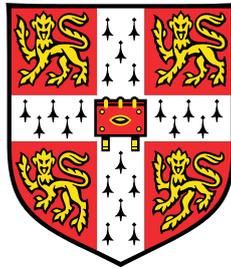


---

CONNECTOMICS OF  
EXTRASYNAPTIC SIGNALLING:  
APPLICATIONS TO THE NERVOUS SYSTEM OF  
CAENORHABDITIS ELEGANS

---



**Barry Bentley**

MRC Laboratory of Molecular Biology  
University of Cambridge

This dissertation is submitted for the degree of  
*Doctor of Philosophy*

Magdalene College

November 2017



*Dedicated to all those who have supported,  
inspired & educated me*



With ideas it is like with dizzy heights you climb: At first they cause you discomfort and you are anxious to get down, distrustful of your own powers; but soon the remoteness of the turmoil of life and the inspiring influence of the altitude calm your blood; your step gets firm and sure and you begin to look – for dizzier heights.

NIKOLA TESLA (1897)



## Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 60,000 words excluding appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

Barry Bentley  
November 2017



## Acknowledgements

Firstly, I would like to thank Dr William R. Schafer for giving me the opportunity to work in his lab, and for his continual support and guidance throughout the course of my PhD. I could not have hoped for a more patient and understanding supervisor, for which I am eternally grateful. Secondly, I would like to thank Dr Petra Vértés for her mentorship. Dr Vértés introduced me to the world of network science, and was pivotal in motivating me to continue my research through many of the difficulties I encountered. Without the support of Bill & Petra the work presented in this thesis would not have been possible. I am also deeply indebted to my second supervisor, Dr M. Madan Babu, and the LMB postgraduate administrator, Amie Blake, who both provided much needed support and motivation early on in my PhD.

I acknowledge the support of the Medical Research Council (MRC) who funded this research; Prof. Scott W. Emmons for the WormWiring connectome data, the analysis of which forms the basis of Chapter 3; Dr Isabel Beets for sharing her knowledge on neuropeptides, which aided in the construction of the network in Chapter 5; Elizabeth Dorey for providing feedback on an early draft of my thesis; and Prof. Edward T. Bullmore for his input and guidance on all of the work presented here. A special thanks is due to Zoric Fieldhouse for his encouragement and proofreading.

Chapters 4 & 5 are based on the paper “*The multilayer connectome of *Caenorhabditis elegans**” (Bentley et al., 2016), which was co-authored with Dr Robyn Branicky, Chris L. Barnes, Dr Yee Lian Chew, Dr Eviatar Yemini, Prof. Edward T. Bullmore, Dr Petra E. Vértés & Dr William R. Schafer. Hive plots were provided courtesy of Chris L. Barnes. Expression analysis of the dopamine receptors was performed by Drs Schafer, Branicky & Chew, and is detailed in Appendix C.

Many persons have contributed to my time at Cambridge and the LMB: Dr Marios Chatzigeorgiou, Dr Robyn Branicky, Dr Denise Walker, and Dr Ruma Raha-Chowdhury all played important roles in teaching me molecular & nematode biology, for which I

am deeply grateful; Dr André Brown, Dr Eviatar Yemini, and Dr Buyun Zhao also provided constructive feedback on much of my early work. I would especially like to thank Marios for helping me find my way around the lab when I first started.

I would also like to express my gratitude to Dr Aubrey de Grey and Dr Stuart R. G. Calimport for their insightful discussions and support at various points; and the many teachers who have been pivotal in developing my thinking throughout the years, notably Prof. Angelo Cangelosi for welcoming me into his lab and providing my first experiences in research; Andrew Chalmers and John Proctor for their broad and inspiring lectures in systems engineering, which continue to deeply influence all of my work; and my school teachers Andrew Sears and Angela Smith, who nurtured my earliest interests in science and technology.

Lastly, I would like to thank my parents, my family, and Tori Darcy, for their unreserved love and support in all of my endeavours.

## Abstract

Connectomics – the study of neural connectivity – is primarily concerned with the mapping and characterisation of wired synaptic links; however, it is well established that long-distance chemical signalling via extrasynaptic volume transmission is also critical to brain function. As these interactions are not visible in the physical structure of the nervous system, current approaches to connectomics are unable to capture them.

This work addresses the problem of missing extrasynaptic interactions by demonstrating for the first time that whole-animal volume transmission networks can be mapped from gene expression and ligand-receptor interaction data, and analysed as part of the connectome. Complete networks are presented for the monoamine systems of *Caenorhabditis elegans*, along with a representative sample of selected neuropeptide systems.

A network analysis of the synaptic (*wired*) and extrasynaptic (*wireless*) connectomes is presented which reveals complex topological properties, including extrasynaptic rich-club organisation with interconnected hubs distinct from those in the synaptic and gap junction networks, and highly significant multilink motifs pinpointing locations in the network where aminergic and neuropeptide signalling is likely to modulate synaptic activity. Thus, the neuronal connectome can be modelled as a multiplex network with synaptic, gap junction, and neuromodulatory layers representing inter-neuronal interactions with different dynamics and polarity. This represents a prototype for understanding how extrasynaptic signalling can be integrated into connectomics research, and provides a novel dataset for the development of multilayer network algorithms.



# Table of contents

<b>List of figures</b>	<b>xvii</b>
<b>List of tables</b>	<b>xxiii</b>
<b>Nomenclature</b>	<b>xxix</b>
<b>I Introduction</b>	<b>1</b>
<b>1 Introduction</b>	<b>3</b>
1.1 Connectomics . . . . .	3
1.1.1 The role of connections . . . . .	4
1.1.2 The state of connectomics . . . . .	6
1.1.3 Non-synaptic interactions . . . . .	8
1.2 Thesis overview . . . . .	11
1.2.1 Project background . . . . .	11
1.2.2 Research objectives . . . . .	12
1.3 Thesis structure & scientific contributions . . . . .	12
<b>2 Complex network theory in connectomics</b>	<b>15</b>
2.1 Introduction . . . . .	15
2.1.1 Network theory and the brain . . . . .	16
2.2 Network concepts . . . . .	17
2.2.1 Connectivity . . . . .	17
2.2.2 Higher-order structural organisation . . . . .	21
2.2.3 Null model reference networks . . . . .	27
2.2.4 Network robustness . . . . .	29
2.2.5 Multilayer networks . . . . .	30
2.3 Foundations & formal definitions . . . . .	32

2.3.1	Graphs . . . . .	33
2.3.2	Adjacency matrices . . . . .	34
2.3.3	Weighted graphs . . . . .	34
2.3.4	Simple graphs & maximum edges . . . . .	35
2.3.5	Node neighbourhood . . . . .	36
2.4	Network measures . . . . .	36
2.4.1	Degree . . . . .	36
2.4.2	Network density . . . . .	37
2.4.3	Connected components . . . . .	38
2.4.4	Path length & efficiency . . . . .	39
2.4.5	Clustering . . . . .	42
2.4.6	Modularity . . . . .	44
2.4.7	Reciprocity . . . . .	48
2.4.8	Assortativity . . . . .	49
2.4.9	Robustness . . . . .	50
2.4.10	Null model networks . . . . .	50
2.4.11	Small-world property . . . . .	60
2.4.12	Scale-free property . . . . .	62
2.4.13	Rich-club organisation . . . . .	64
2.4.14	Network motifs . . . . .	67
2.5	Multilayer networks . . . . .	70
2.5.1	Multilayer reducibility . . . . .	70
2.5.2	Multilink motifs . . . . .	79
2.5.3	Layer interaction . . . . .	80
2.6	Software tools . . . . .	81

## **II The wired connectome 83**

### **3 Comparing the wired connectomes of *C. elegans* 85**

3.1	Introduction . . . . .	85
3.1.1	Chapter overview . . . . .	87
3.2	Materials & Methods . . . . .	87
3.2.1	Connectivity data . . . . .	87
3.2.2	Topological network measures . . . . .	88
3.2.3	Software . . . . .	88
3.2.4	Literature search . . . . .	88

3.3	Results . . . . .	89
3.3.1	Connections omitted in WW . . . . .	92
3.3.2	Asymmetric gap junctions . . . . .	92
3.3.3	Network measures . . . . .	95
3.3.4	Scale-free & small-world properties . . . . .	97
3.3.5	Robustness . . . . .	97
3.3.6	Reciprocity . . . . .	100
3.3.7	Rich-club organisation . . . . .	101
3.4	Discussion . . . . .	104
3.4.1	Synaptic networks . . . . .	104
3.4.2	Gap junction networks . . . . .	105
3.4.3	Validation & future work . . . . .	105
3.4.4	Conclusions . . . . .	107
<b>III</b>	<b>The extrasynaptic connectome</b>	<b>109</b>
<b>4</b>	<b>The monoamine networks of <i>C. elegans</i></b>	<b>111</b>
4.1	Introduction . . . . .	111
4.1.1	Chapter overview . . . . .	112
4.2	Materials & Methods . . . . .	112
4.2.1	Synaptic & gap junction networks . . . . .	112
4.2.2	Monoamine network construction . . . . .	113
4.2.3	Network data . . . . .	117
4.2.4	Topological network measures . . . . .	117
4.2.5	Software . . . . .	117
4.3	Results . . . . .	122
4.3.1	Network structure . . . . .	122
4.3.2	Extrasynaptic signalling . . . . .	125
4.3.3	The multilayer connectome . . . . .	127
4.3.4	Network measures . . . . .	129
4.3.5	Rich-club organisation . . . . .	134
4.3.6	Multilink motifs . . . . .	139
4.4	Discussion . . . . .	141
4.4.1	The monoamine systems . . . . .	141
4.4.2	Multiplex networks . . . . .	146
4.4.3	Future work . . . . .	147

<b>5</b>	<b>Neuropeptide networks of <i>C. elegans</i></b>	<b>149</b>
5.1	Introduction . . . . .	149
5.2	Materials & Methods . . . . .	150
5.2.1	Neuropeptide network construction . . . . .	151
5.2.2	Network data . . . . .	151
5.3	Results . . . . .	153
5.3.1	Network measures . . . . .	160
5.3.2	Rich-club . . . . .	163
5.3.3	Multilink motifs . . . . .	166
5.4	Conclusions . . . . .	169
<b>IV</b>	<b>Conclusions</b>	<b>171</b>
<b>6</b>	<b>Conclusions &amp; discussion</b>	<b>173</b>
6.1	Contributions . . . . .	173
6.1.1	Major results . . . . .	173
6.1.2	Minor results . . . . .	174
6.2	Outlook & future work . . . . .	175
6.2.1	Extending & validating the network . . . . .	175
6.2.2	Network function . . . . .	176
6.2.3	Mapping larger nervous systems . . . . .	176
6.2.4	Outlook . . . . .	177
	<b>References</b>	<b>181</b>
	<b>Appendix A Monoamine expression patterns</b>	<b>223</b>
A.1	Monoamine expression . . . . .	224
A.2	Receptor expression . . . . .	225
	<b>Appendix B Neuropeptide expression patterns</b>	<b>229</b>
	<b>Appendix C Identification of dopamine receptor neurons</b>	<b>233</b>
C.1	Acknowledgements . . . . .	234
C.2	Materials & methods . . . . .	234
C.2.1	For <i>dop-5</i> : . . . . .	234
C.2.2	For <i>dop-6</i> : . . . . .	235
C.2.3	For <i>lgc-53</i> : . . . . .	235

---

C.2.4	Microscopy . . . . .	236
<b>Appendix D</b>	<b>Notes on expression data</b>	<b>237</b>
D.1	Fluorescent reporter lines . . . . .	238
D.2	Gene expression & network construction . . . . .	239
D.3	Data integration & batch effects . . . . .	239
D.4	Underreported expression . . . . .	240
D.5	Conditional expression . . . . .	240
<b>Appendix E</b>	<b>Null model network comparison</b>	<b>241</b>
E.1	Introduction . . . . .	242
E.2	Materials & methods . . . . .	243
E.3	Results . . . . .	243
E.3.1	Wired connectome . . . . .	243
E.3.2	Monoamine connectome . . . . .	247
E.4	Conclusions & discussion . . . . .	250
<b>Appendix F</b>	<b>Neuron classification</b>	<b>251</b>
F.1	Neuron classification schema . . . . .	252
<b>Appendix G</b>	<b>Supplementary results</b>	<b>257</b>
G.1	Hiveplot of individual MA networks . . . . .	258
G.2	Multilink analysis of separate MA networks (incl. <i>dop-5/-6</i> ) . . . . .	259
G.3	Network measures of individual MA layers . . . . .	262
<b>Appendix H</b>	<b>Network analysis software</b>	<b>263</b>



# List of figures

1.1	Cajal’s drawing of cells in the visual cortex (Cajal, 1899). . . . .	3
1.2	Scales of connectomic maps. DTI image of white matter tracts in the human brain; map of GFP labeled axon projections in the mouse brain; EM reconstruction of neurons in the olfactory bulb of a zebrafish larva	6
1.3	3D model of the wired synaptic connectome of <i>C. elegans</i> . Generated using neuroConstruct (Gleeson et al., 2007), with data from the Virtual-Worm and OpenWorm projects (Szigeti et al., 2014). . . . .	7
1.4	Modes of transmission between neurons. . . . .	9
1.5	Partial timeline of the discovery of the main neurotransmitter and neuromodulator systems . . . . .	10
2.1	A small graph composed of eight nodes and ten edges . . . . .	15
2.2	Illustration of the stages of connectomic mapping . . . . .	17
2.3	Degree of connectivity. Diagram depicting a high-degree node ( $k = 8$ ), and a low-degree node ( $k = 2$ ) . . . . .	18
2.4	Example degree distributions . . . . .	19
2.5	Assortativity . . . . .	21
2.6	Schematic diagram showing: functional segregation and the existence of network modules composed from clustered nodes, forming triangles; and functional integration from the connection of modules through hub nodes, forming a central rich-club . . . . .	22
2.7	Small-world networks . . . . .	24
2.8	Motif enumeration . . . . .	26
2.9	Network motifs. Diagram showing all 13 possible three-node motifs with directed edges. . . . .	26
2.10	Diagram showing: a two-layer multiplex network containing four nodes, and the corresponding multilinks between each pair of connected nodes	31
2.11	Representations of neural connectivity at different levels of abstraction	33

2.12	Measures used in the determination of the clustering coefficient . . . . .	43
2.13	Illustration of network modularity . . . . .	48
2.14	Network robustness . . . . .	51
2.15	Edge-swap randomisation . . . . .	57
2.16	Configuration model random network generation . . . . .	60
2.17	Rich-club core . . . . .	65
2.18	Diagram showing a two-layer multiplex network containing four nodes, and a one-layer network formed by layer aggregating . . . . .	72
3.1	Illustration of the manual tracing process used to constructed the original White et al. (1986) connectome . . . . .	86
3.2	Elegance tracing software used to create the updated WormWiring connectome . . . . .	86
3.3	Histogram of degree differences between the aggregate AC & WW networks.	89
3.4	Network wiring diagrams of the synaptic connected components in the AC & WW connectomes. Isolated neurons are identified by name. . . . .	90
3.5	Degree-degree correlation between individual nodes in the AC & WW networks. . . . .	91
3.6	Adjacency matrices showing the differences between the AC and WW synaptic and gap junction networks . . . . .	91
3.7	Sensory and interneurons with asymmetric gap junctions, all with pro- cesses in the ventral nerve cord . . . . .	93
3.8	Comparison of network metrics for the AC & WW aggregate networks	96
3.9	Scale-free properties of the AC & WW aggregate networks (showing best fit for the power-law distribution $P(k) = k^{-\alpha}$ ). . . . .	98
3.10	Robustness and small-worldness of the AC & WW aggregate networks .	98
3.11	Reciprocity for the directed AC & WW synaptic networks . . . . .	100
3.12	Rich-club curves for the undirected AC & WW aggregate networks . .	102
4.1	Monoamine synthesis pathways in <i>C. elegans</i> . . . . .	114
4.2	Monoamine data model. UML class diagram of the conceptual data model used to construct the extrasynaptic monoamine network . . . . .	116
4.3	Multilayer projection of the synaptic, gap junction, and monoamine signalling networks (DA, 5-HT, TA & OA), each represented as a separate layer . . . . .	122
4.4	Adjacency matrix showing the wired synaptic, gap junction, and ex- trasynaptic monoamine networks . . . . .	123

---

4.5	Hive plot showing the wired synaptic, gap junction, and extrasynaptic monoamine connections . . . . .	124
4.6	RIM tyramine releasing neurons, showing outgoing synaptic edges, and neurons expressing one or more of the four tyramine receptors . . . . .	126
4.7	RIC octopamine releasing neurons, showing outgoing synaptic edges, and neurons expressing one or more of the three octopamine receptors . . . . .	126
4.8	Degree-degree correlation matrix . . . . .	128
4.9	Multilayer reducibility dendrogram . . . . .	128
4.10	Comparison of network metrics for the physical synaptic & gap junction network, full monoamine network, and the aggregate physical & monoamine network . . . . .	130
4.11	Reciprocity for the directed synaptic, monoamine, and aggregate synaptic+monoamine networks . . . . .	131
4.12	Scale-free properties of the monoamine network, and aggregate synaptic, gap junction and monoamine network . . . . .	132
4.13	Robustness and small-worldness of the physical, monoamine & aggregate physical+monoamine networks . . . . .	132
4.14	Directed 3-node motif profiles for the aminergic networks . . . . .	133
4.15	Rich-club curve for the directed monoamine network . . . . .	135
4.16	Monoamine rich-club schematic, showing the separate aminergic systems and the volume transmission signalling between them based on receptor expression . . . . .	136
4.17	Connections between the physical & monoamine rich-clubs . . . . .	137
4.18	Monoamine multilink motifs . . . . .	140
4.19	Multilink motif IDs . . . . .	140
5.1	Neuropeptide data model. UML class diagram of the conceptual data model used to construct the extrasynaptic neuropeptide network . . . . .	151
5.2	Multilayer projection of the neuropeptide networks, including physical and monoamine layers . . . . .	154
5.3	Adjacency matrix showing the synaptic and neuropeptide networks . . . . .	156
5.4	Degree-degree correlation matrix . . . . .	157
5.5	Multilayer reducibility dendrogram for the synaptic, gap junction, monoamine, and neuropeptide layers . . . . .	158
5.6	Multilayer reducibility dendrogram showing all of the multiplex network layers . . . . .	159

5.7	Reciprocity for the directed synaptic network, neuropeptide network, and aggregate synaptic+monoamine+neuropeptide network . . . . .	160
5.8	Comparison of network metrics . . . . .	161
5.9	Scale-free properties of the (a) neuropeptide network, and (b) aggregate synaptic, gap junction, monoamine, and neuropeptide network (showing best fit for the power-law distribution $P(k) = k^{-\alpha}$ ). . . . .	162
5.10	Robustness and small-worldness . . . . .	162
5.11	Rich-club curve for the directed neuropeptide network . . . . .	163
5.12	DVA and RIM. Rich-club neurons with the highest degree-rank product	165
5.13	Neuropeptide multilink motifs . . . . .	167
5.14	Multilink motif IDs . . . . .	167
C.1	Expression patterns of the dopamine receptors <i>dop-5</i> , <i>dop-6</i> & <i>lgc-53</i> .	236
E.1	Reciprocity of null model networks generated from the degree sequence of the AC connectome . . . . .	244
E.2	Comparison of network measures for null model networks generated from the degree sequence of the AC connectome . . . . .	245
E.3	Robustness of null model networks generated from the degree sequence of the AC connectome . . . . .	246
E.4	Small-worldness of null model networks generated from the degree sequence of the AC connectome . . . . .	246
E.5	Reciprocity of null model networks generated from the degree sequence of the MA connectome . . . . .	247
E.6	Comparison of network measures for null model networks generated from the degree sequence of the MA connectome . . . . .	248
E.7	Robustness of null model networks generated from the degree sequence of the MA connectome . . . . .	249
E.8	Small-worldness of null model networks generated from the degree sequence of the MA connectome . . . . .	249
G.1	Hiveplot showing individual dopamine, serotonin, tyramine, and octopamine connections . . . . .	258
G.2	Multilink analysis of the aggregate monoamine network including <i>dop-5</i> & <i>dop-6</i> , and the dopamine, serotonin, and octopamine networks . . . .	259
G.3	Multilink analysis of the dopamine network including <i>dop-5</i> & <i>dop-6</i> , and tyramine network . . . . .	260

---

G.4 Multilayer projection including <i>dop-5</i> & <i>dop-6</i> . Network measures for the individual dopamine, dopamine incl. <i>dop-5</i> & <i>dop-6</i> , serotonin, tyramine and octopamine networks; and aggregate MA network including <i>dop-5</i> & <i>dop-6</i> . . . . .	262
--	-----



# List of tables

3.1	Edge count & network density. . . . .	89
3.2	Cell-cell synaptic connections in AC that no longer exist in WW. . . . .	94
3.3	Asymmetric gap junctions in the WW connectome. . . . .	95
3.4	Rich-clubs of the AC and WW aggregate networks at the $\Phi_{norm}(k) \geq 1 + 3\sigma$ level, including newly identified AC rich-club members ( $k = 34$ ) not described by Towlson et al. (2013) . . . . .	103
4.1	Serotonin (5-HT) biosynthetic enzyme, transporter & receptor genes used to construct the extrasynaptic 5-HT network . . . . .	118
4.2	Dopamine (DA) biosynthetic enzyme, transporter & receptor genes used to construct the extrasynaptic DA network. . . . .	119
4.3	Octopamine (OA) biosynthetic enzyme & receptor genes used to construct the extrasynaptic OA network . . . . .	120
4.4	Tyramine (TA) biosynthetic enzyme & receptor genes used to construct the extrasynaptic TA network . . . . .	121
4.5	Monoamine networks nodes & edges . . . . .	123
4.6	The number of monoamine receptor-expressing cells that do not receive synapses from releasing cells, and the number of connections in each layer that are non-synaptic, including connections between neurons within the same class . . . . .	125
4.7	The normalised degree-rank product ( $k_{norm}$ ) showing the neurons that have the highest interaction in all of the network layers . . . . .	137
4.8	Rich-club neurons of the aggregate monoamine network . . . . .	138
4.9	Motif 10. List of neurons connected by motif 10 (i.e. unidirectional MA link, no gap junctions, and reciprocal synapses). . . . .	142
4.10	Multilink motif frequencies for the monoamine, synaptic and gap junction layers . . . . .	143

5.1	Neuropeptide receptor-ligand binding . . . . .	152
5.2	Neuropeptide networks. . . . .	155
5.3	The number of neuropeptide receptor-expressing cells that do not receive synapses from releasing cells, and the number of connections in each layer that are non-synaptic . . . . .	155
5.4	Rich-club neurons of the aggregate neuropeptide network . . . . .	164
5.5	The normalised degree-rank product ( $k_{norm}$ ) showing the neurons that have the highest interaction in all of the network layers . . . . .	165
5.6	Examples of neuropeptide multilink <i>motif 20</i> . . . . .	166
5.7	Multilink motif frequencies for the neuropeptide, synaptic and gap junction layers . . . . .	168
A.1	Serotonin (5-HT) expressing cells . . . . .	224
A.2	Dopamine (DA) expressing cells. . . . .	224
A.3	Octopamine (OA) & tyramine (TA) expressing cells . . . . .	224
A.4	Serotonin (5-HT) receptor expression patterns. . . . .	225
A.5	Octopamine (OA) receptor expression patterns. . . . .	225
A.6	Dopamine (DA) receptor expression patterns. . . . .	226
A.7	Tyramine (TA) receptor expression patterns. . . . .	227
B.1	Neuropeptide expression patterns. . . . .	230
B.2	Neuropeptide receptor expression patterns. . . . .	231
E.1	Comparison of mean network measures computed for 100 randomised networks generated from the directed version of the Albertson-Chklovskii synaptic & gap junction connectome of <i>C. elegans</i> . . . . .	243
E.2	Comparison of mean network measures computed for 100 randomised networks generated from the directed version of the monoamine connectome of <i>C. elegans</i> . . . . .	247
F.1	Neuron classes: A. . . . .	253
F.2	Neuron classes: B – O. . . . .	254
F.3	Neuron classes: P – R. . . . .	255
F.4	Neuron classes: S – V. . . . .	256
G.1	Examples of octopamine multilink <i>motif 9</i> . . . . .	261
G.2	Examples of octopamine multilink <i>motif 11</i> . . . . .	261

H.1 List of software tools that implement network measures or algorithms detailed in Chapter 2 . . . . .	264
---	-----



# List of Algorithms

1	<b>DSP – Dijkstra’s shortest path.</b> Gives the geodesic distances and paths from a given node to every other node (Dijkstra, 1959). . . . .	41
2	<b>MM – Maximised modularity.</b> Greedy algorithm that iteratively agglomerates nodes into modules to find a partitioning that maximise the modularity score $Q$ . First proposed by Newman (2004). . . . .	47
3	<b>ER-Np – Erdős-Rényi <math>G(N, p)</math> model.</b> Generates random networks with $N$ vertices and intervertex connection probability $p$ . Based on the model of random networks first introduced by Solomonoff and Rapoport (1951), and independently developed by Gilbert (1959). . . . .	54
4	<b>ER-NM – Erdős-Rényi <math>G(N, M)</math> model.</b> Generates random networks with $N$ vertices and $M$ edges. Based on the model of random networks introduced by Erdős and Rényi (1959). . . . .	55
5	<b>ES – Edge-swap.</b> Degree-preserving Markov chain Monte Carlo (MCMC) edge-swap rewiring / randomisation algorithm; used for generating null model networks. Based on the procedure described by Maslov and Sneppen (2002); Milo et al. (2003). . . . .	58
6	<b>CM – Configuration model.</b> Stub matching algorithm to generate undirected random graphs with a given degree sequence $\vec{k}$ ; used for generating null model networks. The algorithm can easily be modified to generate directed graphs by connecting out-degree stubs to in-degree stubs. First proposed by Molloy and Reed (1995). . . . .	59
7	<b>ESU – Enumerate subgraphs.</b> Subgraph enumeration algorithm from Wernicke (2005); used for motif detection in FANMOD (Wernicke and Rasche, 2006). . . . .	69



# Nomenclature

## Roman Symbols

$\mathcal{A}$	Multiplex network
$A$	Adjacency matrix
$\mathcal{B}$	Reduced multiplex network
$B$	Reduced multiplex layer
$C$	Clustering coefficient
$c$	Network module
$\mathcal{D}$	Degree matrix
$D_{JS}$	Jensen-Shannon divergence
$D_{KL}$	Kullback-Leibler divergence
$D$	Network density
$d$	Geodesic distance
$dq$	Temporary variable for $\Delta Q$
$d$	Jensen-Shannon distance
$\mathcal{E}$	Global efficiency
$E$	Set of all edges
$e$	Edge (arc, link)
$G$	Graph (network)

---

$g$	Geodesic path
$h$	Motif identity
$I$	Diffusion constant
$J$	Motif frequency
$\vec{k}$	Degree sequence
$K$	Rich-club degree cut-off
$k$	Vertex degree
$\ell$	Network diameter
$\hat{\mathcal{L}}$	Rescaled Laplacian
$\mathcal{L}$	Laplacian matrix
$L$	Characteristic path length
$\mathcal{M}$	Moment
$M$	Number of edges
$m$	Multilink
$\mathbb{N}$	Set of natural numbers
$N$	Number of vertices
$N_h$	Vertex or subgraph neighbourhood
$O$	Asymptotic growth
$P$	Probability
$Q$	Modularity
$q$	Multiplex distinguishability
$Q$	Vertex queue
$\mathbb{R}$	Set of real numbers
$R$	Assortativity; correlation coefficient

---

$r$	Reciprocity
$S$	Small-world index
$s$	Subgraph (motif) size
$\mathbf{S}$	Visited vertices
$t$	Number of triangles
$\text{Tr}$	Trace operator
$V$	Set of all vertices
$v$	Vertex (node)
$W$	Set of all edge weight
$w$	Edge weight
$\mathbb{X}$	Arbitrary set
$X$	Number of multiplex layers
$z$	Z-score

**Greek Symbols**

$\alpha$	Power-law exponent
$\beta$	Multiplex network layer
$\gamma$	Fraction of nodes in the largest connected component
$\Delta$	Difference
$\delta$	Kronecker delta function
$\epsilon$	Efficiency
$\eta$	Source vertex
$\bar{\mathcal{H}}$	Average von Neumann entropy
$\mathcal{H}$	Multilayer von Neumann entropy
$H$	Entropy

---

$\kappa$	Degree rank
$\Lambda$	Ensemble of random networks
$\lambda$	Eigenvalue
$\mu$	Population mean
$\xi$	Edge-swap randomisation factor
$\Pi$	Predecessor vertex
$\varpi$	Random number
$\rho$	Robustness
$\varrho$	Density matrix
$\sigma$	Standard deviation
$\Phi$	Rich-club coefficient
$\phi$	Diffusible quantity
$\chi$	Reducibility
$\psi$	Quantum state
$\Omega$	Motif set
$\omega$	Arbitrary vertex

### Superscripts

$[\beta]$	Layer $\beta$ of multiplex / multilayer network
$\leftrightarrow$	Reciprocal edge
$\rightarrow$	Directed network measure
<i>in</i>	Incoming connections
<i>out</i>	Outgoing connections

### Subscripts

$\eta$	Source vertex
--------	---------------

---

$\omega$	Arbitrary vertex
<i>AC</i>	Albertson-Chklovskii connectome
<i>ag</i>	Aggregate network
<i>CM</i>	Configuration model
<i>ER</i>	Erdős-Rényi random network
<i>ES</i>	Edge-swap randomisation
<i>gj</i>	Gap junction network
JS	Jensen-Shannon divergence / distance
KL	Kullback-Leibler divergence
<i>ma</i>	Monoamine network
max	Maximised value
<i>norm</i>	Normalised
<i>np</i>	Neuropeptide network
<i>opt</i>	Optimum
<i>phys</i>	Physical network; combined synaptic & gap junction
<i>rand</i>	Random network
<i>sub</i>	Subgraph
<i>syn</i>	Synaptic network
shannon	Shannon entropy
temp	Temporary variable
<i>WW</i>	WormWiring connectome

**Acronyms / Abbreviations**

5-HT	Serotonin (5-hydroxytryptamine)
7TM	Seven-transmembrane domain

---

AC	Albertson-Chklovskii connectome
ADR	Adrenergic receptor
ANS	Autonomic nervous system
BAS	Biogenic amine synthesis protein
BFS	Breadth-first search
BOLD	Blood-oxygen-level dependent contrast imaging
CAT	Catecholamine distribution defect associated protein
CCK	Cholecystokinin
CKR	Cholecystokinin receptor homologue protein
CM	Configuration model
CNS	Central nervous system
CRT	Calreticulin / calregulin protein
CSF	Cerebrospinal fluid
DA	Dopamine
DAT	Dopamine transporter
DDC	DOPA decarboxylase
DFS	Depth-first search
DOP	Dopamine receptor protein
DOPA	3,4-dihydroxyphenylalanine
DR	Dopamine receptor
DSP	Dijkstra's shortest path algorithm
DTI	Diffusion tensor imaging
EC <sub>50</sub>	Effective concentration, half-maximal
EEG	Electroencephalography

---

EGL	Egg-laying defect associated protein
EM	Electron microscope
ER	Erdős-Rényi random network model
ES	Edge-swap randomisation
ESU	Enumerate subgraphs (algorithm)
EV	Extracellular vesicle
FIF	Formaldehyde-induced fluorescence
FLP	FMRFamide-like peptide
fMRI	Functional magnetic resonance imaging
FRET	Förster resonance energy transfer
FRPR	FMRFamide-like peptide receptor protein
GABA	$\gamma$ -Aminobutyric acid
Gap	Gap junction network
GFP	Green fluorescent protein
GJ	Gap junction
GLR	Glutamate receptor family (AMPA) protein
GPC	G protein $\gamma$ -subunit protein
GPCR	G protein-coupled receptor
JSD	Jensen-Shannon divergence
KLD	Kullback-Leibler divergence
L-DOPA	Levodopa (L-3,4-dihydroxyphenylalanine)
LGC	Ligand-gated channel
MA	Monoamine
MCMC	Markov chain Monte Carlo process

---

MEG	Magnetoencephalography
MM	Maximised modularity
MOD	Modulation of locomotion defect associated protein
MRI	Magnetic resonance imaging
N2	Bristol wildtype <i>C. elegans</i> strain
Nh	Vertex or subgraph neighbourhood
NLP	Neuropeptide-like protein
NP	Neuropeptide
NPR	Neuropeptide receptor protein
NTC	Nematocin (vasopressin-like peptide)
NTR	Nematocin receptor protein
OA	Octopamine
OCTR	Octopamine receptor protein
PDF	Pigment dispersing factor homologue protein
PDFR	Pigment dispersing factor receptor protein
Phx	Pharynx / pharyngeal nervous system
PPI	Protein-protein interaction
QM	Quantum mechanics
RC	Rich-club network
RNAi	RNA interference
SE	Standard error of the mean
SER	Serotonin / octopamine receptor protein
SLC	Solute carrier; membrane transporter family
Syn	Synaptic network

---

TA	Tyramine
TBH	Tyramine $\beta$ -hydroxylase
TDC	Tyrosine decarboxylase
TEM	Transmission electron microscope
TH	Tyrosine hydroxylase
TPH	Tryptophan hydroxylase
TRH	Thyrotropin releasing hormone
TYRA	Tyramine receptor protein
UML	Unified modeling language
VMAT	Vesicular monoamine transporter
VT	Volume transmission
WBE	Whole brain emulation
WT	Wild type
WW	WormWiring connectome



# Part I

## Introduction



# Chapter 1

## Introduction

### 1.1 Connectomics

ONE of the most defining characteristics of the brain is its complex patterning of cellular connectivity. This observation was first made towards the end of the 19<sup>th</sup> century through the work of Cajal (1888), and lead directly to the formation of the *neuron doctrine* that forms the foundation of modern neuroscience: the concept that the brain is a collection of distinct cells that interact through synapses. Developments in neuroscience and microscopy since Cajal's initial observation have lead to the creation of the modern field of *connectomics*, which aims to map the complete set of these connections within a brain (Sporns, 2015; Sporns et al., 2005).

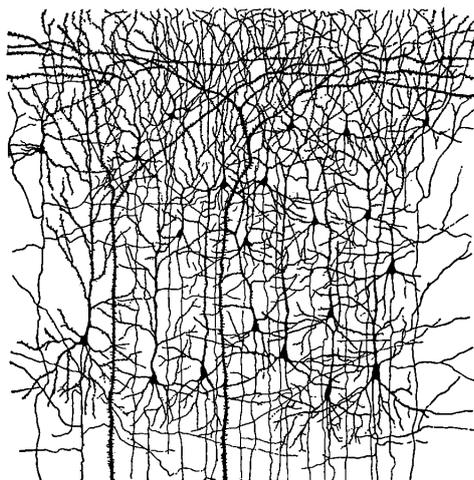


Fig. 1.1 Cajal's drawing of cells in the visual cortex (Cajal, 1899).

While synapses – and the neural projections that form them – are a prominent structural and functional feature of the brain, it is well established that they are not the only route of communication between neurons. As connectomics continues in its venture to map synaptic connectivity, a growing body of research exists suggesting that additional, *non-synaptic*, interactions will also need to be accounted for to understand the operation of neural circuits; specifically, long-distance communication through the release and diffusion of neuromodulatory molecules. The **subject of this thesis** is an attempt to extend the definition of connectomics to include these non-synaptic interactions, by demonstrating that neuromodulatory networks can be mapped at the level of an entire nervous system, and analysed alongside synaptic networks using existing connectomic tools, to yield biologically relevant insights.

The following sections lay the foundation for the presentation of this work in Chapters 4 & 5, by providing an overview of the current state of connectomics, and introducing the nematode worm *Caenorhabditis elegans* in which the work was conducted.

### 1.1.1 The role of connections

All interactions in the nervous system are ultimately grounded in the context of the wider environment. In an energetic and dynamic environment, such as the Earth, life must constantly adapt. This applies not only on evolutionary timescales – where species evolve to maximise their fitness in an environment – but also on a moment-by-moment basis. Organisms must constantly respond to their changing environments to maintain the internal configurations and processes necessary for the continuation of life. This is epitomised by homeostasis (Ashby, 1952; Cannon, 1929).

For unicellular organisms, their immediate environment is the only environment that must be considered. This is not the case, however, for multicellular organisms. As a multicellular organism grows, the number of cells in the internal space begin to outnumber those on the surface, and thus have no way of immediately sensing the larger environment in which they are embedded. Large multicellular organisms therefore require not only a dedicated interface to sense and communicate environmental states to the whole organism, but also mechanisms to coordinate responses for the mutual benefit of all its cells.

In simple and complex animals alike, it is clear that the nervous system is one such adaptive coordinating system: it enables and controls a host of essential processes, from locomotion and feeding, to reproduction, growth, immune responses, and the detection and avoidance of noxious entities. It does so in a way that is both highly robust to failure, and extremely adaptable; capable of responding to novel circumstances, and developing responses on a much faster timescale than the evolutionary process.

The web of cellular connectivity that sits between a nervous system's sensory cells and its effectors clearly play a role in structuring how stimuli are processed and converted into behaviours suited to the environment. The question arises: can we understand and reverse engineer these adaptive processes from the connectivity patterns we observe between cells; and if so, can these patterns also explain the higher-level behaviours and mental states that define our experience as rational, conscious agents? While the second-part of the question is a long way from being answered – some have argued it to be unanswerable (Chalmers, 1995; Schrödinger, 1958) – experiments in simple systems demonstrated early on the power of studying neural connectivity.

Combining Cajal's theories with previous observations that the brain has an electro-chemical basis, starting with the work of Galvani a hundred years before (Galvani, 1791), the neuron doctrine allowed early electrophysiologists to begin describing behaviours in terms of interactions between individual cells, providing a gateway to tractable, reductionist, explanations of the nervous system and the generation of behaviours, with the earliest work resolving the nature of reflexes (e.g. Sherrington, 1906a; 1906b). The power of this model of the nervous system has resulted not only in descriptions of simple behaviours, but also general principles of information processing. Theories of synaptic learning and cell assemblies, where connections are strengthened in response to repeated activation to form functional circuits (Hebb, 1949), have proved capable of describing a great deal of behaviour; while work by McCulloch and Pitts (1943) demonstrated that synaptically connected neurons can perform complex logical operations. Together these theories have provided a foundation for non-symbolic approaches to general computation, where objects and functions are not explicitly modelled, but instead emerge from the dynamic interactions between nodes in a network (Rumelhart and McClelland, 1986). This alternative computational paradigm, known as *connectionism*, has provided insights into the nature of many natural emergent systems and yielded engineering applications outside the realm of neuroscience (Hassabis et al., 2017; Schmidhuber, 2015; Soman et al., 2016).

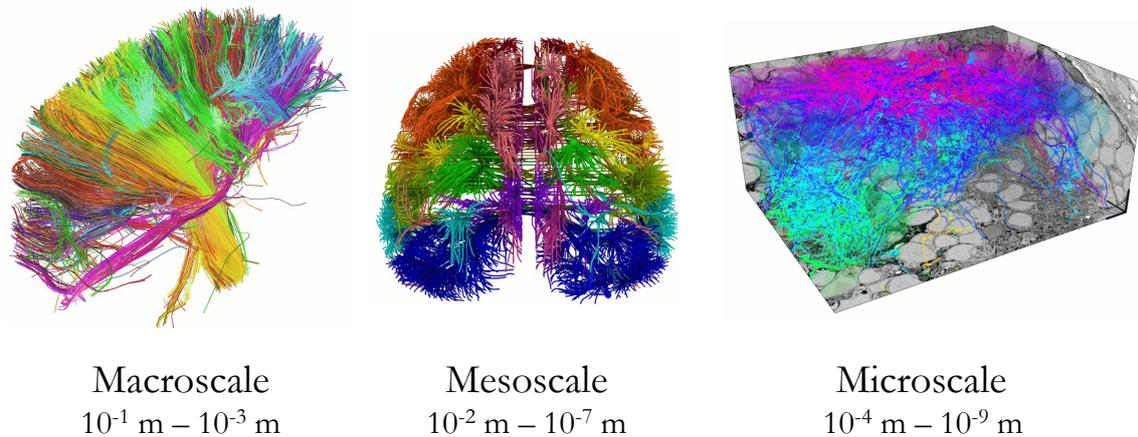


Fig. 1.2 Scales of connectomic maps. From left to right: DTI image of white matter tracts in the human brain, from the NIH Human Connectome Project (Harvard/MGH-UCLA, 2015); map of GFP labeled axon projections in the mouse brain, from The Allen Brain Institute (Oh et al., 2014); EM reconstruction of neurons in the olfactory bulb of a zebrafish larva (Friedrich et al., 2013). Scale ranges taken from Ohno et al. (2015).

### 1.1.2 The state of connectomics

Developments in neuroscience in the intervening century since Cajal and Sherrington have, in recent years, enabled the first serious attempts to describe the complete neural connectivity of an organism, and link connections directly to behaviours. This is only now possible due to the confluence of several disparate technologies. On one side, advances in microscopy, computer imaging, and molecular biology have provided approaches to high-throughput structural and functional imaging of entire nervous systems and their constituent molecules (e.g. MRI, electron / confocal / two-photon microscopy,  $\text{Ca}^{2+}$  imaging, molecular labelling, etc.); on the other, an explosion of activity in complex network research since the mid-1990's has resulted in the theoretical and computational tools necessary for analysing and interpreting the properties of large networks. Together, these technologies provide an approach to overcoming research limits imposed by the size of the brain, permitting the study of neural systems in all of their richness and complexity.

Current work in connectomics can largely be divided into three categories, based on the physical scale and the level of description they provide (Figure 1.2). At the highest level, *macroscale connectomics* seeks to describe the large scale structure of

the brain by mapping white matter tracts and functional interactions between brain regions, using techniques such as MRI and diffusion tensor imaging (DTI). At the smallest scale, *microscale connectomics* charts the morphologies and connections of individual synapses using electron microscopy (EM). In between these two is *mesoscale connectomics* which records the projections of neurons, but not their synaptic details.

Large scale projects exist in all of these areas, and across several species, including mice (Oh et al., 2014), humans (Harvard/MGH-UCLA, 2015), *Drosophila* (Shih et al., 2015), and zebrafish (Friedrich et al., 2013; Hildebrand et al., 2017); however, to date the only organism for which a complete microscale synaptic connectome exists is the nematode worm *Caenorhabditis elegans*; see Figure 1.3 (White et al., 1986). This achievement was largely possible thanks to the worm's small size, with only 302 neurons, and laid the foundation for understanding many organisational and functional properties of its nervous system (Sengupta and Samuel, 2009; Stam and Reijneveld, 2007).

Despite the ongoing success of connectomics in revealing the structural basis of the brain and behaviour, a number of reviews have recently been published highlighting problems and limits with the current approach (Bargmann, 2012; Bargmann and Marder, 2013; Brezina, 2010; Kopell et al., 2014; Marder et al., 2014; Morgan and Lichtman, 2013; Sporns, 2013a). Chief amongst these is the observation that extrasynaptic molecular interactions can radically alter the properties of the synaptic network. Although

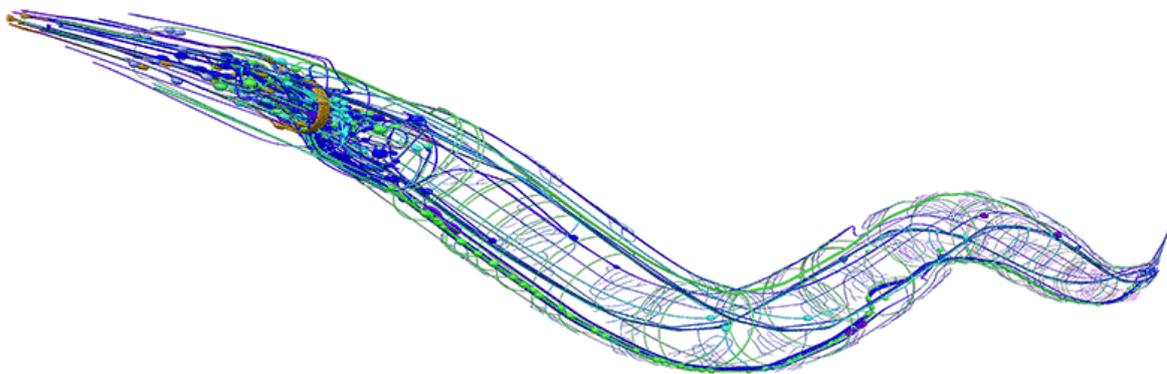


Fig. 1.3 3D model of the wired synaptic connectome of *C. elegans*. Generated using neuroConstruct (Gleeson et al., 2007), with data from the VirtualWorm and OpenWorm projects (Szigeti et al., 2014).

synaptic connections are the most salient feature of the brain, they are far from the only method of communication between cells (Agnati et al., 2010; Jefferys, 1995).

### 1.1.3 Non-synaptic interactions

Several methods of neuronal communication exist (Figure 1.4). Besides from chemical synapses, that form the basis of the wired transmission (WT) system, there are also a number of other known interactions in the nervous system. These interactions are diverse, with spatiotemporal properties that span orders of magnitude.

An additional route of wired transmission, besides synapses, is through the coupling of two or more neurons via gap junctions. Through connecting the cell membranes, gap junctions allow for fast, bidirectional, ionic communication on the order of 100 microseconds, and have also been shown to simultaneously allow the transfer of second messengers and metabolites between cells (Hernandez et al., 2007; Sohl et al., 2005). This form of undirected coupling allows groups of cells to form single functional units with unique computational properties not seen in single neurons (Rabinowitch et al., 2013).

Alongside synapses and gap junctions there exist additional, non-wired, methods of interaction. One of these is ephaptic coupling (Anastassiou et al., 2011), whereby neurons can alter the electrical states of surrounding cells, either through field effects to induce currents or by altering the ionic concentrations of the extracellular space. This is thought to play a role in synchronisation behaviour between groups of neurons (Anastassiou et al., 2011; Jefferys, 1995; Weiss, 2010).

Perhaps one of the most diverse and ancient methods of cellular communication, predating the nervous system itself, is the mechanism of volume transmission (VT). Neurons can release small-molecules, peptides, and gasses that diffuse through the extracellular space to interact with cells over long distances (Agnati et al., 2010). While VT has long been recognised as a function of neurons, it has traditionally been viewed primarily in the context of the autonomic nervous system (ANS), and not as a general principle of neural communication; however, molecular mapping studies of the central nervous system (CNS) have revealed that the density of nerve terminals for particular neurotransmitters often do not align with the location of matching receptors (Taber and Hurley, 2014), providing evidence that some neurotransmitters in the CNS must diffuse over long-distances to reach their targets. For example, this method of com-

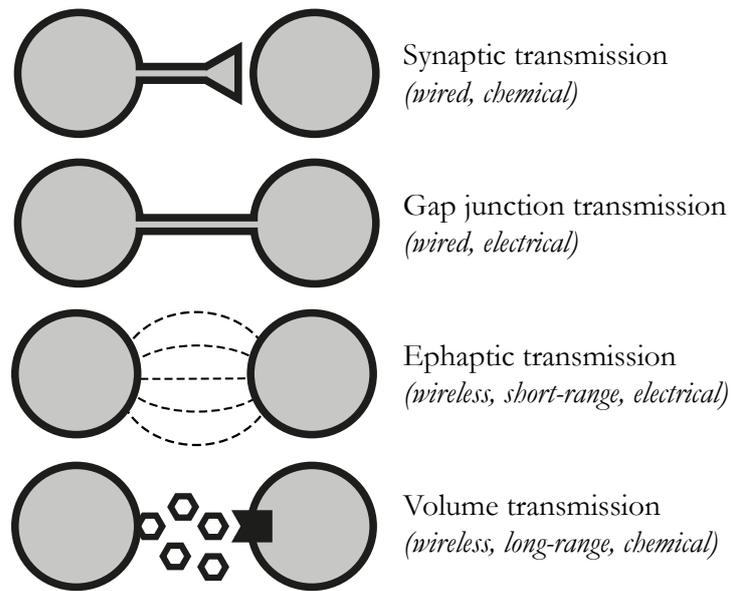


Fig. 1.4 Modes of transmission between neurons.

munication can be seen in the dopamine system (Rice and Cragg, 2008; Rice et al., 2011).

With new molecules constantly being discovered with the ability to both act over long distances and alter the function of neural circuits (e.g. Figure 1.5), it is clear the synaptic network alone will not be sufficient to reverse engineer behaviours (Bargmann and Marder, 2013). Indeed, the connectome of *C. elegans* was first published in 1986, yet the function of many neurons and circuits remain unknown. This is not surprising given the complexity of extrasynaptic signalling; for example, the *C. elegans* genome has so far been identified to encode over 250 distinct neuropeptides from 122 precursor genes, many of which have not been characterised or matched to a receptor (Hobert, 2013; Li and Kim, 2008). To add additional complexity, it is becoming clear that volume transmission can also exist through the release of extracellular vesicles (EV) that can carry a number of payloads over long distances (Agnati and Fuxe, 2014; Budnik et al., 2016; Wang and Barr, 2016).

Although the task of mapping these systems is a difficult one, it is not impossible. Expression profiling and molecular labelling can identify the locations of these molecules in the nervous system, while deorphanization studies can link molecules to receptors and reveal the interaction properties. As will be seen in later chapters, this information can be used to generate maps of extrasynaptic volume transmission networks, and analysed with the same approaches used for traditional wired connectomes.

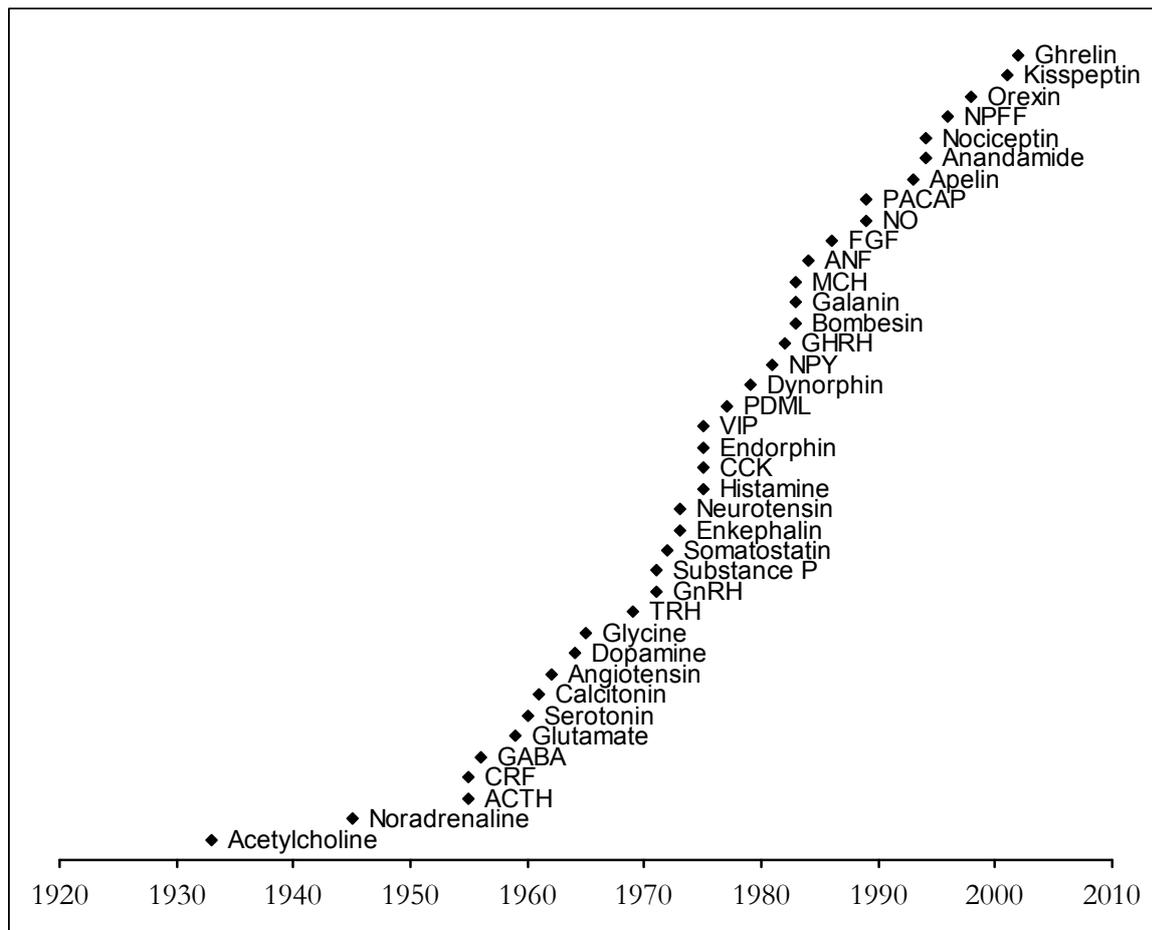


Fig. 1.5 Partial timeline of the discovery of the main neurotransmitter and neuromodulator systems; data after 2000 is underrepresented. Figure adapted from Sandberg and Bostrom (2008).

## 1.2 Thesis overview

### 1.2.1 Project background

This chapter has introduced connectomics as the study of synaptic connectivity and briefly discussed how neurons can interact not only through wired synapses, but also through wireless methods, such as extrasynaptic volume transmission (VT), with the ability to modulate neuronal signalling in multiple ways. As one of the goals of neuroscience is to understand the operations of the nervous system from the patterns of interactions between its constituent neurons, it will be necessary to obtain maps of both synaptic and extrasynaptic interaction networks.

Although synaptic connectomics has developed to a state where the whole-brain reconstruction of small nervous systems is now tractable, the problem of systematically mapping chemical VT networks has not previously been addressed. With many VT networks remaining either uncharacterised or uncharted, the details of how these systems interact with, modulate, and contribute to information processing in the wired synaptic networks are poorly understood, as is their involvement in neuropathologies. This thesis attempts to make progress towards a solution by demonstrating that VT networks can be mapped and analysed as part of the connectome.

*C. elegans* has many properties that make it a useful model organism for the study of VT networks. As well as having a small mapped connectome (White et al., 1986) with a stereotyped cell lineage (Sulston et al., 1983), *C. elegans* is highly amenable to genetic manipulation (Dickinson and Goldstein, 2016; Fay, 2013; Praitis and Maduro, 2011), as well as *in vivo* imaging as a result of its transparent cuticle (Kerr, 2006; Shaham, 2006). With the existence of tools that allow the determination of gene expression patterns at single-cell resolution (Boulin et al., 2006), public repositories of such expression patterns (Bhatla, 2014; Howe et al., 2016), and information on many ligand-receptor interactions, and synthesis and processing pathways for the two main classes of VT molecules – namely, monoamines (e.g. serotonin, dopamine, etc.; Chase and Koelle, 2007) and neuropeptides (e.g. NPY, vasopressin-, oxytocin-, and endorphin-like systems, etc.; Li and Kim, 2008) – it is possible to determine the locations of ligands and receptors used in volume transmission in the nervous system of *C. elegans*, and thus make predictions about which neurons are likely to interact extrasynaptically via VT.

## 1.2.2 Research objectives

This thesis addresses the problem of mapping aminergic and peptidergic VT networks in *C. elegans* by demonstrating that gene expression and ligand-binding data can be combined to chart putative extrasynaptic signalling between neurons; specifically, using the expression patterns of biosynthetic enzymes, peptides, and transporters to identify the broadcasting nodes of the networks; receptor expression patterns to identify the receiving nodes; and ligand-binding data to predict interaction networks between the broadcasting and receiving nodes.

The goals of this thesis are to provide:

1. a proof-of-concept that VT networks can be mapped from gene expression data,
2. a map of the extrasynaptic monoamine networks of *C. elegans*,
3. a partial map of the extrasynaptic neuropeptide networks of *C. elegans*<sup>1</sup>,
4. a graph theoretic analysis of the extrasynaptic monoamine and neuropeptide networks to determine their structural properties and demonstrate that the concept of connectomics can be extended to include VT networks.

## 1.3 Thesis structure & scientific contributions

This thesis is organised into four parts:

**Part I**, containing Chapters 1 & 2, provides an introduction to connectomics. Chapter 2 first gives a broad conceptual overview of the tools from network theory that are used in the analysis of brain networks (covered in § 2.1 – 2.2.5), followed by an in-depth discussion of their usage and the presentation of their formal mathematical definitions (§ 2.3 – 2.5.3). These tools are used throughout the thesis to analyse the neural connectivity of *C. elegans*.

**Part II** (page 83) presents a comparative network analysis of two versions of the synaptic connectome of *C. elegans*: (1) the canonical network derived from manual neuron tracing, and (2) an updated network generated with the aid of computerised

---

<sup>1</sup>Due to the complexity of the neuropeptide networks, with over 250 neuropeptides versus the four monoamines, insufficient data currently exist to map all of the neuropeptide systems in their entirety (see Chapter 5).

tools. As the properties of the extrasynaptic VT networks are compared against the synaptic network, this chapter explores whether there are any differences between the two versions of the synaptic connectome, and if previously published results in the literature are still valid for the new network. The analysis identifies a large increase in the number of synapses in the updated network, but only minor changes in the global architecture and structural properties. Explanations are proposed for the observed differences, and suggestions are made on ways in which the accuracy of connectomic maps could be improved. Due to inconsistencies in the connectivity of the updated connectome, the original is used as the main reference network in all subsequent chapters.

**Part III** (page 109) covers the mapping and analysis of the extrasynaptic networks of *C. elegans*. Chapter 4 focuses on the monoamine networks, while Chapter 5 analyses the neuropeptide networks. These chapters demonstrate that gene expression data can be used to reconstruct VT signalling networks, presenting the first whole-animal map of neuronal monoamine signalling. The network analyses find that both VT networks (i.e. monoamine & neuropeptide) predominantly exhibit star-like structures that broadcast signals throughout the worm, while also containing complex features indicative of internal information processing, such as central rich-clubs of interconnected neurons that are likely to coordinate activity between the various VT systems, and overrepresented connectivity motifs identifying locations where the synaptic and extrasynaptic networks are likely to interact.

**Part IV** (page 171) contains a reflection on the main findings of this thesis, their potential implications, and areas for future work.



# Chapter 2

## Complex network theory in connectomics

### 2.1 Introduction

COMPLEX network theory seeks to describe and understand the properties of large real-world networks through the use of mathematical models (Liu and Tse, 2015; Newman, 2003b). The main construct in network theory is the *graph*: a collection of *nodes* (also referred to as *vertices*) that are connected by *edges* (known as *links* or *arcs*) (Figure 2.1). The mathematical foundations of network science therefore lie in the classical discipline of *graph theory*, originating with the pioneering work of Euler (1736).

Graphs are a natural model for real-world networks due to the ease with which various entities and their relationships can be mapped to nodes and edges. This is especially true for the nervous system, which by its very structure is inherently graph-like, with neurons (nodes) and synapses (edges). Later chapters will demonstrate how extrasynaptic molecular interactions can also be included as edges.

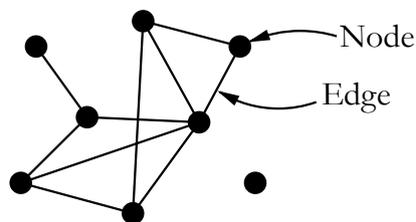


Fig. 2.1 A small graph composed of eight nodes and ten edges. Figure modified from Newman (2003b).

### 2.1.1 Network theory and the brain

The power of applying graph theory to the study of brain networks can be seen in two areas. Firstly, the ability to transform vast webs of interactions into clearly delineated models provides a handle on otherwise overwhelming complexity, aiding interpretation and allowing for comparisons with other networks. Secondly, modelling the brain as a mathematical structure makes it amenable to numerical and logical measurements and manipulations that are difficult or impossible to perform *in vivo*; for example, examining the consequences of rewiring specific circuits.

The same methods can be applied equally well across various scales and data types. On the topic of scale, defining nodes as anatomical regions allows network theory to be applied to macroscale connectomes, to study the structural or functional relationships between brain areas. The same methods can also be applied to microscale circuits, describing single neurons and their synapses. On data types, network theory allows the consideration (and consolidation) of measurements from various sources (Bassett and Sporns, 2017). For example, edges can reflect physical anatomical connections between neurons (as explored in Chapter 3), functional correlations (as from fMRI / BOLD, MEG, EEG, etc.; see Bullmore and Sporns, 2009), effective causal interactions through various routes (Friston, 1994) (as shown with extrasynaptic interactions in Chapters 4 & 5), or other abstract relationships, such as morphological similarity (Kong et al., 2015) or shared gene expression profiles (Lein et al., 2007).

Through the application of the graph formalism to both tangible and intangible systems across numerous spatiotemporal domains – from molecular interactions (Ideker and Sharan, 2008; Tieri et al., 2005) up to global communication networks (Albert et al., 1999; Pastor-Satorras et al., 2001) – network science has revealed that real-world networks of many types share common features (Strogatz, 2001); thus, networks appear to be an essential property of nature, having their own characteristics distinct from both those of their individual constituent elements, and the emergent behaviours of the systems they create. Networks mediate the qualitative transformation between the two scales, and provide insights into the nature of both. This makes network theory especially relevant to the field of neuroscience, which seeks to understand how the properties of relatively simple neurons and molecules can combine to generate complex emergent behaviours (Brown, 2013).

## 2.2 Network concepts

This section gives an overview of the main concepts and tools from network theory – as applied to the nervous system – that are used in subsequent chapters to analyse the synaptic and extrasynaptic connectomes of *C. elegans*. A more detailed introduction including formal definitions is provided in § 2.3.

### 2.2.1 Connectivity

In both network theory and connectomic analysis, the first step is to obtain a graphical model of the system of interest (i.e., a graph; as depicted in Figures 2.1 & 2.2c). In the context of connectomics, this is the set of structural or functional relationships that have been measured between neurons. Such measurements can be made using any one of a number of possible methods, some of which were introduced in the previous chapter and will be explored further later in this thesis.

Once an interaction graph has been constructed, we can use mathematical measures to interrogate the network model and extract pertinent information about the interactions. More specifically, we are often looking to obtain information regarding: (1) the patterns of connectivity that exist within the system of interest (the *structure*), and (2) the properties associated with the particular arrangements of connections we

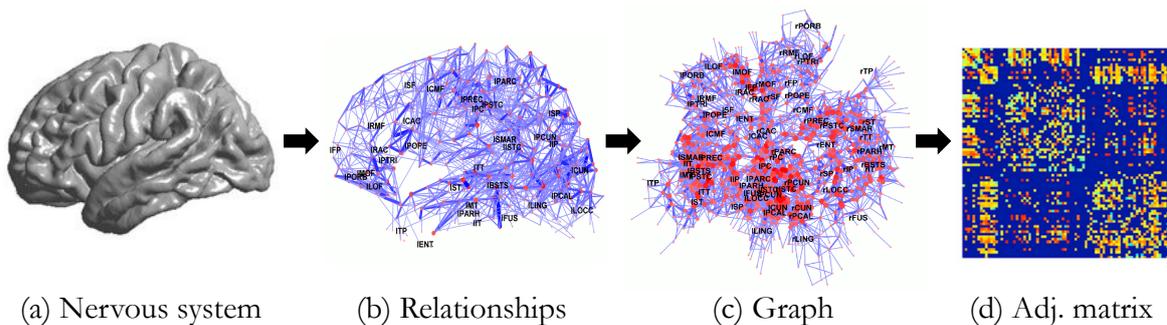


Fig. 2.2 Illustration of the stages of connectomic mapping: (a) measurements are made of a nervous system, and (b) the structural or functional relationships between defined regions are recorded, such as synaptic connections. (c) These relationships can be modelled as a graph, with regions mapping to nodes, and the corresponding measured relationships mapping to edges. (d) The graph can be represented numerically in the form of an adjacency matrix and used to calculate the network's topological properties. Images adapted from: Hagmann et al. (2008); Heuvel and Sporns (2011); Wang et al. (2016).

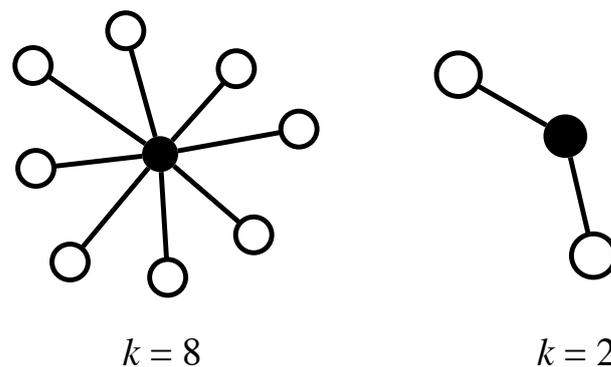


Fig. 2.3 Degree of connectivity. Diagram depicting a high-degree node (left, degree  $k = 8$ ), and a low-degree node (right, degree  $k = 2$ ).

observe (the *function*). The ability to capture and analyse these aspects of a network confers both descriptive and predictive capabilities, and the potential to obtain a deeper understanding of a network's underlying governing principles. At the most basic level, for both points 1 & 2, we are interested in aspects of connectivity; as such, one of the most fundamental measures in network theory is a node's *degree*, which expresses the number of other nodes a given node is connected to (see Figure 2.3).

From an information-processing perspective, a neuron's degree can be thought of as a proxy for its importance in a network: high-degree neurons represent broadcasting hubs that can influence many other neurons (if the connections are primarily outbound), receive and integrate information from many sources (if the connections are inbound), or some combination of the two such as might be required to integrate multiple sensory signals and coordinate an appropriate response. Indeed, this is the case for the high-degree nodes in *C. elegans*, which are almost exclusively premotor interneurons (Morita et al., 2001; Towlson et al., 2013).

Although the degree only directly measures the connectivity of a single node, the *degree distribution* of all the nodes in a network can provide insights into the large-scale structural organisation of the network. The degree distribution can reveal, amongst other things: the probability and randomness of connections, the quantity of high-degree hubs, whether a network has sparse or dense connectivity, and further provide information on the robustness, growth, and dynamical properties of a network (Newman, 2010; Wang et al., 2006).

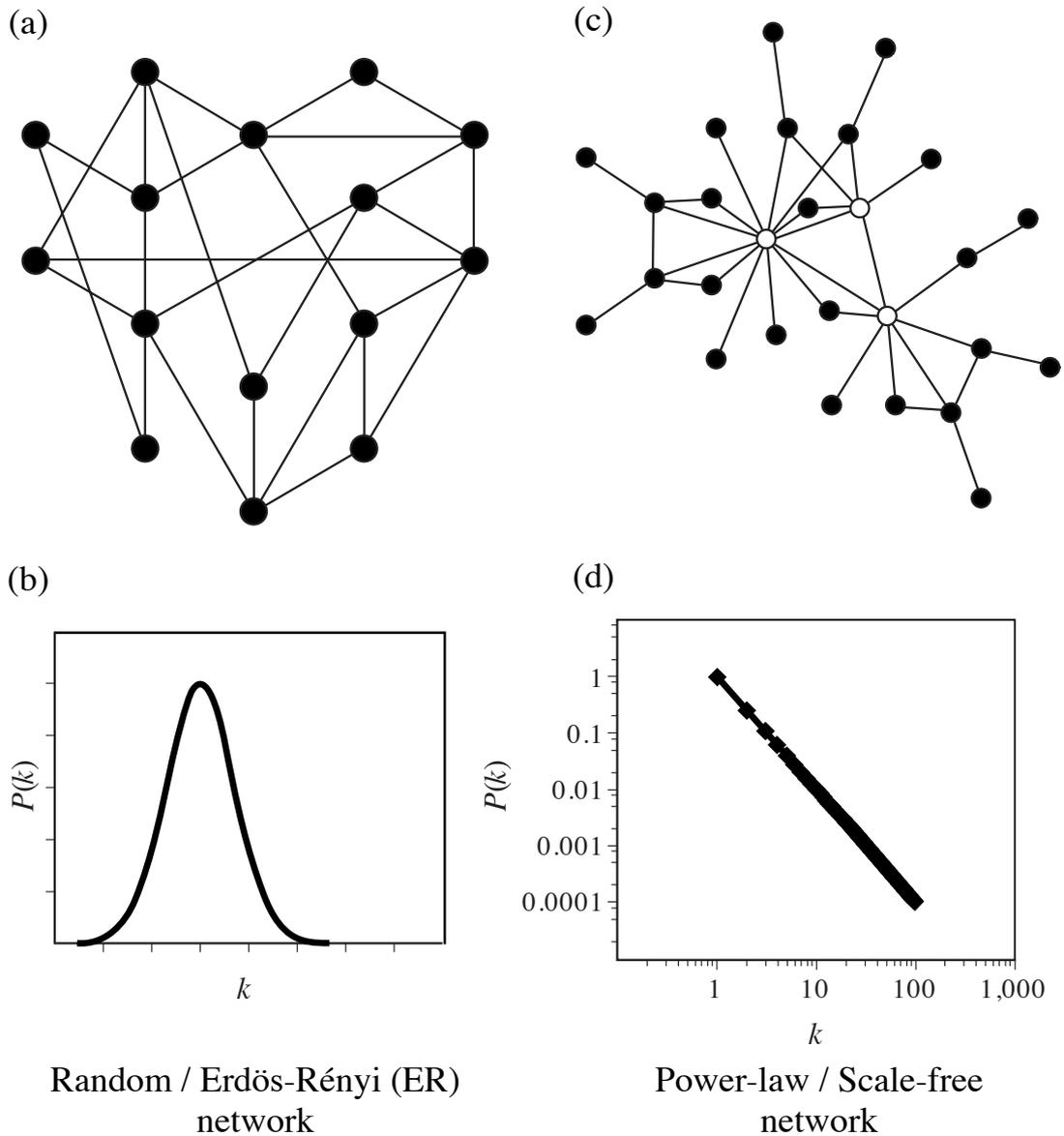


Fig. 2.4 Example degree distributions: (a) structure of a random network generated using the ER model; (b) the degree distribution for a random ER network, following a binomial distribution; (c) example structure of a scale-free network, characterised by many low-degree nodes connected to a small number of high-degree hubs, shown in white; and (d) power-law degree distribution for a scale-free network, following  $P(k) \propto k^{-\alpha}$ , where  $\alpha$  is the power-law exponent (plotted on a log-log scale). Adapted from Barabási and Oltvai (2004).

An illustration of how the degree distribution and structure of a network are related can be seen in Figure 2.4, where two commonly studied distributions are shown. The first, in Figure 2.4a, is a *random network* where the edges connecting nodes are assigned at random from a uniform distribution, known as the *Erdős-Rényi* or ER model. Although the placement of edges is entirely random, for sparse networks the process results in a highly-stereotyped binomial degree distribution where the nodal degrees are tightly distributed around the mean (Erdős and Rényi, 1959). The regularity that emerges from this connectivity results in predictable large-scale properties.

The second network (in Figure 2.4b) is a *scale-free* network, so called because it has no meaningful average degree, emerges from a power-law degree distribution of the form  $P(k) \propto k^{-\alpha}$ , where  $\alpha$  is the power-law exponent typically in the interval  $2 < \alpha < 3$  (Barabási and Albert, 1999; Clauset et al., 2009). Scale-free networks are often observed in real-world systems, and are characterised by many low-degree nodes connected by high-degree hubs. One inherent characteristic of scale-free networks is their high resilience to physical insults, as the loss of nodes or edges at random results in gradual, rather than catastrophic, degradation (Achard et al., 2006; Albert et al., 2000). Although these are just two examples, they illustrate how the degree distribution can reflect the global structural and functional characteristics of a network.

Along with the degree distribution, a second large-scale network property that can be determined from the degree metric is the *degree-degree correlation*, also known as the *assortativity coefficient*. The degree-degree correlation measures the correlation between the degrees of nodes on either side of an edge and describes whether a network has uniform (*assortative* or *homophilic*) or star-like (*disassortative* or *heterophilic*) connectivity; that is to say, whether nodes preferentially attach to nodes with a similar degree or not (Hao and Li, 2011; Newman, 2002, 2003a; Pastor-Satorras et al., 2001). The effect of degree-degree correlation on the large-scale structure of a network can be observed in Figure 2.5, which shows an example of a disassortative network where the low-degree nodes connect only to high-degree nodes (Figure 2.5, left); and an assortative network, with low-degree nodes connecting primarily to other low-degree nodes, and high-degree nodes mutually interconnecting to form a central cluster (Figure 2.5, right).

As with the degree distribution, the degree-degree correlation can also affect the functional characteristics of a network. For example, assortative networks have been shown to be highly robust to failure from circuit degradation (Newman, 2002; Rubinov

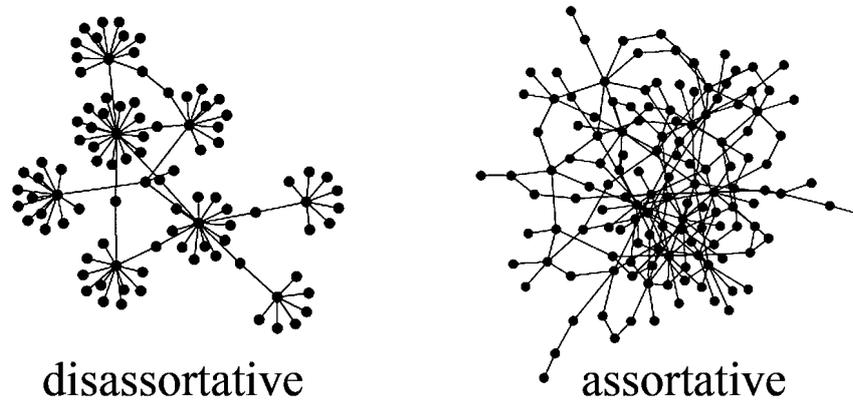


Fig. 2.5 Assortativity. Example networks showing the effects of assortativity on network structure. Adapted from Hao and Li (2011).

and Sporns, 2010; Teller et al., 2014), while disassortativity facilitates communication and coordination between nodes in a network (Perc et al., 2013; Sorrentino et al., 2006; Wang et al., 2008).

So far, we have looked at ways in which the connectivity of individual nodes can be measured (*degree*), how such connections are distributed throughout the nodes in a network (the *degree distribution*), and how the connectivity of a node relates to the connectivity of its immediate neighbours (the *degree-degree correlation* or *assortativity coefficient*). All of these measures can provide insights into the functional properties of the network. Refocusing on the application of network theory to connectomic analysis, the following section will introduce additional measures that allow for the determination of higher-order structures, in turn providing more fine-grained descriptions of network organisation that can be used to study the nervous system.

### 2.2.2 Higher-order structural organisation

To generate useful behaviours, the nervous system must structure the flow of information through its networks. Although basic nodal-connections provide the primary structure, these connections must be arranged into higher-order patterns capable of processing information. In the brain, this is achieved through the combination of two general approaches: *functional segregation* & *functional integration* (Sporns, 2013b; Tononi et al., 1994), visualised in Figure 2.6. As we will see, tools from network theory allow us to study these properties in detail.

