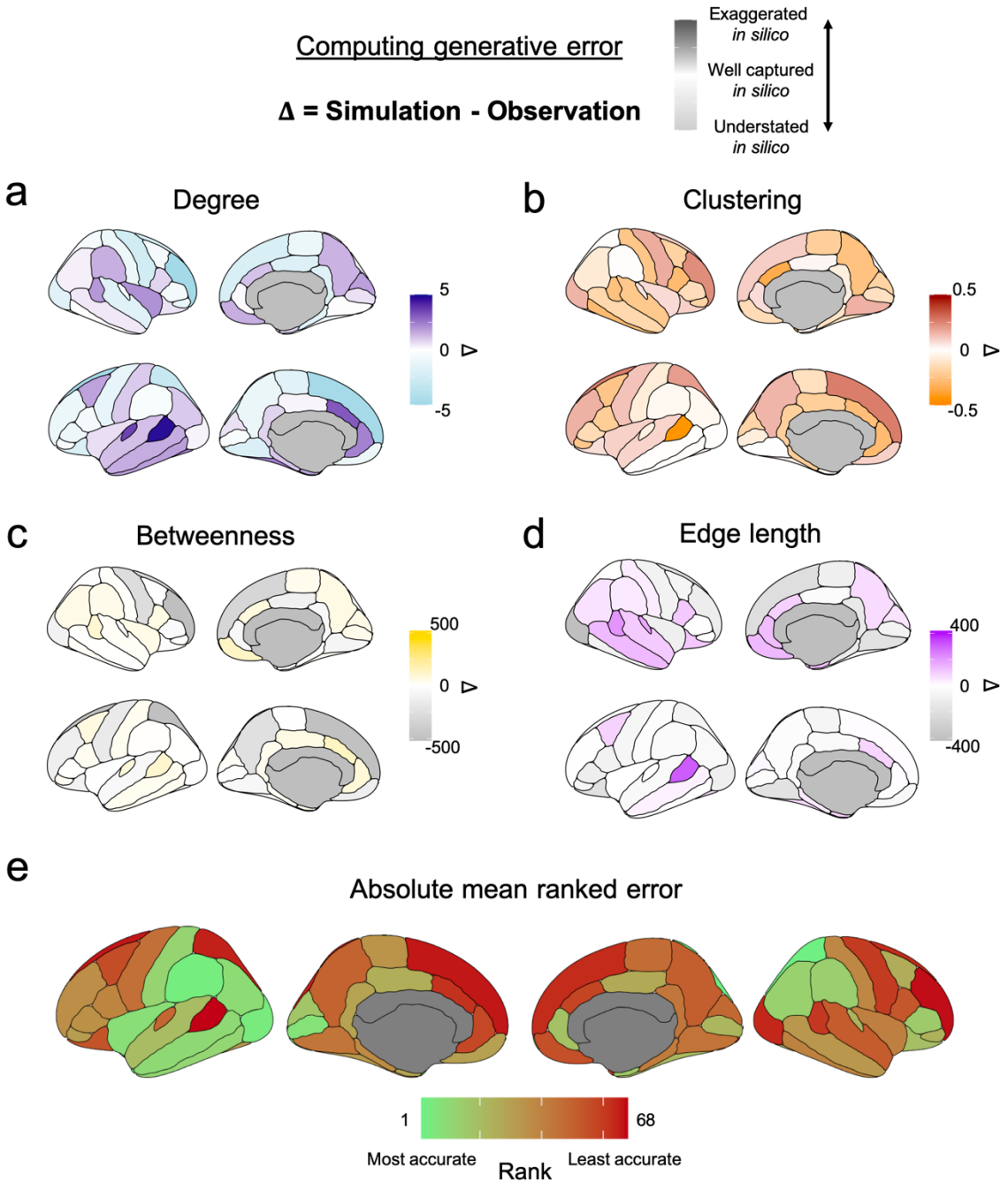


Figure B7. Computing the generative mismatch between low energy simulated networks and those observed. (a) Degree generative errors. (b) Clustering generative errors. (c) Betweenness centrality generative errors. (d) Edge length generative errors. (e) Visualisation of the absolute mean ranked error across the four measures, indicating the accuracy to which regions are approximated via the optimised homophily generative models.

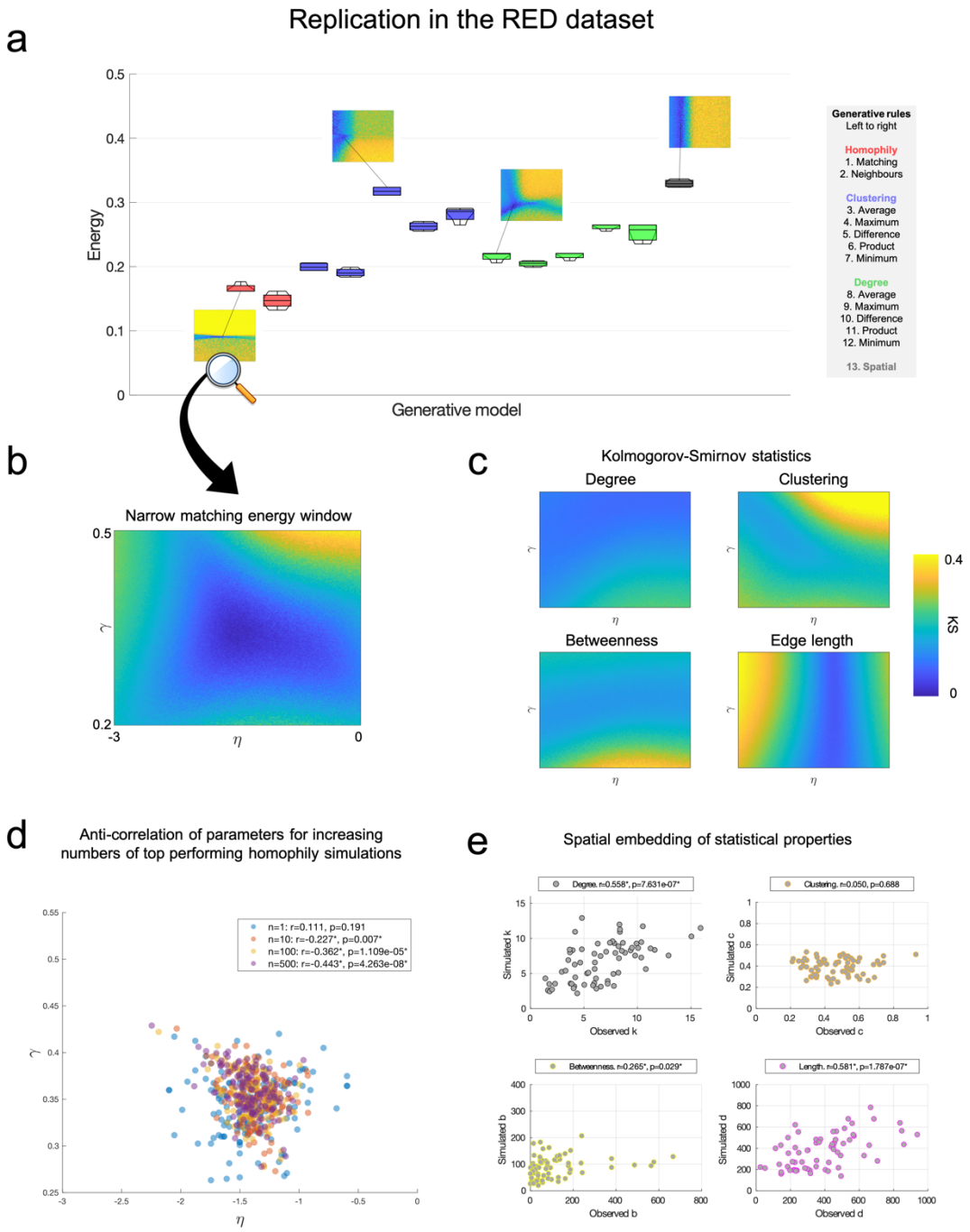


Figure B8. Replication of key findings in the external n=140 RED dataset. (a) Initial generative models were run in the same way as CALM, across thirteen generative rules via an evenly spaced grid-search for 10,000 simulations ($-7 \leq \eta \leq 7$, $-7 \leq \gamma \leq 7$) per rule (130,000 simulations in total). Connectomes were averaged across the sample and thresholded via the same procedure, to produce a single representative connectome on which the generative models were fit. Boxplots represent the highest n=5 performing simulations for this representative connectome. The boxplot presents the median and IQR. Four broad parameter spaces are

shown to highlight how they mirror landscapes as shown in Figure 2.1. (b) The homophily energy window was then computed in the same way as before, with n=50,000 simulations being run for each subject. This visualisation shows the energy window after being averaged over each subject's window. (c) The corresponding Kolmogorov-Smirnov statistics. (d) We find a very similar pattern of parameter anti-correlations in the homophily window, and that negative correlations increase with increasing numbers of high performing parameters (n=1, n=10, n=100 and n=500). Statistics were computed via two-tailed linear correlations, quoting the Pearson's correlation coefficient. (e) We replicate key findings that spatial embedding occurs as a result of the generative homophily model, as shown in Figure 2.2. All statistics were computed via two-tailed linear correlations, quoting the Pearson's correlation coefficient.

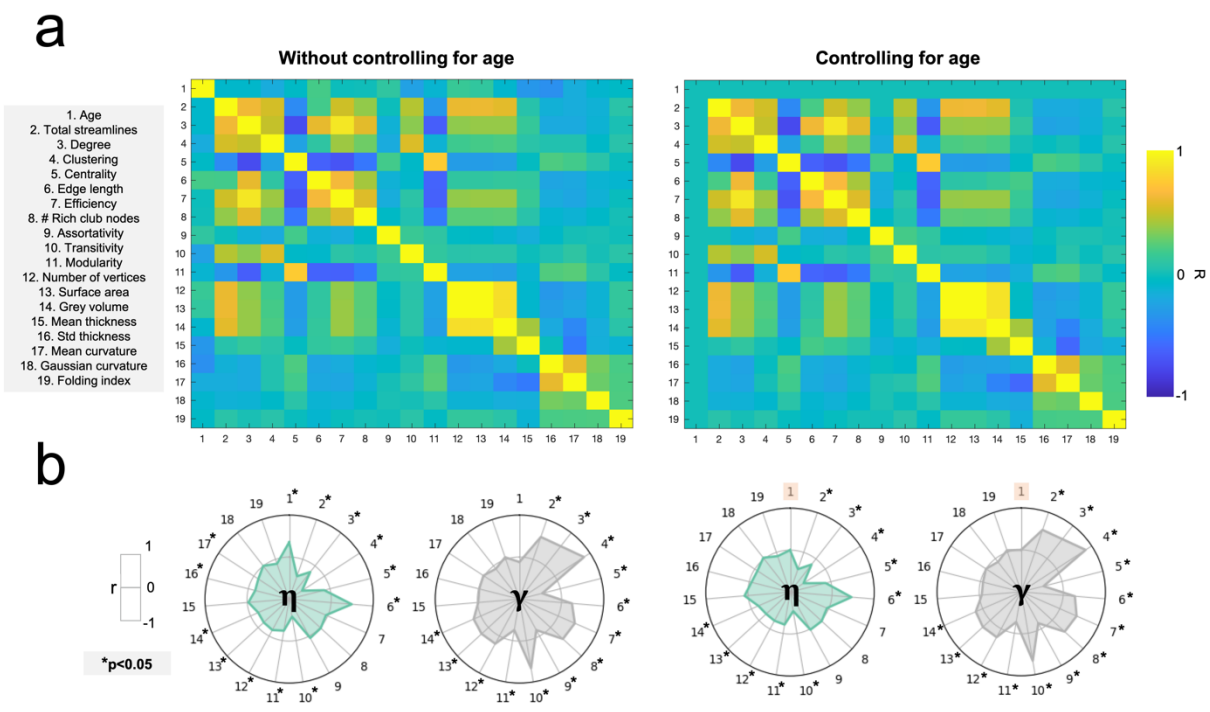


Figure B9. Controlling for age has little influence on the relationships between statistical measures, and how they relate to wiring parameters. (a) Correlation matrices between statistical measures without controlling for age (Left) and when age has been controlled (Right) appear near identical. (b) Radar plots of the correlation coefficients between and and each statistical measure.

Cohort	CALM	RED
n	270	140
Age (years)		
Mean	9.82	9.34
SD	2.15	1.41
Range	5.50–18.60	6.82–12.8
Sex		
Male	178	64
Female	92	76
Referred		
Yes	192	n/a
No	78	n/a

Table B1. Demographic information of cohorts used. Samples include the main CALM and validation RED cohorts.

Rule			Energy					
Name	**Class**	$K_{i,j}$	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Spatial	Spatial	*1*	0.337	0.020	-3.940	0.397	n/a	n/a
Neighbours	Homophily	$\sum_w a_{iw}a_{wj}$	0.165	0.020	-1.995	0.483	0.338	0.045
Matching	Homophily	$\frac{\lvert \Gamma i \setminus j \cap \Gamma j \setminus i \rvert}{\lvert \Gamma i \setminus j \cup \Gamma j \setminus i \rvert}$	0.141	0.021	-1.647	0.490	0.357	0.022
C-Average	Cluster	$\frac{Ci}{2}+\frac{Cj}{2}$	0.181	0.016	-3.070	0.378	-6.203	0.735
C-Minimum	Cluster	$min(C_i, C_j)$	0.252	0.020	-3.626	0.396	-6.131	0.795
C-Maximum	Cluster	$max(C_i, C_j)$	0.182	0.014	-3.095	0.387	-6.153	0.735
C-Difference	Cluster	$\lvert C_i - C_j \rvert$	0.322	0.025	-6.193	0.744	-3.480	2.012
C-Product	Cluster	C_iC_j	0.223	0.024	-3.401	0.393	-5.964	0.926
D-Average	Degree	$\frac{ki}{2}+\frac{kj}{2}$	0.221	0.019	-3.516	0.502	2.419	0.332
D-Minimum	Degree	$min(k_i, k_j)$	0.279	0.016	-4.749	0.749	0.474	0.010
D-Maximum	Degree	$max(k_i, k_j)$	0.216	0.019	-3.557	0.602	2.255	0.332
D-Difference	Degree	$\lvert k_i - k_j \rvert$	0.229	0.018	-3.836	0.669	1.689	0.577
D-Product	Degree	k_ik_j	0.277	0.018	-4.605	0.738	0.399	0.082

Table B2. Generative rules, their formulae and descriptive statistics across the broad parameter space. Descriptive statistics of the η and γ parameters required to achieve the best performing network for each subject, across each of the generative rules over the space defined ($-7 \leq \eta \leq 7$, $-7 \leq \gamma \leq 7$) for 10,000 simulations.

The matching narrow energy window						
Averaged over top n networks	**Energy**					
	Mean	SD	Mean	SD	Mean	SD
***n*=1**	0.090	0.010	-1.565	0.327	0.349	0.033
***n*=10**	0.101	0.011	-1.579	0.334	0.352	0.036
***n*=100**	0.116	0.013	-1.614	0.372	0.355	0.038
***n*=500**	0.1326	0.015	-1.658	0.427	0.357	0.040

Table B3. Descriptive statistics of high performing wiring parameters and the energy of the networks they produce. Parameters were selected using the matching generative rule with 50,000 simulations, evenly spaced across the narrow window ($-3.606 \leq \eta \leq 0.354$, $0.212 \leq \gamma \leq 0.495$). Parameters were averaged across a variable n number of high performing wiring combinations to determine stability

PLS	Response	# Genes	Gene Ontology	Link
PLS1	Parametrised costs	1427	Biological processes and cellular components	https://biit.cs.ut.ee/gplink/l/TYGoJHihSs
	Parameterised values	1584	Biological processes and cellular components	https://biit.cs.ut.ee/gplink/l/H_gK8UIVR1

Table B4. Gene queries submitted to g:Profiler. Ordered gene lists were computed by taking subject-specific significant (p_{corr}<0.05) genes according to their frequency across the sample, up to 10% of the sample.

8. APPENDIX C SUPPLEMENTARY MATERIALS FOR CHAPTER 3

Content included in this chapter was published in bioRxiv, 2022, and is currently in review at eLife: Homophilic wiring principles underpin neuronal network topology in vitro. ***Danyal Akarca****, Alexander W. E. Dunn*, Philipp J. Hornauer, Silvia Ronchi, Michele Fiscella, Congwei Wang, Marco Terrigno, Ravi Jagasia, Petra E. Vértes, Susanna B. Mierau, Ole Paulsen, Stephen J. Eglen, Andreas Hierlemann, Duncan E. Astle† & Manuel Schröter† (*Co-lead authors, †Co-senior authors).*

Figure C1. Tracking neuronal networks at cellular resolution on high-density microelectrode arrays.

Figure C2. Global topological measures of sparse rodent cultures over time.

Figure C3. Distribution of inter-neuronal Euclidean distances across rodent and human datasets.

Figure C4. Generative network modelling results for the sparse rodent primary cortical cultures.

Figure C5. Alternative generative model procedures.

Figure C6. Generative model fits recapitulate observed network statistics.

Figure C7. Comparison of global network statistics across sparse and dense primary cortical rodent networks at DIV14.

Figure C8. Generative network modelling results for the dense rodent primary cortical cultures.

Figure C9. Homophilic generative mechanisms best account for local relationships in developing dense rodent neuronal cultures.

Figure C10. Relationship between model energy and topological fingerprint dissimilarity.

Figure C11. Generative model comparisons between control, gabazine-treated and washout sparse rodent PC cultures.

Figure C12. Spiking dynamics in human monolayer and cerebral organoid cultures compared to rodent cultures.

Table C1. Overview of used antibodies.

Table C2. Statistical comparisons of rodent 50k neuronal culture energy comparisons across generative rules.

Table C3. Statistical comparisons of rodent 100k neuronal culture energy comparisons across generative rules.

Table C4. Statistical comparisons of rodent 50k neuronal culture topological fingerprint dissimilarity comparisons across generative rules.

Table C5. Statistical comparisons of rodent 100k neuronal culture topological fingerprint dissimilarity comparisons across generative rules.

Table C6. Statistical comparisons of DIV28 human iPSC neuronal cultures (glutamatergic neurons, motor neurons and dopaminergic neurons) and human cerebral organoid energy comparisons across generative rules.

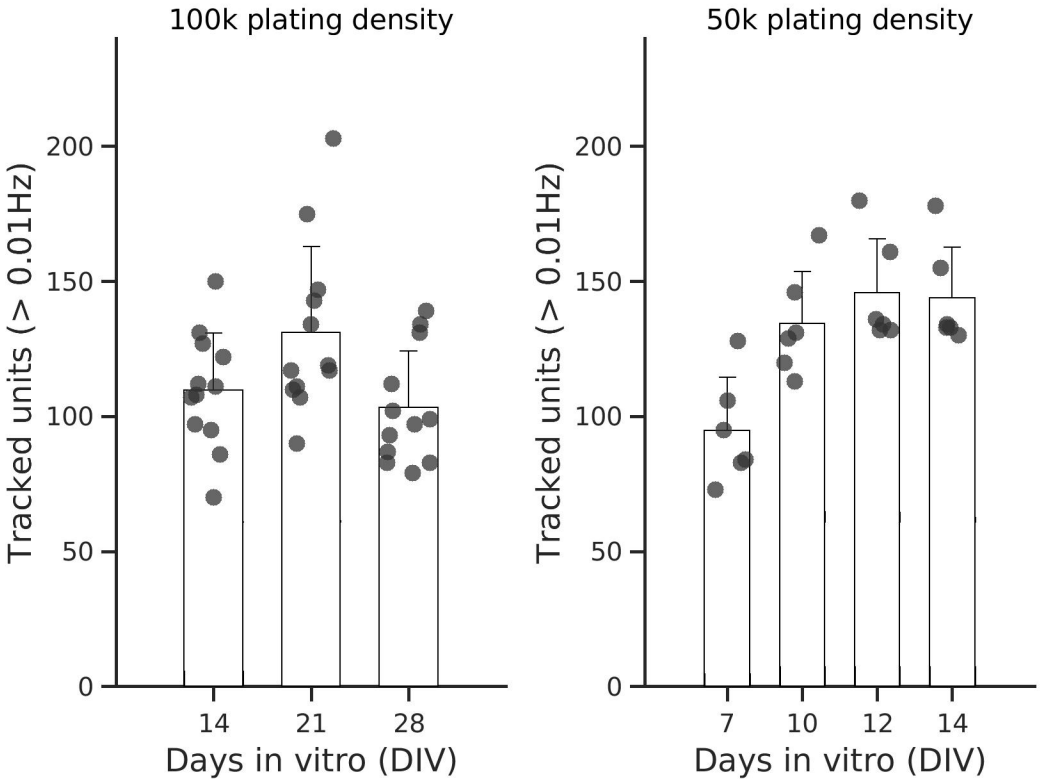


Figure C1. Tracking neuronal networks at cellular resolution on high-density microelectrode arrays. Bar plots depicting the tracking result for the rodent data with 100k (left) and 50k (right) plating density per high-density microelectrode array (HD-MEA).

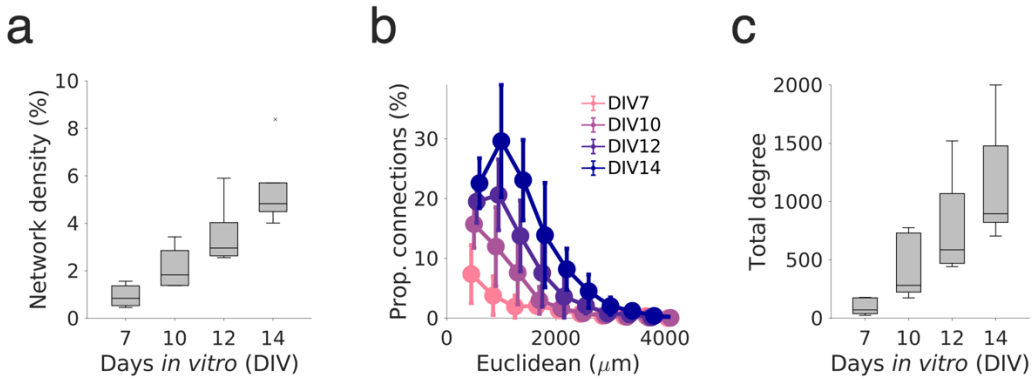


Figure C2. Global topological measures of sparse rodent cultures over time. (a) Panel depicts the statistically inferred network density for the sparse (50,000 cell per well) primary cortical (PC) neuronal networks until days in vitro (DIV)14. (b) Proportion of extant connection by distance and grouped by DIV7 (pink), DIV10 (light purple), DIV12 (dark purple) and DIV14 (dark blue). (c) The total degree for sparse PC networks across development.

