

BENJAMIN F. MAIER

SPREADING PROCESSES IN HUMAN SYSTEMS

Zusammenfassung

Lorem ipsum dolor sit amet¹, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

¹ Ross2009578, 19.

Schlagwörter

Schlagwort1, Schlagwort 2

Abstract

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

Keywords

Keyword1, Keyword2

Contents

1	Introduction	19
I	Human Systems as Networks: Theory & Data	21
2	Networks	23
2.1	Networks as an Abstract Representation of Reality	23
2.2	Network Properties and Effects	23
2.2.1	Adjacency Matrix	23
2.2.2	Node Degree and Its Distribution	25
2.2.3	Clustering Coefficient	25
2.2.4	Path lengths	25
2.2.5	Percolation	25
2.2.6	Average Medium Approximation	25
2.3	Static Network Models	25
2.3.1	Regular Networks and Lattices	25
2.3.2	Erdős–Rényi Random Networks	25
2.3.3	Random Geometric Graphs	26
2.3.4	Watts–Strogatz Small-World Networks	27
2.3.5	Kleinberg Small-World Networks	27
2.3.6	Modular Hierarchical Small-World Networks .	27
2.3.7	Barabasi–Albert Scale-Free Networks	27
2.4	Temporal Networks	27
2.4.1	Temporal Contact Networks	28
2.4.2	Contact- and Inter-Contact Times	28
2.4.3	Time-Varying Group Structure	28
2.4.4	Empirical Temporal Human Proximity Data Sets	28
2.5	Temporal Network Models	35
2.5.1	Edge Activity Model	35
2.5.2	Zhao–Stehlé–Bianconi–Barrat–Model	35
3	Spreading Dynamics	37
3.1	Discrete-Time Random Walks	37
3.1.1	Random Walks on Networks	38
3.1.2	Graph Laplacian	38
3.1.3	Mixing Time	40
3.1.4	(Mean) First Passage Time	40
3.1.5	Global Mean First Passage Time	41
3.1.6	Cover Time	43

3.2	Simple Contagion: The SIS-Model	44
3.2.1	Introduction	44
3.2.2	Mean-Field SIS and the Epidemic Threshold . .	44
3.2.3	SIS on Static Networks	44
3.2.4	SIS on Temporal Networks	44
II Random Walks on Hierarchically Clustered and Other Small-World Networks		45
4	Small Worlds Are All the Same	47
4.1	Self-Similar Modular Hierarchical Random Networks	48
4.1.1	Model Definition and Structural Phases	48
4.1.2	Network Properties	50
4.2	Modified Watts–Strogatz Small-World Networks	53
4.2.1	Structural Definition and Sampling Algorithm	53
4.2.2	Network properties	55
4.3	Power-Law Small-World (PLSW) Networks	57
4.3.1	General Construction of PLSW Networks with Constant Mean Degree	58
4.3.2	Extension to a Finite Set of Distances	63
4.3.3	Extension to Categorical Distances	64
4.3.4	Application to One-Dimensional Lattice PLSW Networks	64
4.4	Structural Relation of Small-World Models	65
4.4.1	Mapping SSMH Networks to One-Dimensional PLSW Networks	65
4.4.2	Mapping modified Watts–Strogatz Networks to One-Dimensional Lattice PLSW Networks	67
4.4.3	All Small-World Network Models Appear to be of Modular Hierarchical Structure	69
4.5	Summary and Discussion	72
5	Passage Time Statistics for Locally Clustered Networks	75
5.1	Average Medium Approximation	76
5.1.1	SSMH First Passage Time Distribution and Pair-Averaged First Passage Time	76
5.1.2	Mixing Time and Small-World Effect in the Modified Watts–Strogatz Model	81
5.2	Actual Network Realizations	83
5.2.1	Numerical Results for Pair-Averaged First Passage Time	84
5.2.2	Heuristic Explanation of Minimum	85
5.3	Summary and Discussion	88
6	Cover Time for Arbitrary Complex Networks	91
6.1	Heuristic Evaluation Method	91
6.1.1	Derivation	91
6.1.2	Cover Time of Networks With Equal GMFPTs .	94
6.2	Numerical Results	95

6.3	Error Analysis	96
6.3.1	Low-Dimensionality Induced Deviations	96
6.3.2	Systematic Error: Continuous-Time Approximation	97
6.3.3	Systematic Error: Approximation of Mean Cover Time Integral	99
6.4	Cover Time on Modular Hierarchical Small-World Models	101
6.4.1	Average Medium Approximation	101
6.4.2	Actual Network Realizations	102
6.5	Summary and Discussion	104
7	Summary & Discussion	105
III Epidemics on Temporal Face-to-Face Contact Networks		
		107
8	Flockworks: Analyzing Temporal Face-to-Face Contact Networks With a Simple Model	109
8.1	The Poissonian Flockwork Model	110
8.1.1	Model Definition	110
8.1.2	Group Sizes and Equilibrium State	111
8.1.3	Equilibrium Configuration Sampling Algorithm	115
8.1.4	Node Degrees	119
8.1.5	Distributions of (Inter-)Contact and Group Life-Times	120
8.1.6	Comparing the Basic Poissonian Model to Real Data	122
8.2	Real-World Systems as Based on the Flockwork Model	124
8.2.1	Inferring Time-Varying Activity Rates and Network State Variables	124
8.2.2	Testing with Synthetic Time-Varying Flockwork Data	128
8.2.3	Application to Real-World Face-to-Face Systems	130
8.3	Summary and Discussion	139
9	The SIS-Model on Temporal Contact Networks	143
9.1	General Considerations	143
9.2	Simulations	145
9.2.1	Edge Activity Model	145
9.2.2	Poissonian Flockwork Model	145
9.2.3	Discussion	148
9.3	The Case of Infinite Infectiousness	148
9.4	Spreading on Real-World Temporal Networks	150
9.5	Discussion	150
10	Summary & Outlook	151
Bibliography		153

A Infomap: A Flow-based Algorithm to Find Modular Hierarchical Network Partitions	165
B Additional Flockwork Evaluations	169
B.1 Group Evolution	169
B.2 Rate Inference	176
C Gillespie's Stochastic Simulation Algorithm	187
C.1 Homogeneous Poisson Processes	187
C.1.1 A Single Event	187
C.1.2 Multiple Events	188
C.2 Inhomogeneous Poisson Processes	190
C.2.1 A Single Event	190
C.2.2 Multiple Events	192
C.2.3 Application to Locally Constant Rates	194
D SIS-Simulations on Temporal Network Models	197
E Fitting Distributions to Models	199

List of Figures

2.1	Example networks and their respective adjacency matrices.	24
2.2	The DTU temporal network data set.	30
2.3	The HT09 temporal network data set.	32
2.4	The HS13 temporal network data set.	33
2.5	Comparison of structural and temporal statistics across the DTU, HT09, and HS13 data sets and the influence of binning.	34
3.1	Illustration of a single time step in a random walk process on an undirected unweighted network.	38
3.2	Illustration of the random walk transition matrix as a weighted directed network and the transition matrix with a sink node.	39
4.1	Schematic hierarchical modular structure of the SSMH model.	48
4.2	The SSMH modular hierarchical phases.	50
4.3	Properties of the SSMH network model for increasing control parameter.	51
4.4	Schematic representation of the modified small-world model.	54
4.5	Network samples from the modified Watts–Strogatz model for increasing structural control parameter.	55
4.6	Schematic figure for the evaluation of the areas of summation to find the expected number of triangles per node in the modified Watts–Strogatz model.	56
4.7	Network properties of the modified Watts-Strogatz model.	57
4.8	How to construct the connection probability to achieve constant mean degree in power-law small-world (PLSW) models.	59
4.9	Pairwise distance distributions of points uniformly distributed in a square and a 2-torus.	60
4.10	Sampling procedure for the unit square PLSW model.	61
4.11	Connection probability and distance distribution of node pairs in the unit square and 2-torus PLSW models.	62
4.12	Network properties of PLSW small-world networks with underlying topology of the unit square and the 2-torus.	63
4.13	Properties of the one-dimensional lattice PLSW network model.	64
4.14	Network samples from the one-dimensional lattice PLSW model.	65
4.15	Method for redistributing the probability mass function of the one-dimensional lattice PLSW network model.	66
4.16	Comparison of the connection probability functions associated with SSMH networks and with one-dimensional lattice PLSW networks.	67

4.17	Topological phases of three small-world models as a function of their respective structural control parameters.	68
4.18	Node properties of the SSMH, the modified WS and the one-dimensional lattice PLSW network model.	69
4.19	Illustration of how the network samples generated in this chapter are analyzed for their modular hierarchical structure.	70
4.20	Results of the modular hierarchical partition detection using <i>Infomap</i> .	71
5.1	Time evolution of the random walker density at hierarchical distance from the sink for the average medium approximation of the SSMH model.	78
5.2	The SSMH average medium arrival time cdf.	79
5.3	The SSMH average medium arrival time pmf	80
5.4	Pair-averaged FPT for the SSMH model and the 1D lattice model computed from simulations.	81
5.5	The small-world effect as illustrated by the modified Watts-Strogatz clustering coefficient and mixing time.	83
5.6	Pair-averaged first passage time computed numerically for the SSMH and the 1D lattice PLSW models.	84
5.7	Pair-averaged FPT for network samples of the main small-world models introduced in the previous chapter.	85
5.8	Illustration of the heuristic method to find the global mean first passage time (FPT) of a sink node with non-zero clustering.	86
5.9	Heuristic result for the pair-averaged FPT.	87
5.10	Comparison between the actual pair-averaged FPT and the heuristic lower bound on a 2-torus PLSW model.	88
6.1	Illustration of the heuristic approach to compute the mean cover time.	92
6.2	The mean cover time of the largest component of Erdős-Rényi, Barabási-Albert, and k -regular networks.	93
6.3	The heuristic cover time evaluation method yielding results with rather large deviations from simulations on lattices.	97
6.4	Cover time pdfs and ccdfs for four example networks.	98
6.5	Mean cover times on the largest component of various real-world networks.	99
6.6	The relative error of the mean cover time heuristic is increasing with increasing mixing time per node.	100
6.7	Relative error between two integrals computing the mean cover time.	100
6.8	The mean cover time as estimated with an average medium approximation of SSMH and 1D lattice PLSW networks.	102
6.9	The mean cover time on actual network realizations of the (modular hierarchical) small-world models.	103
6.10	Relative errors between cover time simulation results and heuristic results for the small-world models.	104
8.1	The possible events of a Flockwork model.	111

- 8.2 Flockwork equilibrium configurations for increasing reconnection probability 112
- 8.3 The temporal evolution of the Flockwork model. 113
- 8.4 Configuration examples following the temporal evolution of the Flockwork model. 114
- 8.5 Flockwork equilibrium group-size distribution for increasing reconnection probability and number of nodes. 115
- 8.6 Mean group size, mean number of components and second moment of the equilibrium group-size distribution compared to asymptotics. 116
- 8.7 The distribution of numbers n_m of m -sized groups in the equilibrium state. 118
- 8.8 Group size distribution as a mean over 10^4 independently constructed equilibrium configurations using Alg. 1 for varying values of P , (**top**) $N = 10$, (**bottom**) $N = 10^3$ nodes and different algorithmic details. For each panel these details are (**top row**) iterating over a decreasingly sorted list M of group sizes m , (**bottom row**) iterating over a shuffled list M of group sizes m , (**left column**) drawing the number n_m of m -sized groups from a binomial distribution for each m , and (**right column**) drawing the number n_m of m -sized groups from a Poisson distribution for each m . The curves obtained from the sampling algorithm are denoted by markers whereas solid lines represent Eqs. (8.3). 119
- 8.9 Flockwork equilibrium degree distributions for increasing reconnection probability and number of nodes. 120
- 8.10 Flockwork equilibrium mean degree as a function of the reconnection probability and number of nodes, compared to the asymptotics. 121
- 8.11 Measurements vs. Eq. ((8.14)). Each data point is a mean over 10 independent simulations of runtime $T = 2 \times 10^6 (N\gamma)^{-1}$. 122
- 8.12 Ergodic group-size and life-time distributions of single Flockwork simulations. 123
- 8.13 Example network behavior for different positions in the Flockwork state space. 127
- 8.14 Inferring time-varying rates and the state-space trajectory of a corresponding Flockwork simulation. 129
- 8.15 Inferred time-varying rates and expected mean degree of the real-world data sets DTU, HT09 and HS13. 131
- 8.16 Inferred state-space trajectories of the real-world data sets DTU, HT09 and HS13. 133
- 8.17 Structural and statistical comparison between the DTU data set and a corresponding time-varying Flockwork surrogate. 134
- 8.18 Structural and statistical comparison between the HT09 data set and a corresponding time-varying Flockwork surrogate. 135
- 8.19 Structural and statistical comparison between the HS13 data set and a corresponding time-varying Flockwork surrogate. 137
- 8.20 The effect of binning on measuring groups. 138
- 8.21 Comparison of binned Flockwork group life-time distributions. 139

9.1	Non-zero ratio of infected in equilibrium as rescaled with the fit parameters for the edge activity model.	145
9.2	Ratio of infected in equilibrium for SIS simulations on the edge activity model.	146
9.3	Ratio of infected in equilibrium for SIS simulations on the edge activity model.	147
9.4	Non-zero ratio of infected in equilibrium as rescaled with the fit parameters for the Flockwork model.	148
B.1	Inference of the state-space trajectory from an artifical varying-rate Flockwork simulation.	176
B.2	Inference of the reconnection rate, disconnection rate, and average mean degree from an artifical varying-rate Flockwork simulation.	177
B.3	Inference of the reconnection rate from the DTU data.	178
B.4	Inference of the expected mean degree from the DTU data.	179
B.5	Inferred trajectory in the state space from the DTU data.	180
B.6	Inference of the reconnection rate from the HT09 data.	181
B.7	Inference of the expected mean degree from the HT09 data.	182
B.8	Inferred trajectory in the state space from the HT09 data.	183
B.9	Inference of the reconnection rate from the HS13 data.	184
B.10	Inference of the expected mean degree from the HS13 data.	185
B.11	Inferred trajectory in the state space from the HS13 data.	186

List of Tables

6.1	Mean cover times of simple discrete-time random walks on the largest component of various real-world networks.	95
6.2	Mean cover times of simple discrete-time random walks on subway networks of big cities.	96
D.1	Chosen values of the equilibration time t_{eq} for SIS simulations on both the Flockwork model as well as the edge activity model.	198

Nomenclature

λ_2	The unnormalized graph Laplacian's second smalled eigenvalue
ω_{N-1}	The transition matrix's second largest eigenvalue
t_{mix}	The random walk mixing time (inverse of the transition matrix's second largest eigenvalue ω_{N-1})
cdf	cumulative distribution function
MFPT	Mean First Passage Time
pdf	probability distribution function
pmf	probability mass function

1

Introduction

this is how you use citations¹.

¹ Ross2009578, 19.

Part I

Human Systems as Networks: Theory & Data

2

Networks

2.1 Networks as an Abstract Representation of Reality

Defining Networks and describing the

2.2 Network Properties and Effects

Networks have certain structural properties which can be used to, e.g. estimate the outcome of dynamic processes running on them, or to distinguish the heterogeneity of nodes in a network. Both will be relevant within the scope of this thesis. To this end, the adjacency matrix will be introduced in this section, based on which the node degree distribution and the global clustering coefficient will be defined. Furthermore, the average shortest path length and the average medium approximation of a network model will be introduced. While other properties and descriptions exist and are useful, their definitions are not of relevance for this thesis and hence the discussion will be constrained to those mentioned before.

2.2.1 Adjacency Matrix

A network is a representation of contacts between *two* entities.¹ One can therefore describe it using a so-called *adjacency matrix*, which is used to keep track of the connections and non-connections between nodes and will prove to be a useful mathematical object later on.

In a network of N nodes, the adjacency matrix A is an $(N \times N)$ -sized matrix. In the simplest case, a network is unweighted and undirected. Then, the adjacency matrix is given as

$$A_{vu} = \begin{cases} 1 & \text{if } v \text{ is connected to } u, \\ 0 & \text{otherwise.} \end{cases}$$

In an undirected network, if an edge (u, v) exists, then the edge points from node u to node v as it points from node v to node u . Hence, the adjacency matrix is symmetric as $A = A^T$. An example is displayed in the upper left panel of Fig. 2.1.

All empirical and synthetic networks discussed in this thesis will be (made) unweighted and undirected. However, since random walk transition matrices can be interpreted to be weighted and directed, the discussion is briefly extended to directed and weighted networks

¹ This distinction has to be made since there are descriptions where contacts of three or more entities are considered called *hypergraphs*, however, these are not of concern in this thesis.

	undirected	directed
unweighted	$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$	$\mathbf{A} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$
weighted	$\mathbf{A} = \begin{pmatrix} 0 & 2 & 0 & 2 \\ 2 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 1 \\ 2 & 0.5 & 1 & 0 \end{pmatrix}$	$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 2 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 \\ 0 & 0.5 & 1 & 0 \end{pmatrix}$

Figure 2.1: Example networks from the space $\{\text{undirected, directed}\} \times \{\text{unweighted, weighted}\}$ and their respective adjacency matrices. Additionally marked in pink is a self-loop which might occur.

² What “strength of the interaction” actually means heavily depends on the context in which a network description is chosen to be applied.

³ Note that the limitation of A_{vu} being positive is relatively arbitrary. It is made here since for most contexts in this thesis, some kind of a spreading probability will be proportional to this weight. In order to properly define and normalize this probability, it is necessary that this weight is non-negative. Other networks which are based on, e.g. correlations or reaction rates, might very well contain negative edge weights.

below. For weighted networks and each pair of nodes (u, v) , the entry A_{vu} of the adjacency matrix describes the strength of the interaction² between node v and node u . Usually for an entry $A_{vu} = 0$ it is implied that there is no connection between u and v . In contrast, if $A_{vu} > 0$, a connection exists and its strength is A_{vu} . ³An example for this kind of adjacency matrix is shown in the lower left panel of Fig. 2.1.

In a directed weighted network, for each pair of nodes (u, v) the entry of the adjacency matrix A_{vu} describes the strength of the interaction on node v given that the source of the interaction is node u . An entry $A_{vu} = 0$ usually implies u cannot influence v or that a commodity cannot spread from node u to node v . In contrast, if $A_{vu} > 0$, node v can be influenced by node u . Hence, the adjacency matrices are not necessarily symmetric such that $A \neq A^T$. An example is shown in the lower right panel of Fig. 2.1.

Empirical networks are often both directed and weighted. In order to analyze the contact structure only, those are often converted to undirected and unweighted networks. This is useful when, for instance, new analysis methods are developed which are usually simpler to evaluate on the latter. Using the Heaviside step function

$$\Theta_t(x) = \begin{cases} 0 & x \leq t, \\ 1 & x > t, \end{cases}$$

a weighted directed network defined by the adjacency matrix A^{WD} can be converted to an unweighted directed network A^{UD} as

$$A_{vu}^{UD} = \Theta_t(A_{vu}^{WD}),$$

Also called a *thresholding procedure* because all weights above a threshold t are converted to unweighted edges and all edge weights below or equal to the threshold will be deleted. Further simplifying the

structure, the network can be made undirected by applying the step function once again to obtain

$$A_{vu}^{\text{UU}} = \Theta_0(A_{vu}^{\text{UD}} + A_{uv}^{\text{UD}}).$$

Then, the matrix A^{UU} is symmetric and its entries consist of ones and zeros only. This procedure will be useful in Ch. 6 where empirical networks will be reduced to their undirected and unweighted counterparts in order to test the newly developed method for computing the mean cover time.

2.2.2 Node Degree and Its Distribution

In networks, nodes can play different roles

2.2.3 Clustering Coefficient

2.2.4 Path lengths

2.2.5 Percolation

2.2.6 Average Medium Approximation

2.3 Static Network Models

2.3.1 Regular Networks and Lattices

2.3.2 Erdős–Rényi Random Networks

One of the simplest stochastic network models is defined as follows: In a network of N nodes, connect each pair of nodes (u, v) with probability p_{ER} .⁴ The ensemble defined by the number of nodes N and this connection probability p_{ER} is called the $\mathcal{G}(N, p_{\text{ER}})$ random graph. Even though this definition was introduced by Gilbert in 1959 [22], it is widely attributed to the names of Erdős and Rényi, who defined a similar ensemble $\mathcal{G}(N, m)$ in the same year [21]. In their model, a fixed number m of all node pairs is chosen to be connected. However, in this thesis, only the $\mathcal{G}(N, p_{\text{ER}})$ model will be used (due to their similarity it suffices to discuss one of them) but referred to as the Erdős–Rényi (ER) random graph, as seems to be the convention.

The random graph is widely used as a reference model to find structural features of other network models and empirically observed static networks which are not caused by noise. An example is provided as follows. The clustering coefficient is relatively large in social networks, whereas the mean degree is small compared to the number of nodes (an upper bound is given by, e.g. the Dunbar number). Following the definition in Sec. 2.2.3, the clustering coefficient is the probability that two neighbors of a node are connected, as well. Since in the Erdős–Rényi model every edge is independently drawn with probability p_{ER} , the clustering coefficient is simply equal to this probability

$$C_{\text{ER}} = p_{\text{ER}}. \tag{2.1}$$

⁴ While this description suffices to define a method to sample networks from the ensemble, a more efficient sampling algorithm is given in [7]. If not stated otherwise, the latter algorithm will be used in this thesis whenever Erdős–Rényi networks are constructed.

Furthermore, a single node has $N - 1$ potential neighbors, each of which it connects to with probability p_{ER} and so its mean degree will be

$$\langle k \rangle = p_{\text{ER}} \times (N - 1). \quad (2.2)$$

Assuming that the mean degree remains constant with growing number of nodes N , the connection probability consequently vanishes and so does the clustering coefficient. Hence, social networks cannot be random.

It is furthermore possible to compute the complete degree distribution. Because a single node has $N - 1$ potential neighbors and connects to each of those with probability p_{ER} , the probability that a node has degree k is equal to the probability of k “successes” and $N - 1 - k$ “failures” in $N - 1$ Bernoulli trials which is

$$P_k = \binom{N-1}{k} p_{\text{ER}}^k (1 - p_{\text{ER}})^{N-1-k}. \quad (2.3)$$

Node degrees of the ER graph therefore follow a binomial distribution.⁵ The node degree variance is consequently given as

$$\text{Var}[k] = N p_{\text{ER}} (1 - p_{\text{ER}}) \stackrel{N \rightarrow \infty}{=} \langle k \rangle, \quad (2.4)$$

which implies that nodes are rather homogeneous concerning their number of connections.

Erdős–Rényi networks are particularly useful as a first step towards discrete systems when starting from “well-mixed” systems. Within these systems, often referred to as “mean-field” approximations, it is assumed that any two entities have uniform probability of encountering within a fixed amount of time. In physical chemistry, this approximation is often used to model the temporal evolution of the number of chemical reactants given their chemical reaction equations [23]. The Erdős–Rényi graph, due to its definitions, presents a “quenched disorder” state of such a system where connections between entities are sampled once and then held forever. It is therefore a useful structure to study how the discretization of a contact system changes the outcome of certain dynamic processes taking their course on the substrate compared to the well-mixed case. For an explicit example of this, see the discussion of epidemic spreading in Sec. 3.2.

In this thesis, the ER model will be of particular use to study how the first passage time of random walks changes when varying the substrate structure between lattices and the random graph in multiple ways (Part II).

2.3.3 Random Geometric Graphs

A two-dimensional random geometric graph (RGG) [15] can be interpreted as a random model for simple two-dimensional lattices, where nodes are connected to nodes in their vicinity. By varying the definition of “vicinity” (i.e. the connection radius), the model can be

⁵For a large number of nodes N , this distribution is asymptotically approaches the Poisson distribution and hence the model is sometimes referred to as the Poisson graph. This will, however, not be of concern within this thesis.

used to model the spread of infectious diseases with varying reach in groups of humans [68].

RGGs can be defined in the following way. Distribute N nodes on the unit square $[0, 1]^2$ where their positions are drawn randomly from a uniform distribution as $x_u, y_u \sim \mathcal{U}(0, 1)$ for each node u . Let $\mathbf{r}_u = (x_u, y_u)$ be the position vector of any node u . Then, an edge between a pair of nodes (u, v) is drawn if and only if $r_{uv} = |\mathbf{r}_u - \mathbf{r}_v| \leq R$ where R is the connection radius. These networks have a homogeneous degree distribution due to the fact that the probability of a node having k neighbors is approximately equal to the probability of exactly k nodes being inside a disc with radius R around the focal node which is of Poissonian shape. As derived in [15], in the limit of large N , the clustering coefficient is given as

$$C_{\text{2DRGG}} = 1 - \frac{1}{\sqrt{\pi}} \times \frac{1}{\Gamma(3/2)} \times \left(\frac{3}{4}\right)^{3/2}. \quad (2.5)$$

Here, $\Gamma(x)$ denotes the gamma function. Note that this result does not depend on the connection radius R . As RGGs are modeling lattices, their average shortest path length is typically higher compared to the average shortest path length of, e.g. Erdős–Rényi networks.

2.3.4 Watts–Strogatz Small-World Networks

2.3.5 Kleinberg Small-World Networks

2.3.6 Modular Hierarchical Small-World Networks

2.3.7 Barabasi–Albert Scale-Free Networks

2.4 Temporal Networks

Naturally, networks do not have to be of static nature exclusively. In fact, many static networks are derived either explicitly or implicitly from data collected over a certain period of time, i.e. they represent time-aggregated temporal data [27, 26]. Examples include social networks of humans or animals where contacts between individuals are recorded for a certain period of time (e.g. [84, 41, 5, 74]), or transportation networks, where in a static picture, edges represent possible connections between stations or stops but actual traversing between those may only occur on times when a train or a bus runs (see e.g. ref. [26], Sec. 2.8). In recent years, interest in analyzing the original temporally resolved data has increased as opposed to analyses of the aggregated static network. This often has been done in the hope of simplifying analyses by the increased amount of information considering the temporal dimension. Naturally, since temporal aggregation projects data on a lower-dimensional space, information about edges is lost when producing aggregated static networks. It has since been found, that, e.g. the inference of community structure is simplified on temporal network data [3], or that during the spread of a disease, not every node is accessible for an infection when respecting the arrow of time in contact data [34], as opposed to a

spreading disease on the corresponding aggregated network. On the other hand, [find an example here where the static picture is sufficient] [?]. In the context of temporal human proximity networks, a central question is how the “temporality” of a network influences the spread of a disease, where “temporality” is ill-defined, but the question might be interpreted to roughly mean “how different is the description as a temporal network compared to a static network for a spreading disease”. The aim of this thesis will be to investigate a new model which replicates certain properties of empirical temporal human proximity networks and to investigate the implications of this model’s properties on the outcome of epidemic spreading. One important insight will come from analysing which certain static limits of temporal networks are sufficient to be studied when investigating epidemic spreading (meaning for which regions of combinations of various important time-scales of both network change and disease spreading).

In this section, temporally resolved human proximity networks and their properties are formally defined and three data sets of empirically recorded human proximity networks are introduced. A collection of important structural and temporal observables of these data sets are measured and described in detail.

- 2.4.1 *Temporal Contact Networks*
- 2.4.2 *Contact- and Inter-Contact Times*
- 2.4.3 *Time-Varying Group Structure*
- 2.4.4 *Empirical Temporal Human Proximity Data Sets*

In recent time, two major series of experiments collected temporally resolved human proximity networks, the Copenhagen network study [74, 68, 69] and the SocioPatterns project [49, 5, 29]. A brief overview of the experiments and the data sets is given in the following; the interested reader is referred to the references cited above.

Copenhagen Network Study

The Copenhagen network study entailed measuring human proximity data, text messages, phone calls, social media contacts using customized Android smart phones which were handed out to 1,000 students of DTU Copenhagen [74, 68]. As the focus of this thesis lies on human proximity networks, these, which was measured using the phones’ built-in Bluetooth sensors. These sensors were active at all times such that once a phone was within range ($\lesssim 10\text{ m}$) of another phone, a contact was recorded and collected on a server as a tuple (i, j, t, σ) with i, j being the pair of phones, t denoting the time of initial contact and σ representing the Bluetooth signaling strength. In one of the experiment’s studies, it was found that definite thresholds can be set for σ to be representative for two phones being within certain ranges $\lesssim r$ [68]. For the data set analyzed in this thesis, the

threshold was chosen such that an adjacency matrix entry $A_{ij}(t) = 1$ means that at time t , the phones i and j were within a distance of $< 1 - 2$ m and an entry $A_{ij}(t) = 0$ in the adjacency matrix implies that at time t , the two phones were located at greater distance. This particular distance was chosen because droplets carrying airborne diseases might typically travel > 1 m when breathing indoors or > 2 m when coughing indoors [83], distances which are expected to decrease outdoors. One might wonder about the effects of under-sampling, since only contacts between those students which carried the custom phone with them were measured. However, in reference [9] it was found that unbiased sampling does not change the statistics of structural and temporal properties of proximity data sets and hence it will be assumed that under-sampling will not cause major deviations when analyzing the data.

The measurement discussed in this thesis started on $t_0 = 0$ which corresponds to Monday, 8th of September 2014, and concluded exactly one week later at $t_{\max} = 168$ h. To achieve a temporal discretization, the contacts were binned with a minimal bin size of $\Delta t = 5$ min, meaning that when a contact was active at time t , it was set to be active during the whole time bin $[n\Delta t, (n + 1)\Delta t]$ with $n\Delta t \leq t \leq (n + 1)\Delta t$ and $n \leq 168$ h/(5 min) – 1 = 2015 being a non-negative integer. Note that the actual measurements on phones were run in a way that an application on the modified smart phones measured the nearby Bluetooth devices and signal strengths for 30 s every 5 min in a desynchronized manner (meaning that the phones did not measure all at the same time). Due to this desynchronized sampling, the data can be considered to be effectively binned over 5 min even though some contacts will have been missed [68]. In total, $N = 412$ individuals participated in the measurement over this period of time. This particular measurement will be referred to as DTU in the following.

Analysing this contact data one obtains the following properties, visually summarized in Fig. 2.2. At any time of the measurement the network consists of single, densely connected groups (see Fig. 2.2a). The size of these groups varies over time with large groups existing during certain short periods of the day (e.g. lunch time at 12 to 1 pm) and smaller groups dominating during the night; however, there are always a non-vanishing number of nodes without contact. Another way to visualize this behavior is by drawing an *edge activity plot*, in which the x -axis represents time, and discrete values on the y -axis represent edge identifiers (IDs). These IDs are drawn in order of time of first appearance, meaning an edge which appeared earlier than another edge will have a smaller identifier number. Now, for each edge, a grey bar can be drawn over each interval for which it was active (see Fig. 2.2b). One generally observes that during certain times of the day, previously unobserved contacts are observed for the first time and are replicated during similar times in a cycle of 24 h. During the night, comparably less “new” contacts are observed, however, contacts which are active during the night are more

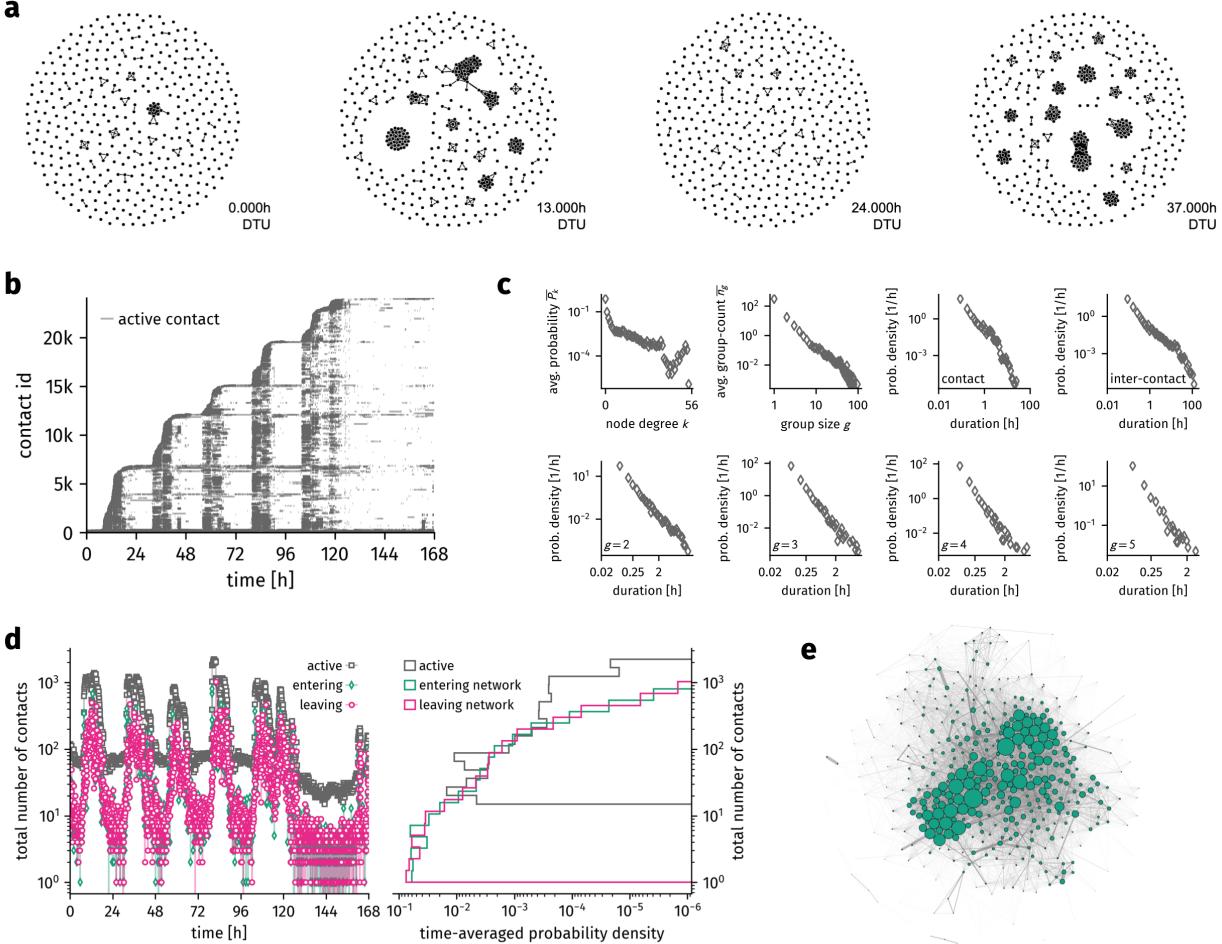


Figure 2.2: The DTU temporal network data set. (a) Snapshots at certain times of measurement. At any time, the network consists of densely connected groups, on average larger during the day than during the night. (b) A similar activity pattern is shown in an edge activity plot. During the day, more edges are active than during the night. Furthermore, more previously unobserved contacts are measured during the day. (c) Almost all structural and temporal observables show heavy-tailed behavior in their time-averaged distributions. An exemption is the degree distribution which seems to decline exponentially. (d) Active contacts and the number of contacts being (de-)activated follow a circadian pattern but are distributed declining faster than a heavy-tailed distribution. (e) The aggregated static network shows “complex” properties.

likely to remain active, or switch between their active and inactive states several times. In total, the network is less “dense” during these times meaning there are less active contacts in the network during the night than there are during the day. This effect is further reflected in Fig. 2.2d which shows the amount of active contacts at time t alongside the number of edges being activated and deactivated in the time interval $[n\Delta t, (n + 1)\Delta t]$. These curves follow a clear circadian pattern where during the day there are more edges activated and deactivated per time bin than during the night, while at the same time there are more edges active in total during the day than there are active during the night. The temporally-averaged distribution of total number of edges, edges being activated per time bin and edges being deactivated per time bin seems to follow a distribution which decays stronger than a heavy-tailed distribution. In Fig. 2.2c, several other time-averaged distributions are shown, where focus is first given to the degree- and group-size distributions as indicators for the average network structure. The time-averaged degree distribution of the network seems to decline exponentially in two regimes, where for smaller degrees it declines faster than for larger degrees. On the other hand, the group-size distribution seems to be heavy-

tailed. Temporal descriptors quantifying the network's change are given by the contact-duration, the inter-contact duration and the lifetime of groups of size m . One observes that all of these temporal distributions seem to be heavy-tailed, however the contact- and group life-time distributions generally decline faster than the inter-contact distribution. This seems to correspond to expectations from everyday life: keeping a long-lasting face-to-face contact requires more attentiveness than just staying alone and thus might be favorable to some individuals, resulting in longer time periods without contacts than periods of being in contact. Note that this point is, however, of anecdotal value at best. Finally, the aggregated network as illustrated in Fig. 2.2e shows properties which qualifies it as being called "complex", i.e. there are nodes which spent considerably more time being in contact with people than other nodes (heterogeneous node properties) and there is non-zero clustering. The network is furthermore "sparse" in the sense that the total number of unique contacts observed during the time of measurement is $\ll N(N - 1)/2$.

SocioPatterns

For SocioPatterns experiments, face-to-face contact data is measured using radio frequency identification (RFID) chips which are clipped to participants' clothes in the general chest area [5, 9]. The chips are programmed to identify radio signals distributed by other chips when they are nearby ($\lesssim 1 - 1.5$ m), subsequently sending information about those mutual occurrences to nearby RFID reading devices with higher detecting range. Since the radio signal from the used chips is shielded by human bodies, the data can approximately be assumed to represent actual face-to-face contact data. For all data sets, the temporal resolution was fixed at $\Delta t = 20$ s using the binning procedure as described above.

The first SocioPatterns data set was collected during the 'Hypertext 2009'-conference where about 75% of all participants volunteered to wear RFID chips on their conference badges, meaning the data set consists of $N = 113$ individuals [29]. The measurement started at $t_0 = 0$, which corresponds to 8am on the 29th of June 2009 and ended at $t_{\max} = 212360$ s ≈ 59 h. This data set will be referred to as HT09 in the following.

The second SocioPatterns data set consists of measured face-to-face contacts of students in a French high school, collected over the course of five days [49]. About 86% of all students present in this particular school setting participated by wearing RFID chips on lanyards, with the data set consisting of $N = 327$ individuals. The measurement started at $t_0 = 0$, which corresponds to 12pm on the 2nd of December 2013 and ended after $t_{\max} = 363580$ s ≈ 111 h. It will be called HS13 in the following.

The same analyses as for the DTU data set can be done for the SocioPatterns data sets, which are shown in Figs. 2.3 and 2.4, revealing statistical patterns which are qualitatively equal to the ones already

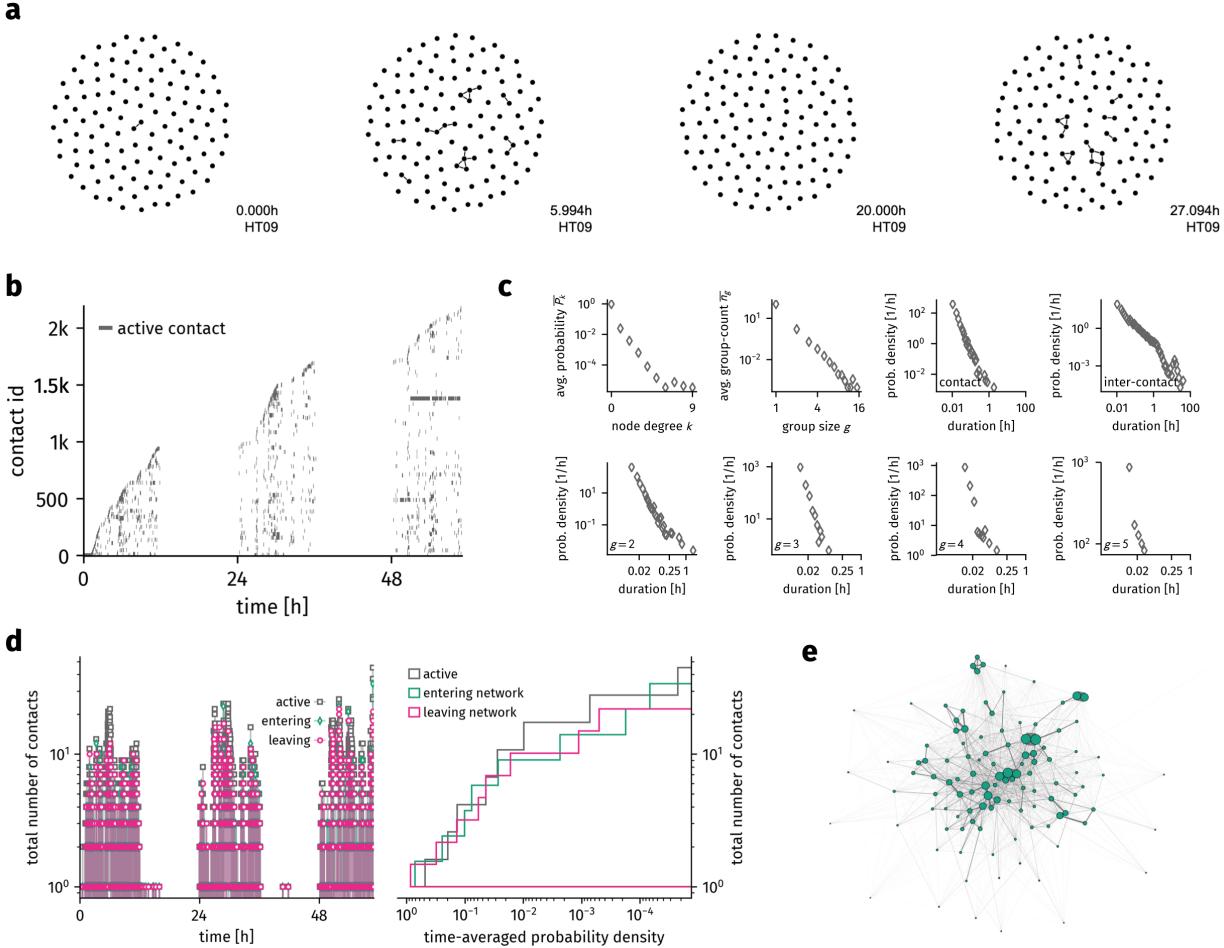


Figure 2.3: The HT09 temporal network data set. (a) Snapshots at certain times of measurement. At any time, the network consists of densely connected groups, on average larger during the day than during the night. Generally, groups are much smaller than in the DTU data set. (b) An edge activity plot. During the day, more edges are active than during the night. Furthermore, more previously unobserved contacts are measured during the day. (c) As for the DTU data, almost all structural and temporal observables show heavy-tailed behavior in their time-averaged distributions. An exemption is the degree distribution which seems to decline exponentially. (d) Active contacts and the number of contacts being (de-)activated follow a circadian pattern but are distributed declining faster than a heavy-tailed distribution. (e) The aggregated static network shows “complex” properties.

described above: The degree distribution decays exponentially, while all other discussed distributions are approximately heavy-tailed. At each time, the network consists of a set of disconnected groups which are densely connected within. Activity in the network follows a circadian pattern. However, compared to the DTU data, the SocioPatterns data sets show no activity during the night. This effect arises due to the spatial restrictiveness of the SocioPatterns experiments: all contact data can only be recorded in the vicinity of RFID readers on-site (i.e. in the school building for the HS13 data set and in the conference venue for the HT09 data set). The aggregated networks show “complex” features, i.e. node and link weight heterogeneity as well as clustering. For the HS13, the aggregated network clearly displays modular clustering, too, which is not surprising given that the students spent most of their time in separate classrooms with the same people [49].

Comparing the distributions of structural and temporal observables across data sets reveals that while the shape of the distributions seem to correspond qualitatively (excluding the average degree distribution), the DTU group size and temporal durations are generally greater than the SocioPatterns durations (see Fig. 2.5a and 2.5c). This

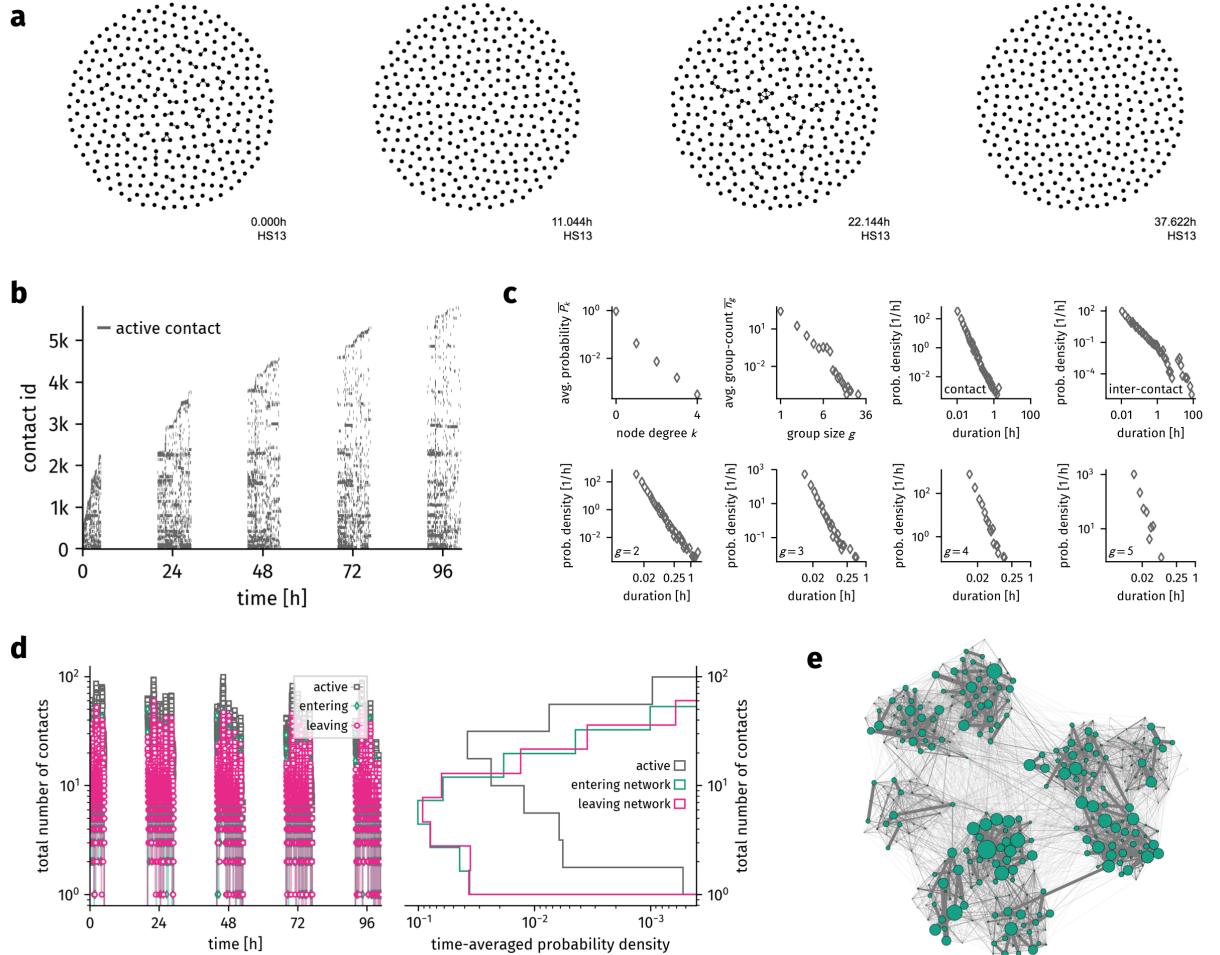
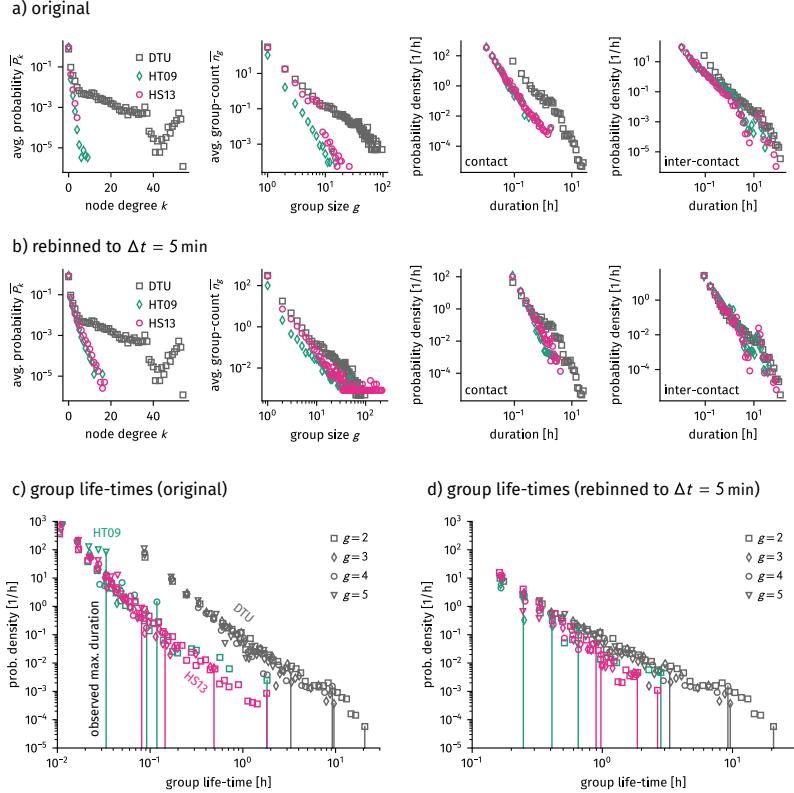


Figure 2.4: The HS13 temporal network data set. Generally, the same properties are observed as in the previously discussed data sets. (a) Snapshots at certain times of measurement. (b) An edge activity plot. During the day, more edges are active than during the night. Furthermore, more previously unobserved contacts are measured during the day. (c) Again, almost all structural and temporal observables show heavy-tailed behavior in their time-averaged distributions. An exemption is the degree distribution which seems to decline exponentially. (d) Active contacts and the number of contacts being (de-)activated follow a circadian pattern but are not distributed according to a heavy-tailed distribution. (e) The aggregated static network shows “complex” properties including modular clustering.

discrepancy vanishes when one regards the measurement process of both experiments, where contacts are counted in discrete time bins, and acknowledges that the binning time in the DTU experiment is one order of magnitude larger than in the SocioPatterns experiments. Hence, when rebinning the SocioPatterns data to the binning time $\Delta t = 5$ min of the DTU experiment, all distributions besides the degree distribution seem to increase correspondence (see Fig. 2.5b and 2.5d). This effect will be discussed in detail in Sec. . All remaining differences are suspected to be caused by the DTU experiment measuring proximity as contacts while the SocioPatterns measures face-to-face contacts. Discrepancies might also be induced by the way in which DTU contacts are measured: phones collect edges for 30 s every 5 min in a desynchronized manner. Considering that each of the experiments was conducted in a different setting with different populations it is indeed surprising that the resulting statistics show the same patterns.

A remarkable result of these analyses is that neither contact, nor inter-contact, nor any group-life time seem to have a typical time-scale as all of them are broadly distributed. In the case of groups, even the group size does not seem to have an influence on the shape of the distribution [85, 9]. One reason that this behavior can be con-

Figure 2.5: Comparison of structural and temporal statistics across DTU, HT09, and HS13 and the influence of binning.



sidered “remarkable” is that it has a direct implication on the outcome of dynamics. In dynamics with definite time-scales like the spread of diseases, one is typically interested in studying how the number of infected changes when infections or recoveries happen “faster” than the network changes. However, this question would be meaningless considering that there is no definite time-scale to which the disease time-scales can be compared to.

Nonetheless, this lack of a time-scale is not entirely accurate: the cut-off duration of the group-lifetime clearly decreases with increasing group-size, thus introducing at least one time-scale (see Fig. 2.5c and 2.5d). Also, the contact and inter-contact duration distributions show cut-off times which are well below the duration of their corresponding experiments. Furthermore, the distributions are broad but display certain deviations from a perfect power-law. All these observations might imply that the observed distributions are superpositions of several distributions associated with varying time-scales – a hypothesis which would be in agreement with the observations of circadian activity in all of the temporal contact data.

Former studies trying to explain the emergence of these heavy-tailed distributions implemented nodes building and cutting contacts using memory kernels [85, 77] where it was found that link-specific, link age-specific, and node-age specific activation and deactivation rates are necessary to generate the statistics explained above. Such time-dependent and link-specific rates are associated with a system having memory – in this case the functional dependence of

the (de-)activation rates had to follow a power-law to produce heavy tails in the contact and inter-contact duration distributions. However, the influence of circadian activity was ignored. In this thesis, instead, memory effects will be ignored and the influence of circadian activity rates on the time-averaged statistics will be explored, which entails the advantage that memory-less systems are mathematically simpler to evaluate. Furthermore, a description where temporal network change is indeed associated with different time scales which change over the course of time would allow for an interpretation of those networks traversing through different structural and temporal states. This in turn would imply that the outcome of dynamics on these networks can indeed be studied comparing network time-scales to dynamic time-scales.

2.5 Temporal Network Models

2.5.1 Edge Activity Model

2.5.2 Zhao-Stehlé-Bianconi-Barrat-Model

3

Spreading Dynamics

- Some text about spreading dynamics in general.
- Make a distinction between random walk where information does not multiply and simple contagion where information multiplies (squared term in density evolution)
- Make a distinction to complex contagion where the course of a disease might be responsible for spreading

3.1 Discrete-Time Random Walks

Random walks have been studied extensively for more than a century and emerged as an efficient descriptive model for spreading and diffusion processes in physics, chemistry, economics, biology, social sciences, epidemiology, and computer science [**Brockmann:2008hb**, 55, 8, 6, 54, 30, 50]. Because of their wide applicability and relevance to dynamic phenomena, random walk processes have become a topic of interest particularly for analyses of dynamics on complex networks [50]. The calculation of a resistor network's effective resistance [54, 82], synchronization phenomena in networks of coupled oscillators [6], the global spread of infectious diseases on the global air traffic network [**Hufnagel:2004kta**, **Brockmann:2013ud**, 28] and ranking the importance of single websites in the world wide web [56] are just a few examples of systems that have been investigated based on concepts derived from random walk theory.

Especially important is the understanding of temporal aspects of stochastic processes and how different network structures influence the equilibration process. Consequently, a lot of theoretical work focused on understanding the connection between network structure and relaxation time scales or first passage times (FPTs), the time it takes a single walker to travel from one node to another [40, 82, 38, 39]. Another important time scale is captured by the cover time of the process which is the time needed to visit *every* node in a network. This quantity has important practical applications from biology to computer science, for instance, for estimating how long it will take to distribute a chemical or a certain commodity to every node in a network or as a measure for navigability in multilayer transportation networks [18].

3.1.1 Random Walks on Networks

A simple discrete-time random walk process on an undirected network without memory is defined as follows. At time point $t \in \mathbb{Z}$, a random walker resides on a node u with degree k_u . It then chooses one of u 's neighbors, each uniform at random with probability $1/k_u$. Subsequently it jumps to the neighbor v to reside there at time $t + 1$. This process is then repeated indefinitely. It is schematically illustrated in Fig. 3.1.

From the description above, one can easily extract that mathematically, this random walk process is a Markov process with transition matrix \mathbf{W} . Given that a random walker resides on node u , the probability for the jump $u \rightarrow v$ is given as the probability that u and v are connected and the probability that v is chosen from all neighbors such that

$$W_{vu} = A_{vu} \times \frac{1}{\sum_{j=1}^N A_{ju}} = \frac{A_{vu}}{k_u}. \quad (3.1)$$

Hence the probability $p_v(t+1)$ for a walker to be on node v at time $t+1$ is given as the probability $p_u(t)$ to be on *any* node u at time t and the probability to transition from u to v . The whole process is thus governed by the master equation

$$p_v(t+1) = \sum_{u=1}^N W_{vu} p_u(t). \quad (3.2)$$

In equilibrium, walker probability has to stay constant on each node meaning that the total walker influx of a node has to equal the total walker outflux. Hence, the equilibrium walker density p_u^* has to meet

$$\underbrace{\sum_{v \in \mathcal{N}(u)} \frac{1}{k_u} p_u^*}_{=\text{outflux}} = \underbrace{\sum_{v \in \mathcal{N}(u)} \frac{1}{k_v} p_v^*}_{=\text{influx}}$$

It is simple to see that this equation is satisfied when $p_u^* = 0$ or $p_u^* \propto k_u$. The first solution is only valid for initial conditions $p_u(t=0) = 0$ for all u since \mathbf{W} is a conserving operator and hence the typical solution with non-zero initial condition is

$$p_u^* = \frac{k_u}{\sum_{v=1}^N k_v} = \frac{k_u}{2m} = \frac{k_u}{N \langle k \rangle} \quad (3.3)$$

where $\sum_{v=1}^N k_v = 2m = N \langle k \rangle$ was used to obtain the total number of edges in the network m .

3.1.2 Graph Laplacian

Another way to find the equilibrium density is by introducing the so-called unnormalized graph Laplacian. In a matrix-vector notation, the entries p_u of the column vector \mathbf{p} represent the random walker density at each node u . Let \mathbf{D} be a diagonal matrix where each

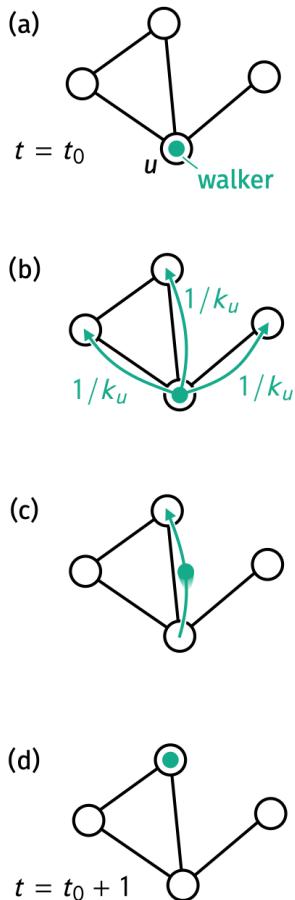
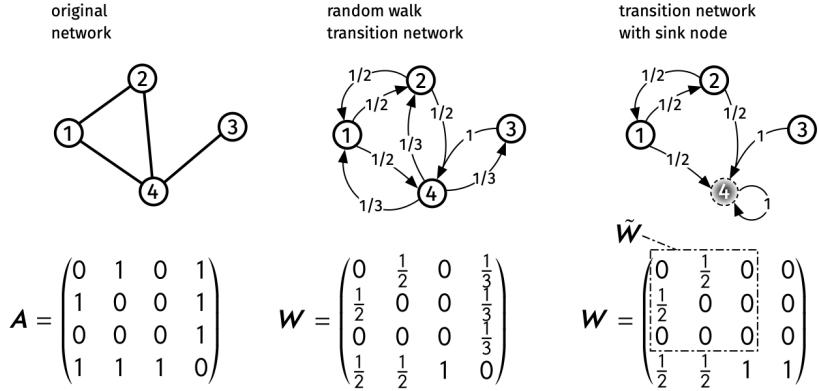


Figure 3.1: Illustration of a single time step in a random walk process on an undirected unweighted network. (a) A walker resides on a node u at $t = t_0$. (b) Each neighbor $v \in \text{Nei}(u)$ has equal probability of being chosen as the walker's target for this step. Since the probability that any of them is chosen has to sum to one, each of them is chosen with probability $1/k_u$. (c) The walker chooses one of the neighbors at random to jump to. (d) The walker now resides on the chosen node at time $t = t_0 + 1$.



diagonal entry is $D_{uu} = k_u$ and each off-diagonal entry is $D_{uv} = 0$. Then the transition matrix is given as

$$\mathbf{W} = \mathbf{AD}^{-1}$$

and the master equation can be written as

$$\mathbf{p}(t+1) = \mathbf{W}\mathbf{p}(t) = \mathbf{AD}^{-1}\mathbf{p}(t) \quad (3.4)$$

such that the equilibrium density is determined by

$$\mathbf{p}^* = \mathbf{AD}^{-1}\mathbf{p}^*.$$

Rewriting this equation gives

$$\begin{aligned} 0 &= (\mathbf{1} - \mathbf{AD}^{-1})\mathbf{p}^* \\ &= (\mathbf{D} - \mathbf{A})\mathbf{D}^{-1}\mathbf{p}^* \\ &= \mathbf{LD}^{-1}\mathbf{p}^*. \end{aligned}$$

Introducing the unnormalized graph Laplacian

$$L_{vu} = \delta_{vu}k_v - A_{vu}, \quad (3.5)$$

one recognizes that $\mathbf{D}^{-1}\mathbf{p}^*$ is an eigenvector of this matrix \mathbf{L} with corresponding eigenvalue 0. Furthermore, it has to be proportional to the column vector $\mathbf{1} = (1, 1, \dots, 1)^T$ since

$$\sum_{u=1}^N L_{vu} = k_v \sum_{u=1}^N \delta_{vu} - \sum_{u=1}^N A_{vu} = 0.$$

Thus, the equilibrium density is given as

$$\mathbf{p}^* \propto \mathbf{D}\mathbf{1}$$

such that normalization Eq. (3.3).