### 2.2.2.1 Functional segregation

Functional segregation provides the brain with the ability to develop specialised processing units (Rubinov and Sporns, 2010); this is characterised by the presence of *modules*. Functional segregation is prominent across various scales in the brain, from microstructural cortical columns, to the separation and specialisation of entire brain regions. In *C. elegans*, this is most apparent in the pharyngeal nervous system, which is almost entirely disconnected from the main somatic network with the exception of a small number of gap junctions, and is dedicated to controlling feeding behaviour (Albertson and Thomson, 1976).

Several measures of functional segregation exist. These include *clustering* (Watts and Strogatz, 1998), which describes the cohesiveness of a group of nodes by measuring the fraction of neighbours shared between nodes (forming triangles); and *modularity* (Newman and Girvan, 2004), which describes the extent to which a network can be divided into modules (defined as groups of nodes that have high within-group connectivity and low between-group connectivity); see Figure 2.6. These two patterns of connectivity are often related: as groups of nodes become selectively more interconnected to one another (giving rise to modules), they also form more common neighbours, and hence

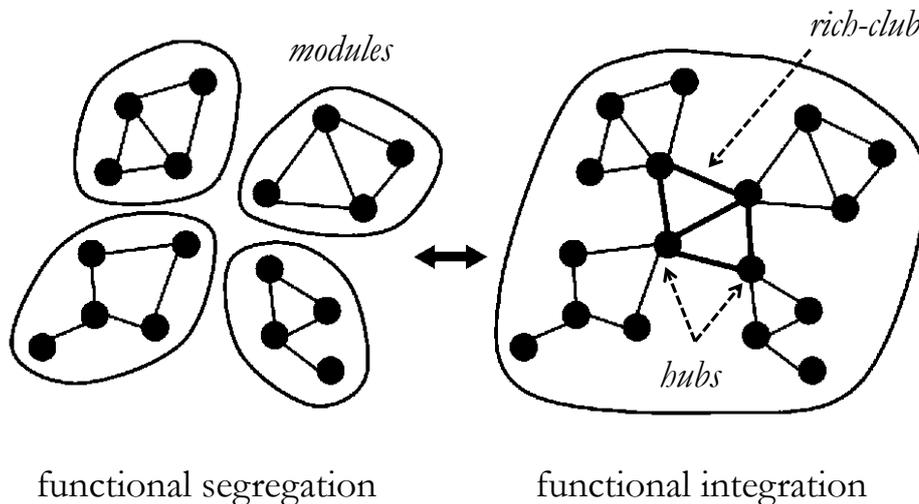


Fig. 2.6 Schematic diagram showing: functional segregation and the existence of network modules composed from clustered nodes, forming triangles (left); and functional integration from the connection of modules through hub nodes, forming a central rich-club (right). Figure modified from Sporns (2013b).

more clustered triangles. These properties allow a group of nodes to exert a greater degree of local influence over one another, effectively reducing their coupling to the rest of the network and allowing functional specialisation to emerge (Sporns and Betzel, 2016).

### 2.2.2.2 Functional integration

While functional segregation is clearly an important organisational feature, the purpose of the nervous system is to *coordinate* behaviour. The nervous system, for the most part, acts as a coherent entity, integrating various modalities to generate a single behavioural trajectory. An important feature of the nervous system is therefore the ability to perform functional integration. At the network level, measures of functional integration estimate the global interconnectedness and the ease with which regions can interact. At the most basic level, this is measured by the number of nodes that are connected into a single *connected component*. If a group of nodes are completely disconnected from the main network, there is no way for them to be integrated into the larger behaviour of the system. Within a connected component, two useful measures of integration are the *characteristic path length* and *reciprocity*.

The characteristic path length gives the average shortest distance between all of the nodes in the network (Watts and Strogatz, 1998), and is related to integration due to the fact that if nodes are highly integrated (i.e. highly connected), then the average path length between those nodes will be low. This measure is also related to a network's *global efficiency* (Latora and Marchiori, 2001): As the distance between nodes decreases, so too does the energetic cost of interacting.

Reciprocity measures the fraction of reciprocal connections between nodes. These connections provide bidirectional couplings that can tightly integrate regions, while also providing highly efficient routes for various control mechanisms; the classical example being the feedback loop. In the nervous system, reciprocal connections play an important role in integrating neural activity from different systems (Edelman and Gally, 2013) and have been shown to have numerous biologically important functions, from the regulation of sensory signals to the provision of associative memory (Dayan and Abbott, 2005; Getting, 1989; Li et al., 2012; Sommer and Wennekers, 2003).

Other structural properties are also indicative of functional integration. These include the presence of highly-connected *hubs* that facilitate communication between

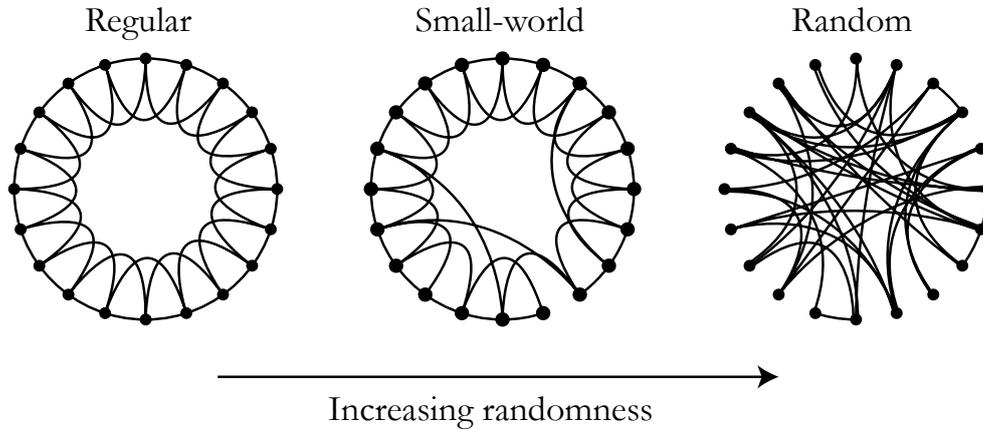


Fig. 2.7 Small-world networks balance the opposing demands of integration & segregation, and are characterised by having higher-than-random clustering and a shorter-than-random characteristic path length. Figure adapted from Watts and Strogatz (1998).

many nodes (Barabási and Albert, 1999), and the interconnection of such hubs into central structures known as *rich-clubs* (Colizza et al., 2006; Zhou and Mondragon, 2004). Hubs and rich-clubs play an important role in linking the functionally segregated modules of a network (see Figure 2.6) and have been observed to be a major topological feature in various brain systems (Harriger et al., 2012; Heuvel and Sporns, 2011; Liang et al., 2017; Reus and Heuvel, 2013; Shih et al., 2015), including the nervous system of *C. elegans* (Morita et al., 2001; Towlson et al., 2013).

### 2.2.2.3 Balancing segregation and integration: small-world structure

As described above, functional segregation and functional integration are both important features for information processing networks, yet they represent conflicting design objectives. As a network becomes more integrated, by definition it also becomes less segregated. The brain must therefore balance these two properties by simultaneously forming segregated modules and integrating them into an efficiently organised global network (Rubinov and Sporns, 2010; Sporns, 2013b; Tononi et al., 1994). A network that exhibits the coexistence of these features is described as being *small-world* (Watts and Strogatz, 1998); see Figure 2.7. Small-world networks feature locally-connected, close-knit modules, with shortcuts between modules that allow for efficient communication across the whole structure. In terms of network measures, this is seen in having higher-than-random clustering and a shorter-than-random characteristic path length (Watts

and Strogatz, 1998). The extent to which a network balances functional segregation and integration can thus be objectively quantified (Humphries and Gurney, 2008; Muldoon et al., 2016; Walsh, 1999; Watts and Strogatz, 1998).

#### 2.2.2.4 Connection motifs

Although the balance of functional segregation and functional integration is important in determining the general shape and characteristics of a network, these measures still only provide coarse-grained descriptions. Identifying the presence of a module does not, in itself, tell us much about its function. It is useful, therefore, to examine circuits at a lower level. One approach that has proven effective in this endeavour is that of *motif analysis*.

Within many complex real-world networks, it has been observed that the same patterns of connectivity are often repeated (Milo, 2002). These patterns are termed *motifs* (see Figures 2.8). A simple example of a motif has already been mentioned above, namely the feedback loop, though many more patterns exist: e.g. Figure 2.9 shows the 13 possible motifs that can be constructed from three connected nodes with directed edges. The properties of many of these motifs have been studied in isolation and found to have well defined functions, giving rise to the hypothesis that motifs form the functional building blocks from which information processing networks are composed. It has therefore been suggested that examining the constituent motifs of a network can provide insights into its functionality and mechanisms of operation (Alon, 2007; Kashtan et al., 2004; Sporns and Kötter, 2004). For example, if a motif has previously been observed to exhibit a well defined function, such as coincidence detection, then the presence of the same motif in a second network would suggest that the second network also performs coincidence detection, or utilises coincidence detection as an elemental operation in a more complex process.

It is clear that this type of functional composition is prevalent across various scales in biology, with repeated subsystems often forming the components of a larger system; indeed, this is an essential feature of the nervous system. As well as providing a means to simplify a system, the reuse of motifs in biological systems has been suggested as a method to increase functional adaptability, as the same elements can be recombined to produce multiple behaviours, and thus increase the responsiveness of an organism to changing environmental demands (Kashtan and Alon, 2005).

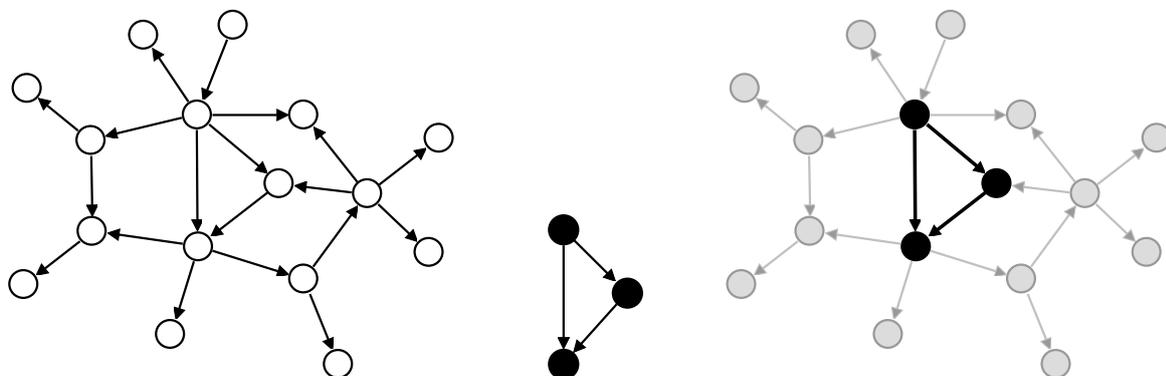


Fig. 2.8 Motif enumeration, showing: example network (left); the motif being searched for (middle, same as *motif 5* in Figure 2.9); and an instance of the motif in the target network (right). Adapted from Schreiber and Schwöbbermeyer (2008).

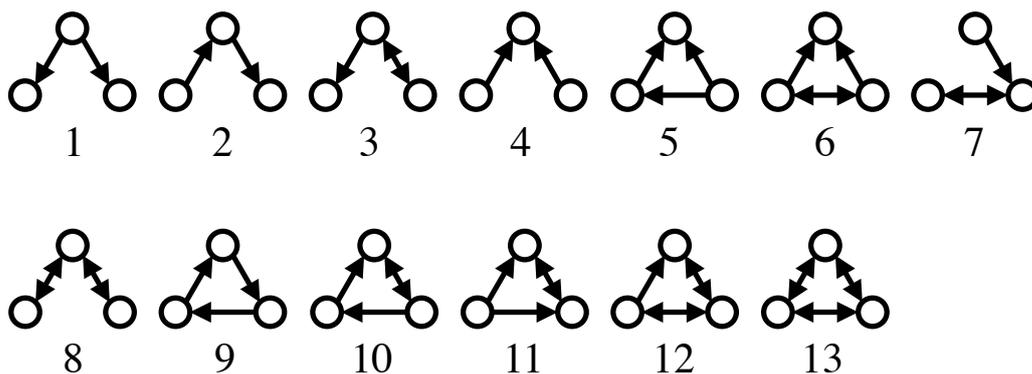


Fig. 2.9 Network motifs. Diagram showing all 13 possible three-node motifs with directed edges.

That repeating patterns of connectivity exist in the brain has been known for some time, many having previously been studied and characterised (Getting, 1989). The existence of connectomic maps now allows us to quantify these patterns and their distributions throughout an entire nervous system, to generate mechanistic theories to explain behaviours associated with particular structures, or predict the function of a structure from the presence and interconnection of motifs (Qian et al., 2011; Sporns and Kötter, 2004). Furthermore, reducing a network to a collection of repeating stereotypical subunits can reduce the complexity of the system and aid in the visualisation and exploration of connectivity data (Dunne and Shneiderman, 2013).

### 2.2.3 Null model reference networks

Despite the power of motif analysis in describing network architectures, care must be taken with interpretation. Motifs will emerge spontaneously in any network with connections, regardless of whether or not there is any underlying function or organising principle. When analysing a large uncharacterised network it is often necessary, therefore, to ask not just which motifs are present in the network, but also whether the identified motifs are statistically over- or under- represented compared to random (Milo, 2002); the aim of this approach is to identify conserved units that are unlikely to exist by chance alone, and exclude random patterns from the analysis. Such a comparison can be performed by normalising the observed frequency of motifs in the network against expected frequencies averaged from multiple comparable random networks. These random networks are termed *null model* (or *null hypothesis*) networks, and a collection of such networks is called a *network ensemble*.

The use of null models applies not only to motif analysis. Many network measures can be influenced by random processes, or by underlying structural features, such as the density of a network (Poisot and Gravel, 2014; Strang et al., 2017). For example, two separate networks with qualitatively similar architectures might appear different on some network measures due only to differences in the number of nodes and edges they contain. As the number of edges increases, the average path length in the network will decrease (as there will be more available paths). Similarly, a network embedded in 3D space will have an increased clustering coefficient simply due to the fact that some nodes will be closer to others. By generating comparable null model networks with the same constraints (such as the number of edges or spacial embedding), we can determine whether the constraints alone are sufficient to explain the properties we observe in the real network. Specifically, we can compare the distribution of results

obtained from a random statistical ensemble of null networks to those observed in a real-world network, and establish whether they are likely to come from the same distribution.

An alternative use of null model networks is in the normalisation of measures to allow for comparisons. We can use the ensemble averages as a baseline against which to normalise the real-world measures, and control for any variability caused by those features captured in the null model, such as the network size. This approach allows us to then directly compare networks with different structures, by controlling for those differences built into the null model.

There are various types of null models, each with different underlying assumptions or constraints that can affect the results of a network analysis (Artzy-Randrup et al., 2004; Betzel and Bassett, 2017; Hosseini and Kesler, 2013; Klimm et al., 2014; Medaglia and Bassett, 2017). The simplest form of null model was mentioned earlier in this chapter, namely the random network, or Erdős-Rényi (ER), model. ER networks are synthetic networks with the same number of nodes as the real network, but with the edges between the nodes assigned at random. Depending on the method of construction, these networks can either be created with the same number of edges, or instead with a specified nodal connection probability.

Although the ER model is conceptually simple, with properties that make it convenient to interrogate analytically (Hofstad, 2017), we have already seen how the random placement of edges leads to a binomial degree distribution that is not a good fit for most real-world networks (see Figure 2.4). As the degree distribution can influence many network properties, such as the number of hubs, it is often desirable to also control for this.

There are many null models that can generate networks with a prescribed degree distribution. Research into the optimal choice of null model is still an ongoing area of research, however two of the most common methods that preserve the degree-distribution of the original network are: (1) the *configuration model* (Molloy and Reed, 1995), and (2) the *edge-swap* or *rewiring* model (Maslov and Sneppen, 2002; Milo et al., 2003).

The configuration model creates synthetic networks in a manner similar to the ER method; however, it includes an additional constraint to ensure that the degree of each node equals the degree of a corresponding node in the original network, thus

resulting in random networks with the same degree distribution. The edge-swap model, on the other hand, creates copies of the original network which are then randomised by swapping the edges between nodes. This is performed by selecting a pair of edges ( $A \rightarrow B$ ) ( $C \rightarrow D$ ) and swapping them to give ( $A \rightarrow D$ ) ( $C \rightarrow B$ ), resulting in randomised networks with the same nodal degrees, and degree distributions, as the original. By parameterising the number of swaps to perform, the edge-swap method can also control the extent of randomisation that is performed.

For most of the analyses presented in later chapters the edge-swap method is used as the primary null model. A comparison of the two degree-preserving methods can be found in Appendix E (page 241), showing no significant difference between the null models for the measures and networks investigated in this thesis.

#### 2.2.4 Network robustness

The brain, as with many biological systems, is capable of exhibiting a high degree of adaptability and fault tolerance, with behaviours and functions often degrading “gracefully” in response to insults, rather than catastrophically; in some cases completely compensating for the lost functionality (Farah and McClelland, 1991; Li et al., 2016; Rumelhart and McClelland, 1986). This is also true for the small nervous system of *C. elegans*: for certain functions the death of neurons can be compensated for through redundant pathways, resulting in minimal behavioural defects (Avery and Horvitz, 1989; Bargmann and Horvitz, 1991; Chung et al., 2013; McIntire et al., 1993; White et al., 2007).

Network theory can provide insights into the resilience of a network by simulating damage and studying how the system is affected. This approach allows estimates to be made about the robustness of the connectivity, and also identify network features that confer either fault tolerance or vulnerability.

The most common means of analysing the robustness of a network is to perform an iterative series of modifications, or attacks, to the structure and observe how the properties of the network change in response. The main measures of interest are usually related to the network’s total cohesiveness and ability to process information, such as its characteristic path length (i.e. the average shortest distance between nodes) (Albert et al., 2000), connection efficiency (Crucitti et al., 2004; Latora and Marchiori, 2001), or the size of the largest connected component (Schneider et al., 2011a,b). As damage

occurs – breaking connections and fragmenting the network – the distance between nodes increases, resulting in lower efficiency, with node isolation reducing the size of the largest connected component. The degradation properties of the network can therefore be quantified and compared to null model networks, providing information on whether the extent and rate of failure in response to damage is higher or lower than expected compared to random connectivity.

A number of attack strategies can be employed when performing robustness analysis, depending on the features of the network under investigation. An analysis can be limited to remove either nodes or edges, with the target selection being at random or according to some metric, such as selectively attacking nodes by their degree in descending order (Holme et al., 2002). This allows different aspects of the network to be investigated. For example, a disassortative star-like network might be highly robust to the removal of nodes or edges at random, but will fragment as the central high-degree hubs are removed (Newman, 2002; Rubinov and Sporns, 2010; Teller et al., 2014).

This method of analysing the robustness of a network from its structure has previously been applied to connectomes, finding that brain networks are frequently more resilient than comparable networks (Achard et al., 2006; Joyce et al., 2013; Kaiser et al., 2007; Vértés et al., 2011).

### 2.2.5 Multilayer networks

So far, we have implicitly only considered static networks (i.e. those with a time-invariant structure) with a single connection type; however, many real-world networks are not so simple. This is especially true for the nervous system. Neurons can communicate through various types of interactions, including chemical synapses, gap junctions, and neuromodulatory volume transmission (see § 1.1.3). Each of these signalling classes can be further subdivided, with the brain containing hundreds of molecular messengers, receptors, and channels, all operating on different timescales, often with non-linear interactions between them. While it is still not possible to measure much of this complexity, it is desirable to be able to represent and analyse those aspects that we are able to measure.

One way to describe these multifarious relationships using network theory is to model the different interaction types as separate layers in a *multilayer* (or *multiplex*) network (Battiston et al., 2014; Boccaletti et al., 2014; Kivela et al., 2014; Vaiana

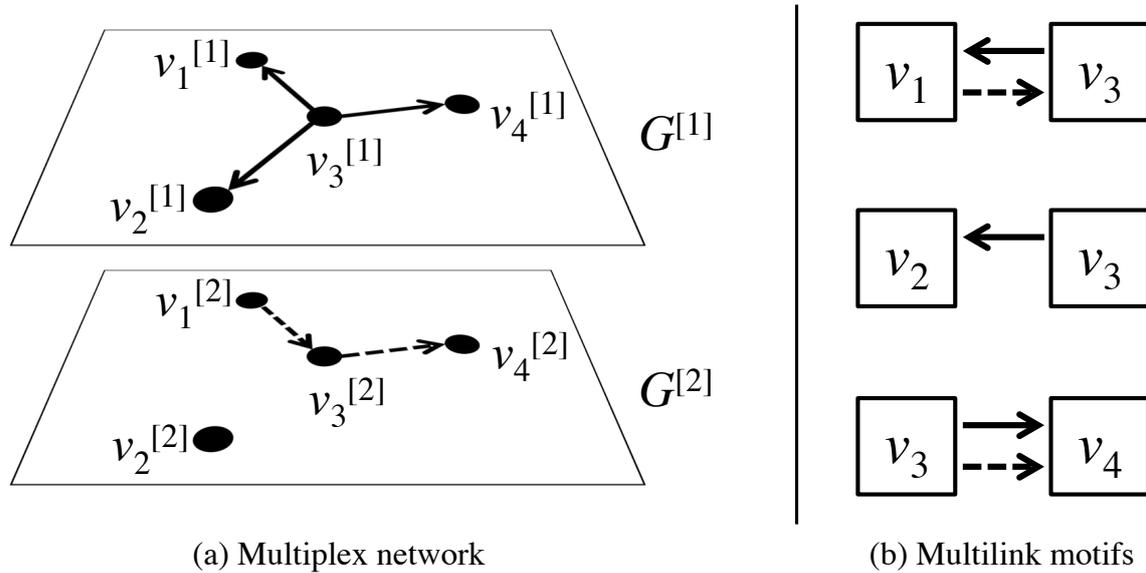


Fig. 2.10 Diagram showing: (a) a two-layer multiplex network containing four nodes, and (b) the corresponding multilinks between each pair of connected nodes. Solid arrows represent edges in layer  $G^{[1]}$  and dashed arrows are edges in layer  $G^{[2]}$ .

and Muldoon, 2017). Here, the real network is decomposed by its interaction types, creating individual graph models for each. These can then be considered alongside one another without the loss of information, and provide insights into how the different interaction types relate to one another. As the structure of a network can change over time – with new connections being established or removed – the multilayer approach can also be used to represent different time points of a network’s development, with each layer containing a snapshot of the structure at a given moment. Information on the dynamics of a network’s connectivity can be obtained from analysing these layers in order, and in relation to one another.

An example of a simple two-layer multiplex network is shown in Figure 2.10a, containing four nodes  $\{v_1, \dots, v_4\}$ . These nodes are the same in each of the two layers, with only the connections varying. These layers could, for example, be used to represent two different relationship or interaction types between the same nodes, or the same network at two different points in time.

Various approaches and tools exist for analysing multilayer networks, including extensions of the various methods already mentioned for monolayer networks. For

example, the degree-degree correlation considered earlier can be applied to measure the interlayer degree correlation (de Arruda et al., 2015), to ask whether a node that has a high degree of connectivity in one layer is also highly connected in another. This can reveal nodes that are important in linking the different interaction layers, facilitating communication between them. The interlayer degree correlation can further identify whether layers overlap (in which case they will be positively correlated), or are distinct (having no correlation, or a negative correlation).

Looking at where connections fall relative to one another across the different layers can reveal structures and patterns that are conserved and reused throughout the network, similar to the concept of motifs introduced earlier. The pattern of edges between two nodes in a multilayer network has been termed a *multilink* (Bianconi, 2013; Menichetti et al., 2014; Mondragon et al., 2017). The arrangement of *multilink motifs* can provide important information on how the different layers are organised in relation to each other. As before, the multilink motif counts can be compared to those from null model networks to identify statistically overrepresented motifs that form the functional building blocks of the network. Example multilink motifs are shown in Figure 2.10b. As we will see in Chapters 4 & 5, multilink motif analysis can be used to study how different signalling channels coordinate their actions in a nervous system.

Another powerful method for investigating the similarity of network layers is multilayer *reducibility analysis* (De Domenico et al., 2014, 2015a, 2016; Wang and Liu, 2017). This method quantifies the similarity or uniqueness of layers by aggregating (or collapsing) layers into a single layer, and determining whether any information is lost in the process, usually by measuring the change in von Neumann entropy. Thus, two network layers that perfectly overlap (i.e. have the same connections in both layers) will have the same entropy once aggregated, and can be said to be redundant and reducible. These concepts will be explored further in § 2.5.1.

## 2.3 Foundations & formal definitions

The previous section introduced some of the concepts from network theory that can be applied to the analysis of neural connectomes. The following sections will expand on these concepts, providing the formal definitions of the data structures and network measures that were introduced. These are used in the subsequent chapters to analyse the synaptic and extrasynaptic connectomes of *C. elegans*.

### 2.3.1 Graphs

As mentioned previously, the network structure of the nervous system can be modelled as a graph, denoted  $G$ . Formally, a graph is defined as a 2-tuple  $G = (V, E)$ , where  $V = \{v_1, \dots, v_N\}$  is the set of all vertices (nodes), and  $E = \{e_1, \dots, e_M\}$  is a set of ordered pairs of vertices ( $E \subseteq V \times V$ ) representing the edges, or connections, between those vertices. An edge between nodes  $v_i$  and  $v_j$  can be written as  $e_{ij} = (v_i, v_j)$ . Applying this to a microscale synaptic connectome,  $V$  would represent the complete set of all neurons, and  $E$  the set of synaptic connections between them (see Figure 2.11). This model, however, can equally be applied to macroscale connectomes or functional maps such that  $V$  is some set of defined anatomical regions, and  $E$  is the structural or functional associations between them.

In the context of *C. elegans*, and the topics covered in this thesis, we will only consider the former microscale definition. As the adult hermaphrodite *C. elegans* has a well defined and largely immutable nervous system consisting of 302 post-mitotic neurons (Sulston and Horvitz, 1977; Sulston et al., 1983; White et al., 1986), the set of vertices  $V$  remains constant; however, the interactions, or edges  $E$ , can include chemical synapses, electrical gap junctions, or any one of a number of molecular volume transmission interactions (see § 1.1.3). The definition of  $E$ , and the interactions being analysed, will be specified in each case. Given this definition of a network, the number of nodes (neurons) is given by  $N = |V|$ , and the number of edges (synapses or VT interactions),  $M = |E|$ .

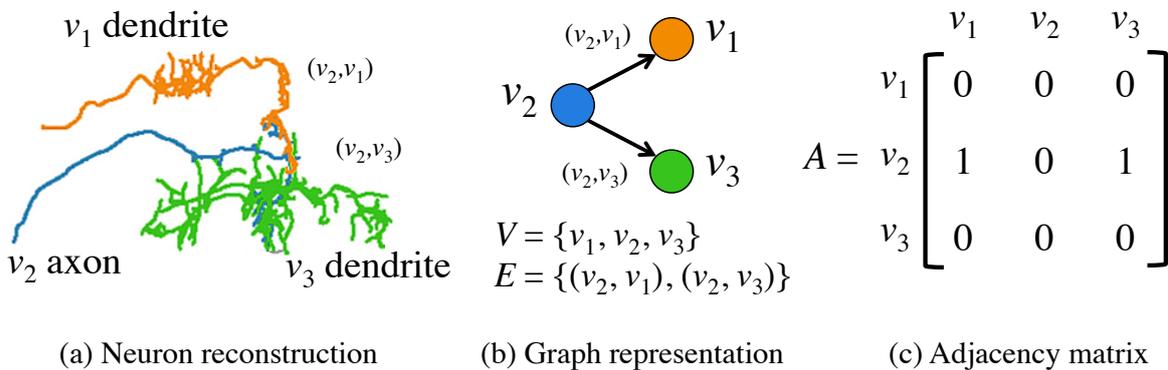


Fig. 2.11 Representations of neural connectivity at different levels of abstraction. Example neuron reconstruction (a) taken from Schneider-Mizell et al. (2016).

### 2.3.2 Adjacency matrices

For computational and analytical purposes, a more convenient representation of a graph can be provided by the *adjacency matrix*, denoted  $A$ .

*Adjacency* refers to two nodes that are connected by an edge; as such, the adjacency matrix is an  $N \times N$  matrix encoding the connectivity of nodes, with the edges between nodes represented as positive values in the respective matrix elements (see Figure 2.11c). For a binary unweighted network (i.e. a network with single unweighted connections between nodes) this takes the form:

$$A_{ij} = \begin{cases} 1 & \text{if } (v_i, v_j) \in E \\ 0 & \text{if } (v_i, v_j) \notin E \end{cases} . \quad (2.1)$$

Thus, given two nodes  $v_i, v_j \in V$ , the adjacency matrix element at  $A_{ij}$  will be  $A_{ij} = 1$  if and only if there is a connection from node  $v_i$  to node  $v_j$ , or  $A_{ij} = 0$  otherwise. Note that for a *directed graph* (or *digraph*) this representation is ordered to encode the directionality of the relationship; e.g. if node  $v_i$  sends a connection to node  $v_j$ , but node  $v_j$  does not return a reciprocal connection to  $v_i$ , the values in the adjacency matrix will be:  $A_{ij} = 1$ ,  $A_{ji} = 0$ .

For analyses where directionality is irrelevant, or for connections that lack directionality such as bidirectional gap junctions, an undirected graph can be used. In this case the adjacency matrix is symmetric such that  $\forall A_{ij} A_{ij} = A_{ji}$ . This also applies to the graph notation: For undirected networks it is implicit that if an edge exists between two nodes, e.g.  $(v_i, v_j) \in E$ , then the corresponding reciprocal connection  $(v_j, v_i) \in E$  also exists, though the two are only ever counted as a single bidirectional edge. In directed networks, these would be counted as two separate connections. The convention adopted in this thesis is to explicitly identify directed graphs and measures with a superscript arrow (e.g.  $G^{\rightarrow}$ ). For all adjacency matrices, self-connections are represented on the diagonal ( $A_{ij}$  where  $i = j$ ).

### 2.3.3 Weighted graphs

The same adjacency matrix structure can also be used to represent multiple connections between nodes (for a *multigraph*), or edge weights representing an edge property (for a *weighted graph*), such as the connection strength, delay, distance, or communication

cost. A multigraph can be modelled as a special instance of a weighted graph, where the weights represent the number of edges. In both cases, the graphs are encoded as adjacency matrices by assigning connection weights to the respective matrix elements. For a set of edge weights  $W$ , where  $w_{ij} \in W$  is the connection weight for the edge  $(v_i, v_j) \in E$ , the adjacency matrix is thus given by:

$$A_{ij} = w_{ij}. \quad (2.2)$$

In connectomic modelling, these weights can be used to represent any measure of interest, with the interpretation of the network changing accordingly. For example, connection weights can be used to represent the number of synapses between neurons, the distance between cells, neuron similarity, the strength or temporal properties of functional associations, or any other desired numerical measure.

In the analyses presented in later chapters, connection weights are disregarded and only binary networks are used. That is to say, we only consider whether a connection exists, and not the number of synapses or receptors expressed. This decision was made to avoid making assumptions about the strength, polarity, linearity, or state-dependent nature of synapses – for which there is little available data – and confine the analyses to the structural organisation of the networks, for which there is a high degree of confidence.

### 2.3.4 Simple graphs & maximum edges

We have now seen that there are different types of graphs (i.e. directed / undirected, binary / weighted). In all of these cases, it is important to note that the maximum number of possible edges can change. The most typical form of graph is the *simple graph*. This is the default form used in this thesis when not specified otherwise.

Simple graphs contain only undirected and unweighted connections between nodes, and have no self-loops (i.e. a node connecting to itself). As such, the maximum number of possible edges is given by the binomial expression:

$$\binom{N}{2} = \frac{N(N-1)}{2}. \quad (2.3)$$

This can be seen when considering that every node is capable of connecting to every other node apart from itself ( $N(N-1)$ ), excluding reciprocal connections (divide by 2).

In the case of directed graphs, where reciprocal edges can exist, the maximum number of edges is simply  $N(N - 1)$ , or  $N^2$  if self-loops are also allowed.

### 2.3.5 Node neighbourhood

Having established the formal definitions for graphs and adjacency matrices, we can begin to look at some of the network measures, tools, and relationships in network theory. A simple relationship, important in many graph algorithms, is that of the node (or subgraph) neighbourhood,  $Nh$  (illustrated in Figure 2.12b). For a given node – or a given set of nodes in the case of a subgraph – the neighbourhood is simply the set of adjacent nodes that it is connected to, while excluding itself. For a subgraph of interest, the neighbourhood of the subgraph node set  $V_{sub} \subset V$  can be written:

$$Nh(V_{sub}) = \{v_j \in [V \setminus V_{sub}] \mid e_{ij} = (v_i \in V_{sub}, v_j \in [V \setminus V_{sub}]) \in E\}. \quad (2.4)$$

To elaborate, this returns those nodes  $v_j$  that are not in the subgraph node set  $V_{sub}$ , but which receive an edge from a node in  $V_{sub}$  and are thus in its neighbourhood.

## 2.4 Network measures

### 2.4.1 Degree

For a node to participate in a network, it must be able to interact with the other nodes in that network through connections; therefore one of the most basic, and important, measures in network theory is a node's *degree*. The degree, denoted  $k$ , measures how many other nodes in the network a specific node is connected to. The degree  $k$  of a node  $v_i$  is therefore:

$$k_i = \#\{v_j \in V \mid (v_i, v_j) \in E\}. \quad (2.5)$$

For an adjacency matrix, this can easily be obtained by summing the matrix elements:

$$k_i = \sum_{j \in V}^N A_{ij}. \quad (2.6)$$

For a directed network, the degree can be further decomposed into a node's in-degree and out-degree, measuring the in-coming (receiving) connections and out-going

(sending) connections, respectively:

$$k_i^{in} = \sum_{j \in V} A_{ji} \quad \text{and} \quad k_i^{out} = \sum_{j \in V} A_{ij}. \quad (2.7)$$

The total degree of a node in a directed network can be found by summing the in-degree and out-degree values:

$$k_i^{\rightarrow} = k_i^{in} + k_i^{out}. \quad (2.8)$$

The degree provides a measure of connectivity for an individual node; though as we saw earlier, the combined degree information of all the nodes in a network, in the form of the degree distribution, can also provide information on the large-scale structure of a network. The degree distribution is the statistical distribution of all the nodal degrees, often expressed as a probability function of a node selected uniformly at random having a given degree,  $P(k)$ .

Two commonly studied degree distributions in network theory are (1) the binomial degree distribution of random Erdős-Rényi (ER) networks, and (2) the power-law distribution of scale-free networks (see Figure 2.4). These are covered in sections § 2.4.10.1 and § 2.4.12, respectively.

## 2.4.2 Network density

A second useful summary measure of a network's global structure is its connection *density*,  $D$ , sometimes also called the *connectance*. Network density is the ratio of actual edges to the maximum possible number of edges. Density can affect many properties of a network by shaping the degree distribution (Poisot and Gravel, 2014) and influencing network measures such as the clustering coefficient and global efficiency (Strang et al., 2017).

For undirected and directed networks the density measure is respectively given by:

$$D(G) = \frac{2M}{N(N-1)} \quad \text{and} \quad D^{\rightarrow}(G^{\rightarrow}) = \frac{M}{N(N-1)}, \quad (2.9)$$

with both variants falling in the interval  $0 \leq D \leq 1$ , where  $D = 0$  corresponds to a network with no connections, and  $D = 1$  corresponds to a fully connected network. A network is considered *sparse* if the number of edges is much less than the maximum, i.e.  $M \ll \frac{1}{2}N(N-1)$ , with  $D \rightarrow 0$  as  $N \rightarrow \infty$ ; whereas a dense network is one where

$D$  approaches some constant as  $N \rightarrow \infty$  (Newman, 2010).

Empirical studies of real-world networks have found that many networks are sparsely connected (Barabási and Pósfai, 2016). While in principle dense connectivity allows for a greater degree of communication, the cost of building and maintaining connections makes it impractical. Dense connectivity is also unnecessary in most cases, as high communication efficiency can be achieved in sparse networks through high-degree hubs that link many nodes, such as in scale-free and small-world networks (Crucitti et al., 2003; Latora and Marchiori, 2001).

As well as being expensive, there is evidence that in some circumstances dense connectivity might also be detrimental to dynamical processes on a networks, such as those that occur in the brain. Results from analytical and computational studies have suggested that densely connected networks have inherently unstable dynamics (Gardner and Ashby, 1970; May, 1972). In the brain, this is supported by the observation that epilepsy and seizures are often associated with hyperconnectivity, and possibly increased neural coupling via gap junctions (Caeyenberghs et al., 2014; Jin and Zhong, 2011; Volman et al., 2011).

### 2.4.3 Connected components

A connected component is a group of nodes that are interconnected by edges, such that in an undirected network a path can be traced between any two nodes in the component. If a node is unreachable from any one of the nodes in the component, then it is not a part of the same component, and instead constitutes an element in a second connected component.

Examining the connected components in a network is useful for identifying separate functional subsystems. This can be seen from considering that if two groups of nodes are completely unconnected, then it is not possible for them to interact to form a single functional unit.

The connected components of a network can be enumerated using any graph traversal algorithm, such as depth-first search (DFS) or breadth-first search (BFS) (Levitin, 2012). By iteratively starting at different nodes in the network, following the edges, and recording which nodes are reachable from which, the number of connected

components and their constituent nodes can be determined (Hopcroft and Tarjan, 1973).

An example of separate connected components can be found in the synaptic connectome of *C. elegans*, where the functionally specialised pharyngeal nervous system has no synaptic connections with the main somatic network (Albertson and Thomson, 1976). Along with several other unconnected neurons, this results in a connectome with multiple connected components (see Figure 3.4, page 90).

#### 2.4.4 Path length & efficiency

An analysis of the connected components within a network can tell us which nodes have the potential to interact and form functional units, but this information on its own does not provide information on the degree of connectedness or extent of functional integration within the connected components of the network. These properties can, however, be estimated by the two related measures of characteristic path length and global efficiency.

The path length describes the shortest distance between two nodes – measured as the number of edges that must be traversed along the shortest path – while efficiency measures the inverse distance. These relationships can be seen from considering that as a network becomes more integrated, the path length between nodes decreases, resulting in increased communication efficiency. These measures can be computed for individual pairs of nodes, or averaged across all possible pairs to provide a description of the large-scale structure of the entire network and an estimate of the network’s total integration and wiring efficiency.

The average network measures are typically referred to as the *characteristic path length*, which is the mean shortest path length between all pairs of nodes; and the *global efficiency*, which is the mean of the inverse shortest path lengths.

To obtain the characteristic path length it is first necessary to compute the *geodesic path* (i.e. shortest path) between each pair of nodes. There are numerous approaches that can be used to compute the geodesic path (see Cormen et al., 2009; Głabowski et al., 2013 for a review). The classical approach is to perform a breadth-first search (BFS) of the network to compose a tree of the shortest paths to each node. One efficient version that can account for edge weights (where the weights encode a non-negative distance measure) is Dijkstra’s shortest path algorithm (DSP) (Cormen et al., 2009;

Dijkstra, 1959), given in Algorithm 1.

Having determined the geodesic path between a pair of nodes  $v_i, v_j$ , it is straightforward to calculate the distance between them. The *geodesic distance* can be found by summing the adjacency matrix elements along the geodesic path  $g$ :

$$d_{ij} = \sum_{(m,n) \in g_{ij}} A_{mn}, \quad (2.10)$$

where  $g_{ij}$  is an ordered set of edges forming the geodesic path between  $v_i$  and  $v_j$ , as returned by **DSP** in Algorithm 1. If there is no connection between  $v_i$  and  $v_j$  then  $d_{ij} = +\infty$ ; however it is common practice to exclude unconnected node pairs from the analysis.

We can then find the characteristic path length of a network by taking the average geodesic distance between each pair of nodes, given by:

$$L(G) = \frac{1}{N(N-1)} \sum_{i \in V} \sum_{j \neq i \in V} d_{ij}. \quad (2.11)$$

As previously stated, the characteristic path length is related to the *global efficiency* of a network, with the cost of communication being inversely proportional to the path lengths (Latora and Marchiori, 2001). The efficiency of a link between vertices  $v_i$  and  $v_j$  is defined as:

$$\epsilon_{ij} = \frac{1}{d_{ij}}, \quad \forall v_i, v_j \in V. \quad (2.12)$$

Unlike the geodesic distance, where unconnected nodes result in an infinite path length, the efficiency provides a more meaningful measure, with unconnected nodes having an efficiency of  $\epsilon = 0$ . The global efficiency of a network is simply the average efficiency of all possible links, given by the expression:

$$\mathcal{E}(G) = \frac{1}{N(N-1)} \sum_{i \in V} \sum_{i \neq j \in V} \epsilon_{ij}, \quad (2.13)$$

with values in the interval  $0 \leq \mathcal{E}(G) \leq 1$ , where  $\mathcal{E}(G) = 0$  corresponds to a completely unconnected graph containing no edges, and  $\mathcal{E}(G) = 1$  corresponds to a fully connected graph.

---

**Algorithm 1: DSP – Dijkstra’s shortest path.** Gives the geodesic distances and paths from a given node to every other node (Dijkstra, 1959).

---

**input** : Graph  $G = (V, E)$ ; index  $\eta$  of source vertex  $v_\eta \in V$  to measure all distances from; and set of non-negative edge weights  $W = \{w_1, \dots, w_M\}$ , where  $\forall w \in W, w \in \{x \in \mathbb{R} \mid x \geq 0\}$ . N.B: If the graph is unweighted ( $W = \emptyset$ ), or the weights represent some measure other than distance, weights must be created or reset to 1 for each edge, such that  $\forall (v_i, v_j) \in E, w_{ij} \in W \leftarrow 1$ .

**output**: Distance  $d$  and geodesic path  $g$  from  $v_\eta$  to every vertex in  $V$ .

```

[1] begin
[2]   /* Set distance to source vertex to 0 */
[3]    $d_{\eta, \eta} \leftarrow 0$ ;  $w_{\eta, \eta} \leftarrow 0$ ;
[4]   /* Initialise all distances  $d$  to the corresponding weights
      in  $W$ . If a vertex  $v_i$  is not adjacent to the source
      vertex ( $(v_\eta, v_i) \notin E$ ), its initial distance is  $d_{\eta i} = \infty$  */
[5]   for  $\forall v_i \in V$  where  $v_i \neq v_\eta$  do
[6]     if  $(v_\eta, v_i) \in E$  then  $d_{\eta i} \leftarrow w_{\eta i}$ ;
[7]     else  $d_{\eta i} \leftarrow \infty$ ;
[8]    $S \leftarrow \emptyset$ ; /* Set of visited vertices; initially empty */
[9]    $Q \leftarrow V$ ; /* Queue of vertices remaining to be visited */
[10]   $\Pi \leftarrow \emptyset$ ; /* Predecessor: used to record shortest path */
[11]  while  $Q \neq \emptyset$  do
[12]    /* Select vertex in  $Q$  with lowest distance from  $v_\eta$  */
[13]    select  $v_i \in Q$  where  $\forall v_j \neq i \in Q, d_{\eta i} \leq d_{\eta j}$ ;
[14]     $S \leftarrow S \cup \{v_i\}$ ;  $Q \leftarrow Q \setminus \{v_i\}$ ; /* Updated visited vertices */
[15]    /* For each vertex  $v_j$  adjacent to  $v_i$  check if the
      distance from  $v_\eta$  is improved by passing through  $v_i$  */
[16]    for  $\forall v_j \in Nh(\{v_i\})$  do Relax( $j, i, w$ );
[17]  for  $\forall v_i \neq \eta \in V$  do
[18]     $g_{\eta i} \leftarrow \text{BuildPath}(i)$ ; /* Get edges for shortest path to  $v_i$  */
[19]  return  $d, g$ ;
[20] function Relax( $j, i, w$ )
[21]   /* If the distance from  $v_\eta$  to  $v_j$  is reduced by passing
      through  $v_i$ , update record. */
[22]   if  $d_{\eta j} > d_{\eta i} + w_{ij}$  then
[23]      $d_{\eta j} \leftarrow d_{\eta i} + w_{ij}$ ;
[24]      $\Pi_j \leftarrow i$ ; /* Shortest path to  $v_j$  was reached via  $v_i$  */
[25] function BuildPath( $i$ ) /* Reconstruct shortest path */
[26]   /* Recursively backtrack  $\Pi_i$  to construct set of edges */
[27]   if  $i \neq \eta$  then
[28]      $g_{\eta i} \leftarrow \text{BuildPath}(\Pi_i)$ ;
[29]     return  $g_{\eta i} \leftarrow g_{\eta i} \cup \{(\Pi_i, v_i)\}$ ;

```

---

Finally, a useful summary measure of integration that can be obtained from the path lengths is a network's diameter  $\ell$ , defined as the maximum shortest path in the network – that is to say, the maximum distance that must be travelled to go from any one node to another, while only considering the most efficient paths. This is written:

$$\ell(G) = \max_{ij} (d_{ij}). \quad (2.14)$$

It is important to note that some papers in the literature also use the term “diameter” to refer to the characteristic path length.

## 2.4.5 Clustering

Recall that functional segregation is a term used to describe the local cohesiveness and interdependence of a group of nodes with the potential to form specialised processing units in the brain (Rubinov and Sporns, 2010).

A common network measure for identifying the presence of mutually interconnected groups of nodes – indicative of functional segregation – is the *clustering coefficient* (Watts and Strogatz, 1998). For a single node, the local clustering coefficient of that node can be described in one of two equivalent ways:

1. as the fraction of triangles formed with its neighbours, as a result of recurrent connections;
2. as the connection density of the node neighbourhood (see Figure 2.12).

This relationship is clear when considering that a neighbourhood with maximum density (forming a *clique*) will also form the maximum number of possible triangles, with the nodes correspondingly exhibiting a high degree of interdependence. Therefore, the presence of triangles – or high neighbourhood density – suggests dependencies between the nodes characteristic of functional segregation. As such, when calculated for all the nodes in a network, the clustering coefficient provides a useful summary measure of network organisation.

In an undirected network, the clustering coefficient  $C$  for a single node  $v_i$  can be computed either as the fraction of triangles  $t_i$  around that node, or as the neighbourhood density  $D$  (see Equations 2.4 & 2.9):

$$C_i = \frac{2t_i}{k_i(k_i - 1)} = D(\text{Nh}(\{v_i\})). \quad (2.15)$$

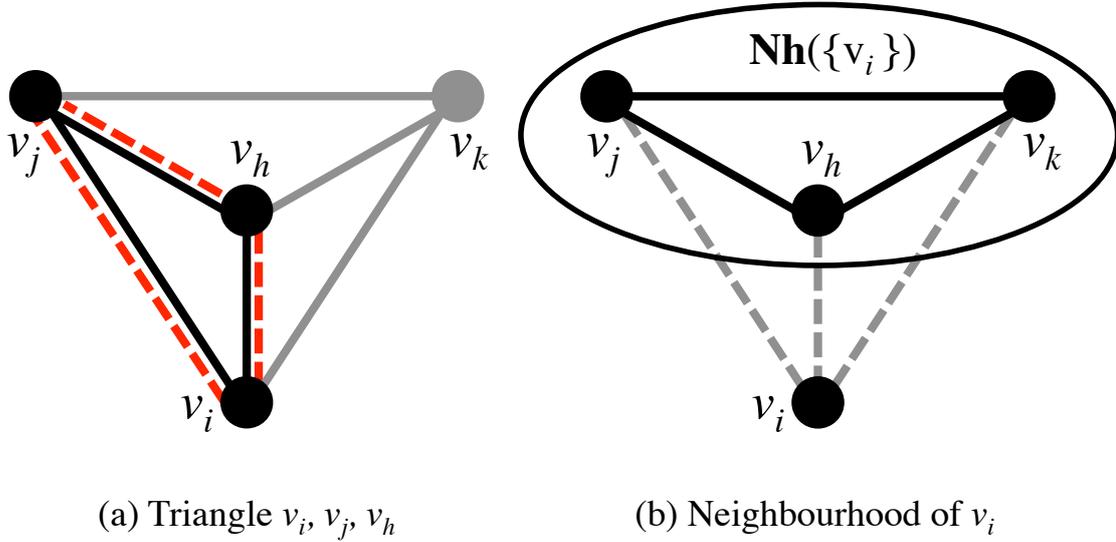


Fig. 2.12 Measures used in the determination of the clustering coefficient: (a) an instance of a triangle around node  $v_i$ ; and (b) the neighbourhood of  $v_i$ , with maximum density,  $D(\text{Nh}(\{v_i\})) = 1$ , and maximum clustering,  $C_i = 1$ .

Thus, the clustering coefficient has values in the interval  $0 \leq C \leq 1$ , with  $C = 1$  corresponding to maximum clustering.

The number of triangles around a node can easily be found from the adjacency matrix:

$$t_i = \frac{1}{2} \sum_{j \in V} \sum_{h \neq j \in V} A_{ij} A_{ih} A_{jh}. \quad (2.16)$$

If an edge is missing between any of the nodes  $v_i, v_j, v_h$ , then the adjacency matrix element for that edge will be zero, with the product of the elements also being zero. Thus, only connected triangles are counted:

$$A_{ij} A_{ih} A_{jh} = \begin{cases} 1 & \text{if triangle } i, j, h \\ 0 & \text{if edge missing} \end{cases}. \quad (2.17)$$

The  $\frac{1}{2}$  in Equation 2.16 accounts for the double counting of equivalent triangles. Using  $(i \leftrightarrow j)$  to denote an undirected edge between  $v_i, v_j$ , it can be seen that the same triangle can be traced via two separate paths from the same source node  $v_i$ :

$$(i \leftrightarrow j \leftrightarrow h \leftrightarrow i) = (i \leftrightarrow h \leftrightarrow j \leftrightarrow i). \quad (2.18)$$

The global clustering properties of an entire network can be determined in one of two ways. The first is to take the mean clustering coefficient ( $C_{avg}$ ) of all the individual nodes in a network (Watts and Strogatz, 1998):

$$C_{avg}(G) = \frac{1}{N} \sum_{i \in V} C_i = \frac{1}{N} \sum_{i \in V} \frac{2t_i}{k_i(k_i - 1)}; \quad (2.19)$$

however, this approach can be disproportionately influenced by low-degree nodes (Rubinov and Sporns, 2010). The alternative is to normalise the number of triangles collectively over the whole network. This second approach is sometimes referred to as *transitivity* (Newman, 2003b; Wasserman and Faust, 1994):

$$C(G) = \frac{\sum_{i \in V} 2t_i}{\sum_{i \in V} k_i(k_i - 1)}. \quad (2.20)$$

Due to the high number of low-degree neurons in the nervous system of *C. elegans*, the transitivity method in Equation 2.20 is used to calculate the global clustering coefficients presented in this thesis.

Slight modifications allow for the clustering coefficient to also be computed for directed networks (Fagiolo, 2007; Rubinov and Sporns, 2010), by considering the in and out connections separately and accounting for reciprocity. The directed version of Equation 2.20 is given by:

$$C^{\rightarrow}(G^{\rightarrow}) = \frac{\sum_{i \in V} t_i^{\rightarrow}}{\sum_{i \in V} [(k_i^{out} + k_i^{in})(k_i^{out} + k_i^{in} - 1) - 2 \sum_{j \in V} A_{ij} A_{ji}]}; \quad (2.21)$$

and likewise for counting triangles in the directed network:

$$t_i^{\rightarrow} = \frac{1}{2} \sum_{j \in V} \sum_{h \neq j \in V} (A_{ij} + A_{ji})(A_{ih} + A_{hi})(A_{jh} + A_{hj}). \quad (2.22)$$

## 2.4.6 Modularity

Modularity describes the extent to which a network can be divided into *modules* (or *communities*): groups of nodes with more within-group connections than between-group connections (see Figure 2.6). Modules can be thought of as distinct highly-clustered structures that allow for specialisation and functional segregation to emerge in a network. The examination of a network's modular composition is therefore a useful

approach to understanding the internal organisation of a network.

To compute the modularity measure  $Q$  for a network, it is first necessary to subdivide the network into non-overlapping modules  $c$  that maximise within-module connectivity and minimise between-module connectivity (see § 2.4.6.1 below). The modularity is then given as the fraction of edges that connect nodes within the same module (intramodular edges), subtracting the fraction expected if the network were wired at random (Clauset et al., 2004; Newman, 2004, 2006; Newman and Girvan, 2004):

$$Q = (\text{fraction of intramodular edges}) - (\text{expected fraction of such edges}). \quad (2.23)$$

The number of intramodular edges can be found using the formula (Newman, 2010):

$$\frac{1}{2} \sum_{i \in V} \sum_{j \in V} A_{ij} \delta(c_i, c_j), \quad (2.24)$$

where  $c_i, c_j$  are the modules respectively containing nodes  $v_i, v_j$ ; and  $\delta$  is the Kronecker delta function:

$$\delta(i, j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}. \quad (2.25)$$

As  $\delta(c_i, c_j) = 1 \iff c_i = c_j$ , multiplying by this ensures that we only count edges between nodes that are in the same module, while the  $\frac{1}{2}$  in Equation 2.24 accounts for double counting of edges.

As there are  $M$  edges in a network, with each edge having two ends (thus contributing to the degree values of two separate nodes), the total number of “edge ends” in a network is  $2M$ . Of those ends, exactly  $k_i$  of them belong to node  $v_i$ , and  $k_j$  of them belong to node  $v_j$ ; thus the expression  $\frac{k_i k_j}{2M}$  gives the expected probability of two nodes mutually connecting at random. We can therefore find the expected number of edges via:

$$\frac{1}{2} \sum_{i \in V} \sum_{j \in V} \frac{k_i k_j}{2M} \delta(c_i, c_j). \quad (2.26)$$

As before, the  $\frac{1}{2}$  accounts for double counting, and  $\delta$  ensures we only count those edges within the same module.

Subtracting Equation 2.26 from Equation 2.24, and dividing by the total number of edges  $M$  in the network, we can then compute the modularity score (Clauset et al., 2004; Newman, 2004, 2006):

$$Q(G, c) = \frac{1}{2M} \sum_{i \in V} \sum_{j \in V} \left( A_{ij} - \frac{k_i k_j}{2M} \right) \delta(c_i, c_j), \quad (2.27)$$

with  $Q \in [-1, 1]$ . Positive values of  $Q$  indicate that the network is modular, with more intramodular edges than expected, and negative values indicate that the network is *anti-modular*, with nodes preferentially connecting to nodes in different modules. A modularity value of  $Q \gtrsim 0.3$  has been proposed as a good indicator of a network having significant modular composition (Clauset et al., 2004).

To account for the directionality of the connections in a directed network, the modularity measure can be modified to consider the out-degrees and in-degrees, with the denominator also changing from  $2M$  to  $M$  to allow for unidirectional edges (Leicht and Newman, 2008):

$$Q^{\rightarrow}(G^{\rightarrow}, c) = \frac{1}{M} \sum_{i \in V} \sum_{j \in V} \left( A_{ij} - \frac{k_i^{\text{out}} k_j^{\text{in}}}{M} \right) \delta(c_i, c_j). \quad (2.28)$$

#### 2.4.6.1 Modular partitioning

As alluded to above, an important step in calculating the modularity is the division of the network into non-overlapping modules  $c$ . The objective is to find a modular partitioning that maximises intramodular connectivity and minimises intermodular connectivity. The modularity score associated with this partitioning is referred to as the *maximised modularity* (see Figure 2.13).

In practice, finding the optimal modular partitioning for a network is non-trivial, and has been shown to be an NP-complete problem (Brandes et al., 2008; Fortunato, 2010); however, heuristic algorithms can be used to find acceptable solutions within a computationally tractable runtime (often locally maximised on a subset of the search space), though these are not guaranteed to find a globally optimal partitioning (Fortunato, 2010).

A computationally efficient, and conceptually simple, method for partitioning a network into modules is given by Newman (2004), listed in Algorithm 2 (**MM**). This

---

**Algorithm 2: MM – Maximised modularity.** Greedy algorithm that iteratively agglomerates nodes into modules to find a partitioning that maximise the modularity score  $Q$ . First proposed by Newman (2004).

---

**input** : Graph  $G = (V, E)$ , where  $V = \{v_1, \dots, v_N\}$  is the set of vertices, and  $E = \{e_1, \dots, e_M\}$  is the set of connecting edges.

**output**: Maximised modularity score  $Q_{\max}$ ; and corresponding set of communities (modules)  $c_{\max}$ .

```

[1] begin
[2]   /* Start with each vertex in its own community */
[3]    $c \leftarrow$  partition of  $N$  singletons;
[4]   /* Get the initial modularity score */
[5]    $Q_{\text{temp}} \leftarrow Q(G, c)$ ;
[6]   /* Begin agglomerating communities and continue until all
[7]     vertices are merged into a single community */
[7]   while  $|c| > 1$  do
[8]     /* Initialise temporary variable used to record  $\Delta Q$  */
[9]      $dq \leftarrow \emptyset$ ;
[10]    /* Compute the change in modularity  $\Delta Q$  from merging
[11]      each pair of communities  $c_i, c_j$ . For efficiency, do
[12]      not consider unconnected communities. */
[11]    foreach connected pair  $c_i, c_j$  do
[12]      /* Temporarily merge pair  $c_i, c_j$  and get the new
[13]        partitioning */
[13]       $c' \leftarrow \text{Merge}(c, i, j)$ ;
[14]      /* Compute change in modularity from merging  $c_i, c_j$  */
[15]       $dq_{ij} \leftarrow \Delta Q(G, c')$ ;
[16]    /* Find the community pair that increased the modularity
[17]      score the most (or decreased it the least) when
[18]      merged, and permanently update the partitioning */
[17]    select  $i, j$  where  $\max_{ij}(dq_{ij})$ ;
[18]     $c \leftarrow \text{Merge}(c, i, j)$ ;
[19]     $Q_{\text{temp}} \leftarrow Q_{\text{temp}} + dq_{ij}$ ;
[20]    /* Record result if the current partitioning is an
[21]      improvement */
[21]    if  $Q_{\text{temp}} > Q_{\max}$  then
[22]       $Q_{\max} \leftarrow Q_{\text{temp}}$ ;
[23]       $c_{\max} \leftarrow c$ ;
[24]  return  $Q_{\max}, c_{\max}$ ;

```

---

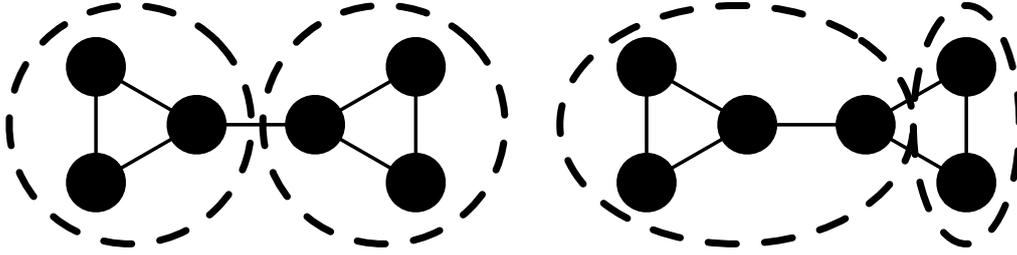


Fig. 2.13 Illustration of network modularity ( $Q$ ) showing: network partitioning giving maximised modularity (left), and a sub-optimal partitioning (right). Figure adapted from Clauset (2016a).

approach starts with each node in its own module, and at each step combines (merges) the two modules that result in the greatest improvement in modularity from being combined ( $\max \Delta Q$ ), until a locally optimal partitioning is found (i.e. a set of modules with a maximised modularity value). Using this approach, reasonable modular partitioning of a network can be achieved (Newman, 2004).

Research into the properties and performance characteristics of network partitioning algorithms is still an ongoing area of research. For discussions and comparisons of other network partitioning algorithms see Lancichinetti and Fortunato (2009); Orman and Labatut (2009); Yang et al. (2016).

### 2.4.7 Reciprocity

As mentioned previously (§ 2.2.2.2) reciprocity measures the fraction of reciprocal connections in a network, providing a useful summary measure of local communication and feedback between nodes.

A reciprocal connection is a connection from a source node to a target node  $(v_i, v_j) \in E$  that has a corresponding return connection from the target node  $(v_j, v_i) \in E$ . Reciprocity  $r^{\leftrightarrow}$  measures the fraction of these reciprocal connections relative to the total number of connections (Newman, 2010). As all of the connections are reciprocal in an undirected network, reciprocity is only meaningful for directed networks.

Where  $|E| = M$  is the total number of edges in a network, and  $|E^{\leftrightarrow}|$  is the number of reciprocal edges, the reciprocity can be written:

$$r^{\rightarrow}(G^{\rightarrow}) = \frac{|E^{\leftrightarrow}|}{|E|}. \quad (2.29)$$

In a binary network the product  $A_{ij}A_{ji} = 1$  if and only if there is a reciprocal edge between  $v_i, v_j$ ; as such, the number of reciprocal connections  $|E^{\leftrightarrow}|$  can easily be computed by summing  $A_{ij}A_{ji}$  over all pairs of adjacency matrix elements:

$$|E^{\leftrightarrow}| = \sum_{i \in V} \sum_{j \in V} A_{ij}A_{ji}, \quad (2.30)$$

while the total number of connections is given by:

$$|E| = \sum_{i \in V} \sum_{j \in V} A_{ij} = M. \quad (2.31)$$

### 2.4.8 Assortativity

The assortativity coefficient measures the extent to which the nodes in a network preferentially attach to other nodes with similar degree values. This can provide useful insights into the large-scale structural organisation of a network (see § 2.2.1 and Figure 2.5, page 21). The assortativity  $R$  can be obtained from the degree-degree correlation, typically measured using the Pearson correlation coefficient (Newman, 2002, 2003a; Pastor-Satorras et al., 2001): In a network with an assortative (*homophilic*) structure, the nodal degrees of a given node's neighbourhood will be highly correlated with its own degree, while in a network with a disassortative (*heterophilic*) structure, they will be anti-correlated.

For an undirected network, the assortativity can be measured using the Pearson correlation coefficient via:

$$R = \left( \frac{M^{-1} \sum_{(i,j) \in E} k_i k_j - \left[ M^{-1} \sum_{(i,j) \in E} \frac{1}{2}(k_i + k_j) \right]^2}{M^{-1} \sum_{(i,j) \in E} (k_i^2 + k_j^2) - \left[ M^{-1} \sum_{(i,j) \in E} \frac{1}{2}(k_i + k_j) \right]^2} \right), \quad (2.32)$$

with values in the range  $R \in [-1, +1]$ . To illustrate, a positive coefficient would indicate that the high-degree nodes preferentially link to other high-degree nodes, (and similarly low-degree to low-degree); while a negative coefficient would indicate a disassortative structure, with low-degree nodes preferentially linking to high-degree nodes and vice versa (as typified by a star network). A correlation coefficient of  $R \approx 0$

indicates a non-assortative structure, with no preferential attachment in either direction.

To measure the correlation between the out-degrees and in-degrees of nodes in a directed network, Equation 2.32 can be modified accordingly:

$$R^{\rightarrow} = \left( \frac{M^{-1} \sum_{(i,j) \in E}^M k_i^{\text{out}} k_j^{\text{in}} - \left[ M^{-1} \sum_{(i,j) \in E}^M \frac{1}{2} (k_i^{\text{out}} + k_j^{\text{in}}) \right]^2}{M^{-1} \sum_{(i,j) \in E}^M \left[ (k_i^{\text{out}})^2 + (k_j^{\text{in}})^2 \right] - \left[ M^{-1} \sum_{(i,j) \in E}^M \frac{1}{2} (k_i^{\text{out}} + k_j^{\text{in}}) \right]^2} \right). \quad (2.33)$$

### 2.4.9 Robustness

In § 2.2.4 (page 29) we discussed some of the main approaches used to estimate the resilience of a network from its structure. The approach used to analyse the networks in this thesis is to subject a network to progressive degradation by removing nodes at random (i.e. random node attack), then measuring the effect of the attack by observing the size of the largest connected component. The results of this can be captured and summarised in a single value, known as the robustness  $\rho$  (Schneider et al., 2011a,b):

$$\rho(G) = \frac{1}{N} \sum_{n=1}^N \gamma(n), \quad (2.34)$$

where  $\gamma(n)$  returns the fraction of nodes in the largest connected component after removing  $n$  nodes, relative to the original number of nodes  $N$ . This is illustrated in Figure 2.14.

The normalisation by  $1/N$  in Equation 2.34 allows for networks of different sizes to be compared, with the range of possible values always being  $\rho \in [1/N, 0.5]$ ,  $1/N$  corresponding to a star network (with high vulnerability), and  $0.5$  to a fully connected graph (with high resilience).

### 2.4.10 Null model networks

In the analysis of complex networks, it is common practice to use samples of random graphs to provide a null model baseline against which to evaluate observations for a network of interest (see § 2.2.3, page 27).

Many types of random network models exist (Betzel and Bassett, 2017; Hosseini and Kesler, 2013; Klimm et al., 2014). The two main classes used in this thesis are

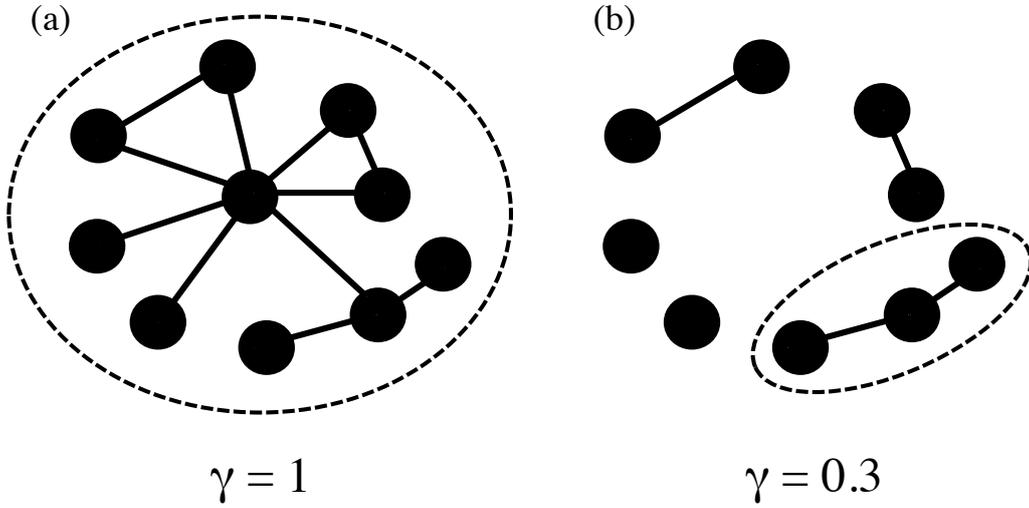


Fig. 2.14 Robustness. (a) A network with a single connected component of 10 nodes. As all of the nodes are in the largest connected component,  $\gamma = 1$ . (b) After removing the central hub, the network fragments and the size of the largest connected component (indicated by the dashed circle) is reduced to 3 nodes, thus the fraction of connected nodes is  $\gamma = 0.3$ .

(1) random Erdős-Rényi networks, and (2) degree-matched random networks. These will be described in the following sections.

For both classes of random networks, the convention used in this thesis is to represent a null model ensemble as a set of graphs:

$$\Lambda = \{G_1, \dots, G_{|\Lambda|}\}, \quad (2.35)$$

where each  $G_i \in \Lambda$  is a unique random network, generated according to one of the models described in the following sections. The exact model being used will be specified in each case.

#### 2.4.10.1 Erdős-Rényi random network model

The classical, and simplest, form of null model is the Erdős-Rényi (ER) model. ER networks are created with a fixed number of nodes  $N$ , with the edges connecting the nodes placed at random while accounting for the desired construction (e.g. directed or undirected edges).

Two variants of the ER model exist. The first, introduced by Solomonoff and Rapoport (1951), assigns edges according to a specified inter-node connection probability, where each pair of nodes has an independent probability  $p$  of being connected, with the corresponding probability of them being unconnected being  $1 - p$  (Gilbert, 1959; Solomonoff, 1952; Solomonoff and Rapoport, 1951). This model is referred to as the  $G_{ER}(N, p)$  model. The second ER model, studied by Erdős and Rényi (1959, 1960, 1961), instead generates random graphs with a given number of edges  $M$ , from a uniform distribution. This is referred to as the  $G_{ER}(N, M)$  model. It is this  $G_{ER}(N, M)$  model that is used in later chapters to normalise the small-world index (see § 2.4.11 below).

To generate random graphs, the  $G_{ER}(N, p)$  model can be implemented as a series of Bernoulli trials for each of the  $\frac{1}{2}N(N - 1)$  unique node pairs. An edge is created between a pair of nodes if the connection probability  $p$  is greater than or equal to a randomly generated number  $\varpi \in (0, 1]$ , as shown in Algorithm 3, **ER-Np** (page 3). The  $G_{ER}(N, M)$  model can instead be implemented by iteratively placing edges between pairs of randomly selected nodes until the desired number of edges  $M$  is reached; Algorithm 4, **ER-NM** (page 4). Although the two variants of the ER model are distinct, they both generate randomly-connected graphs with the same general topological properties.

Using the  $G_{ER}(N, p)$  model, the probability that a graph with  $M$  edges is created – comparable to a graph created via  $G_{ER}(N, M)$  – is given by the binomial distribution (Barabási and Pósfai, 2016; Newman, 2010):

$$P(M) = \binom{\frac{1}{2}N(N - 1)}{M} p^M (1 - p)^{\frac{1}{2}N(N - 1) - M}. \quad (2.36)$$

Decomposing this,  $\binom{\frac{1}{2}N(N - 1)}{M}$  is the number of unique ways  $M$  edges can be placed between the  $\frac{1}{2}N(N - 1)$  node pairs;  $p^M$  is the probability of  $M$  edges existing in the network; and  $(1 - p)^{\frac{1}{2}N(N - 1) - M}$  is the probability of the remaining  $\frac{1}{2}N(N - 1) - M$  potential edges not existing.

The average number of edges in networks generated using this model is:

$$\langle M \rangle = p \frac{N(N - 1)}{2}. \quad (2.37)$$

Using Equation 2.37, the average nodal degree of  $G_{ER}(N, p)$  networks is straightforward to compute. As each edge contributes to the degree value of two separate nodes, this is given by:

$$\langle k \rangle = \frac{2\langle M \rangle}{N} = p(N - 1). \quad (2.38)$$

We can similarly obtain the degree distribution for  $G_{ER}(N, p)$  networks. The probability of a single node connecting to  $k$  nodes, and not to the other  $N - 1 - k$  nodes, is  $p^k(1 - p)^{N-1-k}$ ; and the number of ways  $k$  edges can be placed amongst  $N - 1$  nodes (i.e. all nodes excluding itself) is  $\binom{N-1}{k}$ . Thus, the degree distribution is given by the binomial distribution:

$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}, \quad (2.39)$$

which can be closely approximated by a Poisson distribution when  $N \gg \langle k \rangle$  (Barabási and Pósfai, 2016; Newman, 2010):

$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}. \quad (2.40)$$

For this reason, ER networks are also sometimes referred to in the literature as Poisson random graphs.

As  $M = p \frac{N(N-1)}{2}$  when  $N \rightarrow \infty$ , both the  $G_{ER}(N, p)$  and  $G_{ER}(N, M)$  models become statistically equivalent for large values of  $N$  (Barrat et al., 2013; de Silva and Stumpf, 2005). Both variants of the ER model thus generate random networks with binomial degree distributions.

As previously mentioned, ER networks are a commonly used null model as they are easy to implement and have properties that can be interrogated analytically (Hofstad, 2017); however, as both variants of the ER model generate networks with binomial degree distributions, they cannot recapitulate all of the properties observed in complex real-world networks, which often exhibit power-law degree distributions (see Figure 2.4, page 19). In instances where it is necessary to account for such properties, other null models can be used that generate random networks to a prescribed degree distribution. Two such models are described in the following section.

---

**Algorithm 3: ER-Np – Erdős-Rényi  $G(N, p)$  model.** Generates random networks with  $N$  vertices and intervertex connection probability  $p$ . Based on the model of random networks first introduced by Solomonoff and Rapoport (1951), and independently developed by Gilbert (1959).

---

**input** : Number of nodes  $N$ , and connection probability  
 $p \in \{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$ .

**output**: Random Erdős-Rényi network  $G_{ER} = (V, E)$  where  $|V| = N$ .

```

[1] begin
[2]   /* INITIALISATION */
[3]   /* Initialise node set  $V$  with  $N$  elements */
[4]    $V \leftarrow \{v_1, \dots, v_N\}$ ;
[5]   /* MAIN PROCEDURE */
[6]   /* Iterate over set of nodes */
[7]   for  $\forall v_i \in V$  do
[8]     /* Unique edges are ensured by preventing node  $v_i$  from
[9]        connecting to a node with a lower index ( $j > i$ ). This
[10]        constraint can be removed for directed networks */
[11]     for  $\forall v_j \in V$  where  $j > i$  do
[12]       /* Draw a random number  $\varpi \in (0, 1]$ , and add a
[13]          connection between the current nodes  $v_i, v_j$  if the
[14]          random number is greater than or equal to the
[15]          connection probability  $p$  */
[16]       select  $\varpi$  uniformly at random from  $\{x \in \mathbb{R} \mid 0 < x \leq 1\}$ ;
[17]       if  $p \geq \varpi$  then
[18]          $E \leftarrow E \cup \{(v_i, v_j)\}$ ;
[19]     return  $G_{ER} = (V, E)$ ;

```

---

---

**Algorithm 4: ER-NM – Erdős-Rényi  $G(N, M)$  model.** Generates random networks with  $N$  vertices and  $M$  edges. Based on the model of random networks introduced by Erdős and Rényi (1959).

---

**input** : Number of nodes  $N$ ; and number of edges  $M$  to include, where  $M \leq \frac{1}{2}N(N - 1)$  for undirected networks, or  $M \leq N(N - 1)$  for directed networks.

**output**: Random Erdős-Rényi network  $G_{ER} = (V, E)$ , where  $|V| = N$  and  $|E| = M$ .

```

[1] begin
[2]   /* INITIALISATION                               */
[3]   /* Initialise vertex set  $V$  with  $N$  elements       */
[4]    $V \leftarrow \{v_1, \dots, v_N\}$ ;
[5]   /* Initialise empty edge set  $E$  to populate with edges */
[6]    $E \leftarrow \emptyset$ ;
[7]   /* MAIN PROCEDURE                               */
[8]   /* Add edges to the graph until the specific number  $M$  has
      been reached                                   */
[9]   while  $|E| < M$  do
[10]  | /* Draw two unconnected unique nodes at random      */
[11]  | select  $v_i, v_j \in V$  uniformly at random
[12]  | where  $((v_i, v_j) \notin E) \wedge (v_i \neq v_j)$ ;
[13]  | /* Create an edge to connect the selected nodes    */
[14]  |  $E \leftarrow E \cup \{(v_i, v_j)\}$ ;
[15]  return  $G_{ER} = (V, E)$ ;

```

---

### 2.4.10.2 Degree-matched null models

The degree distribution plays an important role in determining the large-scale structure of a network, and can consequently influence many topological measures. It is therefore often desirable to control for this by creating null model networks with the same degree distributions as the network being studied (see § 2.2.3).