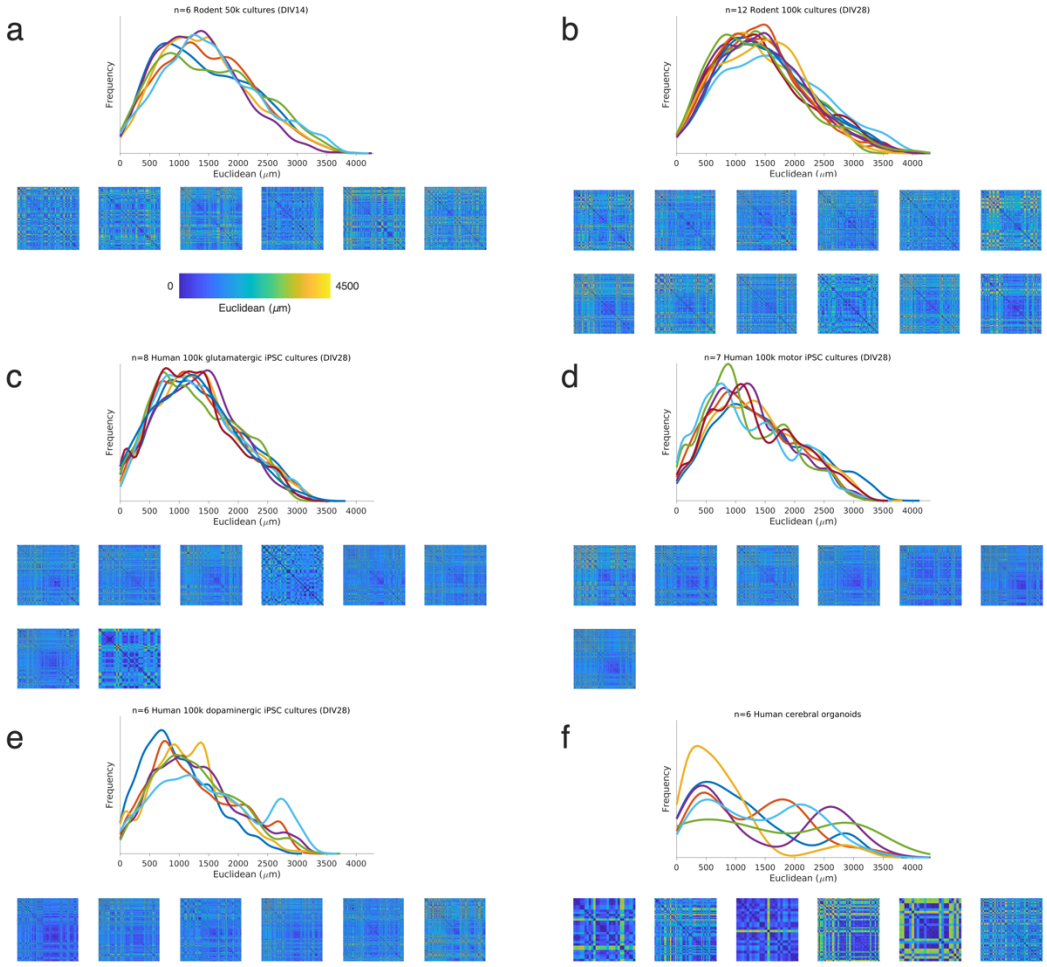


Figure C3. Distribution of inter-neuronal Euclidean distances across rodent and human datasets. (a) Density plot of inter-neuronal Euclidean distances for each culture. Note, each matrix was used as the $D_{i,j}$ term for the generative model constructed for each time point of that culture. Panel a depicts Euclidean distances for the sparse PC rodent networks (50,000 neurons per well); the left panel shows the overall distributions; the right panel shows the individual distance matrices. (b) Euclidean distances for the dense PC rodent networks (100,000 neurons per well). Panels (c-e) show Euclidean distance distribution across the iPSC-derived human neuron lines at DIV28 (c, glutamatergic; d, motor and e, dopaminergic neurons). (f) Euclidean distances for the human cerebral organoid recordings.

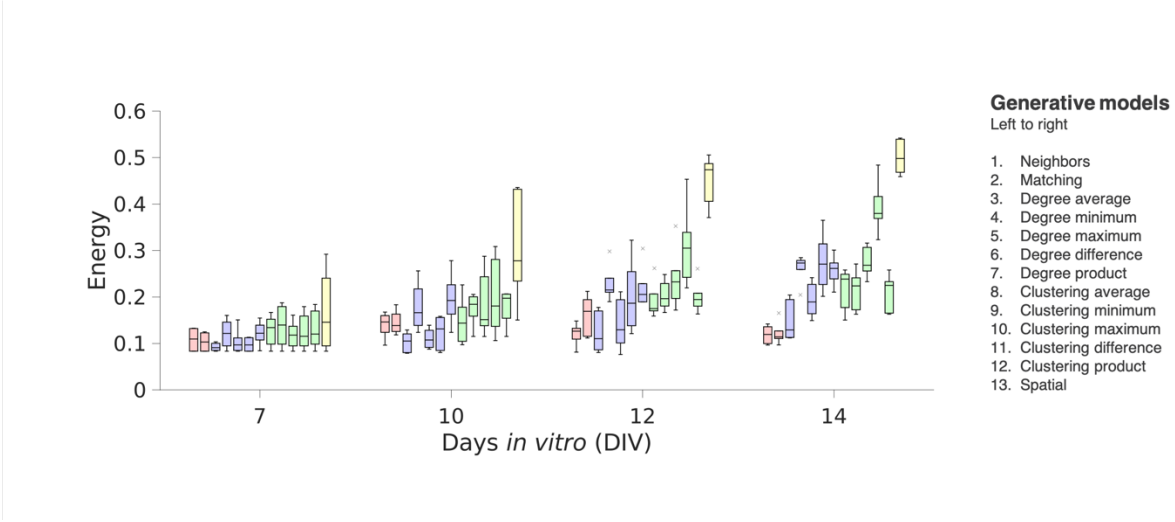


Figure C4. Generative network modelling results for the sparse rodent primary cortical cultures. Generative model fits (energy), including all 13 wiring models, for the sparse primary rodent networks (50,000 cells per well; n=6 cultures) across development. Each boxplot presents the median and IQR. Outliers are demarcated as small black crosses, and are those which exceed 1.5x the interquartile range away from the top or bottom of the box. Generative model performance over time according to the energy equation. In each box, the energy of the top n=1 performing simulations are shown.

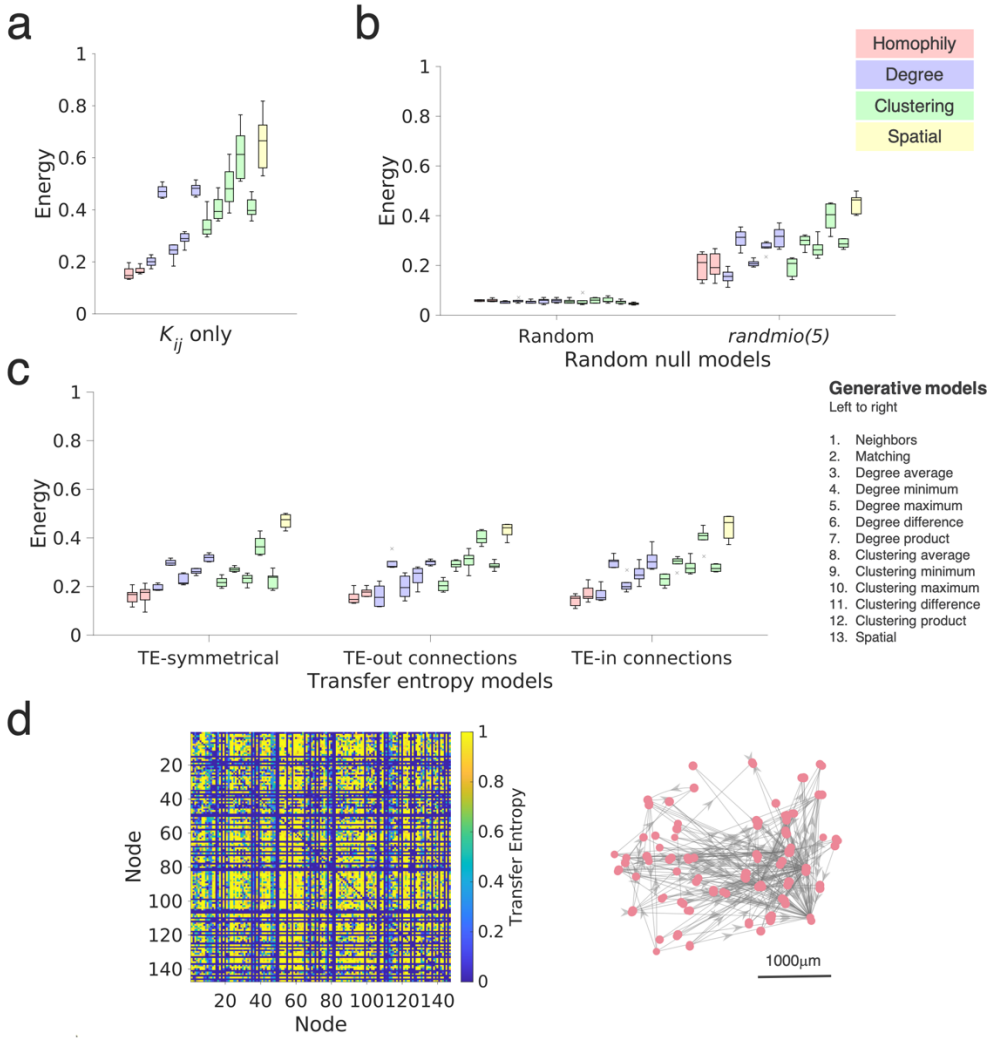


Figure C5. Alternative generative model procedures. Generative model fits (energy), including all 13 wiring models, for 50k-plated primary rodent networks (50,000 cells per well; n=6 cultures) at DIV14. (a) Generative models fit only using the wiring term ($K_{i,j}$) when forming networks. (b) Random null models. Completely randomised networks (size- and density-matched, left) and networks which have been partially randomised via the *randmio_und()* function, using the Maslov-Sneppen algorithm, which rewires each edge approximately five times. (c) Inferred transfer entropy (TE) networks. As the TE is a directed graph, we show the results for the symmetrised network (left); out-connections (middle) and in-connections (right). For each graph, the boxplot presents the median and IQR. Outliers are demarcated as small black crosses, and are those which exceed 1.5x the interquartile range away from the top or bottom of the box. Generative model performance over time according to the energy equation. In each box, the energy of top n=1 performing simulation is shown. (d) A visualisation of the transfer entropy matrix (left) of a single empirical network (right).

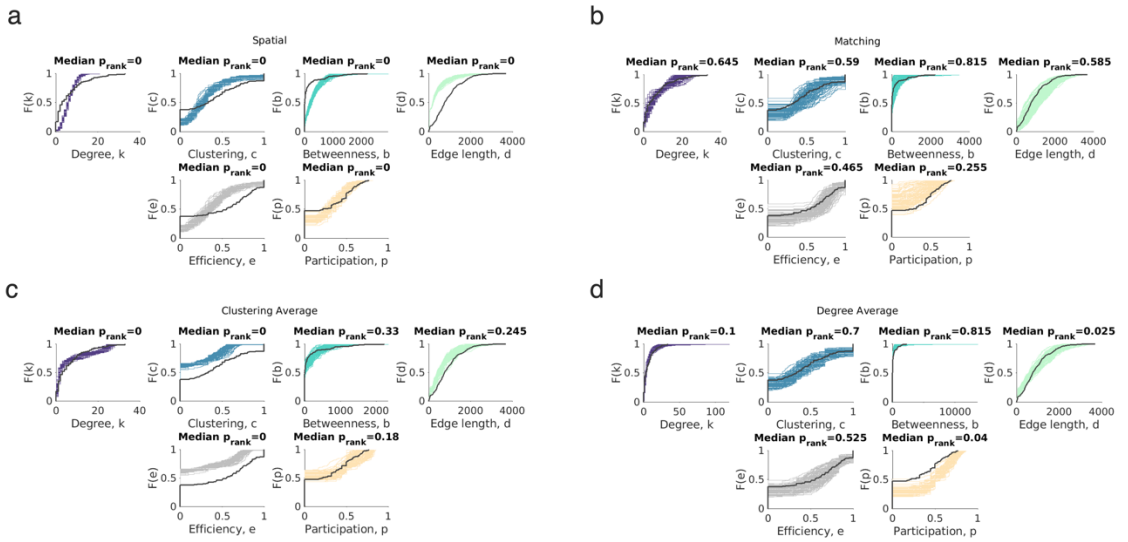


Figure C6. Generative model fits recapitulate observed network statistics. a Cumulative density functions (CDFs) of the top$p$=99 simulations (of the total 20,000 simulations) for the best performing generative models in each class – (a) Spatial. (b) Matching. (c) Clustering average. (d) Degree average – across the four statistics included in the energy equation (top four panels) and two additional metrics - the local efficiency and participation coefficient - not included in the energy equation (bottom two panels). For visualisation, we show only the solutions for a single sparse rodent PC culture. *P*-values were computed using the Monte-Carlo bootstrapping procedure.

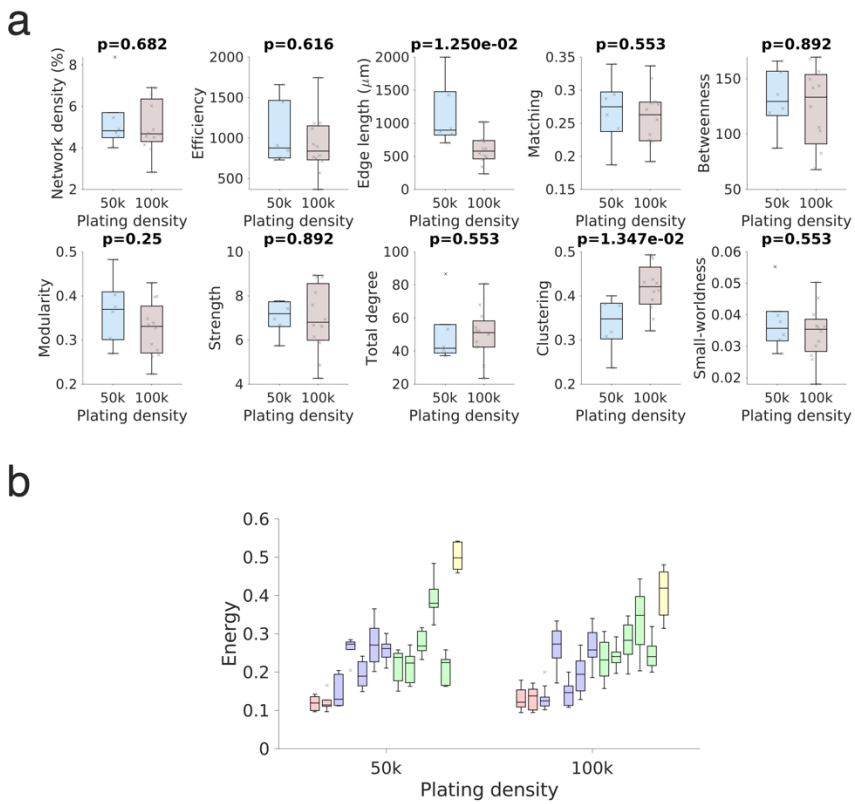


Figure C7. Comparison of global network statistics across sparse and dense primary cortical rodent networks at DIV14. (a) Comparisons were computed for network density, efficiency, edge length, matching, betweenness, modularity, strength, total degree, clustering and small worldness. For each comparison we provide the p-value computed from a Mann-Whitney U test. Sparse networks were plated at 50,000 cells per well, dense networks were plated at 100,000 cells per well. (b) Direct comparisons of the generative model fits (energy), including all 13 wiring models, for the 50k-plated (left) versus 100k-plated (right) primary rodent networks at DIV14. The boxplot presents the median and IQR. Outliers are demarcated as small black crosses, and are those which exceed 1.5x the interquartile range away from the top or bottom of the box. Generative model performance over time according to the energy equation. In each box, the energy of top n=1 performing simulation is shown.

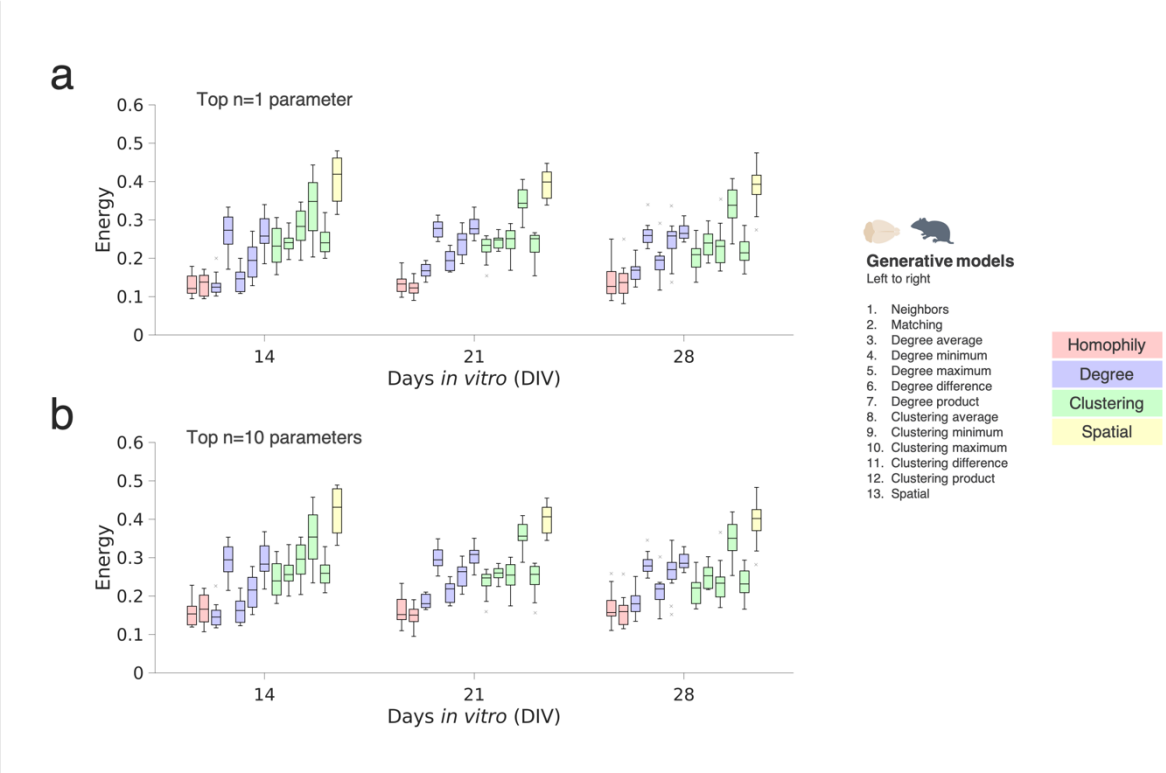


Figure C8. Generative network modelling results for the dense rodent primary cortical cultures. Generative model fits (energy), including all 13 wiring models, for the dense primary rodent networks (100,000 cells per well; n=12 cultures) across development. (a) Top performing n=1 parameter combination. (b) Top performing n=10 parameter combinations. Each boxplot presents the median and IQR. Outliers are demarcated as small black crosses, and are those which exceed 1.5x the interquartile range away from the top or bottom of the box. Generative model performance over time according to the energy equation. In each box, the energy of the top n=1 performing simulations are shown.

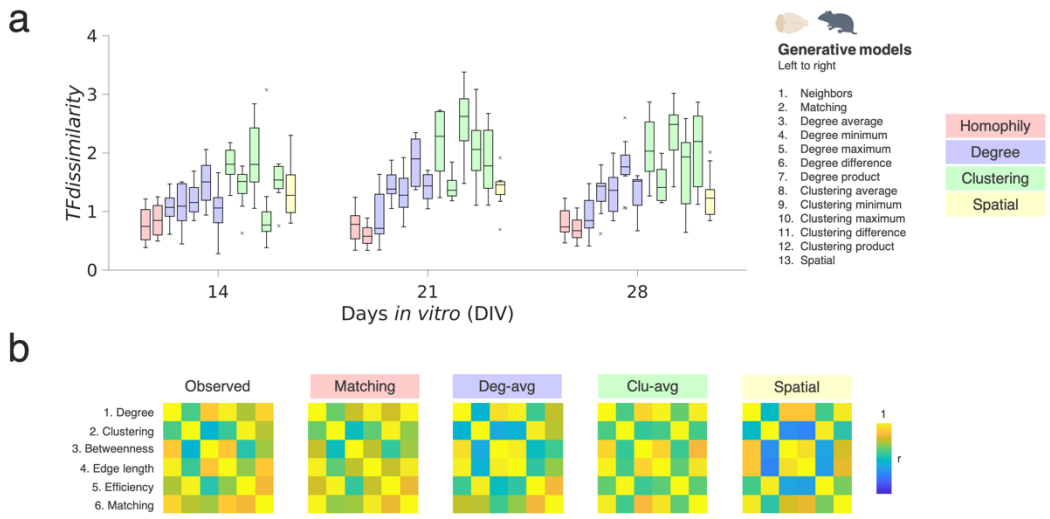


Figure C9. Homophilic generative mechanisms best account for local relationships in developing dense rodent neuronal cultures. (a) Homophily generative models produce the lowest $TF_{dissimilarity}$ across all time-points (DIV14, 21 and 28), suggesting that it can reconstruct local connectivity patterns of in vitro neuronal networks. In total, there are n=12 data points (one per culture) shown in each of the 13 boxplots. Boxplots presents the median and IQR. Outliers are demarcated as small grey crosses, and are those which exceed 1.5 times the interquartile range away from the top or bottom of the box. (b) A visualisation of the averaged topological organisation matrix for the observed (left), matching (middle left), degree-average (middle), clustering-average (middle right), and spatial (right) models at DIV28.