The graph Laplacian obtained its name from the fact that it arises as a discrete generalization of the Laplacian operator in a number of contexts where linear coupling plays a role. Examples are coupled heat baths where it generalizes the heat equation, coupled oscillators where it plays the Laplacian's role in the wave equation, and in coupled resistors where it arises from applying Kirchoff's laws to the meeting points of resistors.

Figure 3.2: Illustration of the random walk transition matrix as a weighted directed network and the transition matrix with a sink node. (Left) Given the original adjacency matrix with entries A_{vu} , (center) the corresponding entry W_{vu} in the random walk transition matrix is computed as the probability to transition from u to v which is given as Eq. (3.1). (Right) Introducing a sink at node 4, a random walker cannot leave anymore once it arrived there, represented by the column $(0, 0, 0, 1)^T$ in the matrix and by a self-loop in the illustration. The reduced transition matrix $\tilde{\mathbf{W}}$ is constructed by dropping both row and column associated with node 4.

As will be shown later, its spectral properties play important roles for all the dynamics described above as well as explicitly for the computation of passage times in simple random walks. The graph Laplacian's eigenvalues are strictly non-negative and ordered as

$$0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N.$$

It is, e.g. possible to show that the multiplicity of the eigenvalue $\lambda_1 = 0$ is equal to the number of disconnected components of the network meaning that in a single-component network one strictly finds $0 = \lambda_1 < \lambda_2$. Furthermore, the graph Laplacian's eigenvectors

$$\mu_i = (\mu_{i1}, \mu_{i2}, \dots, \mu_{iN})^T$$

will be referenced in this work where μ_i is the column eigenvector corresponding to eigenvalue λ_i . As noted above, the vector associated with the smallest eigenvalue $\lambda_1 = 0$ is $\mu_1 = (1, 1, \dots, 1)^T$.

3.1.3 Mixing Time

This section introduces the *mixing time*, a scale measuring how fast a simple random walk approaches the equilibrium solution. The random walk master equation in the form of Eq. (3.4) contains the matrix \mathbf{W} which, as a stochastic matrix, has the property that its largest eigenvalue is always $\omega_N = 1$ if the eigenvalues are ordered as $\omega_1 \leq \cdots \leq \omega_{N-1} < \omega_N$. It was already shown in Sec. 3.1.1 that the corresponding eigenvector w_N is proportional to Eq. (3.3) as $p^* = \alpha_N w_N$. Each initial condition $p(0)$ can thus be expressed as a linear combination of \mathbf{W} 's orthonormal eigenvectors w_i as $p(0) = \sum_{i=1}^N \alpha_i w_i$ such that

$$p(t) = \mathbf{W}^t p(0) = p^* + \sum_{i=1}^{N-1} \omega_i^t \alpha_i w_i.$$

Since all eigenvalues are sorted one can see that the leading term in this equation is proportional to ω_{N-1}^t and that with increasing t , all non-equilibrium contributions vanish. In fact, it is possible to show [53] that the χ^2 -distance between the time-dependent state and the equilibrium state is reduced by a constant factor approximately every

$$t_{\text{mix}} = \frac{1}{1 - \omega_{N-1}} \quad (3.6)$$

time steps, which is the time scale being referred to as *mixing time* in the following.

3.1.4 (Mean) First Passage Time

Central questions for random walks are often connected to first passage times (FPTs), e.g. the mean first passage time τ_{vu} (MFPT) between two nodes u and v , which is the mean number of time steps it takes a random walker to arrive at node v if it started at node u . The following derivations closely follow the argumentation in [54].

It is presented here since the derivations to find the mean first passage time on averaged modular hierarchical networks will be related (Sec.).

The MFPT cdf on any node can be found by introducing a deep trap (or “sink”) at the target node v , which refers to a situation where once a random walker hits the node v , it will stay there forever. Mathematically, this situation can be described by dropping the v -th row and column in the transition matrix W to obtain the adjusted transition matrix \tilde{W} as well as the v -th entry in the corresponding vectors. The adjusted master equation is then

$$\tilde{p}(t+1) = \tilde{W}\tilde{p}(t) = \tilde{W}^{t+1}\tilde{p}(0).$$

Since all walker probability can flow into the deep trap but not out of it, the probability to find a random walker being trapped in node v at time t is

$$p_v(t) = 1 - \sum_{u \neq v} p_u(t) = 1 - \tilde{\mathbf{1}}^T \tilde{p}(t).$$

The probability to hit v first at time t is then $p_v(t) - p_v(t-1)$ such that the mean first passage time $\tilde{\tau}$ is given as

$$\begin{aligned} \tilde{\tau} &= \sum_{t=0}^{\infty} t [p_v(t) - p_v(t-1)] \\ &= \tilde{\mathbf{1}}^T \tilde{D} \tilde{L}^{-1} \tilde{p}(0). \end{aligned}$$

Hence, any mean first passage times are directly related to the graph Laplacian. However, instead of working with the inverse of the reduced Laplacian, it has been shown [75] that the MFPT from start node u to end node v is

$$\tau_{vu} = \frac{1}{2} \sum_{i=1}^N k_i [R_{vu} + R_{vi} - R_{ui}] \quad (3.7)$$

where [82]

$$R_{vu} = \sum_{i=2}^N \frac{1}{\lambda_i} |\mu_{iv} - \mu_{iu}|^2. \quad (3.8)$$

Hence, finding the mean first passage time between any two nodes reduces to finding the diagonalization of the unnormalized graph Laplacian.

3.1.5 Global Mean First Passage Time

Another central time observable in the theory of random walks is the *global mean first passage time* (GMFPT) which can be used as a measure of centrality for node v – a node that is quickly reachable from anywhere in the network might be interpreted to be “important”. Combining Eqs. (3.7) and (3.8) and noticing that the GMFPT is given as the average over all start nodes yields

$$\tau_v = \frac{1}{N-1} \sum_{u \neq v} \tau_{vu} \quad (3.9)$$

$$= \frac{N}{N-1} \sum_{i=2}^N \frac{1}{\lambda_i} \left(2m\mu_{iv}^2 - \mu_{iv} \sum_n k_n \mu_{in} \right), \quad (3.10)$$

a result first derived in [39].

While the result above is exact, it is not particularly informative about how structural features of a network contribute to the value of the GMFPT. A heuristic result was derived under the assumption that a simple random walk equilibrates quickly on a network, meaning that the mixing time is $t_{\text{mix}} \ll N$. Its argumentation is outlined below in greater detail as it will provide the basis for a generalization in Sec. [XXX]. The approach introduced in [38] treats the time dependent walker concentration $p_u(t)$ on any node u at time t as separated distributions

$$p_u(t) \approx p_u p_{\text{total}}(t)$$

accounting for a time-independent distribution p_u of walker density on node u and the total density of walkers spread throughout the network $p_{\text{total}}(t)$. In this scenario, walkers are spread over the whole network quickly reaching equilibrium density and then the total amount of walkers steadily decreases through the sink node, almost equally on all nodes. The total amount of walker density leaving the network to end up in sink v is proportional to the amount of walkers leaving neighbors of i . The outflux of walker density at time step $t \rightarrow t + 1$ is hence

$$\beta_v(t) = \sum_{u \in \text{Nei}(v)} \frac{p_u(t)}{k_u},$$

meaning that the total walker density in the network is changing as

$$\begin{aligned} p_{\text{total}}(t+1) - p_{\text{total}}(t) &= \beta_v(t) = p_{\text{total}}(t) \sum_{u \in \text{Nei}(v)} \frac{p_u}{k_u} \\ &= \beta_v p_{\text{total}}(t) \end{aligned}$$

such that, by recursion,

$$p_{\text{total}}(t) = p_{\text{total}}(0)(1 - \beta_v)^t \stackrel{\beta_v \ll 1}{\approx} p_{\text{total}}(0) \exp(-\beta_v t),$$

where, with the prefactor $p_{\text{total}}(0) = 1$, this coincidentally yields the cumulative distribution function (cdf) of the first passage time at target node i as

$$p_v(t) = 1 - \exp(-\beta_v t).$$

The decay rate is given as

$$\beta_v = \sum_{u \in \text{Nei}(v)} \frac{p_u}{k_u}. \quad (3.11)$$

In order to compute the decay rate β_v and hence the GMFPT as $\tau_v = \beta_v^{-1}$, it is necessary to estimate the sink's neighbors' walker concentration p_u . Lau *et al.* propose that second neighbors i of sink node v with distance $d(v, i) = 2$ are approximated to carry equilibrium density $p_i = p_i^* = k_i/N \langle k \rangle$. Doing so gives

$$p_u = \sum_{i \in \text{Nei}(u) \setminus v} \frac{p_i}{k_i} = \frac{1}{N \langle k \rangle} (k_u - 1). \quad (3.12)$$

This includes the assumption that all walkers flowing into neighbor node u originate from second neighbors i implying that neighbors of v are not connected and hence v has local clustering coefficient $C_v = 0$. Thus, the decay rate of sink node v is given as

$$\beta_v = \frac{1}{N \langle k \rangle} \sum_{u \in \text{Nei}(v)} \frac{k_u - 1}{k_u}.$$

Replacing all k_v neighbors u with average neighbors and using $\langle k_{\text{neigh}}^{-1} \rangle = \langle k \rangle^{-1}$ which is only valid for sufficiently degree-uncorrelated networks [54] yields

$$\beta_v = \frac{k_v}{N \langle k \rangle} \left(1 - \frac{1}{\langle k \rangle} \right).$$

Hence, in this approach, the GMFPT would be given as

$$\tau_v^{\text{es}} \geq \frac{N \langle k \rangle}{k_v} \frac{1}{1 - \langle k \rangle^{-1}}. \quad (3.13)$$

It is merely a lower bound since the time steps it takes for the process to equilibrate are ignored. It was shown that the approximation of the exact GMFPT with this lower bound holds within small relative errors for Erdős–Rényi (ER) and Barabási–Albert (BA) networks, as well as for a variety of real-world networks [38].

3.1.6 Cover Time

Another temporal characteristic with practical relevance is the mean cover time T_u , defined as the mean number of steps it takes a random walker starting at node u to visit every other node at least once. For various network models simple heuristics concerning the asymptotic scaling of the mean cover time as a function of network size N have been derived [12, 13, 40]. For ER, BA and fully connected networks it was shown that

$$\langle T \rangle \sim \alpha N \log N,$$

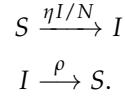
with network specific prefactor α . Here, $A_N \sim B_N$ is chosen to denote $\lim_{N \rightarrow \infty} A_N / B_N = 1$ as in Ref. [13]. Such scaling relationships are useful for comparative analyses, e.g. when networks for different sizes of the same class are compared. They are less helpful when actual expected cover times need to be computed for empirical networks where N is fixed and comparative or relative statements are insufficient. Unfortunately, a general procedure for estimating the actual cover time for arbitrary complex networks, as well as the connection between the mean cover time and FPT observables does not seem to exist.

3.2 Simple Contagion: The SIS-Model

3.2.1 Introduction

3.2.2 Mean-Field SIS and the Epidemic Threshold

For SIS processes, there's two events that can happen, which are



The first describes the spontaneous change of a healthy person to be infected, induced by the force of infection. The force of infection is the joint probability of being in contact with an infected individual $i = I/N$ and to get infected by this individual in time dt which is $P_{\text{inf}} = \eta dt$. The latter describes an event where an infected spontaneously recovers with probability $P_{\text{rec}} = \rho dt$. The respective mean-field equation for the total number of susceptibles is then

$$\begin{aligned} S(t + dt) &= S(t) - S \times i \times P_{\text{inf}} + S \times P_{\text{rec}} \\ &= S(t) - Si\eta dt + S\rho dt. \end{aligned}$$

Which means that the probability for any node to be susceptible changes as

$$\dot{s} = -\eta s(1 - s) + \rho s.$$

where $i = 1 - s$.

3.2.3 SIS on Static Networks

3.2.4 SIS on Temporal Networks

Part II

Random Walks on Hierarchically Clustered and Other Small-World Networks

4

Small Worlds Are All the Same

One of the aims of this thesis is testing the hypothesis that modular hierarchically clustered networks are in a structural optimum for random walk processes. One way to pursue this is to devise a synthetic model whose properties are closely related to the modular hierarchical structures observed in empirical networks. This structure may then be altered using a single control parameter to allow for comparisons to different known structures. Such a model, as we introduce in this chapter, is closely related to the modular hierarchical Watts–Newman network model [80] introduced in Sec. 2.3.6. Since this original model was developed to explain the small-world phenomenon, the discussion will be extended to a variety of small-world network models which all follow the same principles listed below.

1. By adjusting a single control parameter, one interpolates between generative network models which produce locally clustered structures with high regularity and the Erdős–Rényi network model.
2. During interpolation the mean degree remains constant. Since the mean degree is a first-order control parameter for certain dynamic processes, this allows for a thorough comparison of the structural consequences beyond the bias introduced by a varying total number of edges in the network.

While each of the models devised in the following have been introduced in other variants before (see Secs. 2.3.4, 2.3.6, and 2.3.5), the defining feature here will be that while conceptually different, their new definitions will allow for a thorough map between their control parameters which will reveal how, despite their apparent differences, their structural properties are very similar, finally developing a generalized model. It will further be shown that the conceptual differences of the models are irrelevant for the flow-based modular hierarchical partition algorithm *Infomap* [62, 63, 64] (method explained in App. A), suggesting that all of these small-world models produce structures which can be interpreted as being modular hierarchically clustered.

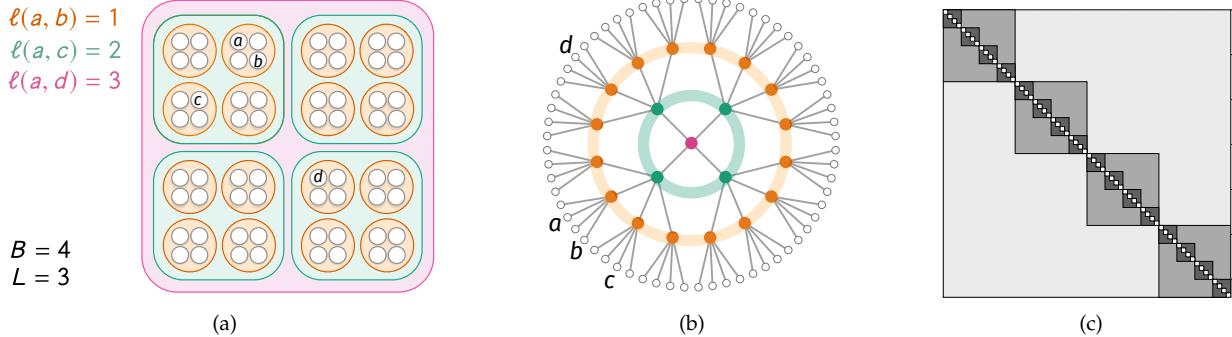


Figure 4.1: Hierarchical modular structure as introduced in Sec. 4.1, using a number of modules per level (base) $B = 4$ and total number of levels $L = 3$ as an example. (a) and (b) represent the hierarchy of modules as embedded structures and a hierarchical tree, respectively. The network nodes (white circles) are grouped into modules of B components at level 1 (orange), which are in turn grouped into B -sized modules (containing B^2 nodes) at level 2 (turquoise). Finally, these are grouped into a single level 3 structure (pink) that also contains B submodules, and thus B^3 nodes in total. The nodes labeled a and b belong to the same level-1 module, those labeled a and c are in the same level-2 module (but not the same level-1 module), and those labeled a and d , in the same level-3 module (but not the same lower-level ones). A hierarchical modular network can then be built by linking with lower probability the nodes that belong to higher level modules only. (c) The structure of the resulting adjacency matrix, with darker shades of grey representing a higher probability of having '1' entries that denote connections.

4.1 Self-Similar Modular Hierarchical Random Networks

In this section a stochastic block model is presented that can generate self-similar random networks with varying “intensity” of modular hierarchical structure (where “intensity” refers to the amount of short-ranged connections an average node has, as will be indicated). It is a variant of the model introduced by Watts *et al.* [80] briefly discussed in Sec. 2.3.6. Its definition will be shown and some of its network properties will be derived. Parts of the text given below are close to the published article [44].

4.1.1 Model Definition and Structural Phases

The model, which was defined in collaboration with C. Huepe and D. Brockmann, assumes that each node belongs to a peer group (often referred to as a module or community) that is part of a bigger community of modules, which is in turn part of a yet larger community of modules of modules, etc., as depicted in Figure 4.1a. This structure of communities and subcommunities can be represented by an underlying hierarchical tree (see Figure 4.1b), where the *hierarchical distance* ℓ between two nodes is defined as the smallest number of levels needed to traverse up the tree to find a common ancestor. Note that, in this picture, only the leaves represent actual nodes in the generated network while the rest of the tree is only used to define a modular hierarchical connectivity structure. Similar to the modified small-world model, the probability of having a connection between two nodes decreases as their distance ℓ increases. This allows for a generation of network structures composed of a hierarchy of modules, where each can be more connected internally than externally.

For simplicity, self-similar modular hierarchical structures are considered, where all modules have the same number of submodules. Thus, SSMH networks are generated starting from a B -ary tree of height L , where B is the module size and L is the total number of hierarchical layers. Then the final network consists of $N = B^L$ nodes,

which could potentially be connected to up to

$$k_\ell^{\max} = B^\ell - B^{\ell-1} = B^{\ell-1}(B-1) \quad (4.1)$$

other nodes of hierarchical distance ℓ . Although this number of potential links to other nodes grows exponentially with ℓ , in order to generate MH structures it seems intuitive to impose a limiting condition in a way that the number of connections to nodes in other modules is typically smaller than, but of similar order as, the number of connections to nodes within the same module. To achieve this, a connection probability p_ℓ is chosen that decreases exponentially with increasing hierarchical distance, given by

$$P_\ell \propto \left(\frac{\xi}{B}\right)^{\ell-1}. \quad (4.2)$$

Here, $0 \leq \xi \leq B$ is defined as the modular hierarchical *structural control parameter*. Hence, choosing different values of ξ allows for the generation of classes of networks with different degrees of hierarchical modularity. As argued before, the mean degree should remain constant for all values of ξ . To ensure this, one first has to find the mean degree of each node with respect to all other nodes at hierarchical distance ℓ , which is given by $\langle k_\ell \rangle = P_\ell k_\ell^{\max}$. The total mean degree is then $\langle k \rangle = \sum_{\ell=1}^L \langle k_\ell \rangle$, which can subsequently be used to normalize the total number of connections, obtaining

$$P_\ell = \frac{\langle k \rangle}{B-1} \left(\frac{1-\xi}{1-\xi^L} \right) \left(\frac{\xi}{B} \right)^{\ell-1}. \quad (4.3)$$

Note that if $\langle k \rangle > B-1$, the modular hierarchical structural control parameter ξ can only be chosen to be larger than or equal to $\xi_{\min} > 0$, where $(B-1)(1-\xi_{\min}^L) = \langle k \rangle(1-\xi_{\min})$, in order to have a connection probability that satisfies $P_\ell \leq 1$ for all hierarchical distances $\ell \geq 1$.

In the case of $\xi = 0$, the network consists of B^{L-1} densely connected Erdős–Rényi random networks of B nodes and connection probability $p = \langle k \rangle / (B-1)$ (note that this is, naturally, only possible if $\xi_{\min} = 0$, as indicated in the last paragraph). As ξ is increased, links are redistributed from the lowest hierarchical layer to higher ones, while keeping the total number of links constant. The system thus goes through the following phases:

- $\xi < \xi_c$: The network does not yet have a largest component of size $\approx B^L$. A numerical analysis revealed that for $\langle k \rangle \geq 4$ and $L \geq 3$, this critical parameter is given¹ as $\xi_c \approx B^{-1}$. It is thus always smaller than ξ_{SW} as defined for the next phase. An exact determination of ξ_c is left for future research.
- $\xi_c < \xi < \xi_{SW}$: The network is in a phase with *strong* hierarchical modularity, in the sense that it (i) has a largest component of size $\approx B^L$ and (ii) an average node has more links to nodes in its own lowest hierarchy module (level $\ell = 1$) than to nodes in all other

¹This can be motivated as follows. In the highest hierarchy layer L , there are B submodules. Each submodule contains B^{L-1} nodes and thus can potentially have $B^{L-1} \times (B^L - B^{L-1})$ connections to other submodules. With $\xi = B^{-1}$, the total number of connections to other submodules is $k_{\text{sub}} = P_L(\xi = B^{-1}) \times B^{L-1} \times (B^L - B^{L-1})$ (using Eq. (4.3)). Hence $k_{\text{sub}} = \langle k \rangle \times (1 - B^{-1}) / (1 - B^{-L}) \approx \langle k \rangle$. This means that on average $\langle k \rangle$ edges are leaving out of a single submodule to connect to other submodules in the highest layer L . Consequently, the larger the value of $\langle k \rangle$, the larger the likelihood that all B submodules are connected to build a large component of size $\approx B^L$ for this value of ξ .

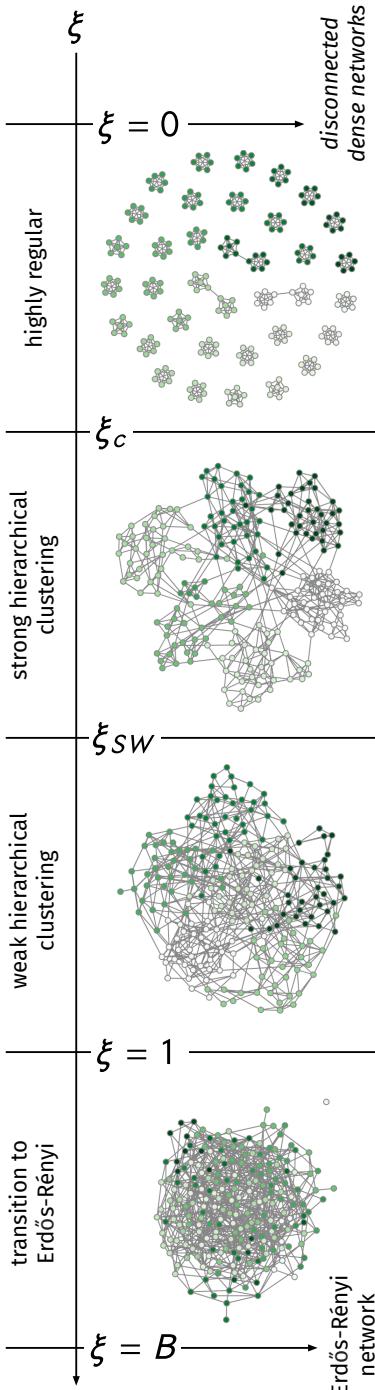


Figure 4.2: SSMH example networks for the modular hierarchical phases defined via Eqs. (4.4–4.6). Used parameters were $B = 6$, $L = 3$ and $\langle k \rangle = 6$.

hierarchy layer groups $\ell > 1$ combined, satisfying

$$\langle k_1 \rangle > \langle k_2 \rangle + \cdots + \langle k_L \rangle. \quad (4.4)$$

The value of ξ_{SW} is given by the solution to equation $0 = \xi_{SW}^L - 2\xi_{SW} + 1$ (with $1/2 \leq \xi_{SW} \leq 1$), which quickly approaches $\xi_{SW} = 1/2$ as L is increased.

- $\xi_{SW} < \xi < 1$: The network is in a phase with *weak* hierarchical modularity, in the sense that the degree at each hierarchical level is smaller than the degree at the next one, that is

$$\langle k_1 \rangle > \langle k_2 \rangle > \cdots > \langle k_L \rangle \quad (4.5)$$

is satisfied, but the condition in Eq. (4.4) is not.

- $1 < \xi < B$: In this phase, only the basic hierarchical modularity probability condition

$$P_1 > P_2 > \dots > P_L \quad (4.6)$$

is satisfied, as the network transitions to a homogeneous random structure.

- $\xi = B$: The network is identical to an Erdős–Rényi random network, as P_ℓ is equal for all layers ℓ .

Example networks for these different phases are displayed in Fig. 4.2.

In order to efficiently generate SSMH networks, an $\mathcal{O}(m)$ algorithm (where m is the total number of edges) is briefly outlined as follows. One begins with the construction of B^{L-1} Erdős–Rényi networks, each with number of nodes B and with connection probability P_1 , using the algorithm described in [7]. Those will be the base modules in layer 1. Subsequently, for every layer $\ell > 1$, one draws a number m_ℓ of edges in this layer from a binomial distribution $m_\ell \sim \mathcal{B}(m_\ell^{\max}, P_\ell)$ with parameters $m_\ell^{\max} = B^L B^{\ell-1} (B-1)/2$ and success probability $p = P_\ell$. For every edge appearing in the layer, a random node u is picked from the set of all nodes, followed by picking a second node v from all $B-1$ modules that this node u can reach in this layer. If there is not yet an edge connecting u and v , the edge is assigned, otherwise a new originating node u is picked. Implementations of this algorithm are publicly available for download at [46].

4.1.2 Network Properties

In this subsection, the number of two-stars per node, number of triangles per node, degree variance and clustering coefficient of this network model are evaluated analytically. Especially the two latter will become of importance to explain an emergent minimum in the pair-averaged first passage time in Ch. 5.

Degree Variance

The second moment of the degree distribution is given by

$$\begin{aligned}\langle k^2 \rangle &= \left\langle \left(\sum_{\ell=1}^L k_{\ell} \right)^2 \right\rangle \\ &= \sum_{\ell=1}^L \langle k_{\ell}^2 \rangle + 2 \sum_{\ell=2}^L \sum_{m=1}^{\ell-1} \underbrace{\langle k_{\ell} k_m \rangle}_{=\langle k_{\ell} \rangle \langle k_m \rangle}.\end{aligned}$$

Since the layer degrees are binomially distributed, the moments of each hierarchical layer are

$$\begin{aligned}\langle k_{\ell} \rangle &= B^{\ell-1}(B-1)P_{\ell} \\ \langle k_{\ell}^2 \rangle &= B^{\ell-1}(B-1)P_{\ell} \left(1 + B^{\ell-1}(B-1)P_{\ell} - P_{\ell} \right).\end{aligned}$$

One therefore has

$$\begin{aligned}\sum_{\ell=1}^L \langle k_{\ell}^2 \rangle &= \langle k \rangle + \langle k \rangle^2 \frac{(1-\xi)(1+\xi^L)}{(1+\xi)(1-\xi^L)} - \\ &\quad - \frac{\langle k \rangle^2}{B-1} \left(\frac{1-\xi}{1-\xi^L} \right)^2 \left(\frac{1-(\xi^2/B)^L}{1-\xi^2/B} \right)\end{aligned}$$

and

$$\sum_{\ell=2}^L \sum_{m=1}^{\ell-1} \langle k_{\ell} \rangle \langle k_m \rangle = \langle k \rangle^2 \left[\frac{1}{1-\xi^L} - \frac{1-\xi^{2L}}{(1-\xi^L)^2} \frac{1}{1+\xi} \right].$$

By combining these results, one obtains

$$\langle k^2 \rangle = \langle k \rangle + \langle k \rangle^2 \left[1 - \frac{1}{B-1} \left(\frac{1-\xi}{1-\xi^L} \right)^2 \left(\frac{1-(\xi^2/B)^L}{1-\xi^2/B} \right) \right], \quad (4.7)$$

from where the variance of the degree is found to be

$$\text{Var}[k] = \langle k \rangle - \frac{\langle k \rangle^2}{B-1} \left(\frac{1-\xi}{1-\xi^L} \right)^2 \left(\frac{1-(\xi^2/B)^L}{1-\xi^2/B} \right), \quad (4.8)$$

which is a monotonically increasing function of the structural control parameter and displayed in Fig. 4.3.

Clustering Coefficient

In the SSMH model, each node is statistically equivalent. Hence, in order to calculate the clustering coefficient as given by its definition in Sec. 2.2.3, where $C = \Delta/\wedge$, one may compute the mean number of triangles Δ per node and the mean number of two-stars \wedge per node without loss of generality for a single focal node which is set to be $i = 1$. The number of triangles per node is then given as

$$\Delta = \sum_{j>1}^{N-1} \sum_{u>j}^N P_{\ell(1,j)} P_{\ell(1,u)} P_{\ell(j,u)}.$$

Hence, for every pair of nodes (j, u) with $j \neq u \neq 1$, one calculates the probability that node 1 is connected to node j , node 1 is connected

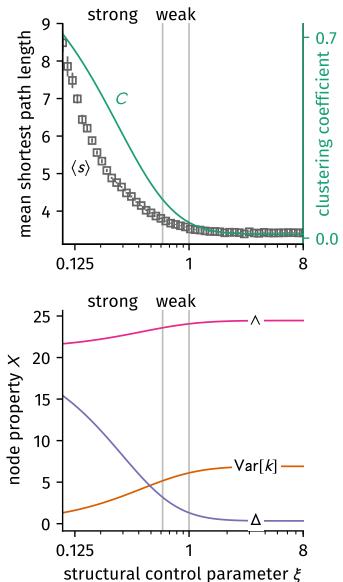


Figure 4.3: Properties of the SSMH network model for increasing control parameter ξ , here with parameters $B = 8$, $L = 3$, and $\langle k \rangle = 7$. **(Top)** The mean shortest path length $\langle s \rangle$ (as an average over the largest components from 10 independently sampled networks for each value of ξ) decreases faster than the clustering coefficient C (given via Eq. (4.11) with increasing structural control parameter ξ). **(Bottom)** The mean number of triangles per node Eq. (4.9), the mean number of two-stars per node Eq. (4.10), and the node degree variance Eq. (4.8). Note that none of the observables change significantly after leaving the weak modular hierarchically clustered phase at $\xi = 1$.

to node u , and node j is connected to node u . Analogously, the expected number of two-stars is the sum over all pairs (j, u) of the probability that node 1 is connected to j and 1 is connected to u (thus building a triad),

$$\wedge = \sum_{j>1}^{N-1} \sum_{u>j}^N P_{\ell(1,j)} P_{\ell(1,u)}.$$

Instead of summing over all node pairs, summing over pairs of hierarchical layers is preferred, as this significantly reduces the complexity of the calculation. Thus, both sums can be split into two contributions each, as

$$\Delta = \pi_S^{(3)} + \pi_L^{(3)} \quad (4.9)$$

$$\wedge = \pi_S^{(2)} + \pi_L^{(2)}, \quad (4.10)$$

where subscripts represent (S)hort-range and (L)ong-range contributions. First, the long range contributions will be evaluated by considering the following scenario. The pair $(1, j)$ has hierarchical distance ℓ_1 , whereas the pair $(1, u)$ has hierarchical distance $\ell_1 < \ell_2$. This means that nodes 1 and j share a subgroup in layer ℓ_1 and thus, if the third node u has distance ℓ_2 to node 1, so has node j distance ℓ_2 to node u . Since the problem is symmetrical, the case $\ell_1 > \ell_2$ will contribute the same amount and hence it suffices to look at $\ell_1 < \ell_2$. Now there are $B^{\ell_1}(B - 1)$ possible target nodes for node 1 in layer ℓ_1 and $B^{\ell_2}(B - 1)$ possible target nodes in layer ℓ_2 . Consequently, the total number of potential pairs in layer combination ℓ_1, ℓ_2 is $B^{\ell_1}(B - 1) \times B^{\ell_2}(B - 1)$ and the total sum evaluates to

$$\begin{aligned} \pi_L^{(3)} &= \sum_{\ell_1=1}^{L-1} \sum_{\ell_2=\ell_1+1}^L (B - 1)^2 B^{\ell_1} B^{\ell_2} P_{\ell_1} P_{\ell_2} P_{\ell_2} \\ &= \frac{\langle k \rangle^3}{B - 1} \left(\frac{1 - \xi}{1 - \xi^L} \right)^3 \frac{1}{1 - \xi^2/B} \times \\ &\quad \times \left[\left(\frac{\xi^2}{B} \right) \frac{1 - (\xi^3/B)^{L-1}}{1 - \xi^3/B} - \left(\frac{\xi^2}{B} \right)^L \left(\frac{1 - \xi^{L-1}}{1 - \xi} \right) \right]. \end{aligned}$$

The short-range contributions stem from pairs of target nodes (j, u) where node 1 has the same distance ℓ to both of them, hence they build a triad with probability P_ℓ^2 . A distinction between two cases can be made.

1. Both j and u have distance ℓ to node 1, and are part of the same subgroup. This means that their distance is $\ell' < \ell$ and that they are connected with probability $P_{\ell'}$. The total number of possible pairs of distance $\ell' < \ell$ is $(1/2) \times B^\ell(B - 1) \times B^{\ell'-1}(B - 1)$. The additional factor $1/2$ emerges to avoid double counting (with once j as source and once u as source).
2. Both j and u have distance ℓ to node 1, but are not part of the same subgroup in $\ell' \leq \ell$. This means that u is at distance ℓ of node j but the number of submodules that j can choose a neighbor from is reduced by two (its own subgroup and the subgroup of

node 1). Hence, the total number of distinct pairs of this type is $(1/2) \times B^{\ell-1}(B-2)B^{\ell-1}(B-1)$. There are $B^{\ell-1}(B-1)$ nodes to pick as first neighbor of 1 and $B^{\ell-1}(B-2)$ nodes to pick as second neighbor of 1. Again, there is an additional factor $1/2$ to avoid double counting.

Considering these cases, the short-range contribution are as follows (where the Kronecker symbol δ_{ij} is used),

$$\begin{aligned}\pi_S^{(3)} &= \frac{1}{2} \sum_{\ell=1}^L P_\ell^2 \sum_{\ell'=1}^\ell B^{\ell-1}(B-1)B^{\ell'-1}(B-1-\delta_{\ell\ell'})P_{\ell'} \\ &= \frac{1}{2} \frac{\langle k \rangle^3}{(B-1)^2} \left(\frac{1-\xi}{1-\xi^L} \right)^3 (B-2) \frac{1-(\xi^3/B)^L}{1-\xi^3/B} + \\ &\quad + \frac{B-1}{1-\xi} \left(\frac{1-(\xi^2/B)^L}{1-\xi^2/B} - \frac{1-(\xi^3/B)^L}{1-\xi^3/B} \right).\end{aligned}$$

The contributions $\pi_L^{(2)}$ and $\pi_S^{(2)}$ can be calculated with an analogous procedure as above by setting $P_{\ell_2} = 1$ and $P_{\ell'} = 1$, respectively, which yields

$$\begin{aligned}\pi_L^{(2)} &= \frac{\langle k \rangle^2 (1-\xi)}{(1-\xi^L)^3} \left[\xi \frac{1-(\xi^2)^{L-1}}{1-\xi^2} - \frac{\xi^L - \xi^{2L-1}}{1-\xi} \right], \text{ and} \\ \pi_S^{(2)} &= \frac{1}{2} \frac{\langle k \rangle^2}{B-1} \left(\frac{1-\xi}{1-\xi^L} \right)^2 \left[(B-1) \frac{1-\xi^{2L}}{1-\xi^2} - \frac{1-(\xi^2/B)^L}{1-\xi^2/B} \right].\end{aligned}$$

As the clustering coefficient is given as $C = \Delta/\wedge$, using Eqs. (4.9) and (4.10) yields

$$C = \frac{\pi_S^{(3)} + \pi_L^{(3)}}{\pi_S^{(2)} + \pi_L^{(2)}}. \quad (4.11)$$

It is a decreasing function of the structural control parameter ξ and shown in Fig. 4.3, alongside the mean number of triangles per node Δ and the mean number of two-stars per node \wedge .

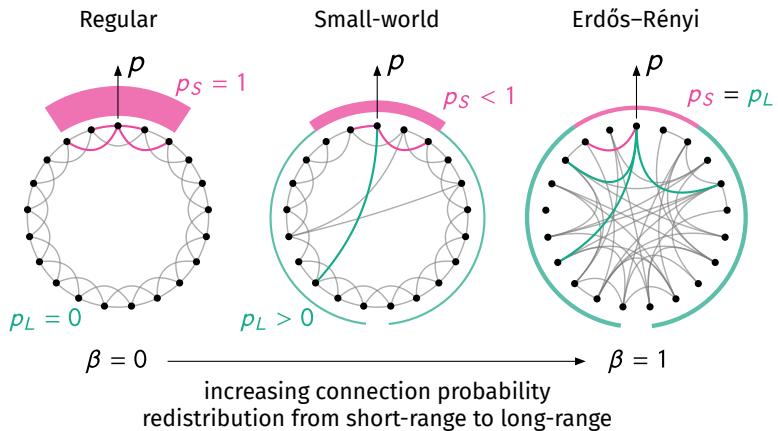
4.2 Modified Watts–Strogatz Small-World Networks

Since a modular hierarchical network model has been introduced first to explain the small-world effect in social networks [80], it seems appropriate to also study the original small-world model which only makes distinctions between short-range and long-range connections in order to study the structural and dynamical differences of the models. This section redefines a variant of the Watts–Strogatz model which happens to be structurally equivalent to a modified model already introduced by Song and Wang [71]. It is redefined below, however, to suit the needs of comparison to the other models introduced in this chapter. This section is close to the text published in [42].

4.2.1 Structural Definition and Sampling Algorithm

As described in Sec. 2.3.4, in the original model N nodes are positioned equidistantly on a ring and subsequently *locally* connected,

Figure 4.4: Schematic representation of the modified small-world model as defined in this section. Much like in the original model, the underlying structure is composed of N nodes placed equidistantly on a ring. However, instead of rewiring a previously defined lattice, each pair of nodes is connected with distance-based probability p_d where d is their minimal distance on the ring. Within distance $d \leq k/2$, nodes are connected with short-range probability p_S . For larger distances, nodes are connected with long-range probability $p_L = \beta p_S$. With increasing redistribution parameter $0 \leq \beta \leq 1$ connection probability is redistributed from the short-range regime to the long-range regime while the mean degree k is kept constant. Hence at $\beta = 0$ the short-range probability is unity while the long-range probability is zero which produces a k -nearest neighbor lattice. With increasing β , long-range “short-cuts” become more probable until at $\beta = 1$ both connection probabilities are equal and thus the model becomes equal to the Erdős–Rényi model.



i.e. connected to nodes in their vicinity with maximal lattice distance $d \leq k/2$ where k is an even positive integer and equal to the degree of every node. For the rewiring process, each node rewrites its connections to its $k/2$ rightmost neighbors to any other node in the network with probability p_r .

Within the model defined below, edges possess an inherent probability to exist, which varies for *short-range* (S) and *long-range* (L) contacts, which already draws similarities to the SSMH model devised above. A potential contact between nodes (i, j) is considered to be short-ranged if their distance in periodic boundary conditions is $d(i, j) \leq k/2$; it exists with probability p_S . It is considered long-range if $d(i, j) > k/2$ and exists with probability p_L . The distance is computed as $d(i, j) = \min(|j - i|, N - |j - i|)$. In short, two nodes with lattice distance d are connected with probability

$$p_d = \begin{cases} p_S, & \text{if } d \leq k/2, \\ p_L, & \text{otherwise.} \end{cases}$$

Hence, if $p_S = 1$ and $p_L = 0$, the model produces a structure which is equal to the original model’s starting point, a one-dimensional k -nearest neighbor lattice. On the other hand, if $p_S = p_L \equiv p$, each edge exists with probability p and hence the model is equal to the Erdős–Rényi network model. The mean degree is held constant by noticing that it is composed of a short-range degree $\langle k_S \rangle$ and a long-range degree $\langle k_L \rangle$. Each node has k potential short-range neighbors and $N - 1 - k$ potential long-range neighbors. Thus, its expected degree is $p_S k + p_L (N - 1 - k) = \langle k_S \rangle + \langle k_L \rangle = \langle k \rangle \equiv k$. To keep the mean degree constant, a *structural control parameter* β is introduced which controls the trade-off of connection probability in the short- and long-range regimes such that $p_L = \beta p_S$. Note that at $\beta = 0$, we have $p_L = 0$ and $p_S = 1$ while at $\beta = 1$ we find $p_L = p_S \equiv p$. In order for the mean degree to be constant, the distance-based probabilities are evaluated to be

$$p_S(\beta) = \frac{1}{1 + \beta(N - 1 - k)/k}, \quad (4.12a)$$

$$p_L(\beta) = \frac{\beta}{1 + \beta(N - 1 - k)/k} = \beta p_S(\beta). \quad (4.12b)$$

The short-range node degree k_S follows a binomial distribution $\mathcal{B}(k, p_S)$ and the long-range node degree k_L follows a binomial distribution $\mathcal{B}(N - 1 - k, p_L)$. A schematic explanation of the model is given in Fig. 4.4. A simple network generation algorithm is given as follows. Each node $0 \leq u \leq N - 1$ connects to each of its $k/2$ rightmost short-range neighbors with probability p_S . Afterwards, m_L long-range edges are drawn, where m_L follows $\mathcal{B}(N(N - 1 - k)/2, p_L)$. For each long-range edge, one chooses a source node u uniform at random from $[0, N - 1]$. This node is then connected to a long-range neighbor $v = (u + k/2 + z) \bmod N$ where the integer z is drawn uniform at random from the interval $[1, N - k - 1]$. If an already existing edge was chosen, repeat the procedure for this long-range edge. This algorithm has complexity $\mathcal{O}(Nk + \langle m_L \rangle)$ for sparse networks. Implementations of the algorithm are available as open source C++/Python packages [46, 47].

A connection to the original model can be drawn by noting that for small β the short-range connection probability p_S should be approximately equal to the probability that an edge has *not* been rewired $1 - p_r$ in the original model. To ensure that $p_r = 1$ for $\beta = 1$ one can set $p_r = [1 - p_S]/[1 - p_{\text{ER}}]$, which would allow to explicitly compare in what way both models produce similar or different structures with respect to observables of dynamic systems or network properties when choosing similar control parameters. Here, $p_{\text{ER}} = k/(N - 1)$ is the corresponding connection probability in an Erdős-Rényi network.

Example networks for increasing structural control parameter are displayed in Fig. 4.5.

4.2.2 Network properties

Since the node degree is given as the superposition of short-range and long-range degree, the degree variance can be found as

$$\text{Var}[k] = \text{Var}[k_S] + \text{Var}[k_L] = kp_S(1 - p_S) + (N - 1 - k)p_L(1 - p_L). \quad (4.13)$$

For increasing β both short-range and long-range variances increase, such that the degree variance is an increasing function of β , as shown in Fig 4.7b. The full degree distribution is computable by noting that any node degree is $k_i = k_{S,i} + k_{L,i}$, such that its distribution evaluates

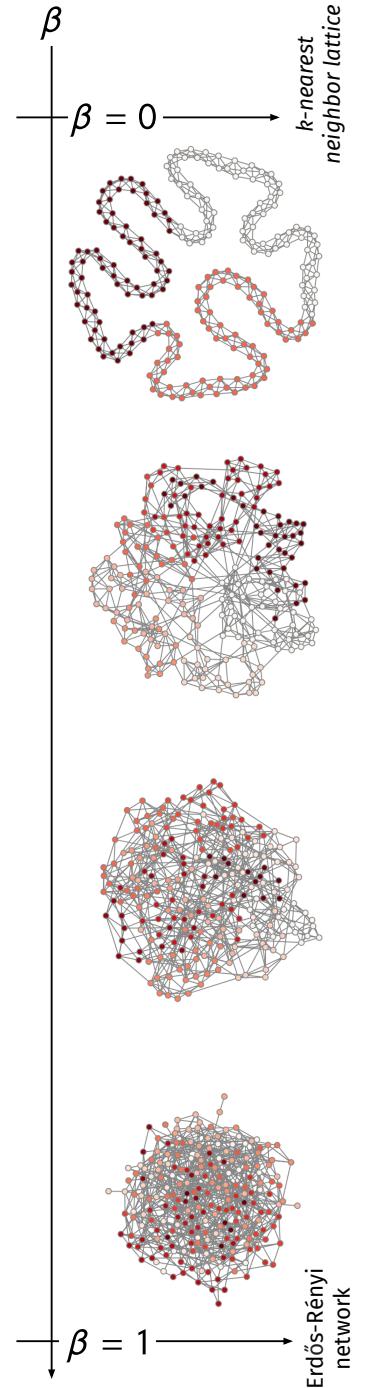
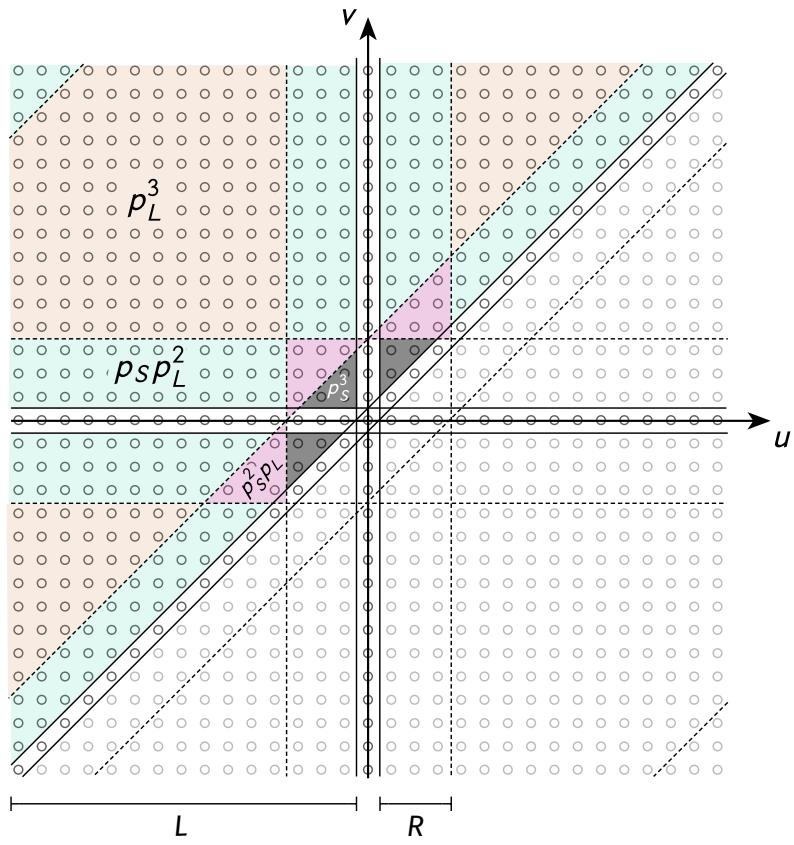


Figure 4.5: Network samples from the modified Watts-Strogatz model for increasing structural control parameter β . Parameters $N = 6^3$ and $k = 6$ were used.

Figure 4.6: Schematic figure for the evaluation of the areas of summation to find the expected number of triangles per node Δ for odd numbers of nodes N as per Eq. (4.16). Note that here, the sum has been shifted to be $\sum_{u=-(N-1)/2+1}^{(N-1)/2} \sum_{v=u+1}^{(N-1)/2} (\cdot)$ such that u and v are equal to their lattice distance to a focal node at $d = 0$. Straight lines mark the border of the areas within which $u \neq v \neq 0$, $|u| \leq k/2$, and $|v| \leq k/2$, and dashed lines mark the border of areas where $|d(u, v)| \leq k/2$.



to

$$\begin{aligned}
 p_{k'} &= \sum_{k_L=0}^{\infty} \sum_{k_S=0}^{\infty} P_S(k_S) P_L(k_L) \delta_{k', k_S + k_L} \\
 &= \sum_{k_S=0}^{\min(k', k)} \binom{k}{k_S} \binom{N-1-k}{k'-k_S} (1-p_S)^{k-k_S} \times \\
 &\quad \times p_S^{k_S} (1-p_L)^{N-1-k-k'+k_S} p_L^{k'-k_S}, \tag{4.14}
 \end{aligned}$$

which is virtually equal to the result derived in [71] (the only difference being that they let the sum run to k) and is shown in Fig. 4.7a.

The following evaluation of the clustering coefficient will rely on its definition in Sec. 2.2.3 where $C = \Delta/\wedge$. To evaluate the expected number of two-stars per node $\wedge = \langle A_{iu} A_{iv} \rangle$ one observes that a node of degree k_v is part of $(1/2)k_v(k_v - 1)$ two-stars such that the mean number of two-stars per node is given as

$$\wedge = (1/2) [\text{Var}[k] + k(k-1)], \tag{4.15}$$

which is illustrated in Fig. 4.7b.

In order to find the expected number of triangles per node Δ one recognizes again that every node is statistically equivalent. Thus, without loss of generality, one may compute the number of triangles that node 1 is part of as the sum over all possible remaining node pairs considering their connection probability based on their distance

as

$$\begin{aligned}\triangle &= \sum_{u=2}^{N-1} \sum_{v=u+1}^N p_{d(u,1)} p_{d(v,1)} p_{d(u,v)} \\ &= F p_S^3 + G p_S^2 p_L + H p_S p_L^2 + I p_L^3.\end{aligned}\quad (4.16)$$

Here, F , G , H , and I are the areas of summation where three, two, one and no node pairs are of short-range distance, respectively. As shown in Fig. 4.6, one can shift the summations to run from lattice distance $-N/2$ to distance $N/2$ around the focal node at $d = 0$ such that by marking the conditions for short-range connections, finding the respective areas reduces to a geometrical exercise. By defining the lengths $L = (N - 1)/2$ and $R = k/2$ one first finds the useful unit of a short-short-long-range area as the triangle $T = (R^2 - R)/2 + R$. Then, the areas of summation are

$$\begin{aligned}F &= 3(T - R) = (3k/8)(k - 2) \\ G &= 3T = (3k/8)(k + 2) \\ H &= 2((L - R)R - T) + T + 2(L - R)R + 2((L - 1)R - T) \\ &= (k/8)(12N - 26 - 11k) \\ I &= (L - R)^2 - 2((L - 1)R - T) - (L - R) + (L - R)^2 - T \\ &= (1/8) [5k^2 - k(12N - 26) + 4(N^2 - 3N + 2)].\end{aligned}$$

The expected number of triangles decreases with increasing β , as expected and as shown in Fig. 4.7b. Considering Eqs. (4.12), the clustering coefficient is consequently given by

$$C(\beta) = p_S^3 \times \frac{F + G\beta + H\beta^2 + I\beta^3}{(1/2)\text{Var}[k] + k(k - 1)}. \quad (4.17)$$

In the respective limits one finds

$$\begin{aligned}C(\beta = 0) &= \frac{3(k - 2)}{4(k - 1)} \\ C(\beta = 1) &= \frac{\sum_{u=2}^{N-1} \sum_{v=u+1}^N p^3}{\sum_{u=2}^{N-1} \sum_{v=u+1}^N p^2} = p,\end{aligned}$$

which are the expected results for both the k -nearest neighbor lattice as well as the Erdős-Rényi graph. Further considering Eqs. (4.12) and (4.17) as well as noting that $\text{Var}[k](\beta \rightarrow 0) = 0$, in the limit of small long-range redistribution one finds

$$\frac{C(\beta \ll 1)}{C(0)} \approx p_S^3 = 1 - 3\beta \frac{N - k - 1}{k} + \mathcal{O}(\beta^2) \quad (4.18)$$

which will be of importance for quantifying the small-world effect in Sec. 5.1.2.

4.3 Power-Law Small-World (PLSW) Networks

In the original Kleinberg model [31], N nodes are embedded in a low-dimensional space and positioned as a lattice, such that they are

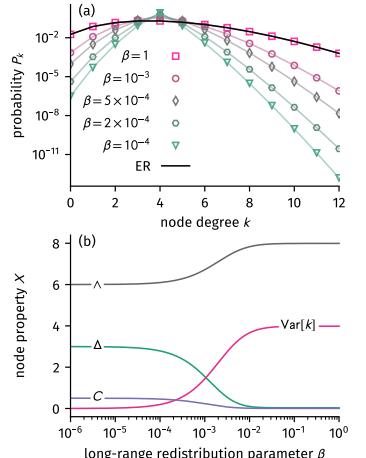


Figure 4.7: Analytic results for (a) degree distribution Eq. (4.14) and (b), expected number of two-stars per node Δ (Eq. (4.15)), expected number of triangles per node \triangle (Eq. (4.16)), clustering coefficient $C = \triangle/\Delta$ and node degree variance Eq. (4.13). While both degree variance and number of two-stars increase with increasing long-range redistribution parameter β , the number of triangles as well as the clustering coefficient decrease. The results shown here were computed for $N = 1001$ and $k = 4$.

connected to their nearest neighbors. Additional long-range links are then added with probability P that decays as a power-law with lattice distance d between two nodes, following $P \propto d^{-\kappa}$, with $\kappa > 0$. This procedure is limited in its application as it constrains the basic topology to low-dimensional lattices and does not account for a constant expected number of generated edges. Furthermore, it does not approach the Erdős–Rényi random graph with $\kappa = 0$, which would be favorable for comparability of those models. For these network models, the power-law exponent $-\kappa$ is the *structural control parameter* which controls the amount of long-range edges and their distance distribution.

In the following, a generalized variant of Kleinberg's model is considered where all pairs of nodes are connected using a connection probability function which decays as a power-law with their distance in an embedding space (thus, neighbors on lattice points are not automatically connected). It will be shown that edge distances do not necessarily have to be distances and thus the embedding topology is not of importance to be known. The model is further generalized to categorical distances where edges can belong to categories associated with positive integers or real numbers which is subsequently used to define a modified Kleinberg model using a one-dimensional ring lattice as an embedding discrete space.

4.3.1 General Construction of PLSW Networks with Constant Mean Degree

In the following, a procedure is described which allows to keep the mean degree fixed when varying a distance-based connection probability $P(r)$. The model initially assumes that nodes are distributed in some arbitrary low-dimensional space. Since nodes will be connected with a distance-based connection probability $P(r)$, it furthermore requires that there exists a metric which allows to compute distances r between nodes, such that the distribution of distances between them is captured by the connection density $f(r)$ where $0 \leq r \leq r_{\max}$ with

$$\int_0^{r_{\max}} dr f(r) = 1.$$

The connection density is used to evaluate the node pair density $\mathcal{M}(r)$ at distance r which is needed to evaluate how many potential edges are expected to be found for each absolute distance r such that the connection probability $P(r)$ can be normed to produce a constant mean degree afterwards. The node pair density has to integrate to the total number of node pairs, which, in a system of N nodes, evaluates to $(1/2)N(N - 1)$, such that

$$\int_0^{r_{\max}} dr \mathcal{M}(r) = \frac{1}{2}N(N - 1)$$

and hence

$$\mathcal{M}(r) = \frac{1}{2}N(N-1)f(r).$$

Now, as per definition of the model, the connection probability $P(r)$ should roughly assume the shape of a power-law $C_\kappa r^{-\kappa}$. Let $\epsilon \ll r_{\max}$ be a minimal radius where nodes within a distance of ϵ will always connect, then the connection probability of a node pair at distance r is defined as

$$P(r) = \begin{cases} 1 & 0 < r \leq \epsilon, \\ C_\kappa r^{-\kappa} & \epsilon < r. \end{cases} \quad (4.19)$$

The prefactor C_κ can be found by fixing the expected total number of edges in the network as given by $\langle m \rangle = 2 \langle k \rangle / N$ with $\langle k \rangle$ being the mean degree. The corresponding integral is

$$\begin{aligned} \langle m \rangle &= \frac{2 \langle k \rangle}{N} = \int_0^{r_{\max}} dr \mathcal{M}(r) P(r) \\ &= \int_0^\epsilon dr \mathcal{M}(r) + \int_\epsilon^{r_{\max}} dr \mathcal{M}(r) C_\kappa r^{-\kappa} \end{aligned} \quad (4.20)$$

which evaluates to

$$C_\kappa = \frac{\frac{\langle k \rangle}{N-1} - \int_0^\epsilon dr f(r)}{\int_\epsilon^{r_{\max}} dr f(r) r^{-\kappa}}. \quad (4.21)$$

In many systems and especially for $\kappa \ll 1$ the prefactor C_κ will produce connection probabilities that exceed unity for small distances. This in turn will cause missing edges because those distances r with generation probability $P(r) > 1$ can not contribute more edges than given by the number of node pairs $\mathcal{M}(r)dr$. In order to avoid that edges cannot be constructed due to this circumstance, a connection probability redistribution procedure is introduced as follows.

For all distances where $P(r) > 1$, the total amount of non-generated edges $\mathcal{M}(r)(P(r) - 1)$ is summed up. This excess probability is then redistributed to higher distances until a critical radius r_c is reached, meaning that within this radius, edges are produced from all node pairs (i.e. with probability 1) while outside of this radius edges are produced according to the demanded power-law with exponent $-\kappa$. The connection probability thus changes to

$$P(r) = \begin{cases} 1, & \text{if } 0 < r \leq r_c, \\ C_\kappa r^{-\kappa}, & \text{otherwise,} \end{cases} \quad (4.22)$$

where

$$r_c = \begin{cases} \epsilon, & \text{if } C_\kappa \epsilon^{-\kappa} > 1, \\ r_{\min}, & \text{otherwise.} \end{cases}$$

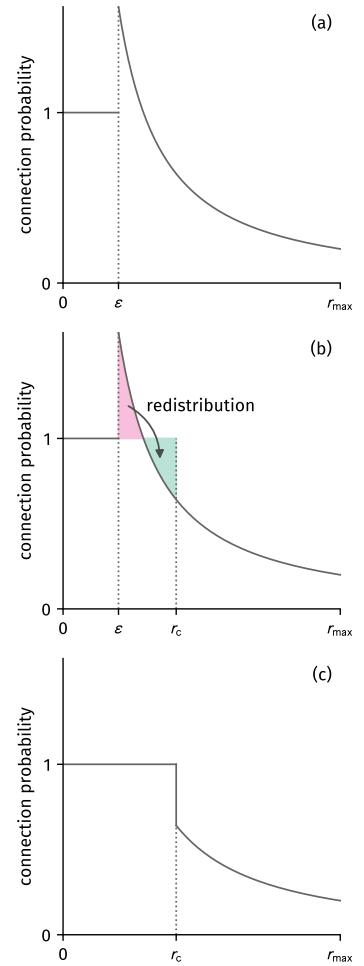


Figure 4.8: How to construct the connection probability such that the expected number of generated edges will be constant for each exponent κ . (a) Start with power-law Eq. (4.19). Note that in the shown case the connection probability exceeds $P = 1$ such that this excess probability cannot generate any edges. (b) Find the critical distance r_c at which the excess number of edges which cannot be generated will be balanced by redistributing their generation probability to higher distances (see Eq. (4.23)); in this particular example a geometry was chosen where the distance distribution $f(r)$ is uniform. (c) The adjusted (final) connection probability.