The two most common methods for creating null model networks that preserve the degree distribution, number of nodes, and number of edges – and thus also the original network’s size and density – are:

1. the *edge-swap* rewiring model (Algorithm 5, **ES**); and
2. the *configuration model* (Algorithm 6, **CM**).

The edge-swap model generates random networks by creating duplicate versions of the original network, which are then rewired via an iterative Markov chain Monte Carlo (MCMC) process that swaps the source or destination nodes between randomly selected pairs of edges, preserving the original nodal degrees (Maslov and Sneppen, 2002; Milo et al., 2003). For example, a pair of edges  $\{(u, v), (x, y)\}$  can be rewired to give  $\{(u, x), (v, y)\}$ . This rewiring process is illustrated in Figure 2.15.

As sufficient edge randomisation will destroy any functionally-relevant structures present in the original network – while preserving the degree distribution, size, and density – the edge-swap method provides a good choice of null model in most cases. Analyses in later chapters use edge-swapped networks that have been randomised with a total  $10 \times M$  edge-swaps, which has previously been shown to be sufficient for full network randomisation (Milo et al., 2003; Ray et al., 2012).

The configuration model (Molloy and Reed, 1995) instead constructs synthetic random networks by first creating  $N$  nodes, each with  $k_i$  stubs, corresponding to the degree of node  $v_i$  in the original network (Figure 2.16a). These unconnected stubs are then linked at random to create a new network with the same degree sequence as the original (Figure 2.16b).

In practice, both of the degree-matched null models mentioned above produce similar random networks. A comparison of the two methods can be found in Appendix E (page 241), showing no significant difference between the null models for the measures and networks investigated in this thesis.

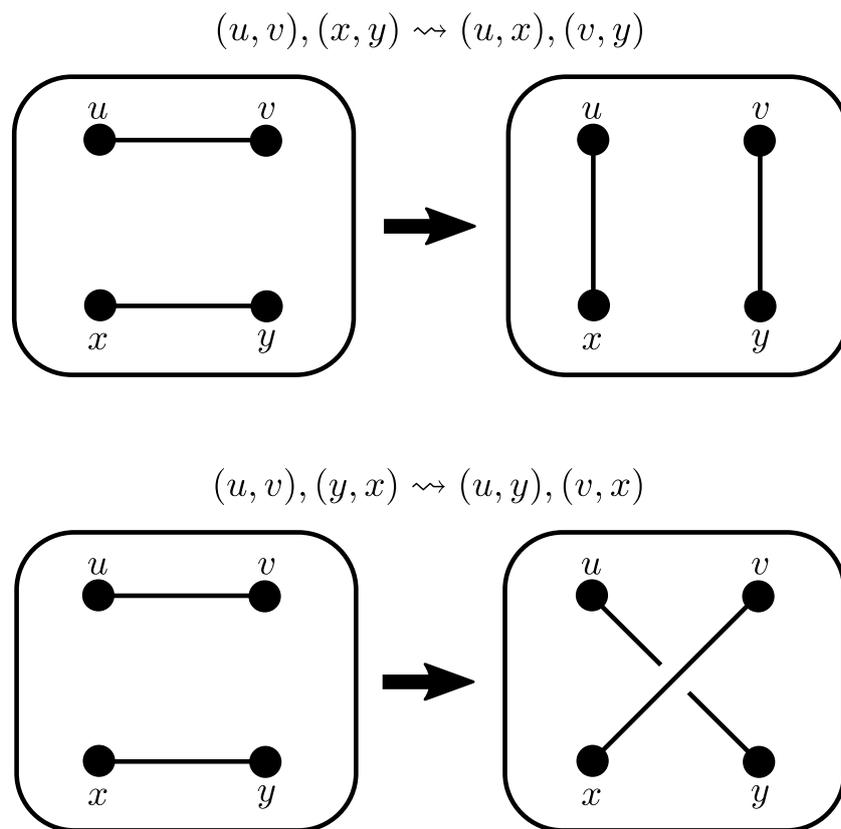


Fig. 2.15 Edge-swap randomisation. The structure of a graph is altered by randomly rewiring connections, preserving the degree distribution. The example shows the two possible outcomes for an undirected graph. Adapted from Fosdick et al. (2016).

---

**Algorithm 5: ES – Edge-swap.** Degree-preserving Markov chain Monte Carlo (MCMC) edge-swap rewiring / randomisation algorithm; used for generating null model networks. Based on the procedure described by Maslov and Sneppen (2002); Milo et al. (2003).

---

**input** : Graph  $G = (V, E)$ , where  $V = \{v_1, \dots, v_N\}$  is the set of vertices, and  $E = \{e_1, \dots, e_M\}$  is the set of connecting edges; and  $\xi$ , the randomisation factor (multiplied by the number of edges to give the total number of edge-swaps), where  $\xi \in \{x \in \mathbb{N} \mid 1 \leq x < \infty\}$ .

**output**: Graph with randomised edges  $G' = (V, E')$ .

```

[1] begin
[2]     /* INITIALISATION */
[3]     /* Create new edge set  $E'$  to rewire */
[4]      $E' \leftarrow E$ ;
[5]     /* Perform  $\xi$  edge-swaps  $\times$  the number of edges  $M$  */
[6]      $swaps \leftarrow \xi \times M$ ;
[7]     /* MAIN PROCEDURE */
[8]     while  $swaps \neq 0$  do
[9]         /* Choose two edges at random  $e_1, e_2 \in E'$  ensuring the
           edges: are unique, do not share source or
           destination vertices, and will not result in
           self-loops if swapped */
[10]        select  $e_1 = (v_A, v_B), e_2 = (v_C, v_D) \in E'$  at random
[11]        where  $(e_1 \neq e_2) \wedge (v_A \neq v_C) \wedge (v_A \neq v_D) \wedge (v_B \neq v_C) \wedge (v_B \neq v_D)$ ;
[12]        /* Note: if these constraints are not met, some
           variants of the algorithm count it as a swap trial
           and decrement the swaps counter accordingly; see
           Miklós and Podani (2004); Milo et al. (2003) */
[13]        /* Create new edges, swapping destination nodes */
[14]         $e'_1 \leftarrow (v_A, v_D); e'_2 \leftarrow (v_C, v_B)$ ;
[15]        /* Remove original edges and replace with the newly
           rewired ones */
[16]         $E' \leftarrow E' \setminus \{e_1, e_2\}$ ;
[17]         $E' \leftarrow E' \cup \{e'_1, e'_2\}$ ;
[18]        /* Decrement loop counter */
[19]         $swaps \leftarrow (swaps - 1)$ ;

[20]     /* Return a graph with the new set of rewired edges  $E'$  */
[21]     return  $G' = (V, E')$ ;

```

---

---

**Algorithm 6: CM – Configuration model.** Stub matching algorithm to generate undirected random graphs with a given degree sequence  $\vec{k}$ ; used for generating null model networks. The algorithm can easily be modified to generate directed graphs by connecting out-degree stubs to in-degree stubs. First proposed by Molloy and Reed (1995).

---

**input** : Number of nodes,  $N$ ; degree sequence,  $\vec{k} = (k_1, \dots, k_N)$ .  
**output** : Random graph  $G_{rand} = (V_{rand}, E_{rand})$  with the degree sequence  $\vec{k}$ , where  $|V_{rand}| = N$ .

```

[1] begin
[2]   /* INITIALISATION */
[3]    $V_{rand} \leftarrow \{v_1, \dots, v_N\}$ ;
[4]    $E_{rand} \leftarrow \emptyset$ ;
[5]   /* Get the number of stubs for each vertex from the degree
      sequence  $\vec{k}$ ; each entry  $k_i$  corresponding to a vertex
       $v_i \in V$  */
[6]    $stubs \leftarrow \vec{k}$ ;
[7]   /* MAIN PROCEDURE */
[8]   /* Continue adding edges while stubs remain */
[9]   while  $\sum_{i=1}^N stubs_i > 0$  do
[10]    /* Select two random nodes  $(v_i, v_j)$  with stubs remaining
      that are not already connected by an edge */
[11]    select  $v_i, v_j \in V_{rand}$  at random where
       $(i \neq j) \wedge (stubs_i > 0) \wedge (stubs_j > 0) \wedge [(v_i, v_j) \notin E_{rand}]$ ;
[12]    /* Create edge between the randomly selected vertices,
      and update the stub counts accordingly */
[13]     $E_{rand} \leftarrow E_{rand} \cup \{(v_i, v_j)\}$ ;
[14]     $stubs_i \leftarrow (stubs_i - 1)$ ;
[15]     $stubs_j \leftarrow (stubs_j - 1)$ ;
[16]  return  $G_{rand} = (V_{rand}, E_{rand})$ ;

```

---

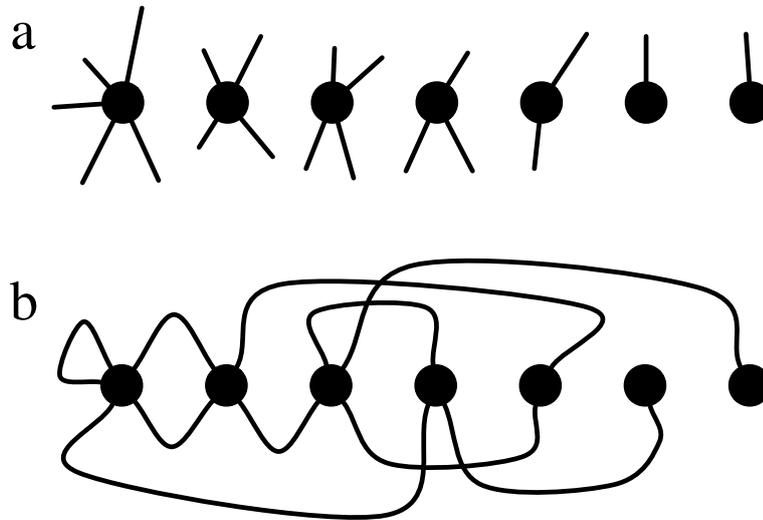


Fig. 2.16 Configuration model random network generation. Showing (a) unconnected stubs assigned to each node, determined from the desired degree sequence; and (b) null model graph generated by randomly connecting stubs to one another. Adapted from Clauset (2016b).

### 2.4.11 Small-world property

The networks of the nervous system must balance the conflicting requirements of functional segregation and functional integration. Networks that exhibit such a balance are described as being small-world, and are formally defined as networks that have a higher-than-random clustering coefficient  $C$  (indicative of functional segregation) and a lower-than-random characteristic path length  $L$  (indicative of functional integration and high wiring efficiency) (Rubinov and Sporns, 2010; Watts and Strogatz, 1998), see § 2.2.2.3 (page 24).

One concise way to measure the “small-worldness” of a network is through the normalised *small-world index*, first proposed by Walsh (1999) and independently developed and expanded by Humphries and Gurney (2008).

To determine whether a given network has a higher-than-random clustering coefficient and lower-than-random characteristic path length, the small-world index  $S$  uses a statistical ensemble of null model networks to normalise the measures. This ensemble of random networks  $\Lambda$  is usually generated using the ER method given in Algorithm 4 ( $\Lambda = \{G_1, \dots, G_{|\Lambda|}\}; \forall G_i \in \Lambda, G_i \leftarrow G_{ER}(N, M)$ ). The properties of these random net-

works can then be analysed and averaged to give representative values for normalisation.

Applying Equation 2.11 to the random networks in  $\Lambda$  can provide us with the mean characteristic path length of the ensemble:

$$\langle L_{rand} \rangle = \frac{1}{|\Lambda|} \sum_{i=1}^{|\Lambda|} L(G_i). \quad (2.41)$$

Similarly, via Equation 2.20 the mean clustering coefficient of the ensemble can also be determined:

$$\langle C_{rand} \rangle = \frac{1}{|\Lambda|} \sum_{i=1}^{|\Lambda|} C(G_i). \quad (2.42)$$

Where  $C$  and  $L$  are the respective clustering coefficient and characteristic path length for the real network of interest, and  $\langle C_{rand} \rangle$  and  $\langle L_{rand} \rangle$  are the mean random values computed from the null model networks in  $\Lambda$ , the normalised small-world index can be computed via:

$$S(G) = \frac{C/\langle C_{rand} \rangle}{L/\langle L_{rand} \rangle}. \quad (2.43)$$

As a greater-than-random clustering coefficient results in a numerator  $> 1$ , and a short-than-random characteristic path length results in a denominator  $< 1$ , the small-world index provides a balanced measure that accounts for both properties, and captures the degree of small-worldness in a single value. Using this measure, a network is defined as small-world if  $S > 1$ .

#### 2.4.11.1 Note on null model usage

Although the small-world index was originally defined to use the  $G_{ER}(N, M)$  null model (Algorithm 4, page 55), some applications in the literature have instead used edge-swapped networks (Harriger et al., 2012). It should be noted that the choice of null model can change the interpretation of the metric: In the case of ER, the small-world index considers whether a given network is small-world relative to networks with completely random connectivity, while for edge-swapped networks (and those generated by the configuration model), it considers whether a network is small-world relative to other networks with the same degree distribution.

As the degree distribution determines whether a network can contain high-degree nodes, the choice of null model can influence the extent to which the comparison networks contain hubs. This is important as hubs act to connect many nodes and

can greatly reduce the characteristic path length. The selection of null model, and resultant change in metric interpretation, also affects every other network measure that uses a random ensemble for normalisation or comparison.

Using the standard definition of the small-world index given by Humphries and Gurney (2008), the analyses presented in this thesis are normalised using ER networks. The reason for this is to determine whether the networks of interest have a small-world structure *at all*, and not whether they are *more* or *less* small-world than similar networks with the same degree distribution (as would be the case if edge-swapped networks were used). All other measures considered here, however, are compared to edge-swapped networks, as the other analyses attempt to address the question of whether the networks have structural features independent of the degree distribution. That is to say, the difference in null model selection is to address two different questions: (1) in the case of the small-world index, is the network small-world compared to random; and (2) what are the characteristics of the network that are not properties of the degree distribution?

### 2.4.12 Scale-free property

In § 2.2.1 and Figure 2.4 we briefly introduced the concept of scale-free networks. These are networks with scale-invariant degree distributions, typically characterised by the presence of many low-degree nodes connected to a small number of high-degree hubs. Scale-free networks have attracted much attention due to their interesting characteristics and abundance in nature, including in the brain.

Formally, a network is defined as being scale-free if its degree distribution  $P(k)$  follows the power-law:

$$P(k) \propto k^{-\alpha}, \quad (2.44)$$

where  $\alpha$  is the power-law exponent. Thus, by fitting this formula to an empirical degree distribution, via a method such as least-squares regression (Clauset et al., 2009), we can determine whether a network is scale-free. Often, a threshold is introduced such that  $P(k)$  is only considered above some minimum value  $k_{min}$  where the power-law holds. This is used to (1) enable a good fit to empirically-derived discrete data, where low values of  $k$  can deviate from the distribution; and (2) allow for the analysis of continuous data, since  $k^{-\alpha} \rightarrow \infty$  as  $k \rightarrow 0$  for  $\alpha > 0$ .

An interesting property of scale-free networks is that their structures are, to a large extent, determined by the power-law exponents  $\alpha$  of their distributions. Most real-world instances of scale-free networks have a power-law exponent in the range  $2 < \alpha < 3$  (Albert and Barabási, 2002; Barabási and Albert, 1999; Clauset et al., 2009), typically resulting in a small-world structure with a higher-than-random clustering coefficient (Albert and Barabási, 2002; Zhou, 2002), and an ultra-small characteristic path length that scales double-logarithmically with the number of nodes (Chung and Lu, 2002; Del Genio et al., 2011):

$$L \propto \log \log N. \quad (2.45)$$

This makes scale-free networks of this type highly efficient, providing a good balance between functional integration and functional segregation. Despite the efficient ultra-small nature of networks in the range  $2 < \alpha < 3$ , it is important to note that not all scale-free networks are small-world (Zhang et al., 2007).

As well as being highly efficient, scale-free networks have also been demonstrated to be highly resilient to random connection loss, although not to targeted attacks. In the case of random attacks, the efficient connectivity results in slow, gradual, network degradation, superior to randomly connected networks (Achard et al., 2006; Albert et al., 2000).

#### 2.4.12.1 Power-laws & statistics

As power-law distributions frequently span over orders of magnitude, the nodes in a scale-free network often have large differences in their nodal degrees. Furthermore, under certain circumstances common summary statistics are not informative. An example of this is when  $\alpha \leq 2$ , the mean degree  $\langle k \rangle$  diverges to infinity as  $N \rightarrow \infty$  (Newman, 2005). Although a finite scale-free network will not have an infinite number of nodes to allow this, the result shows that it is not a well-defined measure, and is unrepresentative of the actual distribution. The same applies to all subsequent moments, with  $\langle k^2 \rangle$  diverging for  $\alpha \leq 3$ , thus having no definite variance or standard deviation; and  $\langle k^3 \rangle$  for  $\alpha \leq 4$ , etc. Generally, where  $\mathcal{M}$  is the moment,  $\langle k^{\mathcal{M}} \rangle$  is only finite and well defined when  $\mathcal{M} < \alpha - 1$ . Where this is satisfied, the  $\mathcal{M}^{\text{th}}$  moment for

an ideal power-law distribution is given by (Newman, 2005):

$$\langle k^{\mathcal{M}} \rangle \propto \left( \frac{\alpha - 1}{\alpha - 1 - \mathcal{M}} \right) k_{min}^{\mathcal{M}}, \quad (2.46)$$

where  $k_{min}$  is the previously-mentioned lower threshold.

#### 2.4.12.2 Scale-free networks in the brain

Power-law degree distributions are found in various types of networks. Examples include metabolic networks (Jeong et al., 2000; Wagner and Fell, 2001), protein-protein interaction networks (Yook et al., 2004), information networks (Albert et al., 1999; Faloutsos et al., 1999), and social networks (Aparicio et al., 2015). Functional brain networks have also been found to exhibit power-law or truncated power-law distributions (Fornito et al., 2010; Hayasaka and Laurienti, 2010; Heuvel et al., 2008), as have connectomes derived from diffusion MRI methods (Gong et al., 2009; Iturria-Medina et al., 2008).

Although the power-law provides a good fit for some connectomic data, it is not the case for all brain regions (Humphries et al., 2006), and alternative distributions have been proposed that might provide a better fit to neural degree distributions (Gastner and Ódor, 2016). Similarly, although the synaptic connectome of *C. elegans* appears to exhibit a degree distribution tail that partially matches a power-law (Figure 3.9, page 98), other distributions have also been proposed (Amaral et al., 2000). Determining and interpreting the degree distributions of connectomes is still an open area of research (Bullmore and Sporns, 2009; Gastner and Ódor, 2016).

One theory for why scale-free networks are so common in nature relates to network growth characteristics. Imposing a simple growth rule, where new nodes have a higher probability of connecting to existing high-degree nodes, than to younger nodes, is sufficient to create a scale-free network with a power-law degree distribution (Barabási and Albert, 1999).

#### 2.4.13 Rich-club organisation

Rich-clubs are subnetworks composed exclusively of high-degree nodes, that serve as the main backbones in a network, linking modules and facilitating functional integration (Figure 2.6 on page 22 for an illustration) (Colizza et al., 2006; Zhou and Mondragon,

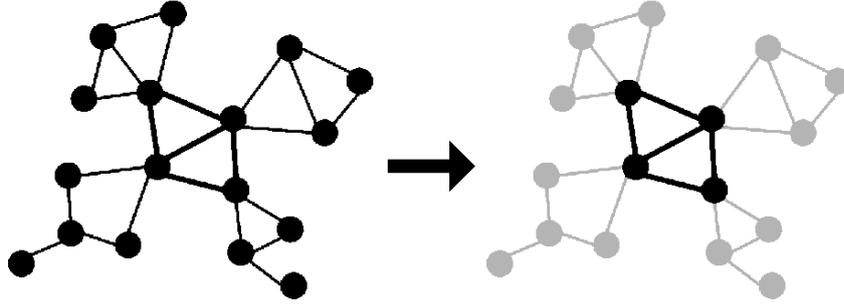


Fig. 2.17 Rich-club core. Rich-clubs can be identified by iteratively removing the lowest degree nodes, and comparing the density of the remaining core to the expected density. The figure on the right reveals a rich-club core after removing nodes  $k \leq 2$  (i.e.  $K = 2$ ).

2004). The name stems from the *rich-club phenomenon* observed in many natural systems, where connection-rich hubs are frequently more interconnected than would be expected by chance (Colizza et al., 2006).

The presence of rich-clubs can be identified by iteratively removing low-degree nodes for a given degree threshold  $K$ , such that nodes with  $k \leq K$  are removed, and comparing the density of the remaining high-degree subnetwork to the expected density from a random ensemble. The subnetwork density for a threshold is referred to as the rich-club coefficient  $\Phi(K)$ , and is the ratio of remaining connections  $M_{>K}$  to the maximum possible number of connections.

If a rich-club exists in a network, then for a certain degree range the hubs will appear more interconnected than expected, and thus have a higher relative density (i.e. normalised rich-club coefficient). Searching for rich-clubs is performed in a stepwise fashion, with the degree threshold  $K$  incrementing with each iteration ( $k \leq 1, \dots, N - 1$ ). An example of a subnetwork for  $K = 2$  is shown in Figure 2.17.

Let  $K$  be the degree threshold, and  $k_i$  the degree of node  $v_i \in V$ , then the subnetwork for a given threshold is:

$$G_{>K} = (V_{>K}, E_{>K}), \quad (2.47)$$

where the set of nodes  $V_{>K}$  and edges  $E_{>K}$  for the subnetwork are:

$$V_{>K} = \{v_i \in V \mid k_i > K\}, \quad (2.48)$$

$$E_{>K} = \{(v_i, v_j) \in E \mid (v_i \in V_{>K}) \wedge (v_j \in V_{>K})\}. \quad (2.49)$$

The number of nodes  $N_{>K}$  and edges  $E_{>K}$  remaining in the network after excluding nodes with a degree  $k \leq K$ , are respectively given by:

$$N_{>K} = |V_{>K}|, \quad (2.50)$$

$$M_{>K} = |E_{>K}|. \quad (2.51)$$

For an undirected network with no self-connections, the rich-club coefficient (i.e. subnetwork density) for a threshold  $K$  can then be calculated:

$$\Phi(K) = D(G_{>K}) = \frac{2M_{>K}}{N_{>K}(N_{>K} - 1)}. \quad (2.52)$$

Similarly, for a directed network:

$$\Phi^{\rightarrow}(K) = D^{\rightarrow}(G_{>K}^{\rightarrow}) = \frac{M_{>K}}{N_{>K}(N_{>K} - 1)}. \quad (2.53)$$

Thus, a fully-connected subnetwork at a given degree threshold  $K$  has a rich-club coefficient  $\Phi(K) = 1$ . To normalise the rich-club coefficient, it can be compared to the mean value from an ensemble of random networks  $\langle \Phi_{rand}(K) \rangle$ :

$$\langle \Phi_{rand}(K) \rangle = \frac{1}{|\Lambda|} \sum_{i=1}^{|\Lambda|} \Phi_{rand_i}(K), \quad (2.54)$$

Here  $|\Lambda|$  is the number of random networks in the ensemble, and  $\Phi_{rand_i}(K)$  is the rich-club coefficient for a random network  $G_i \in \Lambda$ . The normalised rich-club coefficient is then expressed as:

$$\Phi_{norm}(K) = \frac{\Phi(K)}{\langle \Phi_{rand}(K) \rangle}. \quad (2.55)$$

By convention, a significant rich-club is said to exist where the normalised rich-club coefficient is (Harriger et al., 2012; Reus and Heuvel, 2013; Towlson et al., 2013):

$$\Phi_{norm}(K) \geq 1 + 1\sigma, \quad (2.56)$$

with  $\sigma$  being the Standard Deviation of  $\Phi_{rand}(K)$ :

$$\sigma = \sqrt{\frac{1}{|\Lambda| - 1} \sum_{i=1}^{|\Lambda|} [\Phi_{rand_i}(K) - \langle \Phi_{rand}(K) \rangle]^2}. \quad (2.57)$$

### 2.4.14 Network motifs

In §2.2.2.4 we discussed how networks often contain repeating stereotypical connectivity patterns – termed *motifs* – that can form the functional building blocks of a network. Analysing the motifs from which a network is composed can provide deep insights into its structural architecture and computational properties (Alon, 2007; Kashtan et al., 2004; Milo, 2002; Sporns and Kötter, 2004).

Fundamentally, a motif is a small graph  $G' = (V', E')$ . To find the instances of a motif in a network of interest  $G$ , we can search for subgraphs  $G_{sub}$  within the network  $G$  that are *isomorphs* of a template motif  $G'$ ; in other words, we search for parts of the network that have an identical structure to the motif. Formally, a subgraph is a graph  $G_{sub} = (V_{sub}, E_{sub})$  contained within the original graph  $G = (V, E)$ , such that  $V_{sub} \subseteq V$  and  $E_{sub} \subseteq E \cap (V_{sub} \times V_{sub})$ .

Two graphs  $G'$  and  $G_{sub}$  are said to be isomorphic if and only if each node in  $G'$  has a one-to-one correspondence to a node in  $G_{sub}$ , such that each pair of nodes has exactly the same number of edges between them in both  $G'$  and  $G_{sub}$  (Gibbons, 1985; Schreiber and Schwöbbermeyer, 2008). See Figure 2.8 (page 26) for an illustration of motif matching.

The set of possible unique motifs of a given size  $s$  (where  $s$  is the number of vertices  $s = |V'|$ ) can be written:

$$\Omega_s = \{G'_1, \dots, G'_{|\Omega|}\}. \quad (2.58)$$

Using the principle of graph isomorphism introduced above, we can search for instances of each motif type  $G'_h \in \Omega_s$  in a network and record the frequency of each (denoted  $J_h$ ) to provide information on a network's composition (see Figure 2.8).

Although the problem of counting the motif frequency  $J_h$  can be stated simply, in practice it is computationally challenging. Recent progress has been made in the area of determining graph isomorphism – giving this task a sub-exponential ideal running time (Babai, 2015) – yet the process of exhaustively enumerating motifs still grows exponentially, both with the size of the target graph  $G$  and the size of the motif  $s$ . To illustrate: for directed motifs where all of the nodes are connected to at least one other node in the motif, there are only two possible motifs of size  $s = 2$ , as two nodes can only either be connected with a single edge or a recipro-

cal edge ( $|\Omega_2| = 2$ ). For  $s = 3$  this grows to  $|\Omega_3| = 13$  (displayed in Figure 2.9, page 26), while  $|\Omega_4| = 199$ ,  $|\Omega_5| = 9,364$ ,  $|\Omega_6| = 1,530,843$ , and  $|\Omega_7| = 880,471,142$  (Harary and Palmer, 1973; Sloane, 2017). As the search time also grows with the size of the target network being searched, most motif searchers are constrained to  $s \leq 4$ .

Several practical algorithms exist for applying the above concepts to real networks (Wong et al., 2012). The most common approach is to first divide the target graph  $G$  into all of the possible subgraphs  $G_{sub} \subseteq G$  of a given size  $s$  (called the *subgraph census*). A search can then be performed within the subgraph space for motif isomorphs (Wong et al., 2012). One computationally efficient, and widely used, method for this is the ESU algorithm (Wernicke, 2005) given in Algorithm 7 (page 69). This approach works by iterating through the nodes in the network and examining which subgraphs can be grown from each root node, stopping when the path length for each growth attempt reaches  $s$ . These are then collated and categorised by motif isomorphism to compute motif frequencies. The same process is then performed on null model networks for the purposes of normalisation and statistical analysis. For large networks and motif sizes, random sampling of the network can also be applied to estimate motif frequencies (Wong et al., 2012).

The degree to which a motif is over- or under- represented in a network can be determined by comparing the motif frequency observed in the network of interest, to the mean frequency observed in the null model samples, via the motif  $z$ -score (Milo, 2002). For a motif  $G'_h$ , the motif  $z$ -score is given by:

$$z_h = \frac{J_h - \langle J_{rand_h} \rangle}{\sigma_{J_{rand_h}}}, \quad (2.59)$$

where  $\langle J_{rand_h} \rangle$  is the mean occurrence of motif  $G'_h$  across the random networks, and  $\sigma_{J_{rand_h}}$  is the corresponding Standard Deviation.

---

**Algorithm 7: ESU – Enumerate subgraphs.** Subgraph enumeration algorithm from Wernicke (2005); used for motif detection in FANMOD (Wernicke and Rasche, 2006).

---

**input** : Graph  $G = (V, E)$ ; and  $s$ , the size of the subgraphs to enumerate containing exactly  $s$  vertices, where  $s \in \{x \in \mathbb{N} \mid 2 \leq x < N\}$ .

**output**: All size- $s$  subgraphs in  $G$ .

```

[1] begin
[2]   /* Enumerate subgraphs around each vertex  $v_i \in V$  */
[3]   for  $\forall v_i \in V$  do
[4]     /* Get the neighbourhood  $Nh$  of  $v_i$  (adjacent vertices),
[5]        ensuring the indices  $j$  are greater than the current
[6]        root vertex index  $i$  (to prevent double counting) */
[7]      $V_{extension} \leftarrow \{v_j \in Nh(\{v_i\}) \mid j > i\}$ ;
[8]     /* Enumerate the subgraphs around the current root  $v_i$ 
[9]        using the set of neighbouring vertices in  $V_{extension}$  */
[10]    call  $ExtendSubgraph(\{v_i\}, V_{extension}, v_i)$ ;

[11] function  $ExtendSubgraph(V_{sub}, V_{extension}, v_i)$ 
[12]   /* Stop extending the subgraph if it has reached the target
[13]      subgraph size  $s$  */
[14]   if  $|V_{sub}| = s$  then
[15]     /* Create subgraph edge set, and return result */
[16]      $E_{sub} \leftarrow \{(v_i, v_j) \in E \mid v_i, v_j \in V_{sub}\}$ ;
[17]     return  $G_{sub} = (V_{sub}, E_{sub})$ ;
[18]   /* Continue enumerating subgraphs while there are still
[19]      vertices available in the neighbourhood  $V_{extension}$  */
[20]   while  $V_{extension} \neq \emptyset$  do
[21]     /* Get arbitrary vertex  $v_\omega$  to add to the subgraph */
[22]     select  $\{v_\omega\} \subseteq V_{extension}$ ;
[23]      $V'_{sub} \leftarrow V_{sub} \cup \{v_\omega\}$ ;
[24]     /* Remove  $v_\omega$  from the pool of searchable vertices */
[25]      $V_{extension} \leftarrow V_{extension} \setminus \{v_\omega\}$ ;
[26]     /* Get the neighbourhood of the subgraph containing the
[27]        new vertex,  $V'_{sub}$  (i.e. all nodes adjacent to the
[28]        subgraph that are not in the subgraph) and enumerate
[29]        all possible subgraphs that can be grown from it by
[30]        recursively calling  $ExtendSubgraph()$  */
[31]      $V'_{extension} \leftarrow \{v_j \in Nh(V'_{sub}) \mid j > i\}$ ;
[32]     call  $ExtendSubgraph(V'_{sub}, V'_{extension}, v_i)$ ;

```

---

## 2.5 Multilayer networks

As we introduced in § 2.2.5, various interaction types between nodes can be represented as separate layers in a multiplex network. Multiplex networks consist of multiple layers containing the same nodes, but with different arrangements of edges in each layer, corresponding to instances of the interaction type associated with that layer. Such a multiplex network  $\mathcal{A}$  with  $X$  layers can conveniently be represented as a set of  $X$  separate  $N \times N$  adjacency matrices:

$$\mathcal{A} = \{A^{[1]}, A^{[2]}, \dots, A^{[X]}\}. \quad (2.60)$$

It is therefore possible to analyse the properties of each layer using the same network measures previously described, or via variants of the measures that consider all of the layers collectively (Battiston et al., 2014; Boccaletti et al., 2014; Kivelä et al., 2014).

### 2.5.1 Multilayer reducibility

When working with multiplex networks, it is useful to know the extent to which layers differ from one another, and whether each layer is truly a unique representation rather than a redundant (duplicate) description of another layer. One way to address these questions is to measure the information content of the layers and ask whether any information is lost when aggregating, or collapsing, layers; for example, to determine whether a two layer representation contains any more information than a single layer (*monoplex*) representation of the same network. If no information is lost in reducing the network to a single layer, we can say that the multilayer structure adds no information and that the network can be simplified by aggregating the two layers into a single layer (De Domenico et al., 2014, 2015a, 2016; Wang and Liu, 2017).

The approach described above is known as *multilayer reducibility*, and is closely related to edge overlap. Considering two multiplex layers with the same edges in each layer, the overlap between them is 100 %, and indeed no information would be lost aggregating the layers into a single layer with the same edges. The ability to measure the information content of a network therefore allows us to quantify the similarity, overlap, and redundancy of layers, while furthermore enabling us to find the most economical representation of a multilayer network by identifying and aggregating redundant layers.

### 2.5.1.1 Layer aggregation

The first step is to formalise the notion of layer aggregation. As all of the layers in a multiplex network contain the same nodes, the process of layer aggregation is simply performed by taking the edges from two or more layers, and merging them into a single layer (see Figure 2.18). If we are working with a multigraph (which allows for multiple edges between nodes), then the adjacency matrix for the aggregate network can be obtained by summing the adjacency matrices of the various layers (De Domenico et al., 2015a):

$$A_{\text{ag}} = \sum_{\beta=1}^X A^{[\beta]}. \quad (2.61)$$

If a multigraph is not desired – such as the networks analysed in later chapters – then the layer aggregation is performed by assigning exactly one edge between nodes if they are connected by an edge in any of the network layers, else leaving them unconnected (Battiston et al., 2014):

$$A_{\text{ag}_{ij}} = \begin{cases} 1 & \text{if } \exists \beta : A_{ij}^{[\beta]} \\ 0 & \text{otherwise} \end{cases}. \quad (2.62)$$

### 2.5.1.2 Graph entropy

The next requirement for determining multilayer reducibility is to quantify the information content, or complexity, of a layer. This is performed by measuring the entropy of the structure, using a variant of the information entropy measure first introduced by Shannon (1948). This can be viewed as a generalisation of the classical notion of thermodynamic entropy (Bais and Farmer, 2007; Shannon, 1948).

Information entropy essentially describes the predictability of states in the system, and thus the system's information content; this relationship can be seen from considering that simple systems are more predictable than complex ones, and thus encode less information. For a random variable  $Y$  with associated states  $\{y_1, \dots, y_n\}$ , the Shannon entropy can be computed from the probabilities of observing those states  $P$  (Shannon, 1948):

$$H_{\text{shannon}}(Y) = - \sum_{i=1}^n P(y_i) \log_2 P(y_i), \quad (2.63)$$

where by convention  $0 \log_2 0 \equiv 0$ .

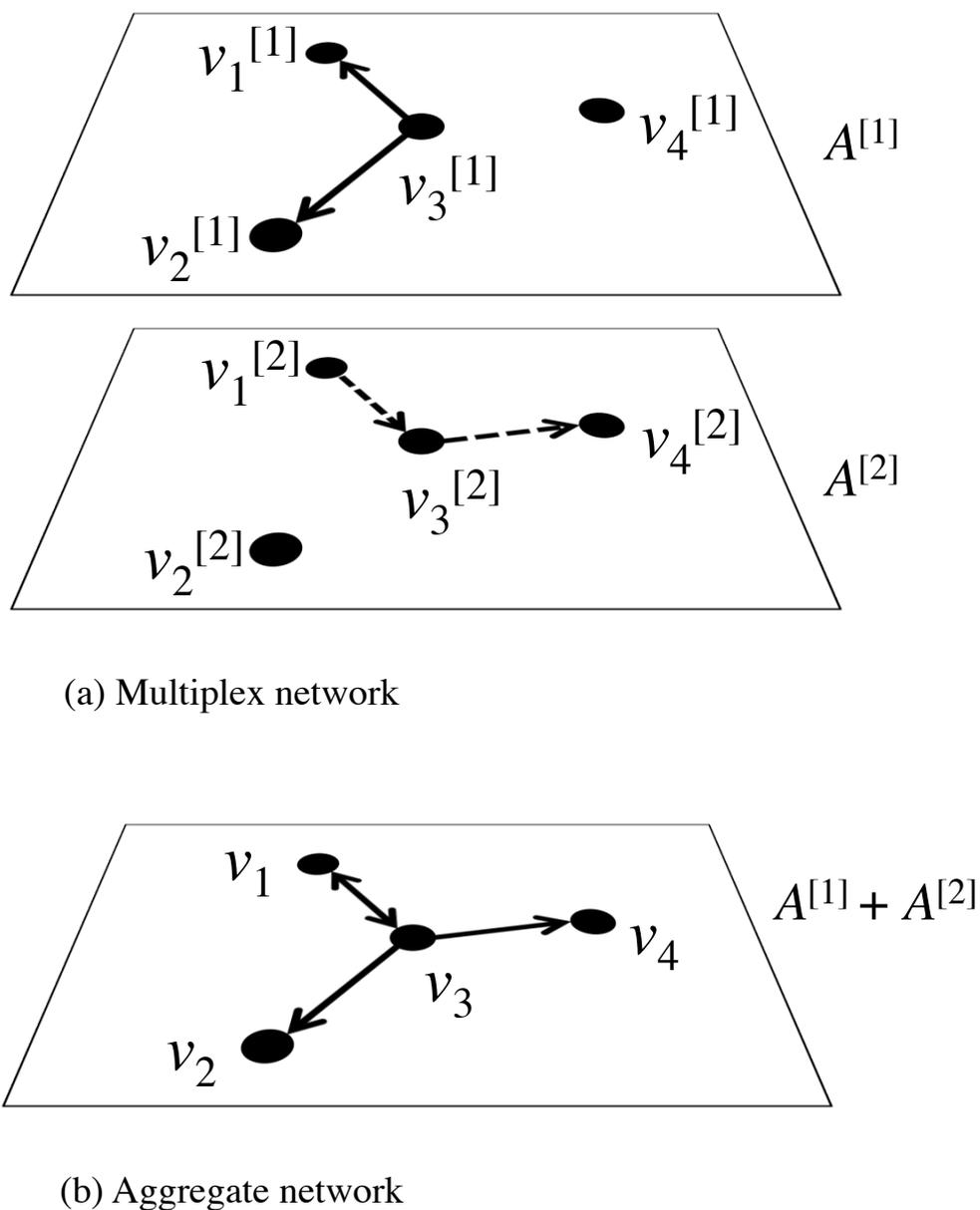


Fig. 2.18 Diagram showing: (a) a two-layer multiplex network containing four nodes, and (b) a one-layer, monoplex, network formed from aggregating the layers of the above multiplex network. Note that the edges from both layers are preserved in the aggregation.

Numerous methods exist for measuring the entropy of a graph structure (reviewed in Dehmer and Mowshowitz, 2011; Anand and Bianconi, 2009) with the most common being the *von Neumann entropy*, originally formulated in the context of quantum mechanics (QM) (Von Neumann, 1955).

In quantum mechanics, a system can either be in a *pure state*, describable by a single state vector, written  $|\psi\rangle$ ; or a *mixed state*, a statistical ensemble of pure states corresponding to a set of state vectors, each with an associated probability of the system being in that state  $p \in [0, 1]$ , satisfying the constraint  $\sum_i p_i = 1$ . This information is usually encoded in a *density matrix*,  $\varrho$ . Where  $\langle\psi_i|$  is the Hermitian transpose (conjugate transpose) of the state vector  $|\psi_i\rangle$ , the density matrix of a system can be expressed as (Han, 2012):

$$\varrho = \sum_i p_i |\psi_i\rangle\langle\psi_i|. \quad (2.64)$$

Taking the density matrix  $\varrho$ , the von Neumann entropy of a quantum system can be defined either in terms of its eigenvalues  $\lambda$  or the trace of the density matrix (Von Neumann, 1955):

$$\begin{aligned} H(\varrho) &= -\sum_i \lambda_i \log_2 \lambda_i \\ &= -\text{Tr}(\varrho \log_2 \varrho). \end{aligned} \quad (2.65)$$

The utility of this measure lies in its ability to express the “mixedness” of a quantum system:  $H = 0$  if and only if the system is in a pure state, else when  $H > 0$  the system is in a mixed state. For a classical (non-quantum) system, this is directly analogous to the Shannon entropy in Equation 2.63.

The applicability of the von Neumann entropy to the structural complexity of graphs has been the topic of several works (Anand et al., 2011; Braunstein et al., 2006; de Beaudrap et al., 2016; Han et al., 2011; Passerini and Severini, 2011). These have shown that the normalised Laplacian matrix of a graph (defined below) corresponds to the density matrix of a quantum system, such that a graph can be described in quantum mechanical terms as being in a pure state if and only if it contains exactly one edge ( $H = 0 \iff M = 1$ ), or as an ensemble of pure states if and only if it has multiple edges (i.e. a mixed state,  $H > 0 \iff M > 1$ ). The von Neumann entropy can thus be used to reliably measure the complexity and information content of a

graph's topological structure.

As stated, to compute the von Neumann entropy of a graph we must first define the normalised graph Laplacian matrix  $\hat{\mathcal{L}}$ . The Laplacian matrix is analogous to a discrete form of the Laplace operator ( $\mathcal{L} = -\nabla^2$ ), and can be used to study diffusion or random walk processes on a graph (Barrat et al., 2013; Newman, 2010). The properties of the Laplacian matrix – such as its relation to the QM density matrix – also make it useful for studying many topological network features not directly related to diffusion (Chung, 1997).

Starting with the concept of diffusion, consider an amount of some diffusible substance at node  $v_i$  of a graph, written  $\phi_i$ , with the diffusion constant  $I$ . The rate of change of  $\phi_i$  diffusing to or from neighbouring nodes in the network can then be written as:

$$\frac{d\phi_i}{dt} = I \sum_j A_{ij}(\phi_j - \phi_i). \quad (2.66)$$

In matrix form, this is equivalent to (Newman, 2010):

$$\frac{d\phi}{dt} = I(A - \mathcal{D})\phi, \quad (2.67)$$

where  $\mathcal{D}$  is a diagonal matrix containing the respective nodal degrees:

$$\mathcal{D} = \begin{pmatrix} k_1 & 0 & \dots & 0 \\ 0 & k_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & k_N \end{pmatrix}. \quad (2.68)$$

The graph Laplacian matrix is then defined as (Chung, 1997; Newman, 2010):

$$\mathcal{L} = \mathcal{D} - A, \quad (2.69)$$

and is constructed with the following elements:

$$\mathcal{L}_{ij} = \begin{cases} k_i & \text{if } i = j \\ -1 & \text{if } (v_i, v_j) \in E, \\ 0 & \text{otherwise} \end{cases} \quad (2.70)$$

allowing us to rewrite Equation 2.67 using the Laplacian matrix in the form:

$$\frac{d\phi}{dt} + I\mathcal{L}\phi = 0. \quad (2.71)$$

To obtain the normalised Laplacian, denoted  $\hat{\mathcal{L}}$ , corresponding to the density matrix  $\rho$  required for the von Neumann entropy, we divide the Laplacian matrix by the sum of the nodal degrees (Braunstein et al., 2006):

$$\hat{\mathcal{L}} = \mathcal{L} / \sum_{i \in V}^N k_i = \mathcal{L} / \text{Tr}(\mathcal{D}). \quad (2.72)$$

This normalised, rescaled, Laplacian has a number of interesting properties, including eigenvalues that satisfy the following:

$$\lambda_i \in [0, 1] \text{ for } \lambda_1, \dots, \lambda_N, \quad (2.73)$$

$$\sum_{i \in V}^N \lambda_i = 1. \quad (2.74)$$

Finally, with the normalised graph Laplacian  $\hat{\mathcal{L}}$ , and its eigenvalues  $\lambda$ , we can extend the definition of the von Neumann entropy (from Equation 2.65) to a single layer  $A^{[\beta]}$  of a multiplex graph  $\mathcal{A}$  (or more generally, to any monoplex adjacency matrix  $A$ ):

$$\begin{aligned} H(A^{[\beta]}) &= - \sum_i^N \lambda_i^{[\beta]} \log_2 \lambda_i^{[\beta]} \\ &= - \text{Tr} \left( \hat{\mathcal{L}}^{[\beta]} \log_2 \hat{\mathcal{L}}^{[\beta]} \right). \end{aligned} \quad (2.75)$$

To measure the total von Neumann entropy of a multiplex network, we can simply sum the entropy of the individual constituent layers, thus measuring the information content of the network:

$$\mathcal{H}(\mathcal{A}) = \sum_{\beta=1}^X H(A^{[\beta]}). \quad (2.76)$$

### 2.5.1.3 Layer reducibility

We have so far established how to aggregate layers in a multiplex network (§ 2.5.1.1), and how to measure the information content of those layers (§ 2.5.1.2). We can now examine methods to find the most economical network representation that maintains the maximum information content, while minimising the number of layers.

As stated in Equations 2.61 & 2.62, it is possible to combine layers in a multiplex network, by either collapsing all of the layers to form an aggregate monoplex network, or a subset of the layers to create a reduced multiplex representation (i.e. a partially aggregated multiplex network). If the original multiplex network  $\mathcal{A}$  has  $X$  layers, we can represent such a reduced multiplex network, with  $X' \leq X$  layers, as:

$$\mathcal{B} = \{B^{[1]}, B^{[2]}, \dots, B^{[X']}\}, \quad (2.77)$$

where  $B^{[\beta]}$  is either identical to a layer from the original multiplex  $A^{[\beta]} \in \mathcal{A}$ , or the aggregation of two or more such layers from  $\mathcal{A}$ . We can then determine the average entropy of the reduced representation via:

$$\bar{\mathcal{H}}(\mathcal{B}) = \frac{\mathcal{H}(\mathcal{B})}{X'} = \frac{\sum_{\beta}^{X'} H(B^{[\beta]})}{X'}. \quad (2.78)$$

Using the average entropy of a multiplex network representation, given by  $\bar{\mathcal{H}}(\mathcal{B})$ , it is possible to measure how distinguishable this representation is from the aggregated monoplex version, derived by collapsing all of the layers in the corresponding original network, denoted  $A_{\text{ag}}$ . The *distinguishability* of the reduced representation with respect to the aggregate network is defined as:

$$q(\mathcal{B}) = 1 - \frac{\bar{\mathcal{H}}(\mathcal{B})}{H(A_{\text{ag}})}. \quad (2.79)$$

The larger the value of  $q(\mathcal{B})$ , the more distinguishable the multilayer representation is from the single layer aggregate representation. If all of the layers in  $\mathcal{B}$  are identical, then the distinguishability is zero ( $q(\mathcal{B}) = 0$ ); while larger values of  $q$  indicate that the multilayer representation is more informative than the reduced one. As such, we can then search for a multiplex representation that maximises the distinguishability, and thus find an optimal reduced network:

$$\mathcal{B}_{\text{opt}} = \underset{\mathcal{B}}{\operatorname{argmax}}(q(\mathcal{B})). \quad (2.80)$$

In practice, enumerating all possible combinations to deterministically find the optimal network representation is often impractical, falling into the class of NP-hard problems (De Domenico et al., 2014, 2015a), where the search time grows at least exponentially with the number of layers. To address this, greedy agglomeration algorithms can be used to search for solutions in a heuristic manner (conceptually identical

to the maximisation method in Algorithm 2, **MM**) (De Domenico et al., 2014, 2015a). The improvement of layer-reduction algorithms is an active, ongoing, area of research (Wang and Liu, 2017).

After determining an optimal multilayer representation  $\mathcal{B}_{opt}$  with the highest distinguishability, we can quantify the degree of reducibility of the original multilayer network, written  $\chi(\mathcal{A})$ , by taking the ratio of the number of layer reductions ( $X - X'_{opt}$ ) to the maximum possible number of layer reductions. This quantity is termed the *reducibility*, and is fully expressed as:

$$\chi(\mathcal{A}) = \frac{X - X'_{opt}}{X - 1}. \quad (2.81)$$

As such,  $\chi(\mathcal{A}) = 0$  if there are no redundant layers that can be reduced, and  $\chi(\mathcal{A}) = 1$  if the network is fully reducible, such that it can be aggregated into a single monoplex layer without the loss of information.

#### 2.5.1.4 Layer distance

We have now examined how to find an optimal (non-redundant) representation of a multilayer network, and how to measure its distinguishability from its fully aggregated counterpart, using von Neumann entropy. It is also desirable, however, to be able to measure the distance (or dissimilarity) between arbitrary individual layers in a multiplex network. As we will see, the ability to quantify these distances can be used to better understand the relationship between layers, and furthermore aid heuristic optimisation approaches to reducibility (as applied to Equation 2.80), mentioned in the preceding section.

In quantum mechanics, the dissimilarity between mixed quantum states can be reliably measured using the *Jensen-Shannon divergence* (JSD) (Briët and Harremoës, 2009; Majtey et al., 2005). We can therefore use the mapping of graphs to quantum systems – via the normalised graph Laplacian corresponding to a QM density matrix (Equation 2.72) – to apply this approach directly to measuring the dissimilarity between individual network layers (De Domenico et al., 2014, 2015a).

To calculate the Jensen-Shannon divergence between two layers, we must first obtain the *relative entropy* of one layer with respect to the other, termed the Kullback-Leibler divergence (KLD) (De Domenico et al., 2014, 2015a; Kullback and Leibler, 1951). For

two multiplex layers ( $A^{[\alpha]}, A^{[\beta]} \in \mathcal{A}$ ) and their corresponding normalised Laplacian matrices ( $\hat{\mathcal{L}}^{[\alpha]}, \hat{\mathcal{L}}^{[\beta]}$ ), the Kullback-Leibler divergence is defined as:

$$D_{\text{KL}}(\hat{\mathcal{L}}^{[\alpha]} \parallel \hat{\mathcal{L}}^{[\beta]}) = \text{Tr} \left( \hat{\mathcal{L}}^{[\alpha]} [\log_2 \hat{\mathcal{L}}^{[\alpha]} - \log_2 \hat{\mathcal{L}}^{[\beta]}] \right). \quad (2.82)$$

Although the Kullback-Leibler divergence provides us with the layers' relative entropy, it does not formally qualify as a distance measure. For an arbitrary set  $\mathbb{X}$ , a function

$$d : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R} \quad (2.83)$$

is a *distance* measure over the set  $\mathbb{X}$ , only if the following conditions are satisfied  $\forall x, y \in \mathbb{X}$  (Lamberti et al., 2008):

$$\text{Positiveness: } d(x, y) > 0 \text{ for } x \neq y; \text{ and } d(x, x) = 0. \quad (2.84)$$

$$\text{Symmetry: } d(x, y) = d(y, x). \quad (2.85)$$

To qualify as a *metric* distance, the function must additionally satisfy the triangle inequality (Burago et al., 2001).  $\forall x, y, z \in \mathbb{X}$ :

$$\text{Triangle inequality: } d(x, z) \leq d(x, y) + d(y, z). \quad (2.86)$$

On the above criteria, the Kullback-Leibler divergence fails both the triangular inequality and symmetry conditions ( $D_{\text{KL}}(\hat{\mathcal{L}}^{[\alpha]} \parallel \hat{\mathcal{L}}^{[\beta]}) \neq D_{\text{KL}}(\hat{\mathcal{L}}^{[\beta]} \parallel \hat{\mathcal{L}}^{[\alpha]})$ ) (Lamberti et al., 2008); it can however be used to construct the aforementioned Jensen-Shannon divergence, which is symmetric, thus meeting all of the conditions of a distance (Lin, 1991). The Jensen-Shannon divergence is defined as:

$$D_{\text{JS}}(\hat{\mathcal{L}}^{[\alpha]} \parallel \hat{\mathcal{L}}^{[\beta]}) = \frac{1}{2} D_{\text{KL}} \left( \hat{\mathcal{L}}^{[\alpha]} \parallel \frac{\hat{\mathcal{L}}^{[\alpha]} + \hat{\mathcal{L}}^{[\beta]}}{2} \right) + \frac{1}{2} D_{\text{KL}} \left( \hat{\mathcal{L}}^{[\beta]} \parallel \frac{\hat{\mathcal{L}}^{[\alpha]} + \hat{\mathcal{L}}^{[\beta]}}{2} \right). \quad (2.87)$$

This distance measure can also be expressed in terms of von Neumann entropy, where it is clear that it measures the difference between the entropy of the combined normalised Laplacians, and the combined entropies of the individual normalised Laplacians:

$$D_{\text{JS}}(\hat{\mathcal{L}}^{[\alpha]} \parallel \hat{\mathcal{L}}^{[\beta]}) = H \left( \frac{\hat{\mathcal{L}}^{[\alpha]} + \hat{\mathcal{L}}^{[\beta]}}{2} \right) - \frac{1}{2} [H(\hat{\mathcal{L}}^{[\alpha]}) + H(\hat{\mathcal{L}}^{[\beta]})]. \quad (2.88)$$

In order to obtain a metric distance, we can take the square root of the Jensen-Shannon divergence, termed the Jensen-Shannon distance,  $d_{\text{JS}}$  :

$$d_{\text{JS}}(\hat{\mathcal{L}}^{[\alpha]}, \hat{\mathcal{L}}^{[\beta]}) \equiv \sqrt{D_{\text{JS}}(\hat{\mathcal{L}}^{[\alpha]} \parallel \hat{\mathcal{L}}^{[\beta]})}, \quad (2.89)$$

which returns values in the interval  $d_{\text{JS}} \in [0, 1]$  and has been shown to satisfy the triangle inequality (Endres and Schindelin, 2003; Lamberti et al., 2008; Oesterreicher and Vajda, 2003).

### 2.5.1.5 Hierarchical clustering

Both the Jensen-Shannon divergence ( $D_{\text{JS}}$ ) and distance ( $d_{\text{JS}}$ ) are suitable for measuring the information-theoretic distance between any arbitrary network layers. Besides from the direct advantage of being able to quantify the similarity between layers, these measures can also be used to compute the pairwise distances for the purpose of hierarchical clustering (De Domenico et al., 2014, 2015a, 2016).

The hierarchical clustering of layers in a multiplex network is typically performed by iteratively aggregating pairs of layers with the least separation distance at each step (measured either as  $D_{\text{JS}}$  or  $d_{\text{JS}}$ ), and proceeding until all of the layers are aggregated, resulting in  $A_{ag}$ . By measuring the distinguishability,  $q$ , at each step we can then identify a partitioning (combination of layer aggregations) that maximises  $q$ , and thus find an economical reduced network representation,  $\mathcal{B}_{opt}$ . Although this approach is not guaranteed to find the global optimum, it does instantiate the optimisation from Equation 2.80 in a computationally efficient manner, avoiding the requirement to evaluate all possible combinations. This approach has the additional benefit of being visualisable as a dendrogram, providing an intuitive way to represent and inspect the aggregation process, layer distances, and optimal multiplex partitioning.

## 2.5.2 Multilink motifs

To interrogate the internal structural properties of a multiplex network, modified versions of the techniques previously introduced for monoplex networks can be applied (Battiston et al., 2014; Boccaletti et al., 2014; Kivela et al., 2014). An example used extensively in later chapters of this thesis is that of multilink motif analysis, where the concept of motifs is extended to consider multiple layers. Just as a motif is a pattern of edges within a network layer, a multilink motif is a pattern of edges across multiple

layers. The analysis of multilink motifs can provide insights into the distribution, interrelation, and overlap of connections in different layers, and insights into a network's functional composition.

A *multilink* describes the arrangement of edges between two nodes in a multiplex network (Bianconi, 2013; Boccaletti et al., 2014; Menichetti et al., 2014; Mondragon et al., 2017). For two nodes  $v_i, v_j$  in an  $X$ -layer network, the multilink is typically expressed as a vector:

$$\vec{m}_{ij} = (m_{ij}^{[1]}, m_{ij}^{[2]}, \dots, m_{ij}^{[X]}). \quad (2.90)$$

where  $m_{ij}^{[\beta]} = A_{ij}^{[\beta]}$ . Therefore,  $m_{ij}^{[\beta]} = 1$  if and only if there is a connection between  $(v_i, v_j)$  in layer  $\beta$ . As such, a multilink  $\vec{m}$  describes the complete set of edges between two nodes in a multiplex network. The different combinations of such edges form the set of possible multilink motifs. An illustration of multilink motifs in an example two-layer network is shown in Figure 2.10b (page 31).

As with motifs in monoplex networks, the multilink motif frequency  $J_h$  can be found by taking a subgraph census and searching for motif isomorphs. The degree to which a multilink motif is under- or over- represented in a network can then be determined by comparing the multilink motif frequency to null model samples and taking the motif  $z$ -score (see § 2.4.14, page 67).

### 2.5.3 Layer interaction

In a multiplex network it is desirable to know which nodes are active in more than one layer, to identify those that are capable of facilitating inter-layer communication. To this end, methods have been developed which quantify the multilayer participation of nodes by measuring the distribution and heterogeneity of their connections across multiple layers (Battiston et al., 2014, 2016).

In later chapters one problem we will attempt to address is to identify those nodes that are the most important in the most layers. A simple and direct approach proposed here is the *normalised degree-rank product*. For a single layer, the degree tells us the importance of a node within that layer. In the absence of data on the relative weighting and linearity of interactions – either within or between layers – ranking the nodes by degree provides a linear ordering of connectivity that can be compared across layers with different interactions types. Although the processes of ranking removes

information, it also prevents high-degree nodes from disproportionately dominating the analysis when layers are compared. The product of this linear rank then measures the importance of nodes, both within layers and between layers, and identifies those with the most potential to act as cross-layer hubs. The full method is as follows: nodes are first dense ranked in ascending order

$$\kappa_i = \text{rank}(k_i) \quad (2.91)$$

such that the lowest degree node has rank  $\kappa = 1$ , and the highest degree has rank  $\kappa \leq N$ . Due to the use of dense ranking, if two nodes exist with the same degree, then they are also assigned the same rank ( $\kappa_i = \kappa_j \iff k_i = k_j$ ), with the rank of the node with the next highest degree incrementing by 1. Thus,  $\max(\kappa) = N$  if and only if each node's degree is unique within that layer:

$$\max(\kappa^{[\beta]}) = N \iff (k_i^{[\beta]} \neq k_j^{[\beta]} \forall v_i^{[\beta]}, v_j^{[\beta]} \in V^{[\beta]} \text{ where } i \neq j). \quad (2.92)$$

As a consequence,  $\max(\kappa)$  can vary between layers, depending on the number of nodes within a layer that share the same degree. To address this, we first rescale the rank to the interval  $[0, 1]$  before taking the product. The normalised degree-rank product can thus be written as:

$$k_{norm_i} = \prod_{\beta=1}^X \left( \frac{\kappa_i^{[\beta]} - \min(\kappa^{[\beta]})}{\max(\kappa^{[\beta]}) - \min(\kappa^{[\beta]})} \right), \quad (2.93)$$

where  $k_{norm_i} \in [0, 1]$ . Computing this for all nodes, we can then find the node with the highest participation across all of the layers:

$$\underset{i}{\text{argmax}}(k_{norm_i}). \quad (2.94)$$

To illustrate, if a node has the highest degree in all of the network layers, thus participating the most across all of the network layers, it would have  $k_{norm} = 1$ .

## 2.6 Software tools

Many of the network measures and algorithms described above have published software implementations available, often in one or more network analysis applications, toolboxes, or software libraries. Many of these are open source. A listing of some of

the most popular software tools for network analysis is included in Appendix H.

The network analyses presented in the following chapters were performed primarily in MATLAB (v8.5, The MathWorks Inc., Natick, MA) using the Brain Connectivity Toolbox (Rubinov and Sporns, 2010) and MATLAB/Octave Networks Toolbox (Bounova, 2014; Bounova and de Weck, 2012). Combined, these two toolboxes implement the majority of network measures described above. Additional custom-written MATLAB functions were used for data handling, network transformations, visualisations / plots, and custom measures.

General network diagrams were created using Cytoscape (Shannon et al., 2003) and Dia<sup>1</sup>. Hive plots (see Krzywinski et al., 2011) were generated using the Python hiveplotter function (v3.5, Python Software Foundation), written and provided courtesy of Barnes (2016), using functionality from the NetworkX package (Hagberg et al., 2008). 3D visualisations of the *C. elegans* connectome were created in neuroConstruct (Gleeson et al., 2007) from model data provided by the OpenWorm (Szigeti et al., 2014) and VirtualWorm projects<sup>2</sup>. Power-law distributions were computed using the maximum-likelihood method from Virkar and Clauset (2014), and Clauset et al. (2009)<sup>3</sup>.

Reducibility analysis, hierarchical layer clustering, and multilayer plots were performed in MuxViz (De Domenico et al., 2015b), using the algorithm described in De Domenico et al. (2015a), and the Ward hierarchical clustering method to visualise layer similarity (Ward, 1963). Finally, network motifs were computed using the ESU algorithm implemented in FANMOD (Wernicke, 2005; Wernicke and Rasche, 2006). URLs are provided in Appendix H, page 263.

---

<sup>1</sup><https://wiki.gnome.org/Apps/Dia/>

<sup>2</sup><http://caltech.wormbase.org/virtualworm/>

<sup>3</sup><http://tuvalu.santafe.edu/~aaronc/powerlaws/>

## Part II

### The wired connectome



# Chapter 3

## Comparing the wired connectomes of *C. elegans*

### 3.1 Introduction

THE wired connectome of *C. elegans* has become the canonical dataset in the study of brain networks, as well as the study of complex networks in general. This is due in part to its completeness. Indeed, thirty years after its publication, *C. elegans* remains the only organism to have had its complete connectome mapped at cellular resolution. The *C. elegans* network has been instrumental in launching the field of connectomics, and insights developed from it have contributed to the understanding of several biological and network properties (Nicosia et al., 2013; Sengupta and Samuel, 2009; Stam and Reijneveld, 2007), as well as providing a test-bed to attempt whole-brain emulation at the neural level (Blau et al., 2014; Kitano et al., 1998; Petrushin et al., 2015, 2016; Szigeti et al., 2014).

While much remains to be learnt about neural processing in the worm, it is desirable to obtain connectomes for other species. The ability to perform *comparative connectomics* would facilitate the investigation of general, evolutionarily-conserved principles of brain organisation common across species (Fiore et al., 2015; Goulas et al., 2014; Heuvel et al., 2016; Meinertzhagen, 2017), and further provide the ability to study more complex behaviours and structures, including topological abnormalities associated with pathologies. With such clinical and scientific objectives in mind, the grand challenge of connectomics is to map the human connectome at the cellular level; however, the size and complexity of the human brain make this extremely challenging.

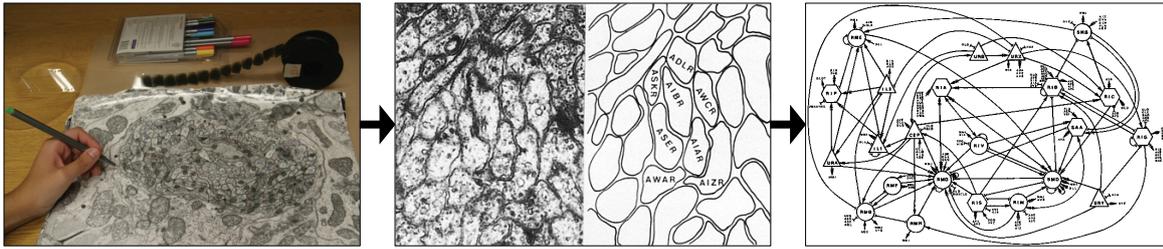


Fig. 3.1 Illustration of the manual tracing process used to constructed the original White et al. (1986) connectome. Figures modified from (Emmons, 2015; White et al., 1983, 1986).

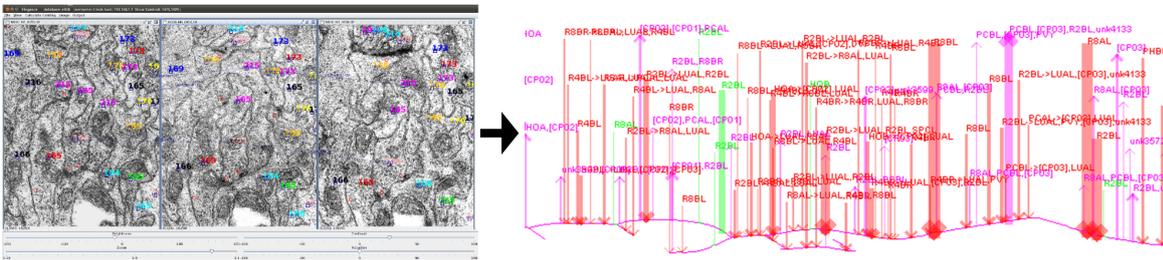


Fig. 3.2 Elegance tracing software used to create the updated WormWiring connectome. Figures modified from Xu et al. (2013).

To illustrate the difficulty of the undertaking, consider it took approximately twenty years to image and manually reconstruct the connectivity of 302 neurons in *C. elegans* by tracing electron micrograph (EM) sections (Emmons, 2015); see Figure 3.1. If the same approach was used for the human brain, with its estimated 100 billion neurons (Williams and Herrup, 1988), and assuming that the time taken scales linearly with the number of cells, it would take  $\sim 7$  billion years to map the whole human brain, or half the current age of the universe.

Clearly the approach used to map the worm cannot be applied to larger brains. A critical task for meso- & micro- scale connectomics then, is to develop a fast accurate method to automatically trace neurons through micrographic sections and identify synaptic contacts. Advances in microscopy and computer vision have begun to make this possible in recent years, and several large-scale projects are now attempting to map the connectomes of both flies and mice (Chklovskii et al., 2010; Osten and Margrie, 2013).

One computer-assisted method, Elegance, has recently been used to reconstruct the previously unmapped tail circuits of a *C. elegans* adult male (Jarrell et al., 2012), enabling the reverse engineering of networks that control worm mating behaviour (LeBoeuf et al., 2014). The same reconstruction software (Xu et al., 2013)<sup>1</sup> has also been applied to the original White et al. (1986) TEM images (Figure 3.2), to produce an updated hermaphrodite connectome using automated methods (Emmons et al., 2015)<sup>2</sup>. This provides a unique opportunity to compare manual and computer-assisted neuron tracing techniques across an entire nervous system – something which is not currently possible in any other organism.

### 3.1.1 Chapter overview

This chapter analyses the first release of the new *C. elegans* connectome, generated using computer-assisted methods, and compares it to the standard worm connectome derived from manual tracing. Topological changes between the two networks are identified, to characterise the differences between manual and automated reconstruction methods, and to assess whether findings in the worm connectome literature are still valid for the new network.

## 3.2 Materials & Methods

### 3.2.1 Connectivity data

Analyses were performed on both the original and updated hermaphrodite *C. elegans* networks, using data taken from three sources. These networks, referred to herein by their two-letter abbreviations, are:

- *Albertson-Chklovskii connectome* (**AC**) — the canonical hermaphrodite connectome, composed from the somatic network of White et al. (1986), updated and released by the Chklovskii lab (Chen, 2007; Chen et al., 2006; Varshney et al., 2011); and the pharyngeal network of Albertson and Thomson (1976), made available by the Cybernetic *Caenorhabditis elegans* Program (CCeP)<sup>3</sup> (Oshio et al., 2003).
- *WormWiring connectome* (**WW**) — the first release of the new hermaphrodite connectome, generated using the computer-assisted method described by Xu et al.

---

<sup>1</sup><https://github.com/emmonslab/>

<sup>2</sup><http://www.wormwiring.org/>

<sup>3</sup><http://ims.dse.ibaraki.ac.jp/ccep/>

(2013). This network is the version provided as of May 2015, courtesy of Scott W. Emmons. The most recent version is available for download from the internet<sup>4</sup>.

The functional classifications referred to in this chapter (i.e. *sensory neuron*, *interneuron*, *motor neuron*) are based on the classification schema used by WormAtlas (Altun et al., 2002) and WormWeb. Full details are given in Appendix F, page 251.

### 3.2.2 Topological network measures

All network measures are the same as those described in Chapter 2. The connected components, edge counts, and adjacency matrices used to identify differential connectivity were generated using binary directed versions of the full datasets. The same networks, excluding self-connections ( $A_{ij} = 0$  if  $i = j$ ), were used to compute the network density and reciprocity measures. All other analyses used binary undirected versions of the networks with self-connections removed, except where specified otherwise.

Network measures are compared to 100 null model networks (shown in the boxplots) generated using the degree-preserving edge-swap randomisation procedure described in Chapter 2 (Algorithm 5), where networks were randomised with  $10 \times M$  edge-swaps. As per the definition given in Humphries and Gurney (2008), the small-world index was normalised against 100 Erdős-Rényi (ER) reference networks containing the same number of nodes and edges as the actual networks.

### 3.2.3 Software

Analyses and data handling were performed using the methods described in § 2.6. URL links to the software tools are provided in Appendix H, page 263.

### 3.2.4 Literature search

To determine if any known function exists for synapses missing from the WW network, a search was performed for each pair of neuron class IDs against the available literature using the Google Scholar search engine<sup>5</sup>. Search results were reviewed manually to assess their relevance.

---

<sup>4</sup>See footnote 2

<sup>5</sup><http://scholar.google.com/>

### 3.3 Results

Profiling the gross structure of the AC and WW networks reveals a large disparity between the number of links. The new WW connectome increases the total number of synaptic edges by 56 % over the original network, with an even greater increase of 96 % for gap junctions, giving a corresponding rise in total network density of 57 % (Table 3.1). Inspecting the distribution of the degree change between the aggregate networks shows that the degree increase approximately follows a Gaussian distribution, with a mean change of +6.83 per node (Figure 3.3).

Measure		AC	WW	$\Delta$
N <sup>o</sup> synaptic edges	$M_{syn}^{\rightarrow}$	2284	3572	1288
N <sup>o</sup> gap junction edges	$M_{gj}$	568	1111	543
N <sup>o</sup> aggregate network edges	$M_{ag}^{\rightarrow}$	3162	5028	1866
Aggregate network density	$D_{ag}^{\rightarrow}$	0.035	0.055	0.02

Table 3.1 Edge count & network density.

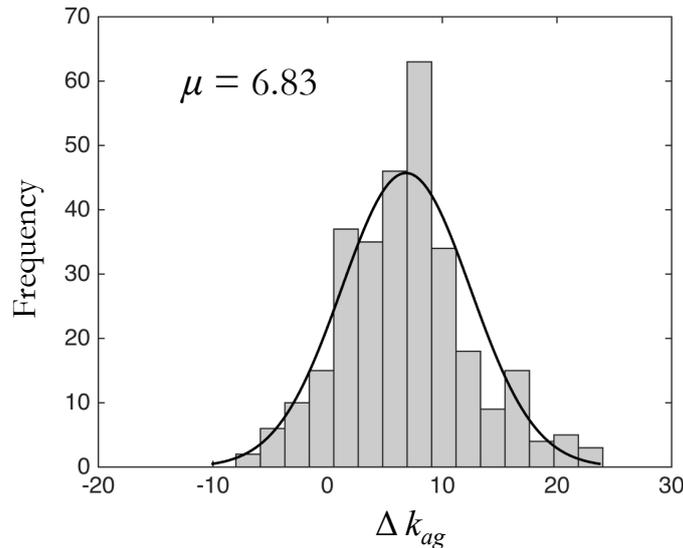


Fig. 3.3 Histogram of degree differences between the aggregate AC & WW networks.

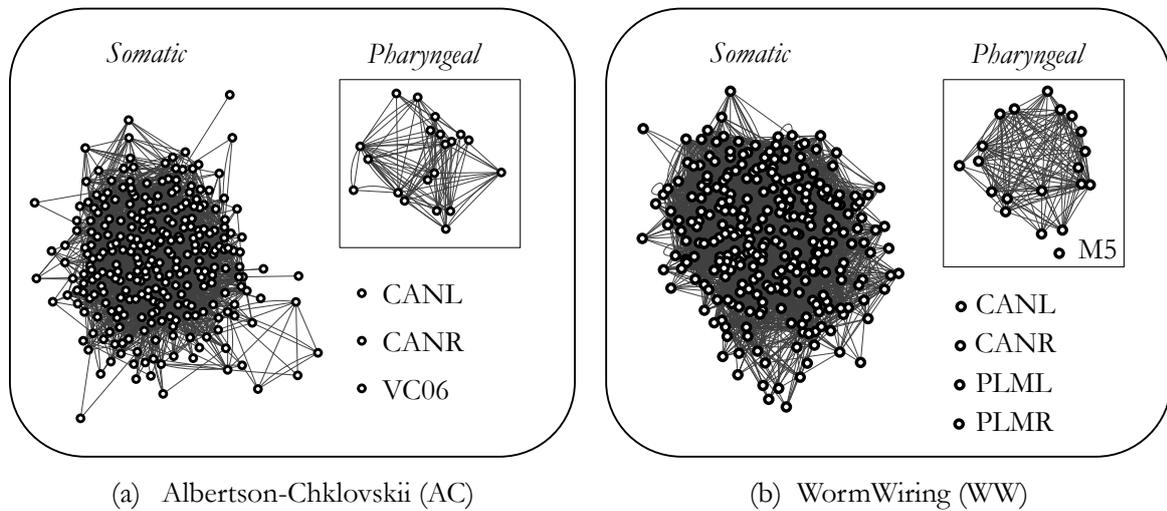


Fig. 3.4 Network wiring diagrams of the synaptic connected components in the AC & WW connectomes. Isolated neurons are identified by name.

Counterintuitively, an examination of the connected components of both synaptic networks reveals that the overall increase in connectivity seen in WW does not correspond with a reduction in network fragmentation (Figure 3.4). Instead, the opposite is observed. Comparing the wiring diagrams of the two networks reveals that while the previously unconnected VC06 motor neuron merges with the main component in WW (Figure 3.4b), three previously connected cells become separated (M5, PLML/R).

Analysing the changes in link types individually, the degree-degree correlations show that the increases seen in WW are uniformly distributed across the synaptic network, yet the changes in gap junction degree are much more diverse (Figure 3.5). Plotting the networks in an adjacency matrix highlights that the variations in gap junction connectivity form two anatomically distinct clusters in the worm (Figure 3.6b): a large increase in connectivity between interneurons in the somatic network (Figure 3.6b, *magenta*), and a decrease in gap junctions within the pharynx (Figure 3.6b, *green*). The adjacency matrices further reveal a widely distributed pattern of synapses and gap junctions in AC that are no longer present in WW (Figure 3.6 a & b, *green*), explaining the observed fragmentation. These total 154 gap junctions, and 295 synapses (listed in Table 3.2).