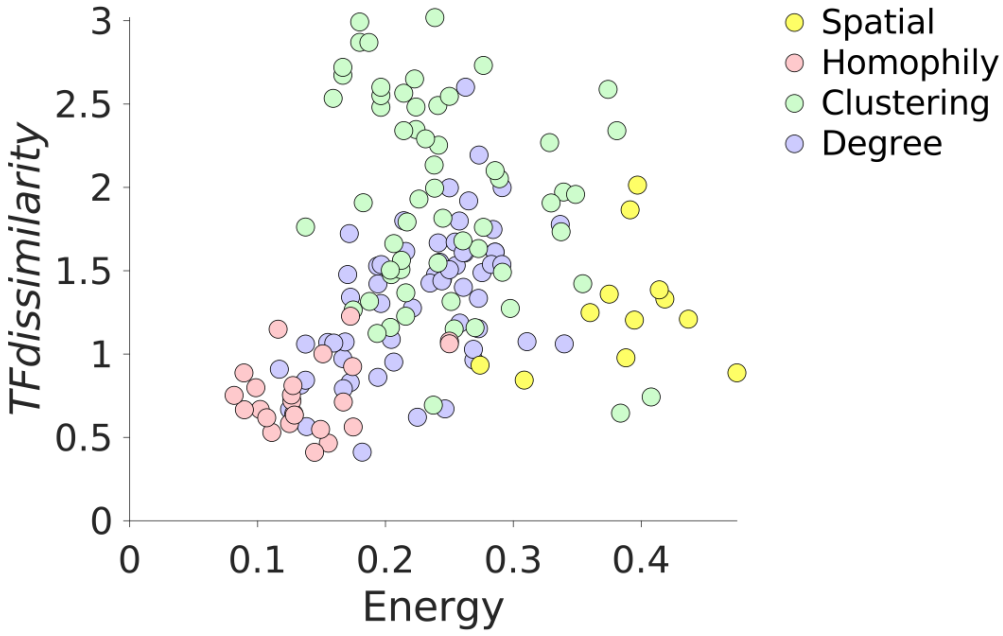


Figure C10. Relationship between model energy and topological fingerprint dissimilarity in developing dense rodent neuronal cultures. Using dense PC rodent network (100,000 cells per MEA), we show n=78 data points (n=6 cultures x n=13 generative model simulations) corresponding to the top n=1 performing simulation's energy and its topological fingerprint dissimilarity performances at DIV14. Distributions are plotted for each, showing homophily to achieve the best fits in both (bottom left).

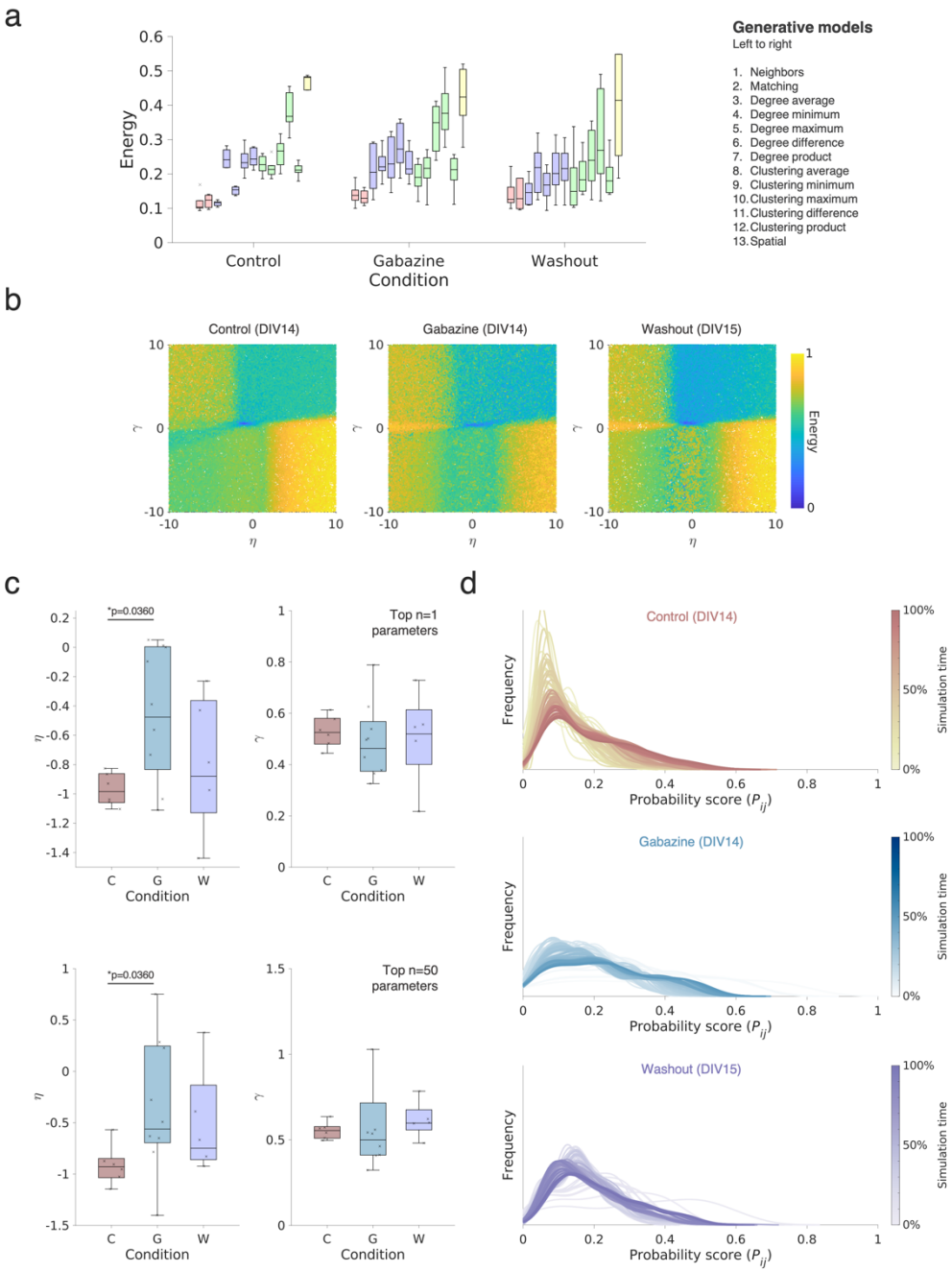


Figure C11. Generative model comparisons between control, gabazine-treated and washout sparse rodent PC cultures. (a) Direct comparisons of the generative model fits (energy), including all 13 wiring models, for control (left) versus gabazine (middle) and washout (right) cultures. The boxplot presents the median and IQR. Outliers are demarcated as small black crosses, and are those which exceed 1.5x the interquartile range away from the top or bottom of the box. Generative model performance over time according to the energy equation. In each box, the energy of top n=1 performing simulation is shown. (b) The energy landscapes of the

matching generative rule for control (left) versus gabazine (middle) and washout (right) cultures. (c) Wiring parameters of control (left; coloured in red), gabazine (middle; coloured in blue) and washout (right; coloured in purple) derived from the top n=1 (top) and n=50 (bottom) performing matching simulation in terms of the wiring equation are shown. (d) The probability distributions for control (top), gabazine (middle) and washout (bottom) networks were computed and scaled at 1% increments throughout the developmental course of the simulations. These distributions were averaged over time to form the mean probability distributions shown in Figure 3.5c.

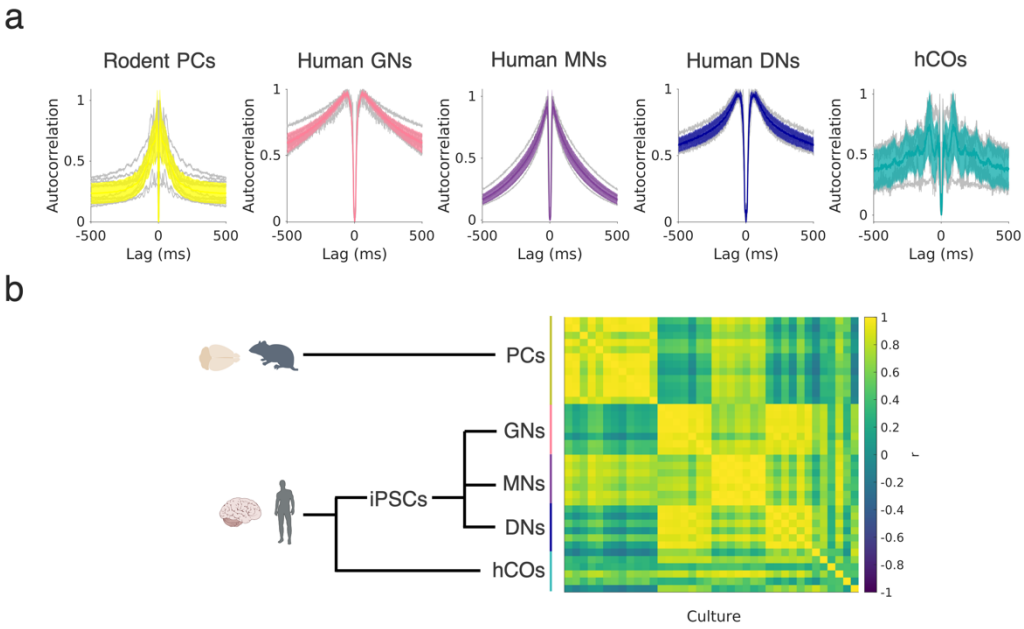


Figure C12. Spiking dynamics in human monolayer and cerebral organoid cultures compared to rodent cultures. (a) For each culture type, autocorrelograms of spiking activity are plotted. Autocorrelograms were constructed by first cumulatively concatenating spike times across all units in a recording into a single vector of spike times. The count of synchronous spikes across the network (in 1ms bins) was quantified across a range of time lags (0 lag was excluded). This indicates the degree of coactivation at various latencies. In all cell culture types, there was much higher spiking synchrony at shorter latencies indicated by peaks close to 0ms lag. However, rodent primary cultures (PCs) and human motor neurons (MNs) showed sharp declines in autocorrelation at longer latencies indicating a strong preference for short latency, high frequency coactivation. Human glutamatergic (GNs) and dopaminergic (DNs) cultures had much greater autocorrelation at lag values further from 0. This indicates coactivation at longer latencies and oscillatory network activity at slower frequencies. Human cerebral organoids (hCOs) were much more variable. (b) The autocorrelation vectors (autocorrelation at each lag value) across all cultures were correlated and a high level of concordance within culture sources in terms of latency preference can be seen, with moderate correlations between rodent PCs and human MNs.

Antibodies (human organoids)	**Type**	**Dilution**	**Catalog number**
Tau	Primary	1:500	#MN1000, ThermoFisher
NeuN	Primary	1:300	#M11954-3, Boster Bio, Pleasanton, CA, USA
GFAP	Primary	1:500	#NB300-141, Novus Biologicals, Englewood, CO, USA
goat anti-mouse IgG, Alexa Fluor Plus 488	Secondary	1:400	#A32723, ThermoFisher
goat anti-rabbit IgG, Alexa Fluor 568	Secondary	1:400	#A11036, ThermoFisher
goat anti-chicken IgY, Alexa Fluor Plus 647	Secondary	1:400	#A32933, ThermoFisher
Antibodies (human iPSC)	Type	Dilution	Catalog number
mouse anti-TH	Primary	1:500	#MAB318, Sigma-Aldrich
chicken anti-MAP2	Primary	1:1000	#CH22103, Neuromics (Edina, MN, USA)
rabbit anti-GFAP	Primary	1:500	#Z0334, Agilent (Santa Clara, CA, USA)
donkey anti-mouse 488	Secondary	1:250	#A-21202, ThermoFisher
goat anti-chicken 647	Secondary	1:500	#A-32933, ThermoFisher
donkey anti-rabbit 568	Secondary	1:250	#A-A10042, ThermoFisher

Antibodies (rodent PC)	Type	Dilution	Catalog number
mouse anti-Synaptophysin	Primary	1:100	#ab8049, Abcam
rabbit anti-NeuN	Primary	1:300	#ab177487, Abcam
chicken anti-beta III Tubulin	Primary	1:1000	#ab41489 Abcam
donkey anti-mouse IgG, Alexa Fluor 488	Secondary	1:500	#ab150105, Abcam
donkey anti-Chicken IgY (IgG)	Secondary	1:400	#703–165-155, Jackson ImmunoResearch, West Grove, USA
donkey anti-rabbit IgG, Alexa Fluor 405	Secondary	1:1000	#ab175651, Abcam

Table C1. Primary and secondary antibodies.

Rule A	Rule B	DIV 7		DIV10		DIV12		DIV14	
		p=1.34x10^{-3}	d	p=1.51x10^{-7}	d	p=3.43x10^{-15}	d	p=3.26x10^{-17}	d
Homophily	Degree	1	0.0482	1	-0.00443	0.222	0.583	4.11x10^{-5}	1.46
Homophily	Clustering	0.286	0.703	0.29	0.704	0.000528	1.52	7.01x10^{-8}	2.94
Homophily	Spatial	0.00592	0.926	1.47x10^{-6}	1.38	3.77x10^{-9}	4.83	3.77x10^{-9}	12
Degree	Clustering	0.127	0.6	0.0909	0.585	0.0242	0.646	0.16	0.42
Degree	Spatial	0.00219	0.675	9.12x10^{-8}	1.5	3.77x10^{-9}	6.3	3.77x10^{-9}	15.3
Clustering	Spatial	0.0802	0.496	2.25x10^{-5}	1.71	3.81x10^{-9}	4.44	3.80x10^{-9}	3.99

Table C2. Statistical comparisons of rodent 50k neuronal culture energy comparisons across generative rules. For each test, we quote the ANOVA p-value across the generative rules and the corresponding Cohen's d if the ANOVA was significant at $p<0.05$. A positive Cohen's d reflects that Rule A has a smaller energy than Rule B, reflecting a better fit. Generative rules have been binned across the generative model class.

Rule A	Rule B	DIV14		DIV21		DIV28	
		p=7.55x10^{-27}	d	p=7.04x10^{-32}	d	p=1.27x10^{-23}	d
Homophily	Degree	1.53x10^{-5}	0.843	3.77x10^{-9}	1.52	4.42x10^{-9}	1.19
Homophily	Clustering	3.77x10^{-9}	2.37	3.77x10^{-9}	3.05	3.77x10^{-9}	1.44
Homophily	Spatial	3.77x10^{-9}	3.33	3.77x10^{-9}	8	3.77x10^{-9}	3.10
Degree	Clustering	3.43x10^{-8}	0.745	0.0146	0.397	0.257	0.281
Degree	Spatial	3.77x10^{-9}	2.8	3.77x10^{-9}	4.14	3.77x10^{-9}	3.59
Clustering	Spatial	3.78x10^{-9}	3.92	3.77x10^{-9}	2.32	3.77x10^{-9}	2.40

Table C3. Statistical comparisons of rodent 100k neuronal culture energy comparisons across generative rules. For each test, we quote the ANOVA p-value across the generative rules and the corresponding Cohen's d if the ANOVA was significant at $p<0.05$. A positive Cohen's d reflects that Rule A has a smaller energy than Rule B, reflecting a better fit. Generative rules have been binned across the generative model class.

Rule A	Rule B	DIV7		DIV10		DIV12		DIV14	
		p=0.0847	d	p=2.6x10^{-9}	d	p=1.81x10^{-5}	d	p=3.9x10^{-7}	d
Homophily	Degree	0.995	-0.084	0.725	-0.272	0.0762	1.15	0.00191	1.34
Homophily	Clustering	0.634	0.244	0.000167	2.39	6.35x10^{-5}	3.58	1.37x10^{-7}	2.59
Homophily	Spatial	0.356	-4.84	1	0.00263	0.000996	1.39	0.0952	0.818
Degree	Clustering	0.328	0.491	5.55x10^{-9}	1.55	0.0189	0.544	0.00749	0.726
Degree	Spatial	0.383	-4.16	0.851	0.704	0.0615	0.733	0.995	-0.0648
Clustering	Spatial	0.0943	-4	0.006	-0.967	0.849	0.291	0.147	-0.515

Table C4. Statistical comparisons of rodent 50k neuronal culture topological fingerprint dissimilarity comparisons across generative rules. For each test, we quote the ANOVA p-value across the generative rules and the corresponding Cohen's d if the ANOVA was significant at $p<0.05$. A positive Cohen's d reflects that Rule A has a smaller energy than Rule B, reflecting a better fit. Generative rules have been binned across the generative model class.

Rule A	Rule B	DIV14		DIV21		DIV28	
		p=1.01x10^{-8}	d	p=1.6x10^{-21}	d	p=2.09x10^{-18}	d
Homophily	Degree	0.00492	0.996	7.87x10^{-9}	1.27	3.89x10^{-6}	1.19
Homophily	Clustering	4.37x10^{-9}	1.79	3.77x10^{-9}	2.87	3.77x10^{-9}	2.76
Homophily	Spatial	0.00332	0.869	0.000125	1.58	0.0138	1.02
Degree	Clustering	0.000195	0.519	3.78x10^{-9}	0.823	3.84x10^{-9}	0.759
Degree	Spatial	0.562	0.428	1	0.0135	0.973	-0.151
Clustering	Spatial	0.693	-0.232	0.000357	-1.02	8.48x10^{-5}	-0.921

Table C5. Statistical comparisons of rodent 100k neuronal culture topological fingerprint dissimilarity comparisons across generative rules. For each test, we quote the ANOVA p-value across the generative rules and the corresponding Cohen's d if the ANOVA was significant at $p<0.05$. A positive Cohen's d reflects that Rule A has a smaller energy than Rule B, reflecting a better fit. Generative rules have been binned across the generative model class.

Rule A	Rule B	Glutamatergic neurons		Motor neurons		Dopaminergic neurons		Cerebral organoids	
		p=2.03x10^{-8}	d	p=4.79x10^{-14}	d	p=2.93x10^{-7}	d	p=2.65x10^{-4}	d
Homophily	Degree	0.00652	0.772	0.0037	0.812	0.866	0.178	0.971	-0.141
Homophily	Clustering	0.000181	1.8	0.00119	1.88	0.054	1.34	0.846	-0.303
Homophily	Spatial	1.22x10^{-8}	2.07	3.77x10^{-9}	10	9.16x10^{-7}	4.34	0.00399	1.82
Degree	Clustering	0.525	0.248	0.969	0.0735	0.0838	0.377	0.959	-0.099
Degree	Spatial	2.05x10^{-5}	2.25	3.78x10^{-9}	8.34	6.27x10^{-7}	4.29	0.000321	2.5
Clustering	Spatial	0.000437	1.83	3.82x10^{-9}	3.96	0.000143	5.48	0.000113	1.57

Table C6. Statistical comparisons of DIV28 human iPSC neuronal cultures (glutamatergic neurons, motor neurons and dopaminergic neurons) and human cerebral organoid energy comparisons across generative rules. For each test, we quote the ANOVA p-value across the generative rules and the corresponding Cohen's d if the ANOVA was significant at $p<0.05$. A positive Cohen's d reflects that Rule A has a smaller energy than Rule B, reflecting a better fit. Generative rules have been binned across the generative model class.

9. APPENDIX D SUPPLEMENTARY MATERIALS FOR CHAPTER 4

Content included in this chapter was published in bioRxiv, 2022, and is currently in review at Nature Machine Intelligence: Spatially-embedded recurrent neural networks reveal widespread links between structural and functional neuroscience findings. Jascha Achterberg, **Danyal Akarca***, DJ Strouse, John Duncan† & Duncan E. Astle† (*Co-lead authors, †Co-senior authors).*

Figure D1. An example matrix representation of a simple graph.

Figure D2. Computing the possible routing paths between two nodes.

Figure D3. Computing the network communicability as a function of the weighted sum of all possible paths.

Figure D4. Weighted communicability and its normalisation.

Figure D5. Comparing communicability terms.

Figure D6. Changing individual weights in a network to determine the effect on the total weighted network communicability.

Figure D7. Generative modelling findings in L1-reguarlised networks.

Figure D8. Replication of spatial-clustering using a weighted variation of the spatial permutation statistical procedure.

Figure D9. Intertwined global structural-functional relationships in seRNNs.

According to graph theory, brain networks can be described as graphs that are composed of nodes (vertices) denoting a neural element (e.g., neuronal unit in an RNN or neuron) that are linked by edges representing their connections (e.g., connection in a RNN or synapse). Beyond the graph representation, it can equally be represented as a simple adjacency matrix, A (**Figure D1**). In this example, we show a symmetrical binary graph because there are no directions in the connections and the edge has no weight – note, this is the same schematic shown in **Figure 4.1**.