The minimal radius r_{\min} is determined by the balance equation

$$\int_{\epsilon}^{r_{\min}} dr f(r) [C_{\kappa} r^{-\kappa} - 1] = 0 \quad (4.23)$$

with the condition $r_{\min} > \epsilon$, which can be solved numerically with Newton's method for finding zeros and Simpson's rule for computing integrals.² This procedure is illustrated in Fig. 4.8.

Generating networks using the connection probability defined above will result in homogeneous networks interpolating between two limit structures. In the first limit $\kappa \rightarrow \infty$, the connection probability is

$$\lim_{\kappa \rightarrow \infty} P(r) = \begin{cases} 1, & \text{if } r \leq r_c \\ 0, & \text{otherwise,} \end{cases}$$

meaning that the generated network will be equal to a random geometric graph (RGG) on the embedding topology with properties $\mathcal{M}(r)$ and r_{\max} (RGGs were introduced in Sec. 2.3.3). In the other limit $\kappa = 0$, it becomes clear that the connection probability becomes distance-independent as $P(r) = C_0 r^0 = C_0$. Furthermore, taking the limit $\epsilon \rightarrow 0$ in Eq. 4.21, yields $C_0 = \langle k \rangle / (N - 1)$ and thus the connection probability is

$$\lim_{\kappa \rightarrow 0, \epsilon \rightarrow 0} P(r) = \frac{\langle k \rangle}{N - 1} = p_{\text{ER}}.$$

Due to its independence of the distance r , the resulting model is the Erdős–Rényi random graph model with uniform connection probability p_{ER} .

Note that as per Eq. (4.20) the distance of generated edges in the sampled network will be distributed according to

$$\mathcal{D}(r) = \frac{2}{N \langle k \rangle} \mathcal{M}(r) P(r). \quad (4.24)$$

Application to Uniform Node Distribution on a 2-Torus and a Square

As a first, simple application, the unit square $[0, 1]^2$ with periodic boundary conditions (i.e. a 2-torus) will be discussed as the embedding space in the following. On this space N nodes are distributed with uniform probability (i.e. their positions x and y follow the same uniform random distribution $x, y \sim \mathcal{U}(0, 1)$). Since the torus is translationally invariant, each node can be assumed to be centered. Hence it suffices to consider a single node and compute the density of nodes around it at distance r . The problem is radially symmetric until radius $r_1 = 1/2$. Then, for $r > r_1$, each remaining corner of the square around the central node is congruent. Hence, the distance density $f(r)$ can be found by considering the split integral

$$1 = \int_0^{1/\sqrt{2}} dr f(r) = 4 \int_0^{1/2} dr \int_0^{\pi/2} d\theta r + 4 \int_{1/2}^{1/\sqrt{2}} dr \int_{\arccos(1/2r)}^{\pi/2 - \arccos(1/2r)} d\theta r$$

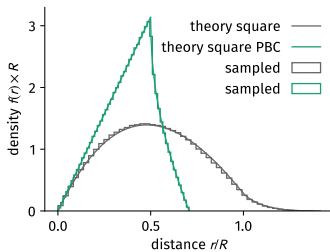
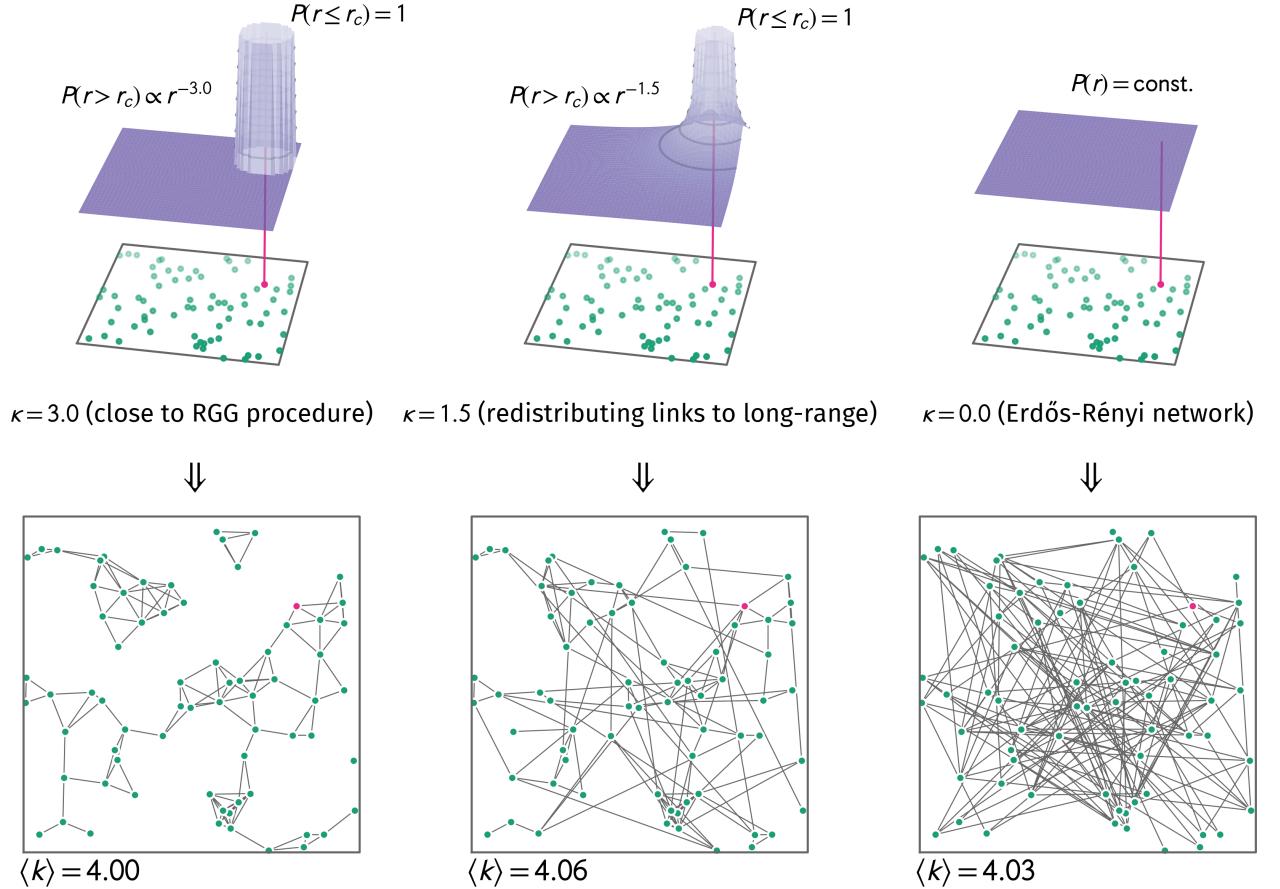


Figure 4.9: Pairwise distance distributions of points uniformly distributed in a square $[0, R]^2$ with and without periodic boundary conditions. Sampling was done for $N = 1000$ points.



which yields

$$f(r) = \begin{cases} 2\pi r, & \text{if } 0 \leq r \leq \frac{1}{2} \\ 2\pi r - 8r \arccos\left(\frac{1}{2r}\right), & \text{if } \frac{1}{2} < r \leq \frac{1}{\sqrt{2}}. \end{cases} \quad (4.25)$$

This distribution is shown in Figure 4.9.

Similarly, the distance distribution of two random points in the unit square can be found by considering no periodic boundary conditions. This distribution was derived in [58] to be

$$f(r) = 2rg(r^2) \quad (4.26)$$

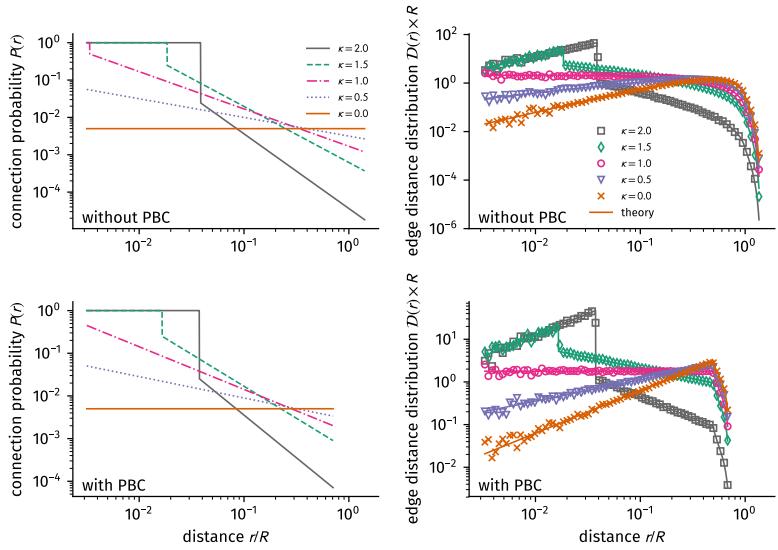
with

$$g(s) = \begin{cases} -4\sqrt{s} + \pi + s & 0 < s \leq 1, \\ -2 - \pi - s + 4\arcsin(s^{-1/2}) + 4\sqrt{s-1} & 1 < s \leq 2. \end{cases}$$

For this embedding geometry, example networks and the connection probability function defined in Sec. 4.3.1 are shown in Fig. 4.10. As one can see for $-\kappa \ll 0$ the procedure generates structures which are essentially equal to samples from the random geometric graph model. With increasing power-law exponent $-\kappa$, the connection probability broadens towards higher distances which introduces more long-range links until each edge is equally probable at $\kappa = 0$ to obtain the Erdős–Rényi model.

Figure 4.10: Sampling procedure for nodes embedded in $[0,1]^2$ without periodic boundary conditions for parameters for $N = 70$, $\langle k \rangle = 4$, $\epsilon = 10^{-9}$ and varying power-law exponent κ . **(Top row)** Illustration of the distance-varying connection probability function for increasing power-law exponent $-\kappa$ for a single node (pink) as defined by Eqs. (4.22), (4.23), and (4.26). **(Bottom row)** Corresponding generated networks. For connection probability strongly decreasing with distance ($-\kappa \ll 0$), the sampling procedure essentially generates a random geometric graph. With increasing power-law exponent $-\kappa$ more long-range links are produced until the network model equals an Erdős–Rényi random graph at $\kappa = 0$.

Figure 4.11: Connection probability for random points in a square of side length R (**top row**) with and (**bottom row**) without periodic boundary conditions. (**Left column**) power-law connection probability $P(r)$ as defined by Eqs. (4.22) and (4.23), (**right column**) edge distance distribution $D(r)$ (see Eq. (4.24)) with distance distributions (**bottom row**) Eq. (4.25) and (**top row**) Eq. (4.26). Parameters of choice were $N = 2000$, $\langle k \rangle = 10$, $\epsilon = 10^{-9}$, and $\kappa \in \{2, 1.5, 1, 0.5, 0\}$. For each parameter combination 20 networks were sampled.



As a numerical validation of the considerations in Sec. 4.3.1, networks were generated using the parameters $N = 2000$, $\langle k \rangle = 10$, $\epsilon = 10^{-9}$, and $\alpha \in \{2, 1.5, 1, 0.5, 0\}$ and the corresponding connection probability functions as defined by Eqs. (4.22) and (4.23), given the distance distributions Eqs. (4.25) and (4.26), respectively. The sampling algorithm was simply as follows. Start with an empty network of N nodes where each node u has coordinates $0 \leq x_u < 1$ and $0 \leq y_u < 1$, both drawn from a uniform distribution $\mathcal{U}(0, 1)$. For each pair of nodes (u, v) add the edge (u, v) to the network with probability $P(r_{uv})$ where $r_{uv} = \sqrt{(x_u - x_v)^2 + (y_u - y_v)^2}$ for the simple unit square and $r_{uv} = \min \left\{ \sqrt{(x_u - l_1 x_v)^2 + (y_u - l_2 y_v)^2} : (l_1, l_2) \in \{-1, 0, 1\}^2 \right\}$ for the 2-torus. For each parameter combination 20 networks were generated, subsequently building a log-binned histogram to find the distribution of distances from the drawn edges. The numerically found edge distance distributions are indeed in close agreement³ with Eq. (4.24), as shown in Fig. 4.11.

Furthermore, the generated networks should reproduce the small-world effect as described in Sec. 2.3.4, meaning that the average shortest path length should decline faster than the mean clustering coefficient with increasing power-law exponent $-\kappa$. Additionally, the clustering coefficient is expected to interpolate between the results for the random geometric graph Eq. (2.5) and the Erdős–Rényi graph Eq. (2.1). To test these hypotheses, the corresponding network properties were computed as averages over 50 network samples for each parameter combination of $N = 1000$, $\langle k \rangle = 10$, $\epsilon = 10^{-9}$ and increasing power-law exponent $-3 \leq -\kappa \leq -1/2$. Both the unit square and the 2-torus were tested as underlying topologies. Note that the average shortest path length $\langle s \rangle$ was only computed on the largest connected component of a network which was always of size $\approx N$.

The numerical results are shown in Fig. 4.12. As expected, the

³ Small discrepancies are observed for short normalised distances r/R and which is due to the small number of network samples and consequently small number of node pairs of short distances.

small-world effect is displayed with increasing power-law exponent, meaning that the average shortest path length decreases faster than the clustering coefficient with increasing introduction of long-range links. Note that the expected theoretical limit results of the clustering coefficient are reached (with higher accuracy for the 2-torus since the periodic boundary conditions prevent geometrical deviations). The mean degree $\langle k \rangle$ remains constant (bar rather small fluctuations) for varying control parameter, as demanded per construction. Surprisingly, both the average number of two-stars per node as well as the degree variance remain constant (small deviations for the square-topology without periodic boundaries, respectively). However, since in the RGG limit the probability to connect to k nodes follows a binomial distribution with the success probability being equal to the disk area of radius r_c around a focal node, the variance might be indeed expected to be equal to that of the Erdős–Rényi model, hinting at the fact that it simply stays constant.

4.3.2 Extension to a Finite Set of Distances

The derivation above is based on the assumption that the embedding topology is known. Assuming, instead, that the geometry of the embedding space is unknown but one has knowledge about a set of nodes and the distance of each pair of nodes, one may proceed as follows. Let $\mathbf{r} \in (\mathbb{R}^+)^{N(N-1)/2}$ denote the increasingly sorted vector of all distances between pairs of nodes r_{uv} .

Then, the normalization condition is

$$\sum_{e=1}^{N(N-1)/2} C_\kappa r_e^{-\kappa} = \frac{N \langle k \rangle}{2}$$

such that the prefactor C_κ evaluates to

$$C_\kappa = \frac{N \langle k \rangle}{2 \sum_{e=1}^{N(N-1)/2} r_e^{-\kappa}}.$$

Again, for small distances and small κ , the connection probability $P(r)$ will exceed unity. Hence, a minimum entry l of the vector \mathbf{r} has to be chosen below which every distance will produce an edge while above this entry, edge creation will follow a power-law with exponent κ . This entry l is given as the last entry meeting the inequality

$$\sum_{e=1}^l [C_\kappa r_l^{-\kappa} - 1] > 0.$$

Subsequently, edges can be drawn according to their index number in the vector \mathbf{r}

$$P_e = \begin{cases} 1 & e \leq l \\ C_\kappa r_e^{-\kappa} & e > l. \end{cases}$$

It is, however, of crucial importance to keep a mapping between the node pair (u, v) and the edge entry in the sorted vector \mathbf{r} . Furthermore, there will be an obvious bias when there's multiple edges of

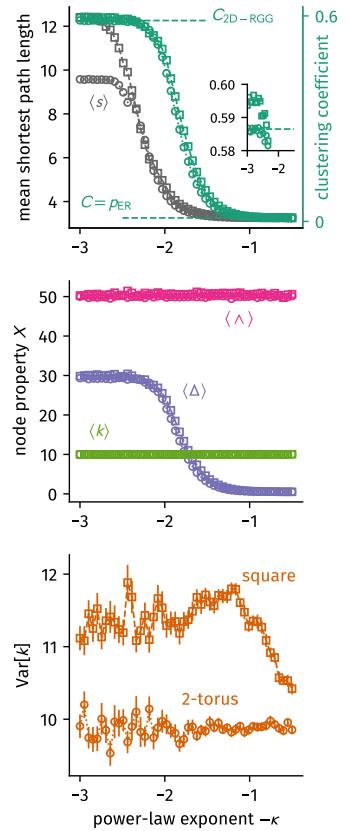


Figure 4.12: Network properties of PLSW small-world networks with underlying topology of the unit square (curves marked with \square) and the 2-torus (curves marked with \circ). Each point is the mean over 50 independent realizations for $N = 1000$, $\langle k \rangle = 10$, $\epsilon = 10^{-9}$ and increasing power-law exponent $-3 \leq -\kappa \leq -1/2$. For the mean shortest path length $\langle s \rangle$, the largest connected component of the generated network was used. **(Top)** The mean shortest path length decreases faster than the clustering coefficient, thus displaying the small-world effect. The expected theoretical values for the clustering coefficient of both limit structures are reached. Inset: The results for the 2-torus better match the theoretical value of Eq. (2.5) than the results for the unit square since for the latter case and finite N the hard boundary skews the overlap of circles for boundary-near nodes. **(Middle)** Mean degree and number of two-stars per node remain constant while the number of triangles per node decreases. **(Bottom)** The degree variance of the networks generated with the 2-torus topology remains constant.

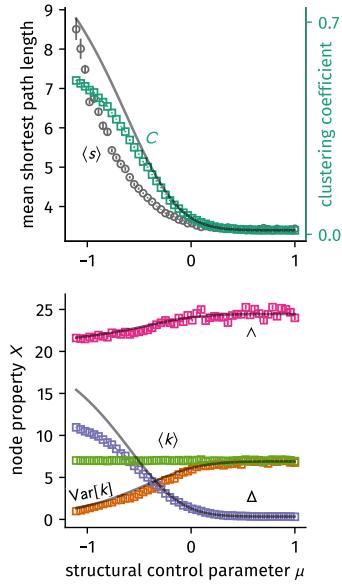


Figure 4.13: Properties of the one-dimensional lattice PLSW network model for increasing control parameter μ , here with parameters $N = 8^3$, and $\langle k \rangle = 7$. The observables displayed here were obtained as an average over 10 independently sampled networks for each value of μ . **(Top)** The mean shortest path length $\langle s \rangle$ was computed on the largest component of the sampled networks and decreases faster than the clustering coefficient $\langle C \rangle$, which is compared to the corresponding SSMH clustering coefficient given via Eq. (4.11) and using the map of the control parameters Eq. (4.31). The SSMH clustering coefficient estimates this PLSW clustering coefficient well for larger values of the control parameters but deviates when approaching the regular structures (since those are fundamentally different). **(Bottom)** The mean number of triangles per node (compared to the SSMH result Eq. (4.9)), the mean number of two-stars per node (compared to the SSMH result Eq. (4.10)), the node degree variance (compared to the SSMH result Eq. (4.8)), and the mean degree $\langle k \rangle$. Note that none of the observables change significantly after leaving the weak modular hierarchically clustered phase at $\mu = 0$.

critical distance r_l . Note that if the pair-wise distances are sufficiently distinct from each other, l can be directly mapped to its distance r_l such that the connection probability may be written down as a dependence on r as

$$P(r) = \begin{cases} 1 & r \leq r_l \\ C_\kappa r^{-\kappa} & r > r_l. \end{cases}$$

At this point it might be important to note that r does not need to be a “distance” anymore in the sense that the values r_{uv} do not have to meet the triangle inequality. Instead, these values could stem from semimetrics but also represent rank values or other positive real numbers, i.e. general edge weights.

4.3.3 Extension to Categorical Distances

For the considerations above it was assumed that distances are positive real numbers and hence distributed according to a probability density function. However, there exist cases where distances are integer numbers (e.g. for lattices) or not even distances but just categorizations of edges and thus follow a probability mass function.

For these cases one may proceed as follows. Assume that pair-wise distances occur in an ordinal manner and there are n categories where m_i is the number of pairs of distance r_i with $i = 1 \dots n$ and strict ordering $r_i < r_j$ if and only if $i < j$. Then the normalization condition is given by

$$\sum_{i=1}^n m_i C_\kappa r_i^{-\kappa} = \frac{N \langle k \rangle}{2}$$

and thus

$$C_\kappa = \frac{N \langle k \rangle}{2 \sum_{i=1}^n m_i r_i^{-\kappa}}.$$

Again, there will be excess probability which has to be redistributed to categories of higher distance. Define j as the last category meeting the inequality

$$\sum_{i=1}^j m_i [C_\kappa r_i^{-\kappa} - 1] > 0.$$

Then there is $\sum_{i=1}^j m_i [C_\kappa r_i^{-\kappa} - 1]$ edges left that have to be redistributed to the $j+1$ -st category. The final connection probability is

$$P_i = \begin{cases} 1 & i \leq j \\ C_\kappa r_{j+1}^{-\kappa} + \frac{1}{m_{j+1}} \sum_{i'=1}^j m_{i'} [C_\kappa r_{i'}^{-\kappa} - 1] & i = j+1 \\ C_\kappa r_i^{-\kappa} & i > j+1. \end{cases} \quad (4.27)$$

This particular result will be of importance to show that the network models introduced in this chapter are structurally equivalent.

4.3.4 Application to One-Dimensional Lattice PLSW Networks

In this subsection, a PLSW network model is introduced which is conceptually comparable to the Kleinberg model in Sec. 2.3.5. Parts of the text given below are close to the published article [44].

The model can be described as follows. Beginning similarly as with the modified Watts–Strogatz (WS) model introduced in Sec. 4.2.1, given two distinct nodes u and v with indices $0 \leq u \leq N - 1$ and $0 \leq v \leq N - 1$ positioned equidistantly on a ring, the shortest distance between them is

$$d(u, v) = \min(|v - u|, N - |v - u|).$$

In contrast to the modified WS model, the connection probability between any two nodes is then defined by a power-law as

$$P_{d(u,v)} = p_0 |d|^{\mu-1}. \quad (4.28)$$

Here, $-\infty < \mu \leq 1$ is the respective structural control parameter that determines the degree of non-locality in the connections (larger μ values imply higher long-range connection probability). Requiring again that the mean degree is fixed, the normalizing constant evaluates to

$$p_0 = \langle k \rangle \left(2 \sum_{d=1}^{\lfloor N/2 \rfloor} d^{\mu-1} + \text{mod}(N, 2)(\lfloor N/2 \rfloor + 1)^{\mu-1} \right)^{-1}, \quad (4.29)$$

where $\text{mod}(N, 2)$ denotes the remainder of the integer division $N/2$ and $\lfloor \cdot \rfloor$ is the floor function.

Note that this model is this definition of the PLSW model above will produce a problem for large negative values of μ , for which the connection probability between close neighbors will exceed unity. To avoid having probability values that are larger than one, this excess probability is redistributed to the nearest neighbors until one runs out of excess probability, as illustrated in Fig. 4.15. This results in producing a $\langle k \rangle$ -nearest neighbor lattice for $\mu \ll 0$, similar to all variants of the Watts–Strogatz small-world model. For a sampling algorithm, loop over all node pairs $0 \leq u \leq N - 2$ and $u + 1 \leq v \leq N - 1$, compute the distance $d(u, v)$ and connect this pair with probability Eq. 4.28. This algorithm is of order $\mathcal{O}(N^2)$.

Example networks for increasing structural control parameter are displayed in Fig. 4.17.

4.4 Structural Relation of Small-World Models

In this section, three of the models introduced in this chapter will be shown to be structurally similar by relating their respective structural control parameters and notion of distance. The three models are the self-similar modular hierarchical (SSMH) small-world model, the one-dimensional lattice PLSW network model and the modified Watts–Strogatz model. It will be shown that all of these models appear to produce modular hierarchical structure.

4.4.1 Mapping SSMH Networks to One-Dimensional PLSW Networks

Parts of the text given below are close to the published article [44]. In order to show that SSMH and one-dimensional lattice PLSW are

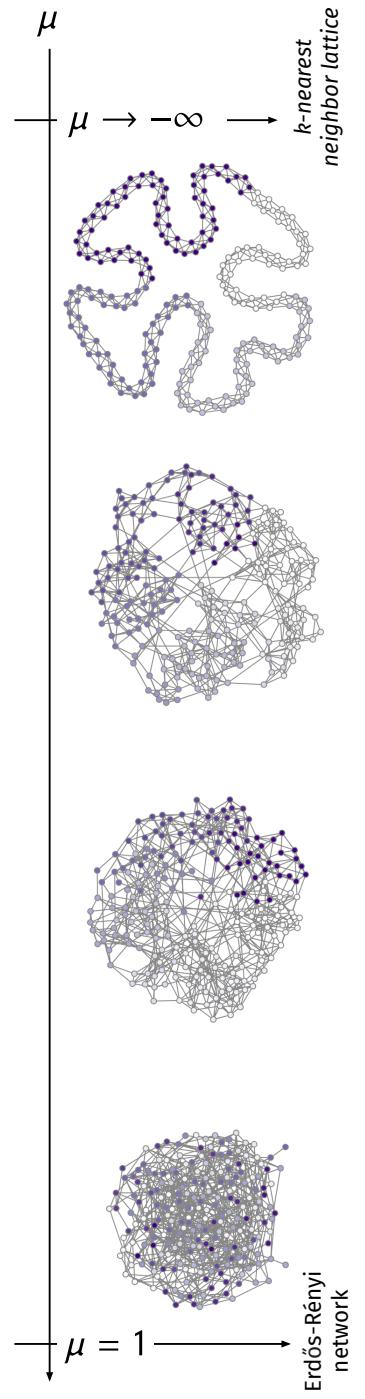
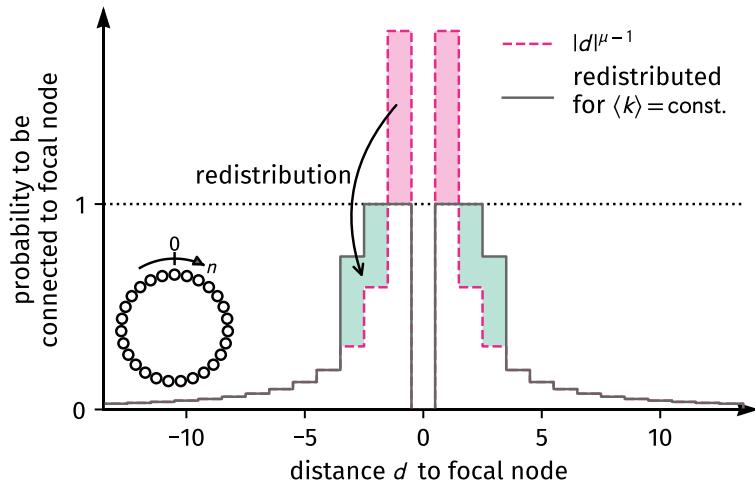


Figure 4.14: Network samples from the one-dimensional lattice PLSW model for increasing structural control parameter μ . Parameters $N = 6^3$ and $k = 6$ were used.

Figure 4.15: Method for redistributing the probability mass function of the one-dimensional lattice PLSW network model such that it does not exceed unity in numerical calculations. The original connection probability (red dashed line) is modified by redistributing the excess probability to its nearest neighbors.



structurally equivalent, the hierarchical distance between nodes in the first model will be mapped to the lattice distance on a ring (i.e. to the defining distance in the latter model). The hierarchical distance from a focal node at layer $\ell = 0$ may first be approximated as a continuous function. Then, the connection probability density in layer space as per Eq. (4.2) is given by

$$p(\ell) = \frac{\xi^\ell}{B^\ell} p_0. \quad (4.30)$$

Note that the prefactor $(\xi/B)^{-1}$ of the original equation has been absorbed to p_0 . The number of the focal node's potential neighbors in layer space is thus given as

$$dn = \frac{dn}{d\ell} d\ell = B^\ell \log B d\ell,$$

where the total number of reachable nodes $n(\ell) = B^\ell$ was used. The number of connections the focal node has to layer ℓ is therefore

$$\begin{aligned} dk &= p(\ell) dn = \frac{\xi^\ell}{B^\ell} B^\ell p_0 \log B d\ell \\ &= \xi^\ell p_0 \log B d\ell. \end{aligned}$$

The connection density (connections per layer) in layer space is defined as

$$dk = \frac{dk}{d\ell} d\ell,$$

so one finally obtains

$$\frac{dk}{d\ell} = \xi^\ell p_0 \log B \equiv p(\ell).$$

Now, with $p(y) = \frac{dy}{dx} p(x)$ and $\ell(n) = \log n / \log B$ one finds

$$\begin{aligned} p(n) &= \frac{1}{n \log B} \xi^{\ell(n)} p_0 \log B = p_0 n^{\frac{\log \xi}{\log B} - 1} \\ &= p_0 n^{\mu-1} \end{aligned}$$

which is the distance-based power-law connection probability in a one-dimensional lattice and thus relates the two respective structural control parameters as

$$\mu = \frac{\log \xi}{\log B}. \quad (4.31)$$

This result can be interpreted as follows. The SSMH connection probability function corresponds approximately to a discretized version of the power-law that is associated to the PLSW connection probability, however, target nodes are grouped in exponentially growing batches (cf. Figure 4.16). In this approximation, the additional embedded nature of the SSMH network structure is lost. It thus becomes clear that SSMH networks are a special case of the general categorized PLSW model of Sec. 4.3.3.

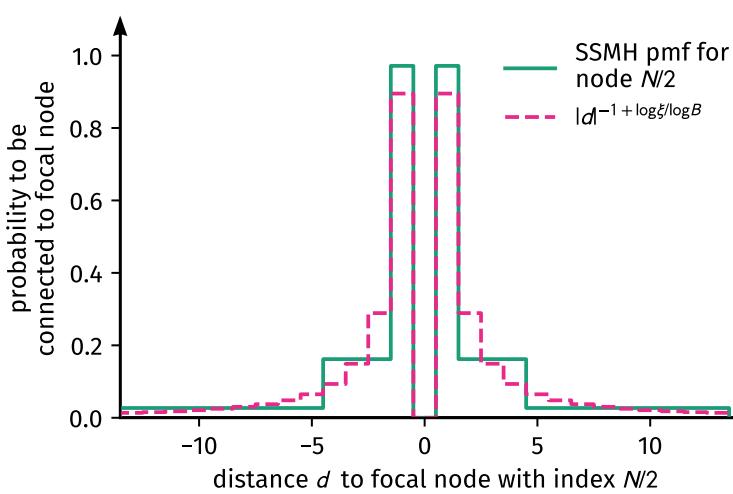


Figure 4.16: Comparison of the connection probability functions associated with SSMH networks and with one-dimensional lattice PLSW networks. Parameter values chosen here are number of submodule per module $B = 3$, hierarchical layers $L = 3$, mean degree $\langle k \rangle = 3.4$ and structural control parameter $\xi = 0.5$ which relates to $\mu \approx -0.63$. The SSMH connection probability function can be interpreted to group nodes of similar distance to the focal node to exponentially growing batches of equal connection probability.

One important structural result of this mapping is that for the SSMH structural control parameter of value $\xi = 1$, where all hierarchical clustering is lost as per Eq. (4.5), the corresponding PLSW control parameter is $\mu = 0$, resulting in the connection probability function $P(n) \propto n^{-1}$ which is the critical connection probability function at which locally informed search processes become optimal (see Secs. 2.3.5 and 2.3.6). A comparison of example networks for equal structural control parameters are shown in Fig. 4.17.

4.4.2 Mapping modified Watts–Strogatz Networks to One-Dimensional Lattice PLSW Networks

In order to relate the structural control parameters of the modified Watts–Strogatz model β and the one-dimensional lattice PLSW model μ , one recognizes first that there are three important structural conditions for the control parameters. As per definition of these small-world models, the two important limits are reached when (a) $\mu \rightarrow -\infty$ and $\beta \rightarrow 0$, which is when both models become structurally equivalent to k -nearest neighbor lattice, and (b) for $\mu = 1$ and $\beta = 1$ which are the values where all network models become equivalent

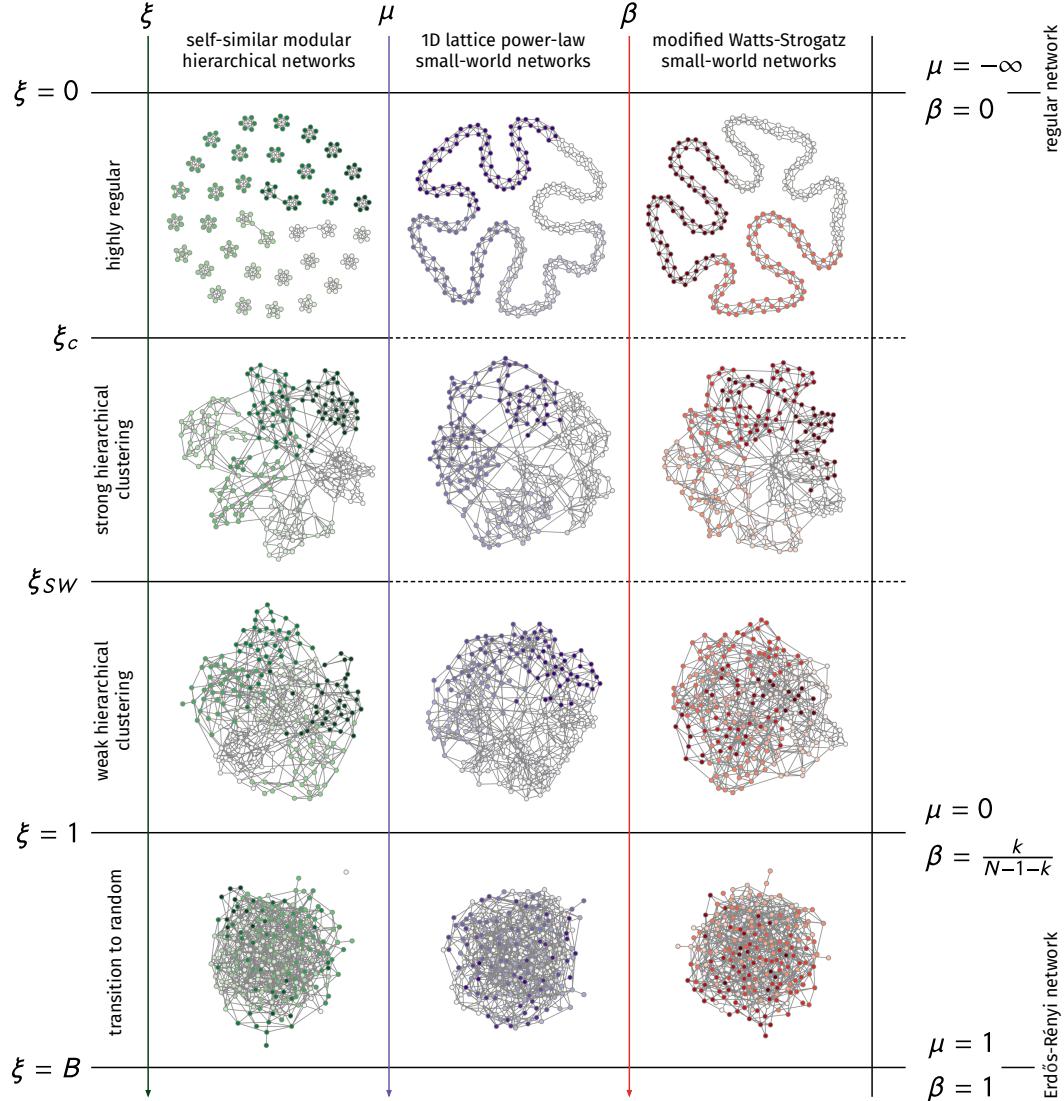


Figure 4.17: Topological phases of three small-world models as a function of their respective structural control parameters ξ , μ , and β as mapped to each other by Eqs. (4.31) and (4.33). Chosen parameters are mean degree $\langle k \rangle = 6$, base number of submodules $B = 6$, and number of hierarchical layers $L = 3$ (number of nodes $N = 6^3$, respectively). For small values of the structural control parameters the sampled networks closely resemble their base structures. With increasing long-range connection probability, all of the network samples seem to display sort of a modular hierarchical structure. While in the SSMH model (green, left) this modular hierarchical structure is given by design, the one-dimensional lattice PLSW (middle, violet) and modified Watts–Strogatz models (right, orange) would display modular hierarchical clustering only as a result from noise (see Sec. 4.4.3). The network figures were generated using a custom interactive visualization tool made by the author which is available online at <https://bit.ly/2TrhUfs>.

to the Erdős–Rényi model. One formulaic way to ensure that these values are reached is

$$\log \beta = A\mu + B. \quad (4.32)$$

This equation can be motivated as follows. Since both models are structurally equivalent for small values of β (large negative values of μ), it seems to be appropriate to demand that the short-range degree k_S of both models corresponds when relating their control parameters. For the modified Watts–Strogatz model one finds $k_S = kp_S = k/[1 + \beta(N - 1 - k)/k]$. For the PLSW model, one may approximate the discrete probability connection function as a continuous function

to obtain the short-range degree

$$\begin{aligned} k_S &\approx k \int_1^{k/2} x^{\mu-1} dx \left/ \int_1^{(N-1)/2} x^{\mu-1} dx \right. \\ &= k \frac{(k/2)^\mu - 1}{((N-1)/2)^\mu - 1} \end{aligned}$$

and thus

$$\beta \approx \frac{k}{N-1-k} \left[\frac{((N-1)/2)^\mu - 1}{(k/2)^\mu - 1} - 1 \right],$$

giving justification to assume an approximate form of Eq. (4.32) for relating the control parameters. Now, since this introduced relation formula Eq. (4.32) already meets two important structural limits, a third one is considered to find the values of the parameters A and B in the following.

As discussed in Sec. 4.5, the SSMH structural control parameter value at which all hierarchical structure is lost is given as $\xi = 1$ (or $\mu = 0$, respectively). This is the structural point at which a random node has, on average, as many connections to nodes in short range as it has to nodes in each of the categories of greater hierarchical distance. For the modified Watts–Strogatz model, there are only two categories: short-range and long-range. Hence, it seems plausible to connect the control parameters in such a way that at $\xi = 1$ (or $\mu = 0$, respectively), the short-range degree is equal to the long-range degree, which as per Eq. (4.12) yields

$$\begin{aligned} kp_S &= (N-1-k)\beta p_S \\ \beta &= k/(N-1-k). \end{aligned}$$

In order for the mapping to reach all structural limits as discussed as well as meet Eq. (4.32), a possible choice is hence

$$\beta = \frac{k}{N-1-k} \left(\frac{N-1-k}{k} \right)^\mu. \quad (4.33)$$

Since usually $N-1-k > k$, a simple check finds the three demanded values $\beta(\mu \rightarrow -\infty) = 0$, $\beta(\mu = 0) = k/(N-1-k)$, and $\beta(\mu = 1) = 1$. A comparison of example networks for equal structural control parameters are shown in Fig. 4.17 and a comparison of network properties are shown in Fig. 4.18. Comparing the sample networks with the naked eye hints at the existence the correspondence of these models, which seems plausible comparing the network properties, as well.

4.4.3 All Small-World Network Models Appear to be of Modular Hierarchical Structure

In the former discussion within this chapter, the claim that the models defined above are structurally equivalent was made a handful of times, supported by the analysis of several network or node properties which indeed behave similarly when varying their respective

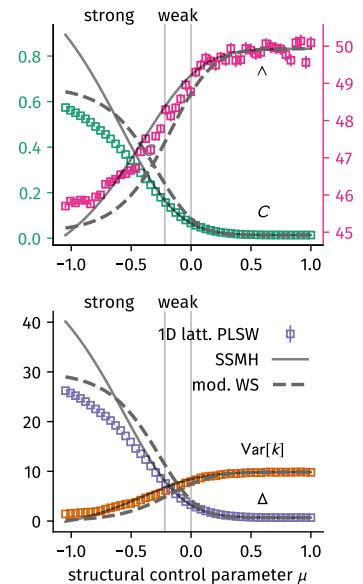
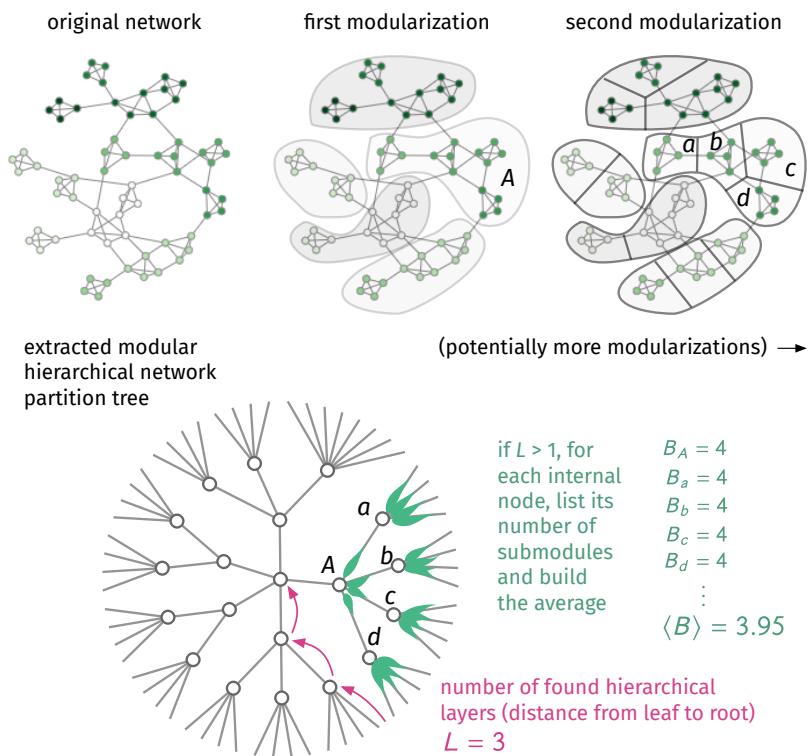


Figure 4.18: Node properties of the SSMH, the modified WS and the one-dimensional lattice PLSW network model for increasing control parameter μ , here with parameters $B = 8$, $L = 3$ (number of nodes $N = 9^3$), and $\langle k \rangle = 10$. The observables displayed as markers were obtained as an average over 100 independently sampled networks for each value of μ . The solid lines represent the theoretical results of the SSMH model and the dashed lines represent the theoretic results of the modified Watts–Strogatz model. Note that none of the observables change significantly after leaving the weak modular hierarchically clustered phase at $\mu = 0$.

Figure 4.19: Illustration of how the network samples generated in this chapter are analyzed for their modular hierarchical structure. Given a network, the hierarchically nested version of *Infomap* is run to find a modular hierarchical network partition (as illustrated in the top row). This network partition can be displayed as a tree of height L where leaves correspond to the nodes of the original network and the tree's internal nodes correspond to the found partition. The number L of found hierarchical layers is one of the measured network observables. Furthermore, each internal node I (or submodule) carries B_I submodules. The measured observable the mean and standard deviation of the number of submodules of all internal nodes (non-leaf nodes).



control parameters. However, one might wonder how similar the networks are in the literal sense of their modular hierarchical structure. There exist several algorithms to partition observed networks in a modular hierarchical way. In the following, one particular algorithm called *Infomap* is used to find optimal modular hierarchical partition of networks generated from the models defined above. *Infomap* finds modular hierarchical network partitions by hierarchically grouping regions of nodes in which a random walker spends a considerable amount of time before entering other regions. The modular hierarchical partition is then given as a tree of height L where leaves represent nodes and internal nodes represent groupings of nodes at different coarse grained levels (see Fig. 4.19). *Infomap* minimizes the minimum description length which can be described as a measure of how compressible the network is given the found modular hierarchical network partition. A more detailed description of the algorithm's functionality is given in App. A.

For the analysis of the modular hierarchical network partitioning of the model networks, network samples were generated for parameters $B = 8$, $L = 3$ and $\langle k \rangle = 8$ for the SSMH, the one-dimensional lattice PLSW, the modified Watts–Strogatz and the 2-torus PLSW model, and their respective structure control parameters were varied. For every parameter combination, 100 networks were sampled from their respective models. Subsequently, *Infomap* was used with default parameters to find an optimal modular hierarchical network partition. The results are shown in Fig. 4.20.

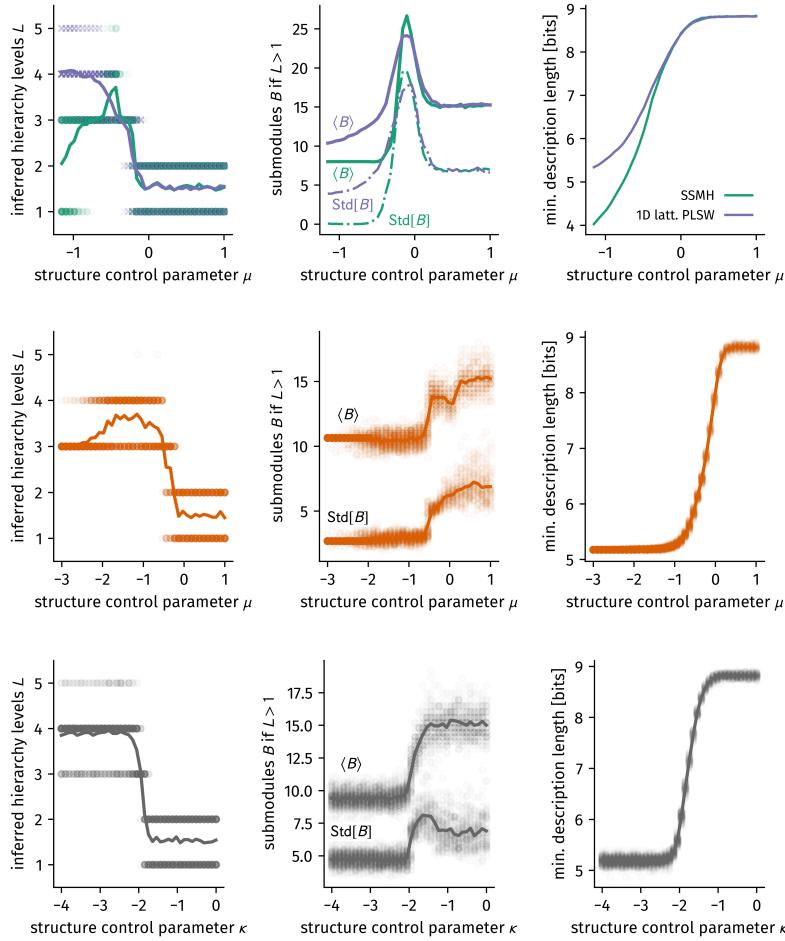


Figure 4.20: Results of the modular hierarchical partition detection using *Infomap* for $B = 8$, $L = 3$ ($N = 8^3$), and $\langle k \rangle = 8$. For each parameter combination, 100 networks were sampled and subsequently analyzed using *Infomap*. Markers represent results for single networks and solid lines are averages over all samples. Dashed lines in the top row are averages of the standard deviation of number of submodules per network sample. **(Top row)** The SSMH (green) and 1D lattice PLSW model (violet). **(Middle row)** The modified Watts–Strogatz model. **(Bottom row)** The 2-torus PLSW model. **(Left column)** The number of measured hierarchy levels L . **(Middle column)** Mean and standard deviation of number of submodules B_I per internal node I . **(Right column)** Minimum description length of the found network partition.

As one can see, all of the networks are found to be of modular hierarchical structure within the corresponding phases of the structural control parameters $\mu < 0$ and $\kappa < -2$, respectively, finding a number of hierarchical layers $L \geq 3$ to best describe their network structures. Similarly, the mean number of submodules is close to $\langle B \rangle = 10$ ($\langle B \rangle = 8$ for the SSMH model, as expected from the construction condition $B = 8$). Remarkably, for the SSMH model, *Infomap* recovers an underlying hierarchical tree which is almost exactly structurally equal to the original underlying hierarchical tree with the standard deviation of number of submodules per internal node being close to zero. For the other models, this standard deviation is greater but approximately half of the mean. A nonzero standard deviation is to be expected since there is no underlying hierarchical tree for the other models and all correspondences to a hidden hierarchical structure are due to noise: By chance, a certain amount of potential edges in the same area will not be sampled while others are, producing randomly appearing blocks of ones in the adjacency matrix which are then found by the extraction algorithm. All network structures seem to no longer be able to be resolved when their structural control parameters come close to some threshold which is determined by the exponent of the power-law connection probability being equal

to their respective embedding dimension (hence, $\kappa = -2$ for the 2-torus PLSW model and $\mu = 0$ as $P_d \propto d^{\mu-1}$ for the rest). This is the point where an average node has as many connections in short-range as it has to any distance in the long-range category (cf. Eq. 4.5). One might expect this effect, as an algorithm cannot resolve a formerly defined group membership if a node does not have more connections to nodes of this group than to nodes in other groups [64]. The modified Watts–Strogatz model seems to lose its resolvable structure a bit sooner than the SSMH and the one-dimensional lattice PLSW models. Hence, one could interpret this model as being somewhat less modular hierarchical. All of the models which are not modular hierarchical by design seem to converge to a number of ≈ 5 bits minimum description length, compared to a compressibility of ≈ 9 bits for the Erdős–Rényi case⁴, whereas the SSMH model approaches ≈ 4 bits compressibility. The SSMH model’s compressibility is smaller than the other models’ compressibility because the SSMH model is truly generated from a hierarchical tree, which can be found, whereas the others are not and modular hierarchical structure is produced by chance, as argued above.

To sum up, all the models seem to generate networks which are interpretable as modular hierarchical, at least for a random walker spending time in different regions of the network as modeled by *Infomap*.

4.5 Summary and Discussion

In this chapter, a modular hierarchical network model was introduced, based on a self-similar hierarchical partition tree. This network model generates modular hierarchically structured networks with a constant mean degree. A single structural control parameter controls the number of long-range edges per node and can be used to interpolate between a structure where all edges are short-ranged (only in modules in the lowest hierarchy module) and the Erdős–Rényi random network model. Clustering coefficient and node degree variance were computed analytically as functions of the structural control parameter. It was shown that different values of the control parameter correspond to different modular hierarchical phases where (a) an average node has more connections to nodes within its lowest-level hierarchy cluster than to nodes of its higher-level clusters combined called the *strong modular hierarchical phase*, (b) an average node has more connections to nodes within its lowest-level hierarchy cluster than to nodes of each of its higher-level clusters, called the *weak modular hierarchical phase* and (c) the network is already close to a random network. As this network model seems similar to two other existing network models, namely the Watts–Strogatz and the Kleinberg model, modified versions of these models were introduced. Their modified definitions allow for a thorough analysis as they both keep their mean degree fixed. In the Kleinberg small-world model, nodes are connected based on a connection probability

⁴ For the ER network, *Infomap* finds a one-level solution where every node is its own module. This means that the per-step minimum description length becomes equal to the Shannon entropy Eq. (A.1) with ergodic node visiting probabilities $p_i = k_i/N\langle k \rangle$. Hence, the compressibility can be computed as $H_{\text{ER}} = -(N\langle k \rangle)^{-1} \sum_{i=1}^N k_i \log_2(k_i/N\langle k \rangle)$. Using the binomial degree distribution of the ER model and noting that $k_i > 0$, since only the largest component has been used, the minimum description length can be estimated as $H_{\text{ER}} \approx -[\langle k \rangle^{-1} \langle k \log_2 k \rangle_{k>0} - \log_2(N\langle k \rangle)]$ which can be numerically computed to find $H_{\text{ER}} \approx 8.9$ for the parameters chosen here.

decaying as a power-law with their distance. This power-law procedure was generalized to find that it can be applied to arbitrary embedding geometries as well as arbitrary edge weights which do not have to be distances. An additional model was devised where nodes are distributed on a 2-torus following a uniform distribution are connected according to a power-law. Contrary to all the other models, this model has not only a constant mean degree but a constant node degree variance, as well while other properties remain.

After the definition of the models, it was shown how they correspond to each other by finding equations mapping their respective structural control parameters. It was subsequently found that their network properties are similar when varying their structural control parameters and that a modular hierarchical network partition algorithm finds similar network partitions for networks generate from all the models.

It therefore has to be assumed that real-world networks which were found to be of modular hierarchical structure do not necessarily have to have been generated using a modular hierarchical generation model. Instead, it seems likely that a considerable amount of these networks can be described as embedded in an abstract low-dimensional space and nodes being connected with a distance-based power-law connection probability. In order to test how “modular hierarchical” a network really is one might proceed as follows. Find the minimum description length of the network. Generate similar networks from the SSMH model where B and L are chosen close to the values of the extracted modular hierarchical tree and scan the modular hierarchical structure parameter ξ . Further generate similar networks from one of the PLSW models and vary their respective modular hierarchical structure parameters. Depending on where the compressibility of the original network is closer to might give an indication on which of the underlying network models are more likely to be responsible to have generated the network.

Another way of testing the power-law small-world hypothesis of real-world network would be to find a modular hierarchical network partition of the network and subsequently trying to find the categorical power-law connection probability Eq. (4.27) based on the nodes’ hierarchical distance.

Such applications to investigate real-world networks for their modular hierarchical or small-world structure, are, however, left for future research. Instead, now that several models are defined which can be interpreted to be of modular hierarchical structure, they will be used to test the hypothesis of them being of optimal topology for fast random walk processes in the following chapters.

5

Passage Time Statistics for Locally Clustered Networks

Modular hierarchical networks are ubiquitous in the real world. In this chapter the hypothesis is tested that such topologies provide an “optimal structure”¹ for certain dynamic processes as a potential explanation of their omnipresence. As argued before, one of the simplest dynamic processes is diffusion as modeled by random walks². Indeed, it has been found that locally informed searches are minimal on a modular hierarchical network model³. Since random walks are a reasonable model for random searches, their temporal observables place a natural upper bound on the corresponding observables of informed searches. Hence, this chapter will evaluate the potential of modular hierarchical networks to provide an optimal structure for random searches, as well, using the network models introduced in Ch. 4.

It will be found that an average medium approximation⁴ can be used to analytically estimate the shape of the pair-averaged first passage time (FPT)⁵ as a function of a model’s respective structural control parameters. This function does not reveal the existence of an optimal structure for fast random searches. A subsequent analysis on actual network realizations will reveal that three of the four network models discussed in Ch. 4 display a minimum in the pair-averaged FPT. This effect is consequently explained by deriving a heuristic lower bound of the pair-averaged FPT for general locally clustered networks, which shows to be composed of a node degree variance distribution as well as a clustering coefficient contribution. As one of them decreases as the other increases with increasing structural control parameters for said three models, they combine to build a minimum. A fourth modular hierarchical network model with constant node degree variance is consequently shown to not possess an optimal structure for random searches. It is therefore argued that modular hierarchical structures do not inherently possess a minimal pair-averaged FPT but that the emergence of this minimum is a pure consequence of network structures with decreasing clustering and increasing node degree variance. Large parts of this chapter are close to the published articles [44, 42].

¹ Naturally, what “optimal” is depends on the application, but in the following it will refer to minimal temporal observables of random walks.

² see Sec. 3.1

³ see Sec. 2.3.6 and [80, 31]

⁴ As introduced in Sec. 5.1.

⁵ see Sec. 3.1.4

5.1 Average Medium Approximation

In the context of random walks on networks, an *average medium approximation* (AMA) can be used to solve diffusion problems by approximating the network topology by an “average structure”.

Here, this AMA is given by connecting all pairs of nodes (u, v) with a link with weight P_{vu} , equal to the probability of connecting both nodes in the original random network. In order to do this, P_{vu} has to be normalized so that, from each node, the total probability to jump to any other node in one time step is equal to one, that is, $\sum_v P_{vu} = 1$. For example, given an Erdős-Rényi network with N nodes, this AMA would connect each node to every other node, except for itself, through a link with weight $P_{vu} = p = 1/(N - 1)$ and hence the mean FPT (MFPT) between any pair of nodes would be equal to the MFPT in a complete graph $\tau_{vu} = N - 1$, a result consistent with the large mean-degree limit of more sophisticated AMAs [72]. In this subsection, the FPT statistics and pair-averaged FPT of the SSMH model introduced in Sec. 4.1 will be evaluated. Furthermore, the mixing time of the modified Watt–Strogatz model introduced in Sec 4.2 will be computed analytically and used to illustrate the small-world effect without the use of the average shortest path length. Both observables are shown to monotonically decrease with increasing structural control parameters ξ and β .

5.1.1 SSMH First Passage Time Distribution and Pair-Averaged First Passage Time

In what follows it will be demonstrated how to compute FPT statistics for structurally averaged SSMH networks, where the possible edges are replaced by the probability for those edges to exist.

Time Evolution of Walker Distribution

Following Equation (4.3), the probability of two nodes to be connected will be denoted as \tilde{P}_ℓ . In the context of this AMA, this means that every node is connected to every other node but edges are weighted with probability $P_\ell = \tilde{P}_\ell / \langle k \rangle$. On this network, a random walk with a sink at an arbitrary node v is investigated. When a walker is positioned at node u , the probability to jump to node v at hierarchical distance $d(u, v) \equiv d$ is

$$P_d = \frac{1}{B-1} \left(\frac{1-\xi}{1-\xi^L} \right) \left(\frac{\xi}{B} \right)^{d-1},$$

as reasoned above. In this context, the following scenarios happen.

1. The walker jumps to node v . This occurs with probability

$$P_1 = P_d.$$

The new distance to node v will be $d' = 0$.