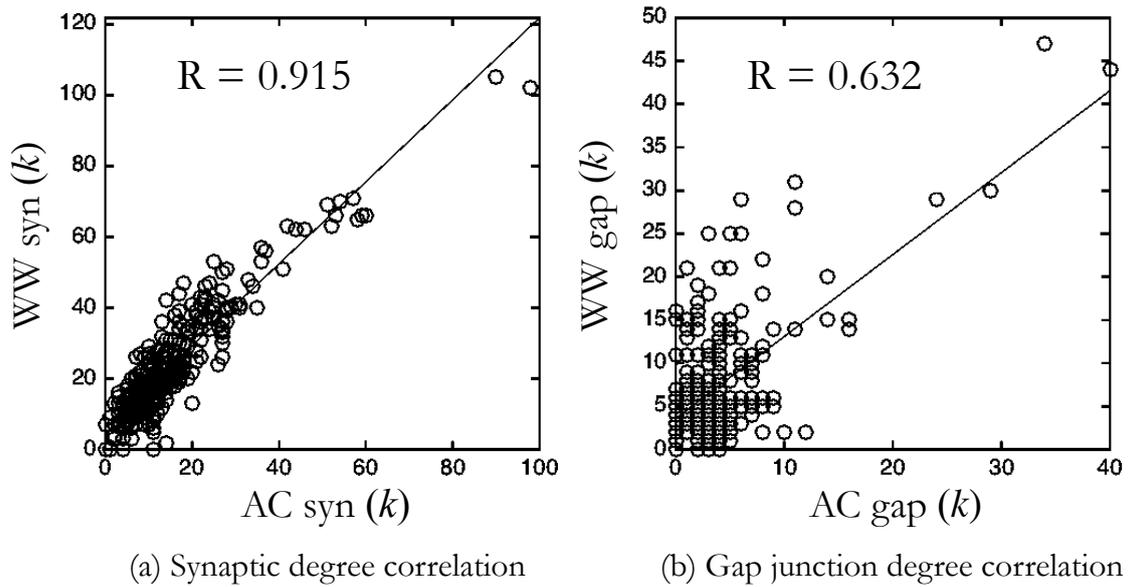


Fig. 3.5 Degree-degree correlation between individual nodes in the AC & WW networks.

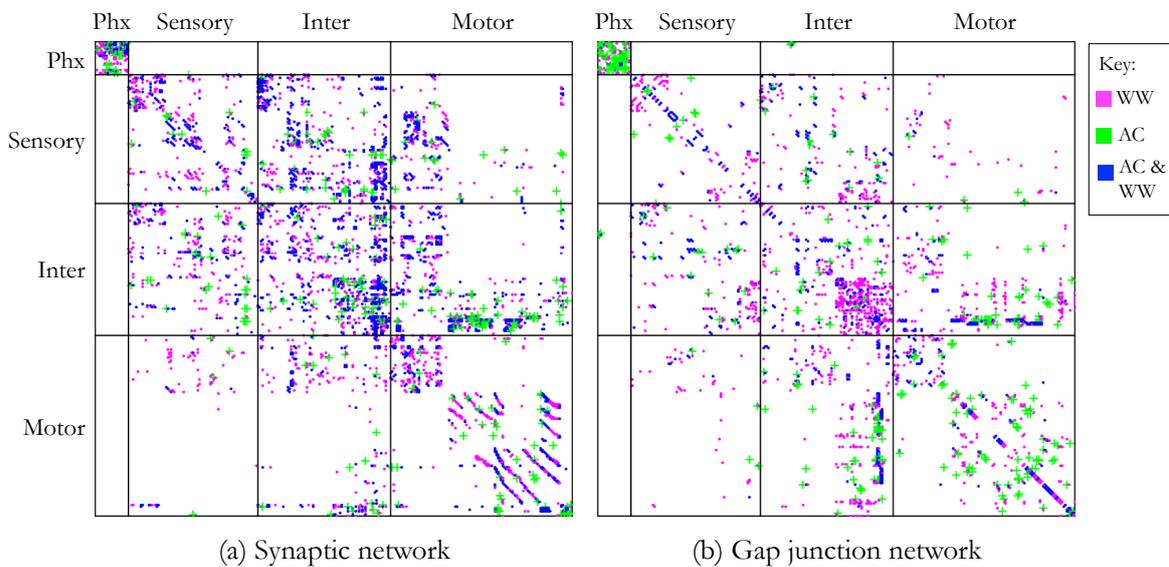


Fig. 3.6 Adjacency matrices showing the differences between the AC and WW synaptic and gap junction networks. New connections in WW that do not exist in AC are shown in magenta, those in AC but not WW are in green, and the overlap between the two networks is shown in blue. Nodes are classified as pharyngeal (Phx), sensory, motor, or interneurons.

### 3.3.1 Connections omitted in WW

To estimate the accuracy of the Elegance method used to generate the new WW connectome, Xu et al. (2013) performed three reconstructions of a single neuron in the male connectome, and evaluated differences in the scoring of chemical synapses. These replications revealed that cells were consistently paired 66 % of the time, suggesting that the majority of connections in the WW connectome derived from this method are likely accurate. Although it is not possible to determine which of the new connections in the WW connectome are real, it is possible to identify omissions that represent likely errors. Of the synapses that are no longer in WW, but which exist in the original network (Table 3.2), the literature suggests that at least some of these omissions are incorrect. For example, considering the egg-laying circuit, it has been shown that reciprocal connections between HSN and VC04/05 play a role in regulating egg-laying behaviour (Zhang et al., 2010), while ALM and PLM are both known to inhibit egg-release through synapses onto HSN (Sawin, 1996; Schafer, 2005; Zhang et al., 2008). Missing connections between I2 and I6 have been independently verified to exist in the pharynx (Bhatla et al., 2015), and multiple connections dropped from WW are within the well characterised locomotory systems: these include connections required to coordinate contralateral muscles groups ( $\{DA, DB\} \rightarrow VD$  &  $\{VA, VB\} \rightarrow DD$ ) (McIntire et al., 1993; Zhen and Samuel, 2015), control forward and backward locomotion (e.g.  $PVC \rightarrow \{DB, VB\}$ ;  $\{AVD, AVE\} \rightarrow DA$ ;  $AVA \rightarrow AVB$ ), and inhibit forward movement in response to posterior body touch ( $PLM \rightarrow \{AVA, AVD\}$ ) (Chalfie et al., 1985; Driscoll and Kaplan, 1997). One missing connection is also thought to be involved in inhibiting movement during sleep-like lethargus ( $ALA \rightarrow AVA$ ) (Fry et al., 2014).

### 3.3.2 Asymmetric gap junctions

Inspecting the weighted directed version of the WW gap junction data reveals that eighteen of the edges are asymmetric ( $A_{ij} \neq A_{ji}$ ), lacking one or more matching reciprocal connections (Table 3.3). Given the inherently bidirectional nature of gap junctions, it seems likely that the observed asymmetries stem from neuron tracing errors. These asymmetries span the length of the worm, but are all localised within the ventral nerve cord. The majority of these connections are to ventral cord motor neurons, with the other neurons being either sensory or interneurons with processes that extend into and along the length of the ventral cord. These neurons are visualised in Figure 3.7.

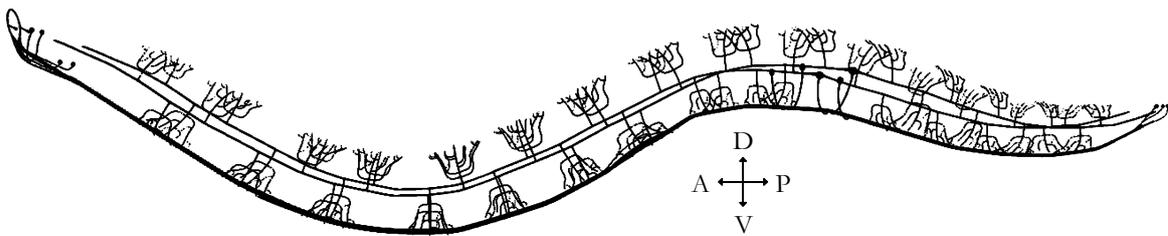


Fig. 3.7 Sensory and interneurons with asymmetric gap junctions, all with processes in the ventral nerve cord. *Visualised using data from OpenWorm (Szigeti et al., 2014).*

I1L→M1	PLMR→AVAR	AVFL→PVQR	PVCL→DA06	RMFL→RMGR
I1L→M2R	PLMR→AVDL	AVFR→ASJL	PVCL→DB05	RIVR→RMEV
I1L→NSMR	PLMR→AVDR	AVFR→ASKL	PVCL→DB06	SMBDL→AVAR
I2L→I6	PLMR→AS06	AVFR→AVJR	PVCL→VB03	SMBVL→PLNL
I2L→MCL	PLMR→HSNR	AVFR→VD11	PVCL→VB05	URADR→RMDVR
I2R→I6	ALNR→RMHR	AVG→AVL	PVCL→VB08	URADR→RMED
I3→I1R	PLNL→SAADL	AVG→PVPR	PVCL→VB10	DA01→VD01
I3→M2R	PLNL→SMBVL	AVG→AVEL	PVCR→PDEL	DA02→VD01
I3→NSML	PLNR→SAADR	AVHL→PVQL	PVCR→DA09	DA04→DA03
I4→M1	PLNR→SMBVR	AVHR→ADLL	PVCR→DB05	DA04→DB03
I5→M5	FLPR→AIBR	AVHR→PVNL	PVCR→DB06	PDA→AS11
I6→M3R	BAGL→RIBL	AVJL→PLMR	PVCR→DB07	DB01→AS03
M1→M3L	URYVR→SIBDR	AVJL→AVFR	PVCR→AS02	DB01→VD01
M1→M3R	ADEL→SIBDR	AVJL→HSNR	PVCR→VB07	DB02→VD05
M1→M5	ADER→ALA	AVJR→PVQR	PVCR→VB08	AS03→DD01
M1→MCL	ADER→RIH	AVKL→PVM	PVCR→VB10	AS03→VD02
M1→M3R	ADER→AVKL	AVKL→PDEL	AVAL→AVHL	AS04→AS05
M4→M1	ADER→AVKR	AVKL→PDER	AVAL→AVBR	AS05→DD02
M5→I5	ADER→PVR	AVKL→RMFR	AVAL→DA06	AS09→AVAR
M5→M1	ADER→SAAVR	AVKL→SMDDR	AVAL→DB05	AS11→PDA
MI→I5	PDEL→PVM	AVKR→SMDDL	AVAL→DB06	AS11→PDB
MI→M3R	PDEL→PDER	AVL→AVFR	AVAL→AS03	DD05→VB08
NSML→I6	PDEL→AVKL	AVL→DA02	AVAR→ADER	DD05→VD09
NSMR→M3L	PDEL→PVR	DVA→PDEL	AVAR→PDEL	VA08→VA09
NSMR→M4	PDEL→VA09	DVA→DB05	AVAR→PDER	VA08→VB08
ADFL→OLQVL	PDEL→VD11	DVA→DB06	AVAR→AVEL	VA09→PVT
ADLL→AIBR	PDER→AVKL	DVC→AVKR	AVAR→SABD	VB01→DVA
AWAR→ASEL	PDER→PVCL	PVNL→RIFL	AVAR→SABVL	VB01→SAADR
AWBR→SMBVR	PDER→PVCR	PVNL→AVFR	AVAR→DB05	VB01→SABD
AWCL→ASEL	PHBR→AVFL	PVNL→AVEL	AVAR→DB06	VB06→VB07
IL1DL→URYDL	PHBR→DA08	PVNR→BDUR	AVAR→AS03	VB07→DD04
IL1DL→RIH	AIAL→HSNL	PVNR→PVNL	AVAR→AS10	VB08→VA09
IL1DL→PVR	AIBL→ASER	PVNR→PVCL	AVAR→VD13	VB08→VB09
IL1L→IL1DL	AIBL→AIYL	PVNR→AVBL	AVBL→AVDL	VB09→VB08
IL1R→IL1DR	AIYR→HSNL	PVNR→DD03	AVBL→AVEL	VD08→DD04
IL1VR→RMER	ADAL→FLPR	PVNR→VD12	AVBL→AS03	VD10→DD05
OLQDL→SIBVL	ADAR→RICL	PVNR→VC03	AVBL→AS06	VD12→VB10
OLQDR→SIBVR	ADAR→AVAL	PVPL→AVER	AVBL→AS09	VD13→PVCL
OLQVL→IL1VL	AIML→AVBR	PVPR→RMGR	AVBL→VA07	HSNL→AVDR
OLQVL→IL2VL	AIML→AVDR	PVPR→RIMR	AVBR→DA05	HSNL→SABVL
OLQVL→RIPL	AIML→AVER	PVQR→DVC	AVBR→AS05	HSNL→VC05
OLQVL→RIH	ALA→ADEL	PVQR→PVT	AVBR→AS06	HSNR→AIBR
OLQVL→SIBDL	ALA→AVAL	PVQR→AVBL	AVBR→AS07	HSNR→AVL
OLQVR→SIBDR	RIPL→OLQDL	PVQR→HSNR	AVBR→VA09	HSNR→VD04
CEPDR→RIBR	RIPR→OLQDL	PVR→IL1DL	AVDL→AS01	HSNR→VC03
CEPDR→RIS	RIAL→CEPVL	PVR→PDEL	AVDL→AS05	VC01→DD03
CEPDR→SIADL	RIAR→SIADL	PVR→PDER	AVDR→DA09	VC02→PVCR
CEPDR→RMDDL	RICL→AVKR	PVR→ADAL	AVDR→AS05	VC02→DD03
CEPDR→SMBDR	RIFR→AVHL	PVR→PVCR	AVDR→VA06	VC03→DVC
CEPVL→ADLL	RIGL→DVC	PVR→AS06	AVEL→DA04	VC03→DD03
AVM→VA01	RIGL→RMFL	PVT→SMBDR	AVER→VA05	VC03→DD04
PVM→PDEL	RIH→CEPDR	PVWL→AVJL	SAADL→RMGR	VC04→AVBL
PVM→PDER	RIS→CEPDR	PVWL→VA12	SAAVL→ALNL	VC04→VC01
PLML→HSNL	RMGL→ALML	PVWR→PVT	SABD→VA02	VC04→VC05
PLMR→PDEL	RMGL→ALNL	PVWR→PVCR	RMDDR→URYDL	VC05→OLLR
PLMR→PDER	RMGL→SIBVL	RID→AS02	RMDL→RMFL	VC05→URBL
PLMR→DVA	RMGR→AVAR	RID→DD03	RMDR→AVKL	VC05→HSNL
PLMR→PVCL	AVFL→PDER	RID→VD13	RMDVR→SIBVR	VC05→VC03
PLMR→AVAL	AVFL→AVJR	PVCL→DA02	RMFL→RIGR	VC05→VC04

Table 3.2 Cell-cell synaptic connections in AC that no longer exist in WW.

Cell A	Cell B	Sent	Received
PVM	DB03	2	0
PVM	VD09	1	0
PVDR	PVCR	2	0
PDEL	VD06	2	0
PDER	AVKL	0	1
DVA	VB04	0	1
AVAL	VA05	4	5
AVAL	VA09	5	4
AVAR	VA05	5	4
AVAR	VA09	4	5
AVBL	DB06	4	3
AVBL	VB08	4	3
AVBL	VB09	3	4
AVBL	VB10	4	3
AVBR	DB06	3	4
AVBR	VB08	3	4
AVBR	VB09	4	3
AVBR	VB10	3	4

Table 3.3 Asymmetric gap junctions in the WW connectome.

### 3.3.3 Network measures

Comparing the global properties of the two networks reveals that, despite the large changes at the level of individual edges, many of the statistical features remain essentially unmodified, including the clustering coefficient, maximised modularity, and characteristic path length (Figure 3.8a-c), all of which correspond with previously reported values (Humphries and Gurney, 2008; Reese et al., 2012; Watts and Strogatz, 1998). All of these measures are also much higher than expected compared to random networks with the same degree distributions. This suggests that at least some parts of the neural network of *C. elegans* are functionally segregated (i.e. highly clustered modules).

The largest change displayed between the two networks is in assortativity. As previously reported, the AC connectome is weakly disassortative (Newman, 2002; Rudolph-Lilith and Muller, 2014), though it was found here to be no more so than random (Figure 3.8d). This is no longer the case in the WW connectome.

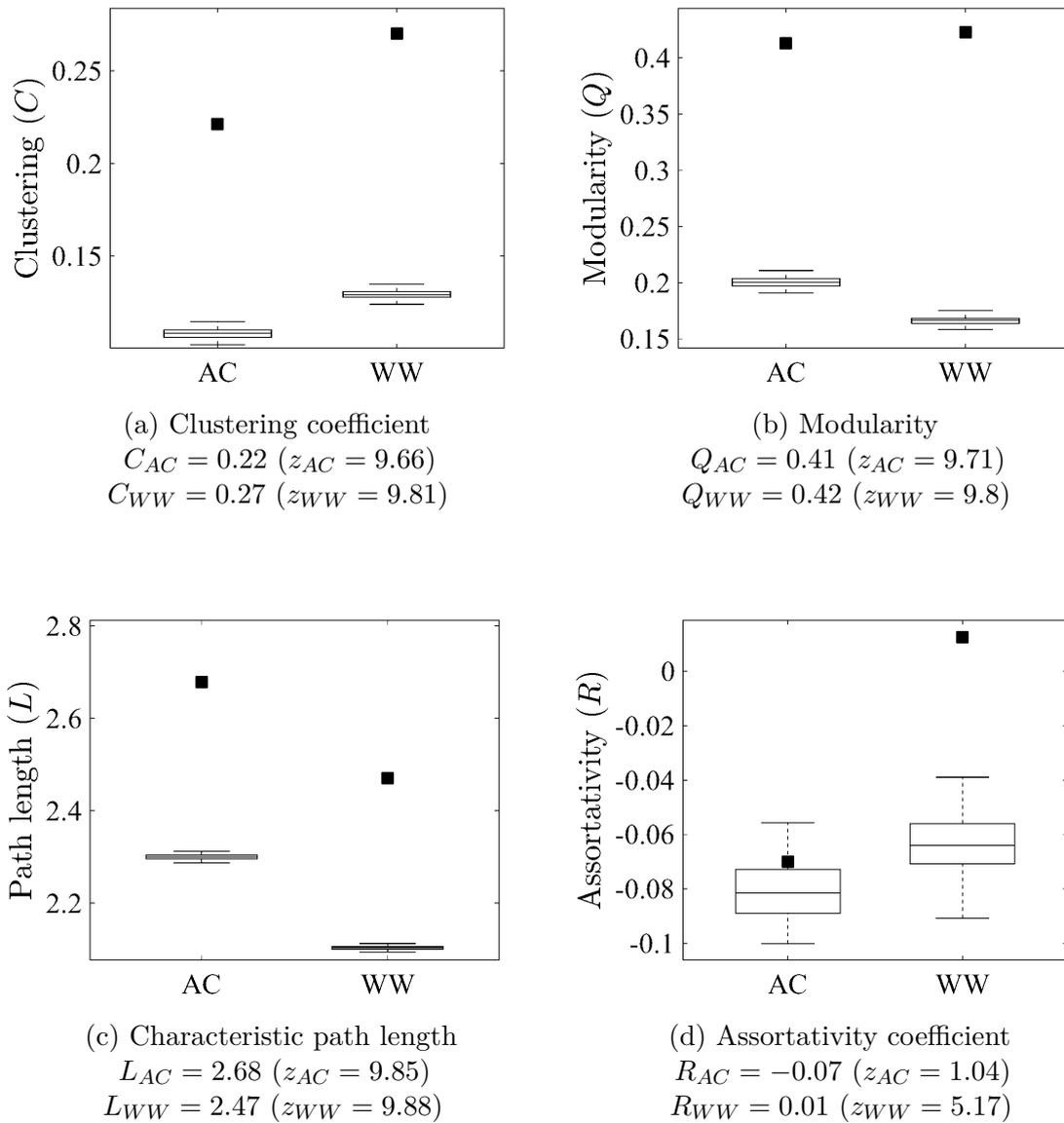


Fig. 3.8 Comparison of network metrics for the AC & WW aggregate networks, showing observed values (filled squares) and expected values from 100 randomised networks (boxplots).

While there is virtually no assortative mixing in WW ( $R_{WW} = 0.01$ ), there is a significant increase in assortativity relative to both the AC network and the null model control ( $z_{WW} = 5.17$  compared to random networks with the same degree distributions). This change can be partially explained by the presence of new connections between interneurons, mentioned previously (Figure 3.6). Consider that disassortativity describes the dominant presence of hub-and-spoke like architectures (central high-degree nodes connected to a low-degree periphery). In the worm, this motif is prominent in connecting sensory neurons to interneurons (e.g. Macosko et al., 2009), and between interneurons and motor neurons (Chalfie et al., 1985). The majority of high-degree to high-degree connectivity is between interneurons, especially those in the nerve ring – as reflected by the rich-club (RC) members (Towlson et al., 2013); therefore, the preferential increase in interneuron connectivity seen in WW is the most likely cause for the change in assortative mixing, which is further reflected in the expansion of the rich-club (§ 3.3.7).

### 3.3.4 Scale-free & small-world properties

Concerning topological arrangement, the AC and WW networks both exhibit heavy-tailed degree distributions in line with previous reports (Rudolph-Lilith and Muller, 2014) (Figure 3.9), and the WW network retains small-world characteristics, though slightly reduced compared to AC (Figure 3.10b).

### 3.3.5 Robustness

Robustness measures the stability of a network to structural damage. As the maximum robustness value is defined as  $\rho = 0.5$  (Schneider et al., 2011a,b), it can be seen that the architectures of both the AC and WW connectomes have a high fault-tolerance (Figure 3.10a), though considerably less so than their randomised counterparts. The high resilience is explained by the reported finding that networks with power-law degree distributions undergo gradual, rather than catastrophic, deterioration in response to lesions (Albert et al., 2000); however, the extent to which this is true is also partially affected by assortativity – components in a highly connected, *assortative*, network are much harder to destroy than those in a hub-and-spoke, *disassortative*, network where the removal of the central hub results in complete dissociation of the network (Newman, 2002; Rubinov and Sporns, 2010; Teller et al., 2014).

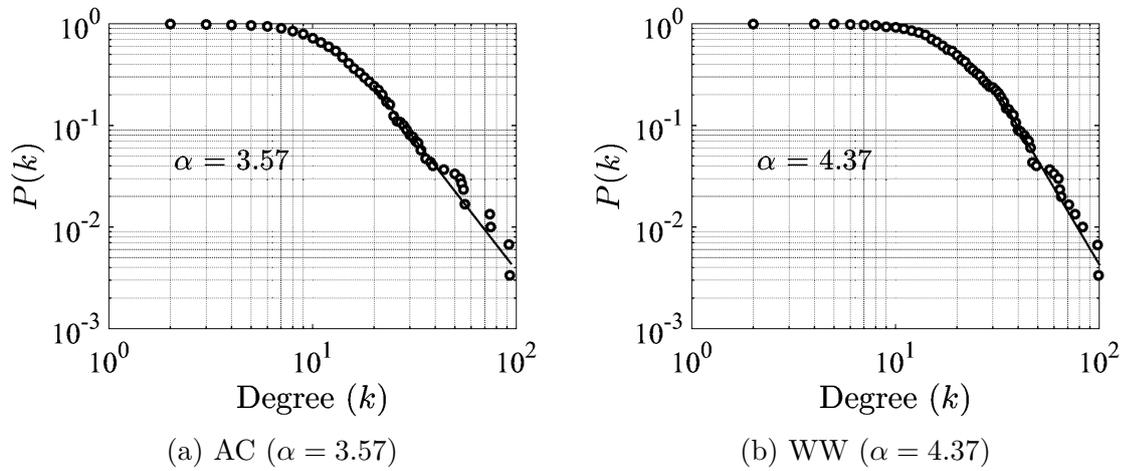


Fig. 3.9 Scale-free properties of the AC & WW aggregate networks (showing best fit for the power-law distribution  $P(k) = k^{-\alpha}$ ).

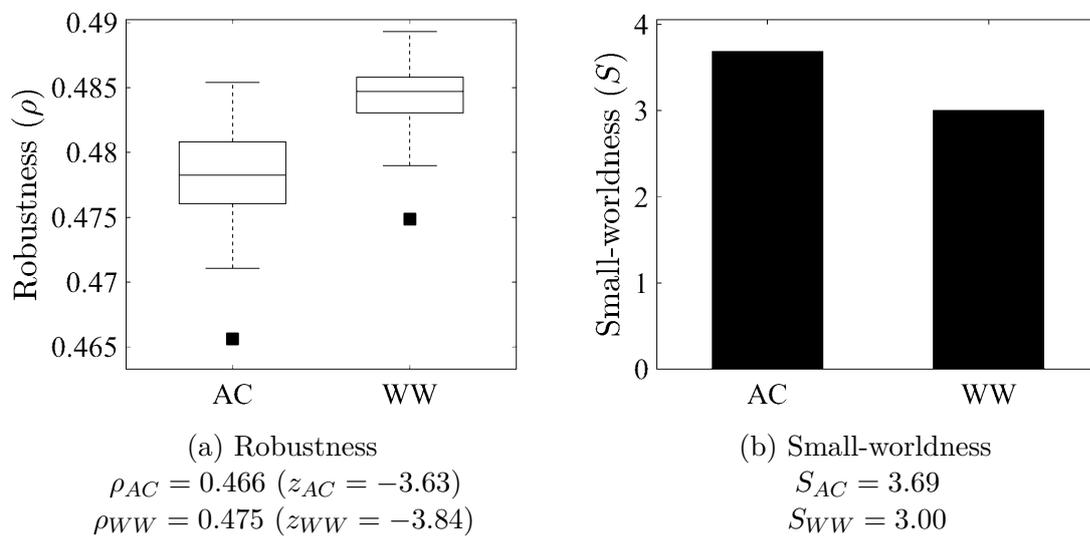


Fig. 3.10 Robustness and small-worldness of the AC & WW aggregate networks. Robustness plot shows observed values (filled squares) and expected values from 100 randomised networks (boxplots).

As mentioned previously (and detailed in § 3.3.7) both of the wired *C. elegans* connectomes exhibit a core-periphery structure with an intrinsic mix of assortative and disassortative circuits. This combination of predominant robustness with localised vulnerability is evidenced by experimental observations. For example, 85 % of the neurons in the worm’s pharynx can be ablated without affecting viability – most of which are interneurons – while the remaining 15 % are essential for normal feeding behaviour (Avery and Horvitz, 1989). Similarly, large portions of the gap junction network can be genetically ablated (Kawano et al., 2011; Simonsen et al., 2014) and whole neurotransmitter systems disrupted without causing lethality (e.g. Jin et al., 1999). This level of robustness also extends to the extrasynaptic networks, where all 2036 monoamine edges can be knocked out without major consequences (Duerr et al., 1999); yet, certain behaviours show comparatively high vulnerability and are known to be dependent on a few peripheral neurons. For example, ablating two classes of sensory cells is sufficient to abolish almost all responses to nose-touch (Kaplan and Horvitz, 1993).

Even though both the AC and WW networks have high robustness, it is surprising that the worm connectomes should be less robust than randomly rewired versions. The fact that random networks are *more robust* than the evolutionarily-designed one suggests the presence of a conflicting design objective, and a trade-off against a more desirable network property. The random counterparts, having the same number of nodes and edges, demonstrate it is unlikely to be the result of an energetic constraint with respect to the structure.

Other biological networks have previously been found to have lower robustness than expected, including the protein-protein interaction (PPI) networks of several organisms. Discussing PPI networks, Schneider et al. (2011a) suggest that reduced resilience might be a consequence of modular structure. Modularity, by definition, reduces global interconnectedness in favour of local connections; this in turn increases a network’s fragility, as modules are vulnerable to becoming disconnected. The robustness metric thus suggests that the *C. elegans* nervous system is preferentially selecting functional segregation at the expense of structural fault-tolerance.

### 3.3.6 Reciprocity

Reciprocity is an important feature in brain networks. Depending on the sign and strengths of the synapses, reciprocal connections can perform several biologically important computations. These include signal integration, associative memory, amplification, gain control, as well as the generation of synchronised, oscillating, or tonic signals (Dayan and Abbott, 2005; Getting, 1989; Sommer and Wennekers, 2003). Reciprocal connections also provide a simple method of circuit switching (e.g. Li et al., 2012), and have been found to be overrepresented in the nervous systems of both worms and mammals (Reigl et al., 2004; Song et al., 2005). In *C. elegans* reciprocal connections are known to be involved in coordinating locomotion (Roberts et al., 2016; White et al., 1986) and regulating male mating behaviours (Correa et al., 2012).

During the original mapping of the worm connectome, White et al. reported the presence of a bias against reciprocal connections in *C. elegans* (White et al., 1983). This was later shown to be false by Reigl et al. (2004).

Analysing the reciprocity of the AC connectome confirms that there are more reciprocal connections than random – with a bias *for*, rather than *against*, reciprocity (Figure 3.11) – however, these only account for 13 % of the total network, which likely lead to the original statement by White et al. (1983).

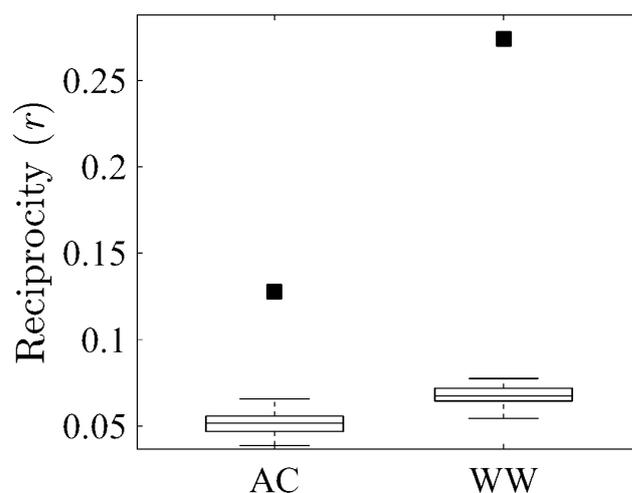


Fig. 3.11 Reciprocity for the directed AC & WW synaptic networks, showing observed values (filled squares) and expected values from 100 randomised networks (boxplots).

In the WW network, the number of reciprocal connections increases to approximately 27 % of all synapses; much higher than in the randomly rewired networks with the same degree distributions ( $z_{WW} = 9.63$ ). This increase in reciprocity over the original AC network is also reflected in the synaptic adjacency matrix, where the WW connections are seen to be more symmetrically distributed along the diagonal (Figure 3.6a).

### 3.3.7 Rich-club organisation

Many neural systems have previously been reported to contain a core rich-club (RC) of densely connected neurons (Harriger et al., 2012; Heuvel and Sporns, 2011; Liang et al., 2017; Reus and Heuvel, 2013; Shih et al., 2015), including in the wired connectome of *C. elegans* (Towlson et al., 2013). Comparing the rich-club coefficients of the AC and WW connectomes reveals that both networks retain this feature, with a high-degree core that is more densely interconnected than random (Figure 3.12).

Examining the neurons within the RCs reveals a large overlap between the two cores (Table 3.4). The AC rich-club is composed entirely of interneurons, most of which are in the head. Three new members are identified here that were not found by Towlson et al. (2013) (at  $k = 34$ ). Two of these neurons are in the same class as existing RC members (AIB, RIB), with the remaining neuron following the established pattern of being a head interneuron (AVKL).

The RC core is found to be expanded in the WW network, including all but three of the RC members from AC, along with an additional seven neurons. Of the new RC neurons, one is the contralateral partner to an existing member (RIAL), three are tail interneurons with processes in the nerve ring (PVT, PVR, PVNR), two are head interneurons (AVHR, RMGL), and the remaining cell is an egg-laying control neuron that is also a member of the monoamine rich-club (HSN; see Chapter 4), suggesting a central role in linking the extrasynaptic and wired transmission networks.

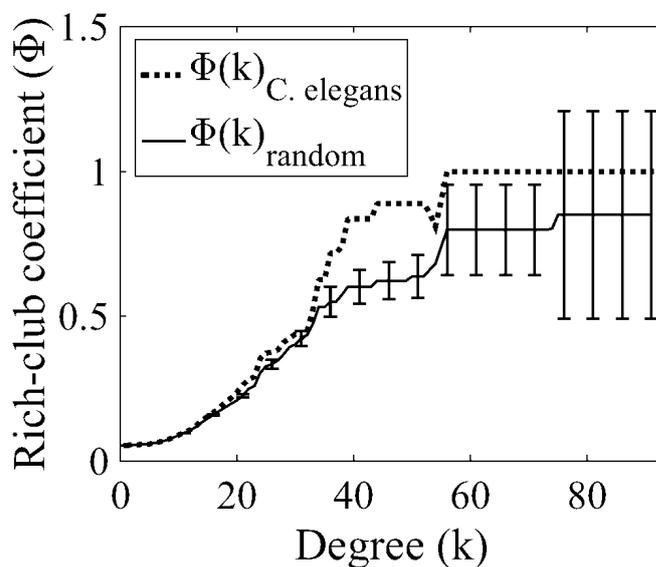
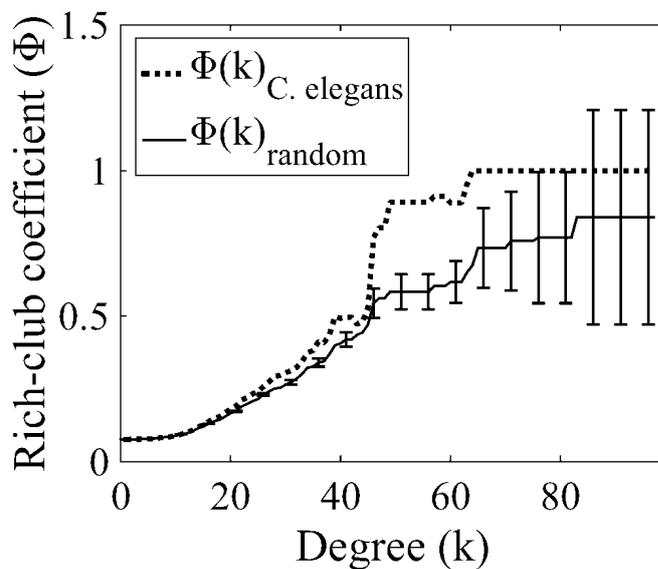
(a) AC network (rich-club at  $k \geq 34$ )(b) WW network (rich-club at  $k \geq 45$ )

Fig. 3.12 Rich-club curves for the undirected AC & WW aggregate networks. Dashed line indicates the rich-club coefficient for the *C. elegans* network and the solid curve represents the average rich-club coefficient of 100 randomised networks (preserving degree distribution) at each value  $k$ .

AC		WW	
rich-club	$k_{agAC}$	rich-club	$k_{agWW}$
AVAR	93	AVAL	99
AVAL	92	AVAR	98
AVBL	75	AVBR	83
AVBR	74	AVBL	76
AVER	56	DVA	71
AVDR	55	AVER	65
AVEL	55	AVDR	64
PVCL	54	PVCR	63
PVCR	53	AVEL	63
DVA	50	PVCL	60
AVDL	44	AVDL	57
AIBR	39	AIBR	49
RIBL <sup>†</sup>	38	PVNR <sup>*</sup>	47
RIAR	36	AIBL	46
—		RIAL <sup>*</sup>	46
<i>AIBL</i>	34	RIAR	46
<i>RIBR</i> <sup>†</sup>	34	AVHR <sup>*</sup>	46
<i>AVKL</i> <sup>†</sup>	34	PVR <sup>*</sup>	46
		RMGL <sup>*</sup>	45
		PVT <sup>*</sup>	45
		HSNR <sup>*</sup>	45

Table 3.4 Rich-clubs of the AC and WW aggregate networks at the  $\Phi_{norm}(k) \geq 1 + 3\sigma$  level, including newly identified AC rich-club members ( $k = 34$ ) not described by Towlson et al. (2013). Neurons in the AC rich-club which are no longer classified as such in WW are identified with <sup>†</sup>. Newly identified rich-club members in WW are marked <sup>\*</sup>.

## 3.4 Discussion

### 3.4.1 Synaptic networks

On the somatic connectivity of the AC network, Varshney et al. (2011) estimate that their connectome is about 90 % complete, and conclude from synaptic antibody staining that only  $\sim 5$  % of the chemical synapses are missing.

The 56 % increase in chemical synapses observed in the WW network suggests one of two possibilities: either the number of synapses in the *C. elegans* nervous system has been vastly underestimated in the past, or the WW network (and the Elegance reconstruction method) has a high false-positive rate.

If the former is true, this might have consequences for existing interpretations of experimental data. Consider, for example, processes described to be cell-autonomous: if experiments have been conducted and interpreted under an incorrect assumption that there are no upstream or parallel synapses that have not been controlled for, these interpretations could be invalid, in turn affecting any larger theories based on them. If, instead, the latter is true and there is a high number of false-positives, this raises concerns about using WW to interpret findings, and has further implications for the accuracy of the male connectome generated using the same method.

In either case, it is worth considering the possibility that the anomalous addition (or omission) of individual connections could lead to erroneous interpretations of biological data, and the generation of false theories. It should therefore be emphasised that no single connectome can be considered authoritative at this stage, and any descriptions of local circuit features should be treated with caution, regardless of the connectivity data used.

While the literature suggests the presence of at least some false-negative mapping errors in WW (i.e. missed connections), it is impossible to quantify the full accuracy of either the AC or WW network from the available data. The highest confidence for a false-negative exists for the missing connections between I2 and I6, due to the additional pharyngeal reconstruction performed by Bhatla et al. (2015).

### 3.4.2 Gap junction networks

The largest increase in connectivity seen in the WW reconstruction is in the gap junction network. Discussing this increase with the authors, Emmons et al. confirm that they score gap junctions more liberally than in the original work by White et al. (Scott W. Emmons, Personal Communication, June 2015).

If the new gap junction mapping is correct, it identifies a cluster of connectivity between the worm's interneurons that has not previously been described. As gap junctions physically couple the internal cytoplasm and electrical potentials of cells, this points to these interneurons as possibly forming a single functional control unit.

The asymmetric connections in the WW gap junction network are strongly indicative of mapping errors, however these form only a small fraction of the total gap junction network. The fact that they are localised to the ventral nerve cord suggests these errors are artefacts specific to this structure, either from poor EM data or difficult-to-score connectivity patterns, rather than a general flaw in the Elegance method.

The localised decrease in gap junctions seen in the pharynx might also be an error specific to the structure or source data. It is known that gap junctions play a role in regulating pharyngeal pumping, and that mutants with gap junction defects have acute feeding phenotypes; however, such phenotypes are likely due to their role in muscles, rather than neurons (Altun et al., 2009; Chalfie et al., 1985; Li et al., 2003; Simonsen et al., 2014; Starich et al., 1996). Further work will need to be conducted to assess the accuracy of the gap junction omissions in WW.

### 3.4.3 Validation & future work

In the absence of additional data, statements can only be made about the relative differences between the AC and WW networks; no statement can be made about the absolute accuracy of either network. This is an issue for any mapping method, manual or automated; however, there are approaches that can be used to improve mapping confidence.

With the development of different reconstruction methods (including non-image-based approaches, such as the RNA barcoding of neurons e.g. Zador et al., 2012; Kebschull et al., 2016), it is foreseeable that a more accurate network could be gener-

ated by combining different systems to build a *consensus network*; for example, using particle filtering or ensemble averaging techniques to minimise errors. This would not completely exclude systematic biases that might exist in the underlying data acquisition or tissue preparation, but it would provide a viable technological approach to increase connectome accuracy without the need for manual validation. Averaged composite networks are already used in the analysis of macroscopic MRI-derived connectomes (Van Essen et al., 2013).

The ultimate validation of any worm connectome will likely only come through the functional testing of synapses. Technologies to non-invasively image and stimulate neurons in freely behaving animals (Leifer et al., 2011; Nguyen et al., 2016; Prevedel et al., 2014; Shipley et al., 2014), coupled with automated worm handling and high-throughput phenotyping (Geng et al., 2003; Husson, 2012; San-Miguel and Lu, 2013; Schwarz et al., 2015; Swierczek et al., 2011; Yemini et al., 2013), provide a feasible route to conduct the systematic characterisation of each individual neuron and synapse in the worm, as well as their links to behaviour. Such an approach, applied to the whole nervous system, will likely only be possible in *C. elegans* due to both its small size, and transparent cuticle that allows for non-invasive imaging and neural stimulation.

Systematic functional validation would not only identify which synapses are true within the connectome, it would also point to missing connections and novel extrasynaptic interactions. Any neural interactions above a certain threshold would be observable either through correlated changes in neural activity, or alterations in behaviour that cannot be explained from existing connections. In support of this statement, pharyngeal pumping experiments have previously identified missing synapses within the worm connectome (Bhatla et al., 2015). This approach would also have the benefit of revealing the inhibitory or excitatory nature of synapses from the sign of the correlations, a property which cannot be determined from the structure of the connectome alone. Molecular approaches have been proposed for inferring the functional properties of neurons, including cell-specific sequencing and *in situ* labelling (Ekstrand et al., 2014; Hawrylycz et al., 2012; Lindner, 2014); however, such profiles are not sufficient to determine the true functional characteristics. For example, certain channels can be both inhibitory and excitatory, in a context dependent manner (Farrant and Kaila, 2007; Melzer et al., 2005); with the same also being true for gap junctions (Pereda et al., 2013; Rabinowitch et al., 2013; Volman et al., 2011).

If systematic functional validation were to be performed on the worm connectome, it could provide a dataset against which to assess the accuracy of automated neural reconstruction methods in other species, provided the effects of different neuron morphologies are sufficiently accounted for.

### 3.4.4 Conclusions

This chapter has shown that the AC and WW networks, respectively derived from manual and computer-assisted tracing methods, have largely comparable statistical properties. Taken together, both maps provide strong evidence that the true *C. elegans* connectome is a small-world network with a heavy-tailed degree distribution, containing a core rich-club of controlling interneurons, and having a structure that supports functional segregation through the presence of clustered modules.

Regardless of whether the new and omitted connections in the WW network are accurate, the observation that the majority of topological features do not change suggests that the Elegance reconstruction method accurately captures the large-scale network structure. Of the differences observed between the AC and WW networks, the WW connectome shows a greater assortativity coefficient than previously reported; however, both networks are still largely non-assortative. The increase in reciprocity seen in the WW network also supports the existence of a bias for reciprocal connectivity, and highlights the importance of local feedback in neural systems.

Generally, both the AC and WW networks share the same core of central neurons (as described by the rich-club). The new additions in WW suggest a more important role for tail interneurons in the connectome, and point to HSN as a central hub connecting the synaptic and aminergic signalling systems. This work has also demonstrated that both the AC and WW connectomes are highly robust to structural failure, though less than expected. This increased fragility to circuit-degradation is likely the result of a trade-off in favour of modular organisation.

The extent to which the existing worm network literature applies to the new network depends on the specific circuits and the level of description. It is to be expected that any global network findings should still hold, as the main structural trends remain unchanged; however, individual synapses and subnetworks would have to be considered on an individual basis due to the large number of small-scale discrepancies that exist.

Xu et al. (2013) estimate that the Elegance method has increased neuron tracing speed by at least an order of magnitude compared to previous methods. This study has shown that it can also reproduce the same global properties found in the manually traced connectome. Although computer-assisted tracing methods are not yet fast enough to map the human brain, the Elegance system has facilitated the mapping of the *C. elegans* male connectome for the first time, and provided a second reference connectome for the hermaphrodite. While differences in the connectomes suggest a need for caution in interpreting their small-scale connectivity, ongoing work to improve EM imaging and neural reconstruction (e.g. Scheffer et al., 2013; Nunez-Iglesias et al., 2014; Mikula and Denk, 2015; Kasthuri et al., 2015) will foreseeably provide more accurate connectomic mapping methods in the future, and eventually lead to technologies capable of reconstructing the microscale connectome of the human brain.

## Part III

### The extrasynaptic connectome



# Chapter 4

## The monoamine networks of *C. elegans*

### 4.1 Introduction

AS OUTLINED in the previous chapters, connectomics has primarily focused on mapping the physical, synaptic, links between neurons. It is, however, well established that chemical synapses are only one of several modes of interaction between neurons. For example, gap junctions, which mediate fast, potentially bidirectional electrical coupling between cells, are widespread in all nervous systems. Likewise, volume transmission and neurohumoral signalling provide means for local or long-range communication between neurons unconnected by synapses. As neuromodulators released through these routes can have profound effects on neural activity and behaviour (Bargmann, 2012; Brezina, 2010; Marder, 2012), a full understanding of neural connectivity requires a detailed mapping of these extrasynaptic pathways.

In *C. elegans*, as in many animals, one important route of neuromodulation is through monoamine signalling. Monoamines are widespread throughout phyla, with evidence that they are one of the oldest signalling systems, evolving at least 1 billion years ago (Roshchina, 2010; Walker et al., 1996). In both humans and *C. elegans*, many neurons expressing aminergic receptors are not post-synaptic to releasing neurons, indicating that monoamine signalling occurs primarily outside the wired connectome (Chase and Koelle, 2007). Monoamines are known to be essential for normal brain function, with abnormal signalling being implicated in numerous neurological and psychiatric conditions including depression, schizophrenia, addiction, Obsessive Compulsive Disorder (OCD), Attention Deficit Hyperactivity Disorder (ADHD), chronic

pain, and Parkinson's Disease (PD) (see Lin et al., 2011). In *C. elegans*, these systems play similarly diverse roles in regulating locomotion, reproduction, feeding states, sensory adaptation, and learning (Chase and Koelle, 2007). Clearly, if the goal of connectomics is to understand communication within the brain, extrasynaptic monoamine interactions must also be mapped, not just the network of wired chemical synapses and gap junctions.

### 4.1.1 Chapter overview

This chapter describes the process used to generate a map of extrasynaptic monoamine signalling in *C. elegans*, based on new and published gene expression data, and finds that the resulting extrasynaptic network exhibits distinct topological properties, including rich-club organisation with interconnected hubs different from those of the synaptic and gap junction networks. Despite the low degree of overlap between the monoaminergic and synaptic networks, highly significant multilink motifs of interaction are identified, pinpointing locations in the network where aminergic signalling is likely to modulate synaptic activity.

This chapter shows that the neuronal connectome can be modelled as a multiplex network with synaptic, gap junction, and neuromodulatory layers representing inter-neuronal interactions with different dynamics and polarity, and provides a prototype for understanding how extrasynaptic signalling can be integrated into a functional connectome, as well as providing a novel dataset for the development of multilayer network algorithms.

## 4.2 Materials & Methods

### 4.2.1 Synaptic & gap junction networks

The physical synaptic and gap junction networks analysed in this chapter are from the Albertson-Chklovskii (AC) hermaphrodite connectome described in Chapter 3. The decision to use the AC network over WormWiring (WW) was based on the identification of missing links in WW, reported in the previous chapter.

As before, the physical network was composed from the somatic connectome of White et al. (1986), updated and released by the Chklovskii lab (Chen, 2007; Chen et al., 2006; Varshney et al., 2011); and the pharyngeal network of Albertson and

Thomson (1976), made available by the Cybernetic *Caenorhabditis elegans* Program (CCeP)<sup>1</sup> (Oshio et al., 2003).

### 4.2.2 Monoamine network construction

To map the aminergic signalling networks of *C. elegans*, a literature search was first performed to identify genes known to be receptors, transporters or synthetic enzymes of monoamines. These were divided into classes based on the primary reported ligand, providing a list of genes which could be used to identify neurons involved in the serotonin (5-HT; Table 4.1), dopamine (DA; Table 4.2), tyramine (TA; Table 4.4), and octopamine (OA; Table 4.3) signalling systems. Genes which could not be definitively assigned to a single pathway were excluded (e.g. *cat-1*, *bas-1*, etc.). The synthesis pathways of these molecules are shown in Figure 4.1.

A further search was performed to collect cell-level expression data for the monoamine associated genes identified in the previous step. Expression patterns were collected primarily from promoter::GFP reporter lines for the appropriate biosynthetic enzymes and vesicular transporters, and verified against immunostaining and formaldehyde-induced fluorescence (FIF) data (Horvitz et al., 1982; Lints and Emmons, 1999; Rivard et al., 2010; Sulston et al., 1975) identifying monoamine-containing neurons, where available (expression patterns and primary sources are listed in Appendix A, page 223). This search was assisted with the curated expression databases of WormBase (Version: WS248)<sup>2</sup> (Howe et al., 2016) and WormWeb (Version date: 2014-11-16)<sup>3</sup> (Bhatla, 2014). The expression patterns for each of five serotonin receptors (*ser-1*, *ser-4*, *ser-5*, *ser-7* and *mod-1*), three octopamine receptors (*octr-1*, *ser-3* and *ser-6*), four tyramine receptors (*ser-2*, *tyra-2*, *tyra-3* and *lgc-55*), and four dopamine receptors (*dop-1*, *dop-2*, *dop-3* and *dop-4*) were compiled from published data (Appendix A). Since these receptors are either ion channels or serpentine receptors predicted to couple to pan-neuronal G-proteins, it was assumed that all neurons expressing monoamine receptors are potential monoamine-responding cells.

Three additional genes encode known or candidate monoamine receptors, but have missing or incomplete expression data (*dop-5*, *dop-6*, and *lgc-53*). Specifically, a ligand-gated chloride channel, *lgc-53*, has been shown to be activated by dopamine

<sup>1</sup><http://ims.dse.ibaraki.ac.jp/cccep/>

<sup>2</sup><http://www.wormbase.org/>

<sup>3</sup><http://wormweb.org/neuralnet/>

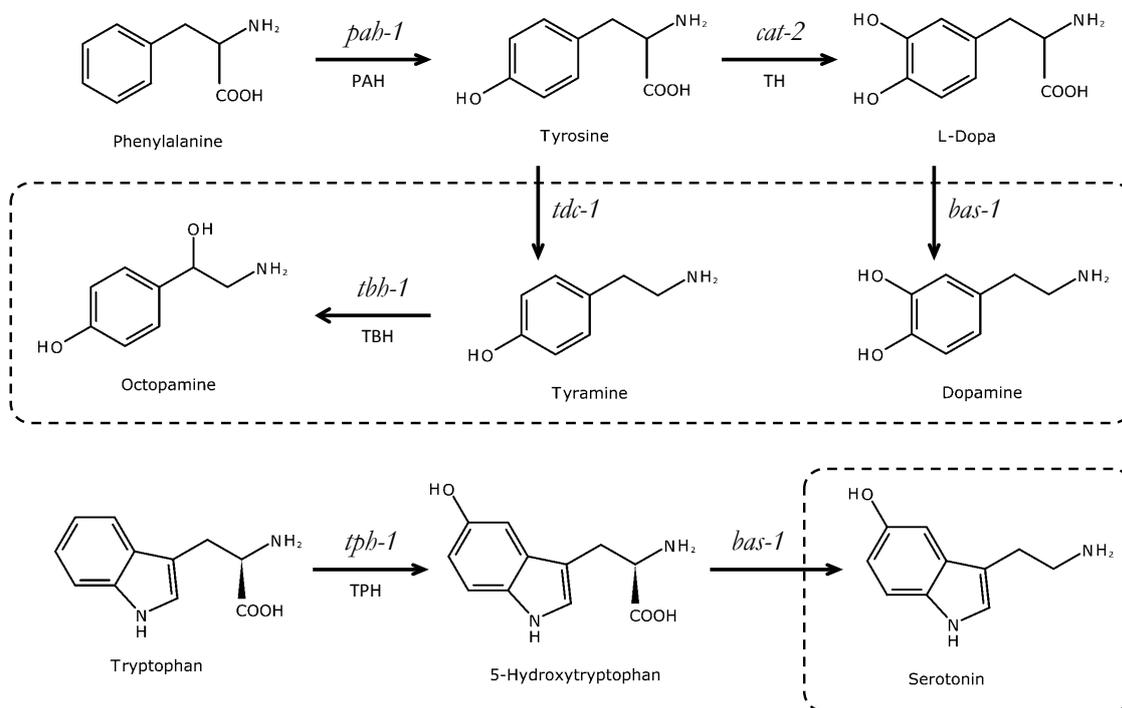


Fig. 4.1 Monoamine synthesis pathways in *C. elegans*. Known signalling molecules are indicated by dashed boxes. Based on Chase and Koelle (2007), Loer (2010) & Hobert (2013).

(Ringstad et al., 2009), but its expression pattern and biological function have not been characterised. Additional expression profiling using a transgenic *lgc-53* reporter line crossed to a series of known reference strains was performed by colleagues (detailed in Appendix C, page 233), which indicated that *lgc-53* is expressed in a small subset of neurons in the head, body, and tail. Together with the published *dop-1*, *dop-2*, *dop-3* and *dop-4*-expressing cells, these were inferred to make up the main domain of dopamine-responding neurons (analysed in this chapter).

In addition, two G-protein coupled receptors, *dop-5* and *dop-6*, have been hypothesised based on sequence homology to *dop-2* and *dop-3* to be dopamine receptors. Using the same approach used for *lgc-53*, expression patterns were collected by colleagues identifying most of the cells with expression of *dop-5* and *dop-6* reporters (Appendix Figure C.1, page 236). As *dop-5* and *dop-6* have not yet been definitively established as functional dopamine receptors, these cells were included in a secondary provisional dopamine network, the analysis of which is presented in Appendix G (page 257).

Using the cell identities derived from gene expression data, a directed graph was constructed with edges linking putative monoamine releasing cells (expressing monoamines, biosynthetic enzymes, or transporters) to those cells expressing a paired receptor. Since biologically-relevant long-distance signalling (e.g. from releasing cells in the head to tail motor neurons) has been experimentally demonstrated in *C. elegans* for both dopamine and serotonin (Chase and Koelle, 2007; Gürel et al., 2012) – while tyramine and octopamine are each released from a single neuronal class (Chase and Koelle, 2007) – edges were not restricted based on the physical distance between nodes. For the serotonin network, only those neurons with strong, consistent expression of serotonin markers such as tryptophan hydroxylase were included (NSM, HSN and ADF, Table A.1). Additional neurons (AIM, RIH, VC4/5) that appear to take up serotonin but not synthesise it (Duerr et al., 1999; Jafari et al., 2011) were not included in the network, since they may function primarily in the homeostatic clearing of serotonin. The ASG neurons were also excluded, which produce serotonin only under hypoxic conditions (Pocock and Hobert, 2010), though they are likely to participate conditionally in the serotonin signalling networks. As tyramine  $\beta$ -hydroxylase (*tbh-1*) converts tyramine to octopamine, neurons expressing both tyrosine decarboxylase (*tdc-1*) and tyramine  $\beta$ -hydroxylase (*tbh-1*) were limited to the octopamine network (RIC, see Table A.3).

For caveats on the use of GFP reporter lines to infer gene expression, as well as a discussion of potential measurement and integration issues, see Appendix D (page 237).

#### 4.2.2.1 Data model

A data model showing the relations and constraints used to construct the network is shown in Figure 4.2.

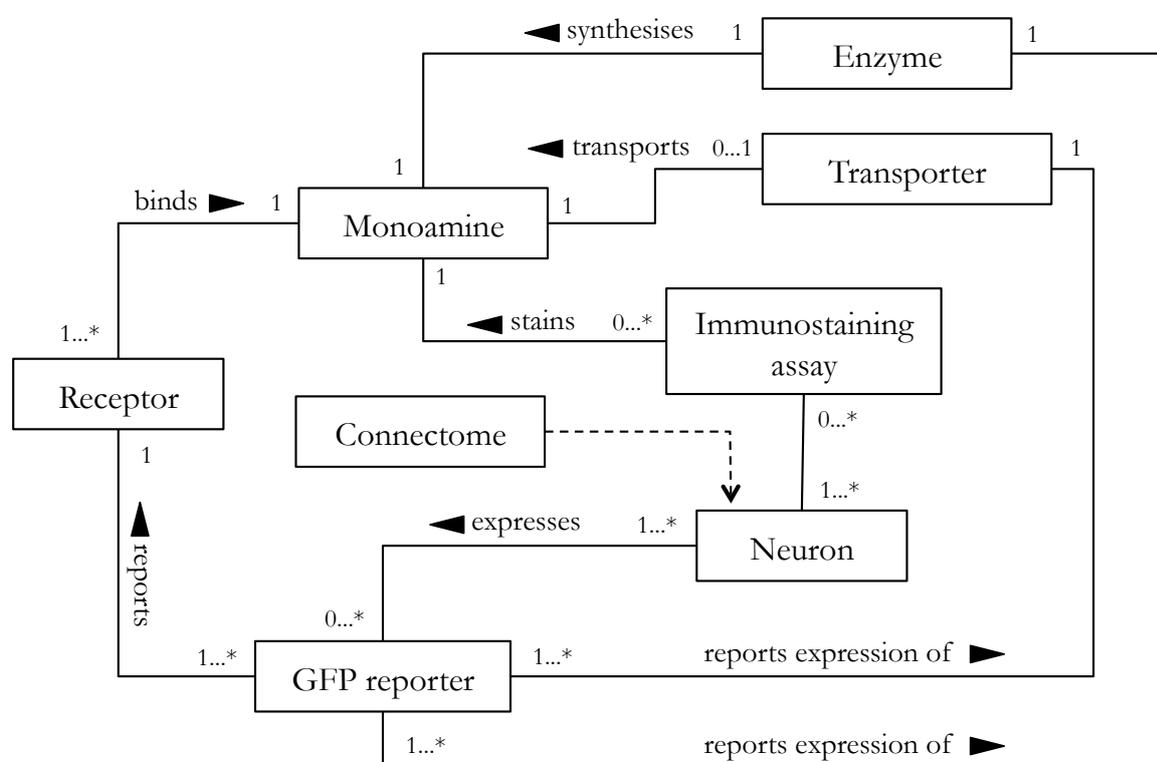


Fig. 4.2 Monoamine data model. UML class diagram (ISO, 2012a,b) of the conceptual data model used to construct the extrasynaptic monoamine network. Each box represents a class of data, with lines depicting their associations. The numbers represent the multiplicity of the constraints assumed in the modelling (e.g. a **Transporter** transports exactly 1 **Monoamine**; while a **Monoamine** can be transported by 0 or 1 **Transporters**). Data which did not match these constraints were excluded (e.g. the *cat-1* **Transporter** transports >1 **Monoamine**, and was therefore not included). Triangles denote the directionality of the relationships, and \* represents an unbounded (i.e.  $\infty$ ) multiplicity.

### 4.2.3 Network data

Edge lists for individual network layers are provided in the supporting information (S1 dataset) of Bentley et al. (2016), and can be accessed online from PLOS<sup>4</sup> or PubMed Central<sup>5</sup>.

### 4.2.4 Topological network measures

All network measures are the same as those described in Chapter 2. Edge counts, adjacency matrices, and reducibility clusters were all computed using binary directed versions of the networks. The same networks, excluding self-connections ( $A_{ij} = 0$  if  $i = j$ ), were used to compute the multilink motifs and reciprocity.

Network measures are compared to 100 null model networks (shown in the box-plots) generated using the degree-preserving randomisation procedure described in Chapter 2 (Algorithm 5), where the networks were randomised by performing  $10 \times M$  edge-swaps. The same null model networks were also used to compute the multilink motif  $z$ -scores, with each layer being randomised independently. As per the definition given in Humphries and Gurney (2008), the small-world index was normalised against 100 Erdős-Rényi (ER) reference networks containing the same number of nodes and edges as the actual networks.

To identify neurons with high-participation in all of the network layers, the normalised degree-rank product was used. This is computed by ranking neurons in each network layer by their degree in descending order (such that the highest degree neuron has a rank of 302), and rescaling to the interval  $[0, 1]$ . The product is then taken of the ranked degrees in each layer. Thus, a neuron that is the highest degree node in all of the network layers has a degree product of 1.

### 4.2.5 Software

Analyses and data handling were performed using the methods described in § 2.6. URL links to the software tools are provided in Appendix H.

---

<sup>4</sup><http://dx.doi.org/10.1371/journal.pcbi.1005283.s004>

<sup>5</sup><https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5215746/bin/pcb.1005283.s004.zip>

Type	Gene	Sequence	Note	Reference
Enzyme	<i>tph-1</i> (TPH2)	ZK1290.2	Tryptophan hydroxylase Catalyses the first step in serotonin biosynthesis	Sze et al. (2000)
Transporter	<i>mod-5</i> (SLC6A4)	Y54E10BR.7	Na <sup>+</sup> /Cl <sup>-</sup> dependent 5-HT transporter Required for 5-HT uptake	Ranganathan et al. (2001)
Receptor	<i>ser-1</i> (5-HT <sub>2A</sub> )	F59C12.2	7TM GPCR (G <sub>αq</sub> ) Coupled to Ca <sup>2+</sup> signalling	Hamdan et al. (1999)
Receptor	<i>ser-4</i> (5-HT <sub>1B</sub> )	Y22D7AR.13	7TM GPCR (G <sub>αi/o</sub> ) Attenuates adenylyl cyclase activity	Olde and McCombie (1997)
Receptor	<i>ser-5</i> (5-HT <sub>6</sub> )	F16D3.7	7TM GPCR (stimulatory) Independent of adenylyl cyclase activity	Hapiak et al. (2009)
Receptor	<i>ser-7</i> (5-HT <sub>7</sub> )	C09B7.1	7TM GPCR (G <sub>αs</sub> ) Increases adenylyl cyclase activity	Hobson et al. (2006)
Receptor	<i>mod-1</i> (GABA <sub>A</sub> )	K06C4.6	5-HT-gated ion channel Cl <sup>-</sup> selective	Ranganathan et al. (2000)

Table 4.1 Serotonin (5-HT) biosynthetic enzyme, transporter & receptor genes used to construct the extrasynaptic 5-HT network. Human orthologues are shown in parentheses below the gene names.

Type	Gene	Sequence	Note	Reference
Enzyme	<i>cat-2</i> (TH)	B0432.5	Tyrosine hydroxylase Rate limiting enzyme in DA biosynthesis	Lints and Emmons (1999)
Transporter	<i>dat-1</i> (SLC6A2)	T23G5.5	Na <sup>+</sup> /Cl <sup>-</sup> dependent DA transporter High DA affinity	Jayanthi et al. (1998)
Receptor	<i>dop-1</i> (DRD1)	F15A8.5	7TM GPCR (G <sub>αq</sub> ) DA receptor	Suo et al. (2002)
Receptor	<i>dop-2</i> (DRD2)	K09G1.4	7TM GPCR (G <sub>αi/o</sub> ) DA receptor	Tsalik et al. (2003)
Receptor	<i>dop-3</i> (DRD2)	T14E8.3	7TM GPCR (G <sub>αi/o</sub> ) Attenuates cAMP formation in response to DA	Chase et al. (2004)
Receptor	<i>dop-4</i> (DRD1)	C52B11.3	7TM GPCR (G <sub>αs</sub> ) Stimulates cAMP accumulation in response to DA	Sugiura et al. (2005)
Receptor	<i>dop-5</i> * (DRD1)	T02E9.3	7TM GPCR Required for normal pharyngeal pumping	Carre-Pierrat et al. (2006)
Receptor	<i>dop-6</i> * (D <sub>2</sub> /D <sub>3</sub> )	C24A8.1	7TM GPCR (G <sub>αi/o?</sub> ) Predicted from paralogy with <i>dop-2</i> & <i>dop-3</i>	Keating et al. (2003)
Receptor	<i>lgc-53</i> (GABA <sub>A</sub> )	T21F2.1	DA-gated ion channel Cl <sup>-</sup> selective	Ringstad et al. (2009)

Table 4.2 Dopamine (DA) biosynthetic enzyme, transporter & receptor genes used to construct the extrasynaptic DA network. Human orthologues are shown in parentheses below the gene names. \* *dop-5* and *dop-6* have not been definitively established as dopamine receptors, and are analysed separately as part of a broader dopamine network in Appendix G (page 257).

Type	Gene	Sequence	Note	Reference
Enzyme	<i>tbh-1</i> (DBH)	H13N06.6	Tyramine $\beta$ -hydroxylase Required for biosynthesis of OA from TA	Alkema et al. (2005)
Receptor	<i>octr-1</i> (ADRA2A)	F14D12.6	7TM GPCR ( $G_{\alpha i/o}$ ) Acts in ASH to inhibit 5-HT stimulation	Harris et al. (2010)
Receptor	<i>ser-3</i> (ADRA1A)	K02F2.6	7TM GPCR ( $G_{\alpha q}$ ) In oocytes SER-3 + OA elicits inward current via $Ca^{2+}$ -gated $Cl^-$ channels	Mills et al. (2012)
Receptor	<i>ser-6</i> (DRD2)	Y54G2A.35	7TM GPCR ( $G_{\alpha s}$ ) Stimulates the release of multiple peptides	Mills et al. (2012)

Table 4.3 Octopamine (OA) biosynthetic enzyme & receptor genes used to construct the extrasynaptic OA network. Human orthologues are shown in parentheses below the gene names.

Type	Gene	Sequence	Note	Reference
Enzyme	<i>tdc-1</i> (DDC)	K01C8.3	Tyrosine decarboxylase Required for biosynthesis of tyramine	Alkema et al. (2005)
Receptor	<i>ser-2</i> (5-HT <sub>1A</sub> )	C02D4.2	7TM GPCR (G <sub>αi/o</sub> ) Inhibits forskolin-stimulated cAMP levels	Rex et al. (2004)
Receptor	<i>tyra-2</i> (5-HT <sub>5A</sub> )	F01E11.5	7TM GPCR (G <sub>αi/o</sub> ) TA receptor	Rex et al. (2005)
Receptor	<i>tyra-3</i> (ADRB1)	M03F4.3	7TM GPCR (G <sub>αq</sub> ) Inhibits 5-HT responses. Independent of adenylyl cyclase	Hapiak et al. (2013)
Receptor	<i>lgc-55</i> (GABA <sub>A</sub> )	Y113G7A.5	TA-gated ion channel Cl <sup>-</sup> selective	Pirri et al. (2009)

Table 4.4 Tyramine (TA) biosynthetic enzyme & receptor genes used to construct the extrasynaptic TA network. Human orthologues are shown in parentheses below the gene names.

## 4.3 Results

### 4.3.1 Network structure

Examining the expression patterns for each of the nineteen monoamine receptors (Tables 4.1–4.4) reveals that the monoamine networks consist of only a few central neurons that broadcast signals to a large number of peripheral neurons. In total, 18 of the 302 neurons in the adult hermaphrodite were found to release monoamines, while 251 neurons (83 %) were found to express one or more monoamine receptors (Table 4.5). This gives the network a star-like topology, which can be directly observed in all of the separate monoamine layers (Figure 4.3). The adjacency matrix further reveals that these monoamine-releasing cells are mostly sensory and motor neurons, with the downstream receptors being distributed throughout the worm (Figure 4.4). As a consequence, the monoamine network exhibits a heavy-tailed degree distribution containing a small number of high-degree hubs (Figure 4.12).

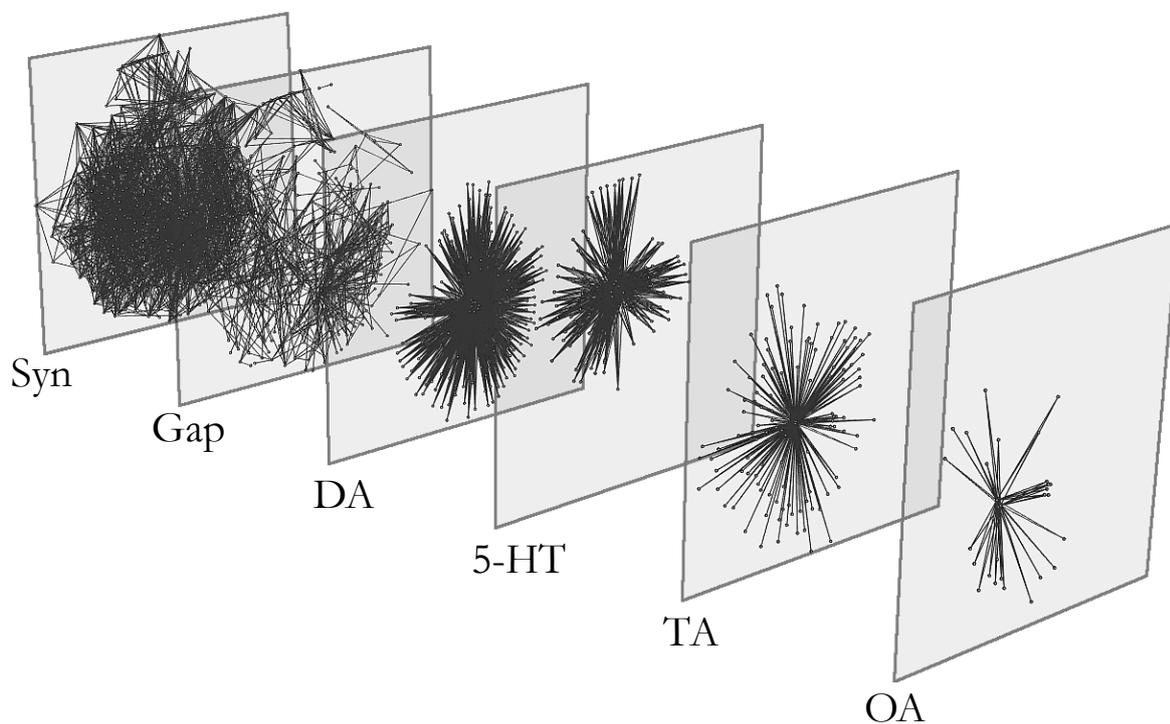
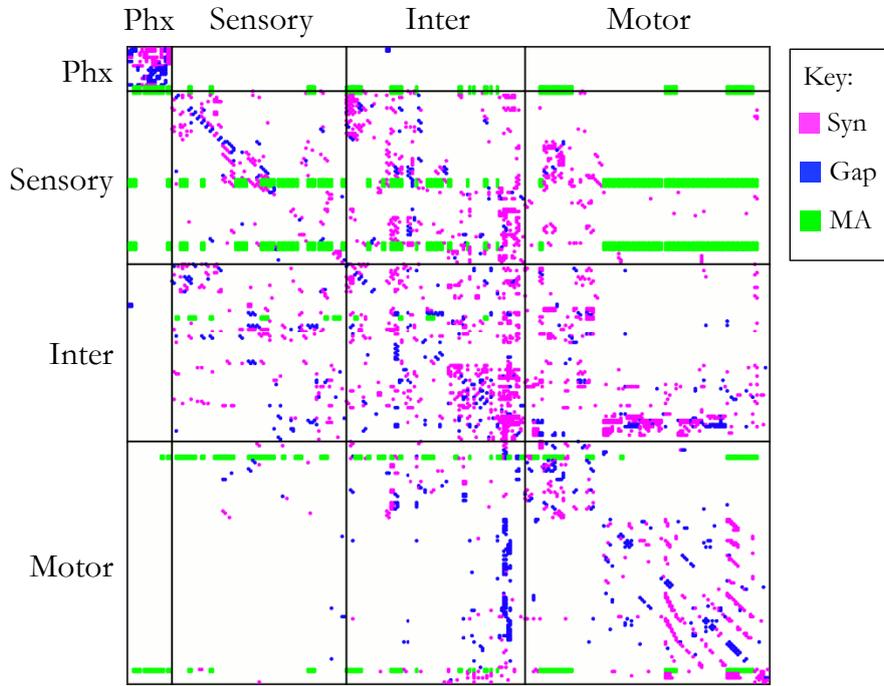


Fig. 4.3 Multilayer projection of the synaptic, gap junction, and monoamine signalling networks (DA, 5-HT, TA & OA), each represented as a separate layer. Node positions are the same in all layers.



Synaptic, gap junction &amp; monoamine networks

Fig. 4.4 Adjacency matrix showing the wired synaptic (magenta), gap junction (blue), and extrasynaptic monoamine networks (green). Nodes are classified as pharyngeal (Phx), sensory, motor, or interneurons.

Network	Nodes $N$	N <sup>o</sup> ligand expressing	N <sup>o</sup> receptor expressing	Edges $M \rightarrow$
Serotonin	86	6	82	490
Dopamine	147 (187)	8	147 (187)	1168 (1488)
Octopamine	28	2	28	54
Tyramine	116	2	114	228
<i>Aggregate</i>	237 (251)	18	235 (251)	1940 (2260)

Table 4.5 Monoamine networks. Table showing the number of nodes and edges in the individual and aggregate monoamine networks. Values for an expanded network including putative *dop-5* and *dop-6* connections are in parentheses.

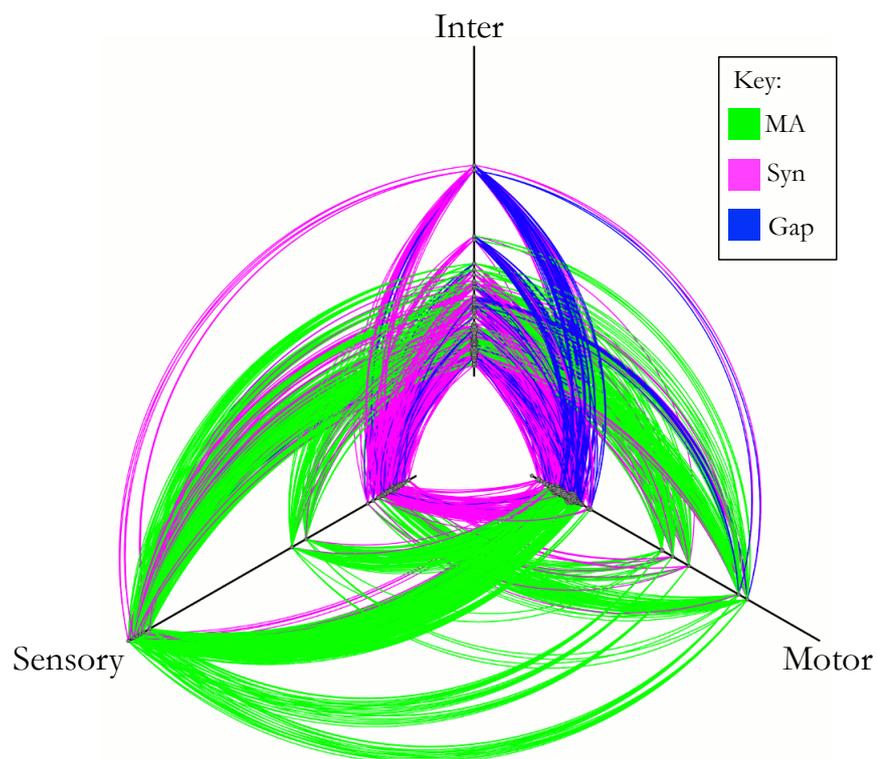


Fig. 4.5 Hive plot showing the wired synaptic (magenta), gap junction (blue), and extrasynaptic monoamine connections (green). Nodes are classified as sensory, motor, or interneurons and are arranged along the three axes according to their degree. Hubs are located further out along the axes. Plot generated using *hiveplotter*, courtesy of Barnes (2016).