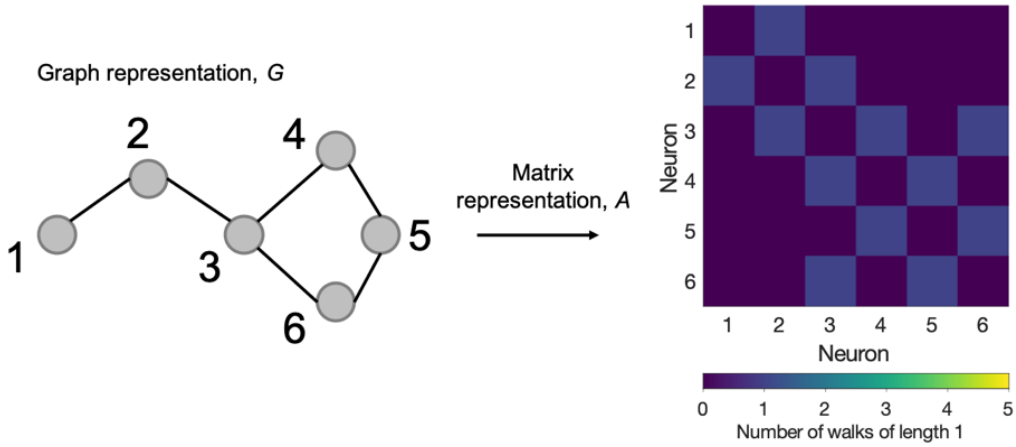


Figure D1. An example matrix representation of a simple graph.

There is an interesting property of graphs like this, where if you raise A to the power of k (i.e., A^k) the result tells you how many walks of length k are possible between any two nodes (by walks we mean simply how many "jumps" can be taken along the graph from one node to the next). For example: A^2 tells you how many walks of length *2* exist between any two nodes, and A^3 tells you how many walks of length *3* exist between any two nodes (**Figure D2**). You can check this by taking any two nodes (e.g., node 5 and node 3) and count how many times you can walk a number of paths (e.g., 2 paths). In this example the answer is 2: the path can go from 5→4→3 or 5→6→3. This is reflected in the matrix $A^2_{5,3}$ which is equal to 2.

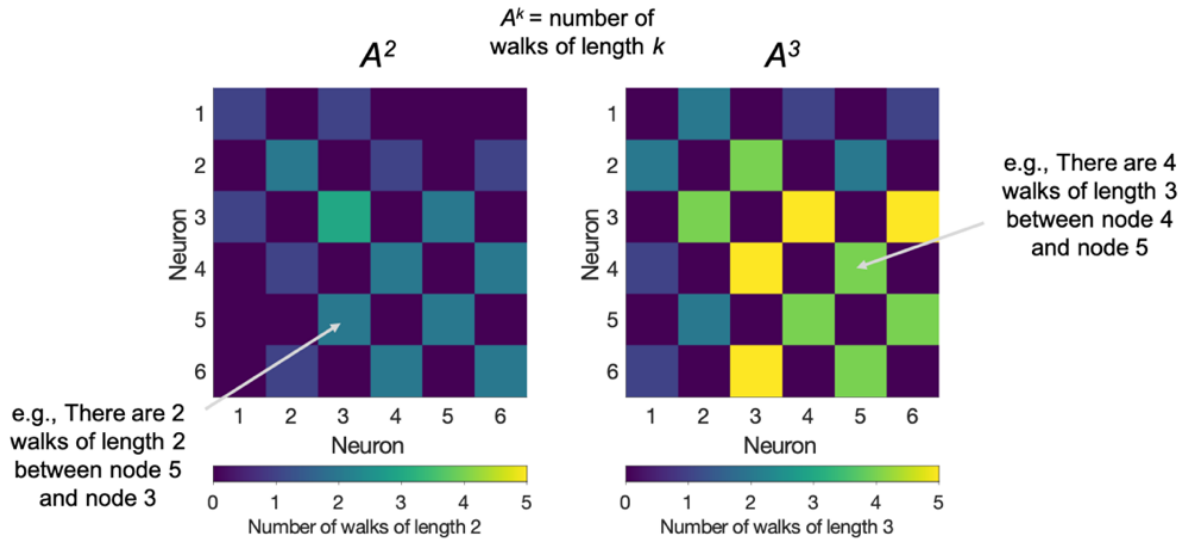


Figure D2. Computing the possible routing paths between nodes.

Why is this relevant? Because if you want to understand the extent to which two regions in a network can *communicate* with each other, you need to model the extent to which information can be passed between nodes on the graph. To calculate this, you can simply take a *weighted sum of all possible walks between two nodes*. Note that it must be a weighted sum because as you extend the number of walks between two regions, it is less likely that any one journey was taken by that route. Given our previous example, you can now see that this would look like this:

$$\sum_{k=0}^{\infty} \frac{A^k}{k!} = \frac{A^0}{0!} + \frac{A^1}{1!} + \frac{A^2}{2!} + \frac{A^3}{3!} + \frac{A^4}{4!} + \frac{A^5}{5!} + \frac{A^6}{6!} + \frac{A^7}{7!} + \cdots \quad (27)$$

This expression is an infinite sum. This expression is equivalent to the Maclaurin series (Taylor series, centred at zero) of e^A:

$$\sum_{k=0}^{\infty} \frac{A^k}{k!} = e^A \quad (28)$$

An illustration of this can be seen in **Figure D3**:

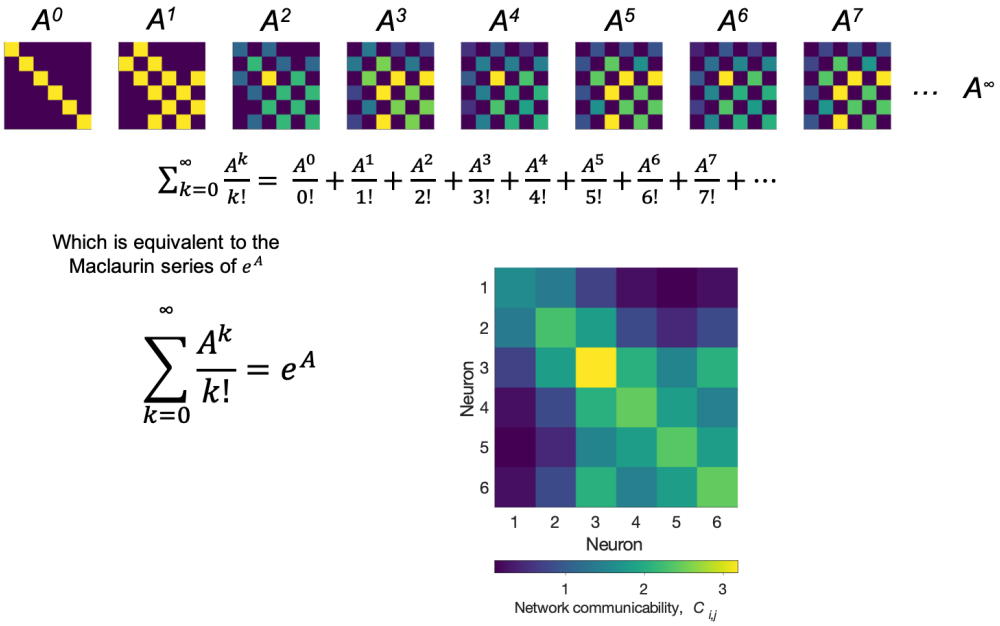


Figure D3. Computing the network communicability as the weighted sum of all possible walks.

Note that A here is the same original matrix we had at the beginning. So, if you raise e to the power of the A matrix, you are computing the weighted sum all possible walks between two regions – and this is what we call the *network communicability, C*. This type of communication actually resembles random diffusion or random-walks across the graph: mirroring what is a likely constraint within biological neural networks – signals propagating through the network and reaching nodes as a function of how communicable they are, based on the underlying connectivity.

The same logic can easily be applied to weighted graphs. The only difference now is that the weights also modulate the extent to which nodes communicate with each other. However, this introduces a slight problem: this equation has the property of biasing walks very heavily to single edges which may be larger – leading to undue influence on the calculation. Crofts & Higham (2009) proposed a solution where you add a normalisation step which controls for each node's strength, leaving the following expression (for an example, see **Figure D4**):

$$C = e^{S^{-\frac{1}{2}}WS^{-\frac{1}{2}}}, \tag{29}$$

Where S is a diagonal matrix denoting the node's strength. This expression is used as a constraint to learning within the seRNN framework.

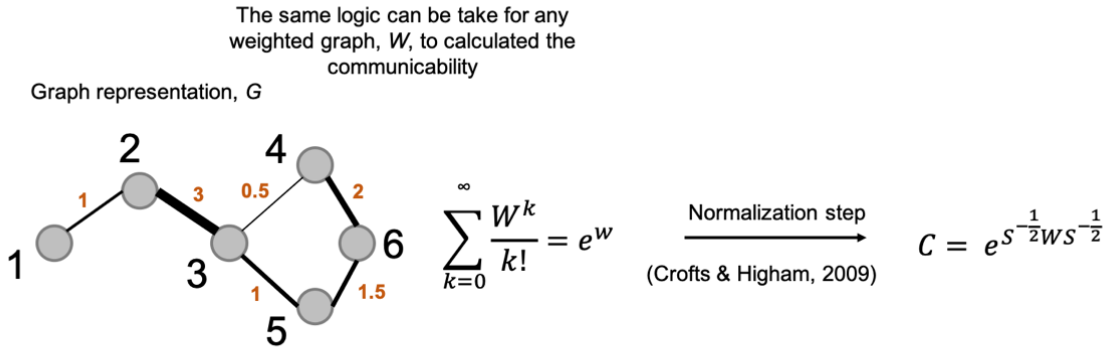


Figure D4. Weighted communicability and its normalisation.

This normalisation step has the effect of increasing the resolution of the communicability metric as shown in **Figure D5** (for more detail, refer to Crofts and Higham, 2009). Whereas in the unweighted case highly communicative connection (e.g., between nodes 2 and 3) dominate the communicability matrix, the weighted communicability is more sensitive to individual edges' contributions to the network communicability.

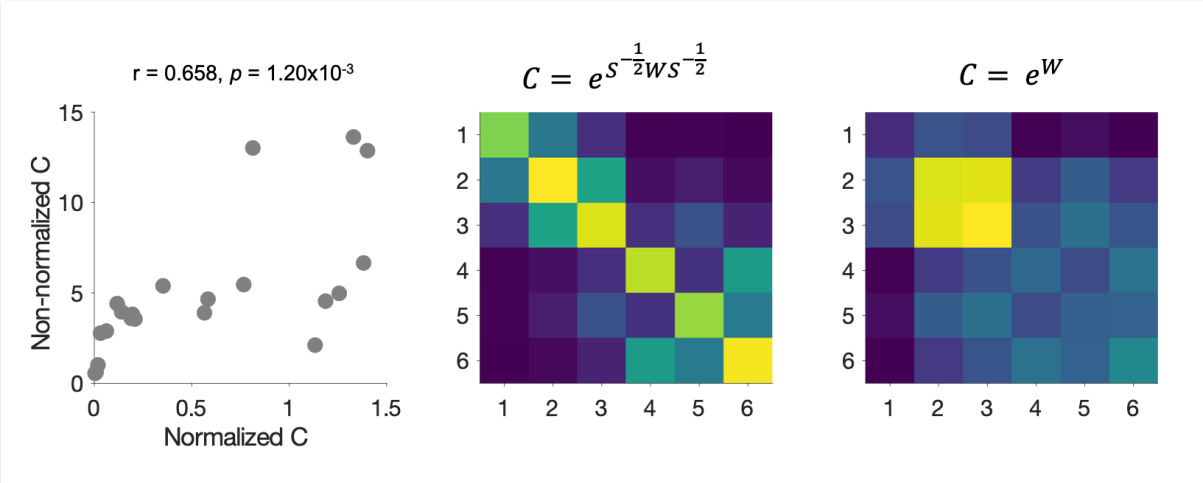


Figure D5. Comparing communicability terms. The left shows that the normalised and non-normalised communicability are positively related to each other. However, from the middle plot you can observe that communication is shared more evenly across the network while on the right, the non-normalised communication becomes bottle-necked within a single edge.

The weighted communicability has a desirable effect on the network optimisation process as shown in show in the experiment in **Figure D6**. In this experiment, individual weights within the example network (top) were increased or decreased in increments of 5% of their value, and the total communicability (C) was calculated at each iteration. As shown in the scatter plot (bottom), as weights are increased (i.e., δw_{ij} is positive – to the right) edges which have high communication due to both their high weight and position within the network (e.g., the orange edge, $w_{\{2,3\}}$) contribute towards the network becoming less communicable in total. This is because communication now gets biased towards this edge and away from other edges. In contrast, as weights are decreased (i.e., δ_{ij} is negative – to the left), the opposite occurs; edges which have low communication (e.g., the dark blue edge, $w_{\{1,2\}}$) now contribute to the network becomes less communicable in total. This means that when seRNNs are optimised for their weighted communicability (i.e., the direction of the optimisation being in the down direction) that highly communicable edges are strengthened and lowly communicable edges are weakened.

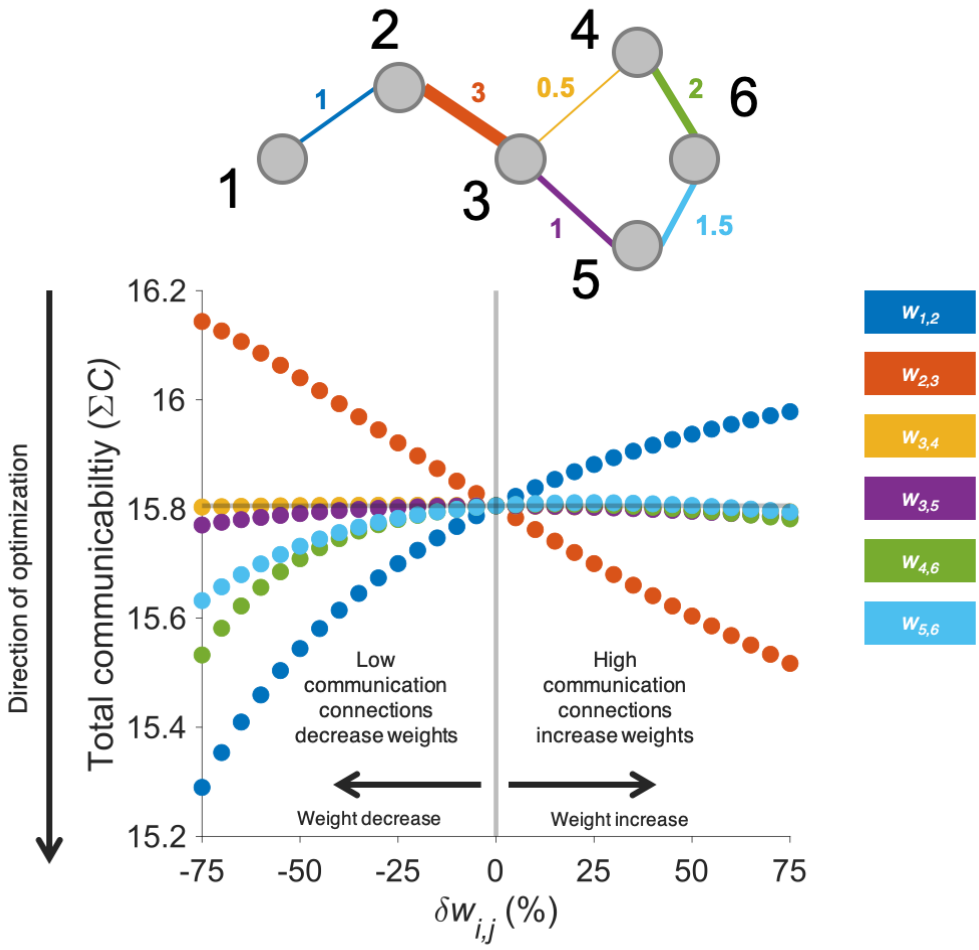


Figure D6. Exploring the effect of changing individual network weights on the total weighted network communicability.

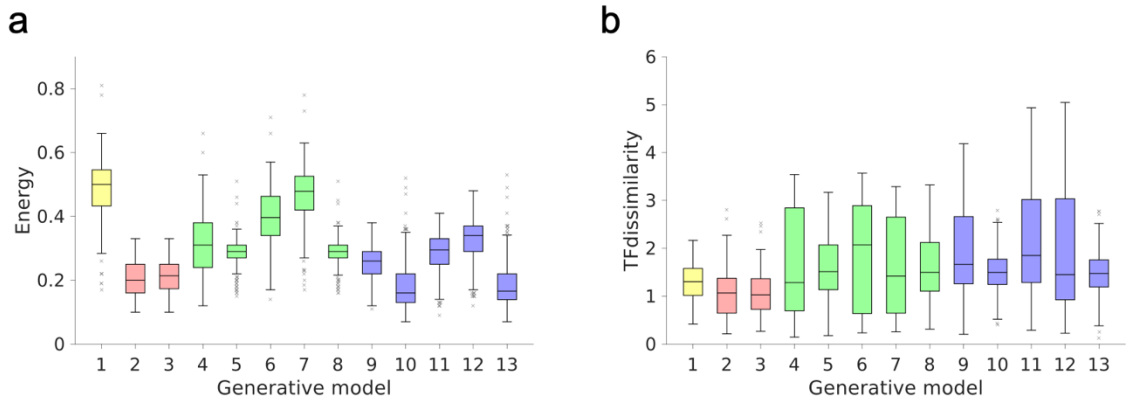


Figure D7. Generative modelling findings in L1-reguarlised networks. (a) The energy attained for L1 networks shows that there is no single generative network model that generates the lowest energy. The top n=1 solutions are shown across the 455 functioning L1 in the evaluated range. (b) The $TF_{dissimilarity}$ attained for L1 networks shows that the topological fingerprint can be approximated by spatial and homophily generative network models. The top n=1 solutions are shown across the 455 functioning L1 in the evaluated range.

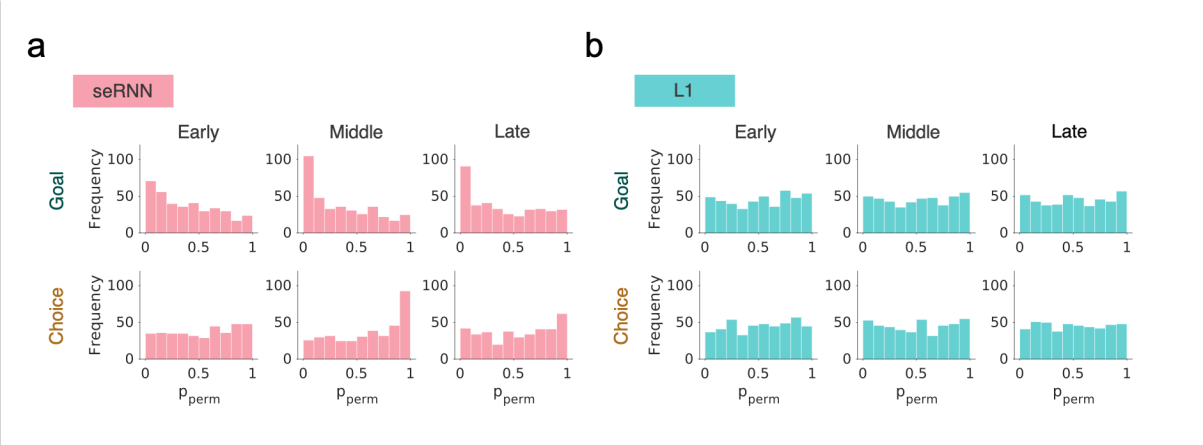


Figure D8. Replication of spatial clustering using a weighted variation of the spatial permutation statistical procedure. (a) seRNNs. (b) L1 networks. Terminology and conventions as Figure 4.4.

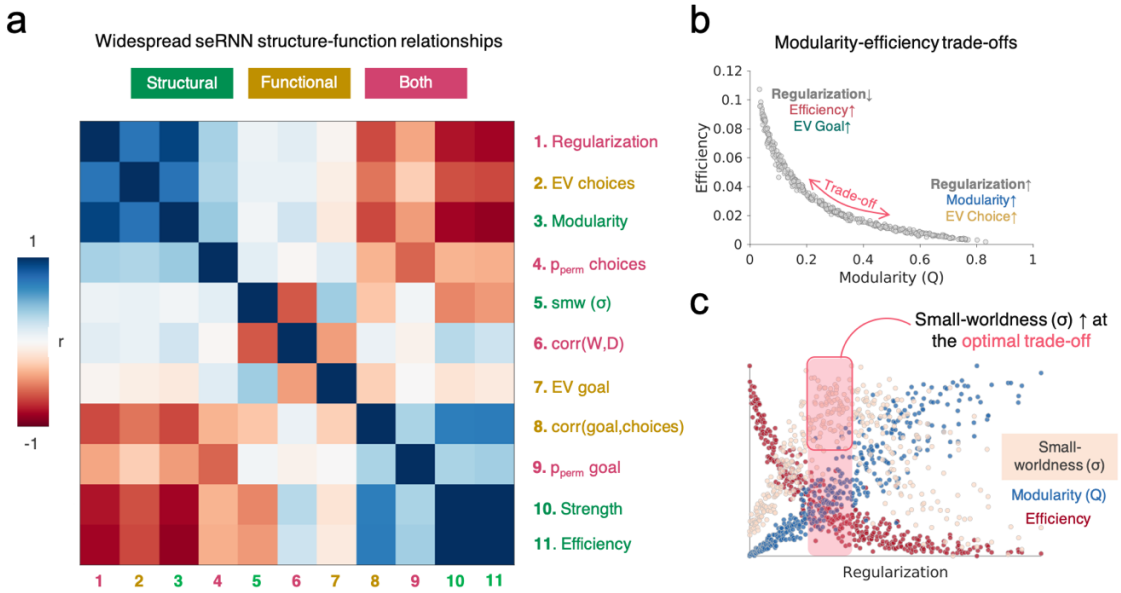


Figure D9. Intertwined global structural-functional relationships in seRNNs. (a) The correlation matrix of structural (green), functional (yellow) and both (pink) measures in the seRNNs shows wide-spread relationships that cluster. (b) A scatter plot showing a negative relationship between global efficiency versus modularity across seRNNs. Each plot in the scatter is a seRNN of a different regularisation strength. seRNNs that have not been regularised as much tend to have higher global efficiency and less modularity, and sit to the left of the curve. These networks tend to have a greater explained variance for the goal. In contrast, seRNNs that have been regularised more tend to have lower global efficiency and higher modularity, and therefore sit to the right of the curve. These networks tend to have a greater explained variance for the choice. (c) A scatter plot of regularisation (x-axis) versus efficiency (red; decreases), modularity (increases; blue) and small-worldness which peaks at the point at which there is a trade-off between modularity and efficiency and then declines thereafter. This inverted-U function peaks within a window of regularisation that is highlighted in purple and continued into the next panel. The y-axis is normalised and therefore the scale is removed.