2. The walker jumps to a node in a lower layer $\ell < d$. The new distance will still be $d' = d$. In each layer $\ell < d$, there will be $B^{\ell-1}(B-1)$ possible target nodes for the walker starting at node u . The probability for this scenario to happen is

$$P_2 = \sum_{\ell=1}^{d-1} P_\ell B^{\ell-1}(B-1) = \frac{1 - \xi^{d-1}}{1 - \xi^L}.$$

3. The walker jumps to a node in $\ell = d$, but misses the branch that v is in and the branch that u is in (because this event is covered in scenario 2). The new distance is still $d' = d$. There are $B^{d-1}(B-2)$ nodes for this event. Consequently, it happens with probability

$$P_3 = P_d B^{d-1}(B-2).$$

4. The walker jumps to a node in $\ell = d$, hits the branch that v is in, but misses v . The new distance is $d' < d$, for each d' with probability

$$P_4 = P_d B^{d'-1}(B-1).$$

5. The walker jumps to a node in $\ell > d$. Consequently, the new distance is $d' > d$. The probability of this happening is, for each d'

$$P_5 = P_{d'} B^{d'-1}(B-1).$$

The transition matrix from a walker being at distance d from the target to being at new distance $0 \leq d' \leq L$ from the target is hence

$$W_{d'd} = \begin{cases} \frac{1}{(B-1)} \frac{1-\xi}{1-\xi^L} \left(\frac{\xi}{B}\right)^{d-1}, & d' = 0 \wedge d > 0 \\ \frac{B^{d'-1}}{B^{d-1}} \xi^{d-1} \frac{1-\xi}{1-\xi^L}, & d' < d \wedge d > 0 \\ \frac{1-\xi^{d-1}}{1-\xi^L} + \frac{1-\xi}{1-\xi^L} \xi^{d-1} \frac{B-2}{B-1}, & d' = d \wedge d > 0 \\ \frac{1-\xi}{1-\xi^L} \xi^{d'-1}, & d' > d \wedge d > 0 \\ 0, & d = 0 \wedge d = 0 \\ 1, & d' = 0 \wedge d = 0. \end{cases} \quad (5.1)$$

Define the vector $\mathbf{p}^{(d_{\text{start}})}(t)$ as the probability to find a single walker at distance d when a random walker started at distance $d_{\text{start}} \equiv d_s$ from the target. This vector has $L+1$ entries, ranging between 0 and L . Let a discrete-time random walk begin with initial conditions given by the vector

$$p_d^{(d_{\text{start}})}(0) = \delta_{dd_{\text{start}}}, \quad (5.2)$$

i.e. there is a single walker at distance d_{start} . The probability to find the random walker at distance d when starting at distance d_{start} after t time steps is then

$$p_d^{(d_{\text{start}})}(t) = (W^t)_{d,d_{\text{start}}} \quad (5.3)$$

as the process is Markovian. This time evolution is displayed in Fig. (5.1) for $B = 8$ and $L = 3$. For a strong hierarchically clustered

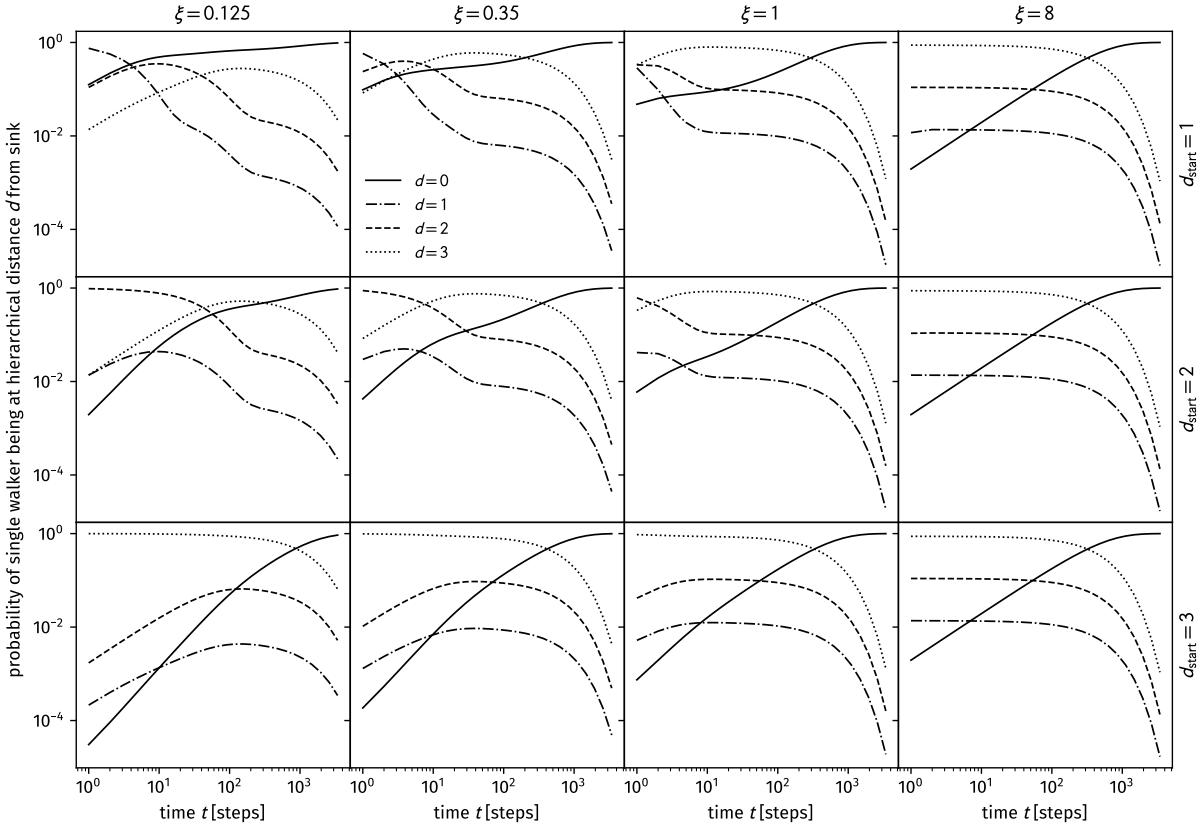


Figure 5.1: Time evolution of the random walker density at hierarchical distance d from the sink when the walk started on a single node at hierarchical distance d_{start} for the average medium approximation as per Eq. (5.3). Parameters used here were $B = 8$ and $L = 3$. Note that the initial condition Eq. (5.2) is not displayed as the plots begin at $t = 1$.

⁶ when corrected for the number of target nodes at distance d

network ($\xi = 1/8$) one clearly sees the influence of the hierarchy layers when starting near the sink at $d_{\text{start}} = 1$. The random walker tends to be either absorbed, stay near ($d = 1$) or jump to a node at hierarchical distance $d = 2$. With progressing time there are phases where the walker tends to stay at certain distances and then leave those again for shorter distances. With growing structure parameter ξ , this behavior is smoothed until being at any distance is equally probable⁶ for the Erdős–Rényi limit at $\xi = 8$. When starting “far away” from the sink at distance $d_{\text{start}} = 3$ for the hierarchically clustered case ($\xi = 1/8$), the walker tends to stay far away, just slowly arriving at modules “nearer” to the sink until being absorbed.

Global Mean First Passage Time

In this section, analytic expressions are derived for the global mean first passage time of a target node in SSMH average medium networks.

Suppose, that a random walk begins with a single walker on every node $u \neq v$. Then the total ratio of walkers that are absorbed up to a

time t is

$$\begin{aligned} p_0(t) &= \frac{1}{B^L - 1} \sum_{d_s=1}^L B^{d_s-1} (B-1) p_0^{(d_s)}(t) \\ &= \frac{1}{B^L - 1} \sum_{d_s=1}^L B^{d_s-1} (B-1) (\tilde{\mathbf{W}}^t)_{0,d_s}. \end{aligned} \quad (5.4)$$

Another way to write down the ratio of walkers absorbed into the sink node at time t is

$$p_0^{(d_s)}(t) = 1 - \sum_{d=1}^L p_d^{(d_s)}(t),$$

since the probability of the walker being at any distance $0 \leq d \leq L$ is equal to one and conserved at all times. At this point of the derivation a few new quantities need to be introduced. As described above, the probability of being at distance d' at time t is

$$p_{d'}(t) = \sum_{d=0}^L P_{d'd} p_d(t-1).$$

However, for $d' > 0$, the column $d = 0$ does not contribute anything to the sum (since it is filled with zeros). Hence, a new vector $\tilde{\mathbf{p}}^{(d_{\text{start}})}$ is defined which is the former probability vector with the 0-th element removed and a new transition matrix $\tilde{\mathbf{W}}$, as the full transition matrix Eq. (5.1) with the 0-th column and row removed. One therefore has

$$\tilde{\mathbf{p}}_d^{(d_{\text{start}})}(t) = (\tilde{\mathbf{W}}^t)_{d,d_{\text{start}}},$$

such that

$$p_0^{(d_s)}(t) = 1 - \sum_{d=1}^L (\tilde{\mathbf{W}}^t)_{d,d_s}.$$

In order to consider the contribution of all starting nodes, these results are combined with Eq. (5.4) to find

$$p_0(t) = 1 - \frac{B-1}{B^L-1} \sum_{d_s=1}^L B^{d_s-1} \sum_{d=1}^L (\tilde{\mathbf{W}}^t)_{d,d_s}.$$

This is the ratio of walkers that have been absorbed in to the sink v up until time t . Further introducing the vector $\tilde{\mathbf{b}}$ containing the fractions of all possible targets at layer d as

$$\tilde{b}_d = \frac{B-1}{B^L-1} B^{d-1}, \quad 1 \leq d \leq L,$$

as well as the vector $\mathbf{1}^T = (1, 1, \dots, 1)$, the absorbed walker density then reduces to

$$p_0(t) = 1 - \mathbf{1}^T \tilde{\mathbf{W}}^t \tilde{\mathbf{b}}. \quad (5.5)$$

This arrival time cdf is exemplarily shown in Fig. 5.2 and compared to the cdfs of the corresponding 1D lattice PLSW model, which was computed using the full transition matrix composed from the 1D lattice PLSW pmf Eq. (4.28) with entries $W_{vu} = P_{d(u,v)} / \langle k \rangle$. The average media of both models produce qualitatively similar arrival time cdfs

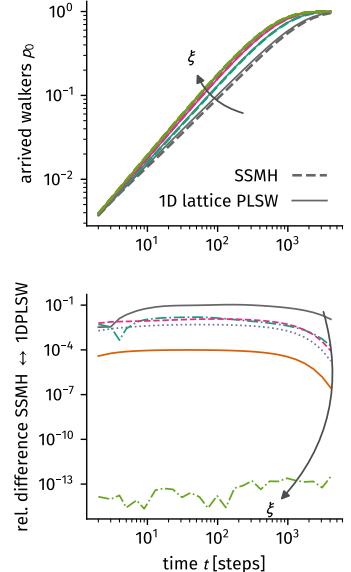


Figure 5.2: (Top) The SSMH average medium arrival time cdf $p_0(t)$ when starting anywhere in the network as per Eq. (5.5). Parameter values $B = 8$, $L = 3$, $\langle k \rangle = 7$ and $\xi \in \{0.125, 0.25, 0.5, 1, 4, 8\}$ were used. The results for the 1D lattice PLSW network model correspond to the SSMH results with increasing ξ as can be seen in the (bottom) panel.

which, with increasing structural control parameter ξ (i.e. increasing long-range connection probability), become even quantitatively similar.

After discussing the arrival time distribution, the focus will be shifted to finding the global mean FPT for the focal node in the following. This quantity can be calculated as

$$\begin{aligned} \tau &= \sum_{t=0}^{\infty} t [p_0(t) - p_0(t-1)] = \sum_{t=0}^{\infty} t [\mathbf{1}^T \tilde{W}^{t-1} \tilde{b} - \mathbf{1}^T \tilde{W}^t \tilde{b}] \\ &= \mathbf{1}^T [\mathbb{1} - \tilde{W}]^{-1} \tilde{b}. \end{aligned} \quad (5.6)$$

Note that this quantity is equal to, both, the global mean FPT and the pair-averaged FPT (that is, the mean global mean FPT), because in this AMA all nodes are equal. This result is similar to the result for arbitrary networks, where instead of $(\mathbb{1} - \tilde{W})^{-1}$ one makes use of the inverse of the reduced unnormalized graph Laplacian, as shown in Sec. 3.1.4. However, using the layer approach, the matrix size can be reduced from $B^L - 1$ to L , a great reduction in degrees of freedom. This makes it possible to obtain analytical expressions for the global mean FPT. As examples, the following results for $L < 4$ were found solving Eq. (5.6) with a computer algebra system:

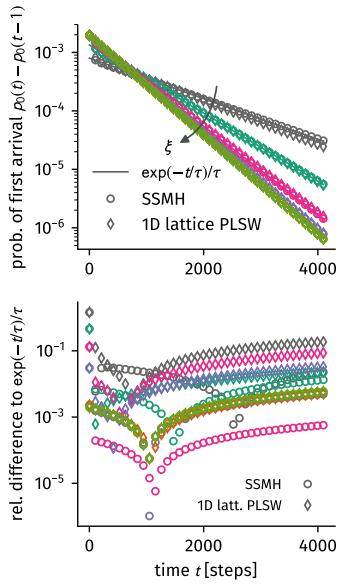


Figure 5.3: **(Top)** The SSMH average medium arrival time pmf $p_0(t) - p_0(t-1)$ as per Eq. (5.5) as well as the one-dimensional lattice PLSW average medium arrival time pmf computed from its full transition matrix. Parameter values $B = 8$, $L = 3$, $\langle k \rangle = 7$ and $\xi \in \{0.125, 0.25, 0.5, 1, 4, 8\}$ were used. The pmfs approximately follow an exponential distribution with mean Eq. (5.7). **(Bottom)** relative difference of the pmf of both models to the approximative relationship.

$$\begin{aligned} \tau_{L=1} &= B - 1, \\ \tau_{L=2} &= \frac{(B-1)B(\xi+1)(B^2\xi + B\xi + B - \xi)}{(B+1)\xi(B\xi + B - \xi)}, \\ \tau_{L=3} &= \left[(B-1)B^2(\xi^2 + \xi + 1) \times \right. \\ &\quad \times \left(B^4\xi^2(\xi+1) + B^3\xi(\xi+1) + \right. \\ &\quad \left. \left. + B^2(\xi+1)^2 - B\xi(2\xi^2 + 3\xi + 2) + \xi^2(\xi+1) \right) \right] \\ &\quad / \left[(B^2 + B + 1) \xi^2(B\xi + B - \xi) \times \right. \\ &\quad \left. \times \left(B(\xi^2 + \xi + 1) - \xi(\xi+1) \right) \right]. \end{aligned} \quad (5.7)$$

Note that for $\xi \rightarrow B$ the global mean FPT approaches the average medium Erdős-Rényi solution $\tau = B^L - 1$, which corresponds to the complete network. In addition, the global mean FPT diverges for $\xi \rightarrow 0$, since the average medium then approaches a state where it consists of B^{L-1} complete networks, each containing B nodes.

The arrival time pmf is exemplarily shown in Fig. 5.3 and compared to the pmfs of the corresponding one-dimensional (1D) lattice PLSW model, which was computed using the full transition matrix composed from the 1D lattice PLSW pmf Eq. ((4.28)) with entries $W_{vu} = P_{d(u,v)}/\langle k \rangle$, as well as the approximative relationship $p_0(t) - p_0(t-1) = \tau^{-1} \exp(-t/\tau)$ where τ is computed from Eq. (5.7). The average media of both models produce qualitatively similar arrival time cdfs which, with increasing structural control parameter ξ (i.e. increasing long-range connection probability), become

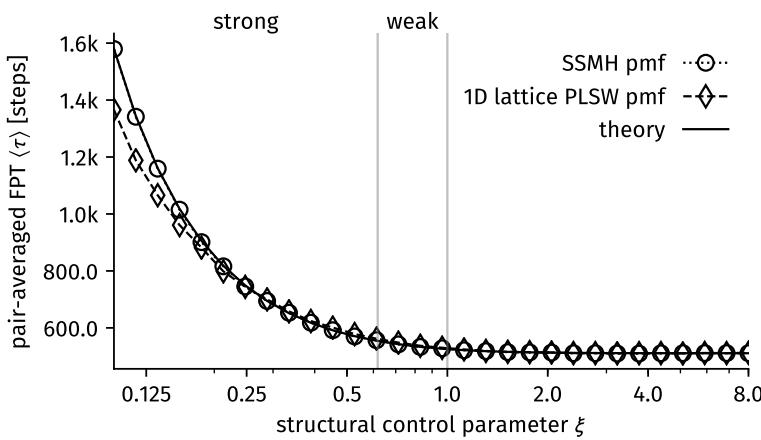


Figure 5.4: Pair-averaged FPT for the SSMH model and the 1D lattice model computed from simulations (markers) and compared to Eq. (5.7). Parameters used here where $B = 8$, $L = 3$ and $\langle k \rangle = 7$.

even quantitatively similar. The computed pmf for the SSMH model corresponds to the approximative relationship, alas showing systematic deviations.

A comparison of simulated pair-averaged FPT and Eq. (5.7) can be seen in Fig. 5.4. Here, the pair-averaged FPT was computed for average media of SSMH and the 1D lattice PLSW model using parameters $B = 8$, $L = 3$, $\langle k \rangle = 7$ and structural control parameter $0.1 \leq \xi \leq 8$. Furthermore, random walk simulations were performed using averaged networks where at initial time $t = 0$ a single walker was placed on every node. The simulation was then run as described in Sec. 3.1.1 until each walker visited each node at least once. At every time step t a walker visited a node it had not visited before, this time t was saved as a FPT. The pair-averaged FPT was computed as the mean of all saved FPTs. The simulations are in agreement with the theory. For small values of $\xi \lesssim 0.2$ the results for SSMH and PLSW begin to differ more strongly. Note that the pair-averaged first passage time is a monotonically decreasing function of the structural control parameter ξ which suggests that no “optimal” structure exists in between the two limiting structures, in the sense that no global minimum is shown between the limits.

5.1.2 Mixing Time and Small-World Effect in the Modified Watts–Strogatz Model

In this section, an average medium approximation will be used to provide an intuitive and analytic illustration of the original small-world effect as introduced in [81]. Parts of this Section are close to the published article [42].

In the original model, the small-world effect was established by comparing the clustering coefficient to the average shortest path length of the networks. While random networks have short path lengths, they possess low clustering, on the other hand regular networks are highly clustered, while nodes are, on average, quite distant from one another. With rewiring only a short amount of edges it was

shown that shorter paths appear immediately while high clustering preservers, explaining the small-world effect. However, in the following, a different approach will be taken.

One of the purposes of the original model was to explain the Milgram small-world experiment [76] where participants had to mail packets⁷ to a stranger by mailing them to a person they did know on a first-name basis and instruct them to pass the packet further. Even though this experiment had methodological flaws [79, 32], it found that some of the packets arrived relatively fast through a short chain of people. In the following the small-world effect will be illustrated by showing that the mixing time, as an upper bound for the delivery time of those messages, decreases much faster than the clustering coefficient with increasing probability of long-range edges.

As argued above, the average medium random walk transition matrix is given as $W_{vu}^{\text{avg}} = \langle A_{vu} \rangle / k = P_{d(v,u)}/k$ where the modified small-world connection probability Eq. (4.12) is used. The time scale with which the equilibrium distribution is approached on this average medium network is given by the eigenvalue gap of the transition matrix W_{vu}^{avg} as $t_{\text{mix}}^{-1} = 1 - \omega_1$ where $\omega_0 = 1$ is the largest eigenvalue and ω_1 is the second largest eigenvalue as introduced in Sec. (3.1.3). The average medium transition matrix is circulant based on the vector

$$\mathbf{w} = k^{-1}(0, \underbrace{p_S, \dots, p_S}_{k/2}, \underbrace{p_L, \dots, p_L}_{N-1-k}, \underbrace{p_S, \dots, p_S}_{k/2}).$$

In this case, the j -th eigenvalue of W_{vu}^{avg} is given as

$$\omega_j = \sum_{v=1}^N w_v \exp(i2\pi v/N)$$

such that the second largest eigenvalue can be computed as $\omega_1 = p_S\Gamma/k - p_L(1 + \Gamma)/k$ where

$$\begin{aligned} \Gamma &= 2 \sum_{j=1}^{k/2} \cos(2\pi j/N) \\ &= k - \left(\frac{\pi}{N}\right)^2 \frac{k}{3} (k/2 + 1)(k + 1) + \mathcal{O}(N^{-4}) \end{aligned}$$

which yields the mixing time

$$t_{\text{mix}}(\beta) = \left[1 - \frac{\Gamma - \beta(1 + \Gamma)}{k + \beta(N - 1 - k)}\right]^{-1}. \quad (5.8)$$

Fig. 5.5 displays how both clustering coefficient and mixing time decrease with increasing long-range redistribution parameter β . In the limits one finds the results from k -regular networks and the average medium approximation of the Erdős-Rényi graph

$$\begin{aligned} t_{\text{mix}}(\beta = 0) &= \left[1 - \frac{\Gamma}{k}\right]^{-1} \xrightarrow{N \gg k/2} \frac{N^2}{\pi^2} \frac{3}{(k/2 + 1)(k + 1)}, \\ t_{\text{mix}}(\beta = 1) &= \left[1 - \frac{1}{N - 1}\right]^{-1} = 1 - \frac{1}{N}. \end{aligned}$$

⁷ also referred to as *messages* in the following

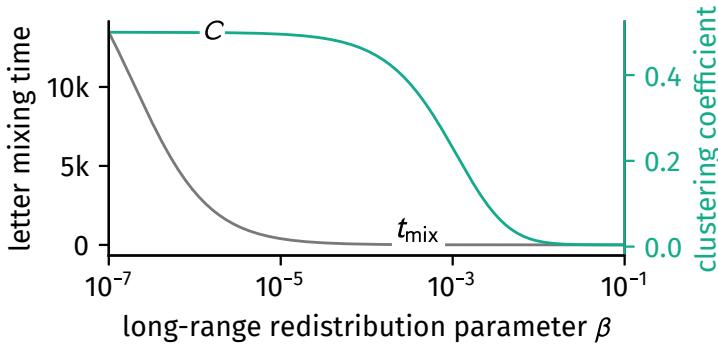


Figure 5.5: The small-world effect as illustrated by the modified Watts-Strogatz clustering coefficient and mixing time. The random walk message delivering mixing time Eq. (5.8) as an upper bound of the real message delivering mixing time decreases rapidly with the introduction of long-range links while clustering Eq. (4.17) preserves. Displayed here are results for $N = 1001$ and $k = 4$.

This implies that for small long-range redistributions the relative mixing time decreases as

$$\frac{t_{\text{mix}}(\beta)}{t_{\text{mix}}(0)} = 1 - \beta \left(\frac{3N^3}{\pi^2 (k/2 + 1) (k + 1) k} - \frac{N}{k} + \frac{1}{k} \right) + \mathcal{O}(\beta^2). \quad (5.9)$$

Comparing Eqs. (4.18) and (5.9), one can easily see that for small β the rate with which the mixing time decreases is of order N^3 while the rate with which the clustering coefficient decreases is of order N , which is a difference of two orders of magnitude. This shows that even with a small amount of long-range connection probability, the delivery time of randomly passed messages declines rapidly while clustering is still preserved. Note that the mixing time does not possess a minimum when varying the model's structural control parameter.

5.2 Actual Network Realizations

So far, theoretical results have been derived for average media, finding that neither mixing time nor pair-averaged FPT seem to be optimal in between the clustered limit and the random limit. In nature, media are not averaged. Instead, one finds structures where connections exist or not. Any information spreading on these networks has to follow discrete paths between nodes to reach other nodes to which its originating node is not connected. In this subsection, network samples from the models introduced in the previous chapter are investigated to see whether the average medium approximation used before fails to describe effects provided by discrete structure.

It will be found that those models from Ch. 4 which are based on lattice node positions indeed show a minimum in the pair-averaged FPT in between both limit structures, while the 2-torus PLSW model does not. A heuristic explanation of this effect is found by deriving a formula to estimate the pair-averaged FPT in a locally clustered environment, which is composed of the clustering coefficient and the degree variance.

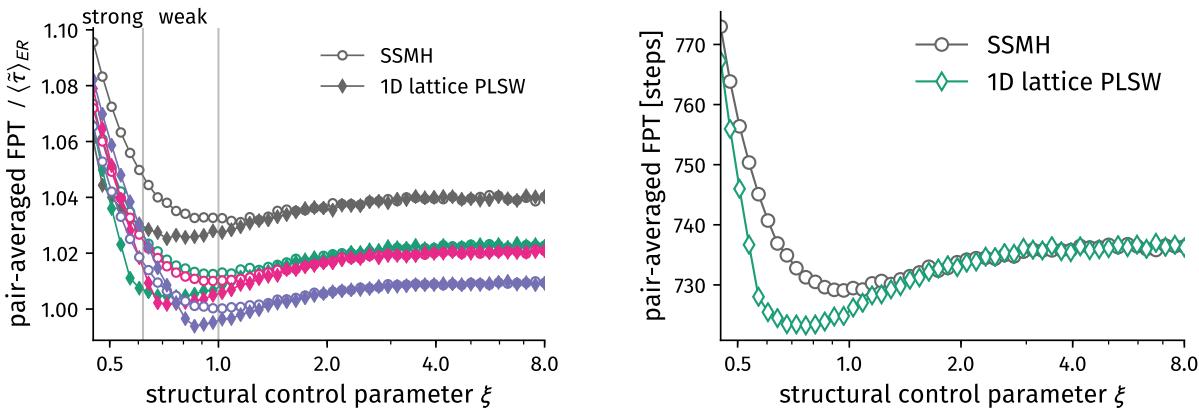


Figure 5.6: Pair-averaged first passage time computed numerically using Eq. (3.10) for the self-similar modular hierarchical (SSMH) and the 1D lattice PLSW network models introduced in the previous chapter. **(Left)** Pair-averaged FPT as a function of the structural control parameter ξ for $B = 8$, $L = 3$, and $\langle k \rangle \in \{6, 7, 8, 10\}$ (top to bottom curves, respectively), averaged over 2000 SSMH and 1D lattice PLSW network realizations and normalized by the lower bound of the pair-averaged FPT on Erdős-Rényi networks given in Eq. (5.12). **(Right)** Unnormalized result for $\langle k \rangle = 7$.

5.2.1 Numerical Results for Pair-Averaged First Passage Time

Similarly to the procedures of the previous sections and the previous chapter, the parameters $B = 8$ and $L = 3$, where chosen, here with mean degree values of $\langle k \rangle \in \{5, 6, 7, 8, 9, 10, 11\}$. Subsequently, different values of the structural control parameter $0.5 \lesssim \xi \leq B$ (or the corresponding μ parameter) were scanned. Note that the lower bound $0.5 \lesssim \xi$ was chosen here because at this point the pair-averaged FPT is already much greater than its value from the corresponding Erdős-Rényi limit, thus no remarkable information can be gained from further decreasing ξ , other than that it will increase even more.

For each parameter set, 2000 different SSMH network realizations were generated with $N = B^L$ nodes, using the algorithm described in Sec. 4.1. Similarly, 2000 PLSW network realizations were sampled with $N = B^L$ nodes, the same mean degree $\langle k \rangle$ values, and the corresponding structural control parameter $\mu = \log \xi / \log B$, using the algorithm described in Sec. 4.3.4. After selecting the largest component of each of the resulting networks, the pair-averaged FPT $\langle \tau \rangle$ was measured using the eigenvalues and eigenvectors of the unnormalized graph Laplacian, as described in Sec. 3.1.5.

As one may observe in Fig. 5.6, a minimum in the pair-averaged FPT emerges in the $\xi \lesssim 1$ region for, both, the SSMH and the 1D lattice PLSW networks. This corresponds to the point where the network structures leave the weak hierarchical clustering regime, where an average node has more connections to nodes in lower hierarchical layers than to nodes in higher layers, but where both quantities are of similar order.

Furthermore, the modified Watts-Strogatz model was tested to display a minimum in the pair-averaged FPT. The parameters $B = 8$, $L = 3$, $\langle k \rangle = 10$ where chosen and the structural control parameter $0.5 \lesssim \xi \leq 8$ was varied. The pair-averaged FPT for the largest component of each network sample was found numerically using

Eq. (3.10) and finally averaged over 10^4 independent samples each.

Finally, the 2-torus PLSW model, while structurally interpretable to be modular hierarchical, does not display a minimum in the pair-averaged FPT. Eq. (3.10) was used on the largest component of 10^4 independent samples each for $N = 8^3$ and varying mean degree $\langle k \rangle \in \{6, 8, 12\}$, as well as increasing structural control parameter $-\kappa$. The result can be seen in Fig. 5.7.

5.2.2 Heuristic Explanation of Minimum

Given that the AMA approach described above fails to reproduce the minimal diffusion time observed numerically at intermediate levels of hierarchical modularity (see Fig. 5.6), a heuristic approach is developed below that will help understand the origin of this effect.

The following calculation is based on the heuristic lower bound of the global mean FPT of a target node v as introduced in Sec. 3.1.5. The base assumption is that the network is locally tree-like (i.e., that it has vanishing clustering coefficient) and that the mean FPT for any target node v is exponentially distributed as $\exp(-t\beta_v)$, with

$$\beta_v = \frac{k_v}{N \langle k \rangle} \left(1 - \frac{1}{\langle k \rangle} \right).$$

Since the mean FPT is only asymptotically exponentially distributed, by using this assumption only a lower bound will be obtained for the corresponding pair-averaged FPT. One thus finds $\langle \tau \rangle \geq N^{-1} \sum_i \beta_i^{-1}$, which yields

$$\langle \tau \rangle \geq N \langle k \rangle \left\langle \frac{1}{k} \right\rangle_{k>0} \frac{1}{1 - \langle k \rangle^{-1}}, \quad (5.10)$$

where $\langle \cdot \rangle_{k>0}$ denotes the average over all nodes with non-zero degree. Even though tree-like approximations are often remarkably useful even in clustered networks, they lose their validity when the average shortest path length becomes too distinct from the average path length in a random network [51], which is the case for the network models from Ch. 4. Hence, an extension of this result is derived in the following, which explicitly considers a non-vanishing clustering coefficient.

The approach taken in [38] treats the time dependent walker concentration $p_u(t)$ on any node u at time t as composed of two separable distributions

$$p_u(t) \approx p_u p_{\text{total}}(t).$$

and finds that the time-independent walker concentration on neighboring nodes u of sink node v follows as

$$p_u = \sum_{u' \in \text{Nei}(u) \setminus v} \frac{p_{u'}}{k_{u'}} = \frac{1}{N \langle k \rangle} (k_u - 1)$$

when the networks are locally tree-like. A better approximation for locally clustered networks is given as follows, where the argumentation follows Sec. 3.1.5 and picks up at Eq. (3.12). It is furthermore illustrated in Fig. 5.8.

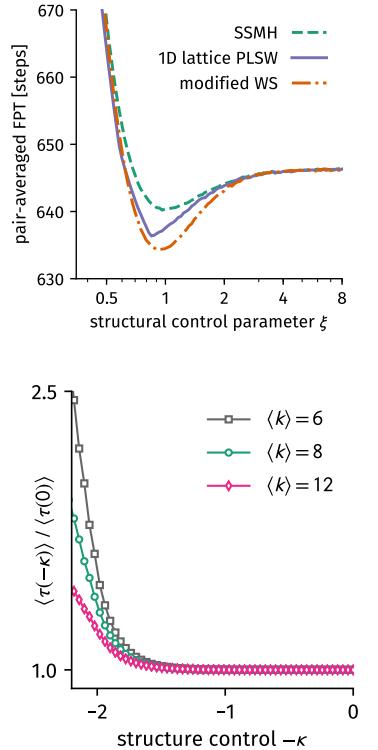
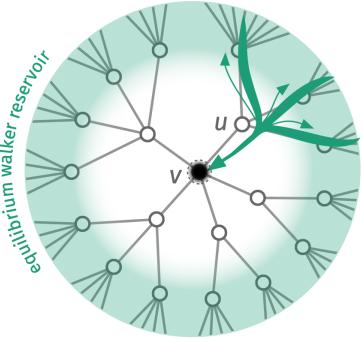
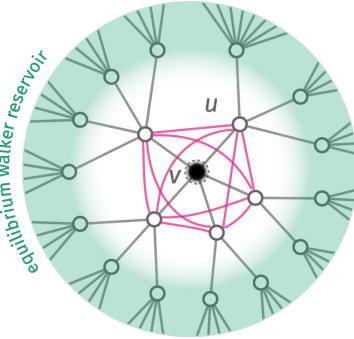


Figure 5.7: Pair-averaged FPT for network samples of the main small-world models introduced in Ch. 4. **(Top)** The modified Watts-Strogatz, self-similar modular hierarchical as well as the 1D lattice power-law small-world model all display a minimum for an intermediate degree of long-range connection probability. **(Bottom)** The 2-torus power-law small-world model does not show a minimum in the pair-averaged FPT.

original heuristic:
tree-like approximation allows for
direct flow into sink



adjusted heuristic:
introducing average links between
neighbors using clustering coefficient C



adjusted heuristic:
flow around sink is possible

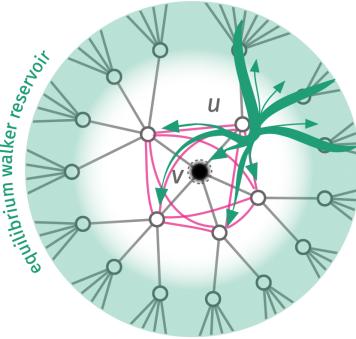


Figure 5.8: Illustration of the heuristic method to find the global mean first passage time (FPT) of a sink node v with non-zero clustering. For both the original and the adjusted heuristic, all nodes which are more than two steps away from the sink node are considered to carry an amount of remaining walkers which is proportional to their equilibrium walker density. **(Left)** In the original heuristic, a locally tree-like network is assumed and hence, walkers can flow from the reservoir into first neighbors, displayed here for an example node u . From there they can either flow into the sink or back into the reservoir. **(Center)** In the adjusted heuristic, local clustering is considered by introducing average links between all neighbors of v using the clustering coefficient C . **(Right)** Due to the added links between first neighbors, walkers can flow around the sink before being absorbed or flowing back into the reservoir, hence increasing the arrival time.

Consider sink node v with degree k_v . On average, a neighbor u of v will have $c_v = C_v(k_v - 1)$ neighbors that are also neighbors of v . The rest of its $k_u - 1 - c_v$ neighbors each contribute an influx of $(N \langle k \rangle)^{-1}$ walkers, such that

$$p_u = \frac{1}{N \langle k \rangle} [k_u - 1 - c_v] + \sum_{u' \in \text{Nei}(u) \setminus v} \frac{p_{u'}}{k_{u'}}. \quad (5.11)$$

Now, every neighbor u is replaced with an average node of degree $\langle k \rangle$. In addition, every node i is assumed to have the same number of edges between neighbors, so $c_v \equiv c = C(\langle k \rangle - 1)$. This transforms Eq. (5.11) into a self-consistent expression given by

$$p_u = \frac{1}{N \langle k \rangle} (\langle k \rangle - 1 - c) + c \frac{p_u}{\langle k \rangle},$$

which is equivalent to

$$p_u = \frac{1}{N} \left(1 - \frac{1}{\langle k \rangle - C[\langle k \rangle - 1]} \right).$$

Using Eq. (3.11), the decay rate is then found to be

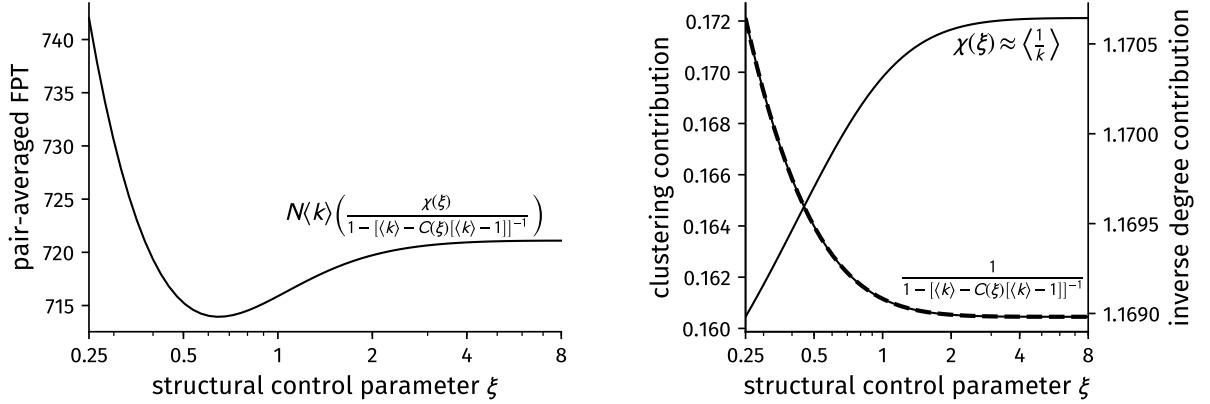
$$\beta_v = \frac{k_v}{N \langle k \rangle} \left(1 - \frac{1}{\langle k \rangle - C[\langle k \rangle - 1]} \right).$$

One therefore finds that the global mean FPT of target node v is approximately lower bounded by

$$\tau_v \geq \frac{N \langle k \rangle}{k_v} \left(\frac{1}{1 - [\langle k \rangle - C[\langle k \rangle - 1]]^{-1}} \right). \quad (5.12)$$

The pair-averaged FPT (i.e. the mean global mean FPT) $\langle \tau \rangle = (1/N) \sum_v \tau_v$ is consequently lower bounded by

$$\langle \tau \rangle \geq N \langle k \rangle \left\langle \frac{1}{k} \right\rangle_{k>0} \left(\frac{1}{1 - [\langle k \rangle - C[\langle k \rangle - 1]]^{-1}} \right). \quad (5.13)$$



In order to quantitatively evaluate this result for the SSMH model one needs to compute the mean inverse degree $\langle k^{-1} \rangle_{k>0}$, which can be numerically intensive. A further approximation of the result is given by expressing this quantity by its second order Taylor expansion around the mean $\langle k \rangle$, such that

$$\begin{aligned} \left\langle \frac{1}{k} \right\rangle_{k>0} &\approx \chi(\xi) = \frac{1}{\langle k \rangle} + \frac{\langle k^2 \rangle}{\langle k \rangle^3} + \Phi_{ER} \\ &= \frac{\langle k^2 \rangle - \langle k^2 \rangle_{ER}}{\langle k \rangle^3} + \left\langle \frac{1}{k} \right\rangle_{ER, k>0}. \end{aligned} \quad (5.14)$$

For the result Eq. (5.13) the expressions for the clustering coefficient C in Eq. (4.11) and for the second moment of the degree distribution in Eq. (4.7) are used. The constant Φ_{ER} contains the error made by considering networks without nodes of degree $k = 0$, which alters the degree distribution in comparison to the second moment $\langle k^2 \rangle$. This issue can be fixed by demanding $\chi(\xi = B) = \langle k^{-1} \rangle_{ER, k>0}$, which can be found by averaging k^{-1} for $k > 0$ using a binomial pmf with $n = B^L$ and $p = k/(B^L - 1)$, yielding

$$\Phi_{ER} = \left\langle \frac{1}{k} \right\rangle_{ER, k>0} - \frac{1}{\langle k \rangle} - \frac{\langle k^2 \rangle_{ER}}{\langle k \rangle^3}.$$

The function Eq. (5.13) is plotted in Figure 5.9a for the SSMH model. Its shape as a function of ξ results from two opposing effects (see Figure 5.9b): the growth of the mean inverse degree and the decay of the clustering coefficient (as derived in Sec. 4.1.2) for increasing ξ . The mean inverse degree as approximated by $\chi(\xi) \propto \langle k^2 \rangle$, grows monotonically with ξ , as shown in Sec. 4.1.2. This results from the fact that the shape of the degree distribution is rather narrow for small ξ values, which leads to a vanishing number of nodes with small degrees. At the same time, this monotonic growth is countered by the clustering behavior because, as the structural control parameter ξ is increased, the network loses its strong clustering and thus the clustering contribution decreases. The combination of these two curves leads to the appearance of a minimum in the pair-averaged FPT at intermediate levels of hierarchical modularity, as shown in

Figure 5.9: Heuristic result (a) for the pair-averaged FPT as a function of the structural control parameter ξ , as given by Eq. (5.12) with parameters $B = 8$, $L = 3$ and $\langle k \rangle = 7$, and (b) the mean inverse degree and clustering coefficient contributions to this curve, showing how the combination of these growing and decaying functions can result in the observed minimum shown in Fig. 5.6.

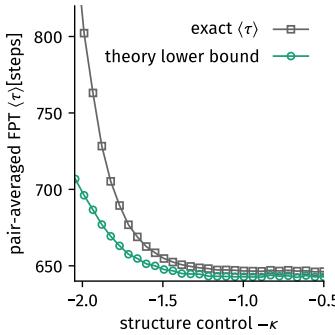


Figure 5.10: Comparison between the actual pair-averaged FPT and the heuristic lower bound on a 2-torus PLSW model with $N = 8^3$ and $\langle k \rangle = 10$. The exact value was obtained using Eq. (3.10) and the lower bound was computed using Eq. (5.13). Markers represent the averages over 1000 independent network samples for each of which the largest component was extracted. The lower bound uses the mean inverse degree $\langle k^{-1} \rangle$ and the mean clustering coefficient C which were computed from those largest components, as well.

Figure 5.9a. The reader should note, however, that the position of the minima given by Eq. (5.12) differs from those in Fig. 5.6, due to the heuristic origin of this equation.

The nature of Eq. (5.12) further explains why the networks generated from the 2-torus PLSW model do not show a minimum in the pair-averaged FPT as can be seen in Fig. 5.7. As argued above, the minimum observed for the other models emerges from a varying degree variance which increases with increasing amount of long-range connection probability. However, it was shown in Sec. 4.3 that the 2-torus PLSW model has constant degree variance over the total range of its structural control parameter (see Fig. 4.12). Hence the minimum cannot exist as only the clustering contribution of Eq. (5.12) plays a role. A comparison of the exact pair-averaged FPT Eq. (3.10) and the lower bound Eq. (5.13) is shown in Fig. 5.10.

One might wonder how necessary the extension of the heuristic to non-zero clustering actually is. Considering the original result Eq. 5.10 it now becomes clear that by considering locally tree-like networks, only the inverse degree contribution would influence the shape of the pair-averaged FPT as a function of the structural control parameters and thus it would remain constant for the 2-torus PLSW. For the remaining models it would increase with increasing structural control parameters because the inverse mean degree increases, directly contradicting the numerical results.

It should finally be noted that even though the adjusted heuristic produces a result which is qualitatively in agreement with the numerically evaluated pair-averaged FPT, the difference between the heuristic lower bound and the exact value increases for all models when approaching the regular clustered limit (i.e. for low values of the structural control parameter where the networks become increasingly lattice-like). This discrepancy is expected to arise since the adjusted heuristic still makes the assumption that the random walk process equilibrates in a few number of time steps $t_{\text{mix}} \ll N$. In networks which are structurally close to lattices, the shortest path length between pairs of nodes is substantially larger compared to more random networks and so the equilibration process will take much longer. One might expect that in a further adjusted heuristic, the increasing path length of lattice-like networks should play a role. This, however, is left for future investigations.

5.3 Summary and Discussion

Within this chapter the hypothesis of modular hierarchical networks being optimal for a fast pair-averaged first passage time in random walks was tested. At first, an average medium approximation was used to find that average modular hierarchical and other small-world models do not possess minimal passage or mixing times. This approximation, however, allowed for the analytic evaluation of those network properties as well as first passage time statistics. It was shown that the SSMH and one-dimensional (1D) lattice PLSW model

closely correspond considering their passage-time statistics, reflecting the result from Ch. 4 that they are of similar structure.

Subsequently it was shown that the pair-averaged first passage time does indeed show a minimum for the SSMH, 1D lattice PLSW and modified Watts-Strogatz model when analyzing actual networks sampled from the models, an effect not replicated by the average medium. In order to explain the minimum, a heuristic formula for the pair-averaged first passage time on networks with non-zero local clustering was derived. Following this heuristic, the minimum emerges as a combined effect of increasing degree variance and decreasing clustering coefficient with increasing long-range connection probability.

The increasing node degree variance, however, seems to be an artifact stemming from discrete node positions in the SSMH, 1D lattice PLSW and modified WS model. In the 2-torus PLSW model, which is equally capable of producing networks describable as hierarchically modular, the node degree variance is constant. Consequently the minimum vanishes for this network model.

Hence, an optimal structure for fast diffusion can be expected of networks which are in a state where both their degree variance and clustering coefficient are low compared to other structures, resulting a minimum pair-averaged FPT. However, no such statement can be made in general for modular hierarchical networks.

In the following chapter, this analysis will expanded to the mean cover time of random walks.

6

Cover Time for Arbitrary Complex Networks

For a discrete-time random walk, the cover time refers to the number of steps a random walker takes to visit each node in a network at least once (see Sec. 3.1.6). It places an upper bound on exhaustive search processes and the time to distribute a commodity within a network. Since neither a simple method nor a heuristic exist for reliably computing this observable, a heuristic method is introduced in the following. In Sec. 6.1, the heuristic approach is explained and used to derive the cover time probability distribution as well as its first moment. Subsequently, the heuristic is applied to compute the mean cover time for various network models and real-world networks in Sec. 6.2. The analysis is finally concluded with a detailed error assessment in Sec. 6.3, alongside which it will be revealed that some of the small-world network models introduced in Ch. possess a minimum cover time in the strong hierarchically clustered regime while the 2-torus PLSW does not. Both will be explained using the heuristic introduced in Sec. 5.2.2. This chapter closely follows the text published in [43].

6.1 Heuristic Evaluation Method

6.1.1 Derivation

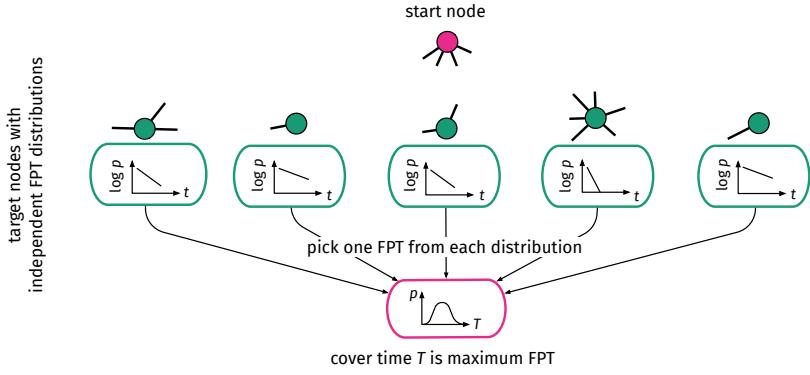
As has been argued in Sec. 3.1.5, given that a random walk process on a network equilibrates quickly, the information about the initial walker position becomes irrelevant rather rapidly. Here, fast equilibration refers to the fact that the initial random walker density on nodes of the network approaches the equilibrium density Eq. (3.3) in a small number of time steps $t_{\text{mix}} \ll N$. As shown in [38] and argued in Sec. 3.1.5, the first passage time at a target node v is then distributed asymptotically according to

$$p_v(\tau) \propto \exp(-\tau/\tau_v) \quad (6.1)$$

where τ_v is the global mean FPT of Eq. (3.9). In this case, the asymptotic behavior refers to larger FPTs τ . The global mean FPT τ_v can vary for each node and only depends on the structural properties of

Figure 6.1: Illustration of the heuristic approach. Each node of the network is represented as an independent object that can be visited by the walker starting at the pink node. Each target node v is associated with its own first passage time (FPT) distribution. This is FPT distribution is asymptotically distributed as $\propto \exp(-\tau/\tau_v)$ (for larger times τ) with its global mean FPT τ_v . In order to compute the cover time, one FPT τ is drawn from every target's distribution. The cover time is then given as the maximum time of all drawn FPTs.

¹ Note that in the following derivations, the FPT decay rate $\beta_v = \tau_v^{-1}$ will be often used instead of the global mean FPT to simplify the notation.



the network.¹

Following the heuristic approach presented below, one can proceed as illustrated in Fig. 6.1 to obtain, first, the cover time distribution and second, the mean cover time, both based on the collection of target FPT distributions $p_v(\tau)$. Ignoring the start node u at which a walker begins its random walk, an FPT t_v is drawn for each potential destination v from their respective target FPT distribution p_v , resulting in the set \mathcal{F} which contains all $N - 1$ target FPTs. The cover time T_u is then given as the maximum element of the set \mathcal{F} . The probability that a time T is an upper bound of \mathcal{F} can be found as the probability that no element of \mathcal{F} is greater than T , which is

$$\begin{aligned} P_u(T) &= P_u("T \text{ is an upper bound of all target FPTs}") \\ &= P_u(t_v \leq T \ \forall t_v \in \mathcal{F}) = \prod_{v \neq u} P(t_v \leq T) \end{aligned} \quad (6.2)$$

and is equal to the cumulative distribution function (cdf) of T . Assuming the validity of Eq. (6.1), the discrete time probability mass function for first passage time τ at node v is evaluated as

$$\begin{aligned} p_{v,\tau} &= \exp(-\beta_v \tau) / \sum_{t=1}^{\infty} \exp(-\beta_v t) \\ &= [1 - \exp(-\beta_v)] \exp(-\beta_v(\tau - 1)), \end{aligned}$$

yielding the cdf

$$\begin{aligned} P(t_v \leq T) &= \sum_{t=1}^T p_{v,t} = (1 - \exp(-\beta_v)) \sum_{t=1}^T \exp(-\beta_v(t - 1)) \\ &= 1 - \exp(-\beta_v T). \end{aligned} \quad (6.3)$$

Eqs. (6.2) and (6.3) yield the cumulative distribution function for the cover time,

$$P_u(T) = \prod_{v \neq u} (1 - \exp(-\beta_v T)).$$

The cdf can be further approximated by assuming a continuous time distribution, which simplifies the following derivations while the outcome does not change significantly, as explained in Sec. 6.3, such

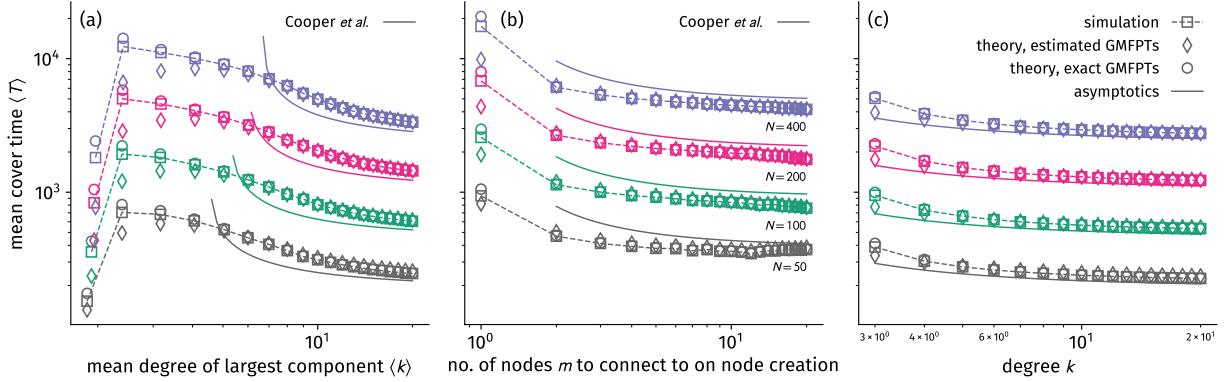


Figure 6.2: The mean cover time $\langle T \rangle$ of the largest component of (a) Erdős-Rényi (ER) networks, (b) Barabási-Albert (BA) networks, and (c) random k -regular networks. Shown are averages measured from 1000 simulations per data point (\square) and both estimations of the mean cover time using (\diamond) estimated GMFPTs from the target nodes' degrees Eq. (6.9) and (\circ) exact GMFPTs Eq. (6.9) computed from the unnormalized graph Laplacian's spectrum in (a) and (b). Respectively, Eq. (6.10) has been used for random k -regular networks in (c). Dashed lines are simulation results and a guide to the eye. Further displayed are the asymptotic results derived for the ER [12] and BA [13] network models as well as the lower bound Eq. (6.11).

that the mean cover time of start node u is approximately

$$T_u \approx \int_0^\infty dT \left[1 - \prod_{v \neq u} (1 - \exp(-\beta_v T)) \right] - \frac{1}{2}. \quad (6.4)$$

Note that a bias of $1/2$ emerges when changing from discrete time to continuous time, as discussed in Sec. 6.3. However, the mean cover time is lower bounded by the cover time of a complete graph which scales as $N \log N$, such that one usually finds $T_u \gg 1/2$. Hence, the bias will be omitted in the following, which introduces relative error of size $(2T_u)^{-1}$. Now, the global mean cover time can be found by taking the average over all target nodes u as

$$\langle T \rangle \approx \int_0^\infty dT \left[1 - P(T) \frac{1}{N} \sum_{u=1}^N \frac{1}{1 - \exp(-\beta_u T)} \right] \quad (6.5)$$

where

$$P(T) = \prod_{v=1}^N (1 - \exp(-\beta_v T)). \quad (6.6)$$

As shown in Sec. 6.3.3, introducing small relative error of order $\mathcal{O}((N \log N)^{-1})$ for the networks discussed in this paper, one can further simplify this integral to find

$$\begin{aligned} \langle T \rangle &\approx \int_0^\infty dT [1 - P(T)] = \int_0^\infty dT \left[1 - \prod_{v=1}^N (1 - \exp(-\beta_v T)) \right] \\ &= \sum_{S \in \mathcal{P}^*(\mathcal{V})} (-1)^{|S|+1} \left(\sum_{v \in S} \beta_v \right)^{-1}, \end{aligned} \quad (6.7)$$

where \mathcal{V} is the set containing all nodes and $\mathcal{P}^*(\mathcal{V})$ is the set of all possible subsets of \mathcal{V} (excluding the empty set). Conceptually, this integral corresponds to a situation where an additional node is inserted on which every random walk starts but which can never be visited again. Even though the integral Eq. (6.7) can be solved analytically to obtain the result above, in practice it is more feasible to solve the integral numerically than iterating over $\mathcal{P}^*(\mathcal{V})$ which has $2^N - 1$ elements and hence becomes very large rather quickly.

Now, the estimation of the global mean cover time reduces to an efficient estimation of the FPT decay rates β_v . There are two ways to estimate the decay rates with the GMFPTs as described in Sec. 3.1.6. Using the estimation of the lower bound Eq. (3.13), the estimated global mean cover time is given by

$$\langle T^{\text{es}} \rangle \geq \int_0^\infty dT \left[1 - \prod_{v=1}^N \left(1 - \exp \left(-Tk_v \frac{1 - \langle k \rangle^{-1}}{N \langle k \rangle} \right) \right) \right]. \quad (6.8)$$

The advantage of this method is that only the network's degree sequence k_v needs to be known in order to estimate the global mean cover time. However, this method can obviously only account for a lower bound. One can also compute the exact GMFPTs using Eq. (3.10). In this case the computed global mean cover time is

$$\langle T^{\text{ex}} \rangle = \int_0^\infty dT \left[1 - \prod_{v=1}^N \left(1 - \exp \left(-\frac{T}{\tau_v^{\text{ex}}} \right) \right) \right]. \quad (6.9)$$

6.1.2 Cover Time of Networks With Equal GMFPTs

Considering a network in which all nodes have approximately the same GMFPT $\tilde{\tau}$ and on which a random walk equilibrates quickly ($t_{\text{mix}} \ll N$), the mean cover time can be estimated using Eq. (6.4) to obtain

$$\begin{aligned} \langle T \rangle (\tilde{\tau}) &\approx \int_0^\infty dT \left[1 - (1 - \exp(-T/\tilde{\tau}))^{N-1} \right] \\ &= \tilde{\tau} [\gamma + \psi(N)], \end{aligned} \quad (6.10)$$

where $\gamma \approx 0.57722$ is the Euler-Mascheroni constant, $\psi(z) = \Gamma'(z)/\Gamma(z)$ and $\Gamma(z)$ the gamma function.

An example for networks fulfilling the conditions above are random k -regular networks as introduced in Sec. 2.3.1 where all nodes have identical degree and the networks possess random structure (as opposed to, e.g. lattice networks on a torus, where all nodes have identical degree but are only connected to their nearest neighbors). The complete graph with N nodes is a $(N-1)$ -regular network and thus a special case of k -regular random networks. The cover time of the complete graph is given as $\langle T \rangle = (N-1)(\log(N-1) + \gamma + \mathcal{O}(N^{-1}))$ [40], a result which is reproduced by Eq. (6.10) since the GMFPT for each node² is $\tilde{\tau} = N-1$ (see Sec.) and $\psi(N+1) = \log N + \mathcal{O}(N^{-1})$. For general random k -regular networks, Eq. (6.10) can be used to find an approximate scaling relation for the lower bound

$$\langle T \rangle \gtrsim \frac{k}{k-1} N \log N \quad (6.11)$$

using the GMFPT lower bound Eq. (3.13), the fact that $k_v = \langle k \rangle = k$ and $\psi(N+1) = \log N + \mathcal{O}(N^{-1})$.

² Suppose a random walker starts at any node u . The probability to reach any other node of the network in one time step is $p = 1/(N-1)$. Looking at a single target node v the probability that v is first passed at time t is given as $p_t = (1-p)^{t-1}p$. Hence, the GMFPT for every target node is $\tilde{\tau} = \sum_{t=1}^\infty t p_t = -p \frac{\partial}{\partial p} \frac{1}{p} = \frac{1}{p} = N-1$.

Network	N	$\langle k \rangle$	$\frac{\langle T^{\text{sim}} \rangle}{N \log N}$	$\frac{\langle T^{\text{es}} \rangle}{N \log N}$	$1 - \frac{\langle T^{\text{sim}} \rangle}{\langle T^{\text{es}} \rangle}$	$\frac{\langle T^{\text{ex}} \rangle}{N \log N}$	$1 - \frac{\langle T^{\text{sim}} \rangle}{\langle T^{\text{ex}} \rangle}$
Intra-org. contacts - Cons. (info) [14]	43	15.3	2.40	2.41	0.004	2.38	0.007
Intra-org. contacts - Cons. (value)	44	16.0	2.00	2.02	0.006	2.07	0.031
Intra-org. contacts - Manuf. (awareness)	77	25.5	3.43	3.47	0.012	3.46	0.009
Intra-org. contacts - Manuf. (info)	76	23.3	2.35	2.29	0.028	2.37	0.007
Matches of the NFL 2009 [1]	32	13.2	1.20	1.27	0.053	1.21	0.008
Connections between 500 largest US airports [11]	500	11.9	12.34	10.30	0.198	13.18	0.063
B. F. Maier's Facebook friends network [45]	329	11.9	11.63	8.45	0.377	12.36	0.059
C. Elegans neural network [81]	297	14.6	8.63	9.15	0.057	8.69	0.007
<i>E. Coli</i> protein interaction [70]	329	2.8	5.55	4.27	0.298	7.24	0.233
Social interaction in dolphins [41]	62	5.1	4.72	4.46	0.059	4.86	0.029
American college football [24]	115	10.7	1.38	1.27	0.081	1.40	0.017
Food web of grassland species [16]	75	3.0	4.66	3.97	0.172	5.17	0.099
Zachary's Karate club [84]	34	4.5	3.04	3.29	0.079	3.06	0.006
Interactions in "Les Misérables" [33]	77	6.6	6.81	6.21	0.097	7.21	0.055
Network of associations between terrorists [35]	62	4.9	4.65	4.47	0.040	4.88	0.046

6.2 Numerical Results

In this section, the predictions of Eqs. (6.8) and (6.9) are compared with simulation results for single component ER, BA and real-world networks, as well as Eq. (6.10) for random k -regular networks. For the simulations, a single walker was placed on every node at time $t = 0$. Subsequently, each walker performed a random walk as described in Sec. 3.1.6. Each walker proceeded until it visited each node at least once, completing total coverage and marking cover time T_u . $\langle T \rangle$ was computed as the average of all T_u .

For both ER and BA networks, networks were generated with $N \in \{50, 100, 200, 400\}$ nodes, ER networks with node connection probability $\{k/(N-1) : k \in \mathbb{N}, 1 \leq k \leq 20\}$, and BA networks with $\{m : m \in \mathbb{N}, 1 \leq m \leq 20\}$. In order to test Eq. (6.10), random k -regular networks were generated using the algorithm given in [73] with $N \in \{50, 100, 200, 400\}$ nodes and node degree $k_u = k \forall u \in \mathcal{V}$, scanning integer degrees $\{k : k \in \mathbb{N}, 3 \leq k \leq 20\}$. Network generation was followed by the extraction of the largest component and a subsequent random walk simulation as described above. The cover time was estimated using Eqs. (6.8), (6.9) and Eq. (6.1), respectively, for 1000 networks each. For Eq. (6.10) and the random k -regular networks, the global mean FPTs $\bar{\tau}^{\text{es}} = N/(1-k^{-1})$ and $\bar{\tau}^{\text{ex}} = N^{-1} \sum_{v=1}^N \tau_v^{\text{ex}}$ were used, respectively.

The theoretic results are in agreement with the simulation results, as can be seen in Fig. 6.2. The relative error decreases with increasing number of nodes N as well as increasing mean degree $\langle k \rangle$ and quickly reaches values below 1%. Unsurprisingly, the heuristic method performs better compared to the results of [12, 13] due to the asymptotic nature of the latter.

In addition to simulations on network models, simulations were performed for the largest component of 15 real-world networks, listed in Tab. 6.1. Initially directed networks were converted to undirected

Table 6.1: Mean cover times of simple discrete-time random walks on the largest component of various networks, in units of the cover time on a complete graph with equal node count. Displayed is the number of nodes N , the mean degree $\langle k \rangle$ of the largest component and the measured mean cover time $\langle T^{\text{sim}} \rangle$ extracted from 50 simulations per network with one walker starting on every node. Additionally shown are both theoretical estimations of the cover time using (a) estimated GMFPTs $\langle T^{\text{es}} \rangle$ from the target nodes' degrees, Eq. (6.8) and (b) exact GMFPTs $\langle T^{\text{ex}} \rangle$ computed from the unnormalized graph Laplacian's spectrum Eq. (6.9). Both estimations are given with their relative error to the simulated mean cover time. Note that all networks have been symmetrized and an unweighted link (u, v) has been created if a weight between two nodes was $w_{uv} > 0$. For the Intra-organizational networks, we created a link if both nodes put something else than "I do not know this person" in their questionnaire. These values are additionally shown in Fig. 6.5 (top).