### 4.3.2 Extrasynaptic signalling

To investigate the extent of extrasynaptic signalling in the *C. elegans* monoamine systems, the expression patterns of monoamine receptors were systematically compared with the postsynaptic targets of aminergic neurons. Analysis of these expression patterns suggests that a remarkably high fraction of monoamine signalling must be extrasynaptic. For example, the two tyraminergetic neurons, RIML and RIMR, are presynaptic to a total of 20 neurons. Yet of the 114 neurons that express reporters for one or more of the four tyramine (TA) receptors, only 7 are postsynaptic to a tyraminergetic neuron (Figure 4.6). Thus, approximately 95 % of tyramine-responsive neurons must respond only to extrasynaptic TA. Similar analyses of the other monoamine systems yield comparable results: 100 % of neurons expressing octopamine receptors receive no synaptic input from octopamine-releasing neurons (Figure 4.7), while 82 % of neurons expressing dopamine receptors, and 76 % of neurons expressing serotonin receptors receive no synaptic input from neurons expressing the cognate monoamine ligand (see Table 4.6 for numbers including releasing cells). Thus, most neuronal monoamine signalling in *C. elegans* appears to occur extrasynaptically, outside the wired physical connectome. The prevalence of extrasynaptic monoamine signalling between neurons unconnected by synapses or gap junctions implies the existence of a large *wireless* component to the functional *C. elegans* connectome, the properties of which have not previously been studied.

Network	Cells with no synaptic input		Non-synaptic edges	
	N <sup>e</sup>	%	N <sup>e</sup>	%
Serotonin	62	75.6	457	93.3
Dopamine	121 (138)	82.3 (73.8)	1133 (1422)	97 (95.6)
Octopamine	28	100	54	100
Tyramine	107	93.9	216	94.7
<i>Aggregate</i>	183 (178)	77.9 (70.9)	1860 (2149)	95.9 (95.1)

Table 4.6 The number of monoamine receptor-expressing cells that do not receive synapses from releasing cells, and the number of connections in each layer that are non-synaptic, including connections between neurons within the same class. Due to a many-to-many relationship between senders and receivers, the fraction of non-synaptic edges can exceed the fraction of non-synaptic cells. Values for the expanded network including putative *dop-5* and *dop-6* connections are in parentheses.



### 4.3.3 The multilayer connectome

With the inclusion of the separate monoamine systems, the full *C. elegans* connectome can be considered as a multiplex or multilayer network (Nicosia and Latora, 2015), with each node representing a neuron and each layer of connections – synaptic, gap junction, and monoamine – characterised by distinct edge properties (Figure 4.3). For example, chemical synapses represent unidirectional wired connections that signal on a fast (ms) time scale, while gap junctions generate reciprocal electrical connections that function on an even faster time scale. In contrast, monoamine connections are wireless (with a single sending cell broadcasting to multiple receivers), slow (acting on a time scale of seconds or longer) and unidirectional (Ezcurra et al., 2011; Gürel et al., 2012).

Prior studies of multiplex networks in non-biological systems – such as communication networks – have tended to find a large degree of overlap between the links observed in distinct layers (De Domenico et al., 2015a). Similarly, the high-degree hubs in each layer are often co-located, unequivocally highlighting certain nodes as key controllers of information flow in the system (Nicosia and Latora, 2015). While the synaptic and gap junction layers are observed to follow this trend, with the same high-degree neurons in both systems (Figure 4.8), the extrasynaptic network exhibits a vastly different structure. Out of 1940 monoamine connections only 80 overlap with physical synapses, meaning 96 % of the monoamine connections are unique to the monoamine layer (Table 4.6). Furthermore, no significant degree-degree correlation is observed between the physical and extrasynaptic layers, indicating that the hubs of the monoamine system are distinct (Figure 4.8), additionally highlighted by the existence of separate rich-clubs (see § 4.3.5 below).

Reducibility analysis (De Domenico et al., 2015a), which clusters the different network layers based on their similarity, provides further support that the monoamine networks have a unique structure, with the monoamines forming a distinct cluster separate from the physical networks (Figure 4.9). This shows that all of the monoamines overlap less with the synaptic and gap junction networks than the synaptic and gap junction networks do with each other.

The analyses above suggest two distinct interpretations for the dissimilarity to the physical network layers. Firstly, monoamines may be functioning as an independent network, with little relation to the faster physical network. Secondly, the dissimilarity between layers might indicate that monoamines have a complementary function that is

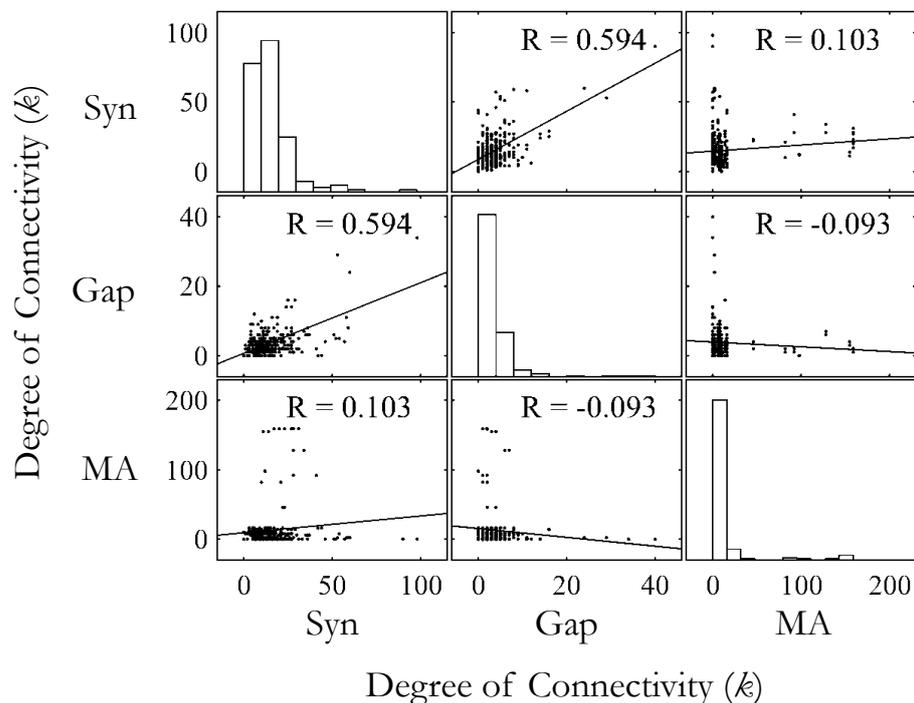


Fig. 4.8 Degree-degree correlation matrix. Off-diagonal panels show the degree-degree correlation between a pair of network layers. Panels on the diagonal show the degree distribution of the individual layers.

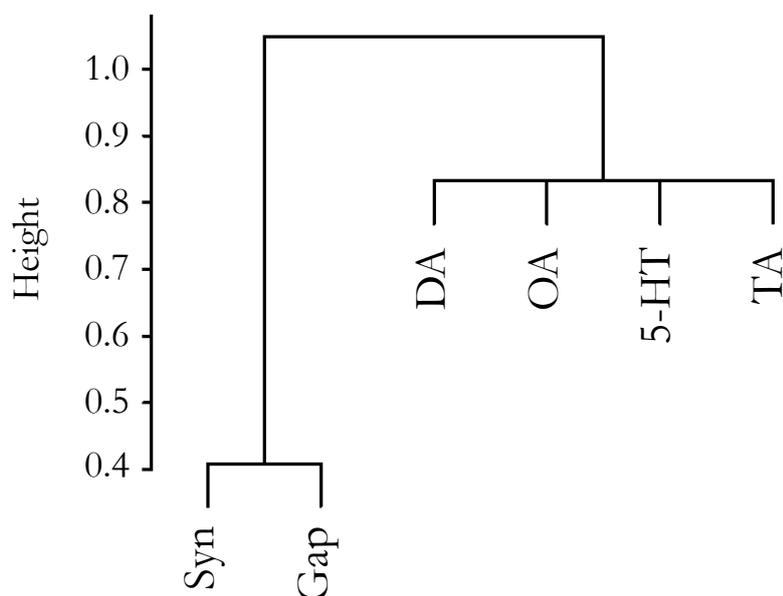


Fig. 4.9 Multilayer reducibility dendrogram. Layers close on the dendrogram have more overlapping edges and are more reducible. Branching height gives the Jensen-Shannon distance between the layers.

nevertheless coupled to that of the physical connections. To address this, the following sections investigate whether the isolated *C. elegans* monoamine network displays the structural organisation required for information processing.

#### 4.3.4 Network measures

As already mentioned, the monoamine network has a star-like architecture that is qualitatively different to the other network layers. This structure is reflected in the topological network measures, with the MA network exhibiting high disassortativity characteristic of star networks (Figure 4.10d), a low number of reciprocal monoamine connections (Figure 4.11), and a heavy-tailed degree distribution (Figure 4.12). The star-like structure of the monoamine layer was also confirmed by three-neuron motif analysis, which revealed the enrichment of a motif consisting of a hub node signalling to two spokes (Figure 4.14).

Disassortativity is known to be relevant in the organisation of collective network dynamics, such as synchronisation (Sorrentino et al., 2006) and cooperation behaviour (Perc et al., 2013; Wang et al., 2008), and is widely observed in other biological and technological networks (Newman, 2003b). The inclusion of these connections into the aggregate connectome has the effect of greatly reducing the overall path length of the network (Figure 4.10c), increasing the efficiency of integrative information processing by providing shortcuts between more segregated subgraphs of the physical wired network (Sporns, 2013b). The adjacency matrix and hiveplot shows that the monoamines provide a direct route of communication between sensory neurons and motor neurons, bypassing the premotor interneurons that play a prominent role in the synaptic and gap junction systems (Towlson et al., 2013) (Figures 4.4 & 4.5). Together, these observations suggest that the monoamines provide a highly efficient shortcut for coordinating behaviour throughout the entire organism. This is an expected consequence of the presence of highly connected hubs directly linking many disparate parts of the network, and is a useful feature given the role of monoamines in signalling physiologically important states relevant to the entire organism, such as food availability (e.g. Ezcurra et al., 2011).

The increased connectivity provided by the monoamines results in a reduction in the aggregate network's modular structure, a consequence of increasing the number of connections between functionally segregated units (Figure 4.10b). The network is, however, still more modular than random, with the monoamine layer also exhibiting

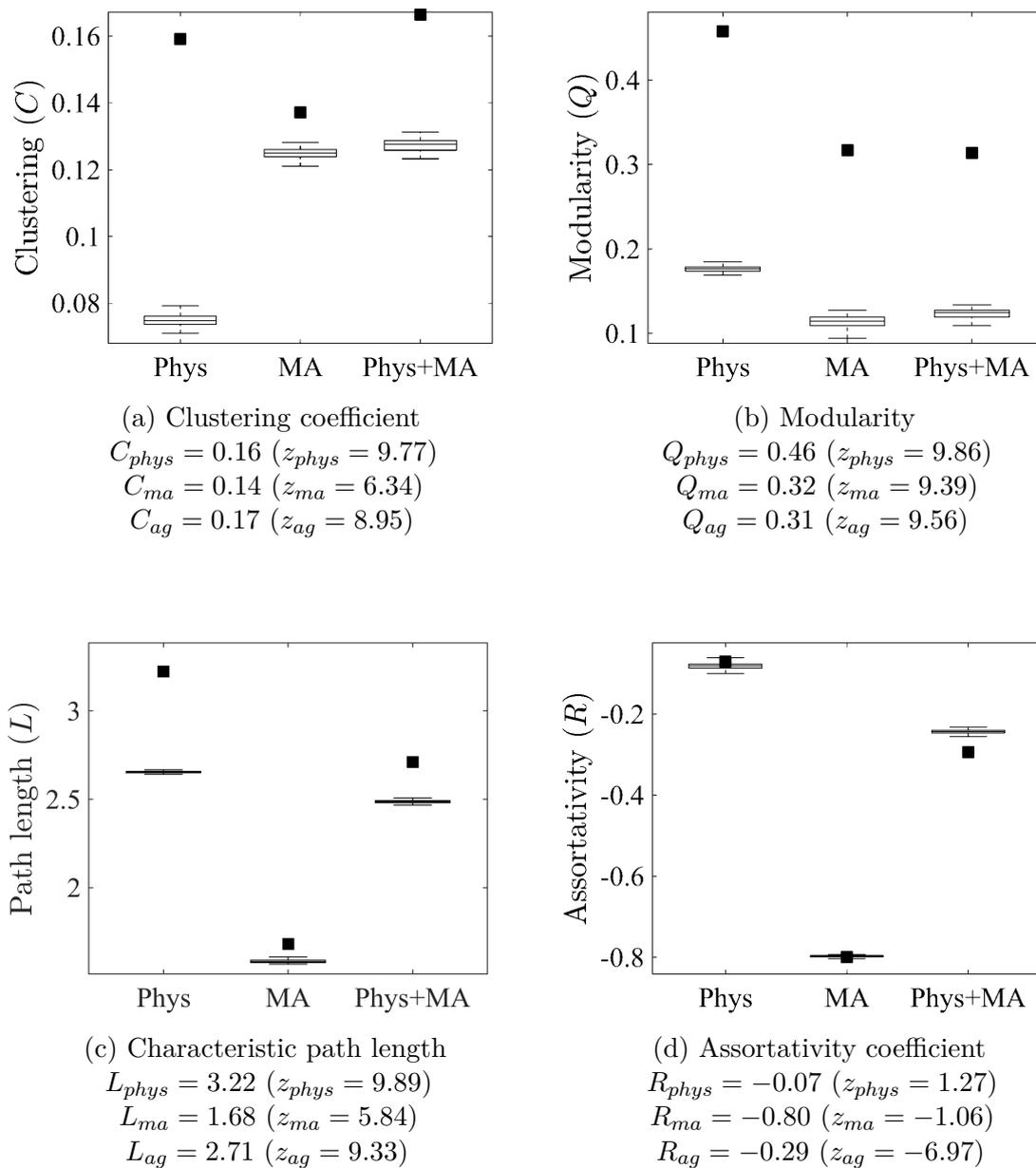


Fig. 4.10 Comparison of network metrics for the physical synaptic & gap junction network (Phys), full monoamine network (MA), and the aggregate physical & monoamine network (Phys + MA). Plots show the observed values (filled squares) and expected values from 100 randomised networks (boxplots). Network measure for individual monoamine networks are presented in Appendix G (page 257).

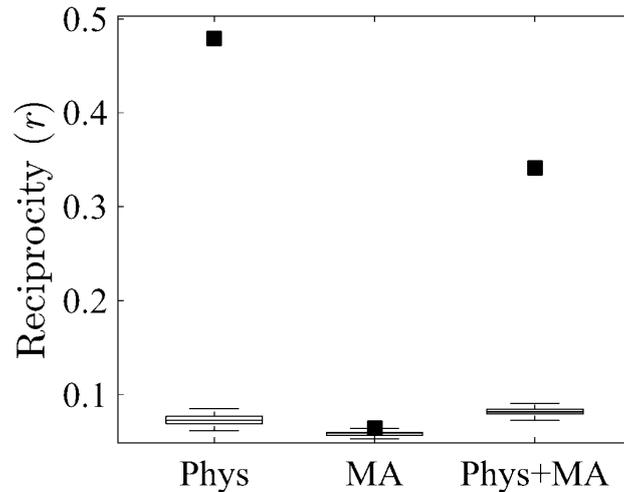


Fig. 4.11 Reciprocity for the directed synaptic, monoamine, and aggregate synaptic+monoamine networks. Plot shows the observed values (filled squares) and expected values from 100 randomised networks (boxplots).

greater-than-random modularity compared to null models that rewire the network edges while preserving degree distribution. This is expected given the monoamine layer's composition from separate signalling systems; indeed the individual monoamine networks considered on their own show very low modularity (Appendix Figure G.4, page 262).

Despite the hub-and-spoke structure of the extrasynaptic network, the monoamine layer exhibits a significant level of global clustering (Figure 4.10a). This observation is explained by two factors. Firstly, the expression of monoamine receptors by releasing neurons creates a central cluster of hub neurons in the network (see Figure 4.16); secondly, as the same monoamines are released by more than one neuron, and many neurons also express more than one monoamine receptor, triangles are formed in the network with a receiving neuron at one vertex, and two transmitting neurons as the others. Indeed, three-neuron motif analysis confirmed that this configuration is overrepresented in all the monoamine networks save tyramine (Figure 4.14). This structure provides a method of dual lateral inhibition, where a releasing neuron can inhibit antagonistic signals from another hub neuron, while simultaneously negating the downstream effects of those signals – this pattern has previously been observed in the OA/TA and 5-HT systems between RIC/RIM & NSM in the aminergic control of feeding behaviours (Li et al., 2012), with similar patterns also existing in individual monoamine layers. For example, the ventral cord motor neurons express both excitatory (*dop-1*) and inhibitory

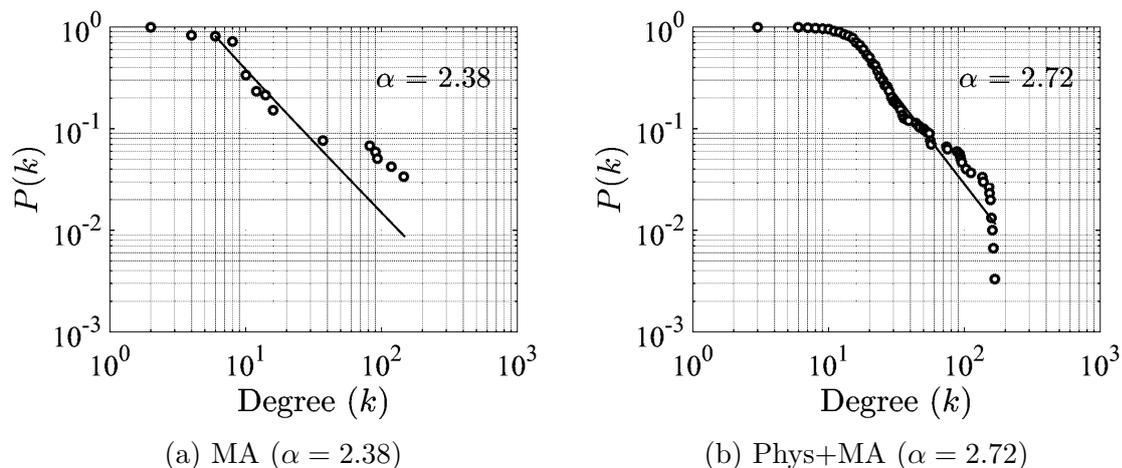


Fig. 4.12 Scale-free properties of the (a) monoamine network, and (b) aggregate synaptic, gap junction and monoamine network (showing best fit for the power-law distribution  $P(k) = k^{-\alpha}$ ).

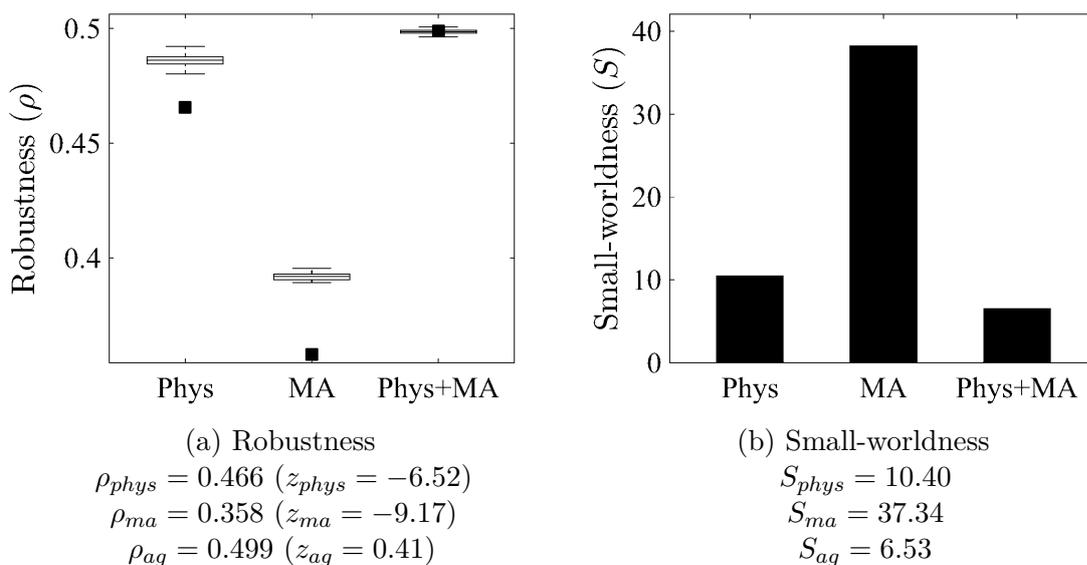


Fig. 4.13 Robustness and small-worldness of the physical, monoamine & aggregate physical+monoamine networks. Robustness plot shows observed values (filled squares) and expected values from 100 randomised networks (boxplots).

(*dop-3*) dopamine receptors, which work antagonistically to regulate locomotion (Chase et al., 2004), while the expression of the inhibitory *dop-2* in dopamine-releasing neurons suggests that the hubs mutually suppress one another to regulate dopamine release.

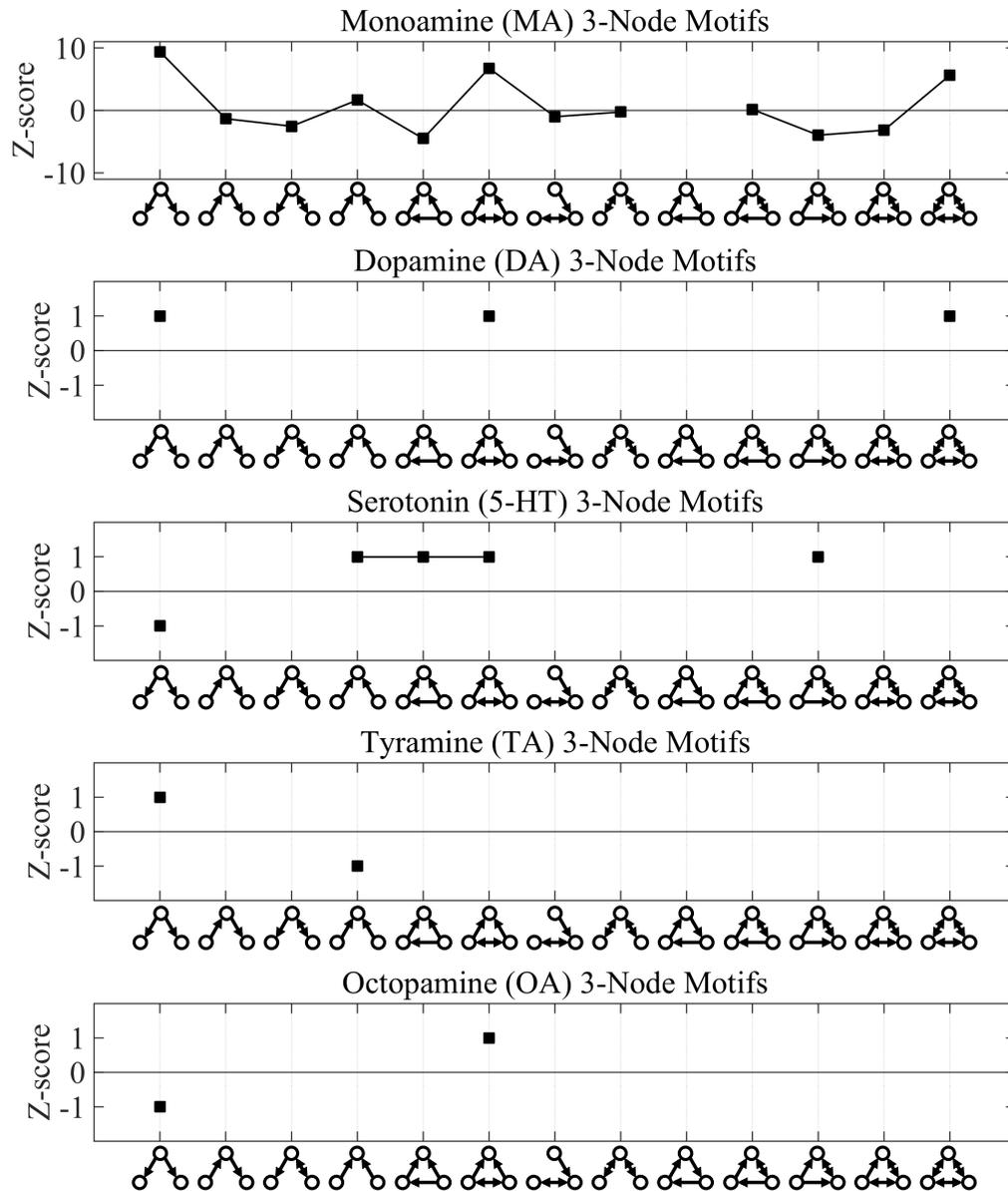


Fig. 4.14 Three-neuron motif analysis. Directed 3-node motif profiles for the aminergic networks, showing all 13 possible combinations with no unconnected nodes.  $Z$ -scores show the level of over- or under- representation for each motif, and were computed relative to a sample of 100 randomised networks generated using the degree-persevering edge-swap procedure with 10 swaps per edge. Motif enumeration was performed using the ESU algorithm (see Algorithm 7, page 69).

Both the monoamine network, and the aggregate network including monoamines and physical connections, display a heavy-tail degree distribution (Figure 4.12) and small-world properties (Figure 4.13b) common in neural systems (Bullmore and Sporns, 2009). As expected, the disassortative star topology of the MA network results in a much lower level of robustness to circuit degradation than the physical network (Figure 4.13a). In contrast, although the aggregate (Phys+MA) network also has reduced assortativity, the addition of the extra connections from the MA network result in a structure that is almost maximally robust. Although the MA network can easily be impaired by destroying the central transmitting neurons, the combination of physical and monoamine signalling is much more structurally resilient to insult than either network individually. This is a result of the non-overlapping connections combining to create a network with fewer gaps and higher overall connectivity. It is important to note, however, that while the aggregate network shows high topological robustness, the different timescales and functional properties of the connection types mean that it is unlikely the networks are any more robust in combination. The existence of an alternative signalling pathway in the network does not mean it can substitute the function of the original.

### 4.3.5 Rich-club organisation

Many neural and brain networks have been shown to exhibit rich-club organisation (Harriger et al., 2012; Heuvel and Sporns, 2011; Liang et al., 2017; Reus and Heuvel, 2013; Shih et al., 2015) in which the most highly-connected nodes are more connected to one another than expected by chance (Colizza et al., 2006). It was previously shown that the *C. elegans* physical connectome includes a rich-club consisting primarily of a small number of premotor interneurons, controlling forward and backward locomotion (see Chapter 3 & Towlson et al., 2013). Subjecting the monoamine connectome to similar analysis, it was found that this network also contains a distinct rich-club (Figure 4.15 & Table 4.8), consisting of dopamine, serotonin, and tyramine-releasing neurons. The rich-club property stems from the fact that most serotonergic neurons contain receptors for both tyramine and dopamine, while dopaminergic and tyraminergetic neurons likewise express receptors for the other two aminergic transmitters (Figure 4.16), suggesting that different monoamines coordinate their actions. This rich-club structure is also reflected in the three-neuron motif analysis, in which the fully-connected motif was overrepresented in the aggregate monoamine layer (Figure 4.14). Interestingly, in contrast to the physical wired rich-club, all of whose members are interneurons, the

monoamine rich-club consists of sensory neurons and motor neurons (Table 4.8).

Despite the distinct structures and topologies of the different neuronal connectome layers, they are likely to interact in functionally significant ways. For example, although the physical and monoamine rich-clubs do not overlap, there are significant links between them (Figure 4.17). To systematically identify neurons that have a role in linking all of the layers, neurons were ordered according to the product of their degree-rank across the synaptic, gap junction and monoamine layers (Table 5.5). The ten highest ranking neurons include three from the monoamine rich-club (RIML, RIMR, and ADEL) and three from the physical rich club (RIBL, RIBR, and DVA). Indeed, the premotor interneuron DVA is a receiver for serotonin, dopamine and tyramine signalling, while the tyramineric RIMs are highly connected to the premotor interneurons of the wired rich-club. As one might expect from their topological role in linking the monoamine and physical network layers, the RIMs have been shown in a number of studies to play a central role in the modulation of sensory pathways in response to feeding states as well as the control of downstream locomotion motor programs (Donnelly et al., 2013; Piggott et al., 2011; Wragg et al., 2007). Similarly RIB, which expresses receptors for serotonin and dopamine, is thought to integrate numerous sensory signals (Mori and Ohshima, 1995; Tsalik and Hobert, 2003) and has been demonstrated to influence reorientation in foraging behaviour (Gray et al., 2005).

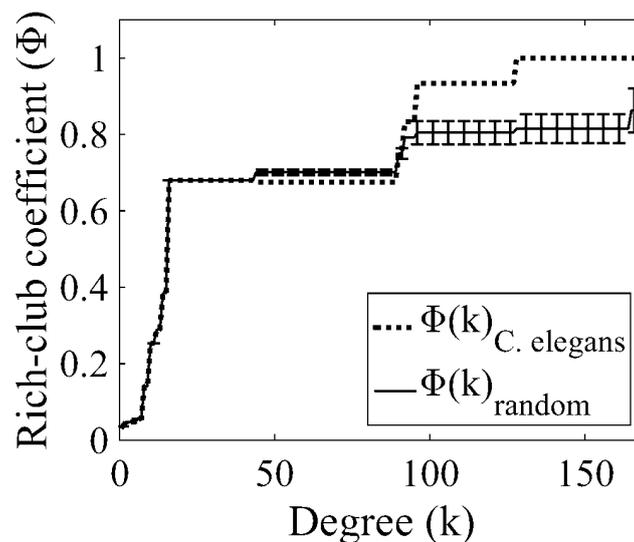


Fig. 4.15 Rich-club curve for the directed monoamine network. Dashed line indicates the rich-club coefficient for the *C. elegans* network and the solid curve represents the average rich-club coefficient of 100 randomised networks (preserving degree distribution).

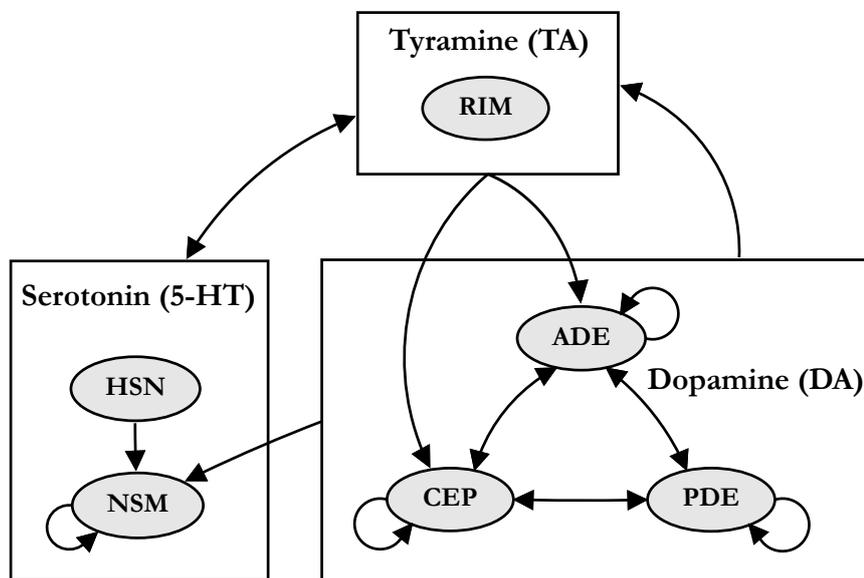


Fig. 4.16 Monoamine rich-club schematic, showing the separate aminergic systems and the volume transmission signalling between them based on receptor expression. Arrows between boxes denote connections between all of the contained neurons.

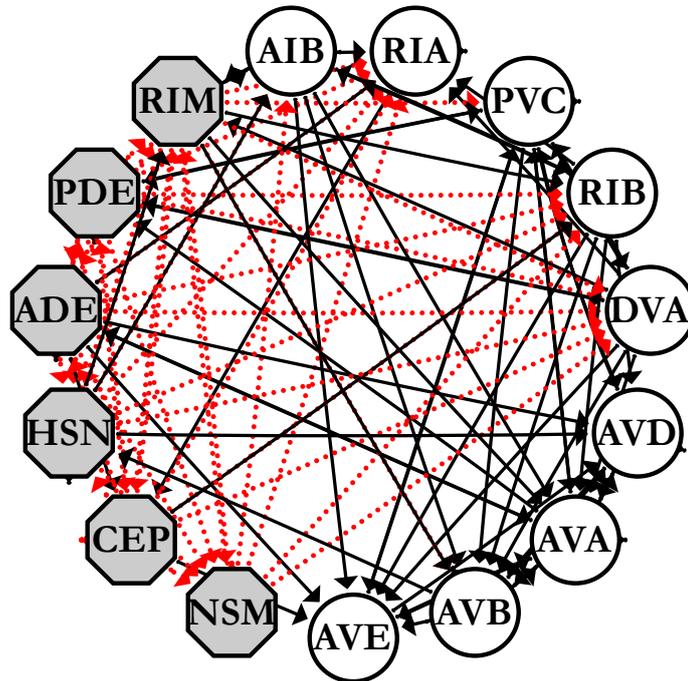


Fig. 4.17 Connections between the physical & monoamine rich-clubs. Aminergic rich-club neurons are represented as grey octagons. Members of the physical rich-club are shown as circles. Dashed red lines are extrasynaptic links. Solid black lines are physical synapses.

Neuron	$k_{norm}$	$k_{syn}$	$k_{gap}$	$k_{ma}$
RIMR*	0.236	34	14	128
RIBL*	0.207	29	30	14
RIBR*	0.178	25	30	14
RIML*	0.171	28	12	128
RIS	0.119	27	16	14
ADEL*	0.073	31	4	157
VD01	0.070	14	16	16
DVA*	0.069	54	10	8
PVQR	0.069	22	10	16
AIBR	0.066	36	16	6

Table 4.7 The normalised degree-rank product ( $k_{norm}$ ) showing the neurons that have the highest interaction across the synaptic, gap junction, and monoamine network layers. Rich-club neurons are indicated with  $\star$ .

Neuron ID	Degree $k_{ma}$	Rich-club $\Phi_{norm}$	MA	Receptors	Type
CEPDL	157	$3\sigma$	DA	<i>dop-2, octr-1, tyra-3</i>	Sensory
CEPDR	157	$3\sigma$	DA	<i>dop-2, octr-1, tyra-3</i>	Sensory
CEPVL	157	$3\sigma$	DA	<i>dop-2, octr-1, tyra-3</i>	Sensory
CEPVR	157	$3\sigma$	DA	<i>dop-2, octr-1, tyra-3</i>	Sensory
ADEL	157	$3\sigma$	DA	<i>dop-2, octr-1, tyra-3</i>	Sensory
ADER	157	$3\sigma$	DA	<i>dop-2, octr-1, tyra-3</i>	Sensory
PDEL	153	$3\sigma$	DA	<i>dop-2</i>	Sensory
PDER	153	$3\sigma$	DA	<i>dop-2</i>	Sensory
RIML	128	$3\sigma$	TA	<i>ser-4, mod-1, dop-1</i>	Motor
RIMR	128	$3\sigma$	TA	<i>ser-4, mod-1, dop-1</i>	Motor
NSML	96	$3\sigma$	5-HT	<i>ser-4, dop-3, ser-2, tyra-2</i>	Pharynx
NSMR	96	$3\sigma$	5-HT	<i>ser-4, dop-3, ser-2, tyra-2</i>	Pharynx
HSNL	92	$1\sigma$	5-HT	<i>lgc-53, lgc-55</i>	Motor
HSNR	92	$1\sigma$	5-HT	<i>lgc-53, lgc-55</i>	Motor

Table 4.8 Rich-club neurons of the aggregate monoamine network. The rich-club column shows the threshold regime to which each neuron belongs, thus  $3\sigma$  indicates  $\Phi_{norm}(K) > 1 + 3\sigma$ , where  $\sigma$  is the standard deviation of the null model samples.

### 4.3.6 Multilink motifs

Multilink motif analysis provides another approach for investigating the interactions between the synaptic, gap junction and monoamine layers (Menichetti et al., 2014). Since each layer contains the same set of nodes but a different pattern of edges, the frequencies with which different combinations of links co-occur between pairs of nodes throughout the multiplex network can be determined. Of the 20 possible multilink motifs (Figure 4.19), six were found to be overrepresented and three underrepresented compared to random networks with similar degree distributions to the real connectome (Figure 4.18 & Table 5.7). The most overrepresented motif is *motif 1*, containing no connections. This indicates that specific link combinations are more likely to co-occur than expected by chance, therefore also increasing the number of ‘empty’ pairs compared to a random distribution of all edges. Three of the overrepresented motifs – reciprocal chemical synapses (*motif 3*) and the co-occurrence of a gap junction with a single or reciprocal chemical synapse (*motifs 5 & 6*) – have been reported in an earlier analysis of the physical network (Varshney et al., 2011). These also align with results from the degree-degree correlation and reducibility (Figures 4.8 & 4.9) indicating that synapses and gap junctions frequently overlap. This is mirrored in the underrepresentation of *motifs 2 & 4* corresponding to synapses or gap junctions alone.

Although the overlap between monoamine and physical connectivity is low (Figure 4.9), multilink motif analysis reveals a few overrepresented motifs involving monoamines. The most interesting of these corresponds to a unidirectional monoamine link coincident with reciprocal synaptic connections (*motif 10*; see Table 4.9). The structure of this motif is well-suited to provide positive or negative feedback in response to experience, suggesting that this may be a functionally important aspect of monoamine activity within the wider network. Indeed, connections of this type have been implicated in a number of *C. elegans* behaviours; for example, *motif 10* connections between ADF and AIY have been shown to be important for the learning of pathogen avoidance (Zhang et al., 2005) and connections between RIM and RMD are important for the suppression of head movements during escape behaviour (Pirri et al., 2009). Putative *Motif 10* connections between PDE and DVA are also thought to play a role in controlling neuropeptide release (Bhattacharya and Francis, 2015).

Intriguingly, most examples of *motif 10* (all except RIMR-RMDR) involve either serotonin or dopamine as the monoamine transmitter. Indeed, when we consider the monoamine networks separately (e.g. Syn-Gap-DA or Syn-Gap-TA multilinks), *motif*

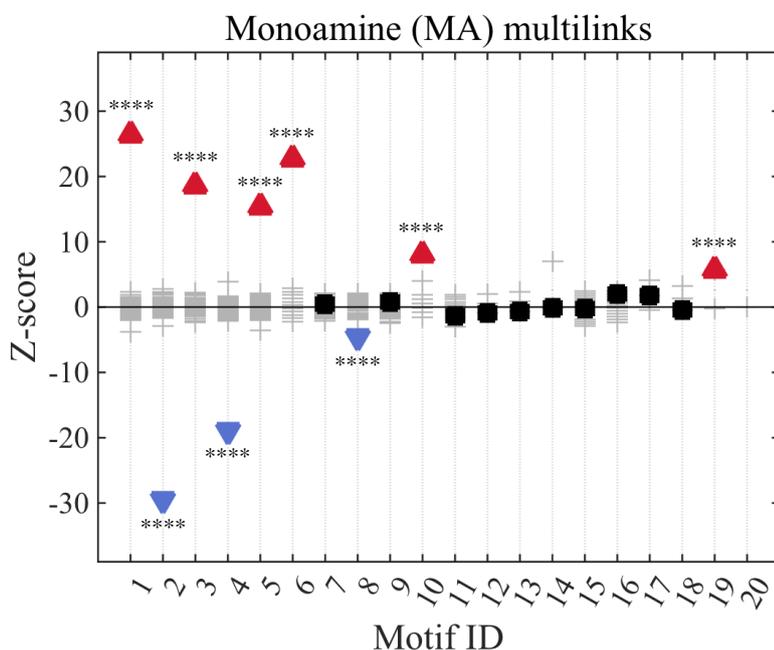


Fig. 4.18 Multilink motifs. Overrepresented motifs are represented by red upward-pointing triangles. Underrepresented motifs are represented by blue downward-pointing triangles. Non-significant motifs are shown by black squares. Values for 100 null model networks are shown as grey crosses. Asterisks report the significance level using the  $z$ -test, with Bonferroni-adjusted  $p$ -values: \*\*\*\* indicates  $p \leq 0.0001$ . Multilink motifs for individual monoamine networks are presented in Appendix G (page 257).

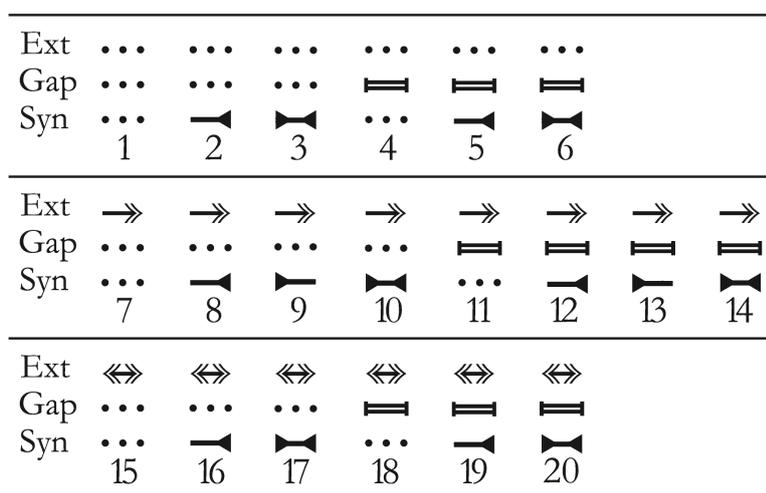


Fig. 4.19 Multilink motif IDs corresponding to all possible configurations of links between two neurons allowing for: no connection (dotted line), extrasynaptic links (Ext, represented as arrows on the top), bidirectional gap junctions (represented as bars in the middle) and synapses (represented as inverted arrows on the bottom line).

*10* is overrepresented for multilinks containing either serotonin or dopamine (Appendix Figure G.2, page 259), but not for tyramine or octopamine (Appendix Figures G.2 & G.3). Two different motifs were found to be overrepresented in the three-layer octopamine network: *motif 9* (a unidirectional synaptic connection coincident with an octopamine connection in the opposite direction) and *motif 11* (a unidirectional octopamine link coincident with a gap junction). Presumably these were not overrepresented in the aggregate network because the octopamine network is much smaller than the networks for the other monoamines. These motifs might serve similar functions to *motif 10* for dopamine and serotonin in providing feedback to modulate wired connections.

The only other overrepresented motif containing a monoamine is *motif 19*, for which there is only a single instance between two neurons of the same class (PDEL & PDER); while the underrepresentation of *motif 8*, consisting of a coincident synapse and monoamine connection, provides further support for the role of monoamines in extrasynaptic signalling, rather than as synaptic transmitters.

## 4.4 Discussion

### 4.4.1 The monoamine systems

This chapter has analysed the properties of an expanded *C. elegans* neuronal connectome, which incorporates a newly-compiled network of extrasynaptic monoamine signalling. Analyses reveal that this extrasynaptic network has a structure distinct from the synaptic network, containing its own rich-club of hub transmitting neurons. Although this system is separate and non-overlapping with the synaptic network, the hubs of both the physical and monoamine networks are highly connected to one another, with multilink motifs showing interaction between the systems at specific points in the network. This suggests that the monoamine system functions both independently – coordinating through its own rich-club – and in unison with the synaptic network through multilayer hubs such as RIM and through the overrepresented multilink motifs identified.

The low degree of overlap between the monoamine and synaptic networks occurs not only because many neurons expressing monoamine receptors are not postsynaptic

Cell A		Cell B
NSM (L/R)	→	I6
ADFR	→	ASHR
ADFR	→	AWBR
ADEL	→	IL2L
ADE (L/R)	→	FLPL
ADER	→	FLPR
ADFR	→	AIYR
CEPDR	→	RIS
RIMR	→	RMDR
HSNL	→	AIAL
HSNR	→	AVJL
HSNR	→	PVQR
CEP (DL/VL)*	→	OLLL*
CEP (DR/VR)*	→	OLLR*
CEPDL	→	RMGL*
ADEL*	→	BDUL*
PDEL*	→	DVA*

Table 4.9 *Motif 10*. List of neurons connected by *motif 10* (i.e. unidirectional MA link, no gap junctions, and reciprocal synapses). Examples involving unconfirmed (i.e. *dop-5* or *dop-6*-mediated) dopamine receptors are listed in the lower section (marked with asterisks \*).

Motif ID	Frequency	Expected	Z-score
1	41298	40974	26.43
2	1543	1991	-29.46
3	178	52	18.61
4	351	491	-18.95
5	154	54	15.39
6	49	4	22.68
7	1703	1698	0.48
8	52	78	-4.56
9	39	35	0.77
10	14	2	8.05
11	8	12	-1.29
12	0	1	-0.85
13	0	0	-0.63
14	0	0	-0.14
15	49	50	-0.25
16	11	6	2.04
17	1	0	1.83
18	0	0	-0.48
19	1	0	5.66
20	0	0	<i>n/a</i>

Table 4.10 Multilink motif frequencies for the monoamine, synaptic and gap junction layers. Motif IDs correspond to the numbers shown in Figure 4.19.

to aminergic neurons, but also because many postsynaptic targets of aminergic neurons do not appear to express monoamine receptors. Some of these synapses could be explained by cotransmission; in particular, tyraminerbic and serotonergic neurons also express either cholinergic or glutaminergic markers, and thus classical transmitters could be used in these wired synapses. However, the dopaminergic and octapaminergic neurons of *C. elegans* are not known to coexpress any classical neurotransmitter. A second possibility is that these synapses could utilise synaptically-released peptides as neurotransmitters. A third possibility is that the postsynaptic cells might express either an unknown monoamine receptor, or a known one at levels too low to be detected using existing reporters. Finally, it is possible that these putative synapses, which have been identified on the basis of electron micrographs, are not really functional synapses. Further work will be necessary to resolve this question.

The importance of extrasynaptic neuromodulation to the function of neural circuits is clearly established, for example from work on crustacean stomatogastric circuits (Marder, 2012). However, systematic attempts to map whole-organism connectomes have focused primarily on chemical synapses, with even gap junctions being difficult to identify using high-throughput electron microscopy approaches (Chklovskii et al., 2010). The incorporation of extrasynaptic neuromodulatory interactions, inferred here from gene expression data, adds a large number of new links largely non-overlapping with those of the physical connectome. Although the valence and strength of these inferred neuromodulatory links are largely unknown (information also lacking for much of the physical connectome), the monoamine networks described here nonetheless provide a more complete picture of potential pathways of communication between different parts of the *C. elegans* nervous system.

How complete is the monoamine network presented here? Since most cell identification in published work has been based on reporter co-expression with well-characterised markers, the rate of false positives (i.e. neurons falsely identified as monoamine receptor expressing) is probably very low. In contrast, the false-negative rate (monoamine receptor-expressing cells not included in the network) is almost certainly somewhat higher.

In some cases reporter transgenes appear to underreport full functional expression domains – e.g. *dop-3* & *dop-4* in ASH (Ezak and Ferkey, 2010; Ezcurra et al., 2011) – while in others (e.g. *ser-5*) it is clear that reporter transgenes are expressed in

unidentified neurons additional to those identified in relevant publications (Harris et al., 2009). With recently developed marker strains (Pereira et al., 2015; Serrano-Saiz et al., 2013), it should be possible to revisit cell identification for these reporters and fill in these missing gaps. In addition, other monoamines such as melatonin (Tanaka et al., 2007) might function as neuromodulators in *C. elegans*. While the monoamine-activated GPCRs in the worm genome have most likely all been identified, it is possible that additional ionotropic monoamine receptors might exist in the worm genome (Hobert, 2013), whose expression is presently uncharacterised. Potentially, some of these receptors might be expressed in postsynaptic targets of aminergic neurons (in particular, those of dopaminergic and octopaminergic neurons, which are not known to express classical neurotransmitters). However, the existence of additional monoamine receptor-expressing cells also means that non-synaptic edges are almost certainly undercounted in the network. Thus, the degree of monoamine releasing hubs – and their importance for neuronal signalling outside the wired connectome – is, if anything, understated by the current findings.

Given the importance of monoamine signalling across different species, it is interesting to note that abolishing monoamine vesicle loading through *cat-1* does not result in serious behavioural impairment in *C. elegans*, whereas synapses and gap junctions are required for normal function (Duerr et al., 1999; Kawano et al., 2011; Richmond, 2007). Duerr et al. (1999) suggest that this might be due to one of several reasons, such as the abundance of food in laboratory conditions reducing the importance of aminergic food responses and foraging; or the existence of parallel pathways, such as alternative vesicle loading mechanisms, or redundant routes of neuronal signalling. Robustness analysis of the combined physical and extrasynaptic networks reveals that the structure of the connectome is almost maximally robust to circuit degradation, supporting the observation that removing monoamine signalling does not result in catastrophic failure. It is worth noting that the high level of robustness seen in *C. elegans* is not unique to its neural network. Indeed, RNAi screens have demonstrated that 85 % of the genes in *C. elegans* can be knocked-down without causing significant changes in phenotypes, suggesting a high level of redundancy across all systems; however, such experiments do result in reduced fitness (Ramani et al., 2012).

Despite enormous differences in scale, the monoamine systems of *C. elegans* and mammals share a number of common properties. As in *C. elegans*, mammalian brains contain a relatively small number of monoamine-releasing neurons that project widely

to diverse brain regions; for example, in humans serotonin is produced by less than 100,000 cells in the raphe nuclei, or one millionth of all brain neurons (Trueta and De-Miguel, 2012). Moreover, extrasynaptic volume transmission is thought to account for much, if not most, monoamine signalling throughout the mammalian brain (De-Miguel and Fuxe, 2012; Fuxe et al., 2012). Parallels between monoamine systems in *C. elegans* and larger nervous systems are not exact; for example, in *C. elegans*, most if not all aminergic neurons appear capable of long-distance signalling, whereas monoamines in larger nervous systems can be restricted by glial diffusion barriers (Owald and Waddell, 2015). Nonetheless, mammalian monoamine-releasing neurons, like their *C. elegans* counterparts, appear to function as high-degree broadcasting hubs with functionally and spatially diverse targets (Trueta and De-Miguel, 2012). Thus, understanding how such hubs act within the context of the completely mapped physical circuitry of *C. elegans*, may provide useful insights into the currently unknown structures of multilayer neuronal networks in larger animals.

#### 4.4.2 Multiplex networks

While network theory has occasionally provided novel insights in *C. elegans* biology, more often the *C. elegans* physical connectome has provided a useful test-bed for validating new network theoretical concepts or their application to larger mammalian brains (Vértes and Bullmore, 2015). In recent years, multilayer complex systems have become an area of intense focus within network science, with a large number of papers dedicated to extending classical network metrics to the multilayer case and to developing new frameworks to understand the dynamical properties of multilayer systems (Kivela et al., 2014).

By definition, multilayer networks contain much more information than simple monoplex networks, leading to significant data-collection challenges. In social networks, for example, large monoplex datasets have been collected describing various types of interactions between people, but these are typically disparate datasets based on different populations. Multiplex datasets combining various edge types into a number of layers are often restricted in size (set of nodes for which data are collected) or in the choice of edges it is possible to consider (interaction types constrained by data availability) (Kivela et al., 2014).

The multiplex connectome of *C. elegans* has the potential to emerge as a gold standard in the study of multilayer networks, much like it has for the study of simple monoplex networks over the last 15 years. The synaptic, gap junction, and monoamine layers already represent a relatively reliable mapping of three distinct connection types. The lack of degree-degree correlation between some of these layers also suggests that they are not just different facets of one true underlying network (such that each edge is essentially duplicated across all layers). Rather, it suggests that the wired and wireless layers provide distinct channels of communication with differing functional roles, which are likely to be coupled in higher-order structures. For example, monoamine-based feedback loops or monoamine-regulated wired interactions. The different time-scales on which each of the layers operate is also likely to allow the emergence of interesting dynamical phenomena. Finally, the large number of distinct extrasynaptic interactions offers the scope for a more refined dataset, each aligned to the same complete set of 302 nodes.

### 4.4.3 Future work

In the future, it should be possible to expand the scope of the multilayer connectome to gain a more complete picture of neuronal functional connectivity. As a first step, reanalysis of reporters for monoamine receptors using recently developed reference strains (Pereira et al., 2015; Serrano-Saiz et al., 2013) could provide a largely complete monoamine signalling network. Additional work, profiling the pharmacodynamics, electrophysiology, and functional properties of each system will provide insights into how the monoamine systems work and modulate the activity of the synaptic network. Obtaining this information, while difficult, is uniquely feasible in *C. elegans* given the small size and precise cellular characterisation of its nervous system. Such a comprehensive multilayer connectome could serve as a prototype for understanding how different modes of signalling interact in the context of neuronal circuitry.

Obtaining extrasynaptic connectomes for larger brains, especially those of mammals, will likely be vastly more complicated than for *C. elegans*, due not only to the increase in size, but also the existence of additional structural and dynamical properties, such as glial barriers, cellular swelling, and arterial pulsations, all of which dynamically alter extracellular diffusion (Syková and Nicholson, 2008; Taber and Hurley, 2014). To address the difference in size, several techniques exist which could allow for such a map to be developed. For example, immunogold labelling was recently used to

simultaneously map synaptic links and neuropeptide locations in EM sections (Shahidi et al., 2015), providing a route to automatically capture elements of both the wired and wireless connectomes. Combined imaging and *in situ* sequencing has also been proposed as a method to concurrently localise synapses and neuromodulatory systems (Marblestone et al., 2014), while existing coarse-grained maps of brain transcriptomes, such as those in the Allen Brain Atlas (Hawrylycz et al., 2012; Lein et al., 2007), provide a useful first approximation for the distribution of molecules involved in volume transmission.

# Chapter 5

## Neuropeptide networks of *C. elegans*

### 5.1 Introduction

IN THE context of the monoamine systems of *C. elegans*, the previous chapter introduced how volume transmission networks can be mapped across an entire nervous system; however, monoamines represent just one class of volume transmission. As mentioned in the introduction, the nervous system of *C. elegans* also expresses 250 distinct neuropeptides from 122 precursor genes, and over 100 putative peptide receptors that form additional volume transmission networks (Hobert, 2013; Li and Kim, 2008). These include homologues of several well-known vertebrate neuropeptide receptors, including those for oxytocin/vasopressin (NTR-1), neuropeptide Y (NPR-1), and cholecystokinin (CKR-2).

Like monoamines, many *C. elegans* neuropeptides are known to function extrasynaptically, often acting at a distance. The high number of neuropeptides is not surprising given the ease with which new peptides can be coded for and synthesised using existing transcription, translation, and splicing mechanisms, adding greatly to the signalling richness of the nervous system. However, this high number of peptides provides a challenge in creating a complete network of their interactions. The receptors for many neuropeptides, and the ligands for many neuropeptide receptors, remain unknown; moreover, the distance over which signalling can occur is uncharacterised for most neuropeptide systems.

Although many neuropeptides have not been characterised (or even classified into families), good data exists for several of the systems. Neuropeptides in *C. elegans* have been associated with feeding behaviours (e.g. Cheong et al., 2015), metabolism (Cohen et al., 2009), mating behaviours (Barrios et al., 2012), as well as learning and memory (McDiarmid et al., 2015). As in other animals, neuropeptide signalling is critical for nervous system function, and frequently involves hormonal or other extrasynaptic mechanisms.

In an approach similar to that used for the monoamines, this chapter uses data from those neuropeptide systems that have been characterised to create a partial and provisional neuropeptide connectome, with the aim of providing insight into the differences between peptide signalling networks, and synaptic, gap junction, and monoamine networks. The analysis focuses on 12 neuropeptide systems with well-established ligands (with biologically-plausible  $EC_{50}$  values in *in vitro* assays) and precisely-characterised expression patterns for both receptor and peptide precursor genes. The results from this partial neuropeptide network show it to be a distinct signalling network, separate from both the physical and monoamine systems; however, multilayer hubs and motifs are identified which suggest that these distinct network layers are coupled, and interact, at several points in the network.

## 5.2 Materials & Methods

As in the previous chapter, the AC connectome was used in all analyses comparing the synaptic and extrasynaptic networks (see Chapter 3). All network methods are the same as those used in the analysis of the monoamine network (see Chapter 2 & 4). Edge counts, adjacency matrices, and reducibility clusters were all computed using binary directed versions of the networks. The same networks, excluding self-connections ( $A_{ij} = 0$  if  $i = j$ ), were used to compute the multilink motifs and reciprocity. As before network measures are compared to 100 null model networks (shown in the boxplots) generated using the degree-preserving randomisation procedure described in Chapter 2 (Algorithm 5), where the networks were randomised by performing  $10 \times M$  edge-swaps. The same null model networks were also used to compute the multilink motif  $z$ -scores, with each layer being randomised independently. As per the definition given in Humphries and Gurney (2008), the small-world index was normalised against 100 Erdős-Rényi (ER) reference networks containing the same number of nodes and edges as the actual networks.

### 5.2.1 Neuropeptide network construction

The neuropeptide network was constructed from published expression data for peptides and receptors, using a similar approach as was used for the monoamines. Only those systems were included for which sufficient expression and ligand-receptor interaction data existed in the literature, with interactions being limited to those with biologically plausible peptide-receptor EC<sub>50</sub> values (shown in Table 5.1). In total, 15 neuropeptides and 12 receptors were matched and included in the network. Networks were classified by receptor, allowing a many-to-many relationship between neuropeptides and receptors (Figure 5.1). For caveats on the gene expression data, see Appendix D (page 237).

### 5.2.2 Network data

Edge lists for individual network layers are provided in the supporting information (S1 dataset) of Bentley et al. (2016), and can be accessed online from PLOS<sup>1</sup> or PubMed Central<sup>2</sup>.

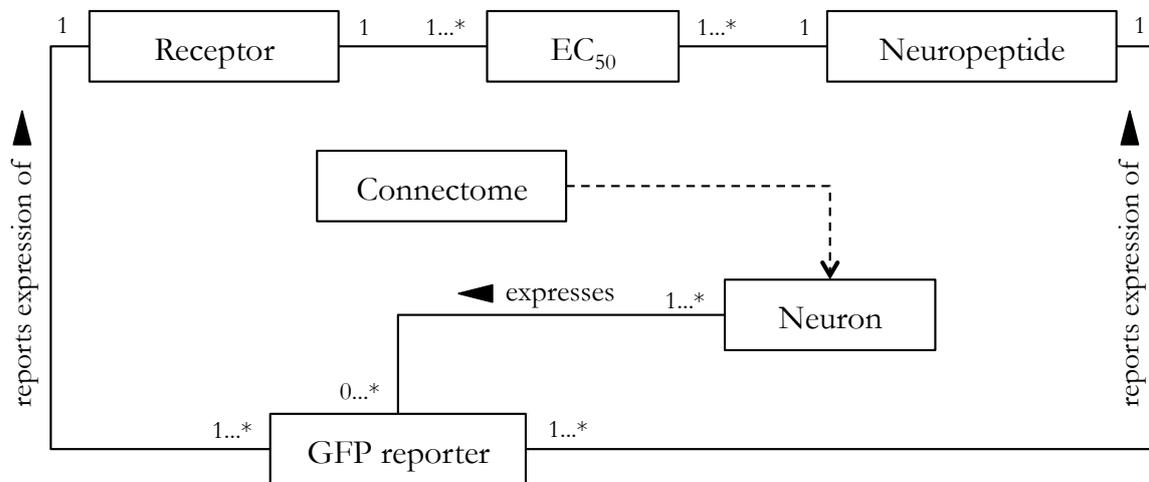


Fig. 5.1 Neuropeptide data model. UML class diagram (ISO, 2012a,b) of the conceptual data model used to construct the extrasynaptic neuropeptide network. Each box represents a class of data, with lines depicting their associations. The numbers represent the multiplicity of the constraints assumed in the modelling. Triangles denote the directionality of the relationships, and \* represents an unbounded (i.e.  $\infty$ ) multiplicity. For more details, see Figure legend 4.2 on page 116.

<sup>1</sup><http://dx.doi.org/10.1371/journal.pcbi.1005283.s004>

<sup>2</sup><https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5215746/bin/pcbi.1005283.s004.zip>

Receptor	Ligand	EC <sub>50</sub>	Reference
<b>NPY(npr) / RFamide receptor group</b>			
NPR-1	FLP-18	~100 nM	Rogers et al. (2003)
	FLP-21	2.5 nM	Kubiak et al. (2003a)
NPR-2	FLP-21	34.4 nM	Ezcurra et al. (2016)
NPR-3	FLP-15	162-599 nM	Kubiak et al. (2003b)
NPR-4	FLP-1	0.4-9 μM	Geary et al. (2002)
	FLP-4	5-80 nM	Geary et al. (2002)
	FLP-18	5 nM-1.2 μM	Cohen et al. (2009)
NPR-5	FLP-18	13.3-117.2 nM	Kubiak et al. (2008)
	FLP-21	267 nM	Kubiak et al. (2008)
NPR-11	FLP-1	1-8 μM	Geary et al. (2002)
	FLP-5	1-8 μM	Geary et al. (2002)
	FLP-18	180-800 nM	Geary et al. (2002)
	FLP-21	1-10 nM	Geary et al. (2002)
	NLP-1	1-100 μM?*	Chalasani et al. (2010)
FRPR-4	FLP-13	67-541 nM	Nelson et al. (2015)
<b>Somatostatin / Urotensin II receptor group</b>			
NPR-17	NLP-24	0.1-1 μM	Cheong et al. (2015)
<b>Gastrin / CCK-like receptor group</b>			
CKR-2	NLP-12	15-30 nM	Janssen et al. (2008b)
<b>Vasopressin-like receptor group</b>			
NTR-1	NTC-1	19 nM	Beets et al. (2012)
<b>Neurotensin / TPH-like receptor group</b>			
EGL-6	FLP-10	11 nM	Ringstad and Horvitz (2008)
	FLP-17	1-28 nM	Ringstad and Horvitz (2008)
<b>Class B / Secretin receptor group</b>			
PDFR-1	PDF-1	0.4-5 μM	Janssen et al. (2008a)
	PDF-2 /	114 nM	Janssen et al. (2008a)
	NLP-37		

Table 5.1 Neuropeptide receptor-ligand binding. \*No EC<sub>50</sub> value reported for NPR-11/NLP-1; strong biological activity seen in the micromolar range.

## 5.3 Results

Including the neuropeptide networks with the existing physical and monoamine layers, a multiplex network can be formed consisting of 18 layers (Figure 5.2). From aggregating the neuropeptide networks into a single layer (NP), it is clear that neuropeptide signalling forms a large component of the nervous system (see Table 5.2), with edges distributed throughout all of the neuron types (Figure 5.3). In just the partial network presented here, 239 neurons are seen to be involved in neuropeptide signalling (out of 302 possible), with 7035 connections between them providing greater connectivity than either the synaptic or monoamine layers alone.

The adjacency matrix reveals that the sending nodes are primarily sensory and interneurons, with the majority of connections between these classes. Few motor neuron to motor neuron connections are observed (Figure 5.3). Results shown in Table 5.3 support that the neuropeptide network is non-synaptic, with almost 60 % of receptor-expressing neurons receiving no synaptic input from cells that express cognate peptides. This table further shows that the majority of edges do not overlap with synapses (97 % non-overlapping), again consistent with a largely extrasynaptic mode of signalling.

Examining the degree-degree correlations of the network types (Figure 5.4) shows that the neuropeptides are neither correlated with the physical layers (synaptic and gap junction) nor the monoamine layer, demonstrating them to be a distinct signalling class. This is further seen in the reducibility results, where the synaptic and gap junction layers are highly-reducible, but not the MA or NP layers (Figure 5.5). Performing the same reducibility analysis on the individual, decomposed, network layers yields a more complex picture, with the physical layers still clustering together and the extrasynaptic layers showing variable outcomes (Figure 5.6). Some neuropeptide systems, in particular CKR-2, overlapped significantly with the networks of monoamine systems, while others, including the neuropeptide F/Y receptors NPR-1/2/5/11, showed little overlap with either the physical or other extrasynaptic networks.

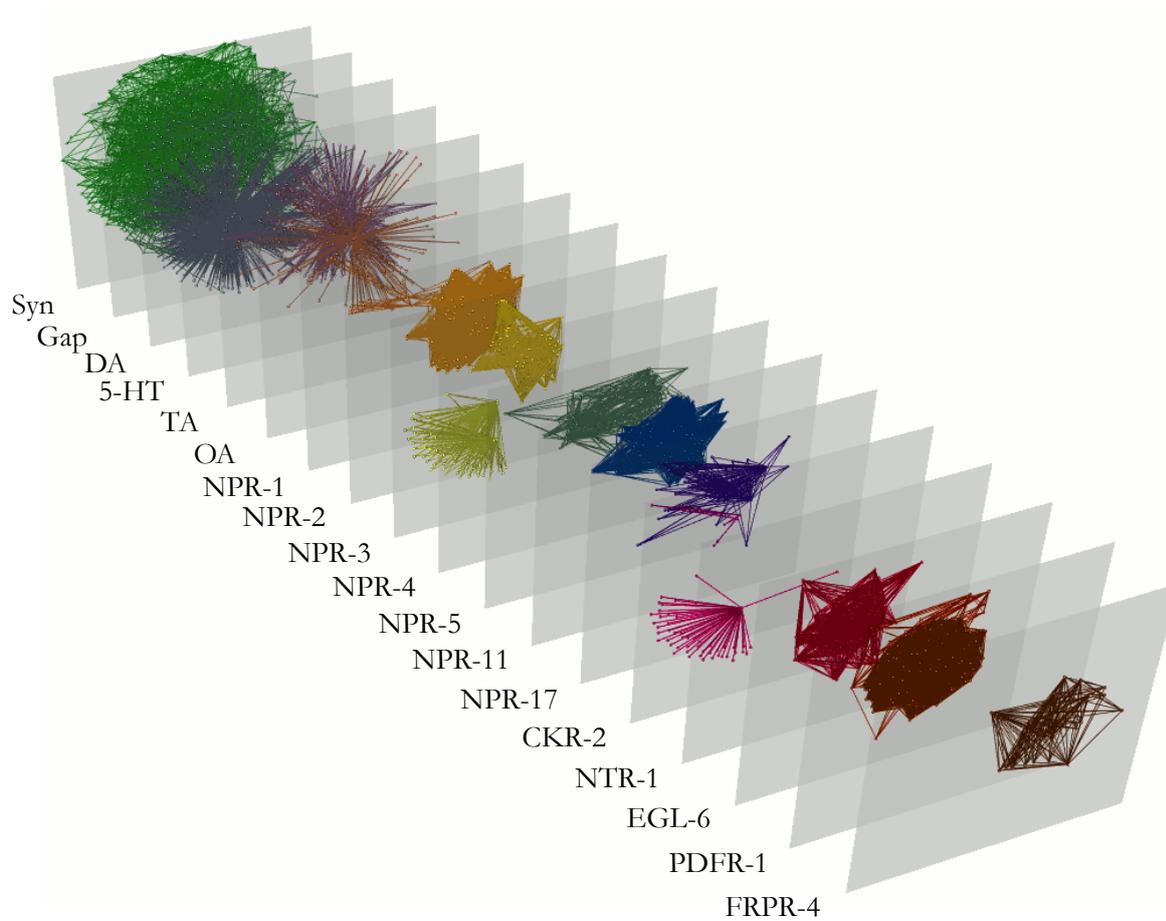


Fig. 5.2 Multilayer projection of the neuropeptide networks, including physical and monoamine layers. Node positions are the same in all layers.

Network	Nodes $N$	Nº ligand expressing	Nº receptor expressing	Edges $M^{\rightarrow}$
NPR-1	82	41	55	2241
NPR-2	44	31	19	583
NPR-3	54	4	50	200
NPR-4	46	39	9	349
NPR-5	59	41	30	1218
NPR-11	60	60	4	236
NPR-17	8	2	8	14
CKR-2	53	1	52	52
NTR-1	57	39	20	778
EGL-6	21	16	5	80
PDFR-1	83	50	44	2189
FRPR-4	27	19	8	152
<i>Aggregate</i>	239	141	187	7035

Table 5.2 Neuropeptide networks.

Network	Cells with no synaptic input		Non-synaptic edges	
	Nº	%	Nº	%
NPR-1	38	69.1	2209	98.6
NPR-2	7	36.8	563	96.6
NPR-3	49	98	199	99.5
NPR-4	4	44.4	319	91.4
NPR-5	16	53.3	1182	97
NPR-11	0	0	209	88.6
NPR-17	8	100	14	100
CKR-2	42	80.8	42	80.8
NTR-1	14	70	771	99.1
EGL-6	1	20	74	92.5
PDFR-1	17	38.6	2096	95.8
FRPR-4	4	50	144	94.7
<i>Aggregate</i>	106	56.7	6802	96.7

Table 5.3 The number of neuropeptide receptor-expressing cells that do not receive synapses from releasing cells, and the number of connections in each layer that are non-synaptic. Due to a many-to-many relationship between senders and receivers, the fraction of non-synaptic edges can exceed the fraction of non-synaptic cells.

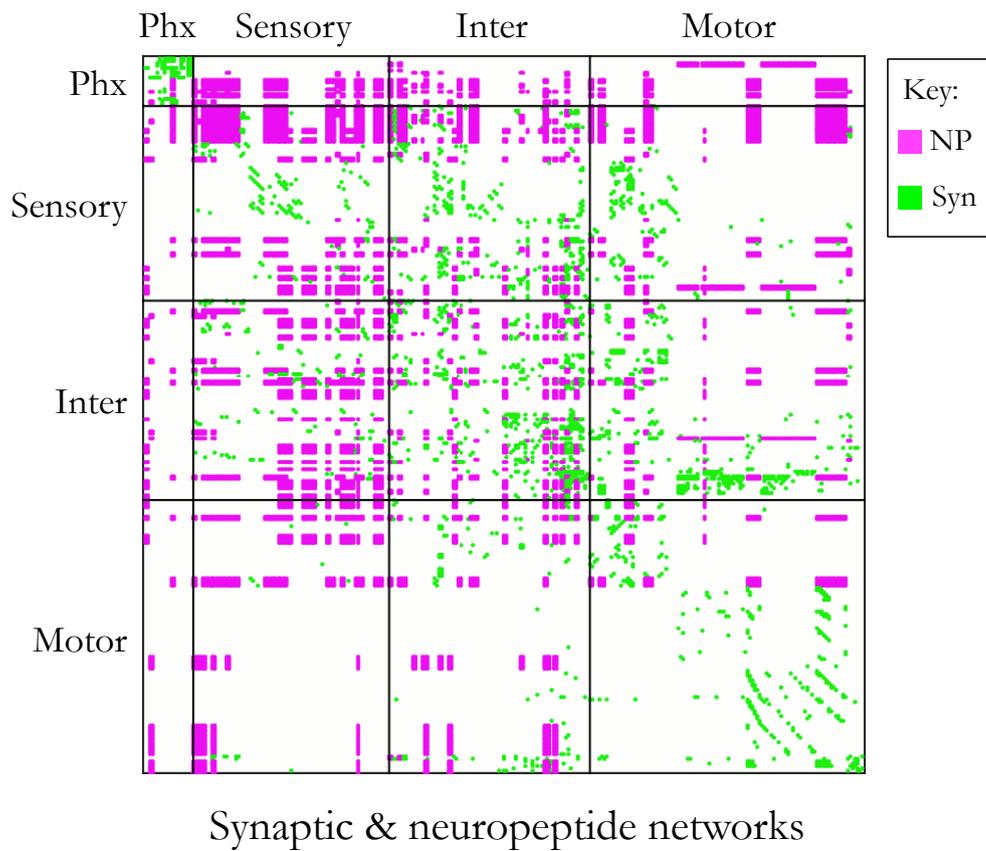


Fig. 5.3 Adjacency matrix showing the synaptic (magenta) and neuropeptide (green) networks.

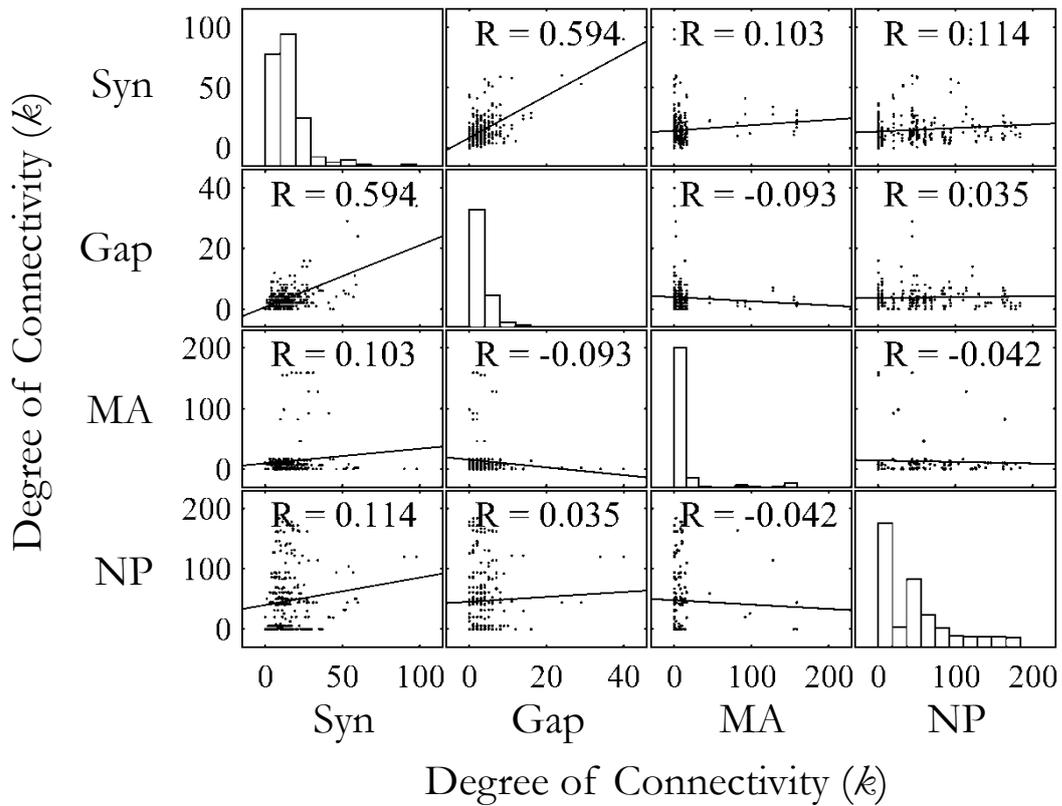


Fig. 5.4 Degree-degree correlation matrix. Off-diagonal panels show the degree-degree correlation between a pair of network layers. Panels on the diagonal show the degree distribution of the individual layers.