10. CODE AND DATA AVAILABILITY

Chapter 2 results were generated using code written in R, Python and Matlab. All code is available at https://github.com/DanAkarca/generativenetworkmodel. The datasets supporting the current study have not been deposited in a public repository because of restrictions imposed by NHS ethical approval, but are available from the corresponding author on request. Requests for access can be made by research-based institutions for academic purposes. Unidentifiable simulated data can be found at https://osf.io/h9px4/?view_only=984260dcff444b59819961ece9c724ec.

Chapter 3 results were generated using code written in Matlab. All code is available at https://github.com/DanAkarca/MEA_generative_models. All data used in this study, along with documentation detailing each dataset, is openly available at https://zenodo.org/record/6109414#.Yid27y-l2J8.

Chapter 4 results were generated using code written in Python (all RNN training was done using TensorFlow) and Matlab. All code is soon to be deposited in https://github.com/DanAkarca/seRNNs

11. BIBLIOGRAPHY

Abdi, H. (2010). Partial least squares regression and projection on latent structure regression (PLS Regression). *WIREs Computational Statistics*, *2*(1), 97–106. https://doi.org/10.1002/wics.51

Achard, S., Salvador, R., Whitcher, B., Suckling, J., & Bullmore, E. (2006). A Resilient, Low-Frequency, Small-World Human Brain Functional Network with Highly Connected Association Cortical Hubs. *Journal of Neuroscience*, *26*(1), 63–72. https://doi.org/10.1523/JNEUROSCI.3874-05.2006

Adams, R. A., Huys, Q. J. M., & Roiser, J. P. (2016). Computational Psychiatry: Towards a mathematically informed understanding of mental illness. *Journal of Neurology, Neurosurgery, and Psychiatry*, *87*(1), 53–63. https://doi.org/10.1136/jnnp-2015-310737

Akarca, D., Dunn, A. W. E., Hornauer, P. J., Ronchi, S., Fiscella, M., Wang, C., Terrigno, M., Jagasia, R., Vértes, P. E., Mierau, S. B., Paulsen, O., Eglen, S. J., Hierlemann, A., Astle, D. E., & Schröter, M. (2022). Homophilic wiring principles underpin neuronal network topology in vitro. *BioRxiv*. https://doi.org/10.1101/2022.03.09.483605

Akarca, D., Vértes, P. E., Bullmore, E. T., & Astle, D. E. (2021). A generative network model of neurodevelopmental diversity in structural brain organization. *Nature Communications*, *12*(1), 4216. https://doi.org/10.1038/s41467-021-24430-z

Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, *74*(1), 47–97. https://doi.org/10.1103/RevModPhys.74.47

Alexander, B., Loh, W. Y., Matthews, L. G., Murray, A. L., Adamson, C., Beare, R., Chen, J., Kelly, C. E., Anderson, P. J., Doyle, L. W., Spittle, A. J., Cheong, J. L. Y., Seal, M. L., & Thompson, D. K. (2019). Desikan-Killiany-Tourville Atlas Compatible Version of M-CRIB Neonatal Parcellated Whole Brain Atlas: The M-CRIB 2.0. *Frontiers in Neuroscience*, *13*. https://www.frontiersin.org/article/10.3389/fnins.2019.00034

Ali, A., Ahmad, N., Groot, E. de, Gerven, M. A. J. van, & Kietzmann, T. C. (2021). *Predictive coding is a consequence of energy efficiency in recurrent neural networks* (p. 2021.02.16.430904). bioRxiv. https://doi.org/10.1101/2021.02.16.430904

Allard, A., & Serrano, M. Á. (2020). Navigable maps of structural brain networks across species. *PLOS Computational Biology*, *16*(2), e1007584. https://doi.org/10.1371/journal.pcbi.1007584

Alloway, T. (2007). *Automated Working Memory Assessment (AWMA)*. Pearson Assessment.

Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2008). Evaluating the validity of the Automated Working Memory Assessment. *Educational Psychology*, *28*(7), 725–734. https://doi.org/10.1080/01443410802243828

Alstott, J., Bullmore, E., & Plenz, D. (2014). powerlaw: A Python Package for Analysis of Heavy-Tailed Distributions. *PLOS ONE*, *9*(1), e85777. https://doi.org/10.1371/journal.pone.0085777

Andersen, H. (2014). A Field Guide to Mechanisms: Part I. *Philosophy Compass*, *9*(4), 274–283. https://doi.org/10.1111/phc3.12119

Andersen, S. L. (2003). Trajectories of brain development: Point of vulnerability or window of opportunity? *Neuroscience & Biobehavioral Reviews*, *27*(1), 3–18. https://doi.org/10.1016/S0149-7634(03)00005-8

Armingol, E., Officer, A., Harismendy, O., & Lewis, N. E. (2021). Deciphering cell–cell interactions and communication from gene expression. *Nature Reviews Genetics*, *22*(2), 2. https://doi.org/10.1038/s41576-020-00292-x

Arnatkevičiūtė, A., Fulcher, B. D., & Fornito, A. (2019). A practical guide to linking brain-wide gene expression and neuroimaging data. *NeuroImage*, *189*, 353–367. https://doi.org/10.1016/j.neuroimage.2019.01.011

Arnatkevičiūtė, A., Fulcher, B. D., Oldham, S., Tiego, J., Paquola, C., Gerring, Z., Aquino, K., Hawi, Z., Johnson, B., Ball, G., Klein, M., Deco, G., Franke, B., Bellgrove, M. A., & Fornito, A. (2021). Genetic influences on hub connectivity of the human connectome. *Nature Communications*, *12*(1), 4237. https://doi.org/10.1038/s41467-021-24306-2

Assem, M., Glasser, M. F., Van Essen, D. C., & Duncan, J. (2020). A Domain-General Cognitive Core Defined in Multimodally Parcellated Human Cortex. *Cerebral Cortex*, *30*(8), 4361–4380. https://doi.org/10.1093/cercor/bhaa023

Assenza, S., Gutiérrez, R., Gómez-Gardeñes, J., Latora, V., & Boccaletti, S. (2011). Emergence of structural patterns out of synchronization in networks with competitive interactions. *Scientific Reports*, *1*(1), 1. https://doi.org/10.1038/srep00099

Astle, D. E., Bathelt, J., Team, T. C., & Holmes, J. (2019). Remapping the cognitive and neural profiles of children who struggle at school. *Developmental Science*, *22*(1), e12747. https://doi.org/10.1111/desc.12747

Astle, D. E., & Fletcher-Watson, S. (2020). Beyond the Core-Deficit Hypothesis in Developmental Disorders. *Current Directions in Psychological Science*, *29*(5), 431–437. https://doi.org/10.1177/0963721420925518

Attwell, D., & Laughlin, S. B. (2001). An Energy Budget for Signaling in the Grey Matter of the Brain. *Journal of Cerebral Blood Flow & Metabolism*, *21*(10), 1133–1145. https://doi.org/10.1097/00004647-200110000-00001

Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., & Gee, J. C. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage*, *54*(3), 2033–2044. https://doi.org/10.1016/j.neuroimage.2010.09.025

Avena-Koenigsberger, A., Goñi, J., Solé, R., & Sporns, O. (2015). Network morphospace. *Journal of The Royal Society Interface*, *12*(103), 20140881. https://doi.org/10.1098/rsif.2014.0881

Avena-Koenigsberger, A., Misic, B., & Sporns, O. (2017). Communication dynamics in complex brain networks. *Nature Reviews. Neuroscience*, *19*(1), 17–33. https://doi.org/10.1038/nrn.2017.149

Avena-Koenigsberger, A., Yan, X., Kolchinsky, A., Heuvel, M. P. van den, Hagmann, P., & Sporns, O. (2019). A spectrum of routing strategies for brain networks. *PLOS Computational Biology*, *15*(3), e1006833. https://doi.org/10.1371/journal.pcbi.1006833

Bakkum, D. J., Frey, U., Radivojevic, M., Russell, T. L., Müller, J., Fiscella, M., Takahashi, H., & Hierlemann, A. (2013). Tracking axonal action potential propagation on a high-density microelectrode array across hundreds of sites. *Nature Communications*, *4*, 2181. https://doi.org/10.1038/ncomms3181

Bakkum, D., Radivojevic, M., Frey, U., Franke, F., Hierlemann, A., & Takahashi, H. (2014). Parameters for burst detection. *Frontiers in Computational Neuroscience*, *7*. https://www.frontiersin.org/article/10.3389/fncom.2013.00193

Ball, G., Aljabar, P., Zebari, S., Tusor, N., Arichi, T., Merchant, N., Robinson, E. C., Ogundipe, E., Rueckert, D., Edwards, A. D., & Counsell, S. J. (2014). Rich-club organization of the newborn human brain. *Proceedings of the National Academy of Sciences*, *111*(20), 7456–7461. https://doi.org/10.1073/pnas.1324118111

Ballini, M., Müller, J., Livi, P., Chen, Y., Frey, U., Stettler, A., Shadmani, A., Viswam, V., Jones, I. L., Jäckel, D., Radivojevic, M., Lewandowska, M. K., Gong, W., Fiscella, M., Bakkum, D. J., Heer, F., & Hierlemann, A. (2014). A 1024-Channel CMOS Microelectrode Array With 26,400 Electrodes for Recording and Stimulation of Electrogenic Cells In Vitro. *IEEE Journal of Solid-State Circuits*, *49*(11), 2705–2719. https://doi.org/10.1109/JSSC.2014.2359219

Barabasi, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science (New York, N.Y.)*, *286*(5439), 509–512. https://doi.org/10.1126/science.286.5439.509

Barabási, D. L., & Barabási, A.-L. (2020). A Genetic Model of the Connectome. *Neuron*, *105*(3), 435-445.e5. https://doi.org/10.1016/j.neuron.2019.10.031

Baram, A. B., Muller, T. H., Whittington, J. C. R., & Behrens, T. E. J. (2018). *Intuitive planning: Global navigation through cognitive maps based on grid-like codes* [Preprint]. Neuroscience. https://doi.org/10.1101/421461

Barral, J., & D Reyes, A. (2016). Synaptic scaling rule preserves excitatory–inhibitory balance and salient neuronal network dynamics. *Nature Neuroscience*, *19*(12), 12. https://doi.org/10.1038/nn.4415

Barron, H. C., Mars, R. B., Dupret, D., Lerch, J. P., & Sampaio-Baptista, C. (2021). Cross-species neuroscience: Closing the explanatory gap. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1815), 20190633. https://doi.org/10.1098/rstb.2019.0633

Barzel, B., & Biham, O. (2009). Quantifying the connectivity of a network: The network correlation function method. *Physical Review E*, *80*(4), 046104. https://doi.org/10.1103/PhysRevE.80.046104

Basser, P. J., Pajevic, S., Pierpaoli, C., Duda, J., & Aldroubi, A. (2000). In vivo fiber tractography using DT-MRI data. *Magnetic Resonance in Medicine*, *44*(4), 625–632. https://doi.org/10.1002/1522-2594(200010)44:4<625::aid-mrm17>3.0.co;2-o

Bassett, D. S., & Bullmore, E. T. (2017). Small-World Brain Networks Revisited. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *23*(5), 499–516. https://doi.org/10.1177/1073858416667720

Bassett, D. S., & Sporns, O. (2017). Network neuroscience. *Nature Neuroscience*, *20*(3), 3. https://doi.org/10.1038/nn.4502

Bathelt, J., Astle, D., Barnes, J., Raymond, F. L., & Baker, K. (2016). Structural brain abnormalities in a single gene disorder associated with epilepsy, language impairment and intellectual disability. *NeuroImage: Clinical*, *12*, 655–665. https://doi.org/10.1016/j.nicl.2016.07.016

Bathelt, J., Barnes, J., Raymond, F. L., Baker, K., & Astle, D. (2017). Global and Local Connectivity Differences Converge With Gene Expression in a Neurodevelopmental Disorder of Known Genetic Origin. *Cerebral Cortex (New York, N.Y. : 1991)*, *27*(7), 3806–3817. https://doi.org/10.1093/cercor/bhx027

Bathelt, J., Gathercole, S. E., Butterfield, S., the CALM team, & Astle, D. E. (2018). Children's academic attainment is linked to the global organization of the white matter connectome. *Developmental Science*, *21*(5), e12662. https://doi.org/10.1111/desc.12662

Bathelt, J., Holmes, J., Astle, D. E., Holmes, J., Gathercole, S., Astle, D., Manly, T., & Kievit, R. (2018). Data-Driven Subtyping of Executive Function–Related Behavioral Problems in Children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *57*(4), 252-262.e4. https://doi.org/10.1016/j.jaac.2018.01.014

Bathelt, J., Scerif, G., Nobre, A. C., & Astle, D. E. (2019). Whole-brain white matter organization, intelligence, and educational attainment. *Trends in Neuroscience and Education*, *15*, 38–47. https://doi.org/10.1016/j.tine.2019.02.004

Batista-García-Ramó, K., & Fernández-Verdecia, C. I. (2018). What We Know About the Brain Structure–Function Relationship. *Behavioral Sciences*, *8*(4), 39. https://doi.org/10.3390/bs8040039

Battista, C., Evans, T. M., Ngoon, T. J., Chen, T., Chen, L., Kochalka, J., & Menon, V. (2018). Mechanisms of interactive specialization and emergence of functional brain circuits supporting cognitive development in children. *Npj Science of Learning*, *3*(1), 1–11. https://doi.org/10.1038/s41539-017-0017-2

Baxter, R. A., & Levy, W. B. (2020). Constructing multilayered neural networks with sparse, data-driven connectivity using biologically-inspired, complementary, homeostatic mechanisms. *Neural Networks*, *122*, 68–93. https://doi.org/10.1016/j.neunet.2019.09.025

Bazinet, V., Hansen, J. Y., Vos de Wael, R., Bernhardt, B. C., van den Heuvel, M. P., & Misic, B. (2022). *Assortative mixing in micro-architecturally annotated brain connectomes* [Preprint]. Neuroscience. https://doi.org/10.1101/2022.05.31.494070

Bazzi, M., Jeub, L. G. S., Arenas, A., Howison, S. D., & Porter, M. A. (2020). A framework for the construction of generative models for mesoscale structure in multilayer networks. *Physical Review Research*, *2*(2), 023100. https://doi.org/10.1103/PhysRevResearch.2.023100

Beasley, W. H., & Rodgers, J. L. (2012). Bootstrapping and Monte Carlo methods. In *APA handbook of research methods in psychology* (Vol. 2, pp. 407–425). https://doi.org/10.1037/13620-022

Beggs, J. M. (2008). The criticality hypothesis: How local cortical networks might optimize information processing. *Philosophical Transactions of the Royal Society A:*

Mathematical, Physical and Engineering Sciences, *366*(1864), 329–343. https://doi.org/10.1098/rsta.2007.2092

Bell, M., Perera, S., Piraveenan, M., Bliemer, M., Latty, T., & Reid, C. (2017). Network growth models: A behavioural basis for attachment proportional to fitness. *Scientific Reports*, *7*(1), 1. https://doi.org/10.1038/srep42431

Benzi, M., & Klymko, C. (2013). Total communicability as a centrality measure | Journal of Complex Networks | Oxford Academic. *Journal of Complex Networks*, *1*, 124–149.

Bernardi, S., Benna, M. K., Rigotti, M., Munuera, J., Fusi, S., & Salzman, C. D. (2020). The Geometry of Abstraction in the Hippocampus and Prefrontal Cortex. *Cell*, *183*(4), 954-967.e21. https://doi.org/10.1016/j.cell.2020.09.031

Bethlehem, R. a. I., Seidlitz, J., White, S. R., Vogel, J. W., Anderson, K. M., Adamson, C., Adler, S., Alexopoulos, G. S., Anagnostou, E., Areces-Gonzalez, A., Astle, D. E., Auyeung, B., Ayub, M., Bae, J., Ball, G., Baron-Cohen, S., Beare, R., Bedford, S. A., Benegal, V., … Alexander-Bloch, A. F. (2022). Brain charts for the human lifespan. *Nature*, *604*(7906), 7906. https://doi.org/10.1038/s41586-022-04554-y

Betzel, R. F. (2020). Organizing principles of whole-brain functional connectivity in zebrafish larvae. *Network Neuroscience*, *4*(1), 234–256. https://doi.org/10.1162/netn_a_00121

Betzel, R. F., Avena-Koenigsberger, A., Goñi, J., He, Y., de Reus, M. A., Griffa, A., Vértes, P. E., Mišic, B., Thiran, J.-P., Hagmann, P., van den Heuvel, M., Zuo, X.-N., Bullmore, E. T., & Sporns, O. (2016). Generative models of the human connectome. *NeuroImage*, *124*(Pt A), 1054–1064. https://doi.org/10.1016/j.neuroimage.2015.09.041

Betzel, R. F., & Bassett, D. S. (2017a). Multi-scale brain networks. *NeuroImage*, *160*, 73–83. https://doi.org/10.1016/j.neuroimage.2016.11.006

Betzel, R. F., & Bassett, D. S. (2017b). Generative models for network neuroscience: Prospects and promise. *Journal of The Royal Society Interface*, *14*(136), 20170623. https://doi.org/10.1098/rsif.2017.0623

Betzel, R. F., & Bassett, D. S. (2018). Specificity and robustness of long-distance connections in weighted, interareal connectomes. *Proceedings of the National Academy of Sciences*, *115*(21), E4880–E4889. https://doi.org/10.1073/pnas.1720186115

Betzel, R. F., Medaglia, J. D., Papadopoulos, L., Baum, G. L., Gur, R., Gur, R., Roalf, D., Satterthwaite, T. D., & Bassett, D. S. (2017). The modular organization of human anatomical brain networks: Accounting for the cost of wiring. *Network Neuroscience*, *1*(1), 42–68. https://doi.org/10.1162/NETN_a_00002

Billeh, Y. N., Cai, B., Gratiy, S. L., Dai, K., Iyer, R., Gouwens, N. W., Abbasi-Asl, R., Jia, X., Siegle, J. H., Olsen, S. R., Koch, C., Mihalas, S., & Arkhipov, A. (2020). Systematic Integration of Structural and Functional Data into Multi-scale Models of Mouse Primary Visual Cortex. *Neuron*, *106*(3), 388-403.e18. https://doi.org/10.1016/j.neuron.2020.01.040

Blankenship, A. G., & Feller, M. B. (2010). Mechanisms underlying spontaneous patterned activity in developing neural circuits. *Nature Reviews. Neuroscience*, *11*(1), 18–29. https://doi.org/10.1038/nrn2759

Blohm, G., Kording, K. P., & Schrater, P. R. (2020). A How-to-Model Guide for Neuroscience. *ENeuro*, *7*(1), ENEURO.0352-19.2019. https://doi.org/10.1523/ENEURO.0352-19.2019

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Bogdan, P., Caetano-Anollés, G., Jolles, A., Kim, H., Morris, J., Murphy, C. A., Royer, C., Snell, E. H., Steinbrenner, A., & Strausfeld, N. (2021). Biological Networks across Scales—The Theoretical and Empirical Foundations for Time-Varying Complex Networks that Connect Structure and Function across Levels of Biological Organization. *Integrative and Comparative Biology*, *61*(6), 1991–2010. https://doi.org/10.1093/icb/icab069

Bohlken, M. M., Mandl, R. C. W., Brouwer, R. M., van den Heuvel, M. P., Hedman, A. M., Kahn, R. S., & Hulshoff Pol, H. E. (2014). Heritability of structural brain network topology: A DTI study of 156 twins. *Human Brain Mapping*, *35*(10), 5295–5305. https://doi.org/10.1002/hbm.22550

Bonetto, G., Belin, D., & Káradóttir, R. T. (2021). Myelin: A gatekeeper of activity-dependent circuit plasticity? *Science*, *374*(6569), eaba6905. https://doi.org/10.1126/science.aba6905

Bonifazi, P., Goldin, M., Picardo, M. A., Jorquera, I., Cattani, A., Bianconi, G., Represa, A., Ben-Ari, Y., & Cossart, R. (2009). GABAergic hub neurons orchestrate synchrony in developing hippocampal networks. *Science (New York, N.Y.)*, *326*(5958), 1419–1424. https://doi.org/10.1126/science.1175509

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network Analysis in the Social Sciences. *Science*, *323*(5916), 892–895. https://doi.org/10.1126/science.1165821

Boyle, E. A., Li, Y. I., & Pritchard, J. K. (2017). An expanded view of complex traits: From polygenic to omnigenic. *Cell*, *169*(7), 1177–1186. https://doi.org/10.1016/j.cell.2017.05.038

Braitenberg, V., & Schüz, A. (1998). Microscopic Evidence of the Spacing of Synapses on Axons. In V. Braitenberg & A. Schüz (Eds.), *Cortex: Statistics and Geometry of Neuronal Connectivity* (pp. 45–49). Springer. https://doi.org/10.1007/978-3-662-03733-1_9

Briggman, K. L., Helmstaedter, M., & Denk, W. (2011). Wiring specificity in the direction-selectivity circuit of the retina. *Nature*, *471*(7337), 183–188. https://doi.org/10.1038/nature09818

Bullmore, E., & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, *10*(3), 3. https://doi.org/10.1038/nrn2575

Bullmore, E., & Sporns, O. (2012). The economy of brain network organization. *Nature Reviews Neuroscience*, *13*(5), 336–349. https://doi.org/10.1038/nrn3214

Buzsáki, G., & Mizuseki, K. (2014). The log-dynamic brain: How skewed distributions affect network operations. *Nature Reviews Neuroscience*, *15*(4), 4. https://doi.org/10.1038/nrn3687

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., Cortes, A., Welsh, S., Young, A., Effingham, M., McVean, G., Leslie, S., Allen, N., Donnelly, P., & Marchini, J. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature*, *562*(7726), 7726. https://doi.org/10.1038/s41586-018-0579-z

Cajal, S. R. y, Swanson, N., & Swanson, L. W. (1995). *Cajal's Histology of the Nervous System of Man and Vertebrates*. Oxford University Press.