Network	N	$\langle k \rangle$	$\frac{\langle T^{\text{sim}} \rangle}{N \log N}$	$\frac{\langle T^{\text{ex}} \rangle}{N \log N}$	$1 - \frac{\langle T^{\text{sim}} \rangle}{\langle T^{\text{ex}} \rangle}$	$\frac{\langle T^{\text{ex}} \rangle}{N \log N}$	$1 - \frac{\langle T^{\text{sim}} \rangle}{\langle T^{\text{ex}} \rangle}$
Barcelona	128	2.2	9.7	2.7	2.6	12.3	0.2
Beijing	104	2.2	10.5	2.7	2.9	16	0.3
Berlin	170	2.1	14.7	2.8	4.3	21.2	0.3
Chicago	141	2.1	14.9	2.7	4.4	20	0.3
Hong Kong	82	2.1	10.1	3.0	2.4	13.1	0.2
London	266	2.3	14.8	2.8	4.3	19.3	0.2
Madrid	209	2.3	14.4	2.7	4.4	20.4	0.3
Mexico	147	2.2	10.9	2.7	3.0	14.8	0.3
Moscow	134	2.3	12.0	2.9	3.1	14.7	0.2
New York	433	2.2	16.2	2.6	5.1	24.3	0.3
Osaka	108	2.3	8.9	2.9	2.1	11.4	0.2
Paris	299	2.4	11.4	2.8	3.0	14.2	0.2
Seoul	392	2.2	19.8	2.6	6.7	31.2	0.4
Shanghai	148	2.1	14.5	2.7	4.4	20	0.3
Tokyo	217	2.4	12.9	2.8	3.7	17.7	0.3

Table 6.2: Same procedure as in Tab. 6.1, but for subway networks of big cities, taken for the year 2009 from [65]. These values are additionally shown in Fig. 6.5 (bottom).

networks replacing every directed link with an undirected link. For weighted networks an undirected link (u, v) was assigned if a weight was $w_{uv} > 0$. For the intra-organizational networks [14], employees had to fill out questionnaires regarding their relationships to co-workers. Here, an undirected link (u, v) was assigned if both u and v marked something else than “I do not know this person”. As can be seen in Tab. 6.1, the heuristic method produces results that are very close to the simulated values (mostly relative errors of $< 10\%$). Exceptions are the computed cover times for the *E. coli* protein interaction network [70] with a relatively high relative error of $\approx 23\%$ and the grassland food web [16] with a relative error of $\approx 10\%$.

Furthermore, some example estimate cover time distributions as per Eq. (6.6) are compared to measured distributions from the simulation. As can be seen in Fig. 6.4, the cover time distributions using the exact GMFPTs predict the covertime pdf and complementary cdf (ccdf) reasonably well. In some cases, even the estimated GMFPTs show reasonable agreement. The last panel in Fig. 6.4 shows some significant deviation from the simulations, which will be further discussed below.

6.3 Error Analysis

6.3.1 Low-Dimensionality Induced Deviations

Concerning the impact of network structure on the error of our heuristic compared to the true mean cover time, networks with a large mixing time seem to consistently be associated with higher errors. Since the results are derived under the assumption that the relaxation time is $t_{\text{mix}} \ll N$, the relative error was measured against the ratio t_{mix}/N

to find

$$\text{rel. err.} \approx 0.19 \times \left(\frac{t_{\text{mix}}}{N} \right)^{0.94} \quad (6.12)$$

as can be seen in Fig. 6.6. Here, t_{mix} was computed using Eq. (3.6). This result indicates that increasing relaxation time indeed increases the error of the heuristic. Additionally, simulations were performed on D -dimensional lattices of dimension $D \in \{1, 2, 3\}$ (chains, squares and cubes) using node numbers $N \in \{(2n+1)^2 : n \in \mathbb{N}, 2 \leq n \leq 12\}$ for $D = 2$ and $N \in \{n^3 : n \in \mathbb{N}, 4 \leq n \leq 8\}$ for $D = 1$ and $D = 3$. For low-dimensional lattice networks with $D \leq 2$, the relaxation time is large compared to a variety of complex networks (see Fig. 1 in [20]). Hence, one might suspect that the method introduced above will not perform well for low-dimensional lattice networks.

Indeed, as can be seen in Fig. 6.3, the relative error between simulation and heuristic results increases with increasing N , up to $\approx 110\%$ for chains and $\approx 10\%$ for square lattices using exact GMFPTs, whereas smaller relative errors of up to $\approx 4\%$ are reached for cube lattices. Similar results are obtained for real-world networks embedded in a two-dimensional space with short-range connection probability such as subway networks [65] (shown in Tab. 6.2 and Fig. 6.5). Here, the estimation from estimated GMFPTs systematically underestimates the cover time while using exact GMFPTs yields an overestimation of the cover time by $\approx 20\% - 40\%$. Generally, the more exact result of GMFPTs calculated via the unnormalized graph Laplacian gives results with lower relative error than using lower bound GMFPTs, as expected.

This result can be even more refined by investigating the relative error of the method for networks generated from the small-world models devised in Ch. 4 as those naturally interpolate between low-dimensional lattice-like networks where the heuristic has large deviations and Erdős–Rényi networks where the heuristic yields satisfying results. Such an analysis will be done in Sec. 6.4.

6.3.2 Systematic Error: Continuous-Time Approximation

Considering discrete time, the mean cover time is given as the series, respectively partial sum

$$T_u = \sum_{T=1}^{\infty} [1 - P_u(T)] = \sum_{T=1}^{\infty} \bar{P}_u(T) \approx \sum_{T=1}^{T_{\max}} \bar{P}_u(T)$$

where the upper boundary was approximated by numerically finding a T_{\max} with $\bar{P}_u(T_{\max}) \leq 10^{-10}$. This partial sum is equal to the trapezoidal approximation of the integral

$$\int_0^{T_{\max}} dT \bar{P}_u(T) = \frac{1}{2} \bar{P}_u(0) + \sum_{T=1}^{T_{\max}-1} \bar{P}_u(T) + \frac{1}{2} \bar{P}_u(T_{\max}) + \Phi$$

with $\Delta T = 1$. Since the function $\bar{P}_u(T)$ has value $\bar{P}_u(0) = 1$, using the integral instead of the sum introduces a systematic error of

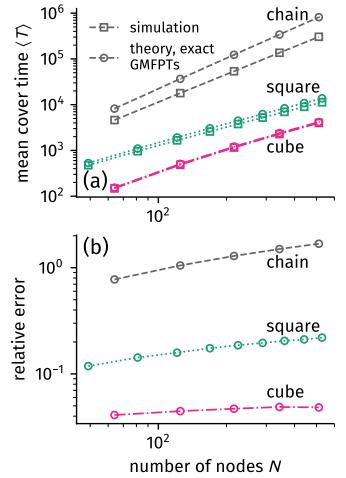


Figure 6.3: Example of the heuristic method yielding results with rather large deviations from simulations. (a) Mean cover time for low-dimensional ($D \leq 2$) lattices as well as lattices in dimension $D = 3$ as (\square) an average over 1000 simulations for each data point and (\circ) theoretical result from exact GMFPTs computed from the unnormalized graph Laplacian's spectrum Eq. (6.9). (b) The relative error is increasing with increasing system size but is comparably lowest for $D = 3$ (cubes).

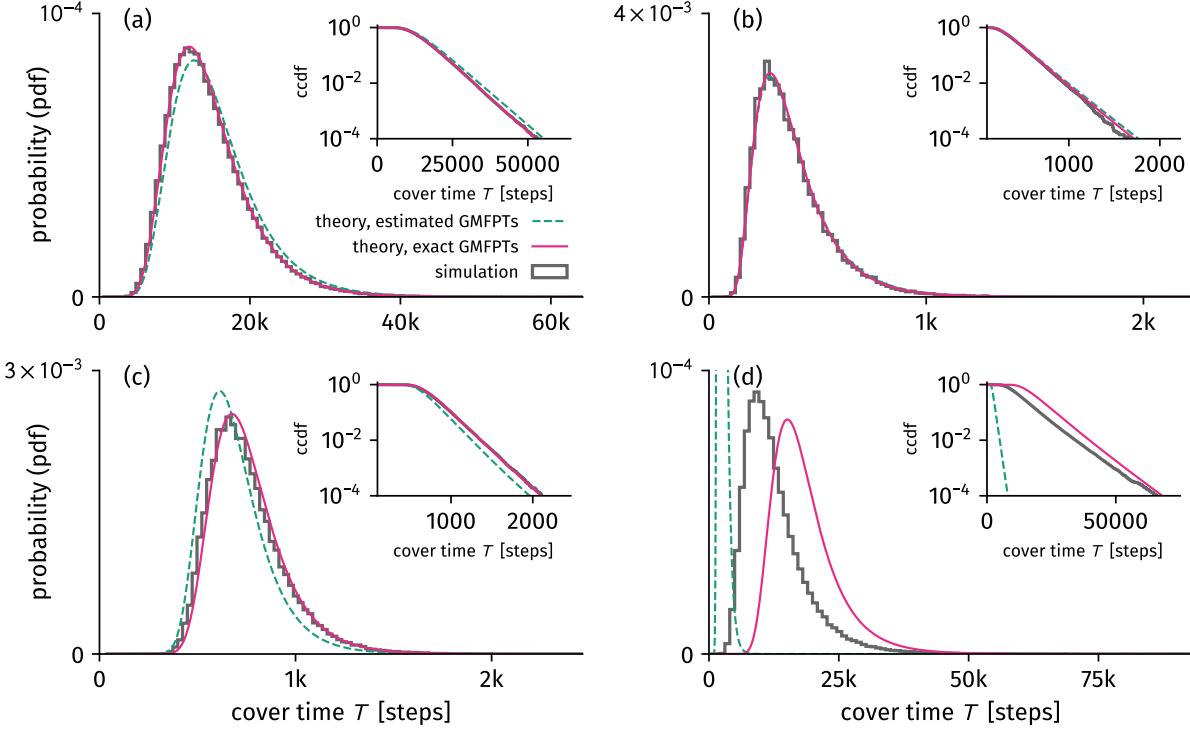


Figure 6.4: Cover time probability distribution functions (pdfs) and complementary cumulative distribution functions (ccdfs, insets) for four example networks (a) C. Elegans, (b) Intra-org. - Consulting (info), (c) College football and (d) Berlin's subway system from 2009. Dashed lines show the distributions from estimated arrival rates as per Eq. (6.8) and solid pink lines show the distributions from exact arrival rates (see Eq. (6.9)).

1/2. Using the first derivative $\bar{P}'_u(T)$, the error Φ emerging from the trapezoidal rule can be asymptotically estimated to be

$$|\Phi| = \frac{\Delta T^2}{12} |\bar{P}'_u(T_{\max}) - \bar{P}'_u(0)|$$

for $T_{\max} \rightarrow \infty$ [4]. With

$$\bar{P}'_u(T) = - \sum_{v \neq u} \beta_v \exp(-\beta_v T) \prod_{w \neq v \neq u} [1 - \exp(-\beta_w T)]$$

one has $\bar{P}'_u(0) = 0$. In another way, analogously to Eq. (6.7) one can find

$$\begin{aligned} \bar{P}_u(T) &= \sum_{\mathcal{S} \in \mathcal{P}^*(\mathcal{V} \setminus \{u\})} (-1)^{|\mathcal{S}|+1} \exp\left(-\sum_{v \in \mathcal{S}} \beta_v T\right) \\ |\bar{P}'_u(T)| &= \sum_{\mathcal{S} \in \mathcal{P}^*(\mathcal{V} \setminus \{u\})} (-1)^{|\mathcal{S}|+1} \left(\sum_{v \in \mathcal{S}} \beta_v\right) \exp\left(-\sum_{v \in \mathcal{S}} \beta_v T\right). \end{aligned}$$

For most nodes the decay rates are $\beta_v \lesssim N^{-1}$ with \lesssim meaning “lower or of similar order”. Then $\sum_{v \in \mathcal{S}} \beta_v \lesssim 1$ and hence $|\bar{P}'_u| \lesssim \bar{P}_u$ such that one can safely assume $|\bar{P}'_u(T_{\max})| \lesssim \bar{P}_u(T_{\max}) \leq 10^{-10}$ yielding absolute error

$$|\Phi| \lesssim 10^{-11},$$

which will be sufficiently small for most applications. For applications where this is not small enough, T_{\max} can be further increased until a satisfying error is reached.

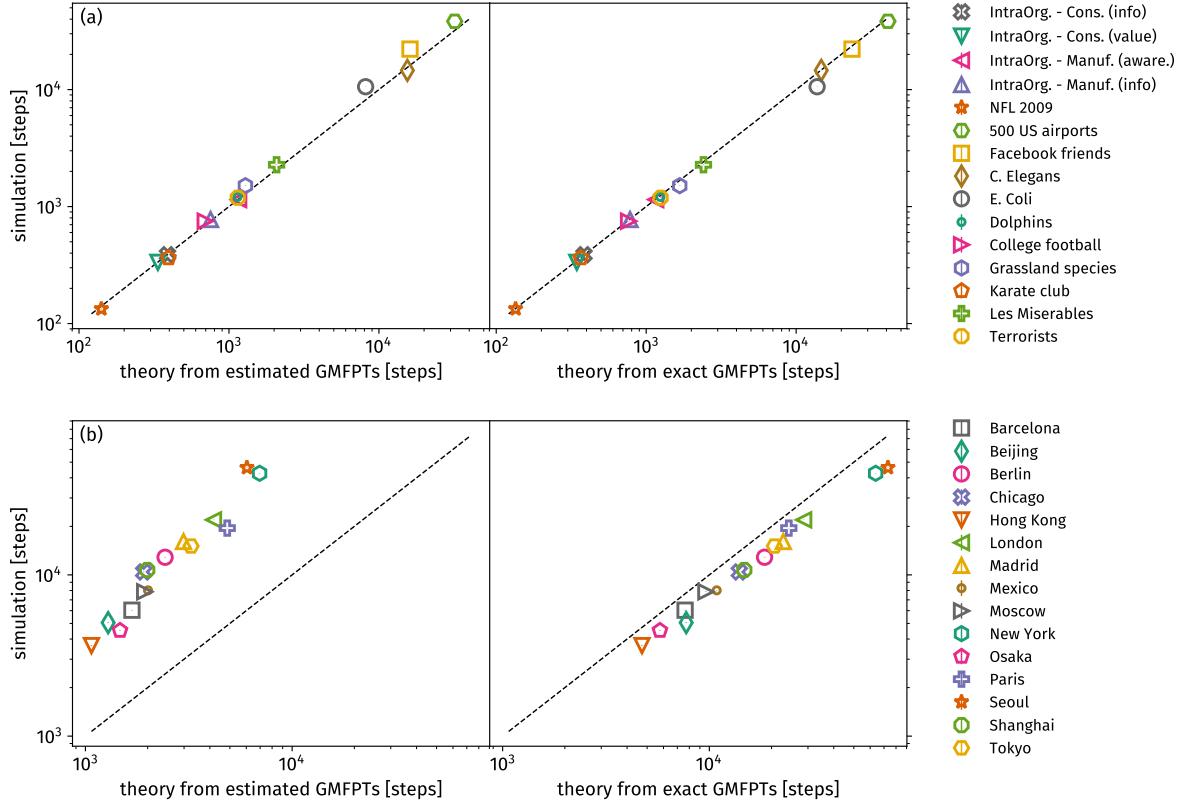


Figure 6.5: Mean cover times of simple discrete time random walks on the largest component of (a) various real-world networks (data sources and relative errors given in Tab. 6.1) and (b) various subway networks (data source and relative errors given in Tab. 6.2). The estimated cover times are compared to the measured cover times (from 50 simulations for each data point). The dashed lines represent the ideal case $\langle T^{\text{sim}} \rangle = \langle T^{\text{res/ex}} \rangle$. Theoretical results are computed from (left) estimated GMFPTs from the target nodes' degrees Eq. (6.8) and (right) exact GMFPTs computed from the unnormalized graph Laplacian's spectrum Eq. (6.9).

6.3.3 Systematic Error: Approximation of Mean Cover Time Integral

In the following, it is demonstrated that instead of solving integral Eq. (6.5), one can safely use Eq. (6.7). The total difference between both is

$$\begin{aligned} \Theta &= \int_0^\infty dT [1 - P(T)] - \\ &\quad - \int_0^\infty dT \left[1 - P(T) \frac{1}{N} \sum_{v=1}^N \frac{1}{1 - \exp(-\beta_v T)} \right] \\ &= \int_0^\infty dT P(T) [Q(T) - 1], \end{aligned} \quad (6.13)$$

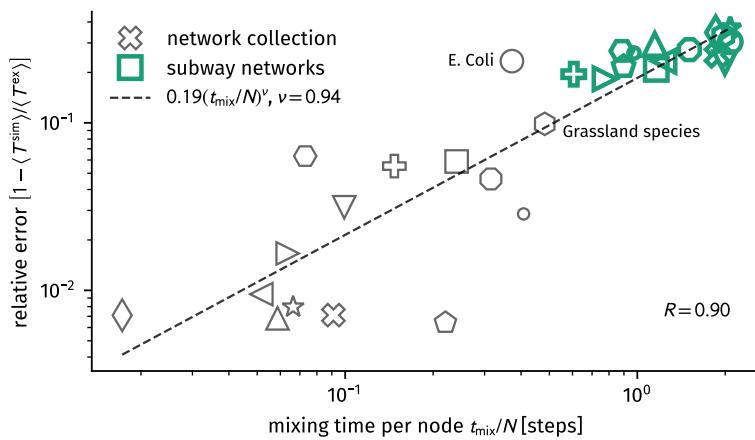
defining $Q(T) = \frac{1}{N} \sum_{v=1}^N (1 - \exp(-\beta_v T))^{-1}$. Note that the cover time cdf $P(T)$ is given by Eq. (6.6), s.t. both

$$\begin{aligned} \lim_{T \rightarrow 0} P(T) &= 0 \\ \lim_{T \rightarrow 0} P(T)Q(T) &= 0 \end{aligned}$$

and

$$\begin{aligned} \lim_{T \rightarrow \infty} P(T) &= 1 \\ \lim_{T \rightarrow \infty} P(T)Q(T) &= 1, \end{aligned}$$

Figure 6.6: The relative error of the mean cover time heuristic is increasing with increasing mixing time per node t_{mix}/N , indicating that the heuristic produces higher deviations for networks with relatively high mixing times as per Eq. (3.6). Networks are marked with the same symbols as in Fig. 6.5.



meaning that for both integration limits, the integrand approaches zero. The following argumentation is based on the assumption that the distribution of decay rates is relatively homogeneous in the region of small rates, implying that there is a low number $1 < n \ll N$ of nodes $i \in \mathcal{V}_{\text{small}}$ with $n = |\mathcal{V}_{\text{small}}|$ that are of the same order as $\beta_{\min} = \min\{\beta_v : v \in \mathcal{V}\}$. This is a relatively safe assumption for most network models and real-world networks as in most cases there are more nodes with small degree (hence small decay rates) than nodes with high degree (hence high decay rates). Now suppose the integration approaches a time where $T \approx \beta_{\min}^{-1}$, implying that, while most terms $1 - \exp(-T\beta_{v \notin \mathcal{V}_{\text{small}}})$ are virtually equal to 1 there are still n terms $1 - \exp(-T\beta_i) < 1$, such that $P(T) \approx \prod_{i \in \mathcal{V}_{\text{small}}} (1 - \exp(-T\beta_i)) \ll 1$. Furthermore, there will already be a majority of terms $1 - \exp(-T\beta_v) \rightarrow 1$ which leads to $Q(T)$ approaching $Q(T) \xrightarrow{N \gg 1} 1$. Hence, one can safely assume that for a network with a larger number of nodes the integrand approaches zero at all times while the global mean cover time grows quickly and thus the relative error of Eq. (6.7) is approaching

$$\frac{\Theta}{\langle T \rangle} \xrightarrow{N \gg 1} 0.$$

In particular, one can calculate the error between the integrals for random networks with constant GMFPT $\tilde{\tau}$ for every node, which is given as

$$\begin{aligned} \tilde{\Theta} &= \tilde{\tau}_{N+1} [\gamma + \psi(N+1)] - \tilde{\tau}_N [\gamma + \psi(N)] \\ &\approx \tilde{\tau}_N \log \left(\frac{N}{N-1} \right) \approx \tilde{\tau}_N \frac{1}{N} \end{aligned}$$

where $\psi(N+1) = \log N + \mathcal{O}(N^{-1})$ and $\tilde{\tau}_N \sim \tilde{\tau}_{N+1}$ was assumed. Consequently, one finds the relative error to be approximately

$$\frac{\tilde{\Theta}}{\langle T \rangle} \approx \frac{1}{N\gamma + N \log(N-1)} \sim \frac{1}{N \log N}. \quad (6.14)$$

Even though this relation is derived for the special case of networks where every node has the same GMFPT, a numerical analysis of

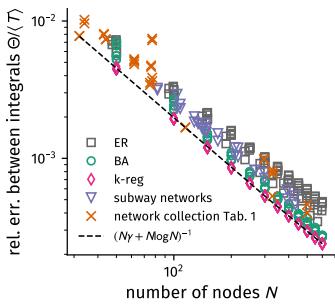


Figure 6.7: Relative error Eq. (6.13) between integrals Eq. (6.5) and Eq. (6.7) for all networks investigated in this chapter besides lattices and the modular hierarchical network models. For each network the relative error for both sets of rates, τ_v^{es} and τ_v^{ex} , is shown. For ER, BA and random k -regular networks, means over the largest components of 100 network realizations for $\langle k \rangle \in \{3, 5, 7, 9\}$ and $m \in \{3, 5, 7, 9\}$ were built, respectively. The measured relative errors are roughly following the scaling relation Eq. (6.14).

Eq. (6.13) reveals that this scaling relation holds approximately for all networks investigated in this chapter, as can be seen in Fig. (6.7).

6.4 Cover Time on Modular Hierarchical Small-World Models

The purpose of studying the cover time on the small-world models introduced in Ch. 4 in the following is two-fold.

First, it raises the opportunity to test the hypothesis that besides the pair-averaged first passage time, which is an upper bound for search times, the mean cover time is minimal between the highly clustered and the random regimes of the models. This would imply that an upper bound for exhaustive searches is optimal on some intermediate modular hierarchical structure.

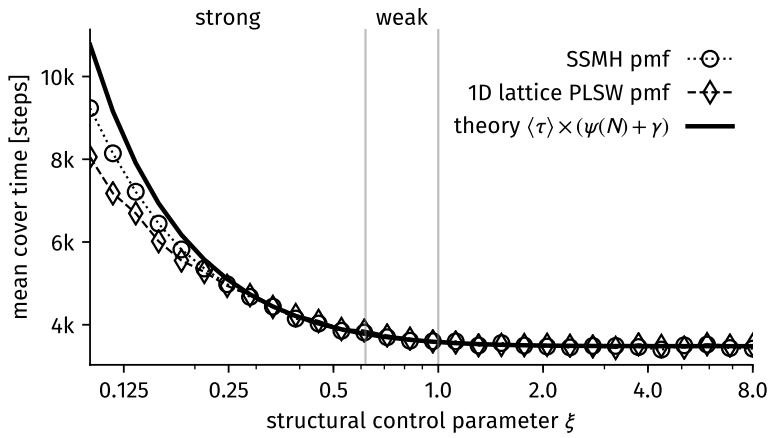
Second, it allows for a more detailed error investigation on when the heuristic begins to fail to predict the cover time accurately as those structures represent a continuous interpolation between the network structures investigated in Secs. 6.2 and 6.3.1.

6.4.1 Average Medium Approximation

As has been shown in Sec. 5.1.1, an average medium approximation (AMA), where a possible edge is replaced with the probability of this edge existing, may provide useful insights into the general mechanisms of random walks on modular hierarchical small-world networks as it allows for thorough analytic considerations. Indeed, one consequence of the AMA is that every node has the exact same GMFPT Eq. (5.6) which allows using Eq. (6.10) to estimate the mean cover time.

In order to test the validity of Eq. (6.10) for average medium self-similar small-world (SSMH) networks, random walk simulations have been performed on averaged SSMH networks and averaged 1D lattice power-law small-world (PLSW) networks; the same simulations as in Sec. 5.1.1. For $B = 8$, $L = 3$ (number of nodes $N = 8^3$) and $\langle k \rangle = 7$, a single walker was placed on each node of each average medium (SSMH and PLSW). Random walk simulations were performed as described in Sec. 3.1.1, where for each walker the simulation was stopped as soon as it visited each node at least once, marking the cover time for this walker, subsequently computing the mean to obtain the mean cover time. The results can be seen in Fig. 6.8. The heuristic Eq. (6.10) seems to predict the cover time reasonably well for moderate degrees of hierarchical clustering and starts to deviate more strongly from the simulation results for structural control parameters $\xi \lesssim 0.25$. This is not surprising as the heuristic evaluation of the mean cover time is based on approximating the first passage time to be distributed according to an exponential distribution. In Sec. 5.1.1 it was found that this approximation is reasonable for values $\xi \gtrsim 0.25$ but is only valid asymptotically for smaller values (see Fig. 5.3) which explains both the monotonically decreasing shape of

Figure 6.8: The mean cover time as estimated with an average medium approximation of self-similar modular hierarchical as well as 1D lattice PLSW networks as described in Sec. 5.1.1. Shown are simulation results on the averaged media as well as Eq. (6.10) with $\tilde{\tau}$ as given by Eq. (5.7).



the mean cover time function (which is determined by the shape of the global mean FPT Eq. (5.7)) and the deviations for small values of ξ .

The result of the AMA is hence that no optimal mean cover time exists between the highly regular structures and the random case.

6.4.2 Actual Network Realizations

Here, the mean cover time is analyzed for actual network samples of the self-similar modular hierarchical (SSMH), the modified Watts–Strogatz (mod. WS), the one-dimensional lattice power-law small world (1D latt. PLSW) and the 2-torus PLSW models. To this end, networks with parameters $B = 8$, $L = 3$ (corresponding to $N = 8^3$) and $\langle k \rangle = 8$ were sampled for varying values of their respective control parameters. The control parameters of the SSMH, the mod WS and the 1D latt. PLSW model were mapped using Eqs. (4.33) and (4.31). For each parameter combination, 740 independent network samples were drawn, subsequently their largest component was extracted to be used for analysis. For each largest component, 100 random walkers were placed on 100 random nodes. Random walk simulations as described in Sec. were started and run until each walker visited each node at least once. When a walker reached its last remaining non-visited node, this time step was saved as its cover time. The mean cover time of that network sample was then computed as the mean of those saved times. The theoretical mean cover time of that network sample was computed using Eq. (6.9). Subsequently, the respective means were built over the 740 obtained single-network means.

The results are shown in Fig. 6.9. Since Eq. (6.9) relies on the global mean FPT which is responsible for a minimum in the pair-averaged FPT, it is not surprising that the mean cover time reveals a minimum, as well (for the SSMH, mod. WS and 1D latt. PLSW models). The position of this optimum in exhaustive search time, lies, however, in the strong hierarchically clustered regime where average nodes are connected to more nodes of short-range distance

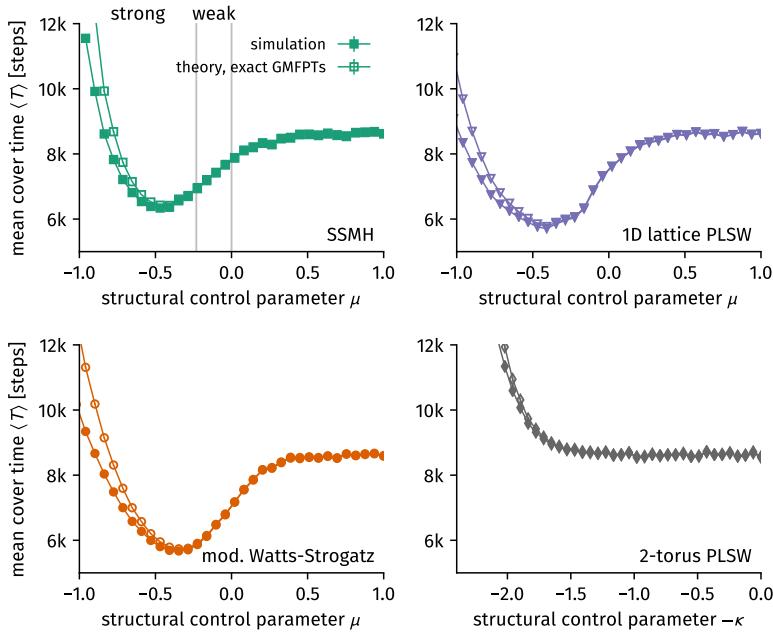


Figure 6.9: The mean cover time on actual network realizations of the (modular hierarchical) small-world models introduced in Ch. 4. Parameters $B = 8$, $L = 3$, and $\langle k \rangle = 8$ were used, markers denote averages over 740 independent network realizations. For each network realization, the cover time was measured from simulations as well as computed from Eq. (6.9). All network models based on discrete node positions display minimal mean cover time in the strong modular hierarchically clustered regime since those display increasing node degree variance with increasing long-range connection probability. The network model based on random node positions (2-torus power-law small-world model) does not as it has constant node degree variance over the full range of its structural control parameter.

than long-ranged nodes. This emergence of the minimum can be explained analogously to the argumentation given in Sec. 5.2.2. For high clustering (and low dimensionality), the cover time becomes large but decreases with increasing long-range connection probability and concurrent decreasing clustering. At the same time the node degree variance increases such that more and more nodes of small degree appear which take longer to visit and hence increase the cover time. This second effect does not contribute to the cover time if the degree distribution is constant when increasing long-range connection probability as is the case for the 2-torus PLSW model. Consequently, a minimum in cover time is not observed for network samples of this model.

As implied, the analyses of this section can also be used to obtain an improved intuition about the error produced by the heuristic evaluation of the cover time. The relative error between simulation results and Eq. (6.9) is shown in Fig. 6.10. As expected, the relative error systematically increases with decreasing structural control parameters as soon as values are reached which produce structures being associated with modular hierarchical clustering. For real-world networks there is no such thing as a measurable structural control parameter, and hence, the mixing time per node t_{mix}/N will be used as a proxy for the randomness of the networks. For all network samples used in the procedure described above, the mixing time was found using Eq. (3.6) and subsequently averaged. As can be seen in the top panel of Fig. 6.10, the functional dependence of the relative error does not correspond ideally to a power-law as with decreasing structural control parameter (i.e. increasing t_{mix}/N), the relative

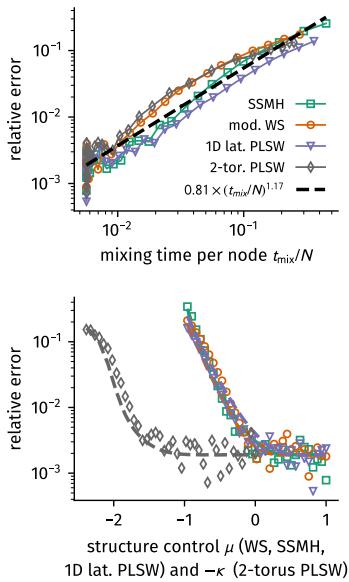


Figure 6.10: Relative errors between simulation results and Eq. (6.9) for the analysis displayed in Fig. 6.9. **(Top)** The relative error as a function of the mixing time per node as computed using Eq. (3.6) for each network sample and subsequent averaging. The dashed line represents the fit Eq. (6.15). **(Bottom)** The relative error as a function of the structural control parameters. The dashed line represents the fit Eq. (6.15) using the respective values of t_{mix}/N used in the top panel.

error seems to flatten off. Nevertheless, the power-law

$$\text{rel. err.} \approx 0.81 \times \left(\frac{t_{\text{mix}}}{N} \right)^{1.17} \quad (6.15)$$

seems to satisfactorily estimate the relative error of the heuristic method (see bottom panel of Fig. 6.10). As this estimation is greater than Eq. (6.12) for larger values of t_{mix}/N it might be more reliable to estimate the relative error using Eq. (6.15).

6.5 Summary and Discussion

A new heuristic method to find the mean cover time of random walks on arbitrary complex networks has been introduced and discussed. It was based on the assumption that arrival times follow an exponential distribution with mean given by the respective target node global mean first passage time. The heuristic was applied to several synthetic networks as well as real-world networks. It was found that the heuristic results become increasingly unreliable when their mixing time becomes increasingly large. For these cases the approximation of exponentially distributed arrival times is only valid asymptotically, the estimated arrival times are consequently wrongly estimated which in turn influences the mean cover time. Ways to estimate the relative error of the method were given, however.

The heuristic was subsequently applied to the modular hierarchical network models defined in Ch. 4. As for the pair-averaged first passage time of Ch. 5 minima in the mean cover time were found for all network models where node degree variance increases with increasing long-range connection probability. The remaining modular hierarchical network model where node degree variance is constant does not display a minimum in mean cover time.

Therefore one cannot state that in general, modular hierarchical networks provide an optimal structure for fast uninformed exhaustive searches.

7

Summary & Discussion

In this part of the thesis, the hypothesis was tested whether networks of modular hierarchical structure provide an optimal topology for fast diffusion and searches as modeled by random walks.

To this end several generative network models were devised which, using a single control parameter, interpolate between highly regular clustered structures and the Erdős–Rényi model while keeping the mean degree fixed. It was shown that these networks correspond to each other by (i) displaying the similarity of their network properties and (ii) comparing modular hierarchical network partitions an algorithm found by analyzing network samples of these models. All network models were found to generate structures which are interpretable to be modular hierarchical.

These models were used to evaluate the pair-averaged first passage time, an observable quantifying how fast randomly passed information can spread on these networks. It was found that those network models whose node degree variance increases as a function of their structural parameters do indeed show a minimum in pair-averaged first passage time. Another model which has constant degree variance, however, shows no minimum.

Similarly, the cover time which was evaluated using a newly found heuristical method as well as by using simulations, shows a minimum for those network models with varying degree variance but no minimum for the remaining model.

Hence it has to be concluded that in general, modular hierarchical networks do not possess an optimal structure for processes connected to random walks. This raises the question why those structures are then still so inherent to the real-world. An argumentation is given as follows.

A considerable number of networks which are said to be of modular hierarchical structure are in actuality networks whose nodes are embedded in some space such that node pairs are associated with distances and consequently connected with a probability following a power-law decreasing with distance $P(r) \propto r^{-\kappa}$. This connection probability can be interpreted to be inversely proportional to the cost of establishing a connection to this distance. Since each long-range connection considerably decreases the time it takes for information to spread, it is favorable to establish them. However, depending on the

nature of the system and the connections, the cost to establish such a long-distance edge can grow as a power-law $C(r) \propto r^k$. Assuming that edges are less likely to be built the more they cost, those networks would then resemble power-law small-world networks and misinterpreted to be of modular hierarchical structure.

In order to test this hypothesis, future research should focus on investigating real-world networks in detail for how they might have been generated and in what way the cost of establishing connections between node pairs can be determined.

It must also be stated that future research using the models defined in Ch. 4 to study their implications for dynamic systems should carefully investigate how their observables depend on the node degree variance since this might heavily influence their results.

Part III

Epidemics on Temporal Face-to-Face Contact Networks

Flockworks: Analyzing Temporal Face-to-Face Contact Networks With a Simple Model

In this part of the thesis it will be investigated how a temporally resolved description of physical contacts between humans influences the spread of diseases in the population as opposed to static networks. Since studying the spread of diseases on empirical temporal networks alone has left open questions or conflicting results, a temporal network model called *Flockwork* will be devised in this chapter. It will reflect several properties of empirical temporal face-to-face contact networks while still being simple enough to evaluate and understand a number of network properties analytically.

As discussed in Sec. 2.4, empirical face-to-face contact networks display the following basic properties.

1. At any time, the network consists of disconnected components, or groups, to which nodes connect or from which they disconnect.
2. Inter-contact life-times, contact life-times, and group life-times follow heavy-tailed distributions. The shape of the group lifetime distribution does not change with group size but the maximal group-life time increases with increasing group size.
3. Network properties follow a circadian activity rhythm.
4. The aggregated networks possess complex structure.

The emergence of these network properties will be discussed while the model's accuracy is increased step by step. The model is can be viewed as a variant of an existing model which postulates nodes being in groups at all times and heterogeneous interaction times [85]. The Flockwork model introduced here, however, will be much simpler and of Poissonian nature. It introduces two node behaviors, each associated with a rate. The first behavior is given by a node actively establishing new connections with another node and its neighborhood with *active reconnection rate* α . Second, nodes can actively decide to stay alone, with rate β , called the *active disconnection rate*. Following these two events, it will be shown that the system always reaches a state in which the network consists of disconnected components which are completely connected within. The size distribution

of this state will be stable. Due to the Poissonian nature of the model, all temporal distributions will be shown to not carry a heavy tail.

In an attempt to increase the realism of the model, temporally varying event rates will be introduced which will lead to a more accurate description of the structural distributions but stays homogeneous in the temporal observables. It will further be shown that that human temporal contact networks might be interpreted to follow a trajectory of equilibrium state in two-dimensional state space.

8.1 The Poissonian Flockwork Model

Here, the basic version of the Flockwork model will be defined. Subsequently, it will be shown that this model always produces an equilibrium state where the network consists of several disconnected components, or groups, which are all complete networks. The size distribution of these groups is computed, along with asymptotic formulas for its first and second moment, the mean number of components, the degree distribution, the asymptotic mean degree and the group life-times.

8.1.1 Model Definition

Given any network at time t_0 , each node can undergo two events.

1. With active reconnection rate α , the node cuts all its links, connects to another node v and connects to all of v 's neighbors $v' \in \text{Nei}(v)$.
2. With active disconnection rate β , the node cuts all its links and stays alone.

These rates are meant to reflect basic human behavior in the sense that people may decide not to have any contacts, or decide to establish a new face-to-face contact, which makes it likely to be in contact with all the other people, who are already in close proximity to the new contact.

Since in both events all links are cut first, these rates can be mapped to a total event rate $\gamma = \alpha + \beta$. Then the model is described as follows. Each node cuts all its links with rate γ . With probability P it reconnects to any node v and all of v 's neighbors. Both descriptions are equivalent. In this section, the γ - P -model will be used because the time scale can be fixed using γ and subsequently the *reconnection probability* P is used as a single parameter to control the structure. For later descriptions, it is useful to provide mapping equations of the control parameters, which are given as

$$\begin{aligned}\gamma &= \alpha + \beta, \\ P &= \alpha / (\alpha + \beta), \\ \alpha &= \gamma P, \\ \beta &= \gamma(1 - P).\end{aligned}$$

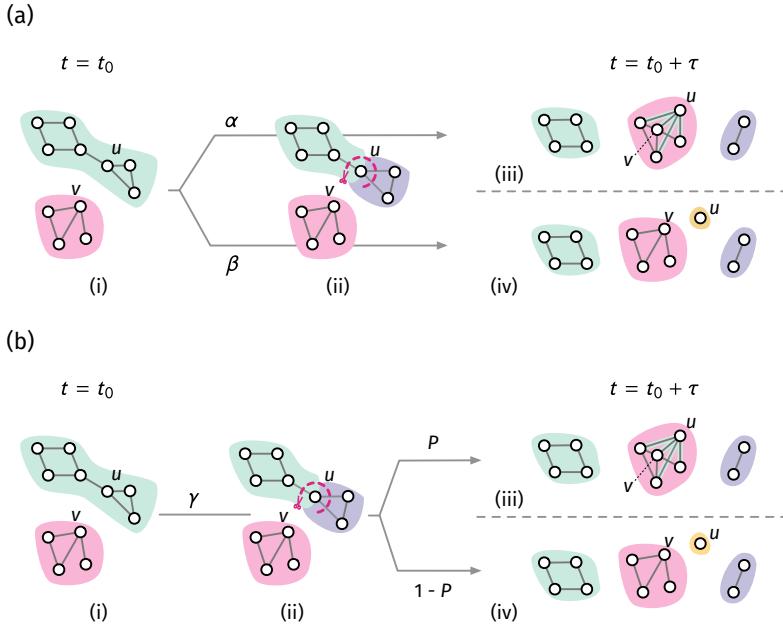


Figure 8.1: The possible events of a Flockwork model for any network state at $t = t_0$. (a) In the α - β -description of the model, each node undergoes a reconnection event with rate α and a disconnection event with rate β . (a.ii) In both cases, the link first cuts all its connections. (a.iii) For an α event it subsequently connects to a random node v and to all of v 's neighbors. (a.iv) For a β -event it stays alone. The updated time is then $t = t_0 + \tau$ where τ follows an exponential distribution $\tau \sim \mathcal{E}(\alpha + \beta)$. (b) in the γ - P -description, each node undergoes a (b.ii) link cutting event with rate γ . Afterwards, it (b.iii) reconnects with probability P to a random node v and all of v 's neighbors. (b.iv) With probability $1 - P$ it stays alone. The updated time is then $t = t_0 + \tau$ where τ follows an exponential distribution $\tau \sim \mathcal{E}(\gamma)$. Both (a) and (b) are equivalent.

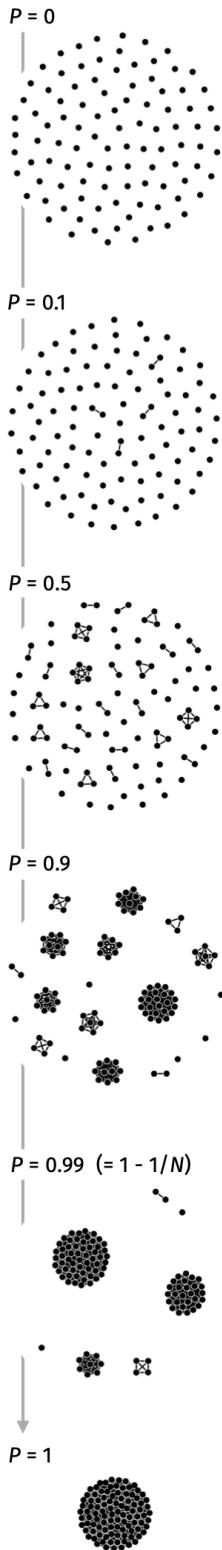
In principle, the number of nodes N in combination with any two of the parameters above define exactly one Flockwork configuration \mathcal{F} . An illustration of the two descriptions is given in Fig. 8.1.

A simulation method is based on Gillespie's stochastic simulation algorithm as described in Sec. XX. Consider any network at time $t = t_0$. An inter-event time τ is sampled from the exponential distribution $\tau \sim \mathcal{E}(\gamma)$. Afterwards, a node u is picked uniform at random and all edges connecting to node u are deleted from the network. Subsequently, the following happens with probability P . A second node v is chosen uniform at random from the remaining $N - 1$ nodes. Then, edges between u and v as well as u and all $v' \in \text{Nei}(v)$ are added to the network. The time is then updated as $t = t_0 + \tau$.

In the following, the structural and statistical consequences of this process are investigated and compared to the observations made in empirical temporal contact networks, beginning with the group-size distribution.

8.1.2 Group Sizes and Equilibrium State

In this section it will first be shown that given a collection of fully connected groups as an initial state, the Flockwork process will approach an equilibrium configuration where the network consists of disconnected components, each of which is fully connected. The group-size distribution is shown to be constant for constant reconnection probability. The discussion is subsequently extended to an arbitrary initial condition (initial network) which is shown to approach the same equilibrium configuration. The group-size distribution, mean group size and mean number of components will be computed analytically. An algorithm is given to construct equilibrium configurations without the need of simulating until an equilib-



rium state is reached.

Temporal Evolution

In order to obtain insight in the temporal evolution of the Flockwork process, one proceeds as follows. Consider an empty network of N nodes as an initial condition. From the model description it becomes clear that any event a node undergoes will lead to the network consisting of fully connected groups: The single nodes will connect to build isolated pairs, then isolated triangles, then isolated fully connected groups of four and so forth. Whenever a node disconnects, it will stay alone, representing a complete group of $N = 1$ nodes, whereas its former group will remain a completely connected component, albeit consisting of one node less than before. It therefore suffices to investigate the group-size distribution of the Flockwork model to obtain a complete picture of the networks the Flockwork model produces over time.

Denote with n_m the number of fully connected groups of size $1 \leq m \leq N$ such that the N -sized column vector \mathbf{n} contains the complete information of the network state. Note that at any time $n_0 = n_{N+1} = 0$, which will be used as boundary conditions below and hence these are not part of the state vector n_m . Furthermore, at any time, the total number of nodes N is conserved and as such, the normalization condition

$$\sum_{m=1}^N mn_m = N \quad (8.1)$$

has to be met. Given that the network is initially in a state which is fully defined by the vector \mathbf{n} , a derivation which considers every event that may happen and is shown in App. B.1, reveals that changes in group sizes by the events of the Flockwork dynamic process are governed by the following ordinary differential equations (ODEs):

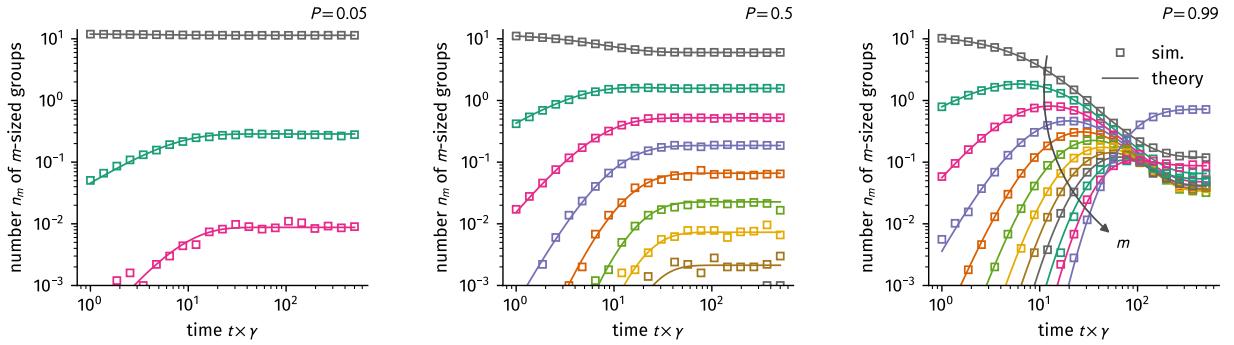
$$\begin{aligned} \gamma^{-1} \partial_t n_1 &= -\frac{2P}{N} n_1 + (1-P) \frac{N-n_1}{N} + \\ &\quad + 2\frac{1-P}{N} n_2 + 2\frac{P}{N} \frac{N-2}{N-1} n_2, \end{aligned} \quad (8.2a)$$

$$\begin{aligned} \gamma^{-1} \partial_t n_{m>1} &= \frac{P}{N(N-1)} \left[-2mn_m(N-m) + \right. \\ &\quad + n_{m-1} \left(N(m-1) - (m-1)^2 \right) + \\ &\quad \left. + n_{m+1} \left(N(m+1) - (m+1)^2 \right) \right] - \\ &\quad - (1-P) \frac{m}{N} n_m + (1-P) \frac{m+1}{N} n_{m+1}. \end{aligned} \quad (8.2b)$$

Figure 8.2: Flockwork equilibrium configurations for increasing reconnection probability P (for the networks shown here, $N = 100$ was used).

¹ meaning that no group size is preferred for any of the two events and hence the probability for each group size count to be affected is equal

Here, the time scale is fixed by the event rate per node γ and therefore the equations describe a single event. They can be interpreted as follows. Since disconnections affect each group size equally,¹ the change equations are (i) linear and (ii) imply that groups of size m can either be reduced by one node or increased by one node. Hence,



the ODE for the number of groups of size m only depends on its own group count (groups of size m), the number of groups of size $m - 1$, and the number of groups of size $m + 1$. The count of groups of size $m = 1$ can increase by considering pure disconnection events where one of the $(N - n_1)$ nodes in groups of size $m > 1$ has a disconnection event and stays alone. The validity of Eqs. (8.2) was tested by comparing their solutions to the expected number of group sizes obtained through simulations, yielding the result that the simulations are in agreement with the theory (see Fig. 8.3).

Since there is no dependence on states other than the current configuration, the Flockwork process is a Markov process. Due to the equations being linear one may build the expectation value of these equations of change and so n_m will be considered to denote the expected number of groups $\langle n_m \rangle$ in the following.

Equilibrium State

For every initial condition describable with the group count vector $\mathbf{n}(0)$, this linear process has a single stationary state \mathbf{n}^* which only depends on the number of nodes N and the reconnection probability P . It is derived in App. B and given as

$$n_1^* = N(1 - P) \quad (8.3a)$$

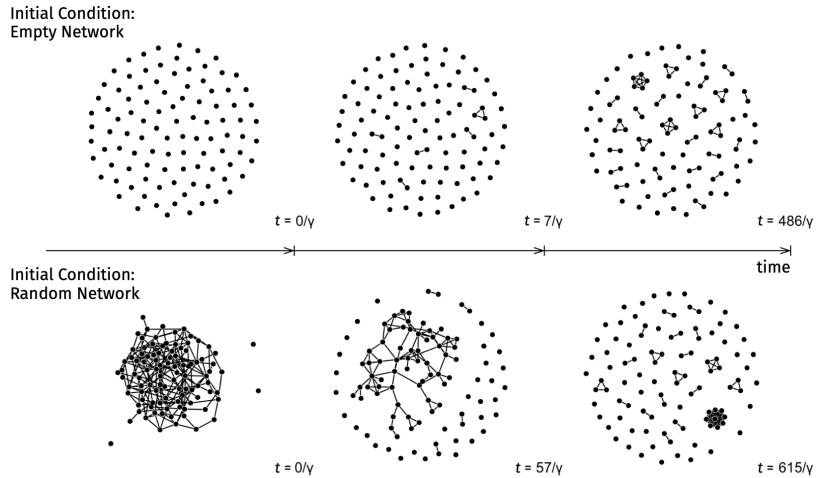
$$n_{1 < m < N}^* = \frac{1}{m} \frac{N!}{(N-m)!} \frac{1-P}{\prod_{j=1}^{m-1} [1 - Pj/(N-1)]} \left(\frac{P}{N-1} \right)^{m-1} \quad (8.3b)$$

$$n_N^* = (N-2)! \frac{P}{\prod_{j=1}^{N-2} [1 - Pj/(N-1)]} \left(\frac{P}{N-1} \right)^{N-1}. \quad (8.3c)$$

In App. B this result is further extended to be valid for *any* initial network. This can be roughly motivated as follows. Any non-fully connected group of size m can transition to produce both fully connected groups and non-fully connected groups through the Flockwork dis- and reconnection events (this includes any connected network of size N). However, no fully connected group of size m can produce any non-fully connected group. Hence, after a sufficient amount of time, any initial condition will transition to a network composed of only completely connected groups. Then, Eqs. (B.3) are applicable and

Figure 8.3: The temporal evolution of the expected number n_m of m -sized groups in the Flockwork model for $N = 12$ nodes and reconnection probability (left) $P = 0.05$, (center) $P = 0.5$, and (right) $P = 0.99$. The initial state was an empty network (i.e. $\mathbf{n}(0) = (N, 0, \dots, 0)^T$) and the event rate was chosen as $\gamma = 1$. For each value of the reconnection probability, 10^4 simulations have been performed using the events described in Sec. 8.1.1. Between each event, an inter-event time τ was sampled from the exponential distribution $\tau \sim \mathcal{E}(\gamma)$. Subsequently the group-size distribution n_m was measured which stayed constant until the next event. Markers represent means over the respective number n_m at definite time points in between $t = \gamma^{-1}$ and $t = 500\gamma^{-1}$. Furthermore, Eqs. (8.2) were integrated using a Runge-Kutta solver of order 4 with adaptive stepsize control (Dormand-Prince method). The results, represented by solid lines, are in agreement with the simulations.

Figure 8.4: Configuration examples for the temporal evolution of a Flockwork model of $N = 100$ nodes and reconnection probability $P = 0.5$ with different initial conditions. (Top) The initial state is the empty network. Small groups start to build until a configuration from the equilibrium is reached. (Bottom) The initial state is an Erdős-Rényi random network with $\langle k \rangle = 4$. Following the Flockwork process, the network fragments into smaller groups until a configuration from the equilibrium is reached.



hence the equilibrium state Eq. (8.3) is approached. An illustration of this effect is shown for exemplary simulations of $N = 100$ and $P = 0.5$ with two different initial conditions in Fig. 8.4. In the following, the approached equilibrium state is analyzed.

Some example equilibrium configurations are shown in Fig. 8.2 for increasing P and various exemplary group-size distributions are displayed in Fig. 8.5. One recognizes that depending on the reconnection probability P , the number of edges present in the network changes. With $P = 0$, the equilibrium state always corresponds to the empty network, whereas for $P = 1$, the equilibrium state will always be a completely connected network (which is both confirmed by considering these values of the reconnection probability in Eqs. (8.3) which yields $n_1^*(P = 0) = N$, $n_{m>1}^*(P = 0) = 0$, $n_N^*(P = 1) = 1$, and $n_{m<N}^*(P = 1) = 0$). Hence, one may interpret the Flockwork model to be able to interpolate between those two well-known network models. It is also important to note that both of those limit states are static. Therefore even though the model is dynamic, its limiting cases are static networks. For increasing values of the reconnection probability $0 < P < 1$, the equilibrium group-size distribution changes. It corresponds to a peaked distribution at $m = 1$ for $P = 0$, becoming broader but decaying quickly with increasing group size for smaller values of P until it reaches the distribution

$$n_m^*(P = 1 - 1/N) = \frac{1}{m}$$

at $P = 1 - 1/N$, as is confirmed by using this value of P in Eqs. (8.3). This will be of further interest below, as empirical networks seem to display broad group-size distributions which might be interpreted as power-laws. Further increasing P , the expected peak at $m = N$ develops.

In order to quantitatively describe the group-size distribution with a single number, the first moment of the group-size distribution is computed below. The mean group size, denoted as $\langle g \rangle \equiv \langle |g(u)| \rangle$ can be evaluated after the group-size distribution has been properly

normalized as

$$\langle g \rangle = \frac{\sum_{m=1}^N mn_m^*}{\sum_{m=1}^N n_m^*} = \frac{\sum_{m=1}^N mn_m^*}{\langle c \rangle} = \frac{N}{\langle c \rangle},$$

where the mean number of disconnected components is given as

$$\langle c \rangle = \sum_{m=1}^N n_m^*.$$

As shown in App. B.1, an exact formula for this quantity can be derived. However, an asymptotic expression for a large number of N shall suffice here, which is conjectured to be

$$\lim_{N \rightarrow \infty} \frac{\langle c \rangle}{N} = -\frac{1-P}{P} \log(1-P), \quad (8.4)$$

and hence the mean group size would be given as

$$\lim_{N \rightarrow \infty} \langle g \rangle = -\frac{P}{(1-P) \log(1-P)}. \quad (8.5)$$

One may further find the second moment of the group-size distribution to follow

$$\lim_{N \rightarrow \infty} \langle g^2 \rangle = \frac{\langle g \rangle}{1-P}, \quad (8.6)$$

which will be motivated when evaluating the degree distribution in Sec. 8.1.4. These expressions and their corresponding exact values are displayed in Fig. 8.6, showing that the asymptotics seem to be valid for $P < 1 - 1/N$. These asymptotic results can be of importance when a parameter P has to be chosen according to given first and/or second moment of the group-size distribution.

After discussing the properties of the equilibrium state, an algorithm to construct equilibrium configurations is introduced below.

8.1.3 Equilibrium Configuration Sampling Algorithm

In order to accurately simulate dynamic systems on a Flockwork model, e.g. the spread of diseases, it is necessary to initiate such a simulation in an already equilibrated system, however, depending on the initial state and the number of nodes, it can take a considerable number of events until the system is equilibrated. Instead, an algorithm to sample equilibrium configurations without having to simulate the Flockwork events is proposed in the following. Note that in contrast to the sections above, a distinction in notation will be made between the discrete observable number n_m of m -sized groups and its corresponding mean $\langle n_m \rangle$ as well as the theoretical mean n_m^* as given by Eqs. (8.3).

The algorithm will work by sampling discrete numbers n_m of groups which meet the normalization condition $\sum_{m=1}^N mn_m = N$ such that the average $\langle n_m \rangle$ of the samples will approach Eqs. (8.3). To this end it is first necessary to find the actual discrete distributions of n_m for varying reconnection probability P . Following the simulations shown in Fig. 8.3, the equilibrium distribution of n_m

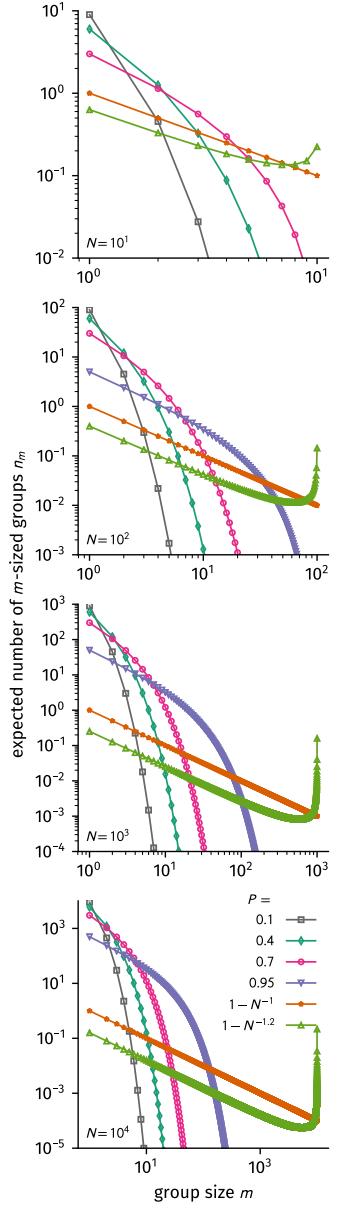
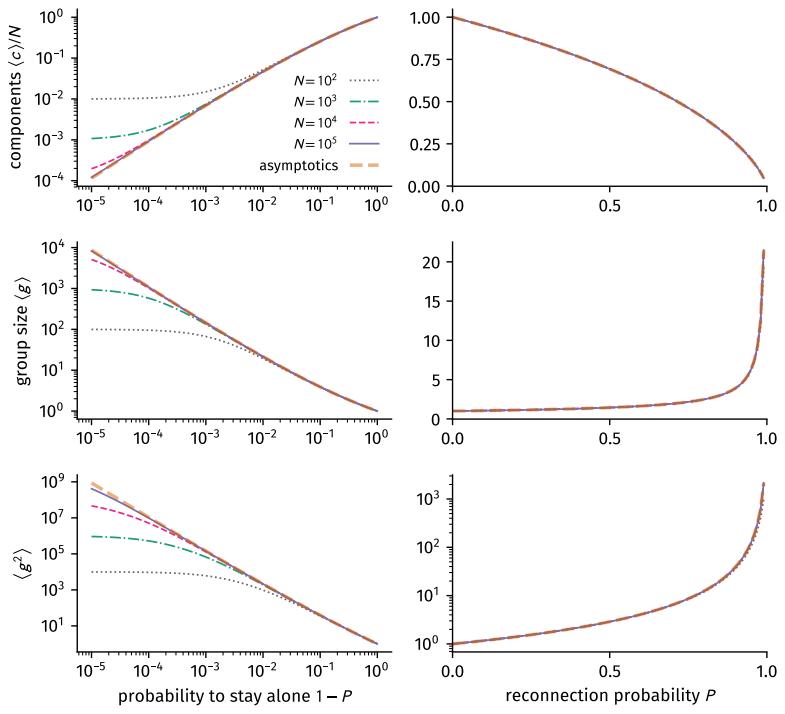


Figure 8.5: Flockwork equilibrium group-size distributions as computed from Eqs. (8.3). Note that the group-size distribution becomes a power-law $n_m^* = m^{-1}$ for $P = 1 - 1/N$.