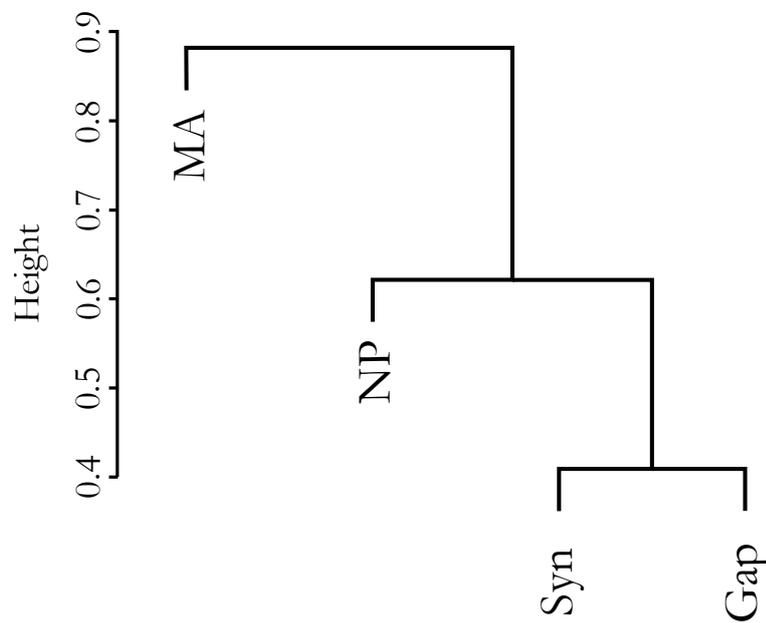


Fig. 5.5 Multilayer reducibility dendrogram for the synaptic, gap junction, monoamine, and neuropeptide layers. Layers close on the dendrogram have more overlapping edges and are more reducible. Branching height gives the Jensen-Shannon distance between the layers.

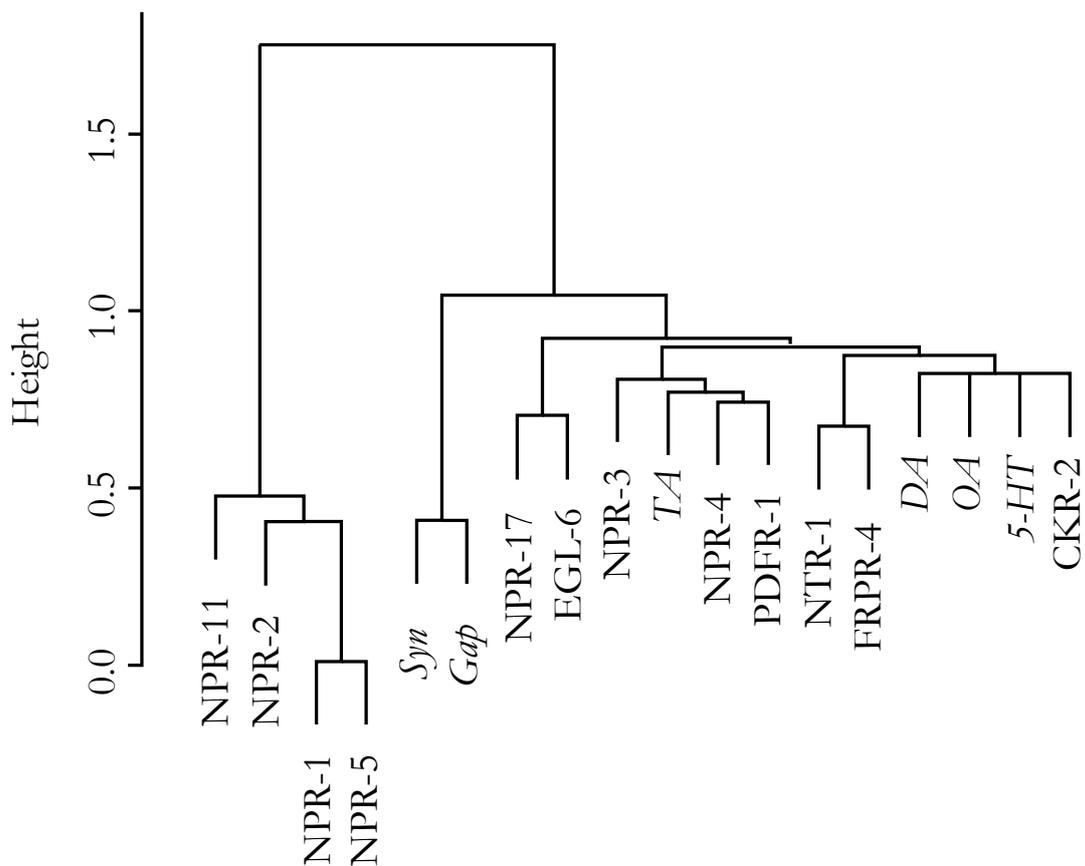


Fig. 5.6 Multilayer reducibility dendrogram showing all of the multiplex network layers. Layers close on the dendrogram have more overlapping edges and are more reducible. Physical and monoamine layers are italicised. Branching height gives the Jensen-Shannon distance between the layers.

### 5.3.1 Network measures

Examining the network measures for the NP network reveal it to have many properties in common with the MA network, including having a shorter characteristic path length and lower modularity and robustness than the physical network (Figure 5.8 & 5.10a). Both networks are also small-world (Figure 5.10b), have heavy-tailed degree distributions (Figure 5.9), and possess distinct rich-club organisation (Figure 5.11); the main differences are that the NP network has much higher reciprocity (Figure 5.7) and clustering (Figure 5.8a), which is an expected consequence of the much greater number of connections in the NP network; however, the addition of the neuropeptide connections to the physical and monoamine network have little effect on the aggregate properties, with the largest difference seen in the degree distribution power-law exponent ( $\alpha_{AC} = 3.57$  vs  $\alpha_{ag} = 5$ ).

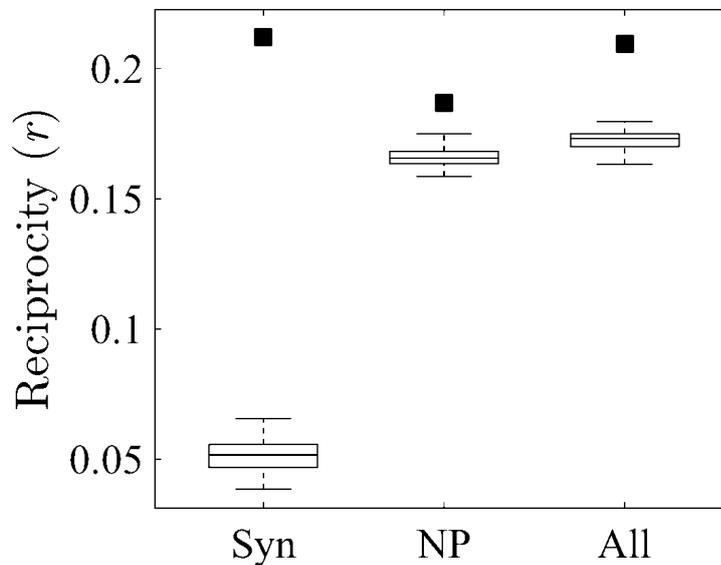


Fig. 5.7 Reciprocity for the directed synaptic network (Syn), neuropeptide network (NP), and aggregate synaptic+monoamine+neuropeptide network (All). Plot shows the observed values (filled squares) and expected values from 100 randomised networks (boxplots).

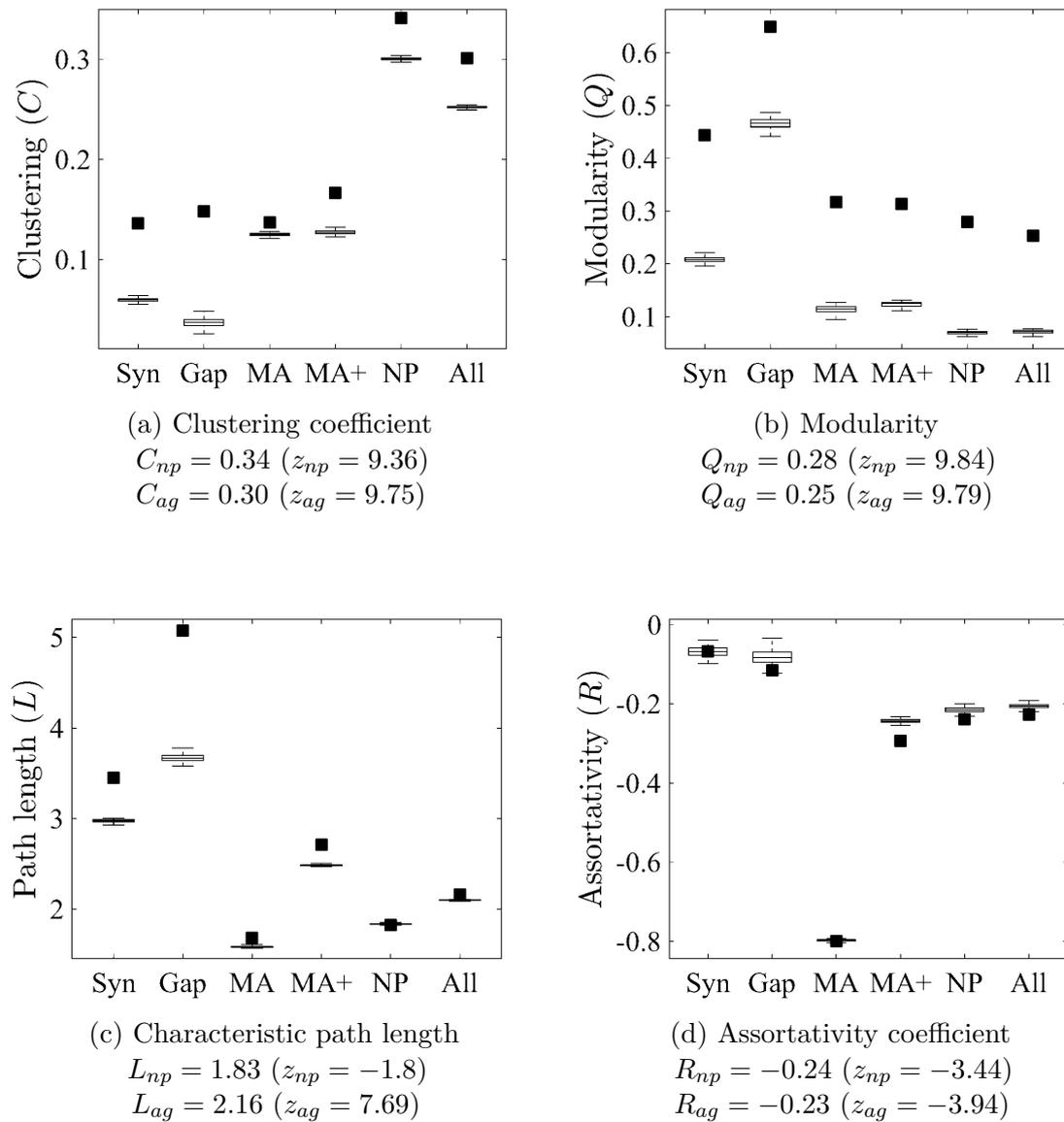


Fig. 5.8 Comparison of network metrics for the synaptic (Syn), gap junction (Gap), monoamine (MA), monoamine+syn+gap (MA+), neuropeptide (NP), and aggregate physical+monoamine+neuropeptide networks (All). Plots show the observed values (filled squares) and expected values from 100 randomised networks (boxplots).

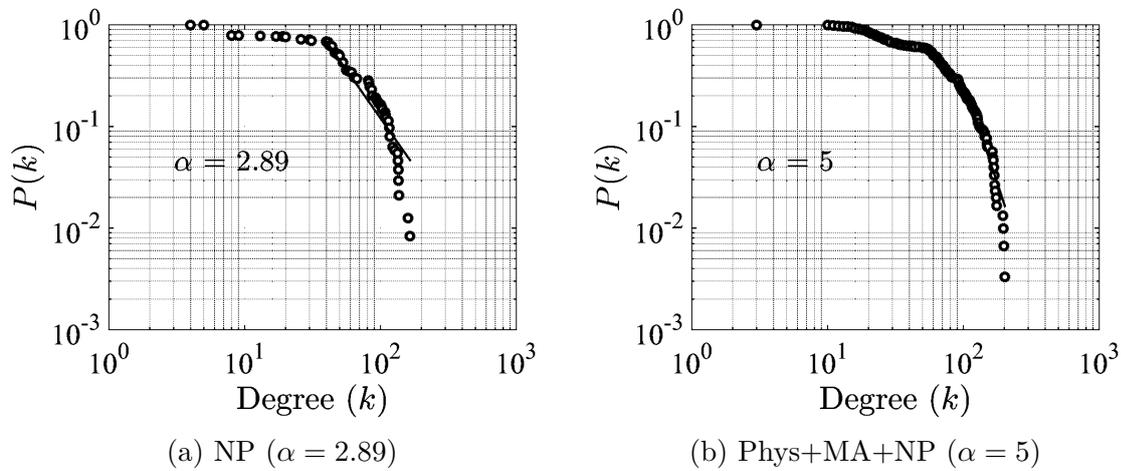


Fig. 5.9 Scale-free properties of the (a) neuropeptide network, and (b) aggregate synaptic, gap junction, monoamine, and neuropeptide network (showing best fit for the power-law distribution  $P(k) = k^{-\alpha}$ ).

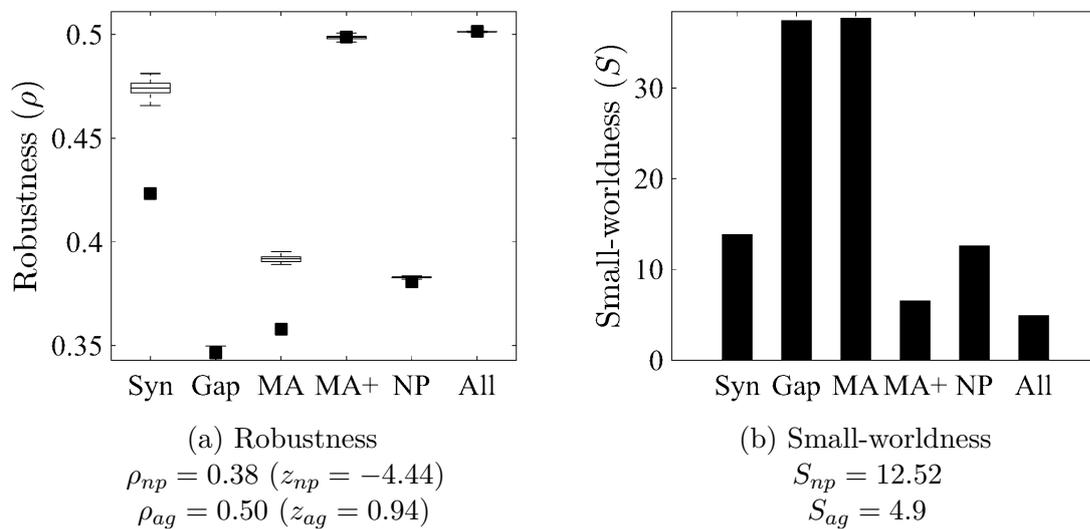


Fig. 5.10 Robustness and small-worldness of the synaptic (Syn), gap junction (Gap), monoamine (MA), monoamine+syn+gap (MA+), neuropeptide (NP), and aggregate physical+monoamine+neuropeptide networks (All). Plots show the observed values (filled squares) and expected values from 100 randomised networks (boxplots).

### 5.3.2 Rich-club

The neuropeptide network can be seen to exhibit its own rich-club of highly-connected hub nodes. This rich-club is mostly composed of sensory neurons (Table 5.4), following the same pattern previously seen with the monoamines. This suggests that the neuropeptides also provide a signalling short-cut from sensory neurons to the rest of the nervous system. Examining the relationships between the high-degree neurons in each layer (using the normalised degree rank-product; Table 5.5) shows that the previously identified RIM and DVA neurons continue to play a central role in linking all of the network layers (Figure 5.12).

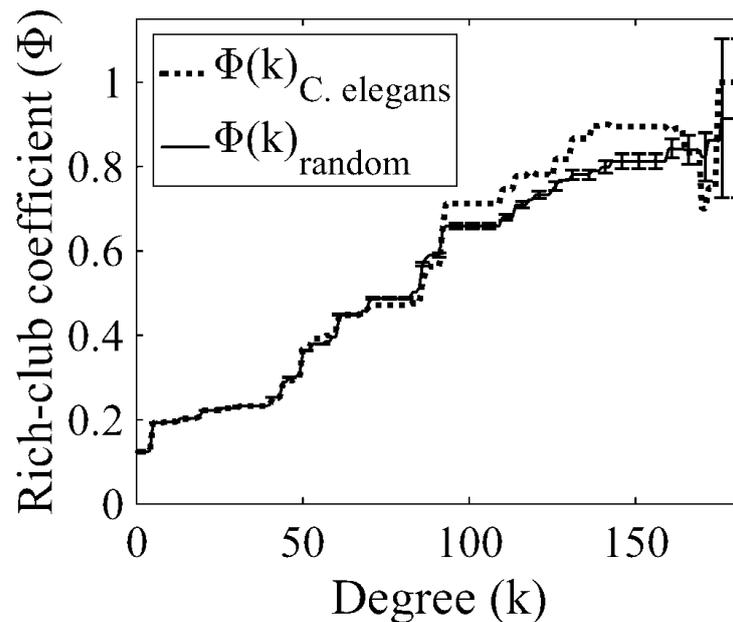


Fig. 5.11 Rich-club curve for the directed neuropeptide network. Dashed line indicates the rich-club coefficient for the *C. elegans* network and the solid curve represents the average rich-club coefficient of 100 randomised networks (preserving degree distribution).

Neuron	$k_{np}$	Receptors	Peptides	Type
PHA	182	NPR-1,-5; PDFR-1	FLP-4,-15; NLP-37	Sensory
ASEL	176	NPR-1,-5; NTR-1	FLP-4,-5,-13,-21	Sensory
PQR	176	NPR-1,-4,-17;NTR-1;PDFR-1	FLP-10;NLP-37;PDF-1	Sensory
ASER	170	NPR-1,-5; NTR-1	FLP-5,-13,-21	Sensory
URX	170	NPR-1; PDFR-1	FLP-10,-11,-21	Sensory
FLP	165	NPR-2; PDFR-1	FLP-4,-21	Sensory
ADF	162	NPR-2,-5; NTR-1	FLP-21	Sensory
ASH	162	NPR-1,-2; NTR-1	FLP-21	Sensory
ASI	160	NPR-5,-17	FLP-10,-21;NLP-1,-24;PDF-1	Sensory
RMG	160	NPR-1	FLP-1,-5,-21; PDF-1	Inter
ASK	159	NPR-5	FLP-13,-21; PDF-1	Sensory
ASG	144	NPR-1,5	FLP-13,-21; NTC-1	Sensory
AIY	140	NPR-11; CKR-2	FLP-1,-18	Inter
PHB	137	NPR-1,-5; PDFR-1	FLP-4;NLP-1;PDF-1;NLP-37	Sensory
ADL	131	NTR-1	FLP-4,-21	Sensory
M3	128	NPR-1	FLP-13,-18	Pharynx
BDU	126	NPR-4; NTR-1	FLP-10; NLP-1,-37	Inter
ASJ	125	NPR-5	FLP-21	Sensory
RIG	120	NPR-1	FLP-1,-18	Inter
AVA	118	NPR-4	FLP-1,-18	Inter
RIM	114	<i>n/a</i>	FLP-18; NLP-37	Motor
PVQ	110	NPR-2,-17;NTR-1;PDFR-1	<i>n/a</i>	Inter
DVA	104	EGL-6; FRPR-4	NLP-12;NTC-1	Inter
I2	93	NTR-1	FLP-5,-15	Pharynx

Table 5.4 Rich-club neurons of the aggregate neuropeptide network. Note that neuron IDs denote all neurons of that classes, with the exception of ASEL & ASER which have different expression profiles.

Neuron	$k_{norm}$	$k_{syn}$	$k_{gap}$	$k_{ma}$	$k_{np}$
RIMR*	0.164	34	14	128	114
RIML*	0.120	28	12	128	114
PVQR	0.047	22	10	16	110
ASHR	0.046	21	12	10	162
DVA*	0.046	54	10	8	104
RIS	0.036	27	16	14	44
VD01	0.033	14	16	16	61
ASHL	0.033	18	10	10	162
ADFR	0.030	21	4	82	162

Table 5.5 The normalised degree-rank product ( $k_{norm}$ ) showing the neurons that have the highest interaction in all of the network layers. Rich-club neurons are indicated with  $\star$ .

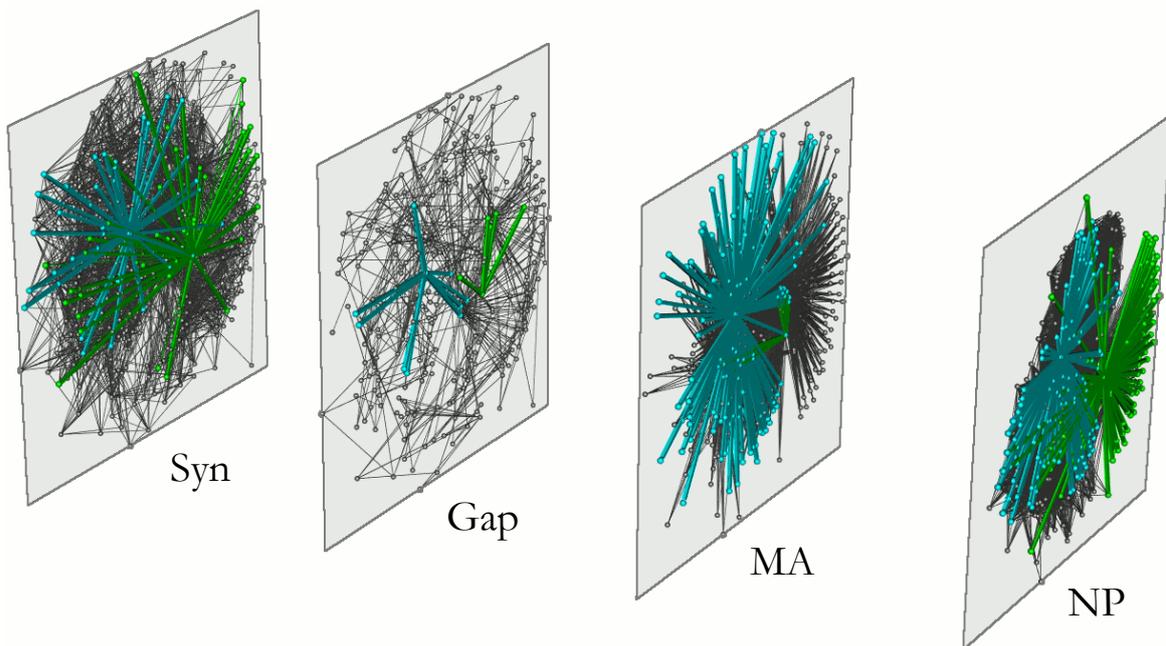


Fig. 5.12 Multilayer projection showing the RIM (blue) & DVA (green) rich-club neurons with the highest degree-rank production, acting as hubs across all four of the network layers.

### 5.3.3 Multilink motifs

Performing multilink analysis with neuropeptides as the extrasynaptic layer, once again reveals a similar pattern to the monoamines: *Motifs 1 - 6*, which consider only the physical networks with no extrasynaptic links, show similar  $z$ -score distribution patterns to before, while *Motif 10* (a unidirectional neuromodulatory connection coincident with a reciprocal synaptic connection) is also significantly overrepresented, as it was previously in the MA network. This supports the notion that this feedback motif plays a key role in extrasynaptic modulation of synaptic computation. Even more highly overrepresented relative to expectation was *motif 20*, reciprocal neuropeptide and synaptic connections coincident with a gap junction. This motif was not overrepresented in the multilink analysis for monoamines, perhaps because of the low reciprocity of the monoamine network. Interestingly, several of the *motif 20* multilinks (Table 5.6) are components of the RMG hub-and-spoke network, which has been implicated in the control of various behaviours including locomotion, aggregation, and pheromone response (Jang et al., 2012; Macosko et al., 2009).

Cell A		Cell B
FLPL	$\leftrightarrow$	FLPR
PHAL	$\leftrightarrow$	PHAR
PHBL	$\leftrightarrow$	PHBR
RMGL	$\leftrightarrow$	URXL
RMGR	$\leftrightarrow$	URXR
RMGR	$\leftrightarrow$	ASHR
PVR	$\leftrightarrow$	DVA
AVAL	$\leftrightarrow$	AVAR

Table 5.6 Examples of neuropeptide multilink *motif 20*. List of neurons connected by *motif 20* (i.e. reciprocal NP link, gap junction, and reciprocal synapses).

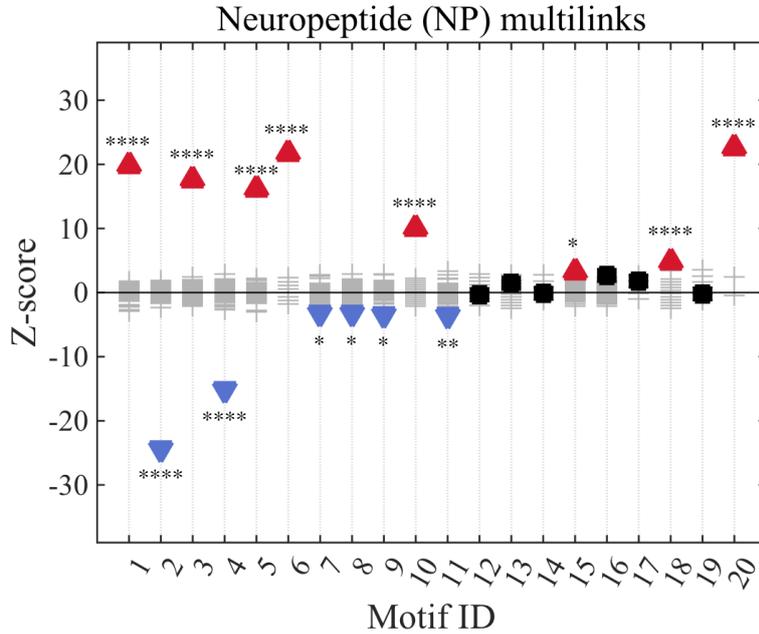


Fig. 5.13 Multilink motifs. Overrepresented motifs are represented by red upward-pointing triangles. Underrepresented motifs are represented by blue downward-pointing triangles. Non-significant motifs are shown by black squares. Values for 100 null model networks are shown as grey crosses. Asterisks report the significance level using the  $z$ -test, with Bonferroni-adjusted  $p$ -values: \* indicates  $p \leq 0.05$  ; \*\* indicates  $p \leq 0.01$ ; \*\*\*\* indicates  $p \leq 0.0001$ .

Ext	...	...	...	...	...	...		
Gap	...	...	...					
Syn	...	┐	┌	...	┐	┌		
	1	2	3	4	5	6		
Ext	⇒	⇒	⇒	⇒	⇒	⇒	⇒	
Gap	...	...	...	...				
Syn	...	┐	┌	┌	...	┐	┌	┌
	7	8	9	10	11	12	13	14
Ext	⇔	⇔	⇔	⇔	⇔	⇔		
Gap	...	...	...					
Syn	...	┐	┌	...	┐	┌		
	15	16	17	18	19	20		

Fig. 5.14 Multilink motif IDs. These correspond to all possible configurations of links between two neurons allowing for: no connection of a given type (dotted line), directed extrasynaptic neuropeptide links (Ext, represented as arrows on the top), bidirectional gap junctions (represented as bars in the middle) and synapses (represented as inverted arrowheads on the bottom line).

Motif ID	Frequency	Expected	Z-score
1	37076	36699	19.84
2	1358	1765	-24.34
3	150	43	17.62
4	298	430	-15.15
5	140	43	16.18
6	40	3	21.67
7	5404	5498	-3.27
8	117	153	-3.22
9	122	158	-3.46
10	40	10	10.02
11	45	67	-3.52
12	4	5	-0.35
13	10	7	1.49
14	1	1	-0.15
15	570	525	3.25
16	48	35	2.71
17	3	1	1.78
18	16	5	4.78
19	1	1	-0.20
20	8	0	22.54

Table 5.7 Multilink motif frequencies for the neuropeptide, synaptic and gap junction layers. Motif IDs correspond to the numbers shown in Figure 5.14.

## 5.4 Conclusions

Although the neuropeptide networks presented here are far from complete, it has been possible to analyse their structural properties to gain information about the patterns of organism-wide neuropeptide signalling for the first time.

While monoamine and neuropeptide signalling both occur extrasynaptically and act on similar timescales, the monoamine and neuropeptide networks show distinct non-overlapping topologies. In particular, the neuropeptide layer shows strikingly high clustering – even taking into account its high density of connections – and higher reciprocity than the monoamine network. These properties suggest the neuropeptide networks are important for cohesiveness within the nervous system. Multilink analysis also identified differences between the extrasynaptic monoamine and neuropeptide networks. In both cases, a unidirectional extrasynaptic connection coincident with a reciprocal synaptic connection (*motif 10*) was overrepresented in the multiplex connectome. This motif is well-suited to provide feedback between linked nodes, and occurs in several microcircuits implicated in learning and memory.

For neuropeptides, a second multilink motif, involving reciprocal neuromodulatory and synaptic connections coincident with a gap junction (*motif 20*) was found to be even more highly overrepresented than *motif 10*. This motif occurs in several places in the RMG-centred hub-and-spoke circuit that plays a key role in control of aggregation and arousal. As more neuropeptide systems become characterised (through expression analysis and deorphanisation), it is reasonable to expect additional examples of this motif will be identified; these may likewise have important computational roles in key neural circuits.

As with the previous analysis of the monoamine network, the RIM and DVA neurons continue to play a central role in linking the various layers of the multiplex network, suggesting they might function as interlayer communication hubs, facilitating and coordinating interaction between the wired, aminergic, and peptidergic transmission systems.

Unlike small-molecules, where gene-expression analysis can only infer their location through the associated synthesis or transport molecules, neuropeptides can be detected more directly, by identifying precisely where they are expressed. Even though there are more neuropeptides than small-molecule ligands, this property makes them more

amenable to large-scale mapping approaches. As mentioned previously, immunogold staining for neuropeptides has been used to map neuropeptides in the context of EM generated connectomic maps (Shahidi et al., 2015), providing an alternate route to mapping these networks in larger animals.

## Part IV

# Conclusions



# Chapter 6

## Conclusions & discussion

### 6.1 Contributions

#### 6.1.1 Major results

VOLUME transmission (VT), through the release and diffusion of molecules such as monoamines (MA) and neuropeptides (NP), plays numerous important roles in neural communication. Dysfunctional aminergic and peptidergic signalling has been implicated in the aetiology of many psychiatric and neurological disorders (Belzung et al., 2006; Lin et al., 2011), yet the ability to analyse the structural properties of these interaction networks has previously been lacking.

This thesis has demonstrated that maps of neuronal extrasynaptic VT can be constructed from gene expression data, presenting the first draft whole-animal map of monoamine signalling for the nervous system of *C. elegans*, along with a partial map of neuropeptide signalling. These represent the first such maps of VT signalling for any animal. The aminergic and peptidergic signalling networks were analysed alongside synaptic connections as part of a multiplex network, fulfilling the project objective to extend the definition of connectomics to include both wired synaptic communication and wireless volume transmission, increasing the information resolution of the *C. elegans* connectome model.

Analysing the topological structures of the aminergic and peptidergic networks, several important results were found:

1. Both the MA and NP systems are distinct networks with a low degree of connection overlap with either the synaptic network, or with each other. This suggests

that these systems do indeed function extrasynaptically and represent unique, non-redundant, channels of communication.

2. The MA and NP networks consist of predominantly disassortative star-like topologies, containing cores of high-degree broadcasting hubs.
3. The broadcasting hubs of the various VT layers are interconnected into a central rich-club, suggesting that the different systems interact to regulate their activities.
4. Several of the broadcasting hubs have a high degree of connectivity in more than one signalling layer. These nodes are likely to play an important role in coordinating the various aminergic, peptidergic, and synaptic signalling systems. The most notable examples are the interneurons RIM and DVA.
5. Performing motif analysis of the multilayer network identified several overrepresented connectivity patterns. One motif, consisting of a unidirectional extrasynaptic connection coincident with a reciprocal synaptic connection, is overrepresented in both the MA and NP networks, suggesting an important role for VT as a feedback mechanism.

### 6.1.2 Minor results

Chapter 3 analysed an updated version of the synaptic network of *C. elegans*, generated using computer-aided neuron tracing and published by the WormWiring project. This was compared to the original connectome derived by manual methods. The analysis of this network revealed a large increase in the number of synapses in the updated network, some of which were identified as having a high likelihood of being false-positives. These findings are being used to inform verification of the WormWiring network to improve its accuracy.

Despite differences in the number of synapses, most statistical measures of the network structure remained essentially unchanged between the original and updated connectomes, supporting the validity of existing network analyses in the literature and the overall accuracy of the Elegance / WormWiring approach to network reconstruction. However, as small-scale discrepancies exist between the two networks at the level of individual synapses, caution should be taken when using either dataset to inform the investigation of specific circuits.

## 6.2 Outlook & future work

### 6.2.1 Extending & validating the network

As developments in high-resolution neuroimaging and circuit reconstruction have begun to make the task of mapping large connectomes a realisable goal, concerns have been raised about the need to account for neuromodulatory VT (Bargmann, 2012; Bargmann and Marder, 2013; Brezina, 2010). This thesis has provided a proof-of-concept that these interactions can be mapped for an entire nervous system, at least in *C. elegans*; however, this work stands only as a first prototype.

As discussed in Chapters 4 & 5 (esp. § 4.4, page 141), work remains to be done in several areas to clarify the details of extrasynaptic interactions and obtain a more complete picture of VT signalling in *C. elegans*. Future work to extend the network might include:

1. additional expression profiling to identify missing neurons in the MA and NP systems;
2. the determination of the valence and strength of extrasynaptic interactions to provide network weights, eventually verifying these properties *in vivo* and exploring possible cross-reactivity between the various systems;
3. the deorphanisation of putative NP receptors;
4. verifying that the observed reporter expression patterns match functional expression, e.g. ensuring that the receptors are active in the cells where the reporters are expressed;
5. the determination of quantitative expression levels, and the relationship between expression levels and activity;
6. the quantification of *in vivo* VT diffusion dynamics and their effects on signalling speed and strength at various locations;
7. the investigation of additional molecules with the potential to act extrasynaptically, such as melatonin (Tanaka et al., 2007) or traditional neurotransmitters;
8. clarifying whether *dop-5* & *dop-6* are dopamine receptors; and
9. investigating conditional expression and state-dependent changes in signalling.

In *C. elegans* the tools already exist to answer many of these questions. With the continued development and application of such tools, the answers should be obtainable in the near future.

## 6.2.2 Network function

Obtaining a more complete map of extrasynaptic interactions will only provide insights into the structures of the VT signalling networks. A complete understanding will also require study into how the networks function and interact with the synaptic network to modulate information processing and behavioural output.

Approaches such as the noninvasive imaging of neural activity, optogenetic neurostimulation, and quantitative behavioural analysis (see § 3.4.3, page 105), coupled with traditional genetic and pharmacological techniques, offer several routes to investigate the functions of the MA and NP networks. Promising starting points for such investigations are the overrepresented multilink motifs (identified in § 4.3.6 & § 5.3.3), as identifying conserved functions performed by these motifs would greatly aid in understanding network organisation, behaviour generation, and furthermore provide a path to technological applications; for example, by revealing ways to more precisely predict and control neural activity. While difficult, such an investigation is uniquely feasible in *C. elegans* given the small size and precise cellular characterisation of its nervous system.

Recent work applying control theory to networks (Badhwar and Bagler, 2015; Gu et al., 2015; Liu et al., 2011; Tang et al., 2012; Yuan et al., 2013) has shown that certain functional properties of neural circuits can be accurately predicted from the network topology alone (Yan et al., 2017). This approach, along with similar ones, provides the means to begin exploring possible functions of the extrasynaptic VT networks in the absence of additional experimental data. Refining and verifying such predictive methods will likely be a necessary step to eventually understanding mammalian brains, due to the inability to exhaustively test the large number of neurons involved.

## 6.2.3 Mapping larger nervous systems

Although much work remains to be done to understand extrasynaptic neuromodulation in the small nervous system of *C. elegans*, it is foreseeable that it will eventually be possible to map the VT networks of larger nervous systems, including the human brain. This endeavour will require the development of new methods and theories to account

for the increased number of neurons, as well as the additional complex structures and dynamics, such as glial diffusion barriers and CSF circulation (Syková and Nicholson, 2008; Taber and Hurley, 2014).

As discussed in § 4.4.3 (page 147), several approaches exist which could allow for extrasynaptic networks to be mapped in humans, including immunolabeling (Shahidi et al., 2015) and *in situ* sequencing (Marblestone et al., 2014). Existing coarse-grained maps of brain transcriptomes (Hawrylycz et al., 2012; Lein et al., 2007) provide a useful first approximation for the distribution of molecules involved in volume transmission.

#### 6.2.4 Outlook

As extrasynaptic MA and NP signalling has been implicated in numerous diseases (Belzung et al., 2006; Lin et al., 2011), with non-synaptic receptors suggested as the primary site of action for many psychotherapeutic drugs (Vizi et al., 2010), it seems likely that understanding the structure and function of VT signalling networks will eventually provide important scientific and clinical insights. This thesis has explored how these networks can be mapped and analysed in the small nervous system of *C. elegans*, and serves to inform the direction of further investigation.



## References



# References

- Abramoff, M. D., Magalhães, P. J., and Ram, S. J. (2004). Image processing with ImageJ. *Biophotonics International*, 11(7):36–42. ISSN: 1081-8693.
- Achard, S., Salvador, R., Whitcher, B., Suckling, J., and Bullmore, E. (2006). A Resilient, Low-Frequency, Small-World Human Brain Functional Network with Highly Connected Association Cortical Hubs. *The Journal of Neuroscience*, 26(1):63–72. doi:10.1523/jneurosci.3874-05.2006.
- Agnati, L. F. and Fuxe, K. (2014). Extracellular-vesicle type of volume transmission and tunnelling-nanotube type of wiring transmission add a new dimension to brain neuro-glial networks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1652). doi:10.1098/rstb.2013.0505.
- Agnati, L. F., Guidolin, D., Guescini, M., Genedani, S., and Fuxe, K. (2010). Understanding wiring and volume transmission. *Brain Research Reviews*, 64(1):137–159. doi:10.1016/j.brainresrev.2010.03.003.
- Albert, R. and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47–97. doi:10.1103/RevModPhys.74.47.
- Albert, R., Jeong, H., and Barabási, A.-L. (1999). Internet: Diameter of the World-Wide Web. *Nature*, 401(6749):130–131. doi:10.1038/43601.
- Albert, R., Jeong, H., and Barabási, A.-L. (2000). Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382. doi:10.1038/35019019.
- Albertson, D. G. and Thomson, J. N. (1976). The pharynx of *Caenorhabditis elegans*. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 275(938):299–325. doi:10.1098/rstb.1976.0085.
- Alkema, M. J., Hunter-Ensor, M., Ringstad, N., and Horvitz, H. R. (2005). Tyramine Functions Independently of Octopamine in the *Caenorhabditis elegans* Nervous System. *Neuron*, 46(2):247–260. doi:10.1016/j.neuron.2005.02.024.
- Alon, U. (2007). Network motifs: theory and experimental approaches. *Nature Reviews Genetics*, 8(6):450–461. doi:10.1038/nrg2102.

- Altun, Z. F., Chen, B., Wang, Z.-W., and Hall, D. H. (2009). High resolution map of *Caenorhabditis elegans* gap junction proteins. *Developmental Dynamics*, 238(8):1936–1950. doi:10.1002/dvdy.22025.
- Altun, Z. F., Herndon, L. A., Wolkow, C. A., Crocker, C., Lints, R., and Hall, D. H. (2002). WormAtlas. <http://www.wormatlas.org>.
- Amaral, L. A. N., Scala, A., Barthelemy, M., and Stanley, H. E. (2000). Classes of small-world networks. *Proceedings of the national academy of sciences*, 97(21):11149–11152.
- Anand, K. and Bianconi, G. (2009). Entropy measures for complex networks: Toward an information theory of complex topologies. *Physical Review E*, 80(4). doi:10.1103/PhysRevE.80.045102.
- Anand, K., Bianconi, G., and Severini, S. (2011). Shannon and von Neumann entropy of random networks with heterogeneous expected degree. *Physical Review E*, 83(3):036109. doi:10.1103/PhysRevE.83.036109.
- Anastassiou, C. A., Perin, R., Markram, H., and Koch, C. (2011). Ephaptic coupling of cortical neurons. *Nature Neuroscience*, 14(2):217–223. doi:10.1038/nn.2727.
- Aparicio, S., Villazón-Terrazas, J., and Álvarez, G. (2015). A Model for Scale-Free Networks: Application to Twitter. *Entropy*, 17(8):5848–5867. doi:10.3390/e17085848.
- Artzy-Randrup, Y., Fleishman, S. J., Ben-Tal, N., and Stone, L. (2004). Comment on "Network Motifs: Simple Building Blocks of Complex Networks" and "Superfamilies of Evolved and Designed Networks". *Science*, 305(5687):1107–1107. doi:10.1126/science.1099334.
- Ashby, W. R. (1952). *Design for a brain*. John Wiley & Sons Inc., New York.
- Avery, L. and Horvitz, H. R. (1989). Pharyngeal pumping continues after laser killing of the pharyngeal nervous system of *C. elegans*. *Neuron*, 3(4):473–485. doi:10.1016/0896-6273(89)90206-7.
- Babai, L. (2015). Graph Isomorphism in Quasipolynomial Time. *arXiv [cs, math]*. arXiv: 1512.03547.
- Badhwar, R. and Bagler, G. (2015). Control of Neuronal Network in *Caenorhabditis elegans*. *PLOS ONE*, 10(9):e0139204. doi:10.1371/journal.pone.0139204.
- Bais, F. A. and Farmer, J. D. (2007). The physics of information. *arXiv*. arXiv: 0708.2837.
- Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512. doi:10.1126/science.286.5439.509.

- Barabási, A.-L. and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nature Reviews Genetics*, 5(2):101–113. doi:10.1038/nrg1272.
- Barabási, A.-L. and Pósfai, M. (2016). *Network science*. Cambridge University Press, Cambridge, UK. ISBN: 978-1-107-07626-6.
- Bargmann, C. I. (2012). Beyond the connectome: How neuromodulators shape neural circuits. *BioEssays*, 34(6):458–465. doi:10.1002/bies.201100185.
- Bargmann, C. I. and Horvitz, H. R. (1991). Control of larval development by chemosensory neurons in *Caenorhabditis elegans*. *Science*, 251(4998):1243. doi:10.1126/science.2006412.
- Bargmann, C. I. and Marder, E. (2013). From the connectome to brain function. *Nature Methods*, 10(6):483–490. doi:10.1038/nmeth.2451.
- Barnes, C. L. (2016). hiveplotter. *GitHub*, <http://github.com/clbarnes/hiveplotter>.
- Barrat, A., Barthélemy, M., and Vespignani, A. (2013). *Dynamical processes on complex networks*. Cambridge University Press, Cambridge, UK, 1st edition. ISBN: 978-0-521-87950-7.
- Barrios, A., Ghosh, R., Fang, C., Emmons, S. W., and Barr, M. M. (2012). PDF-1 neuropeptide signaling modulates a neural circuit for mate-searching behavior in *C. elegans*. *Nature Neuroscience*, 15(12):1675–1682. doi:10.1038/nn.3253.
- Bassett, D. S. and Sporns, O. (2017). Network neuroscience. *Nature Neuroscience*, 20(3):353–364. doi:10.1038/nn.4502.
- Battiston, F., Nicosia, V., and Latora, V. (2014). Structural measures for multiplex networks. *Physical Review E*, 89(3). doi:10.1103/PhysRevE.89.032804.
- Battiston, F., Nicosia, V., and Latora, V. (2016). The new challenges of multiplex networks: measures and models. *arXiv*. arXiv: 1606.09221.
- Beets, I., Janssen, T., Meelkop, E., Temmerman, L., Suetens, N., Rademakers, S., Jansen, G., and Schoofs, L. (2012). Vasopressin/Oxytocin-Related Signaling Regulates Gustatory Associative Learning in *C. elegans*. *Science*, 338(6106):543–545. doi:10.1126/science.1226860.
- Belzung, C., Yalcin, I., Griebel, G., Surget, A., and Leman, S. (2006). Neuropeptides in psychiatric diseases: an overview with a particular focus on depression and anxiety disorders. *CNS & Neurological Disorders-Drug Targets*, 5(2):135–145. doi:10.2174/187152706776359682.
- Bendesky, A., Tsunozaki, M., Rockman, M. V., Kruglyak, L., and Bargmann, C. I. (2011). Catecholamine receptor polymorphisms affect decision-making in *C. elegans*. *Nature*, 472(7343):313–318. doi:10.1038/nature09821.

- Bentley, B., Branicky, R., Barnes, C. L., Chew, Y. L., Yemini, E., Bullmore, E. T., Vértés, P. E., and Schafer, W. R. (2016). The Multilayer Connectome of *Caenorhabditis elegans*. *PLOS Computational Biology*, 12(12):e1005283. <http://dx.doi.org/10.1371/journal.pcbi.1005283>.
- Betz, R. F. and Bassett, D. S. (2017). Generative Models for Network Neuroscience: Prospects and Promise. *arXiv [q-bio]*. arXiv: 1708.07958.
- Bhatla, N. (2014). WormWeb.org - *C. elegans* Interactive Neural Network. *WormWeb*, <http://wormweb.org/neuralnet>.
- Bhatla, N., Droste, R., Sando, S. R., Huang, A., and Horvitz, H. R. (2015). Distinct Neural Circuits Control Rhythm Inhibition and Spitting by the Myogenic Pharynx of *C. elegans*. *Current Biology*, 25(16):2075–2089. doi:10.1016/j.cub.2015.06.052.
- Bhattacharya, R. and Francis, M. M. (2015). In the proper context: Neuropeptide regulation of behavioral transitions during food searching. *Worm*, 4(3):e1062971. doi:10.1080/21624054.2015.1062971.
- Bianconi, G. (2013). Statistical mechanics of multiplex networks: Entropy and overlap. *Physical Review E*, 87(6). doi:10.1103/PhysRevE.87.062806.
- Blau, A., Callaly, F., Cawley, S., Coffey, A., De Mauro, A., Epelde, G., Ferrara, L., Krewer, F., Liberale, C., Machado, P., and others (2014). The *Si elegans* Project—The Challenges and Prospects of Emulating *Caenorhabditis elegans*. In *Conference on Biomimetic and Biohybrid Systems*, pages 436–438. Springer. doi:10.1007/978-3-319-09435-9\_54.
- Boccaletti, S., Bianconi, G., Criado, R., Del Genio, C. I., Gómez-Gardeñes, J., Romance, M., Sendina-Nadal, I., Wang, Z., and Zanin, M. (2014). The structure and dynamics of multilayer networks. *Physics Reports*, 544(1):1–122. doi:10.1016/j.physrep.2014.07.001.
- Boulin, T., Etchberger, J. F., and Hobert, O. (2006). Reporter gene fusions. *WormBook*, doi:10.1895/wormbook.1.106.1.
- Bounova, G. (2014). Octave Networks Toolbox First Release. *GitHub*, doi:10.5281/zenodo.10778.
- Bounova, G. and de Weck, O. (2012). Overview of metrics and their correlation patterns for multiple-metric topology analysis on heterogeneous graph ensembles. *Physical Review E*, 85(1). doi:10.1103/PhysRevE.85.016117.
- Brandes, U., Delling, D., Gaertler, M., Gorke, R., Hofer, M., Nikoloski, Z., and Wagner, D. (2008). On Modularity Clustering. *IEEE Transactions on Knowledge and Data Engineering*, 20(2):172–188. doi:10.1109/tkde.2007.190689.

- Braunstein, S. L., Ghosh, S., and Severini, S. (2006). The laplacian of a graph as a density matrix: a basic combinatorial approach to separability of mixed states. *Annals of Combinatorics*, 10(3):291–317. doi:10.1007/s00026-006-0289-3.
- Brezina, V. (2010). Beyond the wiring diagram: signalling through complex neuro-modulator networks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1551):2363–2374. doi:10.1098/rstb.2010.0105.
- Briët, J. and Harremoës, P. (2009). Properties of Classical and Quantum Jensen-Shannon Divergence. *Physical Review A*, 79(5). doi:10.1103/PhysRevA.79.052311 arXiv: 0806.4472.
- Brown, S. R. (2013). Emergence in the central nervous system. *Cognitive Neurodynamics*, 7(3):173–195. doi:10.1007/s11571-012-9229-6.
- Budnik, V., Ruiz-Cañada, C., and Wendler, F. (2016). Extracellular vesicles round off communication in the nervous system. *Nature Reviews Neuroscience*, 17(3):160–172. doi:10.1038/nrn.2015.29.
- Bullmore, E. and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3):186–198. doi:10.1038/nrn2575.
- Burago, D., Burago, Y., and Ivanov, S. (2001). *A course in metric geometry*. American Mathematical Society. ISBN: 978-0-8218-2129-9.
- Caeyenberghs, K., Powell, H., Thomas, R., Brindley, L., Church, C., Evans, J., Muthukumaraswamy, S., Jones, D., and Hamandi, K. (2014). Hyperconnectivity in juvenile myoclonic epilepsy: A network analysis. *NeuroImage : Clinical*, 7:98–104. doi:10.1016/j.nicl.2014.11.018.
- Cajal, S. R. y. (1888). Estructura de los centros nerviosos de las aves. *Rev. Trim. Histol. Norm. Patol.*, 1:1–10.
- Cajal, S. R. y. (1899). Comparative study of the sensory areas of the human cortex. In *Clark University, 1889–1899: Decennial Celebration*, pages 311–382. Clark University, Worcester, MA.
- Cannon, W. B. (1929). Organization for physiological homeostasis. *Physiological Reviews*, 9(3):399–431.
- Carnell, L. (2005). The G-Protein-Coupled Serotonin Receptor SER-1 Regulates Egg Laying and Male Mating Behaviors in *Caenorhabditis elegans*. *Journal of Neuroscience*, 25(46):10671–10681. doi:10.1523/jneurosci.3399-05.2005.
- Carre-Pierrat, M., Baillie, D., Johnsen, R., Hyde, R., Hart, A., Granger, L., and Ségalat, L. (2006). Characterization of the *Caenorhabditis elegans* G protein-coupled serotonin receptors. *Invertebrate Neuroscience*, 6(4):189–205. doi:10.1007/s10158-006-0033-z.

- Chalasanani, S. H., Kato, S., Albrecht, D. R., Nakagawa, T., Abbott, L. F., and Bargmann, C. I. (2010). Neuropeptide feedback modifies odor-evoked dynamics in *Caenorhabditis elegans* olfactory neurons. *Nature Neuroscience*, 13(5):615–621. doi:10.1038/nn.2526.
- Chalfie, M., Sulston, J. E., White, J. G., Southgate, E., Thomson, J. N., and Brenner, S. (1985). The neural circuit for touch sensitivity in *Caenorhabditis elegans*. *The Journal of Neuroscience*, 5(4):956–964.
- Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2(3):200–219.
- Chase, D. L. and Koelle, M. R. (2007). Biogenic amine neurotransmitters in *C. elegans*. *WormBook*, doi:10.1895/wormbook.1.132.1.
- Chase, D. L., Pepper, J. S., and Koelle, M. R. (2004). Mechanism of extrasynaptic dopamine signaling in *Caenorhabditis elegans*. *Nature Neuroscience*, 7(10):1096–1103. doi:10.1038/nn1316.
- Chen, B. L. (2007). *Neuronal Network of C. elegans: from Anatomy to Behavior*. PhD thesis, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. <http://www.wormatlas.org/images/BethChenThesis.pdf>.
- Chen, B. L., Hall, D. H., and Chklovskii, D. B. (2006). Wiring optimization can relate neuronal structure and function. *Proceedings of the National Academy of Sciences of the United States of America*, 103(12):4723–4728. doi:10.1073/pnas.0506806103.
- Chen, C., Grennan, K., Badner, J., Zhang, D., Gershon, E., Jin, L., and Liu, C. (2011a). Removing Batch Effects in Analysis of Expression Microarray Data: An Evaluation of Six Batch Adjustment Methods. *PLOS ONE*, 6(2):e17238. doi:10.1371/journal.pone.0017238.
- Chen, S. X., Osipovich, A. B., Ustione, A., Potter, L. A., Hipkens, S., Gangula, R., Yuan, W., Piston, D. W., and Magnuson, M. A. (2011b). Quantification of factors influencing fluorescent protein expression using RMCE to generate an allelic series in the ROSA26 locus in mice. *Disease Models & Mechanisms*, 4(4):537–547. doi:10.1242/dmm.006569.
- Cheong, M. C., Artyukhin, A. B., You, Y.-J., and Avery, L. (2015). An opioid-like system regulating feeding behavior in *C. elegans*. *eLife*, 4:e06683. doi:10.7554/eLife.06683.001.
- Chklovskii, D. B., Vitaladevuni, S., and Scheffer, L. K. (2010). Semi-automated reconstruction of neural circuits using electron microscopy. *Current Opinion in Neurobiology*, 20(5):667–675. doi:10.1016/j.conb.2010.08.002.

- Chung, F. and Lu, L. (2002). The average distances in random graphs with given expected degrees. *Proceedings of the National Academy of Sciences*, 99(25):15879–15882. doi:10.1073/pnas.252631999.
- Chung, F. R. K. (1997). *Spectral graph theory*. Number no. 92 in Regional conference series in mathematics. Published for the Conference Board of the mathematical sciences by the American Mathematical Society, Providence, R.I. ISBN: 978-0-8218-0315-8.
- Chung, S., Schmalz, A., Ruiz, R. H., Gabel, C., and Mazur, E. (2013). Femtosecond Laser Ablation Reveals Antagonistic Sensory and Neuroendocrine Signaling that Underlie *C. elegans* Behavior and Development. *Cell Reports*, 4(2):316–326. doi:10.1016/j.celrep.2013.06.027.
- Clauset, A. (2016a). Clustering; Network Analysis and Modeling Lecture Notes. Santa Fe Institute. <http://tuvalu.santafe.edu/~aaronc/courses/5352/>.
- Clauset, A. (2016b). Random graphs with specified degrees; Network Analysis and Modeling Lecture Notes. Santa Fe Institute. <http://tuvalu.santafe.edu/~aaronc/courses/5352/>.
- Clauset, A., Newman, M. E., and Moore, C. (2004). Finding community structure in very large networks. *Physical Review E*, 70(6):066111. doi:10.1103/PhysRevE.70.066111.
- Clauset, A., Shalizi, C., and Newman, M. E. J. (2009). Power-Law Distributions in Empirical Data. *SIAM Review*, 51(4):661–703. doi:10.1137/070710111.
- Coates, J. C. and de Bono, M. (2002). Antagonistic pathways in neurons exposed to body fluid regulate social feeding in *Caenorhabditis elegans*. *Nature*, 419(6910):925–929. doi:10.1038/nature01170.
- Cohen, M., Reale, V., Olofsson, B., Knights, A., Evans, P., and de Bono, M. (2009). Coordinated Regulation of Foraging and Metabolism in *C. elegans* by RFamide Neuropeptide Signaling. *Cell Metabolism*, 9(4):375–385. doi:10.1016/j.cmet.2009.02.003.
- Colizza, V., Flammini, A., Serrano, M. A., and Vespignani, A. (2006). Detecting rich-club ordering in complex networks. *Nature Physics*, 2(2):110–115. doi:10.1038/nphys209.
- Cormen, T. H., Leiserson, C. E., Rivest, R. L., and Stein, C. (2009). *Introduction to algorithms*. The MIT Press, Cambridge, MA, 3rd edition. ISBN: 978-0-262-53305-8.
- Correa, P., LeBoeuf, B., and García, L. R. (2012). *C. elegans* Dopaminergic D2-Like Receptors Delimit Recurrent Cholinergic-Mediated Motor Programs during a Goal-Oriented Behavior. *PLOS Genetics*, 8(11):e1003015. doi:10.1371/journal.pgen.1003015.

- Crucitti, P., Latora, V., Marchiori, M., and Rapisarda, A. (2003). Efficiency of scale-free networks: error and attack tolerance. *Physica A: Statistical Mechanics and its Applications*, 320:622–642. doi:10.1016/S0378-4371(02)01545-5.
- Crucitti, P., Latora, V., Marchiori, M., and Rapisarda, A. (2004). Error and attack tolerance of complex networks. *Physica A: Statistical Mechanics and its Applications*, 340(1-3):388–394. doi:10.1016/j.physa.2004.04.031.
- Cunningham, K. A., Hua, Z., Srinivasan, S., Liu, J., Lee, B. H., Edwards, R. H., and Ashrafi, K. (2012). AMP-Activated Kinase Links Serotonergic Signaling to Glutamate Release for Regulation of Feeding Behavior in *C. elegans*. *Cell Metabolism*, 16(1):113–121. doi:10.1016/j.cmet.2012.05.014.
- Dayan, P. and Abbott, L. F. (2005). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. The MIT Press, Cambridge, MA. ISBN: 978-0-262-54185-5.
- de Arruda, G. F., Cozzo, E., Moreno, Y., and Rodrigues, F. A. (2015). On degree-degree correlations in multilayer networks. *arXiv [cond-mat, physics:physics]*. arXiv: 1507.04550.
- de Beudrap, N., Giovannetti, V., Severini, S., and Wilson, R. (2016). Interpreting the Von Neumann entropy of graph laplacians, and coentropic graphs. *Panor. Math. Pure Appl.*, 658:227–235.
- De Domenico, M., Nicosia, V., Arenas, A., and Latora, V. (2014). Layer aggregation and reducibility of multilayer interconnected networks. *arXiv*. arXiv:1405.0425.
- De Domenico, M., Nicosia, V., Arenas, A., and Latora, V. (2015a). Structural reducibility of multilayer networks. *Nature Communications*, 6:6864. doi:10.1038/ncomms7864.
- De Domenico, M., Porter, M. A., and Arenas, A. (2015b). MuxViz: a tool for multilayer analysis and visualization of networks. *Journal of Complex Networks*, 3(2):159–176. doi:10.1093/comnet/cnu038.
- De Domenico, M., Sasai, S., and Arenas, A. (2016). Mapping multiplex hubs in human functional brain networks. *Frontiers in Neuroscience*, 10:326. doi:10.3389/fnins.2016.00326.
- De-Miguel, F. F. and Fuxe, K. (2012). Extrasynaptic Neurotransmission as a Way of Modulating Neuronal Functions. *Frontiers in Physiology*, 3:16. doi:10.3389/fphys.2012.00016.
- de Silva, E. and Stumpf, M. P. (2005). Complex networks and simple models in biology. *Journal of the Royal Society Interface*, 2(5):419–430. doi:10.1098/rsif.2005.0067.

- Dehmer, M. and Mowshowitz, A. (2011). A history of graph entropy measures. *Information Sciences*, 181(1):57–78. doi:10.1016/j.ins.2010.08.041.
- Del Genio, C. I., Gross, T., and Bassler, K. E. (2011). All Scale-Free Networks Are Sparse. *Physical Review Letters*, 107(17):178701. doi:10.1103/PhysRevLett.107.178701.
- Dernovici, S., Starc, T., Dent, J. A., and Ribeiro, P. (2007). The serotonin receptor SER-1 (5ht2ce) contributes to the regulation of locomotion in *Caenorhabditis elegans*. *Developmental Neurobiology*, 67(2):189–204. doi:10.1002/dneu.20340.
- Dickinson, D. J. and Goldstein, B. (2016). CRISPR-Based Methods for *Caenorhabditis elegans* Genome Engineering. *Genetics*, 202(3):885–901. doi:10.1534/genetics.115.182162.
- Dijkstra, E. W. (1959). A Note on Two Problems in Connexion with Graphs. *Numerische Mathematik*, 1(1):269–271. doi:10.1007/bf01386390.
- Donnelly, J. L., Clark, C. M., Leifer, A. M., Pirri, J. K., Haburcak, M., Francis, M. M., Samuel, A. D. T., and Alkema, M. J. (2013). Monoaminergic Orchestration of Motor Programs in a Complex *C. elegans* Behavior. *PLOS Biology*, 11(4):e1001529. doi:10.1371/journal.pbio.1001529.
- Driscoll, M. and Kaplan, J. (1997). Mechanotransduction. In Riddle, D. L., Blumenthal, T., Meyer, B. J., and Priess, J. R., editors, *C. elegans II*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2nd edition. ISBN: 978-0-87969-532-3.
- Duerr, J. S., Frisby, D. L., Gaskin, J., Duke, A., Asermely, K., Huddleston, D., Eiden, L. E., and Rand, J. B. (1999). The cat-1 Gene of *Caenorhabditis elegans* Encodes a Vesicular Monoamine Transporter Required for Specific Monoamine-Dependent Behaviors. *The Journal of Neuroscience*, 19(1):72–84.
- Dunne, C. and Shneiderman, B. (2013). Motif simplification: improving network visualization readability with fan, connector, and clique glyphs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3247–3256. ACM. doi:10.1145/2470654.2466444.
- Edelman, G. M. and Gally, J. A. (2013). Reentry: a key mechanism for integration of brain function. *Frontiers in Integrative Neuroscience*, 7(63). doi:10.3389/fnint.2013.00063.
- Ekstrand, M. I., Nectow, A. R., Knight, Z. A., Latcha, K. N., Pomeranz, L. E., and Friedman, J. M. (2014). Molecular Profiling of Neurons Based on Connectivity. *Cell*, 157(5):1230–1242. doi:10.1016/j.cell.2014.03.059.
- Emmons, S. W. (2015). The beginning of connectomics: a commentary on White et al. (1986) ‘The structure of the nervous system of the nematode *Caenorhabditis*

- elegans'. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1666):20140309. doi:10.1098/rstb.2014.0309.
- Emmons, S. W., Hall, D. H., Brittin, C., and Cook, S. J. (2015). WormWiring project: Adult Hermaphrodite Connectivity. <http://wormwiring.org/>.
- Endres, D. M. and Schindelin, J. E. (2003). A new metric for probability distributions. *IEEE Transactions on Information theory*, 49(7):1858–1860. doi:10.1109/tit.2003.813506.
- Erdős, P. and Rényi, A. (1959). On random graphs I. *Publicationes Mathematicae*, 6:290–297.
- Erdős, P. and Rényi, A. (1960). On the evolution of random graphs. *Publication of the Mathematical Institute of the Hungarian Academy of Sciences*, pages 17–61.
- Erdős, P. and Rényi, A. (1961). On the strength of connectedness of a random graph. *Acta Mathematica Academiae Scientiarum Hungarica*, 12(1-2):261–267. doi:10.1007/BF02066689.
- Etchberger, J. F., Flowers, E. B., Poole, R. J., Bashllari, E., and Hobert, O. (2009). Cis-regulatory mechanisms of left/right asymmetric neuron-subtype specification in *C. elegans*. *Development*, 136(1):147–160. doi:10.1242/dev.030064.
- Etchberger, J. F., Lorch, A., Sleumer, M. C., Zapf, R., Jones, S. J., Marra, M. A., Holt, R. A., Moerman, D. G., and Hobert, O. (2007). The molecular signature and cis-regulatory architecture of a *C. elegans* gustatory neuron. *Genes & Development*, 21(13):1653–1674. doi:10.1101/gad.1560107.
- Euler, L. (1736). Solutio problematis ad geometriam situs pertinentis. *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, 8:128–140.
- Ezak, M. J. and Ferkey, D. M. (2010). The *C. elegans* D2-Like Dopamine Receptor DOP-3 Decreases Behavioral Sensitivity to the Olfactory Stimulus 1-Octanol. *PLOS ONE*, 5(3):e9487. doi:10.1371/journal.pone.0009487.
- Ezcurra, M., Tanizawa, Y., Swoboda, P., and Schafer, W. R. (2011). Food sensitizes *C. elegans* avoidance behaviours through acute dopamine signalling: Food sensitizes *C. elegans* avoidance behaviours. *The EMBO Journal*, 30(6):1110–1122. doi:10.1038/emboj.2011.22.
- Ezcurra, M., Walker, D. S., Beets, I., Swoboda, P., and Schafer, W. R. (2016). Neuropeptidergic Signaling and Active Feeding State Inhibit Nociception in *Caenorhabditis elegans*. *The Journal of Neuroscience*, 36(11):3157–3169. doi:10.1523/jneurosci.1128-15.2016.
- Fagiolo, G. (2007). Clustering in Complex Directed Networks. *Physical Review E*, 76(2). doi:10.1103/PhysRevE.76.026107 arXiv: physics/0612169.

- Faloutsos, M., Faloutsos, P., and Faloutsos, C. (1999). On Power-law Relationships of the Internet Topology. In *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*, SIGCOMM '99, pages 251–262, New York, NY, USA. ACM. doi:10.1145/316188.316229.
- Farah, M. J. and McClelland, J. L. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology: General*, 120(4):339–357. doi:10.1037/0096-3445.120.4.339.
- Farrant, M. and Kaila, K. (2007). The cellular, molecular and ionic basis of GABAA receptor signalling. In *Progress in Brain Research*, volume 160, pages 59–87. Elsevier. doi:10.1016/S0079-6123(06)60005-8.
- Fay, D. S. (2013). Classical genetic methods. *WormBook*, doi:10.1895/wormbook.1.165.1:1–58.
- Fiore, V. G., Dolan, R. J., Strausfeld, N. J., and Hirth, F. (2015). Evolutionarily conserved mechanisms for the selection and maintenance of behavioural activity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1684):20150053. doi:10.1098/rstb.2015.0053.
- Fornito, A., Zalesky, A., and Bullmore, E. T. (2010). Network Scaling Effects in Graph Analytic Studies of Human Resting-State fMRI Data. *Frontiers in Systems Neuroscience*, 4. doi:10.3389/fnsys.2010.00022.
- Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, 486(3–5):75–174. doi:10.1016/j.physrep.2009.11.002.
- Fosdick, B. K., Larremore, D. B., Nishimura, J., and Ugander, J. (2016). Configuring Random Graph Models with Fixed Degree Sequences. *arXiv*. arXiv: 1608.00607.
- Friedrich, R. W., Genoud, C., and Wanner, A. A. (2013). Analyzing the structure and function of neuronal circuits in zebrafish. *Frontiers in Neural Circuits*, 7:71. doi:10.3389/fncir.2013.00071.
- Friston, K. J. (1994). Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping*, 2(1-2):56–78. doi:10.1002/hbm.460020107.
- Fry, A. L., Laboy, J. T., and Norman, K. R. (2014). VAV-1 acts in a single interneuron to inhibit motor circuit activity in *Caenorhabditis elegans*. *Nature Communications*, 5:5579. doi:10.1038/ncomms6579.
- Frøkjær-Jensen, C. (2015). Transposon-Assisted Genetic Engineering with Mos1-Mediated Single-Copy Insertion (MosSCI). In *C. elegans*, Methods in Molecular Biology, pages 49–58. Humana Press, Totowa, NJ. doi:10.1007/978-1-4939-2842-2\_5.

- Frøkjær-Jensen, C., Davis, M. W., Hopkins, C. E., Newman, B., Thummel, J. M., Olesen, S.-P., Grunnet, M., and Jorgensen, E. M. (2008). Single copy insertion of transgenes in *C. elegans*. *Nature Genetics*, 40(11):1375–1383. doi:10.1038/ng.248.
- Fuxe, K., Borroto-Escuela, D. O., Romero-Fernandez, W., Diaz-Cabiale, Z., Rivera, A., Ferraro, L., Tanganelli, S., Tarakanov, A. O., Garriga, P., Narváez, J. A., Ciruela, F., Guescini, M., and Agnati, L. F. (2012). Extrasynaptic Neurotransmission in the Modulation of Brain Function. Focus on the Striatal Neuronal–Glial Networks. *Frontiers in Physiology*, 3:136. doi:10.3389/fphys.2012.00136.
- Galvani, L. (1791). *De Viribus Electricitatis In Motu Musculari Commentarius*. Typographia Institutii Scientiarum, Bologna, Italy.
- Gardner, M. R. and Ashby, W. R. (1970). Connectance of Large Dynamic (Cybernetic) Systems: Critical Values for Stability. *Nature*, 228(5273):784–784. doi:10.1038/228784a0.
- Garrison, J. L., Macosko, E. Z., Bernstein, S., Pokala, N., Albrecht, D. R., and Bargmann, C. I. (2012). Oxytocin/Vasopressin-Related Peptides Have an Ancient Role in Reproductive Behavior. *Science*, 338(6106):540–543. doi:10.1126/science.1226201.
- Gastner, M. T. and Ódor, G. (2016). The topology of large Open Connectome networks for the human brain. *Scientific Reports*, 6:srep27249. doi:10.1038/srep27249.
- Geary, T., Kubiak, T., Larsen, M., and Lowery, D. (2002). G protein-coupled receptor-like receptors and modulators thereof. *United States Patent*, PCT/US2000/032225.
- Geng, W., Cosman, P., Huang, C., and Schafer, W. R. (2003). Automated worm tracking and classification. In *Signals, Systems and Computers*, volume 2, pages 2063–2068. IEEE. doi:10.1109/acssc.2003.1292343.
- Getting, P. A. (1989). Emerging principles governing the operation of neural networks. *Annual Review of Neuroscience*, 12(1):185–204. doi:10.1146/annurev.ne.12.030189.001153.
- Gibbons, A. (1985). *Algorithmic graph theory*. Cambridge University Press, Cambridge, UK. ISBN: 978-0-521-28881-1.
- Gilbert, E. N. (1959). Random Graphs. *The Annals of Mathematical Statistics*, 30(4):1141–1144.
- Gleeson, P., Steuber, V., and Silver, R. A. (2007). neuroConstruct: A Tool for Modeling Networks of Neurons in 3d Space. *Neuron*, 54(2):219–235. doi:10.1016/j.neuron.2007.03.025.

- Głabowski, M., Musznicki, B., Nowak, P., and Zwierzykowski, P. (2013). Efficiency evaluation of shortest path algorithms. In *AICT 2013, The Ninth Advanced International Conference on Telecommunications*, pages 154–160. ISBN: 978-1-61208-279-0.
- Goh, W. W. B., Wang, W., and Wong, L. (2017). Why Batch Effects Matter in Omics Data, and How to Avoid Them. *Trends in Biotechnology*, 35(6):498–507. doi:10.1016/j.tibtech.2017.02.012.
- Gong, G., He, Y., Concha, L., Lebel, C., Gross, D. W., Evans, A. C., and Beaulieu, C. (2009). Mapping Anatomical Connectivity Patterns of Human Cerebral Cortex Using In Vivo Diffusion Tensor Imaging Tractography. *Cerebral Cortex*, 19(3):524–536. doi:10.1093/cercor/bhn102.
- Goulas, A., Bastiani, M., Bezgin, G., Uylings, H. B. M., Roebroek, A., and Stiers, P. (2014). Comparative Analysis of the Macroscale Structural Connectivity in the Macaque and Human Brain. *PLOS Computational Biology*, 10(3):e1003529. doi:10.1371/journal.pcbi.1003529.
- Gray, J. M., Hill, J. J., and Bargmann, C. I. (2005). A circuit for navigation in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America*, 102(9):3184–3191. doi:10.1073/pnas.0409009101.
- Gu, S., Pasqualetti, F., Cieslak, M., Telesford, Q. K., Yu, A. B., Kahn, A. E., Medaglia, J. D., Vettel, J. M., Miller, M. B., Grafton, S. T., and Bassett, D. S. (2015). Controllability of structural brain networks. *Nature Communications*, 6(8414). doi:10.1038/ncomms9414.
- Gürel, G., Gustafson, M. A., Pepper, J. S., Horvitz, H. R., and Koelle, M. R. (2012). Receptors and Other Signaling Proteins Required for Serotonin Control of Locomotion in *Caenorhabditis elegans*. *Genetics*, 192(4):1359–1371. doi:10.1534/genetics.112.142125.
- Hagberg, A. A., Schult, D. A., and Swart, P. J. (2008). Exploring Network Structure, Dynamics, and Function using NetworkX. In Varoquaux, G., Vaught, T., and Millman, J., editors, *Proceedings of the 7th Python in Science Conference (SciPy 2008)*, pages 11 – 15, Pasadena, CA USA.
- Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., and Sporns, O. (2008). Mapping the Structural Core of Human Cerebral Cortex. *PLOS Biology*, 6(7):e159. doi:10.1371/journal.pbio.0060159.
- Hamdan, F. F., Ungrin, M. D., Abramovitz, M., and Ribeiro, P. (1999). Characterization of a Novel Serotonin Receptor from *Caenorhabditis elegans*. *Journal of Neurochemistry*, 72(4):1372–1383. doi:10.1046/j.1471-4159.1999.721372.x.
- Han, L. (2012). *Graph Generative Models from Information Theory*. PhD thesis, University of York, York, UK.