Camp, J. G., Badsha, F., Florio, M., Kanton, S., Gerber, T., Wilsch-Bräuninger, M., Lewitus, E., Sykes, A., Hevers, W., Lancaster, M., Knoblich, J. A., Lachmann, R., Pääbo, S., Huttner, W. B., & Treutlein, B. (2015). Human cerebral organoids recapitulate gene expression programs of fetal neocortex development. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(51), 15672–15677. https://doi.org/10.1073/pnas.1520760112

Carozza, S., Holmes, J., Vértes, P. E., Bullmore, E., Arefin, T. M., Pugliese, A., Zhang, J., Kaffman, A., Akarca, D., & Astle, D. E. (2022). Early adversity changes the economic conditions of structural brain network organisation. *BioRxiv*. https://doi.org/10.1101/2022.06.08.495303

Casey, B. J., Cannonier, T., Conley, M. I., Cohen, A. O., Barch, D. M., Heitzeg, M. M., Soules, M. E., Teslovich, T., Dellarco, D. V., Garavan, H., Orr, C. A., Wager, T. D., Banich, M. T., Speer, N. K., Sutherland, M. T., Riedel, M. C., Dick, A. S., Bjork, J. M., Thomas, K. M., … Dale, A. M. (2018). The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Developmental Cognitive Neuroscience*, *32*, 43–54. https://doi.org/10.1016/j.dcn.2018.03.001

Castellanos, F. X., Sonuga-Barke, E. J. S., Milham, M. P., & Tannock, R. (2006). Characterizing cognition in ADHD: Beyond executive dysfunction. *Trends in Cognitive Sciences*, *10*(3), 117–123. https://doi.org/10.1016/j.tics.2006.01.011

Chai, X. J., Berken, J. A., Barbeau, E. B., Soles, J., Callahan, M., Chen, J.-K., & Klein, D. (2016). Intrinsic Functional Connectivity in the Adult Brain and Success in Second-Language Learning. *Journal of Neuroscience*, *36*(3), 755–761. https://doi.org/10.1523/JNEUROSCI.2234-15.2016

Changeux, J.-P., Goulas, A., & Hilgetag, C. C. (2021). A Connectomic Hypothesis for the Hominization of the Brain. *Cerebral Cortex*, *31*(5), 2425–2449. https://doi.org/10.1093/cercor/bhaa365

Chechik, G., Meilijson, I., & Ruppin, E. (1998). Synaptic pruning in development: A computational account. *Neural Computation*, *10*(7), 1759–1777. https://doi.org/10.1162/089976698300017124

Chen, B. L., Hall, D. H., & Chklovskii, D. B. (2006). Wiring optimization can relate neuronal structure and function. *Proceedings of the National Academy of Sciences*, *103*(12), 4723–4728. https://doi.org/10.1073/pnas.0506806103

Chklovskii, D. B., Mel, B. W., & Svoboda, K. (2004). Cortical rewiring and information storage. *Nature*, *431*(7010), 782–788. https://doi.org/10.1038/nature03012

Christie, S. T., & Schrater, P. (2015). Cognitive cost as dynamic allocation of energetic resources. *Frontiers in Neuroscience*, *9*, 289. https://doi.org/10.3389/fnins.2015.00289

Cicchetti, D., & Rogosch, F. A. (1996). Equifinality and multifinality in developmental psychopathology. *Development and Psychopathology*, *8*(4), 597–600. https://doi.org/10.1017/S0954579400007318

Cisek, P., & Hayden, B., Y. (2022). *Neuroscience needs evolution. 377*, 20200518.

Conturo, T. E., Lori, N. F., Cull, T. S., Akbudak, E., Snyder, A. Z., Shimony, J. S., McKinstry, R. C., Burton, H., & Raichle, M. E. (1999). Tracking neuronal fiber pathways in the living human brain. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(18), 10422–10427.

Cossart, R. (2014). Operational hub cells: A morpho-physiologically diverse class of GABAergic neurons united by a common function. *Current Opinion in Neurobiology*, *26*, 51–56. https://doi.org/10.1016/j.conb.2013.12.002

Cotterill, E., Charlesworth, P., Thomas, C. W., Paulsen, O., & Eglen, S. J. (2016). A comparison of computational methods for detecting bursts in neuronal spike trains and their application to human stem cell-derived neuronal networks. *Journal of Neurophysiology*, *116*(2), 306–321. https://doi.org/10.1152/jn.00093.2016

Crofts, J. J., & Higham, D. J. (2009). A weighted communicability measure applied to complex brain networks. *Journal of the Royal Society Interface*, *6*(33), 411–414. https://doi.org/10.1098/rsif.2008.0484

Crosby, M., Beyret, B., Shanahan, M., Hernández-Orallo, J., Cheke, L., & Halina, M. (2020). The Animal-AI Testbed and Competition. *Proceedings of the NeurIPS 2019 Competition and Demonstration Track*, 164–176. https://proceedings.mlr.press/v123/crosby20a.html

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. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(43), 14288–14303. https://doi.org/10.1523/JNEUROSCI.2767-14.2014

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction. *NeuroImage*, *9*(2), 179–194. https://doi.org/10.1006/nimg.1998.0395

Damicelli, F., Hilgetag, C. C., & Goulas, A. (2021). *Brain Connectivity meets Reservoir Computing* (p. 2021.01.22.427750). bioRxiv. https://doi.org/10.1101/2021.01.22.427750

Damicelli, F., Hilgetag, C. C., Hütt, M.-T., & Messé, A. (2019). Topological reinforcement as a principle of modularity emergence in brain networks. *Network Neuroscience*, *3*(2), 589–605. https://doi.org/10.1162/netn_a_00085

Davies, G., Lam, M., Harris, S. E., Trampush, J. W., Luciano, M., Hill, W. D., Hagenaars, S. P., Ritchie, S. J., Marioni, R. E., Fawns-Ritchie, C., Liewald, D. C. M., Okely, J. A., Ahola-Olli, A. V., Barnes, C. L. K., Bertram, L., Bis, J. C., Burdick, K. E., Christoforou, A., DeRosse, P., … Deary, I. J. (2018). Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nature Communications*, *9*(1), 1. https://doi.org/10.1038/s41467-018-04362-x

Davis, H., & Anderson, M. (1999). Individual differences and development—One dimension or two? In M. Anderson (Ed.), *The Development of Intelligence.* (pp. 161–191). Psychology Press. https://researchrepository.murdoch.edu.au/id/eprint/17793/

De Schutter, E. (2001). Computational Neuroscience: More Math Is Needed to Understand the Human Brain. In B. Engquist & W. Schmid (Eds.), *Mathematics Unlimited—2001 and Beyond* (pp. 381–391). Springer. https://doi.org/10.1007/978-3-642-56478-9_17

de Wit, J., & Ghosh, A. (2016). Specification of synaptic connectivity by cell surface interactions. *Nature Reviews Neuroscience*, *17*(1), 1. https://doi.org/10.1038/nrn.2015.3

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021

Dorfschmidt, L., Bethlehem, R. A., Seidlitz, J., Váša, F., White, S. R., Romero-García, R., Kitzbichler, M. G., Aruldass, A. R., Morgan, S. E., Goodyer, I. M., Fonagy, P., Jones, P. B., Dolan, R. J., NSPN Consortium, Harrison, N. A., Vértes, P. E., & Bullmore, E. T. (2022). Sexually divergent development of depression-related brain networks during healthy human adolescence. *Science Advances*, *8*(21), eabm7825. https://doi.org/10.1126/sciadv.abm7825

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 Computational Biology*, *8*(5), e1002522. https://doi.org/10.1371/journal.pcbi.1002522

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test*. Pearson Education.

Durston, S., Davidson, M. C., Tottenham, N., Galvan, A., Spicer, J., Fossella, J. A., & Casey, B. J. (2006). A shift from diffuse to focal cortical activity with development. *Developmental Science*, *9*(1), 1–8. https://doi.org/10.1111/j.1467-7687.2005.00454.x

Eickhoff, S. B., Yeo, B. T. T., & Genon, S. (2018). Imaging-based parcellations of the human brain. *Nature Reviews Neuroscience*, *19*(11), 11. https://doi.org/10.1038/s41583-018-0071-7

Engel, T. A., Schölvinck, M. L., & Lewis, C. M. (2021). The diversity and specificity of functional connectivity across spatial and temporal scales. *NeuroImage*, *245*, 118692. https://doi.org/10.1016/j.neuroimage.2021.118692

Ercsey-Ravasz, M., Markov, N. T., Lamy, C., Van Essen, D. C., Knoblauch, K., Toroczkai, Z., & Kennedy, H. (2013). A Predictive Network Model of Cerebral Cortical Connectivity Based on a Distance Rule. *Neuron*, *80*(1), 184–197. https://doi.org/10.1016/j.neuron.2013.07.036

Erdős, P., & Rényi, A. (1959). On Random Graphs. I. *Publicationes Mathematicae*, *6*, 290–297.

Erez, Y., Kadohisa, M., Petrov, P., Sigala, N., Buckley, M. J., Kusunoki, M., & Duncan, J. (2022). Integrated neural dynamics for behavioural decisions and attentional

competition in the prefrontal cortex. *European Journal of Neuroscience*, *56*(4), 4393–4410. https://doi.org/10.1111/ejn.15757

Estrada, E., & Hatano, N. (2008). Communicability in complex networks. *Physical Review E*, *77*(3), 036111. https://doi.org/10.1103/PhysRevE.77.036111

Fahlman, S., & Lebiere, C. (1990). The Cascade-Correlation Learning Architecture. *Advances in Neural Information Processing Systems*, *2*. https://proceedings.neurips.cc/paper/1989/hash/69adc1e107f7f7d035d7baf04342e1ca-Abstract.html

Fair, D. A., Dosenbach, N. U. F., Church, J. A., Cohen, A. L., Brahmbhatt, S., Miezin, F. M., Barch, D. M., Raichle, M. E., Petersen, S. E., & Schlaggar, B. L. (2007). Development of distinct control networks through segregation and integration. *Proceedings of the National Academy of Sciences*, *104*(33), 13507–13512. https://doi.org/10.1073/pnas.0705843104

Fine, P. V. A. (2015). Ecological and Evolutionary Drivers of Geographic Variation in Species Diversity. *Annual Review of Ecology, Evolution, and Systematics*, *46*(1), 369–392. https://doi.org/10.1146/annurev-ecolsys-112414-054102

Finlay, B. L., Darlington, R. B., & Nicastro, N. (2001). Developmental structure in brain evolution. *Behavioral and Brain Sciences*, *24*(2), 263–278. https://doi.org/10.1017/S0140525X01003958

Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images | PNAS. *Proceedings of the National Academy of Sciences*, *97*(20), 11050–11055. https://doi.org/10.1073/pnas.200033797

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., & Dale, A. M. (2002). Whole Brain Segmentation: Automated Labeling of Neuroanatomical Structures in the Human Brain. *Neuron*, *33*(3), 341–355. https://doi.org/10.1016/S0896-6273(02)00569-X

Fiske, A., & Holmboe, K. (2019). Neural substrates of early executive function development. *Developmental Review*, *52*, 42–62. https://doi.org/10.1016/j.dr.2019.100866

Fornito, A., Zalesky, A., & Bullmore, E. T. (Eds.). (2016). Chapter 8—Motifs, Small Worlds, and Network Economy. In *Fundamentals of Brain Network Analysis* (pp. 257–301). Academic Press. https://doi.org/10.1016/B978-0-12-407908-3.00008-X

Frederickson, N., Frith, U., & Reason, R. (1997). *Phonological Assessment Battery*. Manuel and Test Materials.

Friston, K. (2002). Functional integration and inference in the brain. *Progress in Neurobiology*, *68*(2), 113–143. https://doi.org/10.1016/S0301-0082(02)00076-X

Fulcher, B. D., Arnatkeviciute, A., & Fornito, A. (2021). Overcoming false-positive gene-category enrichment in the analysis of spatially resolved transcriptomic brain atlas data. *Nature Communications*, *12*(1), 1. https://doi.org/10.1038/s41467-021-22862-1

Fusi, S., Drew, P. J., & Abbott, L. F. (2005). Cascade Models of Synaptically Stored Memories. *Neuron*, *45*(4), 599–611. https://doi.org/10.1016/j.neuron.2005.02.001

Fusi, S., Miller, E. K., & Rigotti, M. (2016). Why neurons mix: High dimensionality for higher cognition. *Current Opinion in Neurobiology*, *37*, 66–74. https://doi.org/10.1016/j.conb.2016.01.010

Gallos, L. K., Makse, H. A., & Sigman, M. (2012). A small world of weak ties provides optimal global integration of self-similar modules in functional brain networks. *Proceedings of the National Academy of Sciences*, *109*(8), 2825–2830. https://doi.org/10.1073/pnas.1106612109

Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., Van Der Walt, S., Descoteaux, M., & Nimmo-Smith, I. (2014). Dipy, a library for the analysis of diffusion MRI data. *Frontiers in Neuroinformatics*, *8*. https://www.frontiersin.org/article/10.3389/fninf.2014.00008

Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, *349*(6245), 273–278. https://doi.org/10.1126/science.aac6076

Giandomenico, S. L., Mierau, S. B., Gibbons, G. M., Wenger, L. M. D., Masullo, L., Sit, T., Sutcliffe, M., Boulanger, J., Tripodi, M., Derivery, E., Paulsen, O., Lakatos, A., & Lancaster, M. A. (2019). Cerebral organoids at the air-liquid interface generate diverse nerve tracts with functional output. *Nature Neuroscience*, *22*(4), 669–679. https://doi.org/10.1038/s41593-019-0350-2

Gollo, L., Roberts, J., Cropley, V., Di Biase, M., Pantelis, C., Zalesky, A., & Breakspear, M. (2018). Fragility and volatility of structural hubs in the human connectome | Nature Neuroscience. *Nature Neuroscience*, *21*, 1107–1116. https://doi.org/10.1038/s41593-018-0188-z

Gordon, A., Yoon, S.-J., Tran, S. S., Makinson, C. D., Park, J. Y., Andersen, J., Valencia, A. M., Horvath, S., Xiao, X., Huguenard, J. R., Pașca, S. P., & Geschwind, D. H. (2021). Long-term maturation of human cortical organoids matches key early postnatal transitions. *Nature Neuroscience*, *24*(3), 3. https://doi.org/10.1038/s41593-021-00802-y

Gottlieb, G. (1998). Normally occurring environmental and behavioral influences on gene activity: From central dogma to probabilistic epigenesis. *Psychological Review*, *105*(4), 792–802. https://doi.org/10.1037/0033-295x.105.4.792-802

Gottlieb, G. (2005). Probabilistic Epigenesis of Development. In *Handbook of Developmental Psychology* (pp. 3–17). SAGE Publications Ltd. https://doi.org/10.4135/9781848608306

Gottlieb, G. (2007). Probabilistic epigenesis. *Developmental Science*, *10*(1), 1–11. https://doi.org/10.1111/j.1467-7687.2007.00556.x

Goulas, A., Betzel, R. F., & Hilgetag, C. C. (2019). Spatiotemporal ontogeny of brain wiring. *Science Advances*, *5*(6), eaav9694. https://doi.org/10.1126/sciadv.aav9694

Goulas, A., Damicelli, F., & Hilgetag, C. C. (2021). Bio-instantiated recurrent neural networks: Integrating neurobiology-based network topology in artificial networks. *Neural Networks*, *142*, 608–618. https://doi.org/10.1016/j.neunet.2021.07.011

Goulas, A., Majka, P., Rosa, M. G. P., & Hilgetag, C. C. (2019). A blueprint of mammalian cortical connectomes. *PLOS Biology*, *17*(3), e2005346. https://doi.org/10.1371/journal.pbio.2005346

Griffiths, K. R., Grieve, S. M., Kohn, M. R., Clarke, S., Williams, L. M., & Korgaonkar, M. S. (2016). Altered gray matter organization in children and adolescents with ADHD: A structural covariance connectome study. *Translational Psychiatry*, *6*(11), e947. https://doi.org/10.1038/tp.2016.219

Grossberg, S. (2020). A Path Toward Explainable AI and Autonomous Adaptive Intelligence: Deep Learning, Adaptive Resonance, and Models of Perception, Emotion, and Action. *Frontiers in Neurorobotics*, *14*. https://www.frontiersin.org/article/10.3389/fnbot.2020.00036

Grove, J., Ripke, S., Als, T. D., Mattheisen, M., Walters, R. K., Won, H., Pallesen, J., Agerbo, E., Andreassen, O. A., Anney, R., Awashti, S., Belliveau, R., Bettella, F., Buxbaum, J. D., Bybjerg-Grauholm, J., Bækvad-Hansen, M., Cerrato, F., Chambert, K., Christensen, J. H., … Børglum, A. D. (2019). Identification of common genetic risk variants for autism spectrum disorder. *Nature Genetics*, *51*(3), 3. https://doi.org/10.1038/s41588-019-0344-8

Hafizi, H., Nigam, S., Barnathan, J., Ren, N., Stevenson, I. H., Masmanidis, S. C., Newman, E. L., Sporns, O., & Beggs, J. M. (2021). *Inhibition-Dominated Rich-Club Shapes Dynamics in Cortical Microcircuits in Awake Behaving Mice* (p. 2021.05.07.443074). bioRxiv. https://doi.org/10.1101/2021.05.07.443074

Hagmann, P., Kurant, M., Gigandet, X., Thiran, P., Wedeen, V. J., Meuli, R., & Thiran, J.-P. (2007). Mapping Human Whole-Brain Structural Networks with Diffusion MRI. *PLOS ONE*, *2*(7), e597. https://doi.org/10.1371/journal.pone.0000597

Hardt, M., & Recht, B. (2022). *Patterns, Predictions, and Actions*. https://press.princeton.edu/books/hardcover/9780691233734/patterns-predictions-and-actions

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, *106*(3), 491–528. https://doi.org/10.1037/0033-295x.106.3.491

Harriger, L., Heuvel, M. P. van den, & Sporns, O. (2012). Rich Club Organization of Macaque Cerebral Cortex and Its Role in Network Communication. *PLOS ONE*, *7*(9), e46497. https://doi.org/10.1371/journal.pone.0046497

Hassabis, D., Kumaran, D., Summerfield, C., & Botvinick, M. (2017). Neuroscience-Inspired Artificial Intelligence. *Neuron*, *95*(2), 245–258. https://doi.org/10.1016/j.neuron.2017.06.011

Hassan, B. A., & Hiesinger, P. R. (2015). Beyond Molecular Codes: Simple Rules to Wire Complex Brains. *Cell*, *163*(2), 285–291. https://doi.org/10.1016/j.cell.2015.09.031

Hawkins, E., Akarca, D., Zhang, M., Brkić, D., Woolrich, M., Baker, K., & Astle, D. (2020). Functional network dynamics in a neurodevelopmental disorder of known genetic origin. *Human Brain Mapping*, *41*(2), 530–544. https://doi.org/10.1002/hbm.24820

Hawrylycz, M. J., Lein, E. S., Guillozet-Bongaarts, A. L., Shen, E. H., Ng, L., Miller, J. A., van de Lagemaat, L. N., Smith, K. A., Ebbert, A., Riley, Z. L., Abajian, C., Beckmann,

C. F., Bernard, A., Bertagnolli, D., Boe, A. F., Cartagena, P. M., Chakravarty, M. M., Chapin, M., Chong, J., … Jones, A. R. (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature*, *489*(7416), 7416. https://doi.org/10.1038/nature11405

Hebb, D. O. (1949). *The organization of behavior; a neuropsychological theory* (pp. xix, 335). Wiley.

Heide, M., Huttner, W. B., & Mora-Bermúdez, F. (2018). Brain organoids as models to study human neocortex development and evolution. *Current Opinion in Cell Biology*, *55*, 8–16. https://doi.org/10.1016/j.ceb.2018.06.006

Heuvel, M. P. van den, Bullmore, E. T., & Sporns, O. (2016). Comparative Connectomics. *Trends in Cognitive Sciences*, *20*(5), 345–361. https://doi.org/10.1016/j.tics.2016.03.001

Heuvel, M. P. van den, & Sporns, O. (2011). Rich-Club Organization of the Human Connectome. *Journal of Neuroscience*, *31*(44), 15775–15786. https://doi.org/10.1523/JNEUROSCI.3539-11.2011

Heuvel, M. P. van den, & Sporns, O. (2013). Network hubs in the human brain. *Trends in Cognitive Sciences*, *17*(12), 683–696. https://doi.org/10.1016/j.tics.2013.09.012

Hiesinger, P. R., & Hassan, B. A. (2018). The Evolution of Variability and Robustness in Neural Development. *Trends in Neurosciences*, *41*(9), 577–586. https://doi.org/10.1016/j.tins.2018.05.007

Hilger, K., Ekman, M., Fiebach, C. J., & Basten, U. (2017). Intelligence is associated with the modular structure of intrinsic brain networks. *Scientific Reports*, *7*(1), 1. https://doi.org/10.1038/s41598-017-15795-7

Hilgetag, C. C., & Goulas, A. (2020). 'Hierarchy' in the organization of brain networks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1796), 20190319. https://doi.org/10.1098/rstb.2019.0319

Hirokawa, J., Vaughan, A., Masset, P., Ott, T., & Kepecs, A. (2019). Frontal cortex neuron types categorically encode single decision variables. *Nature*, *576*(7787), 7787. https://doi.org/10.1038/s41586-019-1816-9

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Holland, P. W., Laskey, K. B., & Leinhardt, S. (1983). Stochastic blockmodels: First steps. *Social Networks*, *5*(2), 109–137. https://doi.org/10.1016/0378-8733(83)90021-7

Holmes, J., Bryant, A., Gathercole, S. E., & the CALM Team. (2019). Protocol for a transdiagnostic study of children with problems of attention, learning and memory (CALM). *BMC Pediatrics*, *19*(1), 10. https://doi.org/10.1186/s12887-018-1385-3

Hong, S.-J., Vos de Wael, R., Bethlehem, R., Lariviere, S., Paquola, C., Valk, S., Milham, M., Di Martino, A., Margulies, D., Smallwood, J., & Bernhardt, B. (2019). Atypical functional connectome hierarchy in autism | Nature Communications. *Nature Communications*, *10*, 1002.