Figure 8.6: Mean group size, mean number of components, and second moment of the equilibrium group-size distribution, all computed from Eqs. (8.3) and compared to the asymptotic Eqs. (8.4-8.6).



was measured over 10^4 independent simulations for a small number of nodes $N = 12$ and three values of reconnection probability $P \in \{0.05, 0.5, 0.99\}$. The results are shown in Fig. 8.7. One notices that due to the normalization condition, the number of single nodes $n_1 = N - 1$ cannot be realized (because then only a single node would be left, which, by definition, would increase the number of single nodes to $n_1 = N$). For small P , this pattern is repeated in the way that a triangle is highly unlikely, hence $n_1 = N - 3$ has low probability. However, the probability of two pairs occurring is higher, and as such the probability of $n_1 = N - 4$ is higher again. This behavior is difficult to replicate with single model distributions for each m since this would require a multivariate generalized Dirichlet distribution. In the following, the distributions will be approximated by independent discrete distribution.

First, consider a Poisson distribution $n_m \sim \mathcal{P}(n_m^*)$. As one can observe in Fig. 8.7, the Poisson distributions seem to approximate the actual distributions satisfactorily for most cases. However, in the extreme cases, i.e. $P = 0.05$ and $m = 1$, or $P = 0.99$ and $m = N$, the Poisson distribution will substantially differ from the real distribution, because the number of occurrences of m -sized groups is naturally upper bounded by $n_m^{\max} = N \text{ div } m$ which the Poisson distribution does not satisfy. Instead, one might consider a Binomial distribution $n_m \sim \mathcal{B}(n_m^{\max}, n_m^*/n_m^{\max})$, which seems to provide a better approximation of the real distributions in the extreme values of P , albeit ignoring the differences imposed by the normalization condition $\sum_{m=1}^N m n_m = N$ as discussed above. For the algorithm, both distributions will be tested.

```

1:  $\ell \leftarrow N$       ( $\ell$  is the number of nodes left to distribute)
2:  $V \leftarrow$  random permutation of  $\{u : (u \in \mathbb{N}) \wedge (1 \leq u \leq N)\}$ 
3:  $n_m \leftarrow 0 \quad \forall 1 \leq m \leq N$ 
4:  $n_m^* \equiv n_m^*(N, P)$ 
5: while  $\ell > 0$  do
6:   if shuffle then
7:      $M \leftarrow$  permutation of  $(\{m : (m \in \mathbb{N}) \wedge (1 \leq m \leq \ell)\})$ 
8:   else
9:      $M \leftarrow$  decreasingly sorted  $(\{m : (m \in \mathbb{N}) \wedge (1 \leq m \leq \ell)\})$ 
10:  end if
11:  for  $m \in M$  do
12:    if  $n_m^*(N, P) > 0$  and  $\ell \geq m$  then
13:      draw  $n_m^{\text{new}} \sim \mathcal{D}(N \text{ div } m, n_m^*)$ 
14:      if  $n_m^{\text{new}} > n_m$  then
15:         $\Delta n_m \leftarrow n_m^{\text{new}} - n_m$ 
16:        if  $m \times \Delta n_m > \ell$  then  $\Delta n_m \leftarrow \ell \text{ div } m$ 
17:        end if
18:      else if  $\ell = 1$  and  $m = 1$  then
19:         $\Delta n_m \leftarrow 1$ 
20:      else
21:         $\Delta n_m \leftarrow 0$ 
22:      end if
23:    end if
24:    for  $i \in 1, \Delta n_m$  do
25:      Add  $m$ -sized group of node indices  $V_{\ell-m+1} \leq u \leq V_\ell$ 
26:       $n_m \leftarrow n_m + 1$ 
27:       $\ell \leftarrow \ell - m$ 
28:    end for
29:  end for
30: end while

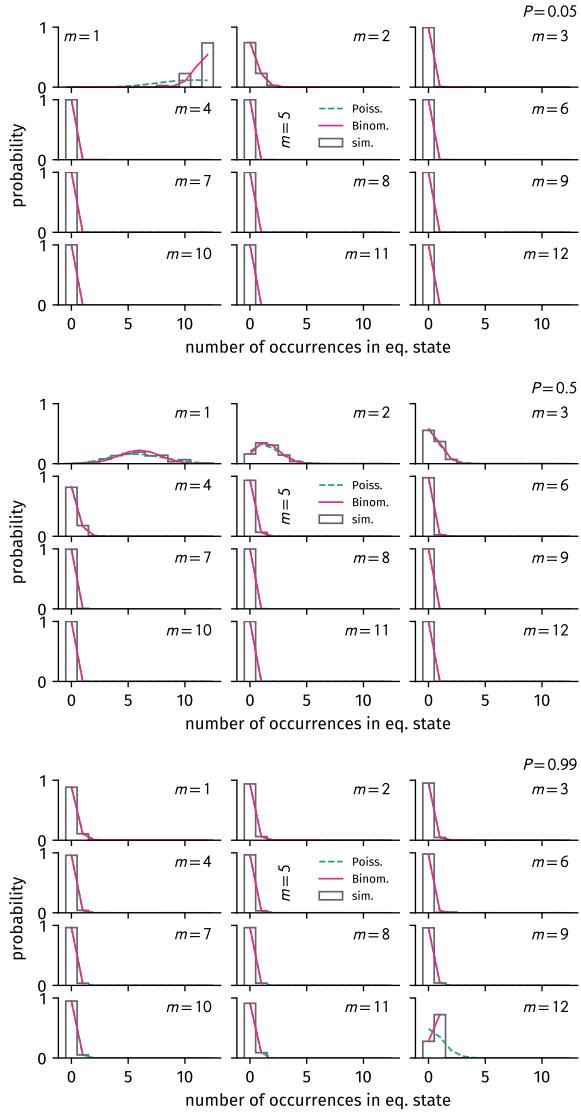
```

Algorithm 1: Sampling approximate equilibrium configurations for a Flockwork model. Note that here, $\mathcal{D}(N \text{ div } m, n_m^*)$ is a placeholder for either of the two distributions $\mathcal{B}(N \text{ div } m, n_m^*/[N \text{ div } m])$ or $\mathcal{P}(n_m^*)$. The algorithm can be adjusted to iterate through group sizes in a randomly permuted or decreasingly sorted way with the boolean variable ‘shuffle’.

The algorithm aims at drawing configurations n_m which satisfy $\sum_{m=1}^N mn_m = N$ as well as $\langle n_m \rangle = n_m^*$. This will be pursued as follows. Consider that initially there are N nodes available for distribution to groups and $n_m = 0$ groups of size m . Going through a permutation of all group sizes $1 \leq m \leq N$, for each m draw a number n_m^{new} from a distribution approximating the real distribution with mean n_m^* (i.e. either the binomial or the Poisson distribution) and add as many groups of size m which are still possible to add considering the number of nodes which are left. Update the new number n_m of m -sized groups accordingly. If, after all m have been visited, there are still nodes left, repeat this procedure. Now, only add Δn_m groups of size m where $\Delta n_m = \max\{0, n_m^{\text{new}} - n_m\}$. Repeat this process until all nodes are distributed to groups. This algorithm can be formalized as shown in Alg. 1.

A numerical analysis shows that the algorithm accurately constructs equilibrium configurations which are consistent with Eqs. (8.3) for a large number N of nodes and for $P < 1 - 1/N$ (cf. Fig. 8.8).

Figure 8.7: The distribution of numbers n_m of m -sized groups in the equilibrium state for increasing reconnection probability P with (top) $P = 0.05$, (middle) $P = 0.5$, and (bottom) $P = 0.99$. Instead of analyzing the expected number of m -sized groups, here the full distributions are shown, as collected from the 10^4 final states of the simulations shown in Fig. 8.3. Additionally shown are Poisson and binomial distributions with the mean n_m^* from Eq. (8.3). For the binomial distributions, the number of trials to draw an m -sized group is given as $N \text{ div } m$.



Both the binomial as well as the Poisson distribution seem to produce similar results. For a smaller number of nodes N , the decreasingly sorted way of group-size iteration yields more accurate configurations as otherwise smaller group sizes are increasingly selected. This changes with increasing number of nodes where the decreasingly sorted way of group-size iteration causes a bias towards smaller group sizes. Hence, the shuffled way of group-size iteration is to be preferred. Furthermore, the sampled configurations become increasingly inaccurate with increasing reconnection probability P . For those values, a strong bias towards smaller group sizes exists while the shape of the distribution is nevertheless roughly reproduced.

When using the algorithm to sample equilibrium configurations it is therefore recommended to first scan all variants of the algorithm to see which of them best reproduce the true equilibrium group-size distributions for the chosen values of N and P . For large values of P one might improve the accuracy by sampling a configuration using Alg. 1 and then letting the system equilibrate to a true equilibrium

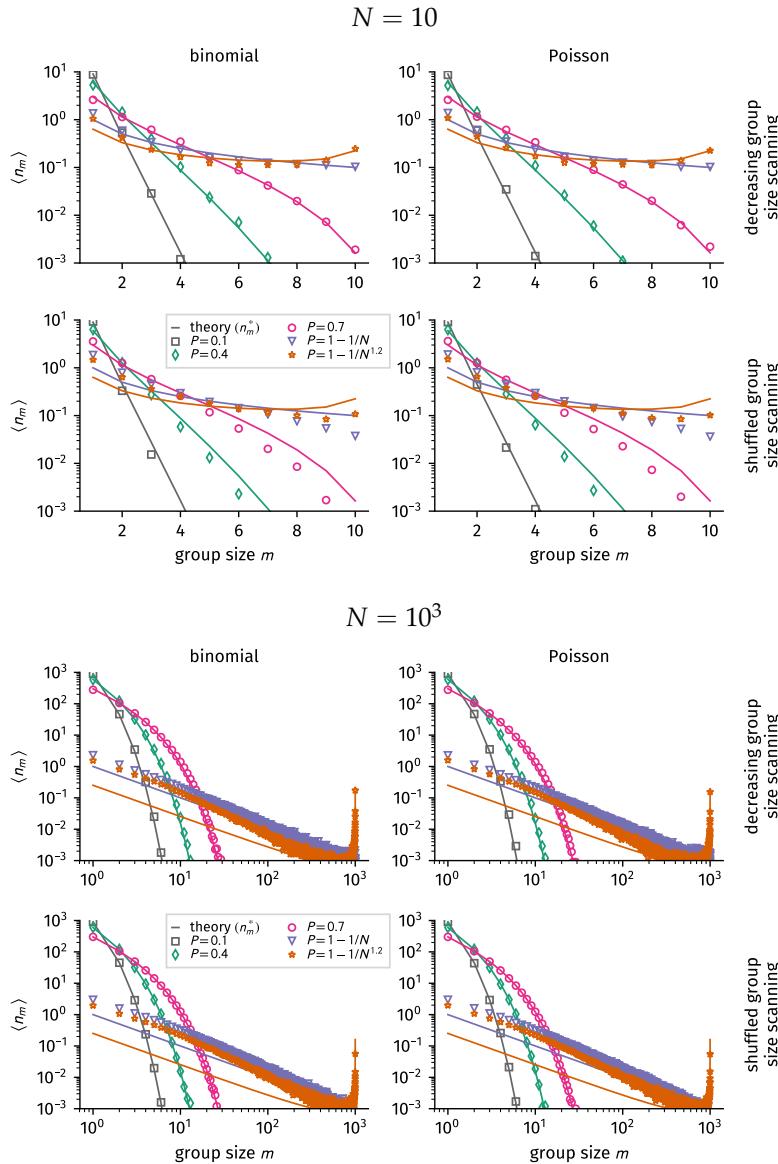


Figure 8.8: Group size distribution as a mean over 10^4 independently constructed equilibrium configurations using Alg. 1 for varying values of P , (top) $N = 10$, (bottom) $N = 10^3$ nodes and different algorithmic details. For each panel these details are (top row) iterating over a decreasingly sorted list M of group sizes m , (bottom row) iterating over a shuffled list M of group sizes m , (left column) drawing the number n_m of m -sized groups from a binomial distribution for each m , and (right column) drawing the number n_m of m -sized groups from a Poisson distribution for each m . The curves obtained from the sampling algorithm are denoted by markers whereas solid lines represent Eqs. (8.3).

configuration using the Flockwork procedure. The equilibration time for this procedure can be found by simulating a small number of configurations and measuring the number of events until the group-size distributions correspond.

Now, after intensively discussing the group sizes and their distribution, the node degrees will be discussed in the following.

8.1.4 Node Degrees

As discussed in Sec. 2.2.2, the node degree is an important simple quantity describing how connected a single node is in the network. The mean node degree $\langle k \rangle$ can influence both random walks and epidemic spreading; a heterogeneous degree distribution can cause the epidemic threshold to be shifted towards lower infection rates and hence be responsible for a more susceptible system. Hence, the

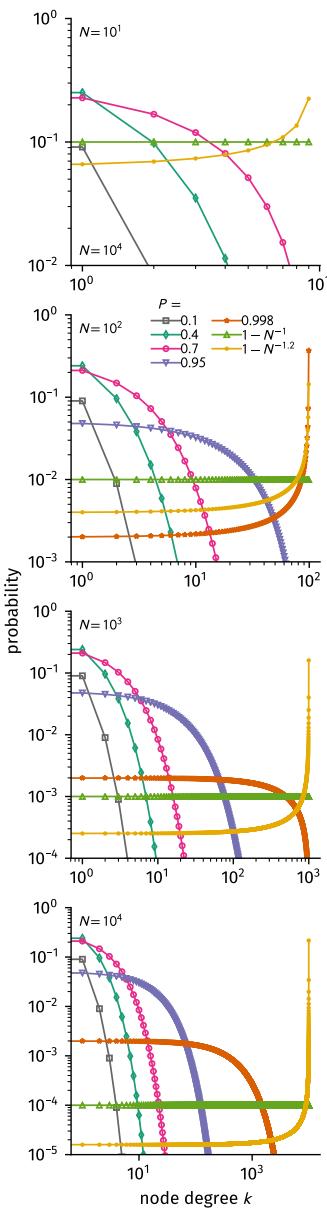


Figure 8.9: Flockwork equilibrium degree distributions as computed from Eq. (8.7) for increasing reconnection probability P and number of nodes N .

degree distribution and mean degree of the Flockwork model will be discussed in the following, which will be of use later on.

Since the Flockwork model always approaches a configuration where it consists of completely connected groups, each node in a group of size m will have degree $k = m - 1$. The probability of a randomly chosen node having degree k is hence equal to the probability of finding a randomly chosen node in a group of size m which is

$$P_k = P[|g(u)| = k + 1] = \frac{k+1}{N} n_{k+1}. \quad (8.7)$$

Hence, the mean degree is given as

$$\begin{aligned} \langle k \rangle &= \sum_{k=0}^{N-1} k P_k = \frac{1}{N} \sum_{k=0}^{N-1} k(k+1) n_{k+1} = \frac{1}{N} \sum_{m=1}^N (m-1)m n_m \\ &= \frac{1}{N} \sum_m m^2 n_m - 1 = \frac{1}{N} \left(\sum_m n_m \right) \frac{\sum_m m^2 n_m}{\sum_m n_m} - 1 = \frac{\langle e \rangle}{N} \langle g^2 \rangle - 1 \\ &= \frac{\langle g^2 \rangle}{\langle g \rangle} - 1. \end{aligned} \quad (8.8)$$

Furthermore, the mean degree can be computed considering how the number $\langle e \rangle(t)$ of edges present in the network change per event. On average, and considering $N \rightarrow \infty$, a γ -event will delete $\langle k \rangle(t)$ edges at time t as an average node cuts all its contacts. Furthermore, with probability P , new edges will be built by connecting to a node of mean degree $\langle k \rangle(t)$ and all its neighbors, thus building $\langle k \rangle(t) + 1$ new edges. Therefore, the temporal evolution of the present number of edges at time t can be approximated as

$$\langle e \rangle(t+1) - \langle e \rangle(t) = -\gamma \langle k \rangle(t) + \gamma P + \gamma P \langle k \rangle(t),$$

which yields the asymptotic equilibrium mean degree

$$\langle k \rangle = \frac{P}{1-P}. \quad (8.9)$$

Combining Eqs. (8.8) and (8.9) gives the asymptotic formula of the second moment of the group-size distribution Eq. (8.6). Exemplary degree distributions are displayed in Fig. 8.9 and the validity of the asymptotic mean degree for $P \lesssim 1 - 1/N$ is confirmed in Fig. 8.10.

8.1.5 Distributions of (Inter-)Contact and Group Life-Times

As in this part of the thesis temporal networks are the focus of attention, this section will discuss the basic model's implications for various time-scales, i.e. the inter-contact time, the contact life-time, and the group life-times for a group of size m .

The inter-contact time is the time a node will stay alone before establishing a new contact. The rate with which a lone node will establish a new connection in the Flockwork model is $2\gamma P = 2\alpha$ (accounting for both an event where the node acts actively as well as an event where it gets passively connected to). Hence, the inter-contact time is exponentially distributed as

$$p(\tau_{IC}) = 2\alpha \exp(-2\alpha \tau_{IC}). \quad (8.10)$$

The group life-times can be evaluated in a similar manner. Suppose a group is of size $m > 1$. Then, three events can lead to a change in group size.

1. One of the m nodes disconnects and stays alone, which happens with rate $m\gamma(1 - P)$.
2. One of the m nodes disconnects and reconnects to a node which is not part of the rest of this group, which happens with rate $m\gamma P(1 - (m - 1)/(N - 1))$. Note that it might be more intuitive to express the fraction of the connecting nodes as $(1 - (m - 1)/(N - 1)) = (N - m)/(N - 1)$.
3. One of the $N - m$ nodes reconnects to one of the m nodes, which happens with rate $(N - m)\gamma Pm/(N - 1)$.

The total rate of change for an m -sized group is thus given as

$$\lambda_m = m\gamma(1 - P) + 2\gamma Pm \frac{N - m}{N - 1}$$

and hence the life-time distribution of groups of size m follows as

$$p(\tau_m) = \lambda_m \exp(-\lambda_m \tau_m). \quad (8.11)$$

The contact decay rate can be estimated as follows. Suppose an edge between two nodes has been established. The nodes are in a group of size $m = k + 1$ and so they have degree k . Now, two events can influence the life-time of the contact

1. One of both nodes disconnects. This happens with rate $2\gamma(1 - P)$.
2. One of both nodes reconnects to a node which did not belong to its group, which happens with rate $2\gamma P[1 - k/(N - 1)]$.

This means that ignoring changes in group size of both nodes, an edge attached to two nodes of degree k decays with rate

$$\omega_k^- = 2\gamma \left(1 - \frac{kP}{N - 1}\right) \quad (8.12)$$

The probability that any random edge drawn from the network is attached to two nodes of degree k is

$$\pi_k = \frac{kp_k}{\langle k \rangle}. \quad (8.13)$$

Hence, the mean decay rate of a single edge can be approximated as

$$\langle \omega^- \rangle = 2\gamma \left(1 - \frac{P}{N - 1} \frac{\langle k^2 \rangle}{\langle k \rangle}\right). \quad (8.14)$$

This result is merely an approximation because while both nodes are connected, their degree may change through group change events. For small-sized groups, the probability that a group size-change event is also an edge decay-event is higher than for larger groups, so for large values of P (higher mean group sizes), Eq. (8.14) should display

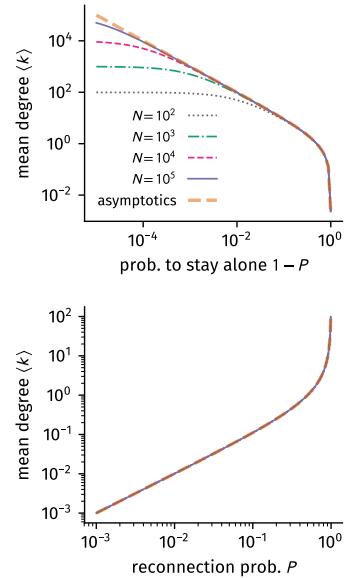


Figure 8.10: Flockwork equilibrium mean degree as a function of the reconnection probability P and number of nodes N , compared to the asymptotics $\langle k \rangle = P/(1 - P)$.

a less accurate behavior. The contact life-time distribution may then be approximated as

$$p(\tau_C) = \langle \omega^- \rangle \exp(-\langle \omega^- \rangle \tau_C). \quad (8.15)$$

First, Eq.(8.14) was tested numerically. For any parameter combination of $N = \{25, 50, 100\}$, and decreasing $0.999 > P > 0$, 100 Flockwork simulations were performed. Each of those was initiated using Alg. 1 with the options “binomial” and “shuffled group size iterations”. Afterwards, the configuration was simulated for $T = 10^5(N\gamma)^{-1}$ to obtain a true equilibrium configuration. This equilibrium configuration was subsequently used as the initial configuration to simulate for a run time of $T = 2 \times 10^6(N\gamma)^{-1}$. From this simulation, the life-time was measured for active periods of each link which began neither at t_0 nor ended at $T + t_0$. Then $\langle \omega^- \rangle$ was obtained as the inverse of the mean of those life times and compared to Eq. (8.14). The results of this analysis are shown in Fig. 8.11. One finds that Eq. (8.14) satisfactorily approximates the contact decay time with a maximum relative error of $< 2\%$ for $P = 0.999$ and $N = 100$. Furthermore, the relative error seems to decrease as a power-law as $\text{rel. err.} < A(1 - P)^{-1}$ where A is a small constant. For larger numbers of N and higher values of P , it is however recommended to test the validity of Eq. (8.14) explicitly for the chosen values of N and P .

In order to map a Flockwork model to the parameters of the edge activity model, one may consequently use Eq. (8.14) and the network density which is

$$p = \frac{\langle k \rangle}{N - 1}.$$

In the α - β -model where $\gamma = \alpha + \beta$ and $P = \alpha / (\alpha + \beta)$ this extends to

$$\langle \omega^- \rangle = 2\alpha \left(1 - \frac{\langle k^2 \rangle}{(N - 1) \langle k \rangle} \right) + 2\beta.$$

Hence, two parameters which fully define an instance of an edge activity model are directly determined by any two parameters which fully define an instance of a Flockwork model.

8.1.6 Comparing the Basic Poissonian Model to Real Data

In order to test Eqs. (8.10)-(8.15), single Flockwork simulations have been performed for $N = 25$ nodes and $P \in \{0.1, 0.7, 1 - 1/N, 0.999\}$. For each value of P , an equilibrium configuration was sampled using Alg. 1 using the “binomial” and “shuffled group sizes” options. Afterwards, the Flockwork process was simulated for a run time of $T = 10^5(N\gamma)^{-1}$ to achieve a true equilibrium configuration. Subsequently, the Flockwork process was run for $T = 3 \times 10^6(N\gamma)^{-1}$ during which the group size distribution, contact and inter-contact time distributions, and group life-time distributions were measured, with the results shown in Fig. 8.12. For each value of P , the theoretical distributions match with the empirical distributions, with the

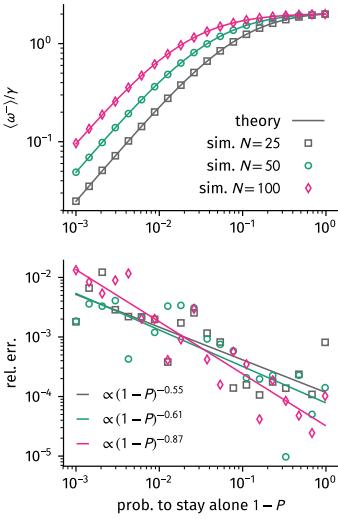


Figure 8.11: Measurements vs. Eq. ((8.14)). Each data point is a mean over 10 independent simulations of runtime $T = 2 \times 10^6(N\gamma)^{-1}$.

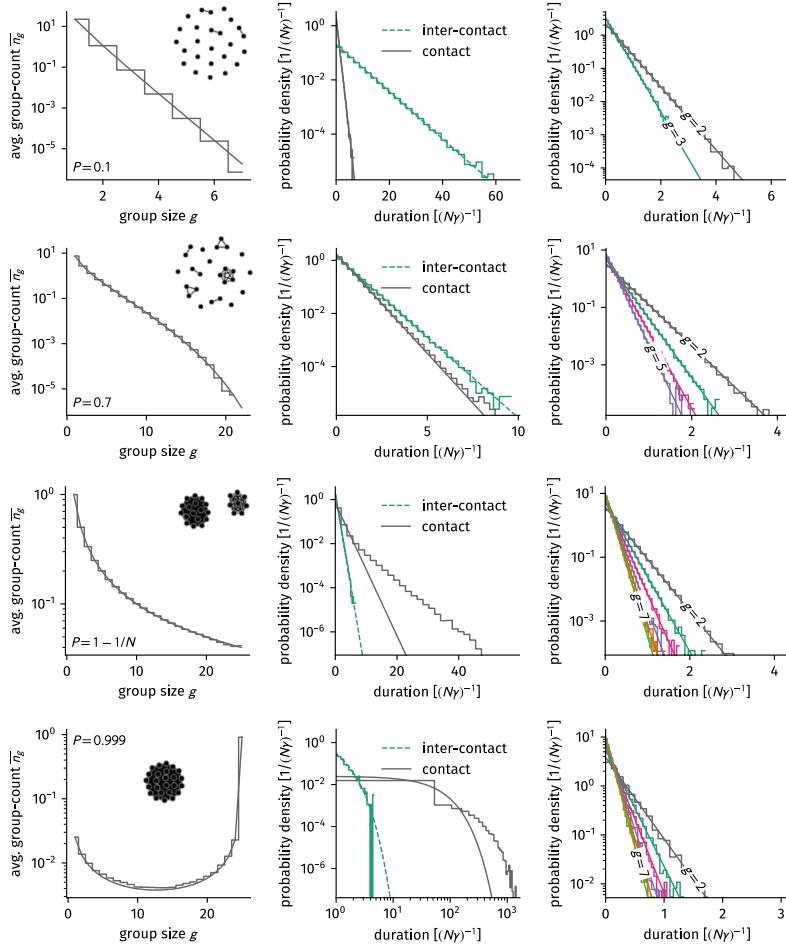


Figure 8.12: Ergodic group-size and life-time distributions of single Flockwork simulations for $N = 25$ and increasing reconnection probability P . Displayed are measurements over a runtime of $T = 3 \times 10^6(N\gamma)^{-1}$ with $\gamma = 1$. Step functions and continuous lines represent the theoretical predictions of the group-size distribution Eqs. (8.3), the inter-contact time distribution Eq. (8.10), the approximated contact time distribution Eq. (8.15) and the group life-time distribution Eq. (0).

exception of the contact-duration distribution which is predicted correctly for short durations but follows a slower exponential decay in the tail. Nevertheless, the distribution is of exponential shape.

As discussed in Sec. 2.4.4, real temporal contact networks show heavy-tailed group-size distributions, as well as heavy-tailed contact time, inter-contact time, and group life-time distributions. For the group life-time distributions, their shape was dependent on the group size, while their cut-off was not. While the Poissonian network model may be tuned to exhibit a power-law group-size distribution, all other behavior is not observed. This is not very surprising as the network model is, as stated, Poissonian. It is, however, well known that Poissonian systems with varying rates can be responsible for heavy tails in inter-event time distributions [48]. Furthermore, real-world contact data shows circadian behavior in their activity. It might therefore be the case that heavy-tailed statistics are a result of circadianly varying node activity rates $\alpha(t)$ and $\beta(t)$. This hypothesis will be explored in the following section.

8.2 Real-World Systems as Based on the Flockwork Model

²The author, spending most of his life in a densely populated urban area, is aware that there are indeed exceptions to this rule, but those will not be of concern here.

All real-world temporal face-to-face networks introduced in Sec. 2.4.4 show that network activity and number of edges follow a time-varying circadian pattern, which is congruent with expectations obtained by reflecting on everyday life: “average” humans are far more likely to connect to various other people during the day than they are during the night.² It therefore seems naïve to assume that any data measured from real-world interactions might be interpretable as generated from a basic Poissonian Flockwork model. Instead, the following picture will be explored in this section. At any time t of the measurement, the system might be interpreted to be described by two state variables: the active reconnection rate per node $\alpha(t)$ and the average mean degree $k_0(t)$. It will be shown that those quantities can be measured from data, which would allow for an interpretation of a human contact system as being locally equilibrated in time and fully described by the state variable pair (α, k_0) , similar to a gas which might be interpreted to go through various equilibrium states in, e.g., a Carnot machine [25] p. 604.

8.2.1 Inferring Time-Varying Activity Rates and Network State Variables

Following the argumentation above, it is reasonable to assume that the activity rates defined in Sec. 8.1.1 are not constant but time-varying such that $\alpha \rightarrow \alpha(t)$ and $\beta \rightarrow \beta(t)$. In the following, a method is derived to infer time-varying rates from real-world networks by counting edges and edge events.

As introduced in Sec. 8.1.1, the rate α quantifies the average number of active reconnection events a node undergoes per unit time. The rate β is the average number of disconnection events per unit time.

Further, the average mean degree k_0 quantifies the configuration mean degree $\langle k \rangle(t) = \sum_{u=1}^N k_u(t)$ as an ergodic mean

$$k_0 = \frac{1}{T} \int_0^T \langle k \rangle(t) dt.$$

Now, adding a time dependency to k_0 , its interpretation is as follows: for any functions $\alpha(t)$ and $\beta(t)$ given a large number of realizations, $k_0(t)$ is the average of $\langle k \rangle(t)$ over all those realizations at time t . The expected number of edges in the system is then given as $e_0 = k_0/(2N)$. Suppose a temporal face-to-face contact network was generated based on a Flockwork process with time-varying rates $\alpha(t)$ and $\beta(t)$. Then, the expected number $e_0(t)$ of edges present in the network will change approximately as

$$\partial_t e_0(t) = \underbrace{-\alpha N k_0(t) - \beta N k_0(t)}_{\text{edges leaving}} + \underbrace{\alpha N (1 + k_0(t))}_{\text{edges being built}}. \quad (8.16)$$

In the following, $\langle e_{\text{out}} \rangle(t, t + \Delta t)$ denotes the expected total number of edges being cut in the time interval $(t, t + \Delta t]$ and $\langle e_{\text{in}} \rangle(t, t + \Delta t)$ will be the expected total number of edges being formed in the time interval $(t, t + \Delta t]$. Then, according to Eq. (8.16), for a single realization of a Flockwork model, one finds

$$\begin{aligned}\langle e_{\text{out}} \rangle(t, t + \Delta t) &= N \int_t^{t+\Delta t} [\alpha(t') + \beta(t')] \langle k \rangle(t') dt' \\ \langle e_{\text{in}} \rangle(t, t + \Delta t) &= N \int_t^{t+\Delta t} \alpha(t') [1 + \langle k \rangle(t')] dt'.\end{aligned}$$

As a simplification, one may assume that $\alpha(t)$ and $\beta(t)$ change every Δt but will be constant in between. Then,

$$\begin{aligned}\langle e_{\text{out}} \rangle(t, t + \Delta t) &= N [\alpha(t, t + \Delta t) + \beta(t, t + \Delta t)] \int_t^{t+\Delta t} \langle k \rangle(t') dt' \\ &= 2[\alpha(t, t + \Delta t) + \beta(t, t + \Delta t)] E(t, t + \Delta t) \\ \langle e_{\text{in}} \rangle(t, t + \Delta t) &= N \alpha(t, t + \Delta t) \int_t^{t+\Delta t} [1 + \langle k \rangle(t')] dt' \\ &= \alpha(t, t + \Delta t) [N \Delta t + 2E(t, t + \Delta t)],\end{aligned}$$

where

$$E(t, \Delta t) = \int_t^{t+\Delta t} e(t') dt' \quad (8.17)$$

with $e(t)$ being the definite number of edges present in this realization of the network at time t . Note that E has the dimension of time. Then one finds the activity rates

$$\begin{aligned}\alpha(t, t + \Delta t) &= \frac{\langle e_{\text{in}} \rangle(t, t + \Delta t)}{N \Delta t + 2E(t, \Delta t)} \\ \beta(t, t + \Delta t) &= \frac{\langle e_{\text{out}} \rangle(t, t + \Delta t)}{2E(t, t + \Delta t)} - \alpha(t, t + \Delta t)\end{aligned}$$

These equations will be the basis for measuring rates. From any temporal network the observables e_{in} and e_{out} can be measured by counting the number of edge-creation (edge-deletion) events of the whole network in the time-interval $(t, t + \Delta t]$. Furthermore, E can be measured by solving the integral Eq. (8.17).

Now, time will be cut into equally-sized intervals of length Δt such that each rate $\lambda \in \{\alpha, \beta\}$ can be expressed as

$$\lambda(t) = \lambda_n \quad n \Delta t < t \leq (n + 1) \Delta t$$

For any non-negative integer n . Hence, the rates will be calculated as

$$\begin{aligned}\alpha_n &= \frac{\langle e_{\text{in}} \rangle(n \Delta t, (n + 1) \Delta t)}{N \Delta t + 2E(n \Delta t, (n + 1) \Delta t)} \\ \beta_n &= \frac{\langle e_{\text{out}} \rangle(n \Delta t, (n + 1) \Delta t)}{2E(n \Delta t, (n + 1) \Delta t)} - \alpha_n.\end{aligned}$$

Note that from the definition above it might occasionally happen that $\beta_n < 0$, in which case one sets $\beta_n = 0$. Furthermore, for networks where at some point $\langle k \rangle(t) = 0$, the calculated rates are often $\alpha = 0$ and $\beta = 0$. This often times produces situations in which the Flockwork will not equilibrate to be an empty network fast enough. In these cases, the disconnection rate β will be artificially increased in order to ensure that every node in the network is disconnected. The following rule is introduced:

If in both, a time bin $(n\Delta t, (n+1)\Delta t]$ and its following time bin $((n+1)\Delta t, (n+2)\Delta t]$ there is no present edges ($e = 0$), and both rates evaluate to $\alpha_n = \beta_n = 0$, and the previous time bin $((n-1)\Delta t, n\Delta t]$ was not used for emptying, redefine the disconnection rate to be

$$\beta_n = \beta^{\text{emptying}} = \frac{\log N}{\Delta t}.$$

The reason for choosing β_n as being proportional to the logarithm of the number of nodes is the following. Picking any node at random, it will on average take $N \log N$ picks to reach any node at least once (this is equal to the mean cover time of a complete network, see Sec. 6.1.2). The total disconnection rate $N\beta_n^{\text{emptying}}$ will therefore ensure that on average every node is disconnected during the time interval Δt and will stay alone afterwards. It now also becomes clear why α and k_0 are chosen to be the state variables describing the networks: β is occasionally artificially set while k_0 evolves over time following Eq. (8.16), closely describing the current network state. Furthermore, each equilibrium mean degree k_0 can be directly mapped to one and only one value of the reconnection probability P . An equilibrium which would be approached by using β_n and α_n for the Flockwork process might not be reached because there is not enough time until α_n and β_n change again for the next time interval. However, k_0 evolves according to Eq. (8.16), and hence the average equilibrium network state at time t is given by the Flockwork equilibrium state determined by $P(t) \approx k_0(t)/(1 + k_0(t))$. One might wonder why the actual mean degree $\langle k \rangle(t)$ of a configuration at time t should not be chosen as a state variable. Doing so, however, does not seem too favourable since this quantity is subject to noise in any Flockwork equilibrium state, while the expected mean degree k_0 is not. Therefore k_0 seems like an ideal choice to describe the average structural state of the network at any time t . The temporal state of the network will be described by the active reconnection rate or network “momentum” α . Note that for any position in the state space (α, k_0) , typical behavior of the temporal network can be described. Please refer to the qualitative regions displayed in Fig. 8.13. For small values of α , a single node will establish a small amount of new contacts within unit time. These contact events will not produce many edges when k_0 is small, so the network will be very sparse with a large amount of nodes being alone and within unit time the network structure will not change too much. In contrast, for small α and large k_0 , larger groups will form, but the network will still appear relatively static when observing for unit time. This changes

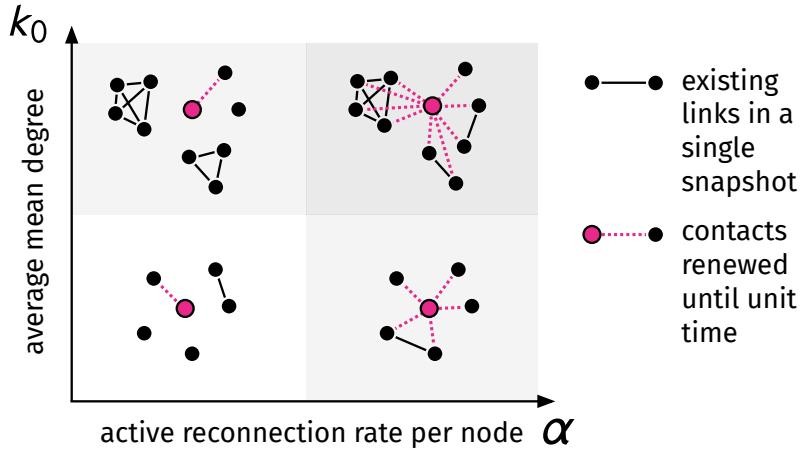


Figure 8.13: Example network behavior for different positions in the Flockwork state space. Solid black lines represent existing links in a single snapshot, dashed pink lines represent new contacts a single focal node will build (and eventually cut) observation of unit time length. Small reconnection rate α implies

for large values of α and small values of k_0 . Here, the network will still consist of many lone nodes, however the network structure will change very much when observing for unit time. Within this time, a single node will establish many new connections but will break them off again within a short time such that k_0 is small at any time. For large values of both α and k_0 large group exist, but the network structure will change drastically while observing for unit time. A real-world system might be interpreted as following a trajectory on this state space.

In order to obtain these state variables from data, an inference procedure is proposed as follows.

1. Choose a time interval size Δt and cut the observation time of a data set in intervals of length Δt . The last interval size $\Delta t_{\text{final}} \leq \Delta t$ may differ to ensure that the whole observation time t_{max} is covered.
2. Measure the number of edges e_{in} being created and the number of edges e_{out} being deleted during each time interval $(n\Delta t, (n+1)\Delta t]$ or $(n\Delta t, t_{\text{max}}]$. For each time interval, integrate the number $e(t)$ of edges present during this interval using Eq. (8.17) to obtain E .
3. For each time interval, calculate the rates using Eqs. (8.18) while ensuring that $\beta_n \geq 0$ and that the empty network may be reached by setting $\beta_n = \beta^{\text{emptying}}$ if applicable as described above. Make sure to replace Δt with Δt_{final} for the last interval.
4. Evaluate the function $k_0(t)$ by integrating Eq. (8.16) using $k_0(0) = 2e(0)/N$ while $e(0)$ is the number of initial edges in the analyzed data set.

Unfortunately, it is unclear how Δt has to be chosen. However, one may assume that this procedure infers the correct rates from a simulated instance of a Flockwork model which was generated using predetermined rates α and β changing their value every Δt . Such a test is done in the following.

8.2.2 Testing with Synthetic Time-Varying Flockwork Data

The procedure described above is tested as follows. A system is defined in which a Flockwork model approaches several states in the (α, k_0) -plane by defining time-dependent rates $\alpha(t)$ and $\beta(t)$ which change their value every $\Delta t = 20$ as step functions. The unit of time will be fixed as $[t] = 1$ in the sense that if $\alpha(t) = 1$ then in one unit of time, there will be N active reconnection events on average. The rates will loop over 4 consecutive state-changes and thus each loop is associated with a non-negative integer $\ell \geq 0$. The number of nodes was chosen as $N = 200$, the time-varying rates were set to

$$\alpha(t) = \begin{cases} 0.1 & 4\ell\Delta t \leq t < (4\ell+1)\Delta t \\ 10 & (4\ell+1)\Delta t \leq t < (4\ell+2)\Delta t \\ 10 & (4\ell+2)\Delta t \leq t < (4\ell+3)\Delta t \\ 0.1 & (4\ell+3)\Delta t \leq t < (4\ell+4)\Delta t, \end{cases} \quad (8.19)$$

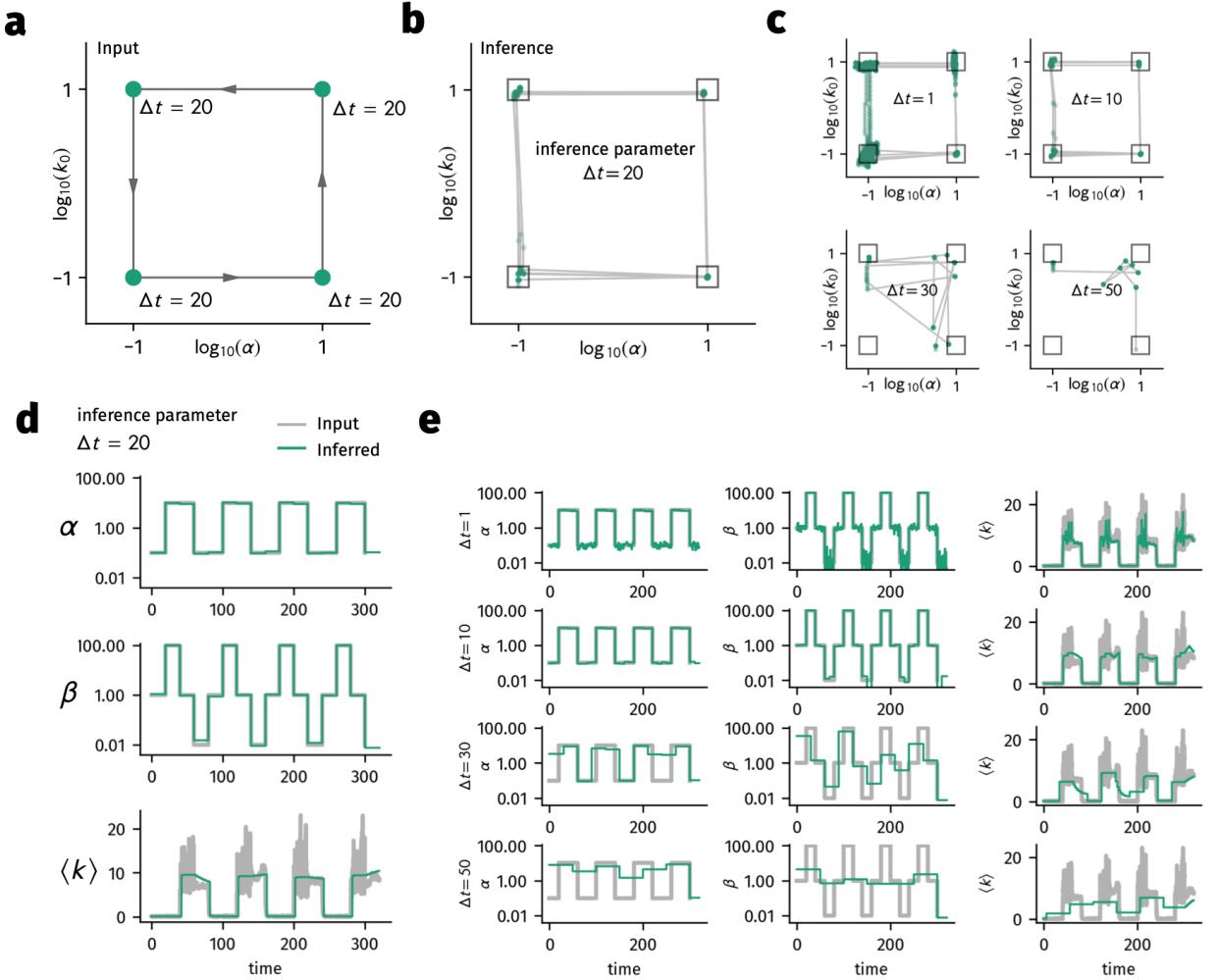
and

$$\beta(t) = \begin{cases} 1 & 4\ell\Delta t \leq t < (4\ell+1)\Delta t \\ 100 & (4\ell+1)\Delta t \leq t < (4\ell+2)\Delta t \\ 1 & (4\ell+2)\Delta t \leq t < (4\ell+3)\Delta t \\ 0.1 & (4\ell+3)\Delta t \leq t < (4\ell+4)\Delta t, \end{cases} \quad (8.20)$$

such that, by using $P = \alpha/(\alpha + \beta)$ and $k_0 \approx P/(1 - P)$, the system is expected to move in the (α, k_0) -plane as $(0.1, 0.1) \rightarrow (10, 0.1) \rightarrow (10, 10) \rightarrow (0.1, 10) \rightarrow (0.1, 0.1)$ and so forth, as shown in Fig. 8.14a. Essentially, the system should stay in each of those states for $\Delta t = 20$ and then transfer to the next state quickly, assuming that they equilibrate rapidly. Then, the parameters α , β , and k_0 should be correctly inferred using the procedure described above with $\Delta t_{\text{inference}} = \Delta t = 20$.

Instances of the Flockwork model with the defined input rates were simulated as follows. A Flockwork configuration was initiated using Alg. 1 with the options “Binomial distribution” and “shuffled group sizes” for $P = 0.1/(1 - 0.1)$ and $N = 200$. Then, the system was simulated using Gillespie’s SSA with rates α and β as defined above until $t = 20$. Each time the rates switched, a new Flockwork simulation was initiated with the new rates and the last active configuration. The system was then simulated until $t = 40$ when the rates switched again. This procedure was repeated until $t = 320$, which means for four loops. Technically, this kind of simulation is “wrong” because when the rates switch, information about the time of the last event and the last rate has to be considered to evaluate the correct inter-event time (see App. C.2.3), this inter-event time when changing from one state to another, is, however, approximately upper bounded by $(N\gamma_{\min})^{-1} = (N\alpha_{\min} + N\beta_{\min})^{-1} \approx 1/(200 \times 0.2) = 1/40$, which is negligible compared to the time-interval length of the rate change $\Delta t = 20$.

Several instances of Flockwork simulations have been performed as described above, subsequently inferring the rates with varying



inference parameter $\Delta t_{\text{inference}}$. The analysis of a single instance is displayed as an example in Fig. 8.14 while a detailed analysis with more values of $\Delta t_{\text{inference}}$ is shown in App. B.2 (see Figs. B.1 and B.2). As one may see in Fig. 8.14b, the correct state-space trajectory is recovered using the input value $\Delta t = 20$ as the inference parameter. The rates $\alpha(t)$ and $\beta(t)$ are correctly inferred, too, as is shown in Fig. 8.14d, alongside the curve of the instance's temporal variation of the configurational mean degree $\langle k \rangle(t)$ and the inferred expected mean degree $k_0(t)$, which is computed using the inferred rate functions and Eq. (8.16). Here, for each inference value $\Delta t_{\text{inference}}$, the expected mean degree $k_0(t)$ was evaluated for 10 equally distant points in each time-interval of length $\Delta t_{\text{inference}}$, hence for $\approx 3200/\Delta t_{\text{inference}}$ time points in total. In order to explore how the inference changes with different values of $\Delta t_{\text{inference}}$, a thorough scan with different values of $\Delta t_{\text{inference}}$ was conducted, four of which are shown in Fig. 8.14c and 8.14e. Generally, values of $\Delta t_{\text{inference}} \leq \Delta t/2$ reproduce a fair picture of the state space trajectory but are subject to noise. The smaller the value of $\Delta t_{\text{inference}}$ is chosen, the more the rates will be fit directly to the equilibrium fluctuations, yielding overfitted results

Figure 8.14: Inferring time-varying rates and the state-space trajectory of a corresponding Flockwork simulation. (a) The input trajectory for the Flockwork simulation as based on the defined rate functions Eqs. (8.19) and (8.20). (b) The rate inference-procedure described in Sec. 8.2.1 yields the correct expected state-space trajectory. (c) Inference with time lengths smaller than the “true” Δt generally reproduce the “correct” state-space trajectory but are subject to noise. Inference using values of Δt greater than the “true” Δt will heavily deviate from the expected state-space trajectory. (d) The rates and expected mean degree are recovered correctly with only minor deviations when using the true interval length as the inference parameter. (e) Using $\Delta t_{\text{inference}} \leq \Delta t/2$ will yield rates fluctuating around the true functions, influenced by overfitting the fluctuations of the Flockwork’s mean degree in its equilibrium state. Values of $\Delta t_{\text{inference}} > \Delta t$ will fail to accurately infer the rate functions and the expected mean degree.

(see Fig. 8.14e). When increased to values $\Delta t_{\text{inference}} > \Delta t$, the procedure yields heavily deviated space-state trajectories.

A thorough and quantitative method to find the optimal inference parameter $\Delta t_{\text{inference}}$ without prior knowledge about the input functions will have to be the subject of future research. For real data, there is no reason for the time-varying rate functions to be step functions changing every Δt , so there is no “true” Δt . Rather, one may learn from this analysis that it may be of value to steadily increase the inference parameter and observe how the trajectories behave. One may then choose a Δt which reduces the noise from the trajectories obtained from small $\Delta t_{\text{inference}}$ while striking features from the noisy trajectories are still present. Furthermore, it seems appropriate to compare the measured curve of $\langle k \rangle(t)$ to the inferred function $k_0(t)$ and choose a Δt which satisfactorily reproduces important features of the original curve while smoothing out heavy fluctuations. This method is used in the following to obtain the state-space trajectories of all data sets introduced in Sec. 2.4.4.

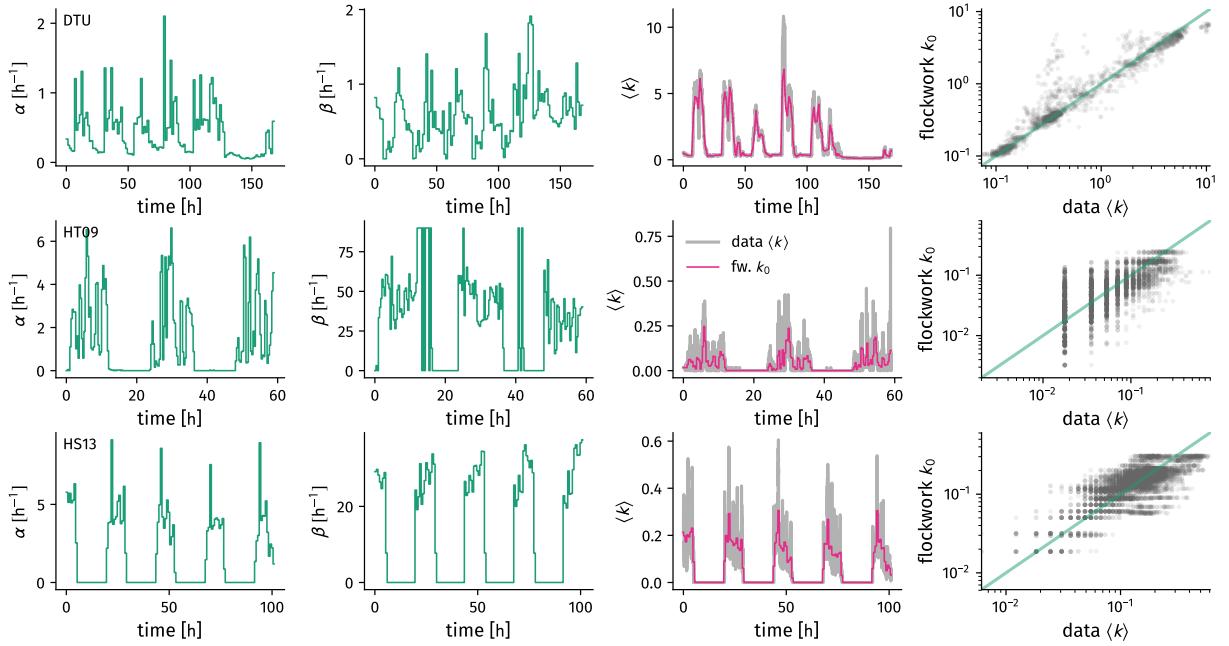
8.2.3 Application to Real-World Face-to-Face Systems

In this section, the DTU, HT09 and HS13 data sets as introduced in Sec. 2.4.4 are investigated for their state-space trajectory in the (α, k_0) -plane. The analysis is based on the assumption that those data sets can be interpreted to have been generated by a time-varying Flockwork model.

All three of the data sets are similar in the sense that they contain contact data which has been binned to discrete time intervals Δt , meaning that if an edge was active during the time-interval $[n\Delta t, (n+1)\Delta t]$ it will be considered to have been active through-out the whole time interval.³ For the DTU data, this value is $\Delta t = 5 \text{ min}$ [68], while for both *SocioPatterns* data sets (HT09 and HS13), the interval length is $\Delta t = 20 \text{ s}$ [5]. One other difference between the data sets is given by the way interactions are recorded. In the *SocioPatterns* experiments, face-to-face contacts actually reflect face-to-face interactions within a radius of $r = 1 - 1.5 \text{ m}$ because the signals which are transferred between chips recording the interaction are shielded by human bodies [5]. In contrast, data from the DTU experiment reflects contacts between all individuals which are within a radius of $r \approx 1.5 \text{ m}$ [68]. Both these differences will be of importance in the following analysis.

For all three data sets, the method introduced in Sec. 8.2.1 was applied for increasing $\Delta t_{\text{inference}}$ where the parameter was chosen as a multiple of the original binning time-interval length. All inferred rate functions as well as the expected mean degree function for varying $\Delta t_{\text{inference}}$ are shown in Fig. B.3-B.11 of App. B.2. Comparing $\alpha(t)$, $\beta(t)$ and $k_0(t)$ for varying $\Delta t_{\text{inference}}$ shows that the position and shape of the state-space trajectory does not drastically vary with the choice of the inference parameter over a wide range, and as such the exact choice of $\Delta t_{\text{inference}}$ does not seem to be too im-

³ with n being a non-negative integer



portant. Hence, based on the criteria established in the previous sections (reduced noise, but stable shape of trajectory), the following choices were made: $\Delta t_{\text{inference}}^{\text{DTU}} = 80 \text{ min}$, $\Delta t_{\text{inference}}^{\text{HT09}} = 20 \text{ min}$, and $\Delta t_{\text{inference}}^{\text{HS13}} = 45 \text{ min}$, which yield the functions shown in Fig. 8.15. For all data sets, the evaluated expected mean degree $k_0(t)$ follows the data while not overfitting to noise. The activity rates differ between the DTU data and the SocioPattern sets, showing active reconnection rate values of $0 < \alpha^{\text{SocPat}} \lesssim 6 \text{ h}^{-1}$ per node and $0 < \alpha^{\text{DTU}} \lesssim 1.5 \text{ h}^{-1}$ per node during the day and values of $\alpha \ll 1 \text{ h}^{-1}$ per node at night, with mean rates of $\overline{\alpha^{\text{DTU}}} = 0.4 \text{ h}^{-1}$, $\overline{\alpha^{\text{HT09}}} = 1.4 \text{ h}^{-1}$, and $\overline{\alpha^{\text{HT09}}} = 1.7 \text{ h}^{-1}$ averaged over the whole measurement. The active disconnection rate β has average values of $\overline{\beta^{\text{DTU}}} = 0.6 \text{ h}^{-1}$, $\overline{\beta^{\text{HT09}}} = 31 \text{ h}^{-1}$, and $\overline{\beta^{\text{HT09}}} = 12.5 \text{ h}^{-1}$ averaged over the whole measurement. These inferred rates yield state-space trajectories which are qualitatively consistent between the SocioPatterns data sets HT09 and HS13 showing behavior in which the system changes from both low values of α and k_0 to higher values of both in a close to linear fashion (see Fig. 8.16a and 8.16b). The DTU trajectory explores less high regions of the active reconnection rate but greater values of the expected mean degree. This discrepancy can be explained by the two striking differences of the experiment: The SocioPattern data is binned in time bins of $\Delta t = 20 \text{ s}$ and only records face-to-face contacts while the DTU experiment bins data in intervals of length $\Delta t = 5 \text{ min}$ and defines contacts via proximity. Binning data with larger time intervals has the effect of building edges in times where no edges have been observed (see Fig. [not yet in, it's supposed to be in Sec 2.4]), which supposedly increases the expected mean degree. Furthermore, contacts which switch on and off within a binning interval will not contribute to the total amount of edges

Figure 8.15: Inferred time-varying rates and expected mean degree of the real-world data sets (**top row**) DTU with $\Delta t_{\text{inference}}^{\text{DTU}} = 80 \text{ min}$, (**middle row**) HT09 with $\Delta t_{\text{inference}}^{\text{HT09}} = 20 \text{ min}$, and (**bottom row**) HS13 with $\Delta t_{\text{inference}}^{\text{HS13}} = 45 \text{ min}$. The activity rates $\alpha(t)$ and $\beta(t)$ as well as the expected mean degree $k_0(t)$ show the expected circadian behavior.

leaving or being created in this time and as such the rate α is hypothesized to decrease by rebinning data in larger time intervals. As displayed in Fig. 8.16c one observes that rebinning the SocioPatterns data set to bins of length $\Delta t = 5$ min and subsequently re-analysing this new data to obtain $k_0(t)$ and $\alpha(t)$ indeed yields decreased rates and increased mean degrees; the trajectories lie closer to the DTU data, however, not reaching the high values of the mean degree. This may be explained by considering the difference between face-to-face data and proximity data. One may approximate the difference as follows. In a disc around a focal individual, face-to-face contacts can only be produced to individuals within roughly a half of this disc, whereas proximity data will count connects to all individuals within the disc, due to signals of the recording RFID chips being absorbed by human bodies. Furthermore, only individuals whose half disc is turned towards the focal individual will be recorded as contacts. Further assuming homogeneous distribution of people, one may estimate that the expected degree of an average node will therefore be $(1/2) \times (1/2) = 1/4$ of a node in proximity data of similar range. This hypothesis was tested by further multiplying the trajectory of the rebinned SocioPattern data with a factor 4 which finally yields trajectories which are of similar shape as the DTU trajectory (see pink curve in Fig. 8.16c). This implies that all three systems might follow similar trajectories. One remarkable feature of the SocioPatterns trajectories is that neither regions where the network is dense but changes slowly, nor regions where the network is sparse but changes fast are explored. It seems that a fast change in network structure is always connected to a more densely connected network and a slowly changing network is associated with a sparse structure. [This will be of importance when analysing the spread of diseases for different time scales and configurational settings in Chap. 9.]

After analysing the supposed trajectory a Flockwork would take if it followed activity rates consistent with real world data sets, it is of course of interest whether actual Flockwork simulations follow the structural and temporal statistics seen in real data. To this end, the naïve varying-rate Flockwork simulation algorithm described in Sec. 8.2.2 was run⁴ using the inferred rates $\alpha(t)$ and $\beta(t)$ for the DTU, HT09, and HS13 data considering the respective number of nodes $N_{DTU} = 412$, $N_{HT09} = 113$, and $N_{HS13} = 327$. To reflect the way in which the original data is measured, the simulated temporal networks were binned in time bins of length $\Delta t = 5$ min for the DTU surrogates and $\Delta t = 20$ s for the SocioPatterns simulations and compared to the statistics obtained from the non-binned original data. Results for single representative measurements can be seen in Figs. 8.17–8.19. Focus is first given to panels a of the referenced figures, where single snapshots of both the original data as well as the binned Flockwork surrogate are shown for two different time points. In general, Flockworks following the varying rates of show similar behavior as the original temporal networks, nodes are organized in densely connected groups whose size distribution varies at different

⁴The reader is reminded that the “naïve” part of this algorithm is that when a rate change happens, the simulation is restarted with the new rate, ignoring all time that has passed before the rate change happened. Using the inference time interval length defined above, this implies that for the DTU data, there will be 126 changes and thus 126 slightly too long inter-event times of a total number of events of approximately $(0.4 + 0.6) h^{-1} \times 168 h \times 412 \approx 69,000$. For the HT09 data set, there is 177 of those slightly skewed inter-event times of 216,000 events and for the HT13 this evaluates to 135 skewed times compared to 270,000 events. Hence, the error made will be negligible for all inferred rates.