- Han, L., Hancock, E. R., and Wilson, R. C. (2011). Characterizing graphs using approximate von Neumann entropy. In *Iberian Conference on Pattern Recognition and Image Analysis*, pages 484–491. Springer. doi:10.1007/978-3-642-21257-4\_60.
- Hao, D. and Li, C. (2011). The Dichotomy in Degree Correlation of Biological Networks. *PLOS ONE*, 6(12):e28322. doi:10.1371/journal.pone.0028322.
- Hapiak, V., Summers, P., Ortega, A., Law, W. J., Stein, A., and Komuniecki, R. (2013). Neuropeptides Amplify and Focus the Monoaminergic Inhibition of Nociception in *Caenorhabditis elegans*. *The Journal of Neuroscience*, 33(35):14107–14116. doi:10.1523/jneurosci.1324-13.2013.
- Hapiak, V. M., Hobson, R. J., Hughes, L., Smith, K., Harris, G., Condon, C., Komuniecki, P., and Komuniecki, R. W. (2009). Dual Excitatory and Inhibitory Serotonergic Inputs Modulate Egg Laying in *Caenorhabditis elegans*. *Genetics*, 181(1):153–163. doi:10.1534/genetics.108.096891.
- Harary, F. and Palmer, E. M. (1973). *Graphical Enumeration*. Academic Press, New York. Pages 124 & 241.
- Harriger, L., van den Heuvel, M. P., and Sporns, O. (2012). Rich Club Organization of Macaque Cerebral Cortex and Its Role in Network Communication. *PLOS ONE*, 7(9):e46497. doi:10.1371/journal.pone.0046497.
- Harris, G., Mills, H., Wragg, R., Hapiak, V., Castelletto, M., Korchnak, A., and Komuniecki, R. W. (2010). The Monoaminergic Modulation of Sensory-Mediated Aversive Responses in *Caenorhabditis elegans* Requires Glutamatergic/Peptidergic Cotransmission. *The Journal of Neuroscience*, 30(23):7889–7899. doi:10.1523/jneurosci.0497-10.2010.
- Harris, G. P., Hapiak, V. M., Wragg, R. T., Miller, S. B., Hughes, L. J., Hobson, R. J., Steven, R., Bamber, B., and Komuniecki, R. W. (2009). Three Distinct Amine Receptors Operating at Different Levels within the Locomotory Circuit Are Each Essential for the Serotonergic Modulation of Chemosensation in *Caenorhabditis elegans*. *The Journal of Neuroscience*, 29(5):1446–1456. doi:10.1523/jneurosci.4585-08.2009.
- Harvard/MGH-UCLA (2015). The Human Connectome Project. <http://www.humanconnectomeproject.org/>.
- Hassabis, D., Kumaran, D., Summerfield, C., and Botvinick, M. (2017). Neuroscience-Inspired Artificial Intelligence. *Neuron*, 95(2):245–258. doi:10.1016/j.neuron.2017.06.011.
- 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., Dalley, R. A., Daly, B. D., Dang, C., Datta, S., Dee, N., Dolbeare, T. A., Faber, V., Feng, D., Fowler, D. R., Goldy, J., Gregor, B. W., Haradon, Z., Haynor, D. R., Hohmann, J. G., Horvath, S., Howard, R. E., Jeromin, A., Jochim, J. M., Kinnunen, M., Lau, C., Lazarz, E. T., Lee, C., Lemon, T. A., Li, L., Li, Y., Morris, J. A., Overly, C. C., Parker, P. D., Parry, S. E., Reding, M., Royall, J. J., Schulkin, J., Sequeira, P. A., Slaughterbeck, C. R., Smith, S. C., Sodt, A. J., Sunkin, S. M., Swanson, B. E., Vawter, M. P., Williams, D., Wohnoutka, P., Zielke, H. R., Geschwind, D. H., Hof, P. R., Smith, S. M., Koch, C., Grant, S. G. N., and Jones, A. R. (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature*, 489(7416):391–399. doi:10.1038/nature11405.
- Hayasaka, S. and Laurienti, P. J. (2010). Comparison of Characteristics between Region- and Voxel-Based Network Analyses in Resting-State fMRI Data. *NeuroImage*, 50(2):499–508. doi:10.1016/j.neuroimage.2009.12.051.
- Hebb, D. (1949). *The organization of behavior: a neuropsychological theory*. John Wiley & Sons Inc., New York.
- Hernandez, V. H., Bortolozzi, M., Pertegato, V., Beltramello, M., Giarin, M., Zaccolo, M., Pantano, S., and Mammano, F. (2007). Unitary permeability of gap junction channels to second messengers measured by FRET microscopy. *Nature Methods*, 4(4):353–358. doi:10.1038/nmeth1031.
- Heuvel, M. P. v. d., Bullmore, E. T., and Sporns, O. (2016). Comparative Connectomics. *Trends in Cognitive Sciences*, 20(5):345–361. doi:10.1016/j.tics.2016.03.001.
- Heuvel, M. P. v. d. and Sporns, O. (2011). Rich-Club Organization of the Human Connectome. *The Journal of Neuroscience*, 31(44):15775–15786. doi:10.1523/jneurosci.3539-11.2011.
- Heuvel, M. P. v. d., Stam, C. J., Boersma, M., and Hulshoff Pol, H. E. (2008). Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain. *NeuroImage*, 43(3):528–539. doi:10.1016/j.neuroimage.2008.08.010.
- Hildebrand, D. G. C., Cicconet, M., Torres, R. M., Choi, W., Quan, T. M., Moon, J., Wetzel, A. W., Scott Champion, A., Graham, B. J., Randlett, O., Plummer, G. S., Portugues, R., Bianco, I. H., Saalfeld, S., Baden, A. D., Lillaney, K., Burns, R., Vogelstein, J. T., Schier, A. F., Lee, W.-C. A., Jeong, W.-K., Lichtman, J. W., and Engert, F. (2017). Whole-brain serial-section electron microscopy in larval zebrafish. *Nature*, 545(7654):345–349. doi:10.1038/nature22356.
- Hobert, O. (2013). The neuronal genome of *Caenorhabditis elegans*. *WormBook*, doi:10.1895/wormbook.1.161.1.

- Hobson, R. J., Hapiak, V. M., Xiao, H., Buehrer, K. L., Komuniecki, P. R., and Komuniecki, R. W. (2006). SER-7, a *Caenorhabditis elegans* 5-HT7-like Receptor, Is Essential for the 5-HT Stimulation of Pharyngeal Pumping and Egg Laying. *Genetics*, 172(1):159–169. doi:10.1534/genetics.105.044495.
- Hofstad, R. v. d. (2017). *Random graphs and complex networks*. Number 43 in Cambridge series in statistical and probabilistic mathematics. Cambridge University Press, Cambridge, UK. ISBN: 978-1-107-17287-6.
- Holme, P., Kim, B. J., Yoon, C. N., and Han, S. K. (2002). Attack vulnerability of complex networks. *Physical Review E*, 65(5). doi:10.1103/PhysRevE.65.056109.
- Hopcroft, J. and Tarjan, R. (1973). Algorithm 447: Efficient Algorithms for Graph Manipulation. *Commun. ACM*, 16(6):372–378. doi:10.1145/362248.362272.
- Horvitz, H. R., Chalfie, M., Trent, C., Sulston, J., and Evans, P. (1982). Serotonin and octopamine in the nematode *Caenorhabditis elegans*. *Science*, 216(4549):1012–1014. doi:10.1126/science.6805073.
- Hosseini, S. M. H. and Kesler, S. R. (2013). Influence of Choice of Null Network on Small-World Parameters of Structural Correlation Networks. *PLOS ONE*, 8(6):e67354. doi:10.1371/journal.pone.0067354.
- Howe, K. L., Bolt, B. J., Cain, S., Chan, J., Chen, W. J., Davis, P., Done, J., Down, T., Gao, S., Grove, C., Harris, T. W., Kishore, R., Lee, R., Lomax, J., Li, Y., Muller, H.-M., Nakamura, C., Nuin, P., Paulini, M., Raciti, D., Schindelman, G., Stanley, E., Tuli, M. A., Van Auken, K., Wang, D., Wang, X., Williams, G., Wright, A., Yook, K., Berriman, M., Kersey, P., Schedl, T., Stein, L., and Sternberg, P. W. (2016). WormBase 2016: expanding to enable helminth genomic research. *Nucleic Acids Research*, 44(D1):D774–D780. doi:10.1093/nar/gkv1217.
- Hu, Z., Pym, E. C. G., Babu, K., Vashlishan Murray, A. B., and Kaplan, J. M. (2011). A Neuropeptide-Mediated Stretch Response Links Muscle Contraction to Changes in Neurotransmitter Release. *Neuron*, 71(1):92–102. doi:10.1016/j.neuron.2011.04.021.
- Humphries, M. D. and Gurney, K. (2008). Network ‘Small-World-Ness’: A Quantitative Method for Determining Canonical Network Equivalence. *PLOS ONE*, 3(4):e0002051. doi:10.1371/journal.pone.0002051.
- Humphries, M. D., Gurney, K., and Prescott, T. (2006). The brainstem reticular formation is a small-world, not scale-free, network. *Proceedings of the Royal Society B: Biological Sciences*, 273(1585):503–511. doi:10.1098/rspb.2005.3354.
- Hunt-Newbury, R., Viveiros, R., Johnsen, R., Mah, A., Anastas, D., Fang, L., Halfnight, E., Lee, D., Lin, J., Lorch, A., McKay, S., Okada, H. M., Pan, J., Schulz, A. K., Tu, D., Wong, K., Zhao, Z., Alexeyenko, A., Burglin, T., Sonnhammer, E., Schnabel,

- R., Jones, S. J., Marra, M. A., Baillie, D. L., and Moerman, D. G. (2007). High-Throughput In Vivo Analysis of Gene Expression in *Caenorhabditis elegans*. *PLoS Biology*, 5(9):e237. doi:10.1371/journal.pbio.0050237.
- Husson, S. J. (2012). Keeping track of worm trackers. *WormBook*, doi:10.1895/wormbook.1.156.1:1–17.
- Ideker, T. and Sharan, R. (2008). Protein networks in disease. *Genome Research*, 18(4):644–652. doi:10.1101/gr.071852.107.
- ISO (2012a). ISO/IEC 19505-1:2012 Information technology - Object Management Group Unified Modeling Language (OMG UML), Infrastructure. <http://www.omg.org/spec/UML/>.
- ISO (2012b). ISO/IEC 19505-2:2012 Information technology - Object Management Group Unified Modeling Language (OMG UML), Superstructure. <http://www.omg.org/spec/UML/>.
- Iturria-Medina, Y., Sotero, R. C., Canales-Rodríguez, E. J., Alemán-Gómez, Y., and Melie-García, L. (2008). Studying the human brain anatomical network via diffusion-weighted MRI and Graph Theory. *NeuroImage*, 40(3):1064–1076. doi:10.1016/j.neuroimage.2007.10.060.
- Jacques, S. L. (2013). Optical properties of biological tissues: a review. *Physics in Medicine & Biology*, 58(11):R37. doi:10.1088/0031-9155/58/11/R37.
- Jafari, G., Xie, Y., Kullyev, A., Liang, B., and Sze, J. Y. (2011). Regulation of Extrasynaptic 5-HT by Serotonin Reuptake Transporter Function in 5-HT-Absorbing Neurons Underscores Adaptation Behavior in *Caenorhabditis elegans*. *The Journal of Neuroscience*, 31(24):8948–8957. doi:10.1523/jneurosci.1692-11.2011.
- Jang, H., Kim, K., Neal, S., Macosko, E., Kim, D., Butcher, R., Zeiger, D., Bargmann, C., and Sengupta, P. (2012). Neuromodulatory State and Sex Specify Alternative Behaviors through Antagonistic Synaptic Pathways in *C. elegans*. *Neuron*, 75(4):585–592. doi:10.1016/j.neuron.2012.06.034.
- Janssen, T., Husson, S. J., Lindemans, M., Mertens, I., Rademakers, S., Donck, K. V., Geysen, J., Jansen, G., and Schoofs, L. (2008a). Functional Characterization of Three G Protein-coupled Receptors for Pigment Dispersing Factors in *Caenorhabditis elegans*. *Journal of Biological Chemistry*, 283(22):15241–15249. doi:10.1074/jbc.M709060200.
- Janssen, T., Meelkop, E., Lindemans, M., Verstraelen, K., Husson, S. J., Temmerman, L., Nachman, R. J., and Schoofs, L. (2008b). Discovery of a Cholecystokinin-Gastrin-Like Signaling System in Nematodes. *Endocrinology*, 149(6):2826–2839. doi:10.1210/en.2007-1772.

- Jarrell, T. A., Wang, Y., Bloniarz, A. E., Brittin, C. A., Xu, M., Thomson, J. N., Albertson, D. G., Hall, D. H., and Emmons, S. W. (2012). The Connectome of a Decision-Making Neural Network. *Science*, 337(6093):437–444. doi:10.1126/science.1221762.
- Jayanthi, L. D., Apparsundaram, S., Malone, M. D., Ward, E., Miller, D. M., Epler, M., and Blakely, R. D. (1998). The *Caenorhabditis elegans* Gene T23g5.5 Encodes an Antidepressant- and Cocaine-Sensitive Dopamine Transporter. *Molecular Pharmacology*, 54(4):601–609.
- Jefferys, J. G. (1995). Nonsynaptic modulation of neuronal activity in the brain: electric currents and extracellular ions. *Physiological Reviews*, 75(4):689–723.
- Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N., and Barabási, A.-L. (2000). The large-scale organization of metabolic networks. *Nature*, 407(6804):651–654. doi:10.1038/35036627.
- Jin, M.-M. and Zhong, C. (2011). Role of gap junctions in epilepsy. *Neuroscience Bulletin*, 27(6):389–406. doi:10.1007/s12264-011-1944-1.
- Jin, Y., Jorgensen, E., Hartweg, E., and Horvitz, H. R. (1999). The *Caenorhabditis elegans* Gene *unc-25* Encodes Glutamic Acid Decarboxylase and Is Required for Synaptic Transmission But Not Synaptic Development. *The Journal of Neuroscience*, 19(2):539–548.
- Joyce, K. E., Hayasaka, S., and Laurienti, P. J. (2013). The Human Functional Brain Network Demonstrates Structural and Dynamical Resilience to Targeted Attack. *PLOS Computational Biology*, 9(1). doi:10.1371/journal.pcbi.1002885.
- Kaiser, M., Martin, R., Andras, P., and Young, M. P. (2007). Simulation of robustness against lesions of cortical networks: Simulation of robustness of cortical networks. *European Journal of Neuroscience*, 25(10):3185–3192. doi:10.1111/j.1460-9568.2007.05574.x.
- Kaplan, J. M. and Horvitz, H. R. (1993). A dual mechanosensory and chemosensory neuron in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences*, 90(6):2227–2231. doi:10.1073/pnas.90.6.2227.
- Karakuzu, O., Wang, D. P., and Cameron, S. (2009). MIG-32 and SPAT-3a are PRC1 homologs that control neuronal migration in *Caenorhabditis elegans*. *Development (Cambridge, England)*, 136(6):943–953. doi:10.1242/dev.029363.
- Kashtan, N. and Alon, U. (2005). Spontaneous evolution of modularity and network motifs. *Proceedings of the National Academy of Sciences of the United States of America*, 102(39):13773–13778. doi:10.1073/pnas.0503610102.
- Kashtan, N., Itzkovitz, S., Milo, R., and Alon, U. (2004). Topological generalizations of network motifs. *Physical Review E*, 70(3):031909. doi:10.1103/PhysRevE.70.031909 arXiv preprint q-bio/0312019.

- Kasthuri, N., Hayworth, K. J., Berger, D. R., Schalek, R. L., Conchello, J. A., Knowles-Barley, S., Lee, D., Vázquez-Reina, A., Kaynig, V., Jones, T. R., Roberts, M., Morgan, J. L., Tapia, J. C., Seung, H. S., Roncal, W. G., Vogelstein, J. T., Burns, R., Sussman, D. L., Priebe, C. E., Pfister, H., and Lichtman, J. W. (2015). Saturated Reconstruction of a Volume of Neocortex. *Cell*, 162(3):648–661. doi:10.1016/j.cell.2015.06.054.
- Kawano, T., Po, M. D., Gao, S., Leung, G., Ryu, W. S., and Zhen, M. (2011). An Imbalancing Act: Gap Junctions Reduce the Backward Motor Circuit Activity to Bias *C. elegans* for Forward Locomotion. *Neuron*, 72(4):572–586. doi:10.1016/j.neuron.2011.09.005.
- Keating, C. D., Kriek, N., Daniels, M., Ashcroft, N. R., Hopper, N. A., Siney, E. J., Holden-Dye, L., and Burke, J. F. (2003). Whole-Genome Analysis of 60 G Protein-Coupled Receptors in *Caenorhabditis elegans* by Gene Knockout with RNAi. *Current Biology*, 13(19):1715–1720. doi:10.1016/j.cub.2003.09.003.
- Kebschull, J., Garcia da Silva, P., Reid, A., Peikon, I., Albeanu, D., and Zador, A. (2016). High-Throughput Mapping of Single-Neuron Projections by Sequencing of Barcoded RNA. *Neuron*, 91(5):975–987. doi:10.1016/j.neuron.2016.07.036.
- Kelly, J. R., Rubin, A. J., Davis, J. H., Ajo-Franklin, C. M., Cumbers, J., Czar, M. J., de Mora, K., Gliberman, A. L., Monie, D. D., and Endy, D. (2009). Measuring the activity of BioBrick promoters using an in vivo reference standard. *Journal of Biological Engineering*, 3:4. doi:10.1186/1754-1611-3-4.
- Kerr, R. A. (2006). Imaging the activity of neurons and muscles. *WormBook*, doi:10.1895/wormbook.1.113.
- Kim, K. and Li, C. (2004). Expression and regulation of an FMR/Famide-related neuropeptide gene family in *Caenorhabditis elegans*. *The Journal of Comparative Neurology*, 475(4):540–550. doi:10.1002/cne.20189.
- Kitano, H., Hamahashi, S., and Luke, S. G. (1998). The perfect *C. elegans* project: An initial report. *Artificial Life*, 4(2):141–156. doi:10.1162/106454698568495.
- Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., and Porter, M. A. (2014). Multilayer networks. *Journal of Complex Networks*, 2(3):203–271. doi:10.1093/comnet/cnu016.
- Klimm, F., Bassett, D. S., Carlson, J. M., and Mucha, P. J. (2014). Resolving Structural Variability in Network Models and the Brain. *PLOS Computational Biology*, 10(3). doi:10.1371/journal.pcbi.1003491.
- Koelle, M. R. (1994). Integrating extrachromosomal arrays into the *C. elegans* chromosomes. *Koelle Lab, Yale School of Medicine*, <https://medicine.yale.edu/lab/koelle/protocols/>.

- Kong, X.-z., Liu, Z., Huang, L., Wang, X., Yang, Z., Zhou, G., Zhen, Z., and Liu, J. (2015). Mapping Individual Brain Networks Using Statistical Similarity in Regional Morphology from MRI. *PLOS ONE*, 10(11):e0141840. doi:10.1371/journal.pone.0141840.
- Kopell, N. J., Gritton, H., Whittington, M., and Kramer, M. (2014). Beyond the Connectome: The Dynome. *Neuron*, 83(6):1319–1328. doi:10.1016/j.neuron.2014.08.016.
- Krzywinski, M., Birol, I., Jones, S. J., and Marra, M. A. (2011). Hive plots—rational approach to visualizing networks. *Briefings in Bioinformatics*, 13(5):627–644. doi:10.1093/bib/bbr069.
- Kubiak, T. M., Larsen, M. J., Bowman, J. W., Geary, T. G., and Lowery, D. E. (2008). FMRamide-like peptides encoded on the flp-18 precursor gene activate two isoforms of the orphan *Caenorhabditis elegans* G-protein-coupled receptor Y58g8a.4 heterologously expressed in mammalian cells. *Biopolymers*, 90(3):339–348. doi:10.1002/bip.20850.
- Kubiak, T. M., Larsen, M. J., Nulf, S. C., Zantello, M. R., Burton, K. J., Bowman, J. W., Modric, T., and Lowery, D. E. (2003a). Differential Activation of “Social” and “Solitary” Variants of the *Caenorhabditis elegans* G Protein-coupled Receptor NPR-1 by Its Cognate Ligand AF9. *Journal of Biological Chemistry*, 278(36):33724–33729. doi:10.1074/jbc.M304861200.
- Kubiak, T. M., Larsen, M. J., Zantello, M. R., Bowman, J. W., Nulf, S. C., and Lowery, D. E. (2003b). Functional Annotation of the Putative Orphan *Caenorhabditis elegans* G-protein-coupled Receptor C10c6.2 as a FLP15 Peptide Receptor. *Journal of Biological Chemistry*, 278(43):42115–42120. doi:10.1074/jbc.M304056200.
- Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86. doi:doi:10.1214/aoms/1177729694.
- Lamberti, P. W., Majtey, A. P., Borrás, A., Casas, M., and Plastino, A. (2008). Metric character of the quantum Jensen-Shannon divergence. *Physical Review A*, 77(5). doi:10.1103/PhysRevA.77.052311.
- Lancichinetti, A. and Fortunato, S. (2009). Community detection algorithms: A comparative analysis. *Physical Review E*, 80(5). doi:10.1103/PhysRevE.80.056117.
- Latora, V. and Marchiori, M. (2001). Efficient Behavior of Small-World Networks. *Physical Review Letters*, 87(19):198701. doi:10.1103/PhysRevLett.87.198701.
- LeBoeuf, B., Correa, P., Jee, C., and García, L. R. (2014). *Caenorhabditis elegans* male sensory-motor neurons and dopaminergic support cells couple ejaculation and post-ejaculatory behaviors. *eLife*, 3:e02938. doi:10.7554/eLife.02938.

- Leek, J. T., Scharpf, R. B., Bravo, H. C., Simcha, D., Langmead, B., Johnson, W. E., Geman, D., Baggerly, K., and Irizarry, R. A. (2010). Tackling the widespread and critical impact of batch effects in high-throughput data. *Nature Reviews Genetics*, 11(10):733–739. doi:10.1038/nrg2825.
- Leicht, E. A. and Newman, M. E. (2008). Community structure in directed networks. *Physical review letters*, 100(11):118703. doi:10.1103/PhysRevLett.100.118703 arXiv:0709.4500.
- Leifer, A. M., Fang-Yen, C., Gershow, M., Alkema, M. J., and Samuel, A. D. T. (2011). Optogenetic manipulation of neural activity in freely moving *Caenorhabditis elegans*. *Nature Methods*, 8(2):147–152. doi:10.1038/nmeth.1554.
- Lein, E. S., Hawrylycz, M. J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A. F., Boguski, M. S., Brockway, K. S., Byrnes, E. J., Chen, L., Chen, L., Chen, T.-M., Chi Chin, M., Chong, J., Crook, B. E., Czaplinska, A., Dang, C. N., Datta, S., Dee, N. R., Desaki, A. L., Desta, T., Diep, E., Dolbeare, T. A., Donelan, M. J., Dong, H.-W., Dougherty, J. G., Duncan, B. J., Ebbert, A. J., Eichele, G., Estin, L. K., Faber, C., Facer, B. A., Fields, R., Fischer, S. R., Fliss, T. P., Frensley, C., Gates, S. N., Glattfelder, K. J., Halverson, K. R., Hart, M. R., Hohmann, J. G., Howell, M. P., Jeung, D. P., Johnson, R. A., Karr, P. T., Kaval, R., Kidney, J. M., Knapik, R. H., Kuan, C. L., Lake, J. H., Laramee, A. R., Larsen, K. D., Lau, C., Lemon, T. A., Liang, A. J., Liu, Y., Luong, L. T., Michaels, J., Morgan, J. J., Morgan, R. J., Mortrud, M. T., Mosqueda, N. F., Ng, L. L., Ng, R., Orta, G. J., Overly, C. C., Pak, T. H., Parry, S. E., Pathak, S. D., Pearson, O. C., Puchalski, R. B., Riley, Z. L., Rockett, H. R., Rowland, S. A., Royall, J. J., Ruiz, M. J., Sarno, N. R., Schaffnit, K., Shapovalova, N. V., Sivisay, T., Slaughterbeck, C. R., Smith, S. C., Smith, K. A., Smith, B. I., Sodt, A. J., Stewart, N. N., Stumpf, K.-R., Sunkin, S. M., Sutram, M., Tam, A., Teemer, C. D., Thaller, C., Thompson, C. L., Varnam, L. R., Visel, A., Whitlock, R. M., Wohnoutka, P. E., Wolkey, C. K., Wong, V. Y., Wood, M., Yaylaoglu, M. B., Young, R. C., Youngstrom, B. L., Feng Yuan, X., Zhang, B., Zwingman, T. A., and Jones, A. R. (2007). Genome-wide atlas of gene expression in the adult mouse brain. *Nature*, 445(7124):168–176. doi:10.1038/nature05453.
- Levitin, A. (2012). *Introduction to the design & analysis of algorithms*. Pearson, Boston, 3rd edition. ISBN: 978-0-13-231681-1.
- Li, C. and Kim, K. (2008). Neuropeptides. *WormBook*, doi:10.1895/wormbook.1.142.1:1–36.
- Li, N., Daie, K., Svoboda, K., and Druckmann, S. (2016). Robust neuronal dynamics in premotor cortex during motor planning. *Nature*, 532(7600):459–464. doi:10.1038/nature17643.
- Li, S., Dent, J. A., and Roy, R. (2003). Regulation of Intermuscular Electrical Coupling by the *Caenorhabditis elegans* Innexin *inx-6*. *Molecular Biology of the Cell*, 14(7):2630–2644. doi:10.1091/mbc.E02-11-0716.

- Li, Z., Li, Y., Yi, Y., Huang, W., Yang, S., Niu, W., Zhang, L., Xu, Z., Qu, A., Wu, Z., and Xu, T. (2012). Dissecting a central flip-flop circuit that integrates contradictory sensory cues in *C. elegans* feeding regulation. *Nature Communications*, 3:776. doi:10.1038/ncomms1780.
- Liang, X., Hsu, L.-M., Lu, H., Sumiyoshi, A., He, Y., and Yang, Y. (2017). The Rich-Club Organization in Rat Functional Brain Network to Balance Between Communication Cost and Efficiency. *Cerebral Cortex*, 26(12). doi:10.1093/cercor/bhw416.
- Lin, J. (1991). Divergence measures based on the Shannon entropy. *IEEE Transactions on Information theory*, 37(1):145–151. doi:10.1109/18.61115.
- Lin, Z., Canales, J. J., Björgvinsson, T., Thomsen, M. M., Qu, H., Liu, Q.-R., Torres, G. E., and Caine, S. B. (2011). Monoamine Transporters: Vulnerable and Vital Doorkeepers. *Progress in Molecular Biology and Translational Science*, 98:1–46. doi:10.1016/B978-0-12-385506-0.00001-6.
- Lindner, A. (2014). Methods, indicators of progress and target values for mapping of the human brain. Report D2.3.1, The Human Brain Project.
- Lints, R. and Emmons, S. W. (1999). Patterning of dopaminergic neurotransmitter identity among *Caenorhabditis elegans* ray sensory neurons by a TGF $\beta$  family signaling pathway and a Hox gene. *Development*, 126(24):5819–5831.
- Liu, X. F. and Tse, C. K. (2015). A General Framework for Complex Network Applications. *arXiv [physics, q-fin]*. arXiv: 1507.05687.
- Liu, Y.-Y., Slotine, J.-J., and Barabási, A.-L. (2011). Controllability of complex networks. *Nature*, 473(7346):167–173. doi:10.1038/nature10011.
- Loer, C. (2010). Neurotransmitters in *Caenorhabditis elegans*. *WormAtlas*, doi:10.3908/wormatlas.5.200.
- Luo, J., Xu, Z., Tan, Z., Zhang, Z., and Ma, L. (2015). Neuropeptide Receptors NPR-1 and NPR-2 Regulate *Caenorhabditis elegans* Avoidance Response to the Plant Stress Hormone Methyl Salicylate. *Genetics*, 199(2):523–531. doi:10.1534/genetics.114.172239.
- Macosko, E. Z., Pokala, N., Feinberg, E. H., Chalasani, S. H., Butcher, R. A., Clardy, J., and Bargmann, C. I. (2009). A hub-and-spoke circuit drives pheromone attraction and social behaviour in *C. elegans*. *Nature*, 458(7242):1171–1175. doi:10.1038/nature07886.
- Majtey, A. P., Lamberti, P. W., and Prato, D. P. (2005). Jensen-Shannon divergence as a measure of distinguishability between mixed quantum states. *Physical Review A*, 72(5):052310. doi:10.1103/PhysRevA.72.052310.

- Marblestone, A. H., Daugharthy, E. R., Kalhor, R., Peikon, I. D., Kebschull, J. M., Shipman, S. L., Mishchenko, Y., Lee, J. H., Kording, K. P., Boyden, E. S., Zador, A. M., and Church, G. M. (2014). Rosetta Brains: A Strategy for Molecularly-Annotated Connectomics. *arXiv [q-bio]*. arXiv: 1404.5103.
- Marder, E. (2012). Neuromodulation of Neuronal Circuits: Back to the Future. *Neuron*, 76(1):1–11. doi:10.1016/j.neuron.2012.09.010.
- Marder, E., O’Leary, T., and Shruti, S. (2014). Neuromodulation of Circuits with Variable Parameters: Single Neurons and Small Circuits Reveal Principles of State-Dependent and Robust Neuromodulation. *Annual Review of Neuroscience*, 37(1):329–346. doi:10.1146/annurev-neuro-071013-013958.
- Maslov, S. and Sneppen, K. (2002). Specificity and stability in topology of protein networks. *Science*, 296(5569):910–913. doi:10.1126/science.1069415.
- May, R. M. (1972). Will a Large Complex System be Stable? *Nature*, 238(5364):413–414. doi:10.1038/238413a0.
- McCulloch, W. S. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4):115–133. doi:10.1007/bf02478259.
- McDiarmid, T. A., Ardiel, E. L., and Rankin, C. H. (2015). The role of neuropeptides in learning and memory in *Caenorhabditis elegans*. *Current Opinion in Behavioral Sciences*, 2:15–20. doi:10.1016/j.cobeha.2014.07.002.
- McDonald, P. W., Hardie, S. L., Jessen, T. N., Carvelli, L., Matthies, D. S., and Blakely, R. D. (2007). Vigorous Motor Activity in *Caenorhabditis elegans* Requires Efficient Clearance of Dopamine Mediated by Synaptic Localization of the Dopamine Transporter DAT-1. *The Journal of Neuroscience*, 27(51):14216–14227. doi:10.1523/jneurosci.2992-07.2007.
- McIntire, S. L., Jorgensen, E., Kaplan, J., and Horvitz, H. R. (1993). The GABAergic nervous system of *Caenorhabditis elegans*. *Nature*, 364(6435):337–341. doi:10.1038/364337a0.
- Medaglia, J. D. and Bassett, D. S. (2017). Network Analyses and Nervous System Disorders. *arXiv*. arXiv: 1701.01101.
- Meinertzhagen, I. A. (2017). Perspective: A New Era of Comparative Connectomics. In *Decoding Neural Circuit Structure and Function*, pages 509–518. Springer, Cham. doi:10.1007/978-3-319-57363-2\_20.
- Melzer, N., Torres-Salazar, D., and Fahlke, C. (2005). A dynamic switch between inhibitory and excitatory currents in a neuronal glutamate transporter. *Proceedings of the National Academy of Sciences of the United States of America*, 102(52):19214–19218. doi:10.1073/pnas.0508837103.

- Menichetti, G., Remondini, D., Panzarasa, P., Mondragón, R. J., and Bianconi, G. (2014). Weighted multiplex networks. *PLOS ONE*, 9(6):e97857. doi:10.1371/journal.pone.0097857.
- Miklós, I. and Podani, J. (2004). Randomization of presence–absence matrices: comments and new algorithms. *Ecology*, 85(1):86–92. doi:10.1890/03-0101.
- Mikula, S. and Denk, W. (2015). High-resolution whole-brain staining for electron microscopic circuit reconstruction. *Nature Methods*, 12(6):541–546. doi:10.1038/nmeth.3361.
- Mills, H., Wragg, R., Hapiak, V., Castelletto, M., Zahratka, J., Harris, G., Summers, P., Korchnak, A., Law, W., Bamber, B., and Komuniecki, R. (2012). Monoamines and neuropeptides interact to inhibit aversive behaviour in *Caenorhabditis elegans*: Monoamines and peptides modulate nociception. *The EMBO Journal*, 31(3):667–678. doi:10.1038/emboj.2011.422.
- Milo, R. (2002). Network Motifs: Simple Building Blocks of Complex Networks. *Science*, 298(5594):824–827. doi:10.1126/science.298.5594.824.
- Milo, R., Kashtan, N., Itzkovitz, S., Newman, M. E., and Alon, U. (2003). On the uniform generation of random graphs with prescribed degree sequences. *arXiv [cond-mat]*. arXiv: 0312028.
- Molloy, M. and Reed, B. (1995). A critical point for random graphs with a given degree sequence. *Random structures & algorithms*, 6(2-3):161–180. doi:10.1002/rsa.3240060204.
- Mondragon, R. J., Iacovacci, J., and Bianconi, G. (2017). Multilink Communities of Multiplex Networks. *arXiv [physics]*. arXiv: 1706.09011.
- Morgan, J. L. and Lichtman, J. W. (2013). Why not connectomics? *Nature Methods*, 10(6):494–500. doi:10.1038/nmeth.2480.
- Mori, I. and Ohshima, Y. (1995). Neural regulation of thermotaxis in *Caenorhabditis elegans*. *Nature*, 376(6538):344–348. doi:10.1038/376344a0.
- Morita, S., Oshio, K.-i., Osana, Y., Funabashi, Y., Oka, K., and Kawamura, K. (2001). Geometrical structure of the neuronal network of *Caenorhabditis elegans*. *Physica A: Statistical Mechanics and its Applications*, 298(3):553–561. doi:10.1016/S0378-4371(01)00266-7.
- Muldoon, S. F., Bridgeford, E. W., and Bassett, D. S. (2016). Small-World Propensity and Weighted Brain Networks. *Scientific Reports*, 6:22057. doi:10.1038/srep22057.
- Murray, J. I., Bao, Z., Boyle, T. J., Boeck, M. E., Mericle, B. L., Nicholas, T. J., Zhao, Z., Sandel, M. J., and Waterston, R. H. (2008). Automated analysis of embryonic

- gene expression with cellular resolution in *C. elegans*. *Nature Methods*, 5(8):703–709. doi:10.1038/nmeth.1228.
- Murray, J. I., Boyle, T. J., Preston, E., Vafeados, D., Mericle, B., Weisdepp, P., Zhao, Z., Bao, Z., Boeck, M., and Waterston, R. H. (2012). Multidimensional regulation of gene expression in the *C. elegans* embryo. *Genome Research*, 22(7):1282–1294. doi:10.1101/gr.131920.111.
- Nathoo, A. N., Moeller, R. A., Westlund, B. A., and Hart, A. C. (2001). Identification of neuropeptide-like protein gene families in *Caenorhabditis elegans* and other species. *Proceedings of the National Academy of Sciences*, 98(24):14000–14005. doi:10.1073/pnas.241231298.
- Nelson, M. D., Janssen, T., York, N., Lee, K. H., Schoofs, L., and Raizen, D. M. (2015). FRPR-4 Is a G-Protein Coupled Neuropeptide Receptor That Regulates Behavioral Quiescence and Posture in *Caenorhabditis elegans*. *PLOS ONE*, 10(11):e0142938. doi:10.1371/journal.pone.0142938.
- Newman, M. E. J. (2002). Assortative Mixing in Networks. *Physical Review Letters*, 89(20). doi:10.1103/PhysRevLett.89.208701.
- Newman, M. E. J. (2003a). Mixing patterns in networks. *Physical Review E*, 67(2). doi:10.1103/PhysRevE.67.026126.
- Newman, M. E. J. (2003b). The Structure and Function of Complex Networks. *SIAM Review*, 45(2):167–256. doi:10.1137/S003614450342480.
- Newman, M. E. J. (2004). Fast algorithm for detecting community structure in networks. *Physical Review E*, 69(6). doi:10.1103/PhysRevE.69.066133.
- Newman, M. E. J. (2005). Power laws, Pareto distributions and Zipf’s law. *Contemporary Physics*, 46(5):323–351. doi:10.1080/00107510500052444.
- Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582. doi:10.1073/pnas.0601602103.
- Newman, M. E. J. (2010). *Networks: an introduction*. Oxford University Press, Oxford, UK. ISBN: 978-0-19-920665-0.
- Newman, M. E. J. and Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, 69(2). doi:10.1103/PhysRevE.69.026113 arXiv: cond-mat/0308217.
- Nguyen, J. P., Shipley, F. B., Linder, A. N., Plummer, G. S., Liu, M., Setru, S. U., Shaevitz, J. W., and Leifer, A. M. (2016). Whole-brain calcium imaging with cellular resolution in freely behaving *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences*, 113(8):E1074–E1081. doi:10.1073/pnas.1507110112.

- Nicosia, V. and Latora, V. (2015). Measuring and modeling correlations in multiplex networks. *Physical Review E*, 92(3). doi:10.1103/PhysRevE.92.032805.
- Nicosia, V., Vertes, P. E., Schafer, W. R., Latora, V., and 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. doi:10.1073/pnas.1300753110.
- Niswender, K. D., Blackman, S. M., Rohde, L., Magnuson, M. A., and Piston, D. W. (1995). Quantitative imaging of green fluorescent protein in cultured cells: Comparison of microscopic techniques, use in fusion proteins and detection limits. *Journal of Microscopy*, 180(2):109–116. doi:10.1111/j.1365-2818.1995.tb03665.x.
- Nunez-Iglesias, J., Kennedy, R., Plaza, S. M., Chakraborty, A., and Katz, W. T. (2014). Graph-based active learning of agglomeration (GALA): a Python library to segment 2d and 3d neuroimages. *Frontiers in Neuroinformatics*, 8:34. doi:10.3389/fninf.2014.00034.
- Oh, S. W., Harris, J. A., Ng, L., Winslow, B., Cain, N., Mihalas, S., Wang, Q., Lau, C., Kuan, L., Henry, A. M., Mortrud, M. T., Ouellette, B., Nguyen, T. N., Sorensen, S. A., Slaughterbeck, C. R., Wakeman, W., Li, Y., Feng, D., Ho, A., Nicholas, E., Hirokawa, K. E., Bohn, P., Joines, K. M., Peng, H., Hawrylycz, M. J., Phillips, J. W., Hohmann, J. G., Wohnoutka, P., Gerfen, C. R., Koch, C., Bernard, A., Dang, C., Jones, A. R., and Zeng, H. (2014). A mesoscale connectome of the mouse brain. *Nature*, 508(7495):207–214. doi:10.1038/nature13186.
- Ohno, N., Katoh, M., Saitoh, Y., and Saitoh, S. (2015). Recent advancement in the challenges to connectomics. *Microscopy*, 65(2):97–107. doi:10.1093/jmicro/dfv371.
- Olde, B. and McCombie, W. R. (1997). Molecular cloning and functional expression of a serotonin receptor from *Caenorhabditis elegans*. *Journal of Molecular Neuroscience*, 8(1):53–62. doi:10.1007/bf02736863.
- Orman, G. and Labatut, V. (2009). A Comparison of Community Detection Algorithms on Artificial Networks. In Gama, J., Costa, V. S., Jorge, A. M., and Brazdil, P. B., editors, *Discovery Science*, volume 5808 of *Lecture Notes in Computer Science*, pages 242–256. Springer, Berlin, Heidelberg. ISBN: 978-3-642-04746-6.
- Oshio, K., Iwasaki, Y., Morita, S., Osana, Y., Gomi, S., Akiyama, E., Omata, K., Oka, K., and Kawamura, K. (2003). Database of Synaptic Connectivity of *C. elegans*. Technical Report of the CCEP 3, Keio University, Tokyo, Japan. <http://ims.dse.ibaraki.ac.jp/ccep/>.
- Osten, P. and Margrie, T. W. (2013). Mapping brain circuitry with a light microscope. *Nature Methods*, 10(6):515–523. doi:10.1038/nmeth.2477.

- Osterreicher, F. and Vajda, I. (2003). A new class of metric divergences on probability spaces and its applicability in statistics. *Annals of the Institute of Statistical Mathematics*, 55(3):639–653. doi:10.1007/bf02517812.
- Owald, D. and Waddell, S. (2015). Olfactory learning skews mushroom body output pathways to steer behavioral choice in *Drosophila*. *Current Opinion in Neurobiology*, 35:178–184. doi:10.1016/j.conb.2015.10.002.
- Passerini, F. and Severini, S. (2011). The von Neumann entropy of networks. *arXiv [cond-mat, physics:quant-ph, q-bio]*. arXiv: 0812.2597.
- Pastor-Satorras, R., Vázquez, A., and Vespignani, A. (2001). Dynamical and Correlation Properties of the Internet. *Physical Review Letters*, 87(25). doi:10.1103/PhysRevLett.87.258701.
- Perc, M., Gómez-Gardeñes, J., Szolnoki, A., Floría, L. M., and Moreno, Y. (2013). Evolutionary dynamics of group interactions on structured populations: A review. *Journal of The Royal Society Interface*, 10(80):20120997–20120997. arXiv: 1301.2247.
- Pereda, A. E., Curti, S., Hoge, G., Cachope, R., Flores, C. E., and Rash, J. E. (2013). Gap junction-mediated electrical transmission: Regulatory mechanisms and plasticity. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1828(1):134–146. doi:10.1016/j.bbamem.2012.05.026.
- Pereira, L., Kratsios, P., Serrano-Saiz, E., Sheftel, H., Mayo, A. E., Hall, D. H., White, J. G., LeBoeuf, B., Garcia, L. R., Alon, U., and Hobert, O. (2015). A cellular and regulatory map of the cholinergic nervous system of *C. elegans*. *eLife*, 4:e12432. doi:10.7554/eLife.12432.
- Petrushin, A., Ferrara, L., and Blau, A. (2015). Past and Recent Endeavours to Simulate *Caenorhabditis elegans*. In *Proceedings of the 3rd International Congress on Neurotechnology, Electronics and Informatics (NEUROTECHNIX 2015)*, pages 133–137, Lisbon, Portugal. SCITEPRESS. doi:10.5220/0005712701330137.
- Petrushin, A., Ferrara, L., and Blau, A. (2016). The Si *elegans* project at the interface of experimental and computational *Caenorhabditis elegans* neurobiology and behavior. *Journal of Neural Engineering*, 13(6):065001. doi:10.1088/1741-2560/13/6/065001.
- Piggott, B., Liu, J., Feng, Z., Wescott, S., and Xu, X. (2011). The Neural Circuits and Synaptic Mechanisms Underlying Motor Initiation in *C. elegans*. *Cell*, 147(4):922–933. doi:10.1016/j.cell.2011.08.053.
- Pirri, J. K., McPherson, A. D., Donnelly, J. L., Francis, M. M., and Alkema, M. J. (2009). A tyramine-gated chloride channel coordinates distinct motor programs of a *Caenorhabditis elegans* escape response. *Neuron*, 62(4):526–538. doi:10.1016/j.neuron.2009.04.013.

- Pocock, R. and Hobert, O. (2010). Hypoxia activates a latent circuit for processing gustatory information in *C. elegans*. *Nature Neuroscience*, 13(5):610–614. doi:10.1038/nn.2537.
- Poisot, T. and Gravel, D. (2014). When is an ecological network complex? Connectance drives degree distribution and emerging network properties. *PeerJ*, 2(e251). doi:10.7717/peerj.251.
- Praitis, V. and Maduro, M. F. (2011). Transgenesis in *C. elegans*. In *Methods in Cell Biology*, volume 106, pages 159–185. Elsevier. doi:10.1016/B978-0-12-544172-8.00006-2.
- Prevedel, R., Yoon, Y.-G., Hoffmann, M., Pak, N., Wetzstein, G., Kato, S., Schrödel, T., Raskar, R., Zimmer, M., Boyden, E. S., and Vaziri, A. (2014). Simultaneous whole-animal 3d imaging of neuronal activity using light-field microscopy. *Nature Methods*, 11(7):727–730. doi:10.1038/nmeth.2964.
- Qian, J., Hintze, A., and Adami, C. (2011). Colored Motifs Reveal Computational Building Blocks in the *C. elegans* Brain. *PLOS ONE*, 6(3):e17013. doi:10.1371/journal.pone.0017013.
- Rabinowitch, I., Chatzigeorgiou, M., and Schafer, W. (2013). A Gap Junction Circuit Enhances Processing of Coincident Mechanosensory Inputs. *Current Biology*, 23(11):963–967. doi:10.1016/j.cub.2013.04.030.
- Ramani, A. K., Chuluunbaatar, T., Verster, A. J., Na, H., Vu, V., Pelte, N., Wannissorn, N., Jiao, A., and Fraser, A. G. (2012). The Majority of Animal Genes Are Required for Wild-Type Fitness. *Cell*, 148(4):792–802. doi:10.1016/j.cell.2012.01.019.
- Ranganathan, R., Cannon, S. C., and Horvitz, H. R. (2000). MOD-1 is a serotonin-gated chloride channel that modulates locomotory behaviour in *C. elegans*. *Nature*, 408(6811):470–475. doi:10.1038/35044083.
- Ranganathan, R., Sawin, E. R., Trent, C., and Horvitz, H. R. (2001). Mutations in the *Caenorhabditis elegans* Serotonin Reuptake Transporter MOD-5 Reveal Serotonin-Dependent and -Independent Activities of Fluoxetine. *The Journal of Neuroscience*, 21(16):5871–5884.
- Ray, J., Pinar, A., and Seshadhri, C. (2012). Are we there yet? When to stop a Markov chain while generating random graphs. In *Algorithms and Models for the Web Graph: 9th International Workshop (WAW 2012)*, pages 153–164, Halifax, NS, Canada. Springer. doi:10.1007/978-3-642-30541-2\_12.
- Reese, T. M., Brzoska, A., Yott, D. T., and Kelleher, D. J. (2012). Analyzing Self-Similar and Fractal Properties of the *C. elegans* Neural Network. *PLOS ONE*, 7(10):e40483. doi:10.1371/journal.pone.0040483.

- Reigl, M., Alon, U., and Chklovskii, D. B. (2004). Search for computational modules in the *C. elegans* brain. *BMC Biology*, 2(1):25. doi:10.1186/1741-7007-2-25.
- Reus, M. A. d. and Heuvel, M. P. v. d. (2013). Rich Club Organization and Intermodule Communication in the Cat Connectome. *The Journal of Neuroscience*, 33(32):12929–12939. doi:10.1523/jneurosci.1448-13.2013.
- Rex, E., Hapiak, V., Hobson, R., Smith, K., Xiao, H., and Komuniecki, R. (2005). TYRA-2 (F01e11.5): a *Caenorhabditis elegans* tyramine receptor expressed in the MC and NSM pharyngeal neurons. *Journal of Neurochemistry*, 94(1):181–191. doi:10.1111/j.1471-4159.2005.03180.x.
- Rex, E., Molitor, S. C., Hapiak, V., Xiao, H., Henderson, M., and Komuniecki, R. (2004). Tyramine receptor (SER-2) isoforms are involved in the regulation of pharyngeal pumping and foraging behavior in *Caenorhabditis elegans*. *Journal of Neurochemistry*, 91(5):1104–1115. doi:10.1111/j.1471-4159.2004.02787.x.
- Rice, M. E. and Cragg, S. J. (2008). Dopamine Spillover after Quantal Release: Rethinking Dopamine Transmission in the Nigrostriatal Pathway. *Brain Research Reviews*, 58(2):303–313. doi:10.1016/j.brainresrev.2008.02.004.
- Rice, M. E., Patel, J. C., and Cragg, S. J. (2011). Dopamine release in the basal ganglia. *Neuroscience*, 198:112–137. doi:10.1016/j.neuroscience.2011.08.066.
- Richmond, J. (2007). Synaptic function. *WormBook*, doi:10.1895/wormbook.1.69.1.
- Ringstad, N., Abe, N., and Horvitz, H. R. (2009). Ligand-gated chloride channels are receptors for biogenic amines in *C. elegans*. *Science*, 325(5936):96–100. doi:10.1126/science.1169243.
- Ringstad, N. and Horvitz, H. R. (2008). FMRFamide neuropeptides and acetylcholine synergistically inhibit egg-laying by *C. elegans*. *Nature Neuroscience*, 11(10):1168–1176. doi:10.1038/nn.2186.
- Rivard, L., Srinivasan, J., Stone, A., Ochoa, S., Sternberg, P. W., and Loer, C. M. (2010). A comparison of experience-dependent locomotory behaviors and biogenic amine neurons in nematode relatives of *Caenorhabditis elegans*. *BMC Neuroscience*, 11:22. doi:10.1186/1471-2202-11-22.
- Roberts, W. M., Augustine, S. B., Lawton, K. J., Lindsay, T. H., Thiele, T. R., Izquierdo, E. J., Faumont, S., Lindsay, R. A., Britton, M. C., Pokala, N., Bargmann, C. I., and Lockery, S. R. (2016). A stochastic neuronal model predicts random search behaviors at multiple spatial scales in *C. elegans*. *eLife*, 5:e12572. doi:10.7554/eLife.12572.
- Rogers, C., Reale, V., Kim, K., Chatwin, H., Li, C., Evans, P., and de Bono, M. (2003). Inhibition of *Caenorhabditis elegans* social feeding by FMRFamide-related peptide activation of NPR-1. *Nature Neuroscience*, 6(11):1178–1185. doi:10.1038/nn1140.

- Roshchina, V. V. (2010). Evolutionary Considerations of Neurotransmitters in Microbial, Plant, and Animal Cells. In Lyte, M. and Freestone, P. P. E., editors, *Microbial Endocrinology*, pages 17–52. Springer New York. doi:10.1007/978-1-4419-5576-0\_2.
- Rubinov, M. and Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, 52(3):1059–1069. doi:10.1016/j.neuroimage.2009.10.003.
- Rudolph-Lilith, M. and Muller, L. E. (2014). Aspects of randomness in neural graph structures. *Biological Cybernetics*, 108(4):381–396. doi:10.1007/s00422-014-0606-6.
- Rumelhart, D. E. and McClelland, J. L. (1986). *Parallel Distributed Processing*, volume 1. MIT Press, Cambridge, MA. ISBN: 978-0-262-68053-0.
- San-Miguel, A. and Lu, H. (2013). Microfluidics as a tool for *C. elegans* research. *WormBook*, doi:10.1895/wormbook.1.162.1:1–19.
- Sandberg, A. and Bostrom, N. (2008). Whole Brain Emulation: A Roadmap. Technical Report 2008-3, Future of Humanity Institute, Oxford University, Oxford, UK. <http://www.fhi.ox.ac.uk/brain-emulation-roadmap-report.pdf>.
- Sanyal, S., Wintle, R. F., Kindt, K. S., Nuttley, W. M., Arvan, R., Fitzmaurice, P., Bigras, E., Merz, D. C., Hébert, T. E., van der Kooy, D., and others (2004). Dopamine modulates the plasticity of mechanosensory responses in *Caenorhabditis elegans*. *The EMBO Journal*, 23(2):473–482. doi:10.1038/sj.emboj.7600057.
- Sawin, E. R. (1996). *Genetic and cellular analysis of modulated behaviors in Caenorhabditis elegans*. PhD thesis, MIT, Cambridge, MA. <http://dspace.mit.edu/handle/1721.1/7582>.
- Sawin, E. R., Ranganathan, R., and Horvitz, H. R. (2000). *C. elegans* Locomotory Rate Is Modulated by the Environment through a Dopaminergic Pathway and by Experience through a Serotonergic Pathway. *Neuron*, 26(3):619–631. doi:10.1016/S0896-6273(00)81199-x.
- Schafer, W. (2005). Egg-laying. *WormBook*, doi:10.1895/wormbook.1.38.1.
- Scheffer, L. K., Karsh, B., and Vitaladevun, S. (2013). Automated Alignment of Imperfect EM Images for Neural Reconstruction. *arXiv [q-bio]*. arXiv: 1304.6034.
- Schmidhuber, J. (2015). Deep Learning in Neural Networks: An Overview. *Neural Networks*, 61:85–117. doi:10.1016/j.neunet.2014.09.003 arXiv: 1404.7828.
- Schneider, C. M., Andrade, R. F. S., Shinbrot, T., and Herrmann, H. J. (2011a). The fragility of protein-protein interaction networks. *EPL (Europhysics Letters)*, 95(1):16001. doi:10.1209/0295-5075/95/16001.

- Schneider, C. M., Moreira, A. A., Andrade, J. S., Havlin, S., and Herrmann, H. J. (2011b). Mitigation of malicious attacks on networks. *Proceedings of the National Academy of Sciences*, 108(10):3838–3841. doi:10.1073/pnas.1009440108.
- Schneider-Mizell, C. M., Gerhard, S., Longair, M., Kazimiers, T., Li, F., Zwart, M. F., Champion, A., Midgley, F. M., Fetter, R. D., Saalfeld, S., and Cardona, A. (2016). Quantitative neuroanatomy for connectomics in *Drosophila*. *eLife*, 5:e12059. doi:10.7554/eLife.12059.
- Schreiber, F. and Schwöbbermeyer, H. (2008). Motifs in biological networks. *Statistical and Evolutionary Analysis of Biological Networks*, pages 45–64.
- Schrödinger, E. (1958). *Mind and matter*. Cambridge University Press, Cambridge, UK.
- Schwarz, R. F., Branicky, R., Grundy, L. J., Schafer, W. R., and Brown, A. E. X. (2015). Changes in Postural Syntax Characterize Sensory Modulation and Natural Variation of *C. elegans* Locomotion. *PLOS Computational Biology*, 11(8):e1004322. doi:10.1371/journal.pcbi.1004322.
- Sengupta, P. and Samuel, A. D. (2009). *Caenorhabditis elegans*: a model system for systems neuroscience. *Current Opinion in Neurobiology*, 19(6):637–643. doi:10.1016/j.conb.2009.09.009.
- Serrano-Saiz, E., Poole, R. J., Felton, T., Zhang, F., De La Cruz, E. D., and Hobert, O. (2013). Modular control of glutamatergic neuronal identity in *C. elegans* by distinct homeodomain proteins. *Cell*, 155(3):659–673. doi:10.1016/j.cell.2013.09.052.
- Shaham, S. (2006). Methods in cell biology. *WormBook*, doi:10.1895/wormbook.1.49.1.
- Shahidi, R., Williams, E. A., Conzelmann, M., Asadulina, A., Verasztó, C., Jasek, S., Bezares-Calderón, L. A., and Jékely, G. (2015). A serial multiplex immunogold labeling method for identifying peptidergic neurons in connectomes. *eLife*, 4:e11147. doi:10.7554/eLife.11147.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423. doi:10.1002/j.1538-7305.1948.tb01338.x.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*, 13(11):2498–2504. doi:10.1101/gr.1239303.
- Sherrington, C. S. (1906a). *The integrative action of the nervous system*, volume xvi of *Yale University Mrs. Hepsa Ely Silliman memorial lectures*. Yale University Press, New Haven, CT.

- Sherrington, C. S. (1906b). Observations on the scratch-reflex in the spinal dog. *The Journal of Physiology*, 34(1-2):1–50. doi:0.1113/jphysiol.1906.sp001139.
- Shih, C.-T., Sporns, O., Yuan, S.-L., Su, T.-S., Lin, Y.-J., Chuang, C.-C., Wang, T.-Y., Lo, C.-C., Greenspan, R., and Chiang, A.-S. (2015). Connectomics-Based Analysis of Information Flow in the Drosophila Brain. *Current Biology*, 25(10):1249–1258. doi:10.1016/j.cub.2015.03.021.
- Shiple, F. B., Clark, C. M., Alkema, M. J., and Leifer, A. M. (2014). Simultaneous optogenetic manipulation and calcium imaging in freely moving *C. elegans*. *Frontiers in Neural Circuits*, 8:28. doi:10.3389/fncir.2014.00028.
- Shyn, S. I. (2003). *Cameleon Reveals a Physiologic Correlate for Alternative Behavioral States in Caenorhabditis elegans Egg-laying*. PhD thesis, University of California. <http://www2.mrc-lmb.cam.ac.uk/groups/wschafer/StanleyShyn.pdf>.
- Simonsen, K. T., Moerman, D. G., and Naus, C. C. (2014). Gap junctions in *C. elegans*. *Frontiers in Physiology*, 5:40. doi:10.3389/fphys.2014.00040.
- Sloane, N. J. A. (2017). Number of connected digraphs with n nodes (A003085). In *The On-Line Encyclopedia of Integer*. <https://oeis.org/A003085>.
- Smith, C. J., Watson, J. D., Spencer, W. C., O'Brien, T., Cha, B., Albeg, A., Treinin, M., and Miller III, D. M. (2010). Time-lapse imaging and cell-specific expression profiling reveal dynamic branching and molecular determinants of a multi-dendritic nociceptor in *C. elegans*. *Developmental Biology*, 345(1):18–33. doi:10.1016/j.ydbio.2010.05.502.
- Soboleski, M. R., Oaks, J., and Halford, W. P. (2005). Green fluorescent protein is a quantitative reporter of gene expression in individual eukaryotic cells. *The FASEB Journal. Federation of American Societies for Experimental Biology*, 19(3):440–442. doi:10.1096/fj.04-3180fje.
- Sohl, G., Maxeiner, S., and Willecke, K. (2005). Expression and functions of neuronal gap junctions. *Nature Reviews Neuroscience*, 6(3):191–200. doi:10.1038/nrn1627.
- Solomonoff, R. (1952). An exact method for the computation of the connectivity of random nets. *The bulletin of mathematical biophysics*, 14(2):153–157. doi:10.1007/bf02477714.
- Solomonoff, R. and Rapoport, A. (1951). Connectivity of random nets. *The bulletin of mathematical biophysics*, 13(2):107–117. doi:10.1007/bf02478357.
- Soman, S., Jayadeva, and Suri, M. (2016). Recent trends in neuromorphic engineering. *Big Data Analytics*, 1(15). doi:10.1186/s41044-016-0013-1.

- Sommer, F. T. and Wennekers, T. (2003). Models of distributed associative memory networks in the brain. *Theory in Biosciences*, 122(1):55–69. doi:10.1007/s12064-003-0037-8.
- Song, S., Sjöström, P. J., Reigl, M., Nelson, S., and Chklovskii, D. B. (2005). Highly Nonrandom Features of Synaptic Connectivity in Local Cortical Circuits. *PLOS Biology*, 3(3):e68. doi:10.1371/journal.pbio.0030068.
- Sorrentino, F., di Bernardo, M., Cuéllar, G. H., and Boccaletti, S. (2006). Synchronization in weighted scale-free networks with degree–degree correlation. *Physica D: Nonlinear Phenomena*, 224(1–2):123–129. doi:10.1016/j.physd.2006.09.030.
- Sporns, O. (2013a). The human connectome: Origins and challenges. *NeuroImage*, 80:53–61. doi:10.1016/j.neuroimage.2013.03.023.
- Sporns, O. (2013b). Network attributes for segregation and integration in the human brain. *Current Opinion in Neurobiology*, 23(2):162–171. doi:10.1016/j.conb.2012.11.015.
- Sporns, O. (2015). Cerebral cartography and connectomics. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 370(1668):20140173. doi:10.1098/rstb.2014.0173.
- Sporns, O. and Betzel, R. F. (2016). Modular Brain Networks. *Annual Review of Psychology*, 67(1):613–640. doi:10.1146/annurev-psych-122414-033634.
- Sporns, O. and Kötter, R. (2004). Motifs in Brain Networks. *PLOS Biology*, 2(11):e369. doi:10.1371/journal.pbio.0020369.
- Sporns, O., Tononi, G., and Kötter, R. (2005). The Human Connectome: A Structural Description of the Human Brain. *PLOS Computational Biology*, 1(4):e42. doi:10.1371/journal.pcbi.0010042.
- Stam, C. J. and Reijneveld, J. C. (2007). Graph theoretical analysis of complex networks in the brain. *Nonlinear Biomedical Physics*, 1(1):3. doi:10.1186/1753-4631-1-3.
- Starich, T. A., Lee, R. Y., Panzarella, C., Avery, L., and Shaw, J. E. (1996). *eat-5* and *unc-7* represent a multigene family in *Caenorhabditis elegans* involved in cell-cell coupling. *The Journal of Cell Biology*, 134(2):537–548. doi:10.1083/jcb.134.2.537.
- Strang, A., Haynes, O., Cahill, N. D., and Narayan, D. A. (2017). Relationships Between Characteristic Path Length, Efficiency, Clustering Coefficients, and Graph Density. *arXiv*. arXiv: 1702.02621.
- Strogatz, S. H. (2001). Exploring complex networks. *Nature*, 410(6825):268–276. doi:10.1038/35065725.

- Sugiura, M., Fuke, S., Suo, S., Sasagawa, N., Van Tol, H. H. M., and Ishiura, S. (2005). Characterization of a novel D2-like dopamine receptor with a truncated splice variant and a D1-like dopamine receptor unique to invertebrates from *Caenorhabditis elegans*. *Journal of Neurochemistry*, 94(4):1146–1157. doi:10.1111/j.1471-4159.2005.03268.x.
- Sulston, J., Dew, M., and Brenner, S. (1975). Dopaminergic neurons in the nematode *Caenorhabditis elegans*. *The Journal of Comparative Neurology*, 163(2):215–226. doi:10.1002/cne.901630207.
- Sulston, J. E. and Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Developmental Biology*, 56(1):110–156. doi:10.1016/0012-1606(77)90158-0.
- Sulston, J. E., Schierenberg, E., White, J. G., and Thomson, J. N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Developmental Biology*, 100(1):64–119. doi:10.1016/0012-1606(83)90201-4.
- Suo, S., Culotti, J. G., and Van Tol, H. H. M. (2009). Dopamine counteracts octopamine signalling in a neural circuit mediating food response in *C. elegans*. *The EMBO Journal*, 28(16):2437–2448. doi:10.1038/emboj.2009.194.
- Suo, S., Kimura, Y., and Tol, H. H. M. V. (2006). Starvation Induces cAMP Response Element-Binding Protein-Dependent Gene Expression through Octopamine–Gq Signaling in *Caenorhabditis elegans*. *The Journal of Neuroscience*, 26(40):10082–10090. doi:10.1523/jneurosci.0819-06.2006.
- Suo, S., Sasagawa, N., and Ishiura, S. (2002). Identification of a dopamine receptor from *Caenorhabditis elegans*. *Neuroscience Letters*, 319(1):13–16. doi:10.1016/S0304-3940(01)02477-6.
- Suo, S., Sasagawa, N., and Ishiura, S. (2003). Cloning and characterization of a *Caenorhabditis elegans* D2-like dopamine receptor. *Journal of Neurochemistry*, 86(4):869–878. doi:10.1046/j.1471-4159.2003.01896.x.
- Swierczek, N. A., Giles, A. C., Rankin, C. H., and Kerr, R. A. (2011). High-throughput behavioral analysis in *C. elegans*. *Nature Methods*, 8(7):592–598. doi:10.1038/nmeth.1625.
- Syková, E. and Nicholson, C. (2008). Diffusion in Brain Extracellular Space. *Physiological Reviews*, 88(4):1277–1340. doi:10.1152/physrev.00027.2007.
- Sze, J. Y., Victor, M., Loer, C., Shi, Y., and Ruvkun, G. (2000). Food and metabolic signalling defects in a *Caenorhabditis elegans* serotonin-synthesis mutant. *Nature*, 403(6769):560–564. doi:10.1038/35000609.
- Szigeti, B., Gleeson, P., Vella, M., Khayrulin, S., Palyanov, A., Hokanson, J., Currie, M., Cantarelli, M., Idili, G., and Larson, S. (2014). OpenWorm: an open-science approach

- to modeling *Caenorhabditis elegans*. *Frontiers in Computational Neuroscience*, 8:137. doi:10.3389/fncom.2014.00137.
- Taber, K. H. and Hurley, R. A. (2014). Volume transmission in the brain: beyond the synapse. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 26(1):iv–4. doi:10.1176/appi.neuropsych.13110351.
- Tanaka, D., Furusawa, K., Kameyama, K., Okamoto, H., and Doi, M. (2007). Melatonin signaling regulates locomotion behavior and homeostatic states through distinct receptor pathways in *Caenorhabditis elegans*. *Neuropharmacology*, 53(1):157–168. doi:10.1016/j.neuropharm.2007.04.017.
- Tang, Y., Gao, H., Zou, W., and Kurths, J. (2012). Identifying Controlling Nodes in Neuronal Networks in Different Scales. *PLOS ONE*, 7(7):e41375. doi:10.1371/journal.pone.0041375.
- Teller, S., Granell, C., Domenico, M. D., Soriano, J., Gómez, S., and Arenas, A. (2014). Emergence of Assortative Mixing between Clusters of Cultured Neurons. *PLOS Computational Biology*, 10(9):e1003796. doi:10.1371/journal.pcbi.1003796.
- Tesla, N. (1897). On Electricity. *Electrical Review*, 30(4):46–47.
- Tieri, P., Valensin, S., Latora, V., Castellani, G. C., Marchiori, M., Remondini, D., and Franceschi, C. (2005). Quantifying the relevance of different mediators in the human immune cell network. *Bioinformatics*, 21(8):1639–1643. doi:10.1093/bioinformatics/bti239.
- Tononi, G., Sporns, O., and Edelman, G. M. (1994). A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proceedings of the National Academy of Sciences*, 91(11):5033–5037.
- Towlson, E. K., Vertes, P. E., Ahnert, S. E., Schafer, W. R., and Bullmore, E. T. (2013). The Rich Club of the *C. elegans* Neuronal Connectome. *Journal of Neuroscience*, 33(15):6380–6387. 10.1523/jneurosci.3784-12.2013.
- Trueta, C. and De-Miguel, F. F. (2012). Extrasynaptic exocytosis and its mechanisms: a source of molecules mediating volume transmission in the nervous system. *Frontiers in Physiology*, 3:319. doi:10.3389/fphys.2012.00319.
- Tsalik, E. L. and Hobert, O. (2003). Functional mapping of neurons that control locomotory behavior in *Caenorhabditis elegans*. *Journal of Neurobiology*, 56(2):178–197. doi:10.1002/neu.10245.
- Tsalik, E. L., Niacaris, T., Wenick, A. S., Pau, K., Avery, L., and Hobert, O. (2003). LIM homeobox gene-dependent expression of biogenic amine receptors in restricted regions of the *C. elegans* nervous system. *Developmental Biology*, 263(1):81–102. doi:10.1016/S0012-1606(03)00447-0.

- Vaiana, M. and Muldoon, S. (2017). Multilayer Brain Networks. *arXiv [physics, q-bio]*. arXiv:1709.02325.
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., and Ugurbil, K. (2013). The WU-Minn Human Connectome Project: An overview. *NeuroImage*, 80:62–79. doi:10.1016/j.neuroimage.2013.05.041.
- Varshney, L. R., Chen, B. L., Paniagua, E., Hall, D. H., and Chklovskii, D. B. (2011). Structural Properties of the *Caenorhabditis elegans* Neuronal Network. *PLOS Computational Biology*, 7(2):e1001066. doi:10.1371/journal.pcbi.1001066.
- Virkar, Y. and Clauset, A. (2014). Power-law distributions in binned empirical data. *The Annals of Applied Statistics*, 8(1):89–119. doi:10.1214/13-AOAS710.
- Vizi, E., Fekete, A., Karoly, R., and Mike, A. (2010). Non-synaptic receptors and transporters involved in brain functions and targets of drug treatment. *British Journal of Pharmacology*, 160(4):785–809. doi:10.1111/j.1476-5381.2009.00624.x.
- Volman, V., Perc, M., and Bazhenov, M. (2011). Gap Junctions and Epileptic Seizures – Two Sides of the Same Coin? *PLOS ONE*, 6(5):e20572. doi:10.1371/journal.pone.0020572.
- Von Neumann, J. (1955). *Mathematical foundations of quantum mechanics*. Princeton University Press, Princeton, NJ. ISBN: 9780691028934.
- Vértes, P. E. and 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. doi:10.1111/jcpp.12365.
- Vértes, P. E., Nicol, R. M., Chapman, S., Watkins, N., Robertson, D. A., and Bullmore, E. T. (2011). Topological Isomorphisms of Human Brain and Financial Market Networks. *Frontiers in Systems Neuroscience*, 5. doi:10.3389/fnsys.2011.00075.
- Wagner, A. and Fell, D. A. (2001). The small world inside large metabolic networks. *Proceedings of the Royal Society B: Biological Sciences*, 268(1478):1803–1810. doi:10.1098/rspb.2001.1711.
- Walker, R. J., Brooks, H. L., and Holden-Dye, L. (1996). Evolution and overview of classical transmitter molecules and their receptors. *Parasitology*, 113(Supplement S1):S3–S33. doi:10.1017/S0031182000077878.
- Walsh, T. (1999). Search in a small world. In *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI '99)*, pages 1172–1177, Stockholm, Sweden. Morgan Kaufmann Publishers Inc., San Francisco, CA. ISBN: 1-55860-613-0.
- Wang, J. and Barr, M. M. (2016). Ciliary Extracellular Vesicles: Txt Msg Organelles. *Cellular and Molecular Neurobiology*, 36(3):449–457. doi:10.1007/s10571-016-0345-4.

- Wang, J., Li, T., Sabel, B. A., Chen, Z., Wen, H., Li, J., Xie, X., Yang, D., Chen, W., Wang, N., Xian, J., and He, H. (2016). Structural brain alterations in primary open angle glaucoma: a 3t MRI study. *Scientific Reports*, 6:18969. doi:10.1038/srep18969.
- Wang, S., Szalay, M. S., Zhang, C., and Csermely, P. (2008). Learning and Innovative Elements of Strategy Adoption Rules Expand Cooperative Network Topologies. *PLOS ONE*, 3(4):e1917. doi:10.1371/journal.pone.0001917.
- Wang, S.-J., Xu, X.-J., Wu, Z.-X., and Wang, Y.-H. (2006). Effects of degree distribution in mutual synchronization of neural networks. *Physical Review E*, 74(4):041915. doi:10.1103/PhysRevE.74.041915.
- Wang, X. and Liu, J. (2017). A layer reduction based community detection algorithm on multiplex networks. *Physica A: Statistical Mechanics and its Applications*, 471:244–252. doi:10.1016/j.physa.2016.11.036.
- Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, 58(301):236. doi:10.2307/2282967.
- Wasserman, S. and Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge University Press. ISBN: 0521387078.
- Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684):440–442. doi:10.1038/30918.
- Weiss (2010). Field effects in the CNS play functional roles. *Frontiers in Neural Circuits*, 4(15). doi:10.3389/fncir.2010.00015.
- Wenick, A. S. and Hobert, O. (2004). Genomic cis-regulatory architecture and trans-acting regulators of a single interneuron-specific gene battery in *C. elegans*. *Developmental cell*, 6(6):757–770. doi:10.1016/j.devcel.2004.05.004.
- Wernicke, S. (2005). A faster algorithm for detecting network motifs. *5th International Workshop on Algorithms in Bioinformatics, Lecture Notes in Bioinformatics*, 3692:165–177. doi:10.1007/11557067\_14.
- Wernicke, S. and Rasche, F. (2006). FANMOD: a tool for fast network motif detection. *Bioinformatics*, 22(9):1152–1153. doi:10.1093/bioinformatics/btl038.
- White, J. G., Southgate, E., Thomson, J. N., and Brenner, S. (1983). Factors that determine connectivity in the nervous system of *Caenorhabditis elegans*. *Cold Spring Harbor Symposia on Quantitative Biology*, 48 Pt 2:633–640.
- White, J. G., Southgate, E., Thomson, J. N., and Brenner, S. (1986). The Structure of the Nervous System of the Nematode *Caenorhabditis elegans*. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 314(1165):1–340. doi:10.1098/rstb.1986.0056.

- White, J. Q., Nicholas, T. J., Gritton, J., Truong, L., Davidson, E. R., and Jorgensen, E. M. (2007). The Sensory Circuitry for Sexual Attraction in *C. elegans* Males. *Current Biology*, 17(21):1847–1857. doi:10.1016/j.cub.2007.09.011.
- Williams, R. W. and Herrup, K. (1988). The Control of Neuron Number. *Annual Review of Neuroscience*, 11(1):423–453. doi:10.1146/annurev.ne.11.030188.002231.
- Wong, E., Baur, B., Quader, S., and Huang, C.-H. (2012). Biological network motif detection: principles and practice. *Briefings in Bioinformatics*, 13(2):202–215. doi:10.1093/bib/bbr033.
- Wragg, R. T., Hapiak, V., Miller, S. B., Harris, G. P., Gray, J., Komuniecki, P. R., and Komuniecki, R. W. (2007). Tyramine and Octopamine Independently Inhibit Serotonin-Stimulated Aversive Behaviors in *Caenorhabditis elegans* through Two Novel Amine Receptors. *The Journal of Neuroscience*, 27(49):13402–13412. doi:10.1523/jneurosci.3495-07.2007.
- Xiao, H., Hapiak, V. M., Smith, K. A., Lin, L., Hobson, R. J., Plenefisch, J., and Komuniecki, R. (2006). SER-1, a *Caenorhabditis elegans* 5-HT<sub>2</sub>-like receptor, and a multi-PDZ domain containing protein (MPZ-1) interact in vulval muscle to facilitate serotonin-stimulated egg-laying. *Developmental Biology*, 298(2):379–391. doi:10.1016/j.ydbio.2006.06.044.
- Xu, K., Tavernarakis, N., and Driscoll, M. (2001). Necrotic Cell Death in *C. elegans* Requires the Function of Calreticulin and Regulators of Ca<sup>2+</sup> Release from the Endoplasmic Reticulum. *Neuron*, 31(6):957–971. doi:10.1016/S0896-6273(01)00432-9.
- Xu, M., Jarrell, T. A., Wang, Y., Cook, S. J., Hall, D. H., and Emmons, S. W. (2013). Computer Assisted Assembly of Connectomes from Electron Micrographs: Application to *Caenorhabditis elegans*. *PLOS ONE*, 8(1):e54050. doi:10.1371/journal.pone.0054050.
- Yan, G., Vértés, P. E., Towlson, E. K., Chew, Y. L., Walker, D. S., Schafer, W. R., and Barabási, A.-L. (2017). Network control principles predict neuron function in the *Caenorhabditis elegans* connectome. *Nature*, 550(7677):519–523. doi:10.1038/nature24056.
- Yang, Z., Algesheimer, R., and Tessone, C. J. (2016). A Comparative Analysis of Community Detection Algorithms on Artificial Networks. *Scientific Reports*, 6:srep30750. doi:10.1038/srep30750.
- Yemini, E., Jucikas, T., Grundy, L. J., Brown, A. E. X., and Schafer, W. R. (2013). A database of *Caenorhabditis elegans* behavioral phenotypes. *Nature Methods*, 10(9):877–879. doi:10.1038/nmeth.2560.
- Yook, S.-H., Oltvai, Z. N., and Barabási, A.-L. (2004). Functional and topological characterization of protein interaction networks. *Proteomics*, 4(4):928–942. doi:10.1002/pmic.200300636.

- Yoshida, M., Oami, E., Wang, M., Ishiura, S., and Suo, S. (2014). Nonredundant function of two highly homologous octopamine receptors in food-deprivation-mediated signaling in *Caenorhabditis elegans*. *Journal of Neuroscience Research*, 92(5):671–678. doi:10.1002/jnr.23345.
- Yuan, Z., Zhao, C., Di, Z., Wang, W.-X., and Lai, Y.-C. (2013). Exact controllability of complex networks. *Nature Communications*, 4(2447). doi:10.1038/ncomms3447.
- Zador, A. M., Dubnau, J., Oyibo, H. K., Zhan, H., Cao, G., and Peikon, I. D. (2012). Sequencing the Connectome. *PLOS Biology*, 10(10):e1001411. doi:10.1371/journal.pbio.1001411.
- Zhang, F., Bhattacharya, A., Nelson, J. C., Abe, N., Gordon, P., Lloret-Fernandez, C., Maicas, M., Flames, N., Mann, R. S., Colón-Ramos, D. A., and Hobert, O. (2014). The LIM and POU homeobox genes *ttx-3* and *unc-86* act as terminal selectors in distinct cholinergic and serotonergic neuron types. *Development*, 141(2):422–435. doi:10.1242/dev.099721.
- Zhang, M., Chung, S. H., Fang-Yen, C., Craig, C., Kerr, R. A., Suzuki, H., Samuel, A. D., Mazur, E., and Schafer, W. R. (2008). A Self-Regulating Feed-Forward Circuit Controlling *C. elegans* Egg-Laying Behavior. *Current Biology*, 18(19):1445–1455. doi:10.1016/j.cub.2008.08.047.
- Zhang, M., Schafer, W. R., and Breitling, R. (2010). A circuit model of the temporal pattern generator of *Caenorhabditis* egg-laying behavior. *BMC Systems Biology*, 4:81. doi:10.1186/1752-0509-4-81.
- Zhang, Y., Lu, H., and Bargmann, C. I. (2005). Pathogenic bacteria induce aversive olfactory learning in *Caenorhabditis elegans*. *Nature*, 438(7065):179–184. doi:10.1038/nature04216.
- Zhang, Z., Zhou, S., and Zou, T. (2007). Self-similarity, small-world, scale-free scaling, disassortativity, and robustness in hierarchical lattices. *The European Physical Journal B*, 56(3):259–271. doi:10.1140/epjb/e2007-00107-6.
- Zhen, M. and Samuel, A. D. (2015). *C. elegans* locomotion: small circuits, complex functions. *Current Opinion in Neurobiology*, 33:117–126. doi:10.1016/j.conb.2015.03.009.
- Zhou, H. (2002). Scaling exponents and clustering coefficients of a growing random network. *Physical Review E*, 66(1). doi:10.1103/PhysRevE.66.016125.
- Zhou, S. and Mondragon, R. J. (2004). The rich-club phenomenon in the Internet topology. *IEEE Communications Letters*, 8(3):180–182. doi:10.1109/lcomm.2004.823426.