Huttenlocher, P. R. (1979). Synaptic density in human frontal cortex—Developmental changes and effects of aging. *Brain Research*, *163*(2), 195–205. https://doi.org/10.1016/0006-8993(79)90349-4

Huttenlocher, P. R., & de Courten, C. (1987). The development of synapses in striate cortex of man. *Human Neurobiology*, *6*(1), 1–9.

Huys, Q. J. M., Maia, T. V., & Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience*, *19*(3), 404–413. https://doi.org/10.1038/nn.4238

Izsak, J., Theiss, S., & Illes, S. (2022). Ontogeny of oscillatory slow-wave and neuronal population activity in human iPSC-3D cortical circuits. *BioRxiv*, 2022.03.14.484311. https://doi.org/10.1101/2022.03.14.484311

Jain, S., Lin, Y., Kurmangaliyev, Y. Z., Valdes-Aleman, J., LoCascio, S. A., Mirshahidi, P., Parrington, B., & Zipursky, S. L. (2022). A global timing mechanism regulates cell-type-specific wiring programmes. *Nature*, *603*(7899), 7899. https://doi.org/10.1038/s41586-022-04418-5

Ji, J. L., Spronk, M., Kulkarni, K., Repovš, G., Anticevic, A., & Cole, M. W. (2019). Mapping the human brain's cortical-subcortical functional network organization. *NeuroImage*, *185*, 35–57. https://doi.org/10.1016/j.neuroimage.2018.10.006

Johnson, A., Bathelt, J., Akarca, D., Crickmore, G., & Astle, D. E. (2021). Far and wide: Associations between childhood socio-economic status and brain connectomics. *Developmental Cognitive Neuroscience*, *48*, 100888. https://doi.org/10.1016/j.dcn.2020.100888

Johnson, M. H. (2000). Functional brain development in infants: Elements of an interactive specialization framework. *Child Development*, *71*(1), 75–81. https://doi.org/10.1111/1467-8624.00120

Johnson, M. H. (2011). Interactive Specialization: A domain-general framework for human functional brain development? *Developmental Cognitive Neuroscience*, *1*(1), 7–21. https://doi.org/10.1016/j.dcn.2010.07.003

Johnson, M. H., & De Haan, M. (2015). *Developmental Cognitive Neuroscience: An Introduction, 4th Edition*. Wiley.

Johnston, W. J., Palmer, S. E., & Freedman, D. J. (2020). Nonlinear mixed selectivity supports reliable neural computation. *PLOS Computational Biology*, *16*(2), e1007544. https://doi.org/10.1371/journal.pcbi.1007544

Jones, J. S., Team, T. C., & Astle, D. E. (2022). Segregation and integration of the functional connectome in neurodevelopmentally 'at risk' children. *Developmental Science*, *25*(3), e13209. https://doi.org/10.1111/desc.13209

Joseph, J. E., Gathers, A. D., & Bhatt, R. S. (2011). Progressive and regressive developmental changes in neural substrates for face processing: Testing specific predictions of the Interactive Specialization account. *Developmental Science*, *14*(2), 227–241. https://doi.org/10.1111/j.1467-7687.2010.00963.x

Ju, H., Colbert, C. M., & Levy, W. B. (2017). Limited synapse overproduction can speed development but sometimes with long-term energy and discrimination penalties. *PLOS Computational Biology*, *13*(9), e1005750. https://doi.org/10.1371/journal.pcbi.1005750

Kaiser, M. (2017). Mechanisms of Connectome Development. *Trends in Cognitive Sciences*, *21*(9), 703–717. https://doi.org/10.1016/j.tics.2017.05.010

Kaiser, M., & Hilgetag, C. C. (2004). Modelling the development of cortical systems networks. *Neurocomputing*, *58–60*, 297–302. https://doi.org/10.1016/j.neucom.2004.01.059

Karcher, N. R., & Barch, D. M. (2021). The ABCD study: Understanding the development of risk for mental and physical health outcomes. *Neuropsychopharmacology*, *46*(1), 1. https://doi.org/10.1038/s41386-020-0736-6

Kasai, H., Ziv, N. E., Okazaki, H., Yagishita, S., & Toyoizumi, T. (2021). Spine dynamics in the brain, mental disorders and artificial neural networks. *Nature Reviews Neuroscience*, *22*(7), 7. https://doi.org/10.1038/s41583-021-00467-3

Kashtan, N., & Alon, U. (2005). Spontaneous evolution of modularity and network motifs. *Proceedings of the National Academy of Sciences*, *102*(39), 13773–13778. https://doi.org/10.1073/pnas.0503610102

Kawai, Y., Park, J., & Asada, M. (2019). A small-world topology enhances the echo state property and signal propagation in reservoir computing. *Neural Networks*, *112*, 15–23. https://doi.org/10.1016/j.neunet.2019.01.002

Keller, N., Hennings, A., & Dunsmoor, J. (2020). Behavioral and neural processes in counterconditioning: Past and future directions. *Behaviour Research and Therapy*, *125*, 103532. https://doi.org/10.1016/j.brat.2019.103532

Kendler, K. S., & Greenspan, R. J. (2006). The Nature of Genetic Influences on Behavior: Lessons From "Simpler" Organisms. *American Journal of Psychiatry*, *163*(10), 1683–1694. https://doi.org/10.1176/ajp.2006.163.10.1683

Khaligh-Razavi, S.-M., & Kriegeskorte, N. (2014). Deep Supervised, but Not Unsupervised, Models May Explain IT Cortical Representation. *PLOS Computational Biology*, *10*(11), e1003915. https://doi.org/10.1371/journal.pcbi.1003915

Kievit, R. A., Davis, S. W., Mitchell, D. J., Taylor, J. R., Duncan, J., & Henson, R. N. A. (2014). Distinct aspects of frontal lobe structure mediate age-related differences in fluid intelligence and multitasking. *Nature Communications*, *5*(1), 5658. https://doi.org/10.1038/ncomms6658

Kool, W., & Botvinick, M. (2018). Mental labour. *Nature Human Behaviour*, *2*(12), 899–908. https://doi.org/10.1038/s41562-018-0401-9

Kriegeskorte, N., & Douglas, P. K. (2018). Cognitive computational neuroscience. *Nature Neuroscience*, *21*(9), 1148–1160. https://doi.org/10.1038/s41593-018-0210-5

Krishnan, S., Watkins, K. E., & Bishop, D. V. M. (2016). Neurobiological Basis of Language Learning Difficulties. *Trends in Cognitive Sciences*, *20*(9), 701–714. https://doi.org/10.1016/j.tics.2016.06.012

Kumar, R., Novak, J., & Tomkins, A. (2006). Structure and Evolution of Online Social Networks. *Models, Algorithms, and Applications*, 337–357.

Latora, V., & Marchiori, M. (2001). Efficient Behavior of Small-World Networks. *Physical Review Letters*, *87*(19), 198701. https://doi.org/10.1103/PhysRevLett.87.198701

Laughlin, S. B., & Sejnowski, T. J. (2003). Communication in neuronal networks. *Science (New York, N.Y.)*, *301*(5641), 1870–1874. https://doi.org/10.1126/science.1089662

Le Magueresse, C., & Monyer, H. (2013). GABAergic Interneurons Shape the Functional Maturation of the Cortex. *Neuron*, *77*(3), 388–405. https://doi.org/10.1016/j.neuron.2013.01.011

Lebel, C., & Deoni, S. (2018). The development of brain white matter microstructure. *NeuroImage*, *182*, 207–218. https://doi.org/10.1016/j.neuroimage.2017.12.097

Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278–2324. https://doi.org/10.1109/5.726791

Li, S.-C., & Lindenberger, U. (1999). Cross-level unification: A computational exploration of the link between deterioration of neurotransmitter systems and dedifferentiation of cognitive abilities in old age. In *Cognitive neuroscience of memory* (pp. 103–146). Hogrefe & Huber Publishers.

Lindfield, G., & Penny, J. (2019). Chapter 2—Linear Equations and Eigensystems. In G. Lindfield & J. Penny (Eds.), *Numerical Methods (Fourth Edition)* (pp. 73–156). Academic Press. https://doi.org/10.1016/B978-0-12-812256-3.00011-7

Liu, X., Si, S., Hu, B., Zhao, H., & Zhu, J. (2020). A Generative Network Model of the Human Brain Normal Aging Process. *Symmetry*, *12*(1), 1. https://doi.org/10.3390/sym12010091

Loewenstein, Y., Kuras, A., & Rumpel, S. (2011). Multiplicative Dynamics Underlie the Emergence of the Log-Normal Distribution of Spine Sizes in the Neocortex In Vivo. *Journal of Neuroscience*, *31*(26), 9481–9488. https://doi.org/10.1523/JNEUROSCI.6130-10.2011

Lord, L.-D., Stevner, A. B., Deco, G., & Kringelbach, M. L. (2017). Understanding principles of integration and segregation using whole-brain computational connectomics: Implications for neuropsychiatric disorders. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *375*(2096), 20160283. https://doi.org/10.1098/rsta.2016.0283

Luppi, A. I., Mediano, P. A. M., Rosas, F. E., Holland, N., Fryer, T. D., O'Brien, J. T., Rowe, J. B., Menon, D. K., Bor, D., & Stamatakis, E. A. (2022). A synergistic core for human brain evolution and cognition. *Nature Neuroscience*, *25*(6), 6. https://doi.org/10.1038/s41593-022-01070-0

Magrans de Abril, I., Yoshimoto, J., & Doya, K. (2018). Connectivity inference from neural recording data: Challenges, mathematical bases and research directions. *Neural Networks*, *102*, 120–137. https://doi.org/10.1016/j.neunet.2018.02.016

Mahon, B. Z., & Cantlon, J. F. (2011). The Specialization of Function: Cognitive and Neural Perspectives. *Cognitive Neuropsychology*, *28*(0), 147–155. https://doi.org/10.1080/02643294.2011.633504

Malkov, Y. A., & Ponomarenko, A. (2016). Growing Homophilic Networks Are Natural Navigable Small Worlds. *PLOS ONE*, *11*(6), e0158162. https://doi.org/10.1371/journal.pone.0158162

Manjón, J. V., Thacker, N. A., Lull, J. J., Garcia-Martí, G., Martí-Bonmatí, L., & Robles, M. (2009). Multicomponent MR Image Denoising. *International Journal of Biomedical Imaging*, *2009*, e756897. https://doi.org/10.1155/2009/756897

Mann, E. O., Kohl, M. M., & Paulsen, O. (2009). Distinct Roles of GABAA and GABAB Receptors in Balancing and Terminating Persistent Cortical Activity. *Journal of Neuroscience*, *29*(23), 7513–7518. https://doi.org/10.1523/JNEUROSCI.6162-08.2009

Mareschal, D., Johnson, M. H., Sirois, S., Spratling, M., Thomas, M. S. C., & Westermann, G. (2007). *Neuroconstructivism - I: How the brain constructs cognition*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198529910.001.0001

Mareschal, D., Plunkett, K., & Harris, P. (1999). A computational and neuropsychological account of object-oriented behaviours in infancy. *Developmental Science*, *2*(3), 306–317. https://doi.org/10.1111/1467-7687.00076

Mareschal, D., & Shultz, T. R. (1996). Generative connectionist networks and constructivist cognitive development. *Cognitive Development*, *11*(4), 571–603. https://doi.org/10.1016/S0885-2014(96)90018-0

Markov, N. T., Ercsey-Ravasz, M., Van Essen, D. C., Knoblauch, K., Toroczkai, Z., & Kennedy, H. (2013). Cortical high-density counterstream architectures. *Science (New York, N.Y.)*, *342*(6158), 1238406. https://doi.org/10.1126/science.1238406

Marshall, N., Timme, N. M., Bennett, N., Ripp, M., Lautzenhiser, E., & Beggs, J. M. (2016). Analysis of Power Laws, Shape Collapses, and Neural Complexity: New Techniques and MATLAB Support via the NCC Toolbox. *Frontiers in Physiology*, *7*. https://www.frontiersin.org/article/10.3389/fphys.2016.00250

Martin, R. E., Silvers, J. A., Hardi, F., Stephano, T., Helion, C., Insel, C., Franz, P. J., Ninova, E., Lander, J. P., Mischel, W., Casey, B. J., & Ochsner, K. N. (2019). Longitudinal changes in brain structures related to appetitive reactivity and regulation across development. *Developmental Cognitive Neuroscience*, *38*, 100675. https://doi.org/10.1016/j.dcn.2019.100675

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, *27*(1), 415–444. https://doi.org/10.1146/annurev.soc.27.1.415

Meunier, D., Lambiotte, R., & Bullmore, E. (2010). Modular and Hierarchically Modular Organization of Brain Networks. *Frontiers in Neuroscience*, *4*. https://www.frontiersin.org/article/10.3389/fnins.2010.00200

Moons, L., & De Groef, L. (2021). Molecular Mechanisms of Neural Circuit Development and Regeneration. *International Journal of Molecular Sciences*, *22*(9). https://doi.org/10.3390/ijms22094593

Morgan, S. E., Achard, S., Termenon, M., Bullmore, E. T., & Vértes, P. E. (2018). Low-dimensional morphospace of topological motifs in human fMRI brain networks. *Network Neuroscience*, *02*(02), 285–302. https://doi.org/10.1162/netn_a_00038

Mori, S., Crain, B. J., Chacko, V. P., & Van Zijl, P. C. M. (1999). Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Annals of Neurology*, *45*(2), 265–269. https://doi.org/10.1002/1531-8249(199902)45:2<265::AID-ANA21>3.0.CO;2-3

Muldoon, S. F., Bridgeford, E. W., & Bassett, D. S. (2016). Small-World Propensity and Weighted Brain Networks. *Scientific Reports*, *6*(1), 1. https://doi.org/10.1038/srep22057

Müller, E. J., Munn, B. R., & Shine, J. M. (2020). Diffuse neural coupling mediates complex network dynamics through the formation of quasi-critical brain states. *Nature Communications*, *11*(1), 1. https://doi.org/10.1038/s41467-020-19716-7

Müller, J., Ballini, M., Livi, P., Chen, Y., Radivojevic, M., Shadmani, A., Viswam, V., Jones, I. L., Fiscella, M., Diggelmann, R., Stettler, A., Frey, U., Bakkum, D. J., & Hierlemann, A. (2015). High-resolution CMOS MEA platform to study neurons at subcellular, cellular, and network levels. *Lab on a Chip*, *15*(13), 2767–2780. https://doi.org/10.1039/C5LC00133A

Murray, J. D., Demirtaş, M., & Anticevic, A. (2018). Biophysical Modeling of Large-Scale Brain Dynamics and Applications for Computational Psychiatry. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, *3*(9), 777–787. https://doi.org/10.1016/j.bpsc.2018.07.004

Navlakha, S., Barth, A. L., & Bar-Joseph, Z. (2015). Decreasing-Rate Pruning Optimizes the Construction of Efficient and Robust Distributed Networks. *PLoS Computational Biology*, *11*(7), e1004347. https://doi.org/10.1371/journal.pcbi.1004347

Nettelbeck, T., & Wilson, C. (1985). A cross-sequential analysis of developmental differences in speed of visual information processing. *Journal of Experimental Child Psychology*, *40*(1), 1–22. https://doi.org/10.1016/0022-0965(85)90063-3

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, *103*(23), 8577–8582. https://doi.org/10.1073/pnas.0601602103

Nicosia, V., Vértes, P. E., Schafer, W. R., Latora, V., & Bullmore, E. T. (2013). Phase transition in the economically modeled growth of a cellular nervous system. *Proceedings of the National Academy of Sciences*, *110*(19), 7880–7885. https://doi.org/10.1073/pnas.1300753110

Novelli, L., Wollstadt, P., Mediano, P., Wibral, M., & Lizier, J. T. (2019). Large-scale directed network inference with multivariate transfer entropy and hierarchical statistical testing. *Network Neuroscience*, *3*(3), 827–847. https://doi.org/10.1162/netn_a_00092

O'Byrne, J., & Jerbi, K. (2022). How critical is brain criticality? *Trends in Neurosciences*, *45*(11), 820–837. https://doi.org/10.1016/j.tins.2022.08.007

Oldham, S., & Fornito, A. (2019). The development of brain network hubs. *Developmental Cognitive Neuroscience*, *36*, 100607. https://doi.org/10.1016/j.dcn.2018.12.005

Oldham, S., Fulcher, B. D., Aquino, K., Arnatkevičiūtė, A., Paquola, C., Shishegar, R., & Fornito, A. (2022). Modeling spatial, developmental, physiological, and topological constraints on human brain connectivity. *Science Advances*, *8*(22). https://doi.org/10.1126/sciadv.abm612

Oliver, A., Johnson, M. H., Karmiloff-Smith, A., & Pennington, B. (2000). Deviations in the emergence of representations: A neuroconstructivist framework for analysing developmental disorders. *Developmental Science*, *3*(1), 1–23. https://doi.org/10.1111/1467-7687.00094

Pachitariu, M., Steinmetz, N. A., Kadir, S. N., Carandini, M., & Harris, K. D. (2016). Fast and accurate spike sorting of high-channel count probes with KiloSort. *Advances in Neural Information Processing Systems*, *29*. https://papers.nips.cc/paper/2016/hash/1145a30ff80745b56fb0cecf65305017-Abstract.html

Park, H.-J., & Friston, K. (2013). Structural and Functional Brain Networks: From Connections to Cognition. *Science*, *342*(6158), 1238411. https://doi.org/10.1126/science.1238411

Perez-Nieves, N., Leung, V. C. H., Dragotti, P. L., & Goodman, D. F. M. (2021). Neural heterogeneity promotes robust learning. *Nature Communications*, *12*(1), 1. https://doi.org/10.1038/s41467-021-26022-3

Pessoa, L. (2014). Understanding brain networks and brain organization. *Physics of Life Reviews*, *11*(3), 400–435. https://doi.org/10.1016/j.plrev.2014.03.005

Popper, K. R. (1959). *The logic of scientific discovery*. Basic Books. http://nukweb.nuk.uni-lj.si/login?url=http://search.ebscohost.com/login.aspx?authtype=ip&direct=true&db=nlebk&AN=143035&site=eds-live&scope=site&lang=sl

Quartz, S. R., & Sejnowski, T. J. (1997). The neural basis of cognitive development: A constructivist manifesto. *The Behavioral and Brain Sciences*, *20*(4), 537–556; discussion 556-596. https://doi.org/10.1017/s0140525x97001581

Reimand, J., Isserlin, R., Voisin, V., Kucera, M., Tannus-Lopes, C., Rostamianfar, A., Wadi, L., Meyer, M., Wong, J., Xu, C., Merico, D., & Bader, G. D. (2019). Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nature Protocols*, *14*(2), 482–517. https://doi.org/10.1038/s41596-018-0103-9

Reimand, J., Kull, M., Peterson, H., Hansen, J., & Vilo, J. (2007). g:Profiler—A web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Research*, *35*(Web Server issue), W193-200. https://doi.org/10.1093/nar/gkm226

Resnik, D. B., & Vorhaus, D. B. (2006). Genetic modification and genetic determinism. *Philosophy, Ethics, and Humanities in Medicine*, *1*, 9. https://doi.org/10.1186/1747-5341-1-9

Rigotti, M., Barak, O., Warden, M. R., Wang, X.-J., Daw, N. D., Miller, E. K., & Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. *Nature*, *497*(7451), 7451. https://doi.org/10.1038/nature12160

Ripke, S., Neale, B. M., Corvin, A., Walters, J. T., Farh, K.-H., Holmans, P. A., Lee, P., Bulik-Sullivan, B., Collier, D. A., Huang, H., Pers, T. H., Agartz, I., Agerbo, E., Albus, M., Alexander, M., Amin, F., Bacanu, S. A., Begemann, M., Belliveau, R. A., … O'Donovan, M. C. (2014). Biological Insights From 108 Schizophrenia-Associated Genetic Loci. *Nature*, *511*(7510), 421–427. https://doi.org/10.1038/nature13595

Ritter, P., Schirner, M., McIntosh, A. R., & Jirsa, V. K. (2013). The virtual brain integrates computational modeling and multimodal neuroimaging. *Brain Connectivity*, *3*(2), 121–145. https://doi.org/10.1089/brain.2012.0120

Roccasalvo, I. M., Micera, S., & Sergi, P. N. (2015). A hybrid computational model to predict chemotactic guidance of growth cones. *Scientific Reports*, *5*(1), 11340. https://doi.org/10.1038/srep11340

Rogers, S. J., & Williams, J. H. G. (2006). Imitation and the Social Mind: Autism and Typical Development. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, *17*(2), 91–93.