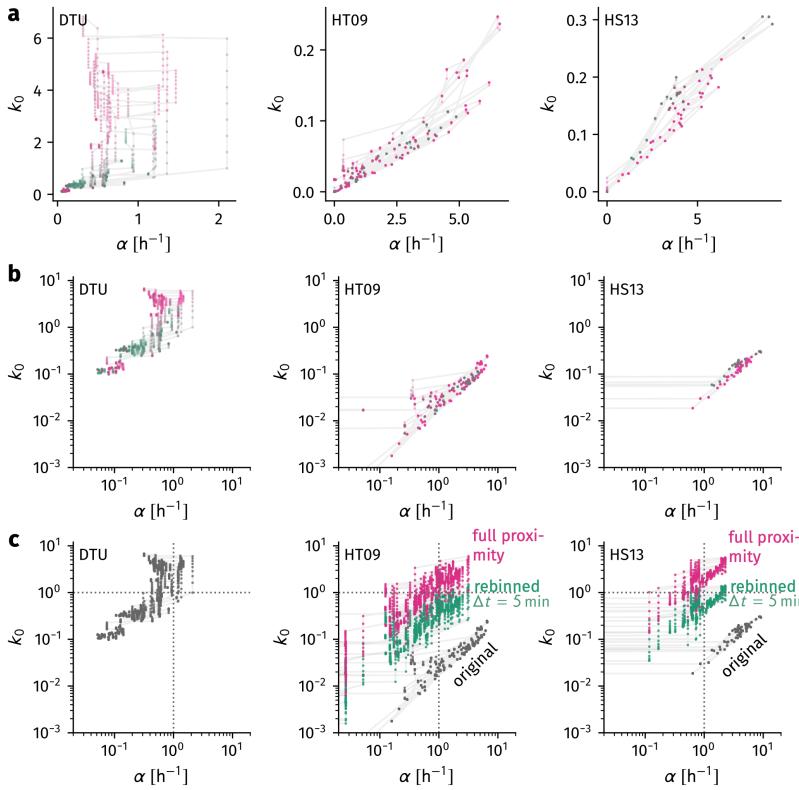


Figure 8.16: Inferred state-space trajectories of the real-world data sets. (**Top row**) trajectories in linear scale, (**middle row**) in log-log scale, and (**bottom row**) showing that binning the SocioPattern data to bins of length $\Delta t = 5$ min and correcting for

times. Furthermore, as can be seen in panel b of the respected figures, the general shape of the number of newly observed contacts as well as the density of edges corresponds between original data and binned Flockwork surrogates. However, many more contacts are explored in the Flockwork surrogates which are not seen in the original data. Additionally, the real data clearly shows that certain pairs of nodes are either in long contact or tend to break shortly just to be rebuilt almost immediately. This behavior is not replicated by the Flockwork model, in which nodes choose new neighbors uniform at random from the set of all other nodes and as such it is highly unlikely that a recently broken contact is rebuilt in a short amount of time. Another effect of nodes choosing neighbors uniform at random is that each link is equally likely to have been formed in Flockwork surrogates, which yields to a homogeneous aggregated network as opposed to the complex ones built from the original data (see panel c of the respective Figs. 8.17-8.19). While the model does not replicate the data in this regard the simple limit case of a homogeneous aggregated network will make it easier to analyze dynamics in aggregated static limit cases on those surrogates.

Attention is now shifted to the statistical properties of structural and temporal observables, i.e. the average degree distribution, average group-size distribution, contact duration distribution, inter-contact duration distribution and group life-time distributions of groups of varying size. First, note that these distributions are well-defined for constant-rate Flockwork equilibrium configurations as

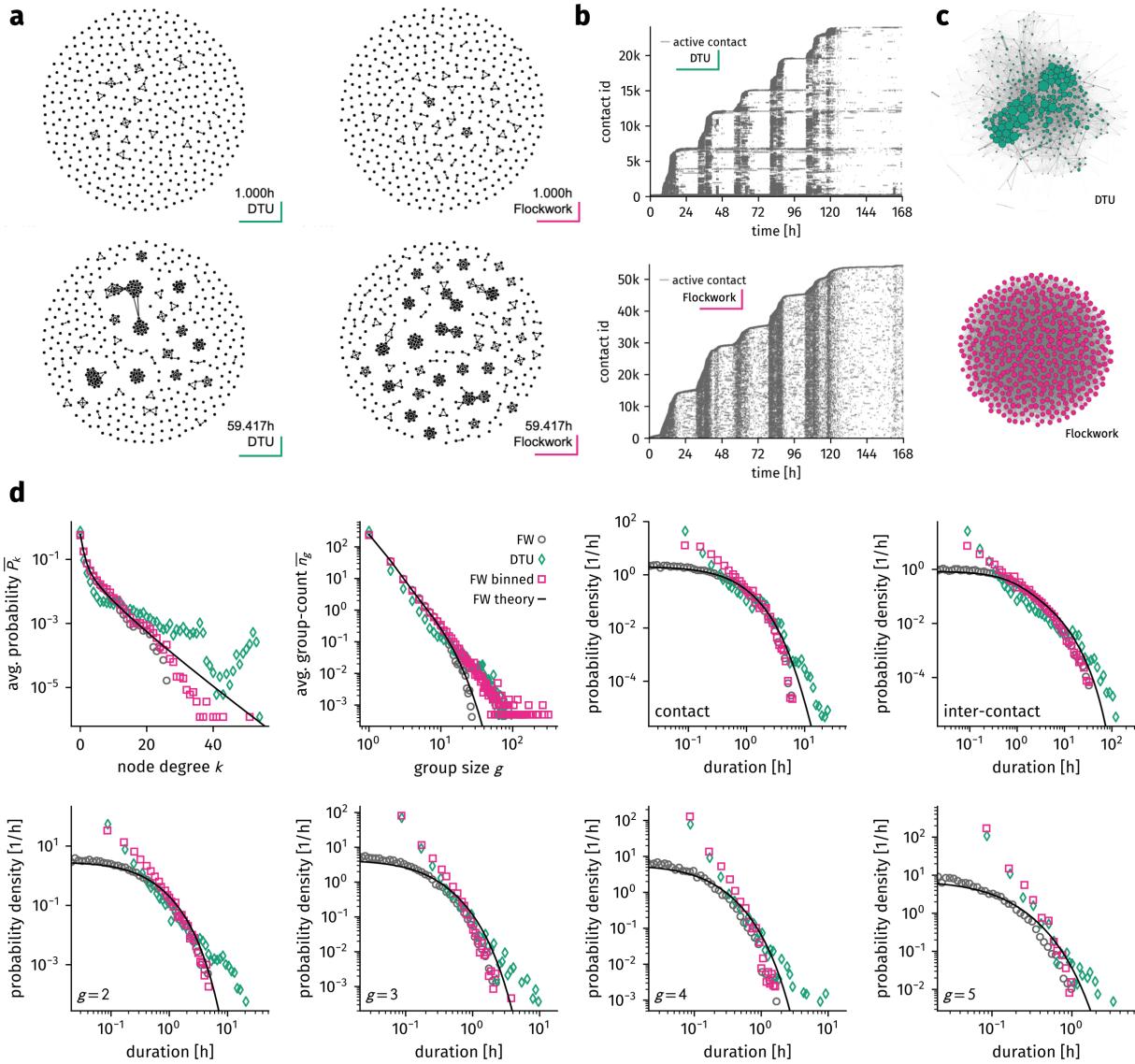


Figure 8.17: Structural and statistical comparison between the DTU data set and a corresponding time-varying Flockwork surrogate. (a) Snapshots of the original data and a corresponding binned Flockwork simulation at two different times. (b) Edge activity plot. The surrogate explores more edges which are not seen in the original data and does not produce contacts which are broken and immediately rebuilt. However, the general shape of how those new contacts are explored corresponds as well as the varied density. (c) The aggregated surrogate network is homogeneous and not “complex” as its original counterpart. (d) Temporal distributions do not generate the heavy tails observed in the real data, however, binning the data produces a heavy tail in group-size distribution and increases correspondence in the degree distribution. Binning further increases the probability of medium-length (inter-)contact life-times. Theoretical predictions Eqs. (8.21)-(8.25) (straight lines) and non-binned data (grey circles) are approximately corresponding.

derived in Sec. 8.1. In the varying-rate Flockwork model, the inferred expected mean degree $k_0(t)$ defines the equilibrium state in a structural sense, which can be approximately mapped to $P(t) = k_0(t)/[k_0(t) + 1]$. The temporal state is defined by the inferred rate $\alpha(t)$, such that the effective total event rate will be approximated as $\gamma(t) = \alpha(t)/P(t)$ (which will be equated to $\gamma(t) = 0$ if $k_0(t) = P(t) = 0$). Hence, assuming that the Flockwork follows a trajectory through the state space, all structural and temporal property distributions should be approximately given by a temporal average over these equilibrium states with $\gamma(t)$ and $P(t)$. For the structural distributions of degree and group sizes, one may use Eq. (8.3) and Eq. (8.7)

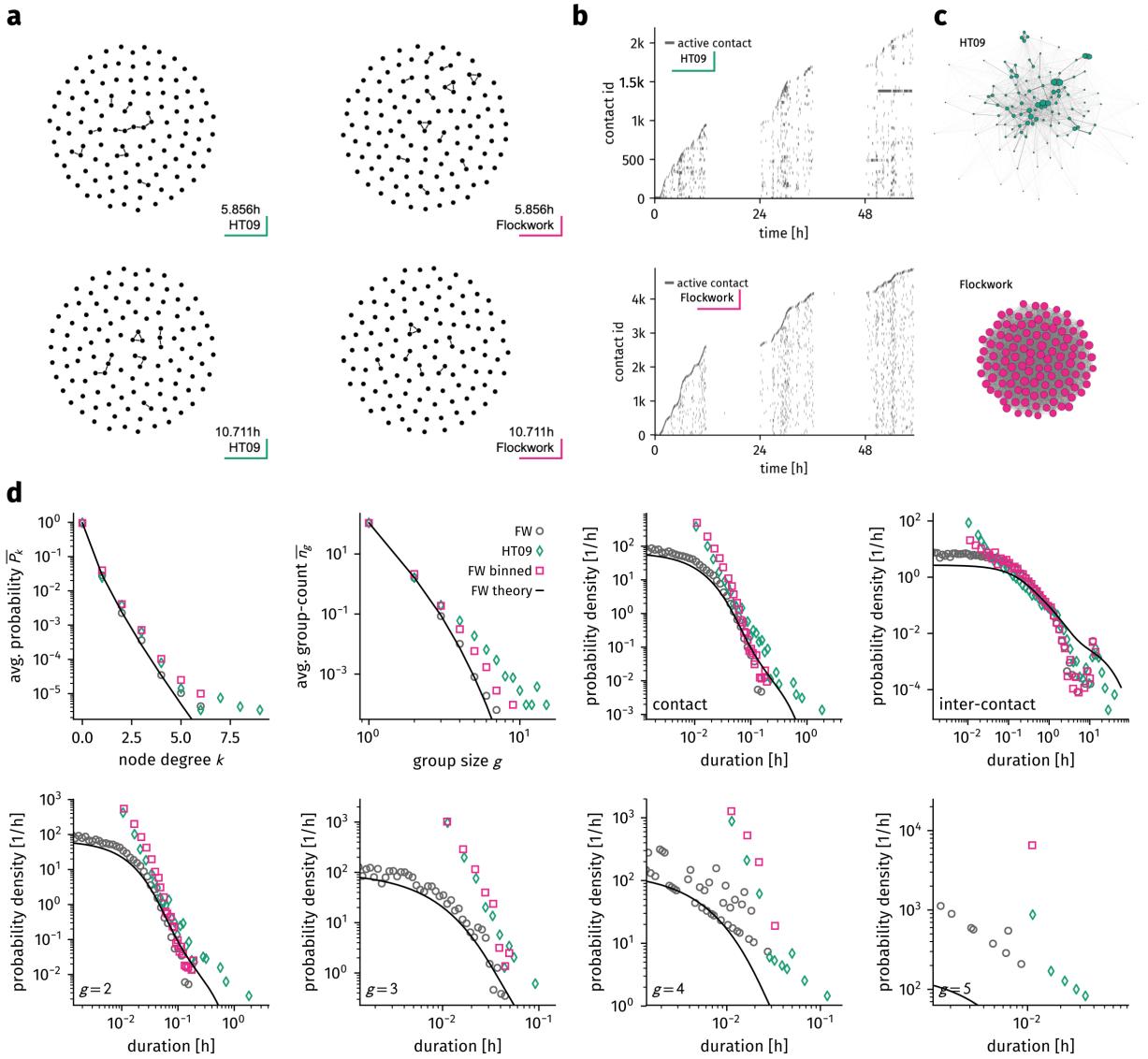


Figure 8.18: Structural and statistical comparison between the HT09 data set and a Flockwork surrogate. (a) Snapshots of original data and Flockwork simulation. (b) Edge activity plot. The surrogate explores more edges which are not seen in the original data and does not produce contacts which are broken and immediately rebuilt. However, the general shape of how those new contacts are explored corresponds as well as the varied density. (c) The aggregated surrogate network is homogeneous, not “complex” as its original counterpart. (d) Temporal distributions do not generate the heavy tails observed in the real data, however, binning the data produces a heavy tail in group-size distribution and increases correspondence in the degree distribution. Binning further increases the probability of medium-length (inter-)contact life-times. For the inter-contact duration distribution, the super-imposed time-scales seem to reproduce a heavy tail. Theoretical predictions Eqs (8.21)-(8.25) (straight lines) and non-binned data (grey circles) are approximately corresponding.

to find

$$\overline{n_m} = \frac{1}{T} \int_0^T dt n_m[P(t)] \quad (8.21)$$

$$\overline{P_k} = \frac{1}{T} \int_0^T dt P_k[P(t)]. \quad (8.22)$$

Using Eq. (8.10), the average inter-contact time distribution might be estimated as

$$\overline{p(\tau_{IC})} = \frac{1}{T} \int_0^T dt 2\alpha(t) \exp [-2\alpha(t)\tau_{IC}], \quad (8.23)$$

while the approximated contact-time distribution defined by Eqs. (8.14) and (8.15) evaluates to

$$\begin{aligned} \overline{p(\tau_C)} = & \frac{1}{T} \int_0^T dt 2\gamma(t) \left(1 - \frac{P(t)}{N-1} \frac{\langle k^2 \rangle [P(t)]}{\langle k \rangle [P(t)]} \right) \times \\ & \exp \left[-2\gamma(t) \left(1 - \frac{P(t)}{N-1} \frac{\langle k^2 \rangle [P(t)]}{\langle k \rangle [P(t)]} \right) \tau_C \right]. \end{aligned} \quad (8.24)$$

Note that here, $\langle \cdot \rangle [P(t)]$ denotes the configurational equilibrium expectation using $P_k[P(t)]$ as per Eq. (8.7). Finally, one may approximate the group life-time distribution as

$$\begin{aligned} \overline{p(\tau_m)} = & \frac{1}{T} \int_0^T dt \left(m\gamma(t)(1-P(t)) + 2\alpha(t)m \frac{N-m}{N-1} \right) \times \\ & \exp \left[-\tau_m \left(m\gamma(t)(1-P(t)) + 2\alpha(t)m \frac{N-m}{N-1} \right) \right]. \end{aligned} \quad (8.25)$$

When the time of measurement T is sufficiently large and the process truly follows a trajectory through equilibrium states, these temporal averages correspond to averages over rate distributions $p(\lambda)$ as

$$\overline{p(\tau)} = \langle p(\tau) \rangle = \int_0^\infty d\lambda p(\lambda) \lambda \exp(-\lambda\tau).$$

However, using the inferred rate functions above, several rates will be $\alpha = \beta = 0$ for a substantial amount of time and thus contribute $p(\tau) = 0$ to the integral for finite times. Intuitively though, the inter-event time between two events will be shifted to the length of the period in which the event rates are zero and thus finite probability for longer inter-event times should contribute to the average above. This results in an obvious discrepancy between the theoretical predictions and the unbinned simulated data.

Comparing the structural statistics first, one notices that the predicted degree and group-size distributions Eqs. (8.21) and (8.22) align rather well with the results from simulated data for all data sets as one may see in panel d of the respective Figs. 8.17, 8.18, and 8.19. The unbinned degree distributions also correspond well with the data degree distributions in the cases of HT09 and HS13, however, they differ in the DTU case which seems to have substantially more nodes of higher degree while still decaying exponentially. Binning the Flockwork surrogates to the time bins of their respective real experiments increases the typical degree-scale, but not drastically. In contrast, the group-size distributions of the unbinned Flockwork data decays faster than a heavy-tailed distributions for all data sets, which changes when binning is introduced – binned Flockworks show indeed heavy tails in the group size distributions. Since individuals can change between groups in a short amount of time, binning will lead to groups being measured as connected where they are actually not, thus increasing the amount of larger groups (see

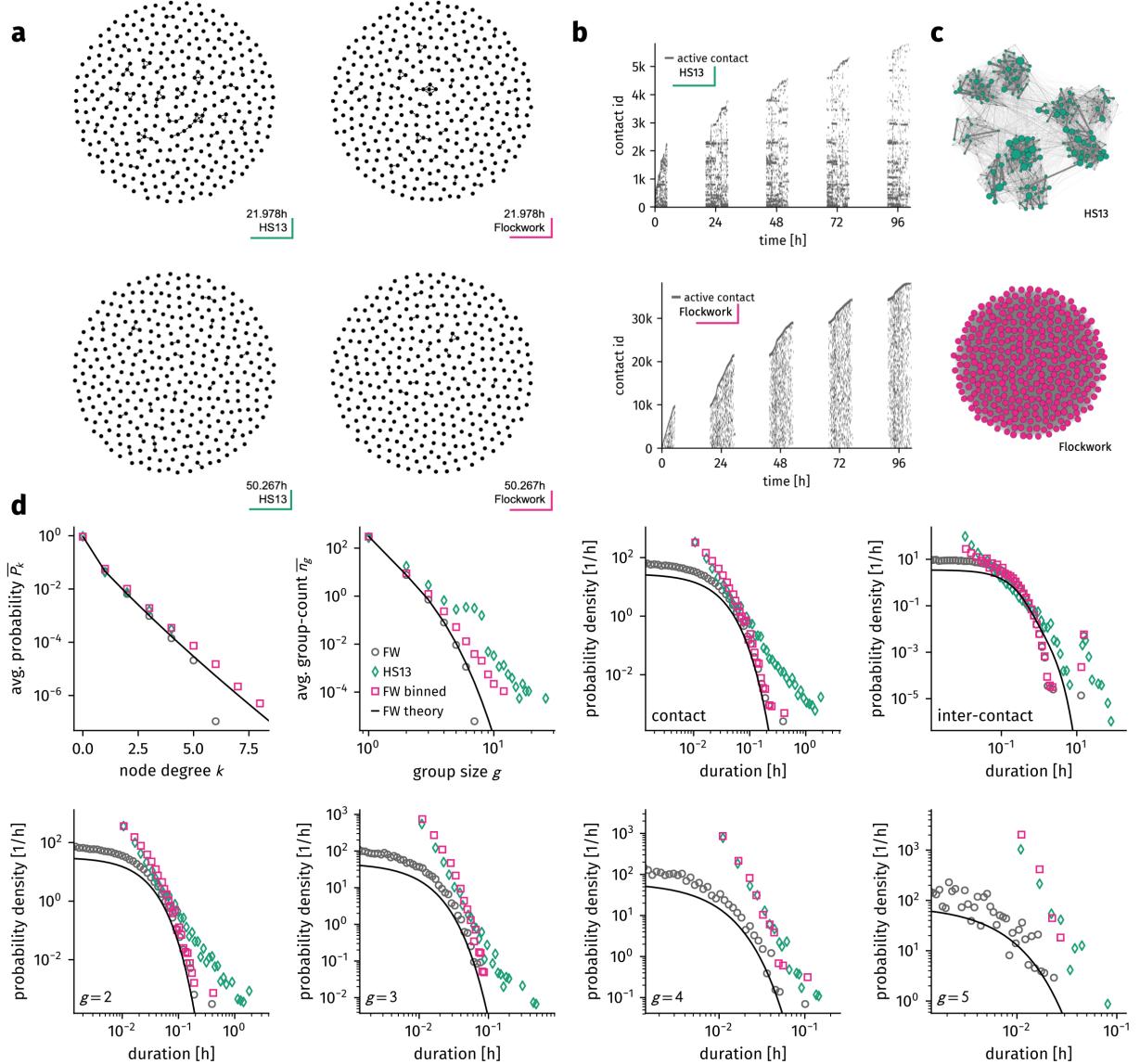


Figure 8.19: Structural and statistical comparison between the HS13 data set and a single corresponding time-varying Flockwork surrogate. (a) Snapshots of the original data and a corresponding binned Flockwork simulation at two different times. (b) Edge activity plot. Again, the surrogate explores more edges which are not seen in the original data and does not produce contacts which are broken and immediately rebuilt. In turn, the general shape of how those new contacts are explored corresponds as well as the varied density. (c) The aggregated surrogate network is homogeneous and not “complex” as its original counterpart. (d) Temporal distributions do not generate the heavy tails observed in the real data. Again, binning the data produces a heavy tail in group-size distribution and increases correspondence in the degree distribution. Binning further increases the probability of medium-length (inter-)contact life-times. Theoretical predictions Eqs. (8.21)-(8.25) (straight lines) and non-binned data (grey circles) are approximately corresponding.

Fig. 8.20) while not strongly changing the shape of the degree distribution. This seems to be true for HT09 and HS13. However, the comparison of degree and group-size distribution of DTU data and binned Flockwork in Fig. 8.17 suggests that binning cannot be the only effect responsible for heavy-tailed group size distributions in this experiment. The larger amount of nodes with higher degrees compared to the binned Flockwork suggests that groups in the original data are indeed larger than in the simulation. It seems that in some single snapshots, the group-size distribution is bimodal with some very large very dense groups and many single nodes. This behavior cannot be replicated with the Flockwork model.

Changing focus to the temporal observables one first notices that the distributions of all unbinned simulations flatten off for short duration, while binning to a fixed time-scale increases the amount of short medium-length durations which leads to distributions which

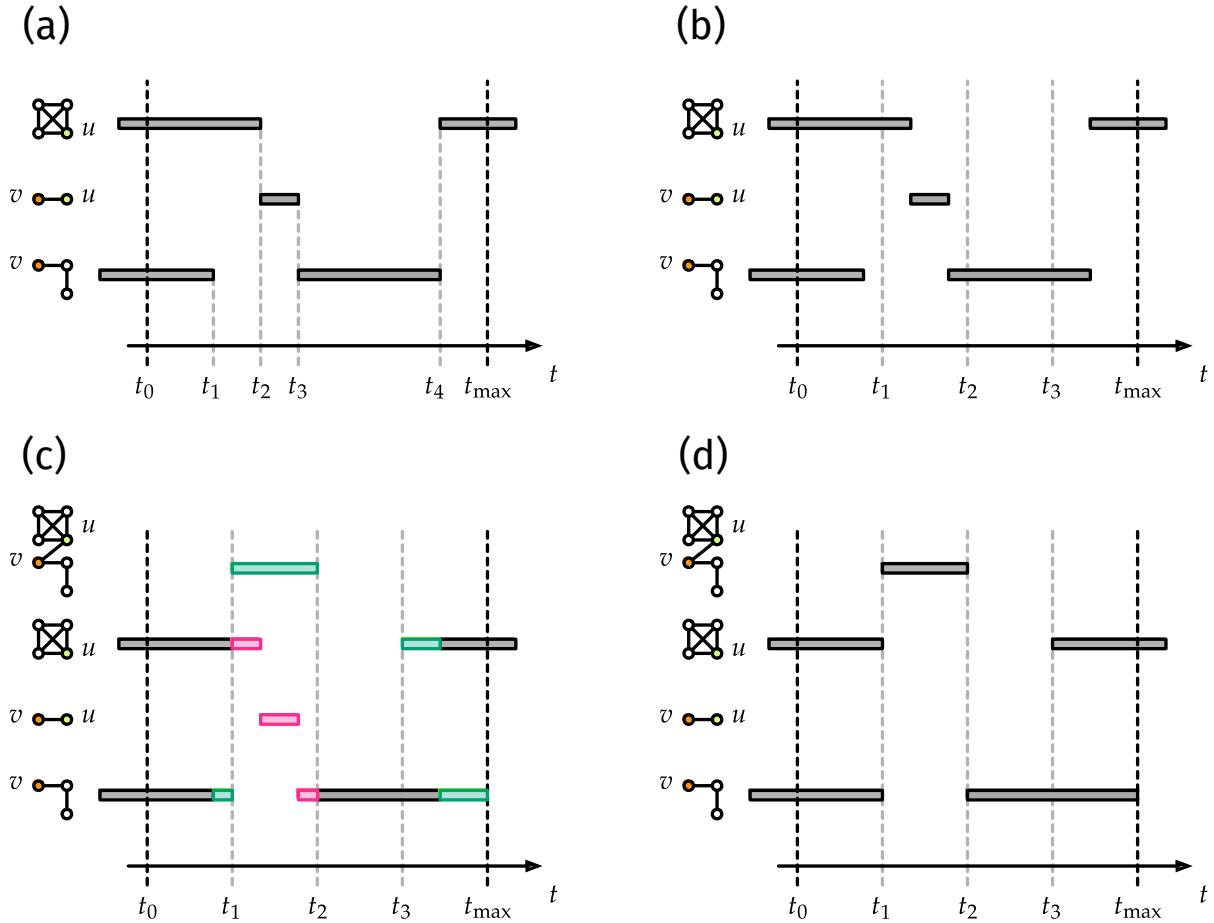


Figure 8.20: The effect of binning on measuring groups. (a) An example temporal network recorded in continuous time displayed as existing groups over time instead of edges. (b) Time is discretized. (c) Edges are binned which leads to certain groups not existing but being merged to new groups. (d) The resulting network contains a large group which did not exist before.

might be mistaken for broad distributions when only analyzing shorter times. However, the tails of the binned Flockwork distributions correspond to the tails of the unbinned Flockwork distributions which seem to decay exponentially as opposed to their real-world counterparts which show heavy tails. It thus seems that time-varying rates are not sufficient to explain heavy-tails in the temporal distributions. Qualitative exceptions are observed in the contact and inter-contact time distributions which contain some larger times in HS13 and HT09. Considering the inter-contact time, indeed, nodes will not interact at night, thus producing those larger inter-contact times. The group life-time distributions do not show strong heavy tails as the original data does, however, binning seems to be the cause for these distributions having similar shape in HT09 and HS13 (see Fig. 8.21). Furthermore, the maximum observed life time of a group decreases with increasing group size, similar to the original data. Note that the group-life time distributions of the DTU data indeed change shape with increasing group size even when binned, as opposed to the distributions obtained from the DTU experiment.

The simulations show that some of the properties of real-world face-to-face systems can be reproduced by Flockwork simulations

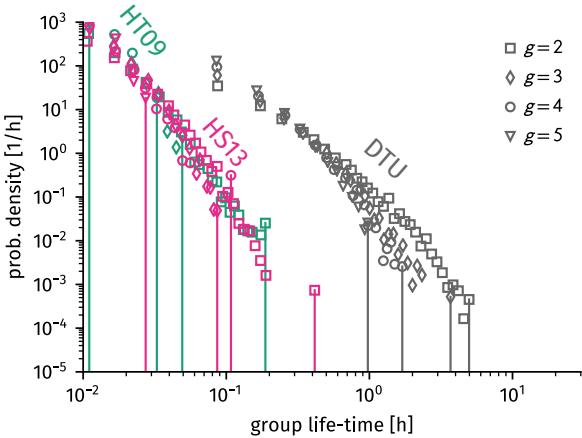


Figure 8.21: Comparison of binned Flockwork group life-time distributions. Both the SocioPatterns surrogates correspond and their shape does not vary significantly with group size. However, the maximal group life-time does indeed decrease with increasing group size g . For the DTU data, the shape of the group life-time distribution varies with group size.

while others cannot. This will be summarized and discussed in detail in the following section.

8.3 Summary and Discussion

In this chapter, a simple Poissonian model called “Flockwork” was introduced, which simulates temporal networks in continuous time. This model entails nodes switching between groups and being alone by events which happen with a constant active reconnection rate and a constant active disconnection rate. It was shown that regardless of the rate choices as well as the initial condition, a statistical equilibrium state will be reached in which the network consists of fragmented completely connected groups. An algorithm was suggested to generate configurations from the equilibrium ensemble. The group-size distribution of this equilibrium state can be computed, as well as the degree distribution, inter-contact duration distribution, group life-time distributions and an estimated contact duration distribution.

Real-world temporal networks, however, are not found to be in any stable equilibrium, which was shown for exemplary data sets from the Copenhagen network study (DTU) [74, 68] and the SocioPatterns experiment (HT09 [29] and HS13[49]). Instead, their activity and density varies over time in a circadian manner and their structural and temporal properties are distributed with heavy tails. Existing theories [78, 85] attribute these heavy tails to memory effects and age-dependent activity rates while the circadian nature of the data is ignored. Memory effects, however, complicate the analysis of temporal systems and as such one might wonder how necessary the consideration of these memory effects is to explain the heterogeneous distributions. In this chapter the hypothesis was explored whether circadian time-varying activity rates are responsible for the heavy tails in the network’s structural and temporal properties, ignoring memory effects. To this end, a method to infer time-varying activity rates from real-world data was developed first, which showed that if real-world systems followed Flockwork models with varying rates,

they would indeed be expressible to follow trajectories of equilibrium states in a newly found two-dimensional state space. One axis of this state space is given by the global active reconnection rate per node which quantifies the average number of active reconnections a single node undergoes per unit time. The second dimension of this state space is quantified by the expected mean degree of the network which directly translates to a group size distribution and corresponds to the network density. For HT09 and HS13 it was found that these trajectories approximately correspond to each other and vary between states where the network is sparse and changes slowly and states where the network is denser and changes faster, ignoring regions where the network is dense but changes slow or is sparse but changes fast. The DTU data seemed to take a shifted, more cluttered trajectory. However, it was shown that the HT09/HS13 systems displayed similar trajectories to the DTU data when rebinned to time bins of 5 min length and considering pure proximity data instead of face-to-face data. This suggests that all the systems analyzed here behave similarly over the course of a few days. Interpreting real-world human contact networks to be embedded in such a low-dimensional state space offers a new perspective and could help understand how diseases spread in face-to-face systems at different times, when the system is in different equilibrium states.

Simulating Flockworks with these inferred time-varying rates it was further shown that heavy-tailed group-size distributions are produced, which is largely attributed to the experimental practice of binning data. Before data was binned, the group-size distributions decayed faster than heavy-tailed. This insight was seemingly previously unexplored since previous models of temporal networks were either simulated in discrete time or analyzed unbinned. Hence, this insight might have implications when recording and analyzing contact data in future experiments. In order to find out whether group sizes actually decay with a heavy-tail, continuous-time measurements could be a key to the answer. In at least one of the experiments, though, the DTU experiment, the large-degree skewed degree distribution suggests that aggregated groups might indeed be distributed with a heavy-tail even in non-binned data. Aggregating the network showed that no complex structure is seen in the aggregated static social networks: nodes and links (representing total amount of time spent together) are homogeneously distributed, as opposed to the aggregated structures from the original data which showed link and node heterogeneity. This is not surprising as node event rates and connection probabilities are completely homogeneous in the Flockwork model even for time-varying rates.

Analyzing the temporal distributions it was found that binning data skews all observed short durations to the mid-length regions. However, heavy tails are not observed. This lack of heavy tails might be attributed to the node-homogeneity of the Flockwork model. Each node undergoes dis- and reconnection events with the same rate, hence, mixing will be relatively fast and contacts will be built and

decay on a homogeneous time-scale during a certain time of constant activity rates. Hence, introducing node-specific dis- and reconnection rates whose averages follow the inferred rates over time might introduce heavy tails in contact and group-life time distributions. This would be induced by long-lasting contacts from nodes who connect once but are reluctant to disconnect due to their reduced activity rates. Second, the inter-contact durations might be influenced by dropping contact likelihood homogeneity. In the current model, nodes are equally likely to connect to every other node. This is, in fact, an unrealistic assumption: humans are more likely to establish contacts to people they already interacted with. Reflecting this in simulations respecting link heterogeneity, a node which has a low node-specific connection rate will not establish any contacts and additionally it might not get connected to by another node, presumably leading to higher inter-contact times. One way to achieve link heterogeneity in a system is certainly given by considering memory kernels which control that nodes reconnect only to nodes with which they had already interacted. However, this might be far fetched to simulate data like DTU, HT09, and HS13, where experiments measure systems with an already established social network: recorded people usually already know other nodes and hence contacts are not produced from memory in observed data but rather enforced from an established static social network (memory from pre-experimental data). Another method to reflect this in a model would thus be to force nodes to reconnect using a pre-defined static social network, which might further reproduce this static social network in the aggregated limit. However, contacts in real-world data still have the property to be relatively likely to be reestablished in short time once they are broken off [77], which would not happen when enforcing an externally given social network. Memory seems unavoidable to reflect this network property. However, one network property which was not discussed yet is that human people are operating in space. Hence, people who break off contacts are very likely to reconnect to people which are in higher proximity which might result in the described effect without introducing explicit memory kernels. Indeed, an already popular model generates temporal networks by letting nodes perform a random walk in two-dimensional space and connecting nodes based on proximity [17] which can reproduce heterogeneous group-size distribution and contact durations, however, circadian activity is once again ignored.

All these hypotheses might be investigated in future studies to be responsible to produce observed statistics once incorporated into the Flockwork model. However, a model's value is not rigorously tied to its ability to reproduce every aspect observed in real data. Rather, simpler models which are known to reproduce certain features of real-world systems but ignore others are useful to single out structural and temporal properties responsible for the outcome of certain dynamics. For instance, real-world static networks, while clustered, often have central well-connected nodes, too. These so-called "hubs"

highly increase the speed of diffusion and thus of distribution. However, in the first part of this thesis it was shown that a variety of clustered network models with homogeneous degree distributions can already speed up diffusion when introducing a small number of long-range links, showing that long path lengths and clustering are responsible for slow diffusion time-scales and it is not necessary to have heterogeneous node properties. In a similar way, the Flockwork model with constant rates may be used in the future to study the effect of group-sized topology on the outcome of dynamic processes such as diffusion or disease spreading, restricting the analyses to a homogeneous time-scale to learn about the effect of group structures only. Further using the time-varying rates, the influence of circadian activity can be studied, without effects imposed by node or link heterogeneity. In the following chapter, one such analysis will be done: the static-rate Flockwork model will be investigated to study the effects of group topologies on the spread of a susceptible-infected-susceptible disease dynamic when the model is in different points in the (α, k_0) -plane.

9

The SIS-Model on Temporal Contact Networks

9.1 General Considerations

The ratio I^* of infected nodes in equilibrium of an SIS process is the observable of interest of this chapter. Its functional dependence on both varying group-size distribution as well as active reconnection rate α is investigated in the following using Gillespie simulations.

In this chapter,

$$k_0 = \frac{1}{T} \int_0^T dt \sum_{i=1}^N k_i(t)$$

is the average mean degree. Since it is both directly mappable to a Flockwork equilibrium configuration with a certain group-size distribution as well as to the network density in an edge activity model, it will be chosen as the parameter controlling the network structure.

The time scale will be fixed by keeping the active reconnection rate α of a Flockwork model constant for all situations, while the recovery rate per node ρ is varied. The control parameter

$$c_0 = \frac{\alpha}{\rho}$$

then describes the number of active reconnections a node performs before it recovers, on average. This parameter will hence be called the *contact-renewal number* and controls the temporal change of the network in comparison to the time scale of the disease. For the edge activity model, the temporal control parameter corresponding to the active reconnection rate α of the Flockwork model is given by the global link decay rate ω^- which is determined by Eq. (8.14).

Further, one is interested in keeping the infectiousness of the diseases comparable while varying the network structure. As discussed in Sec. , a non-severe disease (i.e. with small infection rate η) will die out on a sparse network while it might persist on a densely connected network. Hence, the basic reproduction number

$$R_0 = k_0 \frac{\eta}{\rho}$$

will be the control parameter of the infectiousness, correcting for an increased number of edges in the system.

In total, the ratio I^* of infected nodes in equilibrium will be studied as a function of the structural control parameter k_0 (the average mean degree), the temporal control parameter c_0 (the contact renewal number) and the disease control parameter R_0 (basic reproduction number), such that

$$I^* = I^*(k_0, c_0, R_0).$$

For $k_0 \leq 1$ both the edge activity model as well as the Flockwork model produce network configurations composed of disconnected groups. When increasing the average mean degree to values $k_0 > 1$ only the Flockwork model produces group size distributions while the edge activity model possesses a giant component. Hence, SIS-simulations will be studied on both models to investigate how the disconnected state of temporal face-to-face contacts influences the disease spreading process on different time scales.

A few considerations can be done already. If $c_0 \rightarrow \infty$, an infected node re-connects to a high number of other nodes before it recovers. Hence, it is expected that independent of the structure of the network, the number of infected will follow the mean-field result

$$I^*(k_0, c_0 \rightarrow \infty, R_0) = 1 - \frac{1}{R_0}$$

because in both the Flockwork as well as the edge activity model, each link is equally likely to be activated when considering a large observation time. For real-world networks this corresponds to a situation where modeling the spread on the aggregated network will be appropriate. When $c_0 \ll 1$, the network will appear static in contrast to the evolving disease, i.e. it might suffice to study the spread on an Erdős-Rényi model of degree k_0 for the edge activity model or a number disconnected components for the Flockwork model.

In the case of $k_0 \rightarrow N - 1$, the network will for both models approach a completely connected state and hence

$$I^*(k_0 \rightarrow \infty, c_0, R_0) = 1 - \frac{1}{R_0}.$$

An observable of interest is furthermore the epidemic threshold. Usually this refers to the critical basic reproduction number necessary to sustain the disease within the population. Since here, the ratio I^* of infected in equilibrium takes a functional form depending on k_0 , c_0 and R_0 , each of those can be “critical” when the other two are fixed. In general, the threshold will be given as the surface defined by

$$I^*(k_0, c_0, R_0) = 0.$$

In the following, these situations will be investigated numerically, using Gillespie simulations on the Flockwork model and the edge activity model.

9.2 Simulations

9.2.1 Edge Activity Model

Simulations to obtain the ratio I^* of infected nodes in equilibrium on an edge activity model have been performed as described in App. D. The results of these simulations are shown in Fig. 9.2. A first result of this analysis is that all curves are describable using a functional form of

$$I^* = A - \frac{B}{(R_0 - D)^C}. \quad (9.1)$$

which means that for fixed k_0 and c_0 , the critical reproduction number $R_{0,c}$ can be computed as

$$R_{0,c} = \left(\frac{A}{B}\right)^{-1/C} + D. \quad (9.2)$$

Here, the parameters A , B , C , and D are found using a least-squares fit with initial values $A = 1$, $B = 1$, $C = 1$, and $D = 0$ as these are the corresponding parameters for the mean-field response. For fixed k_0 and c_0 , the fits were performed on all simulation data pairs of R_0 and I^* where $I^* > 0$. The goodness of fit was tested by recognizing the following. Suppose that the data is indeed describable using Eq. (9.1) with the parameters found by the fit process. If this is true, one may rescale the data of any simulation using coordinates $x = (R_0 - D)/B^{1/C}$ and $y = (A - I^*)^{1/C}$ (where A , B , C and D are the corresponding fit parameters found for fixed k_0 and c_0). Then $y = x^{-1}$ and as such, plotting $\log x$ against $\log y$, all data should collapse to a straight line falling with slope -1 . As can be seen in Fig. , this is indeed the case. Consequently the epidemic threshold as defined by the critical basic reproduction number can be found using Eq. (9.2). The response curves and interpreted fit parameters are discussed in the following.

One first recognizes that for smaller values of $k_0 \lesssim 1$ and $c_0 \lesssim 1$ the network is sparse or an infected node might not reach many other nodes before recovering, respectively. As such, the maximum number of nodes which can be infected is limited by the parameter A which decreases with decreasing k_0 and c_0 . This implies that even an infinitely infectious disease will not be able to infect the whole network for these finite numbers of c_0 and k_0 , which will be further discussed in Sec. . Second, with $B > 1$, this further implies that the critical basic reproduction number will shift to higher values.

9.2.2 Poissonian Flockwork Model

Simulations to obtain the ratio I^* of infected nodes in equilibrium have been performed as described in App. D. The results of this exercise may be seen in Fig. .

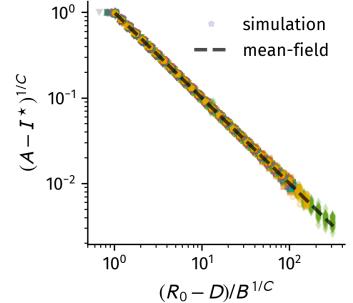


Figure 9.1: Non-zero ratio of infected in equilibrium I^* as rescaled with the fit parameters of Eq. (9.1). Markers represent all data from the simulations shown in Fig. 9.2. The parameters A , B , C , and D change with varying mean degree k_0 and contact-renewal c_0 . For mean-field we have $I^* = 1 - 1/R_0$ and as such $A = 1$, $B = 1$, $C = 1$, and $D = 0$.

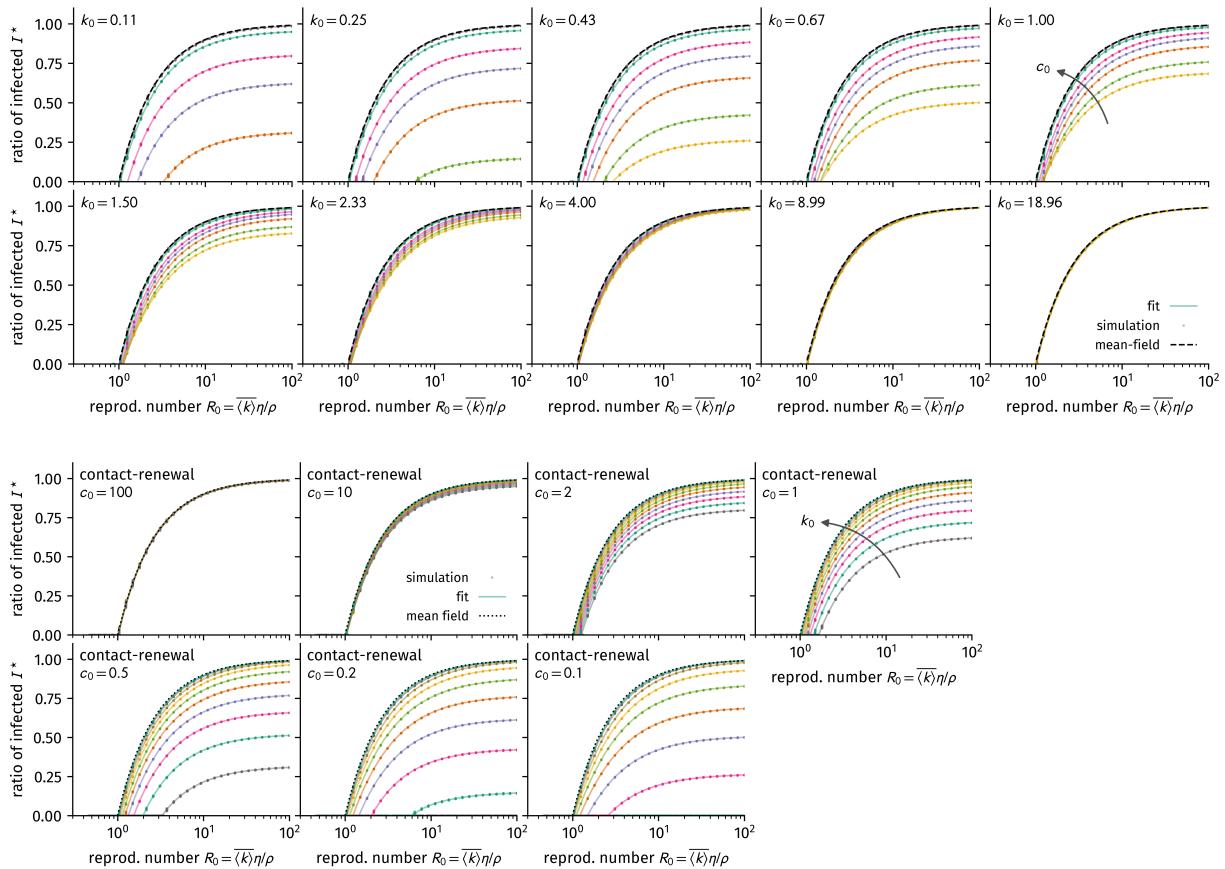


Figure 9.2: Ratio I^* of infected in equilibrium for SIS simulations on the edge activity model. Here, the mean degree k_0 , contact-renewal number $c_0 = \alpha/\rho$, and the basic reproduction number R_0 are varied. Shown are simulation results as well as fits using Eq. (9.1).

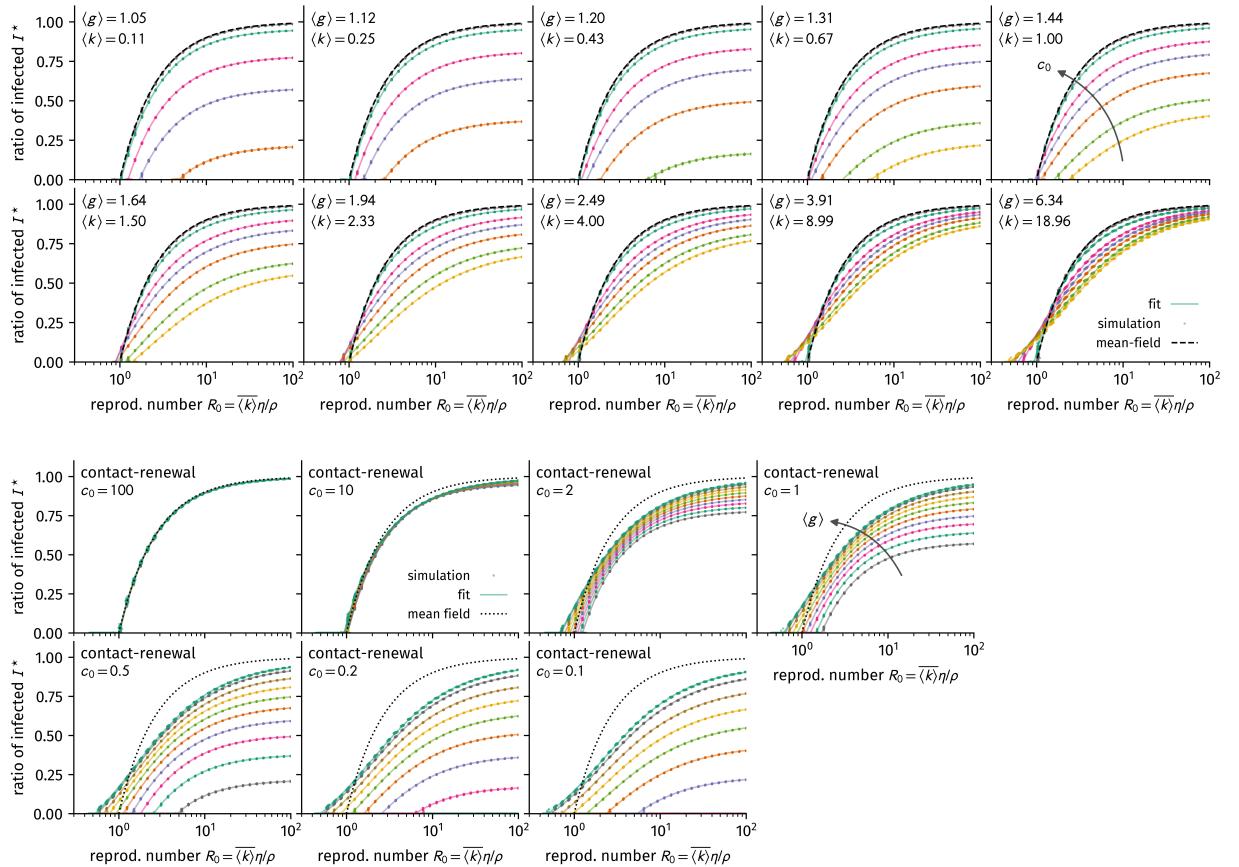


Figure 9.3: Ratio I^* of infected in equilibrium for SIS simulations on the Flockwork model. Here, the mean degree k_0 , contact-renewal number $c_0 = \alpha/\rho$, and the basic reproduction number R_0 are varied. Shown are simulation results as well as fits using Eq. (9.1).

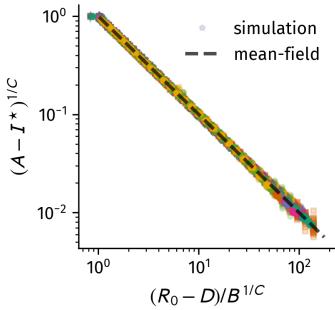


Figure 9.4: Non-zero ratio of infected in equilibrium I^* for the Flockwork model as rescaled with the fit parameters of Eq. (9.1). Markers represent all data from the simulations shown in Fig. 9.3. The parameters A , B , C , and D change with varying mean degree k_0 and contact-renewal c_0 . For mean-field we have $I^* = 1 - 1/R_0$ and as such $A = 1$, $B = 1$, $C = 1$, and $D = 0$.

9.2.3 Discussion

9.3 The Case of Infinite Infectiousness

Here, a special case will be considered for infinitely infectious diseases $\eta \rightarrow \infty$. That means whenever an infected node connects to a susceptible node, it will infect every susceptible in a group before any other structural event may happen. This way, any reconnection event between an infected and any group of susceptibles is likewise an infection event (as well as every reconnection event between a susceptible and group of infected, but less severe since only the reconnected susceptible node becomes infected). Hence, the probabilities for group size evolution of infected groups by reconnection only within infected groups are analogous to the evolution before, with the probabilities

$$\begin{aligned}
P_1 &\equiv P[|g(u)| = m, |g(v)| = m, g(u) \neq g(v)] = P \frac{mI_m}{N} \times \frac{m(I_m - 1)}{N - 1} \\
P_2 &\equiv P[|g(u)| = m - 1, |g(v)| = m - 1, g(u) \neq g(v)] = P \frac{(m - 1)I_{m-1}}{N} \times \frac{(m - 1)(I_{m-1})}{N - 1} \\
P_3 &\equiv P[|g(u)| = m + 1, |g(v)| = m + 1, g(u) \neq g(v)] = P \frac{(m + 1)I_{m+1}}{N} \times \frac{(m + 1)(I_{m+1})}{N - 1} \\
P_4 &\equiv P[|g(u)| = m, |g(v)| \neq m, |g(v)| \neq m - 1] = P \frac{mI_m}{N} \times \left(\frac{I - 1}{N - 1} - \frac{mI_m - 1}{N - 1} \right) \\
P_5 &\equiv P[|g(u)| = m + 1, |g(v)| \neq m + 1, |g(v)| \neq m] = P \frac{(m + 1)I_{m+1}}{N} \times \left(\frac{I - 1}{N - 1} - \frac{m}{N} \right) \\
P_6 &\equiv P[|g(u)| \neq m - 1, |g(u)| \neq m, |g(v)| = m - 1] = P \left(\frac{I}{N} - \frac{(m - 1)I_{m-1}}{N} - \frac{mI_m}{N} \right) \\
P_7 &\equiv P[|g(u)| \neq m + 1, |g(u)| \neq m, |g(v)| = m] = P \left(\frac{I}{N} - \frac{(m + 1)I_{m+1}}{N} - \frac{mI_m}{N} \right) \\
P_8 &\equiv P[|g(u)| = m] = (1 - P) \frac{mI_m}{N} \\
P_9 &\equiv P[|g(u)| > 1] = (1 - P) \frac{1}{N} (m + 1) I_{m+1}.
\end{aligned}$$

The group size evolution is happening in the same way as before. Now there's also infection-susceptible couplings that can happen

1. The first chosen node is in group $g(u) = m$ and infected. it chooses a susceptible group of size $g(v) = n \neq m - 1$. The change in infected group sizes is

$$\Delta I_m = -1 \quad \Delta I_{n+1} = +1 \quad \Delta I_{m-1} = +1$$

2. The first chosen node is in group $g(u) = m$ and infected. it chooses a susceptible group of size $g(v) = m - 1$. The change in infected group sizes is

$$\Delta I_{m-1} = +1$$

3. The first chosen node is in group $g(u) = n$ and susceptible. It chooses an infected group of size $g(v) = m$. The change in infected group sizes is

$$\Delta I_m = -1 \quad \Delta I_{m+1} = +1$$

This implies

$$\begin{aligned} \Delta I_m / \Delta t = & P[|g(u)| = m, |g(v)| = m, g(u) \neq g(v)] \times (-2) \\ & + P[|g(u)| = m-1, |g(v)| = m-1, g(u) \neq g(v)] \times (+1) \\ & + P[|g(u)| = m+1, |g(v)| = m+1, g(u) \neq g(v)] \times (+1) \\ & + P[|g(u)| = m, |g(v)| \neq m, |g(v)| \neq m-1] \times (-1) \\ & + P[|g(u)| = m+1, |g(v)| \neq m, |g(v)| \neq m+1] \times (+1) \\ & + P[|g(u)| \neq m, |g(u)| \neq m-1, |g(v)| = m-1] \times (+1) \\ & + P[|g(u)| \neq m+1, |g(u)| \neq m, |g(v)| = m] \times (-1) \\ & + P[|g(u)| = m] \times (-1) \\ & + P[|g(u)| > 1] \times (+1) \\ & + P[u \in \mathcal{I}, |g(u)| = m, v \notin \mathcal{I}, g(v) \neq m-1] \times (-1) \\ & + P[u \in \mathcal{I}, |g(u)| = m+1, v \notin \mathcal{I}] \times (+1) \\ & + P[u \in \mathcal{I}, |g(u)| \neq m, v \notin \mathcal{I}, g(v) = m-1] \times (+1) \\ & + P[u \notin \mathcal{I}, v \in \mathcal{I}, g(v) = m] \times (-1) \\ & + P[u \notin \mathcal{I}, v \in \mathcal{I}, g(v) = m-1] \times (+1) \end{aligned}$$

with

$$\begin{aligned} P[u \in \mathcal{I}, |g(u)| = m, v \notin \mathcal{I}, g(v) \neq m-1] &= P \frac{mI_m}{N} \left(\frac{N-I}{N-1} - \frac{(m-1)(N_{m-1} - I_{m-1})}{N-1} \right) \\ P[u \in \mathcal{I}, |g(u)| = m+1, v \notin \mathcal{I}] &= P \frac{(m+1)I_{m+1}}{N} \left(\frac{N-I}{N-1} \right) \\ P[u \in \mathcal{I}, |g(u)| \neq m, v \notin \mathcal{I}, g(v) = m-1] &= P \left(\frac{I-mI_m}{N} \right) \frac{(m-1)(N_{m-1} - I_{m-1})}{N-1} \\ P[u \notin \mathcal{I}, v \in \mathcal{I}, g(v) = m] &= P \left(\frac{N-I}{N} \right) \frac{mI_m}{N-1} \\ P[u \notin \mathcal{I}, v \in \mathcal{I}, g(v) = m-1] &= P \left(\frac{N-I}{N} \right) \frac{(m-1)I_{m-1}}{N-1}. \end{aligned}$$

$$\begin{aligned}
P_1 &= -P \frac{N-I}{N} \frac{mI_m}{N-1} \\
P_2 &= +P \frac{N-I}{N} \frac{(m-1)I_{m-1}}{N-1} \\
P_3 &= +2P \frac{(m+1)I_{m+1}}{N} \frac{(m-1)(N_{m-1} - I_{m-1})}{N-1} \\
P_4 &= +P \frac{(m+1)I_{m+1}}{N} \left(\frac{N-I - (m-1)(N_{m-1} - I_{m-1})}{N-1} \right) \\
P_5 &= -P \frac{mI_m}{N} \left(\frac{N-I - (m-1)(N_{m-1} - I_{m-1})}{N-1} \right) \\
P_6 &= +P \left(\frac{I - mI_m - (m+1)I_{m+1}}{N} \right) \frac{(m-1)(N_{m-1} - I_{m-1})}{N-1}
\end{aligned}$$

Additionally, the group of infected loners can recover with rate ρ , such that

$$\partial_t I_1^{\text{rec}}(t) = -\rho \frac{I_1}{N}.$$

So the final equations sum up to

$$\begin{aligned}
\partial_t I_1 &= - \underbrace{\rho I_1 / N}_{\text{recovering}} + \underbrace{\frac{\gamma}{N}(1-P)(I - I_1 + 2I_2)}_{\text{nodes leaving infected groups}} + \\
&\quad + 2 \underbrace{\frac{\gamma P}{N(N-1)} [I_2(I-2) - I_1(I-1)]}_{\text{reconnections between infected}} + \\
&\quad + \underbrace{2 \frac{\gamma P}{N(N-1)} (N-I) [I_2 - I_1]}_{\text{reconnections between infected and susceptibles}} \\
\partial_t I_{m>1} &= \frac{\gamma P}{N(N-1)} \left[(m-1)I_{m-1}(I - (m-1)) + \right. \\
&\quad \left. + (m+1)I_{m+1}(I - (m+1)) - 2mI_m(I - m) \right] + \\
&\quad + \frac{\gamma}{N}(1-P) [(m+1)I_{m+1} - mI_m] + \\
&\quad + \frac{\gamma P}{N(N-1)} \left[I(m-1)(n_{m-1} - I_{m-1}) + \right. \\
&\quad \left. (N-I)(-2mI_m + (m+1)I_{m+1} + (m-1)I_{m-1}) \right].
\end{aligned}$$

9.4 Spreading on Real-World Temporal Networks

The

9.5 Discussion

10

Summary & Outlook

At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, At accusam aliquyam diam diam dolore dolores duo eirmod eos erat, et nonumy sed tempor et et invidunt justo labore Stet clita ea et gubergren, kasd magna no rebum. sanctus sea sed takimata ut vero voluptua. est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur

Bibliography

- [1] Christopher Aicher, Abigail Z. Jacobs, and Aaron Clauset. "Learning latent block structure in weighted networks". In: *Journal of Complex Networks* 3.2 (June 2015), pp. 221–248. ISSN: 2051-1310. DOI: 10.1093/comnet/cnu026. URL: <https://academic.oup.com/comnet/article-abstract/3/2/221/375946/Learning-latent-block-structure-in-weighted> (visited on 04/05/2017).
- [2] Alex Arenas, Albert Díaz-Guilera, and Conrad J. Pérez-Vicente. "Synchronization reveals topological scales in complex networks". In: *Physical Review Letters* 96.11 (2006), p. 114102. DOI: 10.1103/PhysRevLett.96.114102. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.96.114102> (visited on 02/10/2016).
- [3] Ulf Aslak, Martin Rosvall, and Sune Lehmann. "Constrained information flows in temporal networks reveal intermittent communities". In: *Physical Review E* 97.6 (June 2018), p. 062312. DOI: 10.1103/PhysRevE.97.062312. URL: <https://link.aps.org/doi/10.1103/PhysRevE.97.062312> (visited on 03/11/2019).
- [4] Kendall E. Atkinson. *An introduction to numerical analysis*. 2nd. New York: Wiley, 1989. ISBN: 978-0-471-62489-9.
- [5] A. Barrat, C. Cattuto, A. E. Tozzi, P. Vanhems, and N. Voirin. "Measuring contact patterns with wearable sensors: methods, data characteristics and applications to data-driven simulations of infectious diseases". English. In: *Clinical Microbiology and Infection* 20 (Jan. 2014), pp. 10–16. DOI: 10.1111/1469-0691.12472.
- [6] Alain Barrat, Marc Barthelemy, and Alessandro Vespignani. *Dynamical processes on complex networks*. OCLC: ocn231581094. Cambridge, UK ; New York: Cambridge University Press, 2008. ISBN: 978-0-521-87950-7.
- [7] Vladimir Batagelj and Ulrik Brandes. "Efficient generation of large random networks". In: *Physical Review E* 71.3 (2005), p. 036113. DOI: 10.1103/PhysRevE.71.036113. URL: <http://link.aps.org/doi/10.1103/PhysRevE.71.036113> (visited on 02/09/2015).
- [8] Howard C. Berg. *Random walks in biology*. Princeton, NJ: Princeton University Press, 1993. ISBN: 978-0-691-00064-0.