# Appendices



# Appendix A

## Monoamine expression patterns

## A.1 Monoamine expression

Marker	WormBase ID	Neurons	Reference
<i>tph-1</i>	Expr959	RIH <sup>†</sup> , AIM <sup>†</sup> , ADF, NSM, HSN	Sze et al. (2000)
	Expr12176	ASG <sup>†</sup>	Pocock and Hobert (2010)
<i>mod-5</i>	Expr9350	AIM <sup>†</sup> , NSM, ADF, RIH <sup>†</sup>	Jafari et al. (2011)
Immuno-staining	N/A	VC04 <sup>†</sup> , VC05 <sup>†</sup>	Duerr et al. (1999)
	N/A	PHB <sup>†</sup> , I5 <sup>†</sup>	Sawin et al. (2000)

Table A.1 Serotonin (5-HT) expressing cells. Cells with weak or conditional expression are marked †.

Marker	WormBase ID	Neurons	Reference
<i>cat-2</i>	Expr2619	ADE, PDE, CEP	Suo et al. (2003)
<i>dat-1</i>	Expr8327	ADE, PDE, CEP	McDonald et al. (2007)

Table A.2 Dopamine (DA) expressing cells.

Marker	WormBase ID	Neurons	Reference
<b>Octopamine</b> <i>tbh-1</i>	Expr3721	RIC	Alkema et al. (2005)
<b>Tyramine</b> <i>tdc-1</i>	Expr3722	RIM, RIC*	Alkema et al. (2005)

Table A.3 Octopamine (OA) & tyramine (TA) expressing cells. \*RIC is excluded from the TA network due to co-expression of *tbh-1* which converts TA to OA.

## A.2 Receptor expression

Receptor	WormBase ID	Neurons	Reference
<i>ser-1</i>	Expr7825	RIA, RIC, PVT, DVC, URY	Dernovici et al. (2007)
	Expr8282	PVQ	Carnell (2005)
	Expr3962	RMD, RMF, RMH	Xiao et al. (2006)
<i>ser-4</i>	Expr2710	RIB, PVT, DVC, DVA, RIS	Tsalik et al. (2003)
	Expr10554	AIB, NSM	Gürel et al. (2012)
	N/A	M1, RIM	Shyn (2003)
<i>ser-5</i>	Expr12174	ASH, AWB	Harris et al. (2009)
	Expr12172	AVJ	Cunningham et al. (2012)
<i>ser-7</i>	Expr3759	MC, M2, M3, M4, M5, I2, I3, I4, I6	Hobson et al. (2006)
<i>mod-1</i>	Expr10023	RIM, RID, RIC, AIZ, AIY, AIB, AIA	Li et al. (2012)
	Expr10553	RME, DD, VD	Gürel et al. (2012)

Table A.4 Serotonin (5-HT) receptor expression patterns.

Receptor	WormBase ID	Neurons	Reference
<i>octr-1</i>	Expr7846	ASH, ASI, AIY, ADE, CEP	Wragg et al. (2007)
<i>ser-3</i>	Expr8275	PVQ, PHB, PHA, SIA	Suo et al. (2006)
	Expr10640	ASH	Mills et al. (2012)
<i>ser-6</i>	Expr10641	AWB, ASI, ADL	Mills et al. (2012)
	Expr11709	RIC, SIA	Yoshida et al. (2014)

Table A.5 Octopamine (OA) receptor expression patterns.

Receptor	WormBase ID	Neurons	Reference
<i>dop-1</i>	Expr2882	AUA, RIM, ALM, RIB, PLM, PHC	Sanyal et al. (2004)
	Expr2708	AVM, ALN, PVQ, PLN, RIS	Tsalik et al. (2003)
	Expr3047	PVD, VA, VB, AS, DA, DB	Chase et al. (2004)
<i>dop-2</i>	Expr2618	ADE, PDE, CEP	Suo et al. (2003)
	Expr2709	RID, RIA, PDA, SIB, SIA	Tsalik et al. (2003)
<i>dop-3</i>	Expr3048	PVD, VA, VB, AS, DA, DB, DD, VD	Chase et al. (2004)
	Expr7939	ASE	Etchberger et al. (2007)
	Expr8667	RIC, SIA	Suo et al. (2009)
	Expr11452	NSM	Zhang et al. (2014)
	Expr12177	ASK	Ezak and Ferkey (2010)
<i>dop-4</i>	Expr3687	AVL, ASG, PQR, I2, I1, CAN	Sugiura et al. (2005)
<i>dop-5</i>	Expr7939 N/A	ASE MI, M5, BDU, RIB, PHA, PHB, DVA, AIM, ADA, DVC, ASI, RMG, PVC, RIF, URX, AIY, PVT	Etchberger et al. (2007) (Appendix C)
<i>dop-6</i>	Expr11993 N/A	OLL AUA, RID, RMD, RIB, ASI, PHA, IL2, PVQ, URA, AVF, ADF, RIH, URX	Smith et al. (2010) (Appendix C)
<i>lgc-53</i>	N/A	HSN, PVD, CAN, IL2, VA, AIM, FLP, AVF, URY	(Appendix C)

Table A.6 Dopamine (DA) receptor expression patterns.

Receptor	WormBase ID	Neurons	Reference
<i>ser-2</i>	Expr2707	BDU, AVH, AUA, ALN, RID, RIC, AIZ, RIA, AIY, PVT, PVD, PVC, OLL, NSM, LUA, DVA, DA09, CAN, SIA, SDQ, SAB, RME	Tsalik et al. (2003)
	Expr3206	PVD	Rex et al. (2004)
	Expr10758	VD	Donnelly et al. (2013)
<i>tyra-2</i>	Expr3415	ASI, ASH, ASG, ASE, ALM, PVD, NSM, MC, CAN	Rex et al. (2005)
<i>tyra-3</i>	Expr11003	BAG, AWC, AUA, ASK, AIM, AFD, ADL, OLQ, CEP, SDQ	Bendesky et al. (2011)
	Expr6415	PVT	Hunt-Newbury et al. (2007)
	Expr12173	ADE	Wragg et al. (2007)
<i>lgc-55</i>	Expr8613	AVB, ALN, IL1, HSN, SMD, SDQ, RMD	Pirri et al. (2009)
	Expr8997	AVM, ALM	Ringstad et al. (2009)

Table A.7 Tyramine (TA) receptor expression patterns.



# Appendix B

## Neuropeptide expression patterns

Receptor	WBID	Neurons	Reference
<i>flp-1</i>	Expr3003	AVK, AVE, AVA, RIG, AIY, AIA, M5, RMG	Kim and Li (2004)
<i>flp-4</i>	Expr3006	AWC, AVM, ASEL, ADL, PVD, PHB, PHA, NSM, I5, I6, FLP	Kim and Li (2004)
<i>flp-5</i>	Expr3007	ASE, PVT, M4, I4, I2 RMG	Kim and Li (2004)
<i>flp-10</i>	Expr3011	AIM, ASI, AUA, BAG, BDU, DVB, PQR, PVR, URX	Kim and Li (2004)
<i>flp-13</i>	Expr3014 Expr12005	ASE, ASG, ASK, BAG, DD, I5, M3, M5 ALA	Kim and Li (2004) Nelson et al. (2015)
<i>flp-15</i>	Expr3015	PHA, I2	Nelson et al. (2015)
<i>flp-17</i>	Expr3016	BAG, M5	Kim and Li (2004)
<i>flp-18</i>	Expr3017	AVA, AIY, RIG, RIM, M2, M3	Kim and Li (2004)
<i>flp-21</i>	Expr3020 Expr12181	ASI, ASH, ASE, ADL, MC, M4, FLP, URA RMG, ASJ, URX, M2, ASK, ASG, ADF	Kim and Li (2004) Macosko et al. (2009)
<i>nlp-1</i>	Expr1686 Marker88	ASI, AWC, PHB, BDU HSN	Nathoo et al. (2001) Karakuzu et al. (2009)
<i>nlp-12</i>	Expr8057	DVA	Janssen et al. (2008b)
<i>ntc-1</i>	Expr11371 Expr11368	AVK, RIC, AIZ, AFD, NSM, M5, DVA, DD, VD, VC ASG	Garrison et al. (2012) Beets et al. (2012)
<i>pdf-1</i>	Expr11002 Expr9958	AVB, ASK, AIM, AFD, PVT, PVP, PVN, LUA, SIA, SAA, RMG ASI, RID, ADA, ADE, PQR, PHB, PHA, RME	Barrios et al. (2012) Janssen et al. (2008b)
<i>pdf-2 / nlp-3</i>	Expr9959	BDU, AVG, AVD, RIM, AQR, RID, AIM, PVT, PVP, PQR, PHB, PHA, RIS	Janssen et al. (2008b)
<i>nlp-24</i>	Expr1717	ASI	Nathoo et al. (2001)

Table B.1 Neuropeptide expression patterns.

Receptor	WormBase ID	Neurons	Reference
<i>npr-1</i>	Expr2257	AUA, ASH, ASG, ASE, AQR, RIG, PQR, PHB, PHA, OLQ, IL2L, IL2R, URX, SMBDL, SMBDR, RMG, RIV, DD, VD, M3, SAADL, SAADR, SDQ	Coates and de Bono (2002)
<i>npr-2</i>	Expr12242	ADF, AIZ, ASH, FLP, OLQ, PVD, PVQ, SAB	Luo et al. (2015)
<i>npr-3</i>	Expr2766	AS, DA, DB, VA, VB	Keating et al. (2003)
<i>npr-4</i>	Expr8975	BDU, BAG, AVA, PQR, RIV	Cohen et al. (2009)
<i>npr-5</i>	Expr8976	AWB, AWA, AUA, ASK, ASJ, ASI, ASG, ASE, AIA, ADF, PHB, PHA, IL2	Cohen et al. (2009)
<i>npr-11</i>	Expr12179	AIA, AIY	Chalasani et al. (2010)
<i>frpr-4</i>	N/A	RIA, PVM, AVE, I1, DVA	Nelson et al. (2015)
<i>npr-17</i>	Expr12182	AVG, ASI, PVP, PVQ, PQR	Harris et al. (2010)
<i>ckr-2</i>	Expr10065	AIY	Wenick and Hobert (2004)
	Expr12178	AS, DA, DB, VA, VB	Hu et al. (2011)
<i>ntr-1</i>	Expr11372	ASH, RIC, ADL, ADF, PVW, PVR, PVQ, I2	Garrison et al. (2012)
	Expr11369	BDU, ASE, PQR	Beets et al. (2012)
<i>egl-6</i>	Expr8338	HSN, DVA, SDQ	Ringstad and Horvitz (2008)
<i>pdfr-1</i>	Expr10592	AVM, AVD, RIF, ALM, PVW, PVQ, PVM, PVC, PQR, PLM, PHA, OLL, DB2, URY, URX, RME, AVF	Barrios et al. (2012)
	Expr8177	PHB, OLQ, I1 FLP	Janssen et al. (2008a)

Table B.2 Neuropeptide receptor expression patterns.



## Appendix C

### Identification of dopamine receptor neurons

## C.1 Acknowledgements

Courtesy of Robyn Branicky, Yee Lian Chew & William R. Schafer.

## C.2 Materials & methods

The expression patterns of the dopamine receptors were determined using the reporter strains DA1646 *lin-15B* & *lin-15A(n765)* *X*; *adEx1646 [lin-15(+)* *T02E9.3(dop-5)::GFP*], BC13771 *dpy-5(e907)* *I*; *sEX13771 [rCesC24A8.1(dop-6)::GFP + pCeh361]*, and FQ78 *wzIs26 [lgc-53::gfp; lin-15(+)]*; *lin-15B* & *lin-15A(n765)* (kindly provided by Niels Ringstad).

The neurons expressing the receptors were identified based on the position and shape of the cell bodies and in most cases co-labelling with other markers. The reporter strains were all crossed with the cholinergic reporter (Pereira et al., 2015) OH13646 *pha-1(e2123)* *III*; *him-5(e1490)* *V*; *otIs544 [cho-1(fosmid)::SL2::mCherry::H2B + pha-1(+)]* and the glutamatergic reporter (Serrano-Saiz et al., 2013) OH13645 *pha-1(e2123)* *III*; *him-5(e1490)* *V*; *otIs518 [eat-4(fosmid)::SL2::mCherry::H2B + pha-1(+)]* (both kindly provided by Oliver Hobert), and dye-filled with DiI using standard procedures (Shaham, 2006). Strains were also crossed to AQ3072 *ljEx540[cat-1::mcherry]* and PT2351 *him-5(e1490)* *V*; *myEx741 [pdf-1(3kb)::NLS::RFP + unc-122::GFP]*, which label cells expressing the vesicular monoamine transporter and the PDFR-1 receptor, respectively. When ambiguous, reporter strains were crossed with the additional strains listed below. Reporter expression in individual neurons was confirmed with the following crosses:

### C.2.1 For *dop-5*:

- AIM and ADF were confirmed based on coexpression with *cat-1*.
- URX, PVC, RIF, RIB, AIY, M5, and DVA were identified based on position and coexpression with *cho-1* (Pereira et al., 2015).
- MI, DVC, ASE (previously identified in Etchberger et al. (2009)) and ADA were confirmed based on position and coexpression with *eat-4* (Serrano-Saiz et al., 2013).
- ASI, PHA and PHB were confirmed based on costaining with DiI.

- PVT, RMG and BDU were identified based on cell body position and shape alone.

### C.2.2 For *dop-6*:

- RIH and ADF were confirmed based on coexpression with *cat-1* (Duerr et al., 1999).
- ASI and PHA were confirmed based on costaining with DiI.
- AQ3499 *ljEx805 [sra-6::mcherry + PRF4]* was used to confirm expression in PVQ.
- AQ3682 *ljEx921[flp-8::mcherry + unc-122::gfp]* was used to confirm expression in URX and AUA.
- IL2, RIB, RMD and URA were identified based on position and coexpression with *cho-1*.
- AVF was identified based on position and failure to coexpress *eat-4* and *cho-1*.
- RID was identified based on position relative to URX and morphology.

### C.2.3 For *lgc-53*:

- AIM was confirmed based on coexpression of *cat-1*.
- AVF was confirmed based on coexpression with *pdfr-1* and failure to coexpress *eat-4* and *cho-1*.
- URY was confirmed based on position, coexpression with *eat-4*, and lack of coexpression with *ocr-4*.
- AQ3526 *ljEx822 [klp-6::mcherry + pRF4]* was used to confirm IL2 expression.
- AQ3535 *ljEx828 [unc-4::mcherry + pRF4]* was used to confirm VA expression.
- FLP was confirmed based on position, morphology, and coexpression with *eat-4*.
- HSN, CAN and PVD expression were identified based on position and morphology.

### C.2.4 Microscopy

Strains were examined using a Zeiss Axioskop. Images were taken using a Zeiss LSM780 confocal microscope. Worms were immobilised on 3 % agarose pads with 2.5 mM levamisole. Image stacks were acquired with the ZEN 2010 software and processed with Image J (Abramoff et al., 2004).

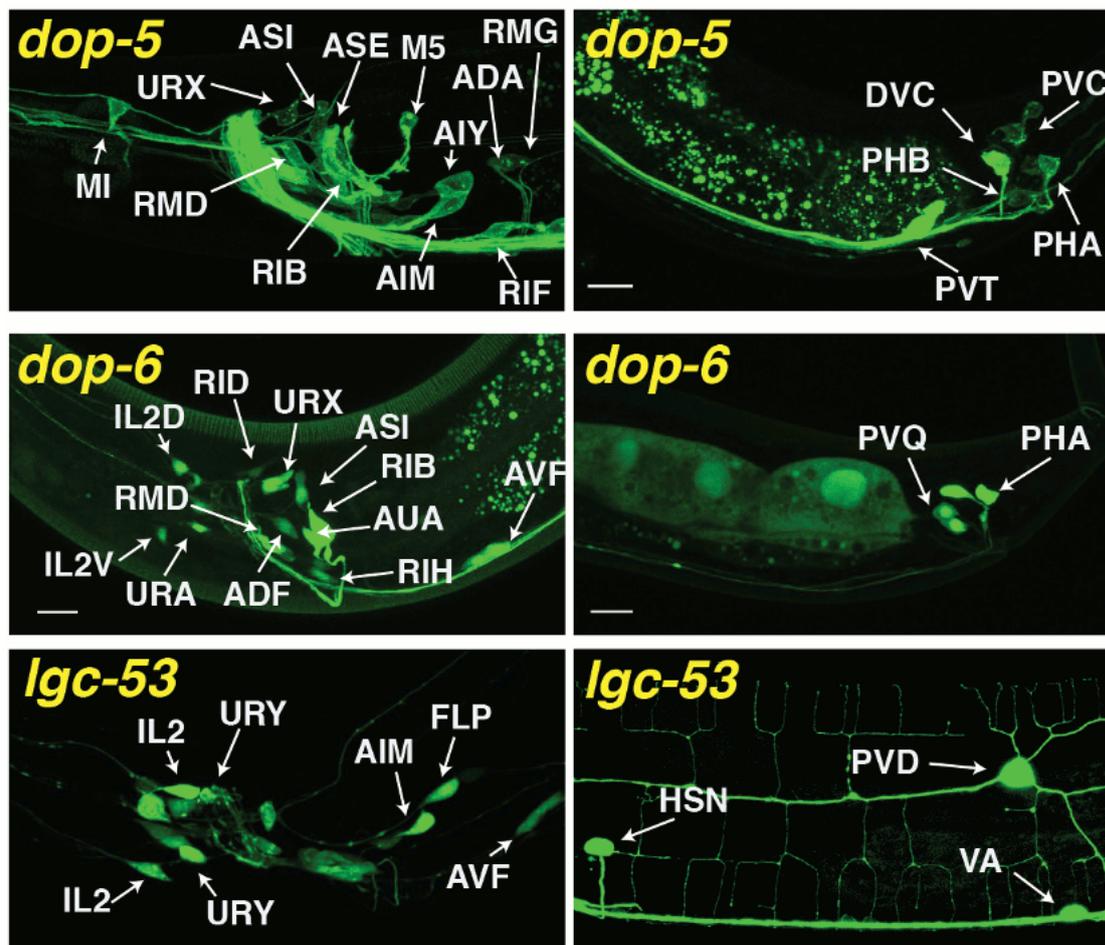


Fig. C.1 Expression patterns of the dopamine receptors *dop-5*, *dop-6* & *lgc-53*. Shown are representative images showing expression of GFP reporters under the control of indicated receptor promoters in the head (left panels) or tail/posterior body (right panels). Identified neurons are labelled; procedures for confirmation of cell identities are described in methods. In all panels, dorsal is up and anterior is to the left. In addition to the neurons indicated, dopamine receptor reporters were detected in the following neurons: *dop-5*: BDU (some animals); *lgc-53*: CAN (some animals). Courtesy of RB, YLC & WRS (Bentley et al., 2016).

# Appendix D

## Notes on expression data

## D.1 Fluorescent reporter lines

**M**OST of the expression patterns presented in this thesis are inferred from GFP reporter fusions. As such, a number of caveats apply.

The construction and analysis of the monoamine and neuropeptide VT networks – covered in Chapters 4 & 5, respectively – included only those neurons that were identified as having strong, consistent, expression of the reporter genes of interest. The expression results in the literature are predominantly qualitative measures (e.g. strong / weak expression), due to the inherent difficulties of quantifying absolute expression from reporter fluorescence.

The expression level of a reporter is only a proxy measure of the true expression of the endogenous gene, due to potential differences in regulatory information between the transgene and the endogenous counterpart (Boulin et al., 2006). Furthermore, the signal from a fluorophore reporter is also subject to confounding factors such as molecular interactions (quenching), reduced fluorescence from previous light exposure and prolonged imaging (photobleaching), reporter copy number variation, and variable brightness due to protein accumulation, aggregation, and the site of localisation. Many of these are highly sensitive to environmental and inter-animal variation. Although some variables can be partially controlled for – for example, by ensuring single copy chromosomal integration (Frøkjær-Jensen, 2015; Frøkjær-Jensen et al., 2008) and using a reference marker or standard intensity signal to calculate relative expression (Chen et al., 2011b; Kelly et al., 2009) – additional problems are often introduced, such as epigenetic and position effects due to the site of integration (Koelle, 1994), and measurement difficulties from fluorescence energy transfer (FRET), bleedthrough, interference from background autofluorescence (Niswender et al., 1995), and variability or nonlinearity in the optical properties of tissues (Jacques, 2013).

While work has been done to demonstrate that GFP reporters can be used to quantitatively estimate absolute expression levels (Murray et al., 2008, 2012; Soboleski et al., 2005), this is not common practice in standard experiments to analyse expression patterns in *C. elegans* (Boulin et al., 2006). The reason for this is that most investigations are primarily concerned with determining the locations of expression, rather than the quantitative expression levels.

## D.2 Gene expression & network construction

Fortunately, precise measurements of gene expression levels are not required for the objectives of this thesis: to construct and analyse extrasynaptic signalling networks. Identifying which cells are involved as either releasing or receiving nodes in the network is sufficient to construct an unweighted directed interaction map.

The lack of information on absolute expression levels means that it is not possible to make predictions about the strength of interactions, but it is also not necessary. As demonstrated in this thesis, the network topology can still be analysed on the basis of which cells have the capacity to interact, as predicted from qualitatively strong expression signals. This is similar to the current level of information available for the wired connectome, where much remains unknown about the functional interaction strengths and polarities of synapses. The lack of data on connection strengths has not significantly hindered the ability to study synaptic circuits or large scale organisation in the nervous system of *C. elegans*.

## D.3 Data integration & batch effects

Usually, integrating gene expression data from different sources can present additional unique problems. Variations in the genetic background of strains (as a result of genetic drift), as well as differences in protocols, reagents, environmental conditions, and animal handling techniques, are likely to result in batch effects, where results vary between different batches of the same experiment, as well as between different experiments and laboratories (Chen et al., 2011a; Goh et al., 2017; Leek et al., 2010). Such batch effects are less important when working with descriptive measures of expression levels.

The course-grained approach of binary filtering used to construct the network in this thesis (i.e. a neuron is included if it has strong expression, else it is excluded) removes the issue of minor measurement noise and expression variability, as we are not dealing with precise expression levels that are highly sensitive to experimental conditions. Furthermore, many of the cell identifications used in the construction of the network have been validated using more than one reporter line, and in the case of the monoamines, additional support is provided by immunostaining (Duerr et al., 1999; Sawin et al., 2000) and formaldehyde-induced fluorescence (FIF) experiments (Horvitz et al., 1982; Lints and Emmons, 1999; Rivard et al., 2010; Sulston et al., 1975).

## D.4 Underreported expression

Reporter lines typically provide a good approximation of the expression locations of endogenous genes; however, under certain circumstances reporters can fail to identify cells where expression has been demonstrated via other methods, such as phenotypic rescue in mutants, or evidence from cell-specific RNAi knockdown (e.g. Xu et al., 2001). Reasons for such a discrepancy might include missing regulatory sequences in the reporter construct, or expression being below the level of detection.

Examples of potential expression underreporting relevant to this thesis include evidence of the *dop-3* and *dop-4* functioning in ASH (Ezak and Ferkey, 2010; Ezcurra et al., 2011). These receptors, along with others that have only been identified as having with weak expression, were excluded from the network construction to prevent misreporting and false-positives. It is therefore likely that, at least in some instances, the networks presented in this thesis understate the real extent of connectivity in *C. elegans*. For further discussion, see § 4.4.1 on page 141.

## D.5 Conditional expression

An important point to consider is that the approach and data used in this thesis only focus on neurons that constitutively express the genes of interest. As mentioned in Chapter 4, some cells can express VT-associated ligands or receptors conditionally, as in the case of ASG and serotonin under hypoxic conditions (Pocock and Hobert, 2010). More data will be required before such expression dynamics can be considered.

# Appendix E

## Null model network comparison

## E.1 Introduction

AS DISCUSSED in Chapter 2, network measures can be influenced by many of the underlying structural features of the network, such as the density of the network and its degree distribution (Strang et al., 2017); for example, two separate networks with qualitatively similar architectures might appear different on some network measures due only to differences in the number of nodes and edges they contain. As such, as well as studying the raw network measures, it is also useful to compare the properties of a given network to comparable null hypothesis, or *null model*, networks which can provide a baseline reference for a set of structural assumptions or constraints, and determine whether the properties observed in a network can be explained by a random process. That is to say, we can compare the distribution of results obtained from a random statistical ensemble of artificial (null) networks to those observed in a real-world network, and establish whether they are likely to come from the same distribution. Alternatively, we can use the average ensemble measures to normalise the real-world measures, and control for any variability caused by those features captured in the null model, such as the network size. This approach allows us to then directly compare networks with different structures.

The two most common methods for generating null model reference networks that preserve the degree distribution and the number of nodes and edges (and thus the network's size and density), are the Markov chain Monte Carlo *edge-swap* (**ES**) algorithm – where networks are randomised by iteratively swapping edges at random while preserving the degrees of nodes, based on the procedure described by Maslov and Sneppen (2002); Milo et al. (2003) (see Algorithm 5, page 58) – and the *configuration model* (**CM**), which constructs synthetic random networks from a prescribed degree sequence (Molloy and Reed, 1995) (Algorithm 6, page 59).

The edge-swap (ES) null model is used as the main reference model throughout this thesis, in the analysis of the wired synaptic networks as well as the extrasynaptic monoamine and neuropeptide networks. To determine whether the results in this thesis are likely to be influenced by the choice of null model, a comparison was performed on network measures for both ES and CM null model network ensembles, generated to match the degree sequence of (1) the Albertson-Chklovskii (AC) aggregate synaptic and gap-junction connectome, and (2) the aggregate monoamine (MA) network.

## E.2 Materials & methods

Four sets of null model networks were generated and compared on the standard set of network measures. For the aggregate AC synaptic + gap-junction wired connectome of *C. elegans*, 100 ES networks were generated by randomising the original network with  $10 \times M$  edge-swaps, and 100 CM networks were generated using the degree sequence of the AC wired connectome. The same methods were used to construct 100 ES networks and 100 CM networks for the aggregate monoamine connectome.

All of the analyses were performed on directed versions of the networks, with multiple edges removed (i.e. unweighted networks). Network measures are the same as those described in Chapter 2.

## E.3 Results

### E.3.1 Wired connectome

	Measure	ES mean	$Z_{\text{ES}}$	CM mean	$Z_{\text{CM}}$	Observed
$r$	Reciprocity	0.07	(9.86)	0.08	(9.85)	0.48
$Q$	Modularity	0.18	(9.86)	0.18	(9.82)	0.46
$C$	Clustering	0.07	(9.77)	0.08	(9.71)	0.16
$L$	Path length	2.65	(9.89)	2.66	(9.87)	3.22
$R$	Assortativity	-0.08	(1.27)	-0.07	(0.07)	-0.07
$\rho$	Robustness	0.486	(-6.52)	0.486	(-6.33)	0.466

Table E.1 Comparison of mean network measures computed for 100 randomised networks generated from the degree sequence of the directed version of the Albertson-Chklovskii synaptic & gap junction connectome of *C. elegans* (see Chapter 3); showing results for the edge-swap rewiring procedure (ES), the configuration model (CM), and the observed values from *C. elegans*.  $Z$ -scores comparing the observed values to the expected are shown in parentheses.

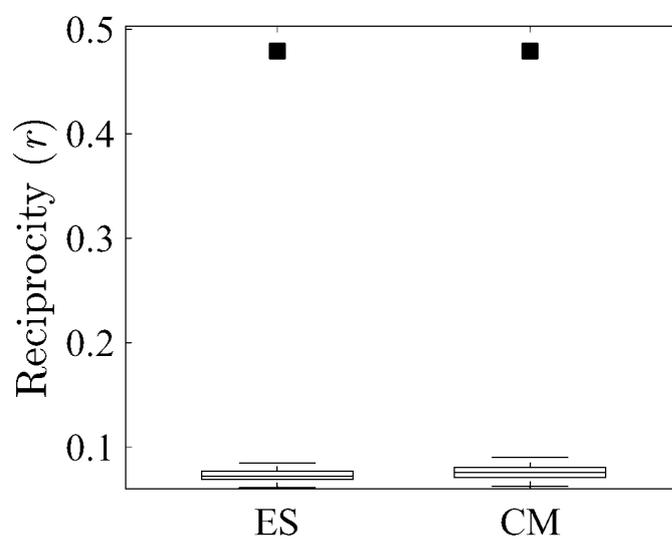


Fig. E.1 Reciprocity ( $r_{ES} = 0.07$ ,  $r_{CM} = 0.08$ ). Comparison of reciprocity for null model networks generated from the degree sequence of the AC connectome, using either the edge-swap method (ES) or the configuration model (CM). Plots show the observed values for the original AC connectome (filled squares) and expected values from 100 null model networks (boxplots).

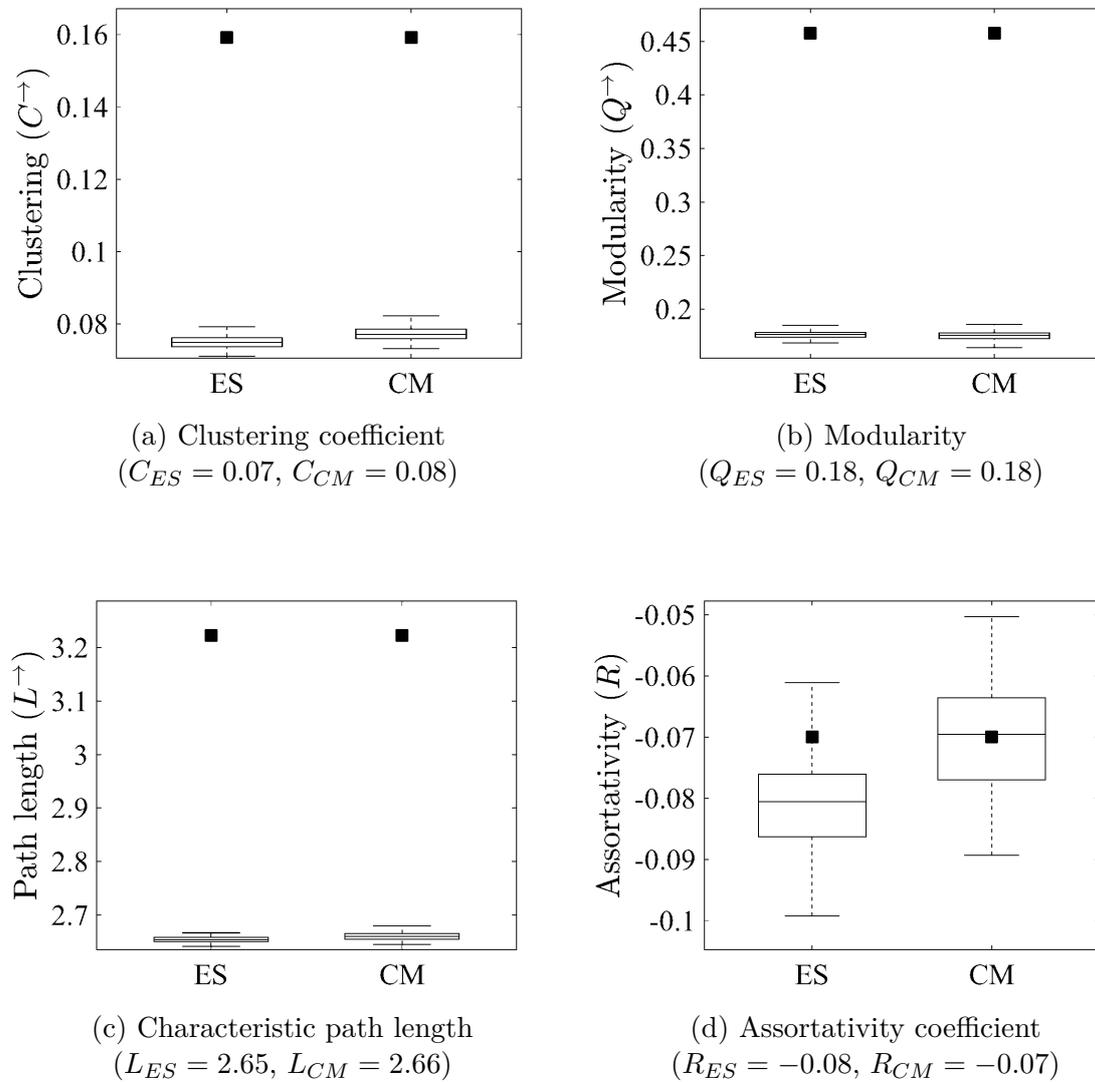


Fig. E.2 Comparison of network measures for null model networks generated from the degree sequence of the AC connectome, using either the edge-swap method (ES) or the configuration model (CM). Plots show the observed values for the original AC connectome (filled squares) and expected values from 100 null model networks (boxplots).

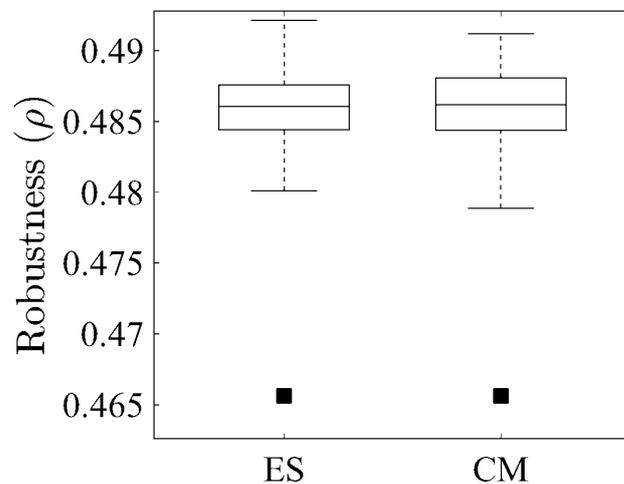


Fig. E.3 Robustness ( $\rho_{ES} = 0.486$ ,  $\rho_{CM} = 0.486$ ). Comparison of reciprocity for null model networks generated from the degree sequence of the AC connectome, using either the edge-swap method (ES) or the configuration model (CM). Plots show the observed values for the original AC connectome (filled squares) and expected values from 100 null model networks (boxplots).

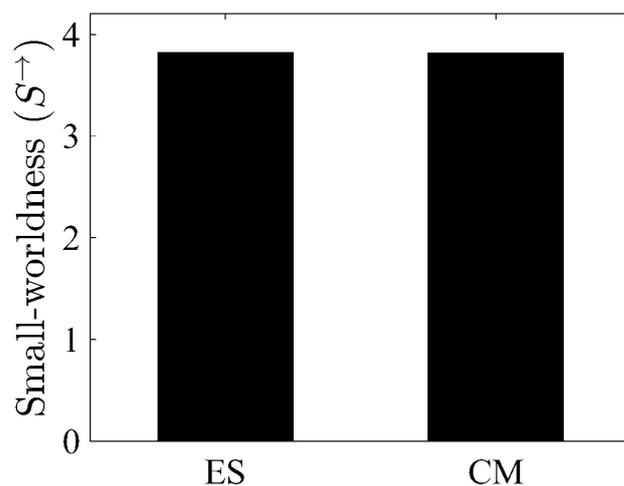


Fig. E.4 Small-worldness ( $S_{ES} = 3.82$ ,  $S_{CM} = 3.81$ ). Comparison of small-worldness of the AC connectome normalised using 100 null model networks generated using either the edge-swap method (ES) or the configuration model (CM).

### E.3.2 Monoamine connectome

	Measure	ES mean	$Z_{ES}$	CM mean	$Z_{CM}$	Observed
$r$	Reciprocity	0.058	(2.1)	0.058	(1.94)	0.064
$Q$	Modularity	0.11	(9.38)	0.11	(9.44)	0.32
$C$	Clustering	0.125	(6.34)	0.125	(6.71)	0.137
$L$	Path length	1.58	(5.84)	1.59	(5.99)	1.68
$R$	Assortativity	-0.797	(-1.06)	-0.797	(-1.23)	-0.8
$\rho$	Robustness	0.39	(-9.17)	0.39	(-9.21)	0.36

Table E.2 Comparison of mean network measures computed for 100 randomised networks generated from the degree sequence of the directed version of the monoamine connectome of *C. elegans* (see Chapter 4); showing results for the edge-swap rewiring procedure (ES), the configuration model (CM), and the observed values from *C. elegans*.  $Z$ -scores comparing the observed values to the expected are shown in parentheses.

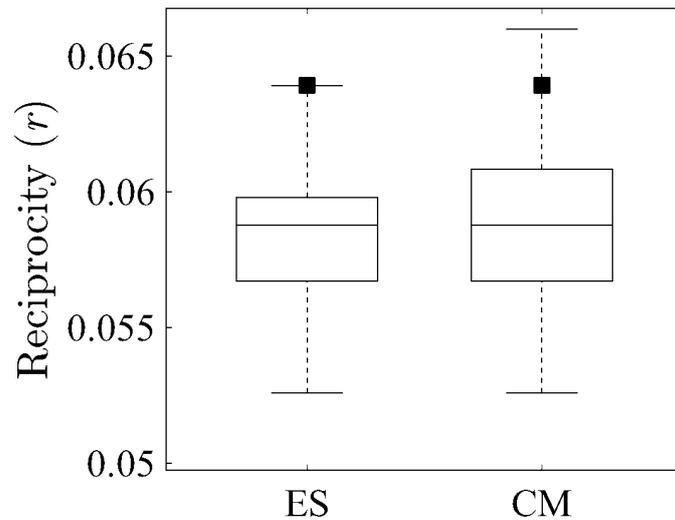


Fig. E.5 Reciprocity ( $r_{ES} = 0.058$ ,  $r_{CM} = 0.058$ ). Comparison of reciprocity for null model networks generated from the degree sequence of the MA connectome, using either the edge-swap method (ES) or the configuration model (CM). Plots show the observed values for the original MA connectome (filled squares) and expected values from 100 null model networks (boxplots).

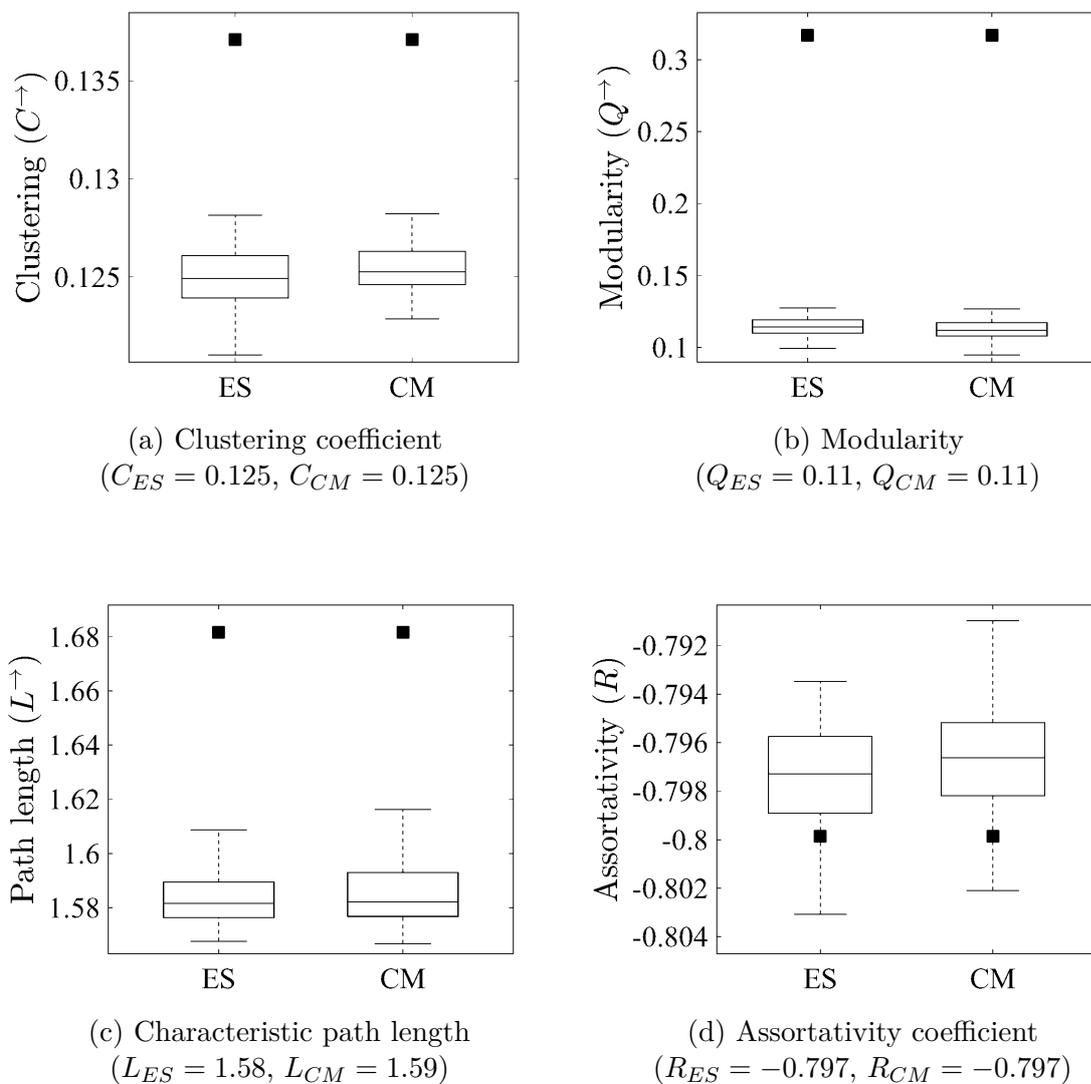


Fig. E.6 Comparison of network measures for null model networks generated from the degree sequence of the MA connectome, using either the edge-swap method (ES) or the configuration model (CM). Plots show the observed values for the original MA connectome (filled squares) and expected values from 100 null model networks (boxplots).

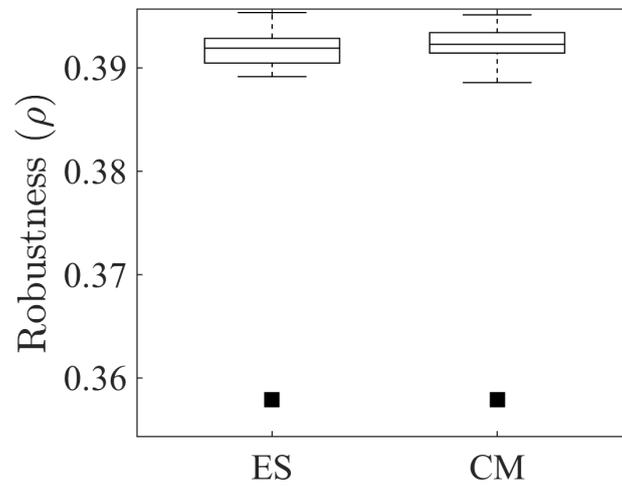


Fig. E.7 Robustness ( $\rho_{ES} = 0.39$ ,  $\rho_{CM} = 0.39$ ). Comparison of reciprocity for null model networks generated from the degree sequence of the MA connectome, using either the edge-swap method (ES) or the configuration model (CM). Plots show the observed values for the original MA connectome (filled squares) and expected values from 100 null model networks (boxplots).

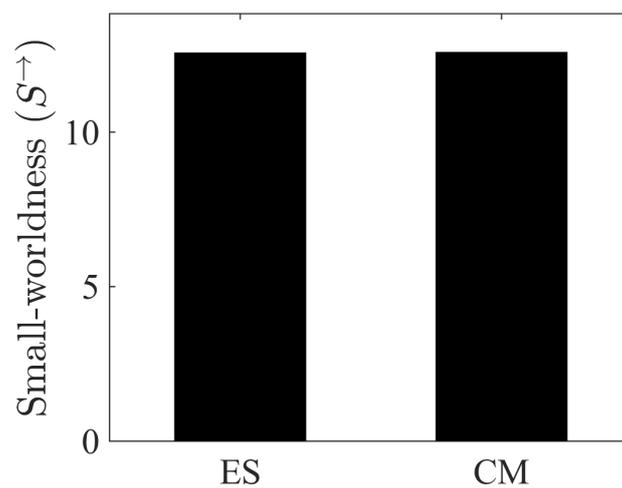


Fig. E.8 Small-worldness ( $S_{ES} = 12.56$ ,  $S_{CM} = 12.57$ ). Comparison of small-worldness of the MA connectome normalised using 100 null model networks generated using either the edge-swap method (ES) or the configuration model (CM).

## E.4 Conclusions & discussion

No major differences were observed on any of the main network metrics between the ES and CM null models, using either the synaptic + gap junction network ( $AC_{ag}$ ) or the wireless monoamine network ( $MA_{ag}$ ).

The maximum difference identified on any of the mean network measures between the ES and CM null models was  $\pm 0.01$  (Tables E.1 & E.2). This was the case for both the AC and MA conditions.

Currently, edge-swap randomisation and random network generation via the configuration model represent two of the most common degree-preserving null network models. The results presented here suggest that for the networks under investigation the decision to use either the ES or CM model should have negligible effects on the outcome of network measures. This is not surprising considering that both methods are based on similar assumptions about the underlying network structures, and preserve the same topological features; namely the number of nodes, edges, and their degree sequences. As these are known to have large effects on other network properties (Strang et al., 2017), they are useful variables to control for. It is likely that other null models, based on different assumptions, would generate different results (Artzy-Randrup et al., 2004; Hosseini and Kesler, 2013; Klimm et al., 2014); however, as research into the generation, parameterisation, and applicability of null models is still an ongoing area of research, the determination of the optimal null model for any given network is still an open question.

# Appendix F

## Neuron classification

## F.1 Neuron classification schema

IN THE analyses presented in this thesis, neurons are classified into one of four categories based on their function and position in the network. These functional classes are: pharyngeal neurons, sensory neurons, interneurons, and motor neurons. These classes are derived from more complex classifications published by WormAtlas (Altun et al., 2002) and the WormWiring project (Emmons et al., 2015). As the base classification schemata used by WormAtlas and WormWiring are more detailed, having more than four classes (see Tables F.1 – F.4 for the source data from WormWiring), the neurons were here reclassified using a reduced schema with the following mapping:

### Sensory neurons:

- Amphid sensory neurons
- Labial and cephalic sensory neurons
- Mechanosensory neurons
- Phasmid neurons

### Interneurons:

- Amphid interneurons
- Ring interneurons
- Ventral cord interneurons

### Motor neurons:

- Head motor neurons
- Bodywall motor neurons
- Hermaphrodite motor neurons

Pharyngeal neurons: unchanged.

---

ADAL	Amphid Interneuron	ASEL	Amphid sensory
ADAR	Amphid Interneuron	ASER	Amphid sensory
ADEL	Mechanosensory	ASGL	Amphid sensory
ADER	Mechanosensory	ASGR	Amphid sensory
ADFL	Amphid sensory	ASHL	Amphid sensory
ADFR	Amphid sensory	ASHR	Amphid sensory
ADLL	Amphid sensory	ASIL	Amphid sensory
ADLR	Amphid sensory	ASIR	Amphid sensory
AFDL	Amphid sensory	ASJL	Amphid sensory
AFDR	Amphid sensory	ASJR	Amphid sensory
AIAL	Amphid Interneuron	ASKL	Amphid sensory
AIAR	Amphid Interneuron	ASKR	Amphid sensory
AIBL	Amphid Interneuron	AUAL	Mechanosensory
AIBR	Amphid Interneuron	AUAR	Mechanosensory
AIML	Amphid Interneuron	AVAL	Ventral cord Interneuron
AIMR	Amphid Interneuron	AVAR	Ventral cord Interneuron
AINL	Amphid Interneuron	AVBL	Ventral cord Interneuron
AINR	Amphid Interneuron	AVBR	Ventral cord Interneuron
AIYL	Amphid Interneuron	AVDL	Ventral cord Interneuron
AIYR	Amphid Interneuron	AVDR	Ventral cord Interneuron
AIZL	Amphid Interneuron	AVEL	Ventral cord Interneuron
AIZR	Amphid Interneuron	AVER	Ventral cord Interneuron
ALA	Amphid Interneuron	AVFL	Ventral cord Interneuron
ALML	Mechanosensory	AVFR	Ventral cord Interneuron
ALMR	Mechanosensory	AVG	Ventral cord Interneuron
ALNL	Mechanosensory	AVHL	Ventral cord Interneuron
ALNR	Mechanosensory	AVHR	Ventral cord Interneuron
AQR	Mechanosensory	AVJL	Ventral cord Interneuron
AS01	Bodywall motor neuron	AVJR	Ventral cord Interneuron
AS02	Bodywall motor neuron	AVKL	Ventral cord Interneuron
AS03	Bodywall motor neuron	AVKR	Ventral cord Interneuron
AS04	Bodywall motor neuron	AVL	Ventral cord Interneuron
AS05	Bodywall motor neuron	AVM	Mechanosensory
AS06	Bodywall motor neuron	AWAL	Amphid sensory
AS07	Bodywall motor neuron	AWAR	Amphid sensory
AS08	Bodywall motor neuron	AWBL	Amphid sensory
AS09	Bodywall motor neuron	AWBR	Amphid sensory
AS10	Bodywall motor neuron	AWCL	Amphid sensory
AS11	Bodywall motor neuron	AWCR	Amphid sensory

---

Table F.1 Neuron classes: A.

---

BAGL	Mechanosensory	I1L	Pharynx
BAGR	Mechanosensory	I1R	Pharynx
BDUL	Amphid Interneuron	I2L	Pharynx
BDUR	Amphid Interneuron	I2R	Pharynx
CANL	Amphid Interneuron	I3	Pharynx
CANR	Amphid Interneuron	I4	Pharynx
CEPDL	Labial / cephalic sensory	I5	Pharynx
CEPDR	Labial / cephalic sensory	I6	Pharynx
CEPVL	Labial / cephalic sensory	IL1DL	Labial / cephalic sensory
CEPVR	Labial / cephalic sensory	IL1DR	Labial / cephalic sensory
DA01	Bodywall motor neuron	IL1L	Labial / cephalic sensory
DA02	Bodywall motor neuron	IL1R	Labial / cephalic sensory
DA03	Bodywall motor neuron	IL1VL	Labial / cephalic sensory
DA04	Bodywall motor neuron	IL1VR	Labial / cephalic sensory
DA05	Bodywall motor neuron	IL2DL	Labial / cephalic sensory
DA06	Bodywall motor neuron	IL2DR	Labial / cephalic sensory
DA07	Bodywall motor neuron	IL2L	Labial / cephalic sensory
DA08	Bodywall motor neuron	IL2R	Labial / cephalic sensory
DA09	Bodywall motor neuron	IL2VL	Labial / cephalic sensory
DB01	Bodywall motor neuron	IL2VR	Labial / cephalic sensory
DB02	Bodywall motor neuron	LUAL	Ventral cord Interneuron
DB03	Bodywall motor neuron	LUAR	Ventral cord Interneuron
DB04	Bodywall motor neuron	M1	Pharynx
DB05	Bodywall motor neuron	M2L	Pharynx
DB06	Bodywall motor neuron	M2R	Pharynx
DB07	Bodywall motor neuron	M3L	Pharynx
DD01	Bodywall motor neuron	M3R	Pharynx
DD02	Bodywall motor neuron	M4	Pharynx
DD03	Bodywall motor neuron	M5	Pharynx
DD04	Bodywall motor neuron	MCL	Pharynx
DD05	Bodywall motor neuron	MCR	Pharynx
DD06	Bodywall motor neuron	MI	Pharynx
DVA	Ventral cord Interneuron	NSML	Pharynx
DVB	Ventral cord Interneuron	NSMR	Pharynx
DVC	Ventral cord Interneuron	OLLL	Labial / cephalic sensory
FLPL	Mechanosensory	OLLR	Labial / cephalic sensory
FLPR	Mechanosensory	OLQDL	Labial / cephalic sensory
HSNL	Herm motor neuron	OLQDR	Labial / cephalic sensory
HSNR	Herm motor neuron	OLQVL	Labial / cephalic sensory
		OLQVR	Labial / cephalic sensory

---

Table F.2 Neuron classes: B – O.

---

PDA	Bodywall motor neuron	RIBR	Ring Interneuron
PDB	Bodywall motor neuron	RICL	Ring Interneuron
PDEL	Mechanosensory	RICR	Ring Interneuron
PDER	Mechanosensory	RID	Ventral cord Interneuron
PHAL	Phasmid	RIFL	Ring Interneuron
PHAR	Phasmid	RIFR	Ring Interneuron
PHBL	Phasmid	RIGL	Ring Interneuron
PHBR	Phasmid	RIGR	Ring Interneuron
PHCL	Phasmid	RIH	Ring Interneuron
PHCR	Phasmid	RIML	Head motor neuron
PLML	Mechanosensory	RIMR	Head motor neuron
PLMR	Mechanosensory	RIPL	Ring Interneuron
PLNL	Mechanosensory	RIPR	Ring Interneuron
PLNR	Mechanosensory	RIR	Ring Interneuron
PQR	Mechanosensory	RIS	Ring Interneuron
PVCL	Ventral cord Interneuron	RIVL	Head motor neuron
PVCR	Ventral cord Interneuron	RIVR	Head motor neuron
PVDL	Mechanosensory	RMDDL	Head motor neuron
PVDR	Mechanosensory	RMDDR	Head motor neuron
PVM	Mechanosensory	RMDL	Head motor neuron
PVNL	Ventral cord Interneuron	RMDR	Head motor neuron
PVNR	Ventral cord Interneuron	RMDVL	Head motor neuron
PVPL	Ventral cord Interneuron	RMDVR	Head motor neuron
PVPR	Ventral cord Interneuron	RMED	Head motor neuron
PVQL	Ventral cord Interneuron	RMEL	Head motor neuron
PVQR	Ventral cord Interneuron	RMER	Head motor neuron
PVR	Ventral cord Interneuron	RMEV	Head motor neuron
PVT	Ventral cord Interneuron	RMFL	Head motor neuron
PVWL	Ventral cord Interneuron	RMFR	Head motor neuron
PVWR	Ventral cord Interneuron	RMGL	Ring Interneuron
RIAL	Ring Interneuron	RMGR	Ring Interneuron
RIAR	Ring Interneuron	RMHL	Head motor neuron
RIBL	Ring Interneuron	RMHR	Head motor neuron

---

Table F.3 Neuron classes: P – R.

---

SAADL	Head motor neuron	VA04	Bodywall motor neuron
SAADR	Head motor neuron	VA05	Bodywall motor neuron
SAAVL	Head motor neuron	VA06	Bodywall motor neuron
SAAVR	Head motor neuron	VA07	Bodywall motor neuron
SABD	Head motor neuron	VA08	Bodywall motor neuron
SABVL	Head motor neuron	VA09	Bodywall motor neuron
SABVR	Head motor neuron	VA10	Bodywall motor neuron
SDQL	Ring Interneuron	VA11	Bodywall motor neuron
SDQR	Ring Interneuron	VA12	Bodywall motor neuron
SIADL	Ring Interneuron	VB01	Bodywall motor neuron
SIADR	Ring Interneuron	VB02	Bodywall motor neuron
SIAVL	Ring Interneuron	VB03	Bodywall motor neuron
SIAVR	Ring Interneuron	VB04	Bodywall motor neuron
SIBDL	Ring Interneuron	VB05	Bodywall motor neuron
SIBDR	Ring Interneuron	VB06	Bodywall motor neuron
SIBVL	Ring Interneuron	VB07	Bodywall motor neuron
SIBVR	Ring Interneuron	VB08	Bodywall motor neuron
SMBDL	Head motor neuron	VB09	Bodywall motor neuron
SMBDR	Head motor neuron	VB10	Bodywall motor neuron
SMBVL	Head motor neuron	VB11	Bodywall motor neuron
SMBVR	Head motor neuron	VC01	Herm motor neuron
SMDDL	Head motor neuron	VC02	Herm motor neuron
SMDDR	Head motor neuron	VC03	Herm motor neuron
SMDVL	Head motor neuron	VC04	Herm motor neuron
SMDVR	Head motor neuron	VC05	Herm motor neuron
URADL	Head motor neuron	VC06	Herm motor neuron
URADR	Head motor neuron	VD01	Bodywall motor neuron
URAVL	Head motor neuron	VD02	Bodywall motor neuron
URAVR	Head motor neuron	VD03	Bodywall motor neuron
URBL	Ring Interneuron	VD04	Bodywall motor neuron
URBR	Ring Interneuron	VD05	Bodywall motor neuron
URXL	Mechanosensory	VD06	Bodywall motor neuron
URXR	Mechanosensory	VD07	Bodywall motor neuron
URYDL	Mechanosensory	VD08	Bodywall motor neuron
URYDR	Mechanosensory	VD09	Bodywall motor neuron
URYVL	Mechanosensory	VD10	Bodywall motor neuron
URYVR	Mechanosensory	VD11	Bodywall motor neuron
VA01	Bodywall motor neuron	VD12	Bodywall motor neuron
VA02	Bodywall motor neuron	VD13	Bodywall motor neuron
VA03	Bodywall motor neuron		

---

Table F.4 Neuron classes: S – V.

# Appendix G

## Supplementary results

## G.1 Hiveplot of individual MA networks

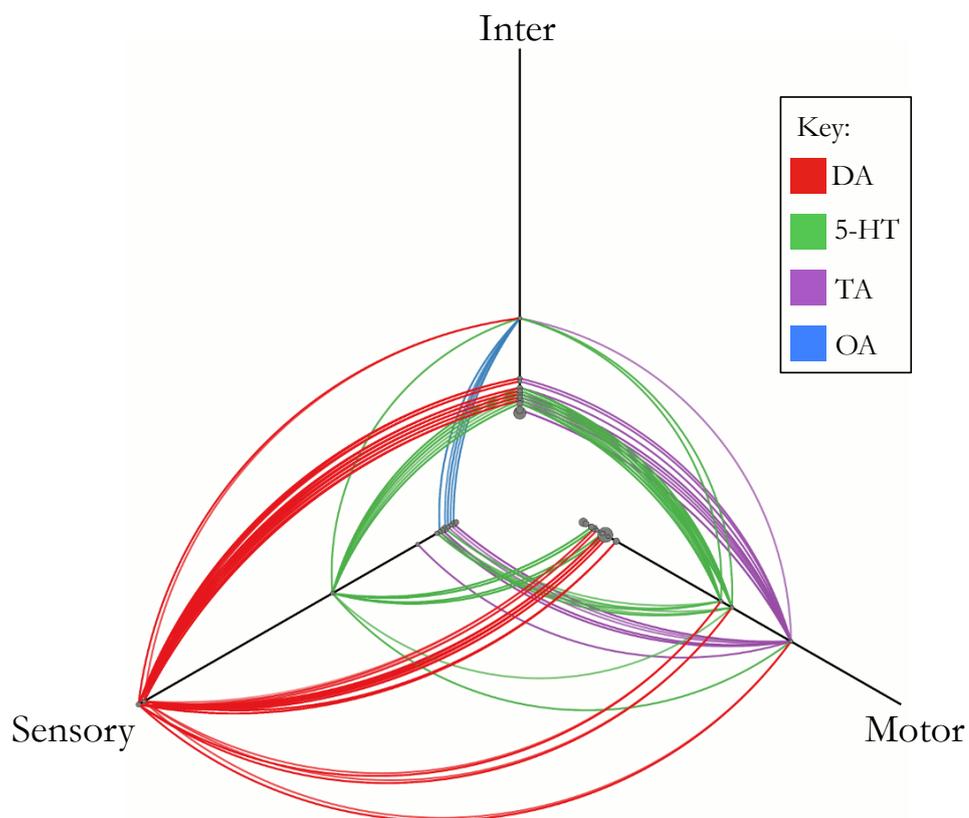


Fig. G.1 Hiveplot showing individual dopamine (red), serotonin (green), tyramine (purple), and octopamine connections (blue). Nodes are classified as sensory, motor, or interneurons and are arranged along the three axes according to their degree. Hubs are located further out along the axes. Plot generated using *hiveplotter*, courtesy of Barnes (2016).

## G.2 Multilink analysis of separate MA networks (incl. *dop-5/-6*)

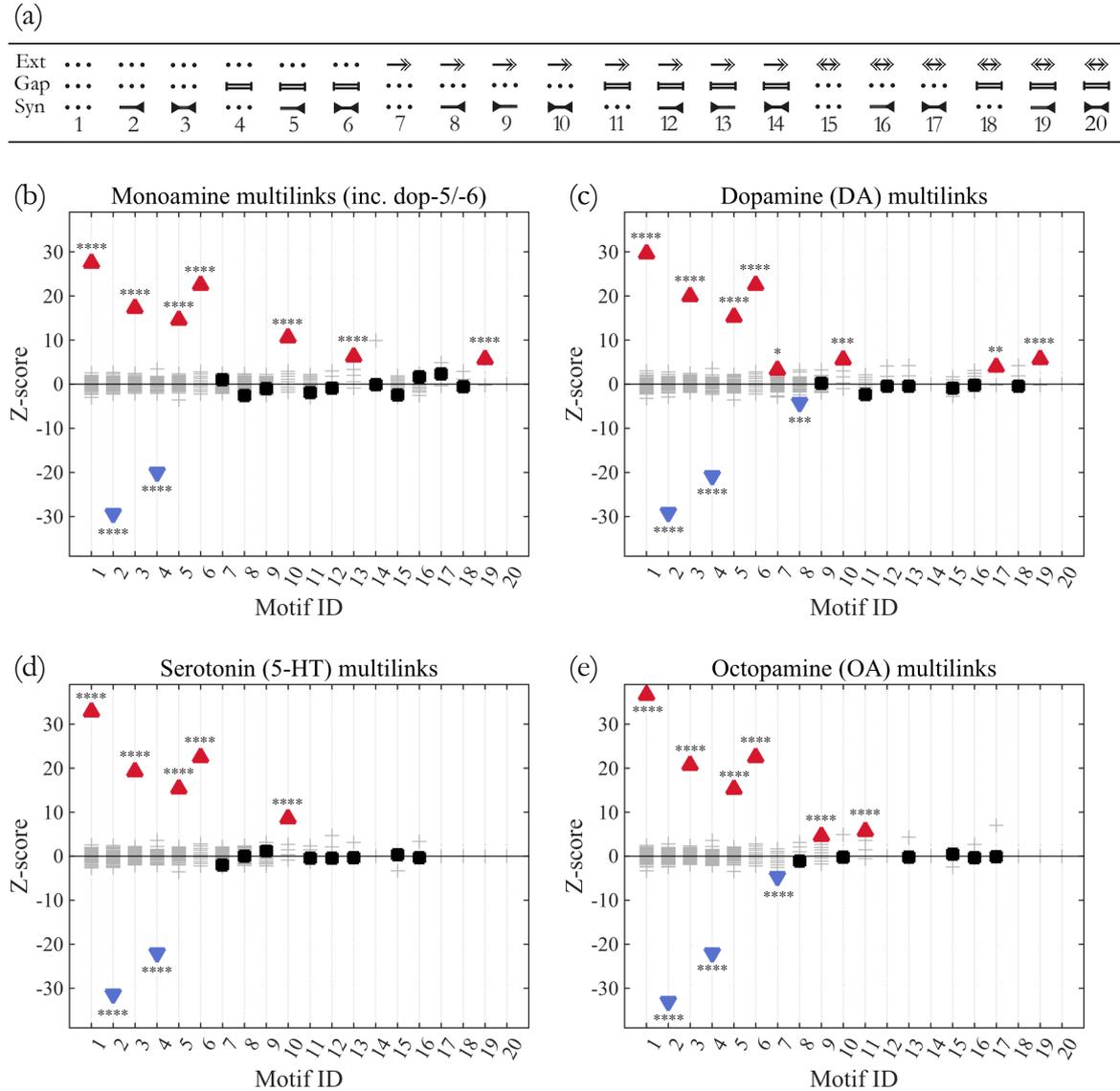


Fig. G.2 Multilink motifs. (a) Motif IDs, as described in Figure 4.19, page 140, (b) aggregate monoamine network including *dop-5* & *dop-6*, (c) dopamine, (d) serotonin, and (e) octopamine networks. Overrepresented motifs are represented by red upward-pointing triangles. Underrepresented motifs are represented by blue downward-pointing triangles. Non-significant motifs are shown by black squares. Values for 100 null model networks are shown as grey crosses. Asterisks report the significance level using the  $z$ -test, with Bonferroni-adjusted  $p$ -values: \* indicates  $p \leq 0.05$ ; \*\* indicates  $p \leq 0.01$ ; \*\*\* indicates  $p \leq 0.001$ ; \*\*\*\* indicates  $p \leq 0.0001$ .

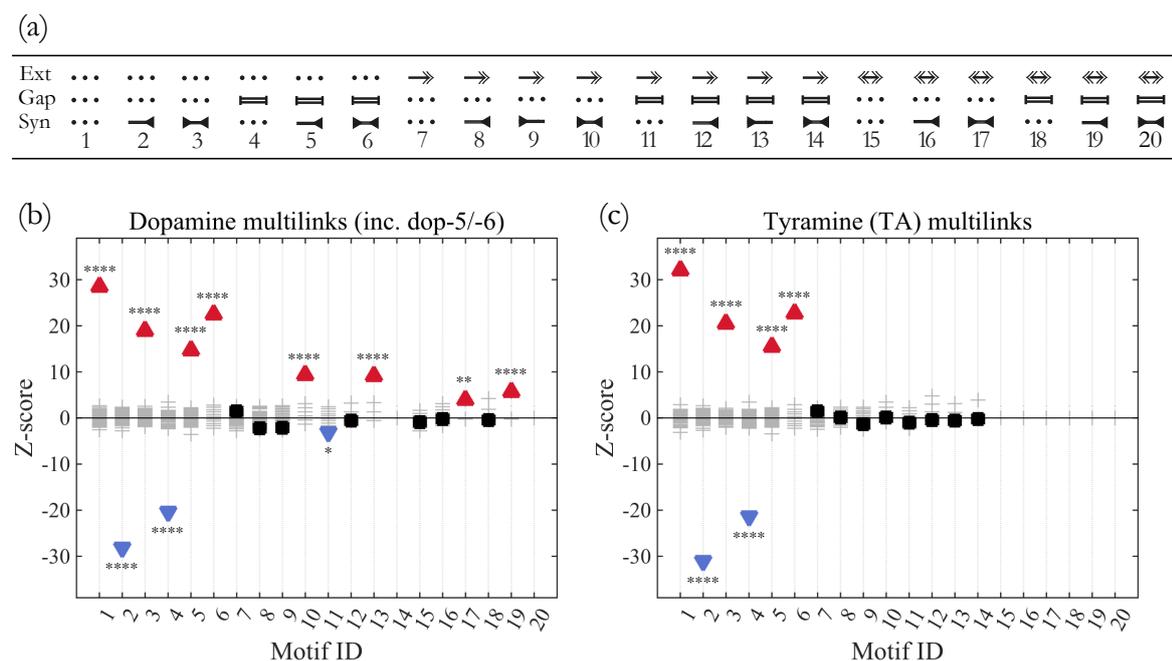


Fig. G.3 Multilink motifs. (a) Motif IDs, as described in Figure 4.19, page 140, (b) dopamine network including *dop-5* & *dop-6*, and (c) tyramine network. Overrepresented motifs are represented by red upward-pointing triangles. Underrepresented motifs are represented by blue downward-pointing triangles. Non-significant motifs are shown by black squares. Values for 100 null model networks are shown as grey crosses. Asterisks report the significance level using the  $z$ -test, with Bonferroni-adjusted  $p$ -values: \* indicates  $p \leq 0.05$ ; \*\* indicates  $p \leq 0.01$ ; \*\*\* indicates  $p \leq 0.001$ ; \*\*\*\* indicates  $p \leq 0.0001$ .

Cell A		Cell B
RICL	$\leftrightarrow$	ADLL
RICR	$\leftrightarrow$	ADLR
RICL	$\leftrightarrow$	CEP (V/D/L/R)
RICR	$\leftrightarrow$	CEP (V/D/L/R)

Table G.1 Examples of octopamine multilink *motif 9*. List of neurons connected by *motif 9* (i.e. unidirectional OA link and synapse in reverse direction).

Cell A		Cell B
RICL	$\leftrightarrow$	ASHL
RICR	$\leftrightarrow$	ASHR
RICL	$\leftrightarrow$	AWBR

Table G.2 Examples of octopamine multilink *motif 11*. List of neurons connected by *motif 11* (unidirectional OA link coincident with gap junction).

### G.3 Network measures of individual MA layers

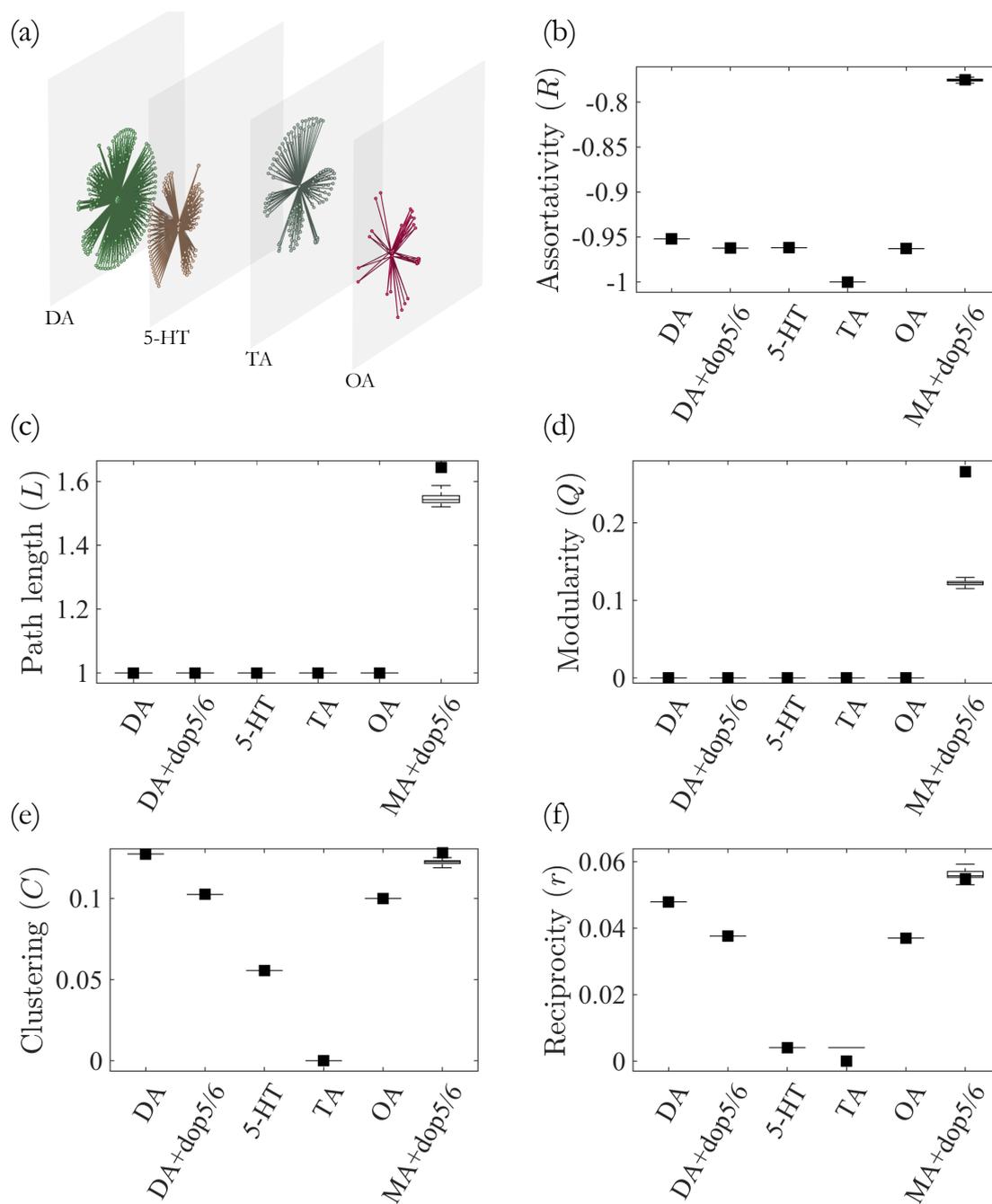


Fig. G.4 (a) Multilayer projection including *dop-5* & *dop-6*. (b-f) Network measures for the individual dopamine (DA), dopamine incl. *dop-5* & *dop-6* (DA+dop-5/-6), serotonin (5-HT), tyramine (TA) and octopamine networks (OA); and aggregate MA network including *dop-5* & *dop-6* (MA+dop-5/-6).

# Appendix H

## Network analysis software

Name / URL	Description
Pajek <a href="http://mrvar.fdv.uni-lj.si/pajek/">http://mrvar.fdv.uni-lj.si/pajek/</a>	Application: general analysis
Gephi <a href="https://gephi.org/">https://gephi.org/</a>	Application: general analysis
Cytoscape* <a href="http://cytoscape.org/">http://cytoscape.org/</a>	Application: general analysis
NetMiner <a href="http://www.netminer.com/">http://www.netminer.com/</a>	Application: general analysis
Network Workbench <a href="http://nwb.cns.iu.edu/">http://nwb.cns.iu.edu/</a>	Application: general analysis
GraphStream <a href="http://graphstream-project.org/">http://graphstream-project.org/</a>	Application: dynamical nets
muxViz* <a href="http://muxviz.net/">http://muxviz.net/</a>	Application: multilayer nets
FANMOD* <a href="http://theinf1.informatik.uni-jena.de/motifs/">http://theinf1.informatik.uni-jena.de/motifs/</a>	Application: motif analysis
MAVisto <a href="http://mavisto.ipk-gatersleben.de/">http://mavisto.ipk-gatersleben.de/</a>	Application: motif analysis
mfinder <a href="https://www.weizmann.ac.il/mcb/UriAlon/">https://www.weizmann.ac.il/mcb/UriAlon/</a>	Application: motif analysis
NetworkX* <a href="https://networkx.github.io/">https://networkx.github.io/</a>	Library: Python
Graph-tool <a href="https://graph-tool.skewed.de/">https://graph-tool.skewed.de/</a>	Library: Python
igraph <a href="http://igraph.org/">http://igraph.org/</a>	Library: Python, R, C/C++
JUNG <a href="http://jung.sourceforge.net/">http://jung.sourceforge.net/</a>	Library: Java
BCT* <a href="http://www.brain-connectivity-toolbox.net/">http://www.brain-connectivity-toolbox.net/</a>	Library: Matlab
Octave Network Toolbox* <a href="https://github.com/aeolianine/octave-networks-toolbox">https://github.com/aeolianine/octave-networks-toolbox</a>	Library: Matlab / Octave

Table H.1 List of software tools that implement network measures or algorithms detailed in Chapter 2. Those used in the network analyses presented in this thesis are denoted by an asterisk \*.