Roine, U., Roine, T., Salmi, J., Nieminen-von Wendt, T., Tani, P., Leppämäki, S., Rintahaka, P., Caeyenberghs, K., Leemans, A., & Sams, M. (2015). Abnormal wiring of the connectome in adults with high-functioning autism spectrum disorder. *Molecular Autism*, *6*(1), 65. https://doi.org/10.1186/s13229-015-0058-4

Romero-Garcia, R., Seidlitz, J., Whitaker, K. J., Morgan, S. E., Fonagy, P., Dolan, R. J., Jones, P. B., Goodyer, I. M., Suckling, J., Bullmore, E., Dolan, R., Goodyer, I., Fonagy, P., Jones, P., Vaghi, M., Moutoussis, M., Hauser, T., Neufeld, S., Romero-Garcia, R., … Bullmore, E. T. (2020). Schizotypy-Related Magnetization of Cortex in Healthy Adolescence Is Colocated With Expression of Schizophrenia-Related Genes. *Biological Psychiatry*, *88*(3), 248–259. https://doi.org/10.1016/j.biopsych.2019.12.005

Ronchi, S., Buccino, A. P., Prack, G., Kumar, S. S., Schröter, M., Fiscella, M., & Hierlemann, A. (2021). Electrophysiological Phenotype Characterization of Human iPSC-Derived Neuronal Cell Lines by Means of High-Density Microelectrode Arrays. *Advanced Biology*, *5*(3), e2000223. https://doi.org/10.1002/adbi.202000223

Rosen, B. Q., & Halgren, E. (2022). An estimation of the absolute number of axons indicates that human cortical areas are sparsely connected. *PLOS Biology*, *20*(3), e3001575. https://doi.org/10.1371/journal.pbio.3001575

Rubinov, M. (2016). Constraints and spandrels of interareal connectomes. *Nature Communications*, *7*(1), 1. https://doi.org/10.1038/ncomms13812

Rubinov, M. (2022). *Circular and unified analysis in network neuroscience*. OSF Preprints. https://doi.org/10.31219/osf.io/mdqak

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, *52*(3), 1059–1069. https://doi.org/10.1016/j.neuroimage.2009.10.003

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536. https://doi.org/10.1038/323533a0

Rumelhart, D. E., McClelland, J. L., & Group, P. R. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations* (Vol. 1). A Bradford Book.

Russell, G., Rodgers, L. R., Ukoumunne, O. C., & Ford, T. (2014). Prevalence of Parent-Reported ASD and ADHD in the UK: Findings from the Millennium Cohort Study. *Journal of Autism and Developmental Disorders*, *44*(1), 31–40. https://doi.org/10.1007/s10803-013-1849-0

Sanes, J. R., & Zipursky, S. L. (2020). Synaptic Specificity, Recognition Molecules, and Assembly of Neural Circuits. *Cell*, *181*(3), 536–556. https://doi.org/10.1016/j.cell.2020.04.008

Saupe, E. E., Myers, C. E., Townsend Peterson, A., Soberón, J., Singarayer, J., Valdes, P., & Qiao, H. (2019). Spatio-temporal climate change contributes to latitudinal diversity gradients. *Nature Ecology & Evolution*, *3*(10), 10. https://doi.org/10.1038/s41559-019-0962-7

Schiavi, S., Ocampo-Pineda, M., Barakovic, M., Petit, L., Descoteaux, M., Thiran, J.-P., & Daducci, A. (2020). A new method for accurate in vivo mapping of human brain

connections using microstructural and anatomical information. *Science Advances*, *6*(31), eaba8245. https://doi.org/10.1126/sciadv.aba8245

Schröeter, 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. *Journal of Neuroscience*, *35*(14), 5459–5470. https://doi.org/10.1523/JNEUROSCI.4259-14.2015

Schröter, M., Paulsen, O., & Bullmore, E. T. (2017). Micro-connectomics: Probing the organization of neuronal networks at the cellular scale. *Nature Reviews. Neuroscience*, *18*(3), 131–146. https://doi.org/10.1038/nrn.2016.182

Seguin, C., Heuvel, M. P. van den, & Zalesky, A. (2018). Navigation of brain networks. *Proceedings of the National Academy of Sciences*, *115*(24), 6297–6302. https://doi.org/10.1073/pnas.1801351115

Seguin, C., Mansour L, S., Sporns, O., Zalesky, A., & Calamante, F. (2022). Network communication models narrow the gap between the modular organization of structural and functional brain networks. *NeuroImage*, *257*, 119323. https://doi.org/10.1016/j.neuroimage.2022.119323

Seguin, C., Tian, Y., & Zalesky, A. (2020). Network communication models improve the behavioral and functional predictive utility of the human structural connectome. *Network Neuroscience*, *4*(4), 980–1006. https://doi.org/10.1162/netn_a_00161

Seidlitz, J., Váša, F., Shinn, M., Romero-Garcia, R., Whitaker, K. J., Vértes, P. E., Wagstyl, K., Kirkpatrick Reardon, P., Clasen, L., Liu, S., Messinger, A., Leopold, D. A., Fonagy, P., Dolan, R. J., Jones, P. B., Goodyer, I. M., Raznahan, A., & Bullmore, E. T. (2018). Morphometric Similarity Networks Detect Microscale Cortical Organization and Predict Inter-Individual Cognitive Variation. *Neuron*, *97*(1), 231-247.e7. https://doi.org/10.1016/j.neuron.2017.11.039

Sengupta, B., Stemmler, M., Laughlin, S. B., & Niven, J. E. (2010). Action Potential Energy Efficiency Varies Among Neuron Types in Vertebrates and Invertebrates. *PLOS Computational Biology*, *6*(7), e1000840. https://doi.org/10.1371/journal.pcbi.1000840

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research*, *13*(11), 2498–2504. https://doi.org/10.1101/gr.1239303

Shorten, D. P., Spinney, R. E., & Lizier, J. T. (2021). Estimating Transfer Entropy in Continuous Time Between Neural Spike Trains or Other Event-Based Data. *PLOS Computational Biology*, *17*(4), e1008054. https://doi.org/10.1371/journal.pcbi.1008054

Siugzdaite, R., Bathelt, J., Holmes, J., & Astle, D. E. (2020). Transdiagnostic Brain Mapping in Developmental Disorders. *Current Biology*, *30*(7), 1245-1257.e4. https://doi.org/10.1016/j.cub.2020.01.078

Son, J.-Y., Bhandari, A., & FeldmanHall, O. (2021). Cognitive maps of social features enable flexible inference in social networks. *Proceedings of the National Academy of Sciences*, *118*(39). https://doi.org/10.1073/pnas.2021699118

Song, H. F., Kennedy, H., & Wang, X.-J. (2014). Spatial embedding of structural similarity in the cerebral cortex. *Proceedings of the National Academy of Sciences*, *111*(46), 16580–16585. https://doi.org/10.1073/pnas.1414153111

Song, S., Miller, K. D., & Abbott, L. F. (2000). Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nature Neuroscience*, *3*(9), 919–926. https://doi.org/10.1038/78829

Sperry, M. M., Telesford, Q. K., Klimm, F., & Bassett, D. S. (2017). Rentian scaling for the measurement of optimal embedding of complex networks into physical space. *Journal of Complex Networks*, *5*(2), 199–218. https://doi.org/10.1093/comnet/cnw010

Sporns, O. (2013a). Network attributes for segregation and integration in the human brain. *Current Opinion in Neurobiology*, *23*(2), 162–171. https://doi.org/10.1016/j.conb.2012.11.015

Sporns, O. (2013b). Structure and function of complex brain networks. *Dialogues in Clinical Neuroscience*, *15*(3), 247–262.

Sporns, O., & Betzel, R. F. (2016). Modular Brain Networks. *Annual Review of Psychology*, *67*, 613–640. https://doi.org/10.1146/annurev-psych-122414-033634

Sporns, O., & Kötter, R. (2004). Motifs in Brain Networks. *PLOS Biology*, *2*(11), e369. https://doi.org/10.1371/journal.pbio.0020369

Sporns, O., & Zwi, J. D. (2004). The small world of the cerebral cortex. *Neuroinformatics*, *2*(2), 145–162. https://doi.org/10.1385/NI:2:2:145

Srivastava, P., Nozari, E., Kim, J. Z., Ju, H., Zhou, D., Becker, C., Pasqualetti, F., Pappas, G. J., & Bassett, D. S. (2020). Models of communication and control for brain networks: Distinctions, convergence, and future outlook. *Network Neuroscience*, *4*(4), 1122–1159. https://doi.org/10.1162/netn_a_00158

Sternberg, R. J. (1983). Components of human intelligence. *Cognition*, *15*(1), 1–48. https://doi.org/10.1016/0010-0277(83)90032-X

Stöckl, C., Lang, D., & Maass, W. (2021). *Structure induces computational function in networks with diverse types of spiking neurons* [Preprint]. Neuroscience. https://doi.org/10.1101/2021.05.18.444689

Stokes, M. G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., & Duncan, J. (2013). Dynamic Coding for Cognitive Control in Prefrontal Cortex. *Neuron*, *78*(2), 364–375. https://doi.org/10.1016/j.neuron.2013.01.039

Stroud, J. P., Watanabe, K., Suzuki, T., Stokes, M. G., & Lengyel, M. (2021). *Optimal information loading into working memory in prefrontal cortex* [Preprint]. Neuroscience. https://doi.org/10.1101/2021.11.16.468360

Suárez, L. E., Richards, B. A., Lajoie, G., & Misic, B. (2021). Learning function from structure in neuromorphic networks. *Nature Machine Intelligence*, *3*(9), 9. https://doi.org/10.1038/s42256-021-00376-1

Südhof, T. C. (2018). Towards an Understanding of Synapse Formation. *Neuron*, *100*(2), 276–293. https://doi.org/10.1016/j.neuron.2018.09.040

Symmonds, M., Moran, C. H., Leite, M. I., Buckley, C., Irani, S. R., Stephan, K. E., Friston, K. J., & Moran, R. J. (2018). Ion channels in EEG: Isolating channel dysfunction in

NMDA receptor antibody encephalitis. *Brain: A Journal of Neurology*, *141*(6), 1691–1702. https://doi.org/10.1093/brain/awy107

Szebényi, K., Wenger, L. M. D., Sun, Y., Dunn, A. W. E., Limegrover, C. A., Gibbons, G. M., Conci, E., Paulsen, O., Mierau, S. B., Balmus, G., & Lakatos, A. (2021). Human ALS/FTD brain organoid slice cultures display distinct early astrocyte and targetable neuronal pathology. *Nature Neuroscience*, *24*(11), 1542–1554. https://doi.org/10.1038/s41593-021-00923-4

Takemura, S., Xu, C. S., Lu, Z., Rivlin, P. K., Parag, T., Olbris, D. J., Plaza, S., Zhao, T., Katz, W. T., Umayam, L., Weaver, C., Hess, H. F., Horne, J. A., Nunez-Iglesias, J., Aniceto, R., Chang, L.-A., Lauchie, S., Nasca, A., Ogundeyi, O., … Scheffer, L. K. (2015). Synaptic circuits and their variations within different columns in the visual system of Drosophila. *Proceedings of the National Academy of Sciences*, *112*(44), 13711–13716. https://doi.org/10.1073/pnas.1509820112

Talaga, S., & Nowak, A. (2020). Homophily as a Process Generating Social Networks: Insights from Social Distance Attachment Model. *Journal of Artificial Societies and Social Simulation*, *23*(2), 6. https://doi.org/10.18564/jasss.4252

Tamnes, C. K., Herting, M. M., Goddings, A.-L., Meuwese, R., Blakemore, S.-J., Dahl, R. E., Güroğlu, B., Raznahan, A., Sowell, E. R., Crone, E. A., & Mills, K. L. (2017). Development of the Cerebral Cortex across Adolescence: A Multisample Study of Inter-Related Longitudinal Changes in Cortical Volume, Surface Area, and Thickness. *Journal of Neuroscience*, *37*(12), 3402–3412. https://doi.org/10.1523/JNEUROSCI.3302-16.2017

Taylor, S. R., Santpere, G., Weinreb, A., Barrett, A., Reilly, M. B., Xu, C., Varol, E., Oikonomou, P., Glenwinkel, L., McWhirter, R., Poff, A., Basavaraju, M., Rafi, I., Yemini, E., Cook, S. J., Abrams, A., Vidal, B., Cros, C., Tavazoie, S., … Miller, D. M. (2021). Molecular topography of an entire nervous system. *Cell*, *184*(16), 4329-4347.e23. https://doi.org/10.1016/j.cell.2021.06.023

Thomas, M., & Karmiloff-Smith, A. (2002). Are developmental disorders like cases of adult brain damage? Implications from connectionist modelling. *The Behavioral and Brain Sciences*, *25*(6), 727–750; discussion 750-787. https://doi.org/10.1017/s0140525x02000134

Thomas, M., & Karmiloff-Smith, A. (2003). Connectionist models of development, developmental disorders, and individual differences. In *Models of intelligence: International perspectives* (pp. 133–150). American Psychological Association.

Tian, Y., Margulies, D. S., Breakspear, M., & Zalesky, A. (2020). Topographic organization of the human subcortex unveiled with functional connectivity gradients. *Nature Neuroscience*, *23*(11), 11. https://doi.org/10.1038/s41593-020-00711-6

Tooley, U. A., Bassett, D. S., & Mackey, A. P. (2021). Environmental influences on the pace of brain development. *Nature Reviews Neuroscience*, *22*(6), 372–384. https://doi.org/10.1038/s41583-021-00457-5

Towlson, E. K., Vértes, P. E., Ahnert, S. E., Schafer, W. R., & Bullmore, E. T. (2013). The Rich Club of the C. elegans Neuronal Connectome. *Journal of Neuroscience*, *33*(15), 6380–6387. https://doi.org/10.1523/JNEUROSCI.3784-12.2013

Tymofiyeva, O., Hess, C. P., Xu, D., & Barkovich, A. J. (2014). Structural MRI connectome in development: Challenges of the changing brain. *The British Journal of Radiology*, *87*(1039), 20140086. https://doi.org/10.1259/bjr.20140086

van den Heuvel, M. P., Sporns, O., Collin, G., Scheewe, T., Mandl, R. C. W., Cahn, W., Goñi, J., Hulshoff Pol, H. E., & Kahn, R. S. (2013). Abnormal Rich Club Organization and Functional Brain Dynamics in Schizophrenia. *JAMA Psychiatry*, *70*(8), 783–792. https://doi.org/10.1001/jamapsychiatry.2013.1328

Váša, F., Romero-Garcia, R., Kitzbichler, M. G., Seidlitz, J., Whitaker, K. J., Vaghi, M. M., Kundu, P., Patel, A. X., Fonagy, P., Dolan, R. J., Jones, P. B., Goodyer, I. M., the NSPN Consortium, Vértes, P. E., Bullmore, E. T., Bullmore, E., Dolan, R., Goodyer, I., Fonagy, P., … Kievit, R. (2020). Conservative and disruptive modes of adolescent change in human brain functional connectivity. *Proceedings of the National Academy of Sciences*, *117*(6), 3248–3253. https://doi.org/10.1073/pnas.1906144117

Vértes, P. E., Alexander-Bloch, A., & Bullmore, E. T. (2014). Generative models of rich clubs in Hebbian neuronal networks and large-scale human brain networks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1653), 20130531. https://doi.org/10.1098/rstb.2013.0531

Vértes, P. E., Alexander-Bloch, A. F., Gogtay, N., Giedd, J. N., Rapoport, J. L., & Bullmore, E. T. (2012). Simple models of human brain functional networks. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(15), 5868–5873. https://doi.org/10.1073/pnas.1111738109

Vértes, P. E., & Bullmore, E. T. (2015). Annual Research Review: Growth connectomics – the organization and reorganization of brain networks during normal and abnormal development. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *56*(3), 299–320. https://doi.org/10.1111/jcpp.12365

Vitrac, A., & Cloëz-Tayarani, I. (2018). Induced pluripotent stem cells as a tool to study brain circuits in autism-related disorders. *Stem Cell Research & Therapy*, *9*, 226. https://doi.org/10.1186/s13287-018-0966-2

Vivanti, G., & Rogers, S. J. (2014). Autism and the mirror neuron system: Insights from learning and teaching. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1644), 20130184. https://doi.org/10.1098/rstb.2013.0184

Vogt, G. (2015). Stochastic developmental variation, an epigenetic source of phenotypic diversity with far-reaching biological consequences. *Journal of Biosciences*, *40*(1), 159–204. https://doi.org/10.1007/s12038-015-9506-8

Vydrova, R., Komarek, V., Sanda, J., Sterbova, K., Jahodova, A., Maulisova, A., Zackova, J., Reissigova, J., Krsek, P., & Kyncl, M. (2015). Structural alterations of the language connectome in children with specific language impairment. *Brain and Language*, *151*, 35–41. https://doi.org/10.1016/j.bandl.2015.10.003

Waddington, C. H. (1957). *The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology. With an Appendix by H. Kacser*. Allen & Unwin.

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), 11. https://doi.org/10.1186/1471-2202-7-11

Waggoner, M. R., & Uller, T. (2015). Epigenetic Determinism in Science and Society. *New Genetics and Society*, *34*(2), 177–195. https://doi.org/10.1080/14636778.2015.1033052

Wallach, A., Melanson, A., Longtin, A., & Maler, L. (2022). Mixed selectivity coding of sensory and motor social signals in the thalamus of a weakly electric fish. *Current Biology*, *32*(1), 51-63.e3. https://doi.org/10.1016/j.cub.2021.10.034

Wang, J. X., Kurth-Nelson, Z., Kumaran, D., Tirumala, D., Soyer, H., Leibo, J. Z., Hassabis, D., & Botvinick, M. (2018). Prefrontal cortex as a meta-reinforcement learning system. *Nature Neuroscience*, *21*(6), 860–868. https://doi.org/10.1038/s41593-018-0147-8

Waskom, M. L., & Wagner, A. D. (2017). Distributed representation of context by intrinsic subnetworks in prefrontal cortex. *Proceedings of the National Academy of Sciences*, *114*(8), 2030–2035. https://doi.org/10.1073/pnas.1615269114

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, *393*(6684), 440–442. https://doi.org/10.1038/30918

Wechsler, D. (1996). *Wechsler Objective Numerical Dimensions*. Psychology Corporation.

Wechsler, D. (2005). *Wechsler Individual Achievement Test* (2nd ed.). Pearson Assessment.

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence* (2nd ed.). Pearson Assessment.

Wei, Y., Liao, X., Yan, C., He, Y., & Xia, M. (2017). Identifying topological motif patterns of human brain functional networks. *Human Brain Mapping*, *38*(5), 2734–2750. https://doi.org/10.1002/hbm.23557

Whitaker, K. J., Vértes, P. E., Romero-Garcia, R., Váša, F., Moutoussis, M., Prabhu, G., Weiskopf, N., Callaghan, M. F., Wagstyl, K., Rittman, T., Tait, R., Ooi, C., Suckling, J., Inkster, B., Fonagy, P., Dolan, R. J., Jones, P. B., Goodyer, I. M., the NSPN Consortium, & Bullmore, E. T. (2016). Adolescence is associated with genomically patterned consolidation of the hubs of the human brain connectome. *Proceedings of the National Academy of Sciences*, *113*(32), 9105–9110. https://doi.org/10.1073/pnas.1601745113

Wig, G. S. (2017). Segregated Systems of Human Brain Networks. *Trends in Cognitive Sciences*, *21*(12), 981–996. https://doi.org/10.1016/j.tics.2017.09.006

Witvliet, D., Mulcahy, B., Mitchell, J. K., Meirovitch, Y., Berger, D. R., Wu, Y., Liu, Y., Koh, W. X., Parvathala, R., Holmyard, D., Schalek, R. L., Shavit, N., Chisholm, A. D., Lichtman, J. W., Samuel, A. D. T., & Zhen, M. (2021). Connectomes across development reveal principles of brain maturation. *Nature*, *596*(7871), 7871. https://doi.org/10.1038/s41586-021-03778-8

Wong, R. O., Meister, M., & Shatz, C. J. (1993). Transient period of correlated bursting activity during development of the mammalian retina. *Neuron*, *11*(5), 923–938. https://doi.org/10.1016/0896-6273(93)90122-8

Woodcock, R. W., Mather, N., McGrew, K. S., & Wendling, B. J. (2001). *Woodcock-Johnson III tests of cognitive abilities*. Riverside Publishing Company.

Yang, G., Pan, F., & Gan, W.-B. (2009). Stably maintained dendritic spines are associated with lifelong memories. *Nature*, *462*(7275), 920–924. https://doi.org/10.1038/nature08577

Yang, G. R., Joglekar, M. R., Song, H. F., Newsome, W. T., & Wang, X.-J. (2019). Task representations in neural networks trained to perform many cognitive tasks. *Nature Neuroscience*, *22*(2), 2. https://doi.org/10.1038/s41593-018-0310-2

Yogev, S., & Shen, K. (2014). Cellular and molecular mechanisms of synaptic specificity. *Annual Review of Cell and Developmental Biology*, *30*, 417–437. https://doi.org/10.1146/annurev-cellbio-100913-012953

Zador, A. M. (2019). A critique of pure learning and what artificial neural networks can learn from animal brains. *Nature Communications*, *10*(1), 1. https://doi.org/10.1038/s41467-019-11786-6

Zamora-López, G., Zhou, C., & Kurths, J. (2010). Cortical Hubs Form a Module for Multisensory Integration on Top of the Hierarchy of Cortical Networks. *Frontiers in Neuroinformatics*, *4*, 1. https://doi.org/10.3389/neuro.11.001.2010

Zhang, X., Braun, U., Harneit, A., Zang, Z., Geiger, L. S., Betzel, R. F., Chen, J., Schweiger, J. I., Schwarz, K., Reinwald, J. R., Fritze, S., Witt, S., Rietschel, M., Nöthen, M. M., Degenhardt, F., Schwarz, E., Hirjak, D., Meyer-Lindenberg, A., Bassett, D. S., & Tost, H. (2021). Generative network models of altered structural brain connectivity in schizophrenia. *NeuroImage*, *225*, 117510. https://doi.org/10.1016/j.neuroimage.2020.117510

Zheng, M., Allard, A., Hagmann, P., Alemán-Gómez, Y., & Serrano, M. Á. (2020). Geometric renormalization unravels self-similarity of the multiscale human connectome. *Proceedings of the National Academy of Sciences*, *117*(33), 20244–20253. https://doi.org/10.1073/pnas.1922248117

Zhou, D., Lynn, C. W., Cui, Z., Ciric, R., Baum, G. L., Moore, T. M., Roalf, D. R., Detre, J. A., Gur, R. C., Gur, R. E., Satterthwaite, T. D., & Bassett, D. S. (2022). Efficient coding in the economics of human brain connectomics. *Network Neuroscience*, 1–41. https://doi.org/10.1162/netn_a_00223