- [9] Ciro Cattuto, Wouter Van den Broeck, Alain Barrat, Vittoria Colizza, Jean-François Pinton, and Alessandro Vespignani. "Dynamics of Person-to-Person Interactions from Distributed RFID Sensor Networks". en. In: *PLOS ONE* 5.7 (July 2010), e11596. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0011596. URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0011596> (visited on 03/12/2019).
- [10] Aaron Clauset, Cristopher Moore, and M. E. J. Newman. "Hierarchical structure and the prediction of missing links in networks". en. In: *Nature* 453.7191 (2008), pp. 98–101. ISSN: 0028-0836. DOI: 10.1038/nature06830. URL: <http://www.nature.com/nature/journal/v453/n7191/full/nature06830.html> (visited on 01/21/2015).
- [11] Vittoria Colizza, Romualdo Pastor-Satorras, and Alessandro Vespignani. "Reaction-diffusion processes and metapopulation models in heterogeneous networks". en. In: *Nature Physics* 3.4 (Apr. 2007), pp. 276–282. ISSN: 1745-2473. DOI: 10.1038/nphys560. URL: <https://www.nature.com/nphys/journal/v3/n4/abs/nphys560.html> (visited on 04/05/2017).
- [12] Colin Cooper and Alan Frieze. "The cover time of sparse random graphs". en. In: *Random Structures and Algorithms* 30.1-2 (Jan. 2007), pp. 1–16. ISSN: 10429832, 10982418. DOI: 10.1002/rsa.20151. URL: <http://doi.wiley.com/10.1002/rsa.20151> (visited on 04/10/2017).
- [13] Colin Cooper and Alan Frieze. "The cover time of the preferential attachment graph". In: *Journal of Combinatorial Theory, Series B* 97.2 (2007), pp. 269–290. ISSN: 0095-8956. DOI: 10.1016/j.jctb.2006.05.007. URL: <http://www.sciencedirect.com/science/article/pii/S0095895606000840> (visited on 04/10/2017).
- [14] R. Cross and A. Parker. *The hidden power of social networks*. Boston, MA: Harvard Business School Press, 2004.
- [15] Jesper Dall and Michael Christensen. "Random geometric graphs". In: *Physical Review E* 66.1 (July 2002), p. 016121. DOI: 10.1103/PhysRevE.66.016121. URL: <https://link.aps.org/doi/10.1103/PhysRevE.66.016121> (visited on 01/24/2019).
- [16] Hassan Ali Dawah, Bradford A. Hawkins, and Michael F. Claridge. "Structure of the Parasitoid Communities of Grass-Feeding Chalcid Wasps". In: *Journal of Animal Ecology* 64.6 (1995), pp. 708–720. ISSN: 0021-8790. DOI: 10.2307/5850. URL: <http://www.jstor.org/stable/5850> (visited on 04/05/2017).
- [17] Josep Diaz, Dieter Mitsche, and Xavier Perez. "Dynamic Random Geometric Graphs". In: *arXiv:cs/0702074* (Feb. 2007). arXiv: cs/0702074. URL: <http://arxiv.org/abs/cs/0702074> (visited on 09/23/2018).

- [18] Manlio De Domenico, Albert Solé-Ribalta, Sergio Gómez, and Alex Arenas. "Navigability of interconnected networks under random failures". en. In: *Proceedings of the National Academy of Sciences* 111.23 (June 2014), pp. 8351–8356. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1318469111. URL: <http://www.pnas.org/content/111/23/8351> (visited on 08/10/2017).
- [19] Manfred Eigen, John McCaskill, and Peter Schuster. "Molecular quasi-species". In: *J. Phys. Chem.* 92.24 (1988), pp. 6881–6891. DOI: 10.1021/j100335a010. URL: <http://dx.doi.org/10.1021/j100335a010>.
- [20] Robert Elsässer and Thomas Sauerwald. "Tight bounds for the cover time of multiple random walks". In: *Theoretical Computer Science*. Selected Papers from 36th International Colloquium on Automata, Languages and Programming (ICALP 2009) 412.24 (May 2011), pp. 2623–2641. ISSN: 0304-3975. DOI: 10.1016/j.tcs.2010.08.010. URL: <http://www.sciencedirect.com/science/article/pii/S0304397510004275> (visited on 09/12/2017).
- [21] Paul Erdős and Albert Rényi. "On random graphs I". In: *Publicationes Mathematicae (Debrecen)* 6 (1959), pp. 290–297.
- [22] E. N. Gilbert. "Random Graphs". EN. In: *The Annals of Mathematical Statistics* 30.4 (Dec. 1959), pp. 1141–1144. ISSN: 0003-4851, 2168-8990. DOI: 10.1214/aoms/1177706098. URL: <https://projecteuclid.org/euclid.aoms/1177706098> (visited on 02/06/2019).
- [23] Daniel T. Gillespie. "Exact stochastic simulation of coupled chemical reactions". In: *The Journal of Physical Chemistry* 81.25 (1977), pp. 2340–2361. ISSN: 0022-3654. DOI: 10.1021/j100540a008. URL: <http://dx.doi.org/10.1021/j100540a008> (visited on 07/01/2016).
- [24] M. Girvan and M. E. J. Newman. "Community structure in social and biological networks". In: *Proceedings of the National Academy of Sciences of the United States of America* 99.12 (June 2002), pp. 7821–7826. ISSN: 0027-8424. DOI: 10.1073/pnas.122653799. URL: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC122977/> (visited on 04/05/2017).
- [25] David Halliday, Robert Resnick, Jearl Walker, and Stephan W. Koch. *Physik*. ger. 1. korrig. Nachdr. OCLC: 297583264. Weinheim: Wiley-VCH, 2005. ISBN: 978-3-527-40366-0.
- [26] Petter Holme. "Modern temporal network theory: a colloquium". en. In: *The European Physical Journal B* 88.9 (Sept. 2015), p. 234. ISSN: 1434-6036. DOI: 10.1140/epjb/e2015-60657-4. URL: <https://doi.org/10.1140/epjb/e2015-60657-4> (visited on 03/05/2019).

- [27] Petter Holme and Jari Saramäki. "Temporal networks". In: *Physics Reports*. Temporal Networks 519.3 (2012), pp. 97–125. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2012.03.001. URL: <http://www.sciencedirect.com/science/article/pii/S0370157312000841> (visited on 02/24/2016).
- [28] Flavio Iannelli, Andreas Koher, Dirk Brockmann, Philipp Hövel, and Igor M. Sokolov. "Effective distances for epidemics spreading on complex networks". In: *Physical Review E* 95.1 (Jan. 2017), p. 012313. DOI: 10.1103/PhysRevE.95.012313. URL: <https://link.aps.org/doi/10.1103/PhysRevE.95.012313> (visited on 05/22/2017).
- [29] Lorenzo Isella, Juliette Stehlé, Alain Barrat, Ciro Cattuto, Jean-François Pinton, and Wouter Van den Broeck. "What's in a crowd? Analysis of face-to-face behavioral networks." English. In: *Journal of Theoretical Biology* 271 (Feb. 2011), pp. 166–180. DOI: 10.1016/j.jtbi.2010.11.033.
- [30] J. Klafter and Igor M. Sokolov. *First steps in random walks: from tools to applications*. OCLC: ocn714724924. Oxford ; New York: Oxford University Press, 2011. ISBN: 978-0-19-923486-8.
- [31] Jon Kleinberg. "The Small-world Phenomenon: An Algorithmic Perspective". In: *Proceedings of the Thirty-second Annual ACM Symposium on Theory of Computing*. STOC '00. New York, NY, USA: ACM, 2000, pp. 163–170. ISBN: 978-1-58113-184-0. DOI: 10.1145/335305.335325. URL: <http://doi.acm.org/10.1145/335305.335325> (visited on 08/22/2016).
- [32] Judith Kleinfield. "Six Degrees: Urban Myth?" In: *Psychology Today* (Mar. 2002). URL: <https://www.psychologytoday.com/intl/articles/200203/six-degrees-urban-myth> (visited on 02/05/2019).
- [33] D. E. Knuth. *The Stanford GraphBase: A platform for combinatorial computing*. Reading, MA: Addison-Wesley, 1993.
- [34] Andreas Koher, Hartmut H. K. Lentz, Philipp Hövel, and Igor M. Sokolov. "Infections on Temporal Networks—A Matrix-Based Approach". en. In: *PLOS ONE* 11.4 (Jan. 2016), e0151209. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0151209. URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0151209> (visited on 03/11/2019).
- [35] Vladis Krebs. "Mapping networks of terrorist cells". In: *Connections* 24 (2002), pp. 43–52.
- [36] Lambiotte, R. *Hierarchical networks: (information) diffusion and mapping*. 2010.
- [37] Andrea Lancichinetti, Santo Fortunato, and Janos Kertesz. "Detecting the overlapping and hierarchical community structure of complex networks". In: *New Journal of Physics* 11.3 (Mar. 2009). arXiv: 0802.1218, p. 033015. ISSN: 1367-2630. DOI: 10.1088/1367-2630/11/3/033015. URL: <http://arxiv.org/abs/0802.1218> (visited on 01/21/2015).

- [38] H. W. Lau and K. Y. Szeto. "Asymptotic analysis of first passage time in complex networks". en. In: *EPL (Europhysics Letters)* 90.4 (2010), p. 40005. ISSN: 0295-5075. DOI: 10.1209/0295-5075/90/40005. URL: <http://stacks.iop.org/0295-5075/90/i=4/a=40005> (visited on 02/28/2017).
- [39] Yuan Lin, Alafate Julaiti, and Zhongzhi Zhang. "Mean first-passage time for random walks in general graphs with a deep trap". eng. In: *The Journal of Chemical Physics* 137.12 (Sept. 2012), p. 124104. ISSN: 1089-7690. DOI: 10.1063/1.4754735.
- [40] L. Lovász. "Random walks on graphs: a survey". In: *Combinatorics, Paul Erdős is Eighty*. Ed. by D. Miklós, V. T. Sós, and T. Szőnyi. Vol. 2. Budapest: János Bolyai Mathematical Society, 1996, pp. 353–398.
- [41] David Lusseau, Karsten Schneider, Oliver J. Boisseau, Patti Haase, Elisabeth Slooten, and Steve M. Dawson. "The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations". en. In: *Behavioral Ecology and Sociobiology* 54.4 (Sept. 2003), pp. 396–405. ISSN: 0340-5443, 1432-0762. DOI: 10.1007/s00265-003-0651-y. URL: <https://link.springer.com/article/10.1007/s00265-003-0651-y> (visited on 04/05/2017).
- [42] Benjamin F. Maier. "Analytic evaluation of the small-world effect on a model approaching the Erdős-Rényi random graph". In: *arXiv:1901.02381 [physics]* (Jan. 2019). arXiv: 1901.02381. URL: <http://arxiv.org/abs/1901.02381> (visited on 02/01/2019).
- [43] Benjamin F. Maier and Dirk Brockmann. "Cover time for random walks on arbitrary complex networks". In: *Physical Review E* 96.4 (Oct. 2017), p. 042307. DOI: 10.1103/PhysRevE.96.042307. URL: <https://link.aps.org/doi/10.1103/PhysRevE.96.042307> (visited on 11/01/2017).
- [44] Benjamin F. Maier, Cristián Huepe, and Dirk Brockmann. "Modular hierarchical and power-law small-world networks bear structural optima for minimal first passage times and cover time". In: *arXiv:1808.00240 [cond-mat, physics:physics]* (Aug. 2018). arXiv: 1808.00240. URL: <http://arxiv.org/abs/1808.00240> (visited on 09/25/2018).
- [45] Benjamin Frank Maier. *B.F. Maier's FB friends network*, <https://github.com/benmaier/BFMaierFBnetwork>. Repository. Apr. 2017. URL: <https://github.com/benmaier/BFMaierFBnetwork>.
- [46] Benjamin Frank Maier. *cMHRN - A C++/Python/MATLAB package to generate SSMH and PLSW networks in a fast manner*, <https://github.com/benmaier/cMHRN>. Repository. Mar. 2018. URL: <https://github.com/benmaier/cMHRN>.

- [47] Benjamin Frank Maier. *smallworld - A Python package for generating and analyzing networks from the alternative small-world model*, <https://github.com/benmaier/smallworld>. Repository. Nov. 2018. URL: <https://github.com/benmaier/smallworld>.
- [48] R. Dean Malmgren, Daniel B. Stouffer, Adilson E. Motter, and Luís A. N. Amaral. "A Poissonian explanation for heavy tails in e-mail communication". en. In: *Proceedings of the National Academy of Sciences* 105.47 (Nov. 2008), pp. 18153–18158. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.0800332105. URL: <https://www.pnas.org/content/105/47/18153> (visited on 03/13/2019).
- [49] Rossana Mastrandrea, Julie Fournet, and Alain Barrat. "Contact Patterns in a High School: A Comparison between Data Collected Using Wearable Sensors, Contact Diaries and Friendship Surveys". en. In: *PLOS ONE* 10.9 (Jan. 2015), e0136497. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0136497. URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0136497> (visited on 09/23/2018).
- [50] Naoki Masuda, Mason A. Porter, and Renaud Lambiotte. "Random walks and diffusion on networks". In: *arXiv:1612.03281 [cond-mat, physics:physics]* (Dec. 2016). arXiv: 1612.03281. URL: <http://arxiv.org/abs/1612.03281> (visited on 04/10/2017).
- [51] Sergey Melnik, Adam Hackett, Mason A. Porter, Peter J. Mucha, and James P. Gleeson. "The unreasonable effectiveness of tree-based theory for networks with clustering". In: *Physical Review E* 83.3 (Mar. 2011), p. 036112. DOI: 10.1103/PhysRevE.83.036112. URL: <https://link.aps.org/doi/10.1103/PhysRevE.83.036112> (visited on 02/05/2019).
- [52] David Meunier, Renaud Lambiotte, Alex Fornito, Karen D. Ernsche, and Edward T. Bullmore. "Hierarchical modularity in human brain functional networks". In: *Frontiers in Neuroinformatics* 3 (Oct. 2009). ISSN: 1662-5196. DOI: 10.3389/neuro.11.037.2009. URL: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2784301/> (visited on 01/22/2015).
- [53] Bojan Mohar. "Some applications of Laplace eigenvalues of graphs". en. In: *Graph Symmetry*. NATO ASI Series. Springer, Dordrecht, 1997, pp. 225–275. ISBN: 978-90-481-4885-1 978-94-015-8937-6. URL: https://link.springer.com/chapter/10.1007/978-94-015-8937-6_6 (visited on 09/13/2017).
- [54] Mark Newman. *Networks: An introduction*. New York, NY, USA: Oxford University Press, Inc., 2010. ISBN: 0-19-920665-1 978-0-19-920665-0.
- [55] Bernt Øksendal. *Stochastic differential equations: An introduction with applications*. Universitext. OCLC: 246776666. Berlin: Springer, 1992. ISBN: 978-3-540-53335-1.

- [56] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. *The PageRank Citation Ranking: Bringing Order to the Web*, <http://ilpubs.stanford.edu:8090/422/>. Technical Report 1999-66. Stanford InfoLab, Nov. 1999. URL: <http://ilpubs.stanford.edu:8090/422/>.
- [57] Tiago P. Peixoto. "Hierarchical block structures and high-resolution model selection in large networks". In: *Physical Review X* 4.1 (2014), p. 011047. DOI: 10.1103/PhysRevX.4.011047. URL: <http://link.aps.org/doi/10.1103/PhysRevX.4.011047> (visited on 01/21/2015).
- [58] Johan Philip. "The Probability Distribution of the Distance between two Random Points in a Box". Department of Mathematics, KTH, Dec. 2007.
- [59] E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A.-L. Barabási. "Hierarchical Organization of Modularity in Metabolic Networks". en. In: *Science* 297.5586 (Aug. 2002), pp. 1551–1555. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1073374. URL: <http://www.sciencemag.org/content/297/5586/1551> (visited on 01/21/2015).
- [60] Erzsébet Ravasz and Albert-László Barabási. "Hierarchical organization in complex networks". In: *Physical Review E* 67.2 (Feb. 2003), p. 026112. DOI: 10.1103/PhysRevE.67.026112. URL: <http://link.aps.org/doi/10.1103/PhysRevE.67.026112> (visited on 01/21/2015).
- [61] P. A. Robinson, J. A. Henderson, E. Matar, P. Riley, and R. T. Gray. "Dynamical reconnection and stability constraints on cortical network architecture". In: *Physical Review Letters* 103.10 (Sept. 2009), p. 108104. DOI: 10.1103/PhysRevLett.103.108104. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.103.108104> (visited on 01/21/2015).
- [62] M. Rosvall, D. Axelsson, and C. T. Bergstrom. "The map equation". en. In: *The European Physical Journal Special Topics* 178.1 (Nov. 2009), pp. 13–23. ISSN: 1951-6401. DOI: 10.1140/epjst/e2010-01179-1. URL: <https://doi.org/10.1140/epjst/e2010-01179-1> (visited on 01/28/2019).
- [63] Martin Rosvall and Carl T. Bergstrom. "Maps of random walks on complex networks reveal community structure". en. In: *Proceedings of the National Academy of Sciences* 105.4 (Jan. 2008), pp. 1118–1123. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.0706851105. URL: <https://www.pnas.org/content/105/4/1118> (visited on 01/28/2019).
- [64] Martin Rosvall and Carl T. Bergstrom. "Multilevel compression of random walks on networks reveals hierarchical organization in large integrated systems". In: *PLoS ONE* 6.4 (Apr. 2011), e18209. DOI: 10.1371/journal.pone.0018209. URL: <http://dx.doi.org/10.1371/journal.pone.0018209> (visited on 02/12/2016).

- [65] Camille Roth, Soong Moon Kang, Michael Batty, and Marc Barthelemy. "A long-time limit for world subway networks". In: *Journal of The Royal Society Interface* (2012). ISSN: 1742-5689. DOI: 10.1098/rsif.2012.0259. URL: <http://rsif.royalsocietypublishing.org/content/early/2012/05/15/rsif.2012.0259>.
- [66] Marta Sales-Pardo, Roger Guimerà, André A. Moreira, and Luís A. Nunes Amaral. "Extracting the hierarchical organization of complex systems". In: *Proceedings of the National Academy of Sciences of the United States of America* 104.39 (Sept. 2007), pp. 15224–15229. ISSN: 0027-8424. DOI: 10.1073/pnas.0703740104. URL: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2000510/> (visited on 01/21/2015).
- [67] Somwrita Sarkar, James A. Henderson, and Peter A. Robinson. "Spectral characterization of hierarchical network modularity and limits of modularity detection". In: *PLoS ONE* 8.1 (Jan. 2013), e54383. DOI: 10.1371/journal.pone.0054383. URL: <http://dx.doi.org/10.1371/journal.pone.0054383> (visited on 01/21/2015).
- [68] Vedran Sekara and Sune Lehmann. "The Strength of Friendship Ties in Proximity Sensor Data". en. In: *PLOS ONE* 9.7 (July 2014), e100915. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0100915. URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0100915> (visited on 11/02/2018).
- [69] Vedran Sekara, Arkadiusz Stopczynski, and Sune Lehmann. "The fundamental structures of dynamic social networks". In: *arXiv:1506.04704 [physics]* (June 2015). arXiv: 1506.04704. URL: <http://arxiv.org/abs/1506.04704> (visited on 02/24/2016).
- [70] Shai S. Shen-Orr, Ron Milo, Shmoolik Mangan, and Uri Alon. "Network motifs in the transcriptional regulation network of *Escherichia coli*". en. In: *Nature Genetics* 31.1 (2002), pp. 64–68. ISSN: 1061-4036. DOI: 10.1038/ng881. URL: <http://www.nature.com/ng/journal/v31/n1/full/ng881.html> (visited on 04/05/2017).
- [71] H. Francis Song and Xiao-Jing Wang. "Simple, distance-dependent formulation of the Watts-Strogatz model for directed and undirected small-world networks". In: *Physical Review E* 90.6 (Dec. 2014), p. 062801. DOI: 10.1103/PhysRevE.90.062801. URL: <https://link.aps.org/doi/10.1103/PhysRevE.90.062801> (visited on 01/09/2019).
- [72] V. Sood, S. Redner, and D. ben-Avraham. "First-passage properties of the Erdős–Renyi random graph". en. In: *Journal of Physics A: Mathematical and General* 38.1 (2005), p. 109. ISSN: 0305-4470. DOI: 10.1088/0305-4470/38/1/007. URL: <http://stacks.iop.org/0305-4470/38/i=1/a=007> (visited on 07/31/2018).

- [73] A. Steger and N. C. Wormald. "Generating random regular graphs quickly". In: *Comb. Probab. Comput.* 8.4 (July 1999), pp. 377–396. ISSN: 0963-5483. DOI: 10.1017/S0963548399003867. URL: <http://dx.doi.org/10.1017/S0963548399003867> (visited on 08/10/2017).
- [74] Arkadiusz Stopczynski, Vedran Sekara, Piotr Sapiezynski, Andrea Cuttone, Mette My Madsen, Jakob Eg Larsen, and Sune Lehmann. "Measuring Large-Scale Social Networks with High Resolution". In: *PLOS ONE* 9.4 (Apr. 2014), e95978. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0095978. URL: <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0095978> (visited on 02/24/2016).
- [75] Prasad Tetali. "Random walks and the effective resistance of networks". en. In: *Journal of Theoretical Probability* 4.1 (Jan. 1991), pp. 101–109. ISSN: 0894-9840, 1572-9230. DOI: 10.1007/BF01046996. URL: <https://link.springer.com/article/10.1007/BF01046996> (visited on 05/22/2017).
- [76] Jeffrey Travers and Stanley Milgram. "An Experimental Study of the Small World Problem". In: *Sociometry* 32.4 (1969), pp. 425–443. ISSN: 0038-0431. DOI: 10.2307/2786545. URL: <https://www.jstor.org/stable/2786545> (visited on 11/25/2018).
- [77] Christian L. Vestergaard and Mathieu Génois. "Temporal Gillespie Algorithm: Fast Simulation of Contagion Processes on Time-Varying Networks". en. In: *PLOS Computational Biology* 11.10 (Oct. 2015), e1004579. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1004579. URL: <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004579> (visited on 09/23/2018).
- [78] Christian L. Vestergaard, Mathieu Génois, and Alain Barrat. "How memory generates heterogeneous dynamics in temporal networks". In: *Physical Review E* 90.4 (Oct. 2014), p. 042805. DOI: 10.1103/PhysRevE.90.042805. URL: <https://link.aps.org/doi/10.1103/PhysRevE.90.042805> (visited on 03/12/2019).
- [79] Duncan J. Watts. "How small is the world, really?" In: *Medium* (Feb. 2016). URL: <https://medium.com/@duncanjwatts/how-small-is-the-world-really-736fa21808ba> (visited on 02/05/2019).
- [80] Duncan J. Watts, Peter Sheridan Dodds, and M. E. J. Newman. "Identity and Search in Social Networks". In: *Science* 296.5571 (May 2002), pp. 1302–1305. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1070120. URL: <http://science.sciencemag.org/content/296/5571/1302> (visited on 02/21/2017).
- [81] Duncan J. Watts and Steven H. Strogatz. "Collective dynamics of 'small-world' networks". en. In: *Nature* 393.6684 (June 1998), pp. 440–442. ISSN: 0028-0836. DOI: 10.1038/30918. URL: <http://www.nature.com/nature/journal/v393/n6684/full/393440a0.html> (visited on 02/12/2016).

- [82] F. Y. Wu. "Theory of resistor networks: the two-point resistance". en. In: *Journal of Physics A: Mathematical and General* 37.26 (2004), p. 6653. ISSN: 0305-4470. DOI: 10.1088/0305-4470/37/26/004. URL: <http://stacks.iop.org/0305-4470/37/i=26/a=004> (visited on 11/29/2016).
- [83] X. Xie, Y. Li, A. T. Y. Chwang, P. L. Ho, and W. H. Seto. "How far droplets can move in indoor environments—revisiting the Wells evaporation-falling curve". eng. In: *Indoor Air* 17.3 (June 2007), pp. 211–225. ISSN: 0905-6947. DOI: 10.1111/j.1600-0668.2007.00469.x.
- [84] W. W. Zachary. "An information flow model for conflict and fission in small groups". In: *Journal of Anthropological Research* 33 (1977), pp. 452–473.
- [85] Kun Zhao, Juliette Stehlé, Ginestra Bianconi, and Alain Barrat. "Social network dynamics of face-to-face interactions". English. In: *Physical review. E, Statistical, nonlinear, and soft matter physics* 83 (2011). DOI: 10.1103/PhysRevE.83.056109.

Appendices

A

Infomap: A Flow-based Algorithm to Find Modular Hierarchical Network Partitions

Real-world static networks are often claimed to be of *modular hierarchical structure* [10, 60, 59, 66, 2, 36, 37, 52, 57, 67, 61], where this term usually refers to the fact that the nodes of a network can be grouped to form modules based on some sort of similarity which is computed from their connections. Then, recursively, these modules may again be grouped together to form an even coarser grained structure and so forth. A single realization of such a modular hierarchical grouping scheme is called a *modular hierarchical network partition* (MHNP). While there exist descriptions allowing for *overlapping* modules where nodes and submodules may be part of multiple groups, these will not be of concern within this work since they do not add any particular value to the qualitative investigation of the consequence of modular hierarchical clustering on dynamic systems.

Extraction algorithms to find “optimal” such partitions usually search the space of all partitions of a particular observed network to find a single one or multiple which minimize a certain cost function (or maximize a certain likelihood function). Some of these depend on the description of the network as generated by a stochastic block model and maximize the likelihood that a nested stochastic block model defined by a modular hierarchical network partition will generate the observed network while minimizing the information theoretic description length of the model to prevent overfitting [10, 57]. However, this thesis mainly aims at exploring the consequence of modular hierarchical clustering on the outcome of dynamic processes, more accurately on the outcome of random walk observables. To this end, it seems plausible to deploy a modular hierarchical partition algorithm which finds partitions based on the similarity of nodes concerning their role in random walks. One such algorithm is *Infomap*, which searches for partitions minimizing the so-called *map equation* [62, 63, 64]. In the following, the underlying mechanism of *Infomap* is briefly outlined to give an intuition about its inner workings, while a complete description would go beyond the scope of this work due to it only playing a minor role overall.

At the base of the algorithm lies the trajectory of a single random

¹ The codenames are chosen *prefix-free* such that no delimiting symbols are necessary.

walker traversing through the network as described in Sec. 3.1.1. Giving each node a unique codename formed by the bits “1” and “0”, the trajectory of the walker can be completely encoded by the succession of these codenames, called a *trajectory sentence* in the subsequent description.¹ *Infomap* then aims to find the minimal description length of this trajectory sentence by introducing a modular hierarchical network partition.

Following the principle of Huffman coding, nodes which are visited more often than others are given shorter names such that their overall space in the trajectory sentence is minimized. This implies that node names can become rather long with increasing number of nodes N leading to a long trajectory sentence of any ergodic random walk. The idea behind *Infomap* is that (short) codenames for single nodes can be reused when nodes are grouped into modules in which the random walker spends a considerable amount of time steps before it exits the module to enter another module. Then, short module names can be used in the trajectory sentence, each indicating that it has been entered by the random walker. The codenames following this entry symbol consequently belong to the nodes of this module, even though their names can be duplicates of node names in other modules. Additionally, exit code names are used to indicate that the current module has been left.

Instead of actually simulating a random walker and subsequently encoding its trajectory, the description length of a trajectory of n steps can be lower bounded by $n\mathcal{L}(M)$ where $\mathcal{L}(M)$ is the lower bound average description length of a single step given that the coding is based on the modular hierarchical network partition M . The lower bound can be evaluated using Shannon’s source coding theorem which states that when a series of entities i is drawn randomly, each with probability p_i , from an ensemble \mathcal{R} , then the average length of the codename of such a symbol cannot be less than the ensemble’s information entropy which is given as

$$H(\mathcal{R}) = - \sum_{i=1}^{|\mathcal{R}|} p_i \log_2 p_i. \quad (\text{A.1})$$

The trajectory sentence of the random walk can be approximated as a randomly generated sentence of codenames associated with their probability of occurrence. For a two-level modular network partition, nodes are grouped to modules, but those modules are not further grouped. In such a partition the following events with corresponding codenames may occur.

1. Each module is associated with an entry symbol. When there are m such modules, there have to be m entry codes, each occurring with conditional probability $q_{i\curvearrowright} / \sum_{j=1}^m q_{j\curvearrowright}$ (the probability that module i is exited given that any module switching event happened). The probability $q_{i\curvearrowright}$ is proportional to the number of edges leading out of the module. This ensemble of events shall be called \mathcal{Q} in the following.

2. Each visiting node in a module is associated with a visiting symbol. The conditional probability that node u of module i is chosen given that any event of module i is chosen is $p_u / [q_i + \sum_{u \in i} p_u]$. The visiting probability p_u of node u is chosen as the equilibrium probability to find a random walker on node u as given in Eq. (3.3). Furthermore, each module i is additionally associated with an exit symbol, where $q_{i\sim}$ is the probability that module i is exited, hence, the corresponding conditional probability that the exit happened given that an event of module i happened is $q_{i\sim} / [q_{i\sim} + \sum_{u \in i} p_u]$. This particular ensemble of module- i events shall be called $\mathcal{P}^{(i)}$.

Considering these events and their ensembles \mathcal{Q} and $\mathcal{P}^{(i)}$, the weighted averaged length of a codename used within a single step is lower bounded by

$$\mathcal{L}(M) = \left(\sum_{i=1}^m q_{i\sim} \right) H(\mathcal{Q}) + \sum_{i=1}^m \left(q_{i\sim} + \sum_{u \in i} p_u \right) H(\mathcal{P}^{(i)})$$

which is called the *map equation* (here, *map* is a synonym for two-level modular network partition). This average single-step codename description length can be computed in a fast manner for any two-level partition M . In principle, any standard algorithm to minimize a cost function can be used to find an optimal partition. For a more detailed explanation of the map equation the reader is referred to [63, 62].

For modular hierarchical network partitions, the map equation takes a recursive form. Any modular hierarchical network partition M consists, at first, of m submaps M^i . For each of these submaps the map equation is valid in the same form, where now exit probabilities refer to the exit of coarse grained modules. A detailed explanation of the hierarchical map equation and an algorithm to find an optimal hierarchical map is given in [64]. In this thesis, the C++-implementation of the method described and implemented in [64] was used with default parameter choices.

B

Additional Flockwork Evaluations

In Chap. 8, the temporal network model of “Flockworks” is introduced, however several detailed analyses are shortened to increase comprehensibility. These analysis are shown here. In Sec. B.1, the temporal group size evolution of the model is derived. Sec. shows more detailed analysis results of the state-space trajectory inference method. Finally, Sec. shows how to derive the evolution equations of an SIS-model with infinite infection rate on the Flockwork model.

B.1 Group Evolution

In Sec. 8.1.2, sethe temporal evolution of group sizes is discussed. The detailed considerations of how those group sizes change considering the events that can take place are, however, rather tedious and do not provide any particular insight. Hence, they have been banned to this appendix and are presented in the following.

In a Flockwork process defined by the event rate γ and the reconnection probability P as introduced in Sec. 8.1, the change in group sizes can be found for a single time step for both reconnection and disconnection events using P as a single control parameter because the time scale can be fixed by γ^{-1} for a single event.

First, the reconnection events are discussed, which each node undergoes with probability P in a single time step. The initial condition is that the network state can be expressed by the group size vector \mathbf{n} with $0 \leq m \leq N + 1$ where each entry n_m contains the total number of completely connected groups of size m . The group function $g(u) = \text{Nei}(u) + \{u\}$ is used.¹ These are the possible reconnection events.

1. A node u detaches from its group $g(u)$ with group size $m = |g(u)|$. It then reattaches to another group of size m , which is not the same group as its original one. The number of groups of size $m, m - 1$ and $m + 1$ change as

$$\Delta n_m = -2, \quad \Delta n_{m-1} = 1, \quad \Delta n_{m+1} = 1.$$

2. A node u detaches from its group $g(u)$ with group size $m = |g(u)|$. It then reattaches to the same group of size m . The number

¹ Note that the groups are of course time dependent, so it should be written as $g(u, t)$, but in the following, the group size’s time dependence is omitted in the notation.

of groups of size m change as

$$\Delta n_m = 0.$$

3. A node u detaches from its group $g(u)$ with group size $m = |g(u)|$. It then reattaches to a group of size $m - 1$. The number of groups of size m and $m - 1$ change as

$$\Delta n_m = 0, \quad \Delta n_{m-1} = 0.$$

4. A node u detaches from its group $g(u)$ with group size $m = |g(u)|$. It then reattaches to any group of size $i \neq m$ and $i \neq m - 1$. The numbers of groups change as

$$\Delta n_m = -1, \quad \Delta n_i = -1, \quad \Delta n_{m-1} = 1, \quad \Delta n_{i+1} = 1.$$

From these events, the following changes in the number of groups of size $1 \leq m \leq N$ have to be considered, where the superscript (r) denotes a reconnection event.

$$\begin{aligned} \Delta n_m^{(r)} = & P^{(r)} \left[|g(u)| = m, |g(v)| = m, g(u) \neq g(v) \right] \times (-2) \\ & + P^{(r)} \left[|g(u)| = m - 1, |g(v)| = m - 1, g(u) \neq g(v) \right] \times (+1) \\ & + P^{(r)} \left[|g(u)| = m + 1, |g(v)| = m + 1, g(u) \neq g(v) \right] \times (+1) \\ & + P^{(r)} \left[|g(u)| = m, |g(v)| \neq m, |g(v)| \neq m - 1 \right] \times (-1) \\ & + P^{(r)} \left[|g(u)| = m + 1, |g(v)| \neq m, |g(v)| \neq m + 1 \right] \times (+1) \\ & + P^{(r)} \left[|g(u)| \neq m, |g(u)| \neq m - 1, |g(v)| = m - 1 \right] \times (+1) \\ & + P^{(r)} \left[|g(u)| \neq m + 1, |g(u)| \neq m, |g(v)| = m \right] \times (-1). \end{aligned} \tag{B.1}$$

Note that here, the acting node is named u and node that u connects to is named v . The probabilities are proportional to the probabilities that from all N nodes, node u is picked *and* it reconnects *and* node v is picked from the remaining $N - 1$ nodes afterwards. They are hence given as

$$\begin{aligned}
P^{(r)}[|g(u)| = m, |g(v)| = m, g(u) \neq g(v)] &= \\
&= P \frac{mn_m}{N} \times \frac{m(n_m - 1)}{N - 1} \\
P^{(r)}[|g(u)| = m - 1, |g(v)| = m - 1, g(u) \neq g(v)] &= \\
&= P \frac{(m - 1)n_{m-1}}{N} \times \frac{(m - 1)(n_{m-1} - 1)}{N - 1} \\
P^{(r)}[|g(u)| = m + 1, |g(v)| = m + 1, g(u) \neq g(v)] &= \\
&= P \frac{(m + 1)n_{m+1}}{N} \times \frac{(m + 1)(n_{m+1} - 1)}{N - 1} \\
P^{(r)}[|g(u)| = m, |g(v)| \neq m, |g(v)| \neq m - 1] &= \\
&= P \frac{mn_m}{N} \times \left(1 - \frac{mn_m - 1}{N - 1} - \frac{(m - 1)n_{m-1}}{N - 1}\right) \\
P^{(r)}[|g(u)| = m + 1, |g(v)| \neq m + 1, |g(v)| \neq m] &= \\
&= P \frac{(m + 1)n_{m+1}}{N} \times \left(1 - \frac{mn_m}{N - 1} - \frac{(m + 1)n_{m+1} - 1}{N - 1}\right) \\
P^{(r)}[|g(u)| \neq m - 1, |g(u)| \neq m, |g(v)| = m - 1] &= \\
&= P \left(1 - \frac{(m - 1)n_{m-1}}{N} - \frac{mn_m}{N}\right) \times \frac{(m - 1)n_{m-1}}{N - 1} \\
P^{(r)}[|g(u)| \neq m + 1, |g(u)| \neq m, |g(v)| = m] &= \\
&= P \left(1 - \frac{(m + 1)n_{m+1}}{N} - \frac{mn_m}{N}\right) \times \frac{mn_m}{N - 1}.
\end{aligned}$$

Now, the events at which nodes disconnect and subsequently stay alone have to be considered. These happen to every node with probability $1 - P$.

1. A node u from group $g(u)$ with group size $m = |g(u)| > 1$ detaches. The change in group sizes is

$$\Delta n_m = -1 \quad \Delta n_1 = +1 \quad \Delta n_{m-1} = +1.$$

2. A node u of group size $m = 1$ detaches. This does not change the group size distribution.

$$\Delta n_m = 0.$$

The new probabilities for a change in group size m are

$$\begin{aligned}
P^{(d)}[|g(u)| = m] &= (1 - P) \frac{mn_m}{N} \\
P^{(d)}[|g(u)| > 1] &= (1 - P) \frac{N - n_1}{N} = \frac{1 - P}{N} \sum_{m=2}^N mn_m.
\end{aligned}$$

So the change in the number of groups of size m in a single time step are

$$\begin{aligned}
\Delta n_{m>1}^{(d)} &= P^{(d)}[|g(u)| = m] \times (-1) + P^{(d)}[|g(u)| = m + 1] \times (+1) \\
\Delta n_1^{(d)} &= P^{(d)}[|g(u)| > 1] \times (+1) + P^{(d)}[|g(u)| = 2] \times (+1).
\end{aligned}$$

Combining Eqs. (B.1) and (B.2) one finds the evolution equations

$$N\Delta n_1 = -2Pn_1 + 2Pn_2 \frac{N-2}{N-1} + (1-P)(N-n_1) + (1-P)2n_2, \quad (\text{B.3a})$$

$$\begin{aligned} N\Delta n_{1 < m \leq N} = & \frac{P}{N-1} \left[-2mn_m(N-m) \right. \\ & + n_{m-1} \left(N(m-1) - (m-1)^2 \right) + \\ & \left. + n_{m+1} \left(N(m+1) - (m+1)^2 \right) \right] - \\ & - (1-P)mn_m + (1-P)(m+1)n_{m+1}. \end{aligned}$$

Furthermore, no groups of size $m = 0$ and $m = N+1$ can exist, so $\Delta n_0 = \Delta n_{N+1} = 0$. Using Eq. (8.1), the term $N - n_1$ in Eq. (B.3a) can be expressed as $N - n_1 = \sum_{m=2}^N mn_m$ and hence the temporal evolution of the expected number of groups of size m can also be expressed using a transition matrix \mathbf{W} such that

$$\mathbf{n}(t+1) = \mathbf{W}\mathbf{n}(t), \quad (\text{B.4})$$

where

$$W_{mm'} = \begin{cases} 1 - 2P/N & m = 1, m' = 1 \\ 2\frac{P}{N} \frac{N-2}{N-1} + \frac{4}{N}(1-P) & m = 1, m' = 2 \\ \frac{m'}{N}(1-P) & m = 1, m' > 2 \\ P \frac{N(m-1)-(m-1)^2}{N(N-1)} & 1 < m < N, m' = m-1 \\ 1 - 2P \frac{m(N-m)}{N(N-1)} - (1-P)\frac{m}{N} & 1 < m < N, m' = m \\ (1-P)\frac{m+1}{N} + P \frac{N(m+1)-(m+1)^2}{N(N-1)} & 1 < m < N, m' = m+1 \\ P/N & m = N, m' = N-1 \\ P & m = N, m' = N \\ 0 & \text{otherwise.} \end{cases} \quad (\text{B.5})$$

The transitions of this process are discrete but the inter-event time follows an exponential distribution with the relevant time scale given by the inverse event rate γ^{-1} . Hence, the temporal evolution of \mathbf{n} can be expressed by the continuous-time ordinary differential equation

$$\partial_t \mathbf{n} = \gamma (\mathbf{W} - \mathbb{1}) \mathbf{n}.$$

These are the temporal evolution equations presented in the main text Sec. 8.1.2. In the following, the equilibrium state of this process will be determined by considering that the overall change in a single time step vanishes (hence $\Delta\mathbf{n} = 0$) and recursively solving Eqs. (B.3). Consider Eq. (B.3a) with $\Delta n_1 = 0$ for the number of single nodes (group size $m = 1$). This gives the equilibrium state of the number of groups of size $m = 2$ as

$$n_2^* = \frac{N(P-1) + n_1(1+P)}{2P(N-2)/(N-1) + 2(1-P)}.$$

Now for every $2 \leq m \leq N - 1$ the equations given by $\Delta n_m = 0$ can be solved to obtain an equation for the equilibrium value n_{m+1}^* which only depends on the equilibrium value n_1^* , recursively given as

$$\begin{aligned} n_{m+1}^* = & \left[m^2 P (2n_m^* - n_{m-1}^*) + \right. \\ & + m ((N+2)n_{m-1}^* P - n_m^* (NP + N + P - 1)) - \\ & \left. - (N+1)n_{m-1}^* P \right] / [(m+1)(mP - N + 1)]. \end{aligned}$$

This process can be repeated until one arrives at $m = N - 1$ and so obtains an equation for n_N^* that only depends on n_1^* , i.e. it is of the form

$$n_N^* = an_1^* + b$$

with a, b determined by the recursion relation. Subsequently, one uses Eq. (B.3a) for $m = N$ to obtain

$$n_N^* = \frac{P}{(1-P)N} n_{N-1}^*,$$

an equation which also only depends on n_N^* and n_1^* considering the previous recursive equations such that $n_{N-1}^* = cn_1^* + d$. Now, both relations can be equated to find

$$n_1^* = \frac{d - b}{a - c}$$

and hence all values of n_m^* by applying the recursive equations. Here, a computer algebra system was used to find the values of n_m^* in this manner, which, after applying the normalization condition $N = \sum_{m=1}^N mn_m$, are given as

$$n_1^* = N(1 - P) \tag{B.6a}$$

$$n_{1 < m < N}^* = \frac{1}{m} \frac{N!}{(N-m)!} \frac{1-P}{\prod_{j=1}^{m-1} [1 - Pj/(N-1)]} \left(\frac{P}{N-1} \right)^{m-1} \tag{B.6b}$$

$$n_N^* = (N-2)! \frac{P}{\prod_{j=1}^{N-2} [1 - Pj/(N-1)]} \left(\frac{P}{N-1} \right)^{N-1}. \tag{B.6c}$$

This result was derived under the assumption that the initial network can be described as a collection of components which are alle fully connected. Below, it will be shown that this equilibrium distribution will be approached for *any* initial condition.

Consider two group count vectors, one counting the number n_m of groups of size $1 \leq m \leq N$ which are completely connected, and one counting the number \bar{n}_m of groups of size $3 \leq m \leq N$ which are *not* fully connected. Hence, the state vector $(n, \bar{n})^T$ can completely contain any network (even though the explicit composition of the groups counted in the vector \bar{n} are ignored). Analyzing complete groups, the following events can happen.

1. A node disconnects from a fully connected group. It hence will still be a fully connected group containing one node less. Furthermore, the disconnect node will stay alone with probability $1 - P$, itself forming a complete network of size $m = 1$.
2. A node connects to a fully connected group. It hence will still be a fully connected group containing one node more.

It thus becomes clear that once a group is counted within the vector \mathbf{n} , it will stay in this category. The transition matrix describing the changes in counts of complete groups will be called \mathbf{W}' in the following (compared to Eq. (B.4)). Now, consider groups which are not complete and thus are counted in the vector $\bar{\mathbf{n}}$. The following events may happen.

1. A node connects to a non-fully connected group. It hence will still not be fully connected but contain one group more.
2. A node disconnects from a non-fully connected group. A distinction between two cases has to be made.
 - (a) The remaining group is now either fully connected and counted in the vector \mathbf{n} or it is still non-fully connected and counted in the vector $\bar{\mathbf{n}}$.
 - (b) Due to a node connecting multiple components of this group, the original group splits in multiple groups and each of those are now either counted in the vector $\bar{\mathbf{n}}$ or the vector \mathbf{n} .

Transition rates from groups in $\bar{\mathbf{n}}$ to fully connected groups in \mathbf{n} will be contained in the matrix $\bar{\mathbf{W}}_1$ which has non-zero values in the upper triangle (because non-connected groups of size m can influence the count of fully connected groups of any size $m' < m$ but no fully connected groups of size $m' \geq m$). Transition rates from groups in $\bar{\mathbf{n}}$ to other groups in $\bar{\mathbf{n}}$ will be contained in the matrix $\bar{\mathbf{W}}_2$. This matrix has non-zero values in the upper triangle, the diagonal and the first diagonal shifted to the lower left, because non-fully connected groups of size m can influence the count of non-fully connected groups of size $m' < m$, non-fully connected groups of size $m' = m + 1$ (by reconnection) or $m' = m$. Then, the full time evolution is given as

$$\begin{pmatrix} \mathbf{n}(t+1) \\ \bar{\mathbf{n}}(t+1) \end{pmatrix} = \begin{pmatrix} \mathbf{W}' & \bar{\mathbf{W}}_1 \\ 0 & \bar{\mathbf{W}}_2 \end{pmatrix} \begin{pmatrix} \mathbf{n}(t) \\ \bar{\mathbf{n}}(t) \end{pmatrix}.$$

The shape of this equation implies that for any initial condition $(\mathbf{n}(0), \bar{\mathbf{n}}(0))^T$ the group count of $\bar{\mathbf{n}}$ will feed into the group count of \mathbf{n} until $\bar{\mathbf{n}} = 0$. When this point is reached, the full temporal evolution is determined by Eq. (B.4) and hence the equilibrium Eq. (8.3) is reached eventually.

One may compute the expected number of components and the expected size of those components from the group-size distribution.

The mean group size is

$$\langle g \rangle = \frac{\sum_{m=1}^N mn_m}{\sum_{m=1}^N n_m} = \frac{\sum_{m=1}^N mn_m}{\langle c \rangle} = \frac{N}{\langle c \rangle}.$$

with $\langle c \rangle$ being the mean number of components. Using Eq. (B.6) one may find² the exact result

$$\begin{aligned} \frac{\langle c \rangle}{N} &= 1 - P + \underbrace{\frac{1}{N} \frac{(-1)^N P \Gamma(N-1) \Gamma\left(\frac{-N+P+1}{P}\right)}{\Gamma\left(\frac{-N+P+1}{P} + N - 2\right)}}_{A_1} + \frac{1-P}{P^2} \left[\right. \\ &\quad + \underbrace{\frac{\Gamma\left(\frac{-N+P+1}{P}\right) ((-1)^N (N-1) P^2 \Gamma(N-1))}{\Gamma\left(\frac{PN-N+1}{P}\right) N}}_{A_2} + \\ &\quad + \underbrace{\frac{\Gamma\left(\frac{-N+P+1}{P}\right)}{\Gamma\left(\frac{-N+2P+1}{P}\right)} \times (N-1-P) \times}_{A_3} \\ &\quad \left. \times \underbrace{\left(P - \frac{N-1}{N} \psi\left(\frac{1-N}{P}\right) + \frac{N-1}{N} \psi\left(\frac{N(P-1)+1}{P}\right) \right)}_{A_4} \right], \end{aligned}$$

² Here, this was done using a computer algebra system.

where $\Gamma(x)$ is the Gamma function and $\psi(x) = \Gamma'(x)/\Gamma(x)$ is the polygamma function. This result is incredibly unhelpful. However, it is possible to conjecture an asymptotic result for large numbers of nodes $N \gg 1$, which is done separately for each of the terms above.

For the first two terms, a numerical test shows that the conjectures $\lim_{N \rightarrow \infty} A_1/N = 0$ and $\lim_{N \rightarrow \infty} A_2 = 0$ seem appropriate. For A_3 , one finds

$$A_3 \xrightarrow{N \gg 1} -P/(N-1).$$

In the last term, one might use the fact that $\psi(x) \approx \log x$ for $x \gg 1$. Even though the arguments are negative here and hence this approximation should not be valid, one finds.

$$\begin{aligned} A_4 &\approx P + \frac{N-1}{N} \log\left(\frac{N(P-1)+1}{P}\right) - \\ &\quad - \frac{N-1}{N} \log\left(\frac{1-N}{P}\right) \\ &= P + \frac{N-1}{N} \log\left(1 - P \frac{N}{N-1}\right) \\ &\xrightarrow{N \rightarrow \infty} P + \log(1-P). \end{aligned}$$

Summing up, this yields the asymptotic formula

$$\begin{aligned}
\lim_{N \rightarrow \infty} \frac{\langle c \rangle}{N} &= 1 - P + (P + \log(1 - P)) \frac{1 - P}{P^2} \\
&\quad \lim_{N \rightarrow \infty} \left(-\frac{P}{N-1} \right) (N-1-P) \\
&= 1 - P - P(P + \log(1 - P)) \frac{1 - P}{P^2} \\
&= \frac{P - P^2}{P} - \frac{P - P^2 + (1 - P) \log(1 - P)}{P} \\
&= -\frac{1 - P}{P} \log(1 - P),
\end{aligned}$$

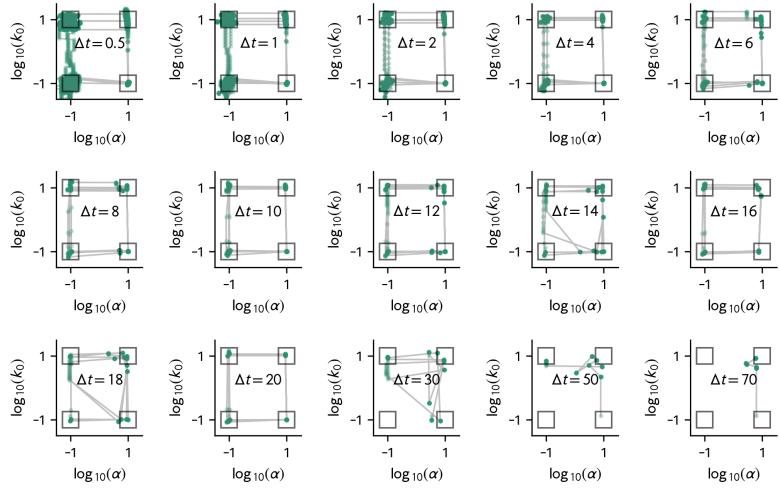
which also means that

$$\lim_{N \rightarrow \infty} \langle g \rangle = -\frac{P}{(1 - P) \log(1 - P)}.$$

B.2 Rate Inference

In Sec. 8.2, a method to infer the state-space trajectory of time-varying rate Flockwork simulations is introduced and applied to both artificial Flockworks and real-world data. The full analyses for increasing Δt is shown in the following.

Figure B.1: Inference of the (α, k_0) -state-space trajectory from a single Flockwork simulation as described in Sec. 8.2.2 for increasing inference parameter Δt .



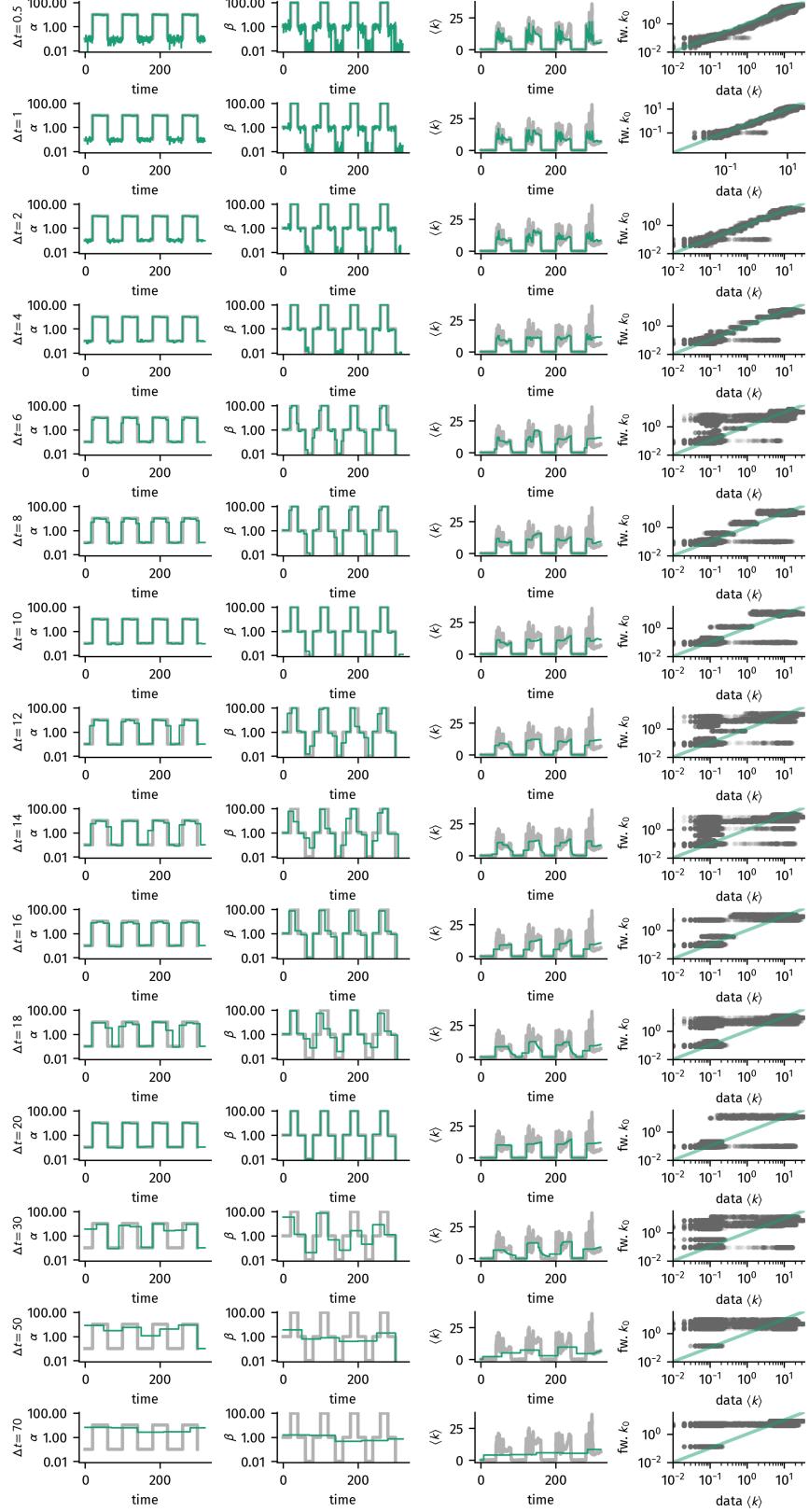


Figure B.2: Inference of the reconnection rates α and β from a single Flockwork simulation as described in Sec. 8.2.2 for increasing inference parameter Δt .

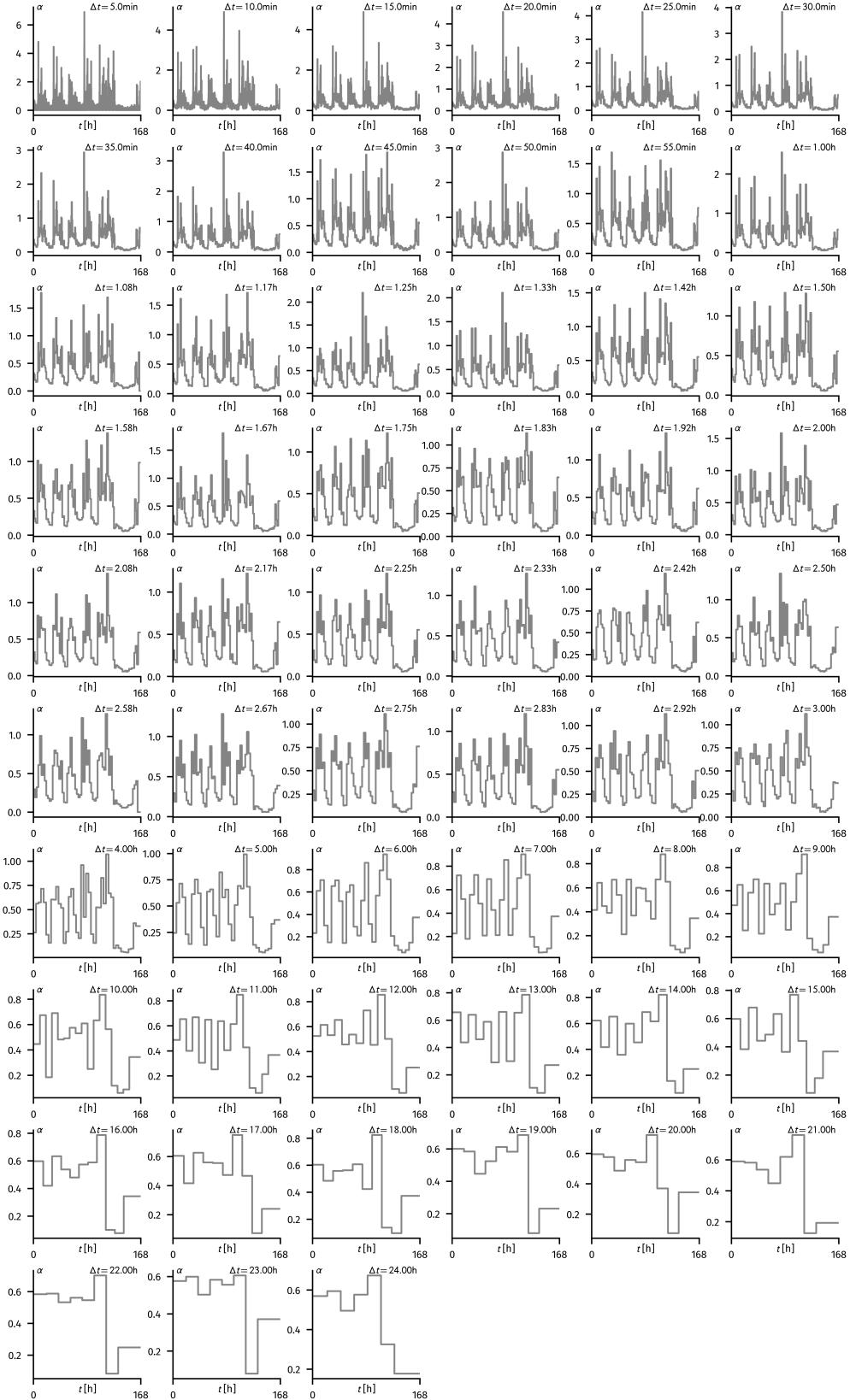


Figure B.3: Inference of the reconnection rate α from the DTU data for increasing inference parameter Δt .

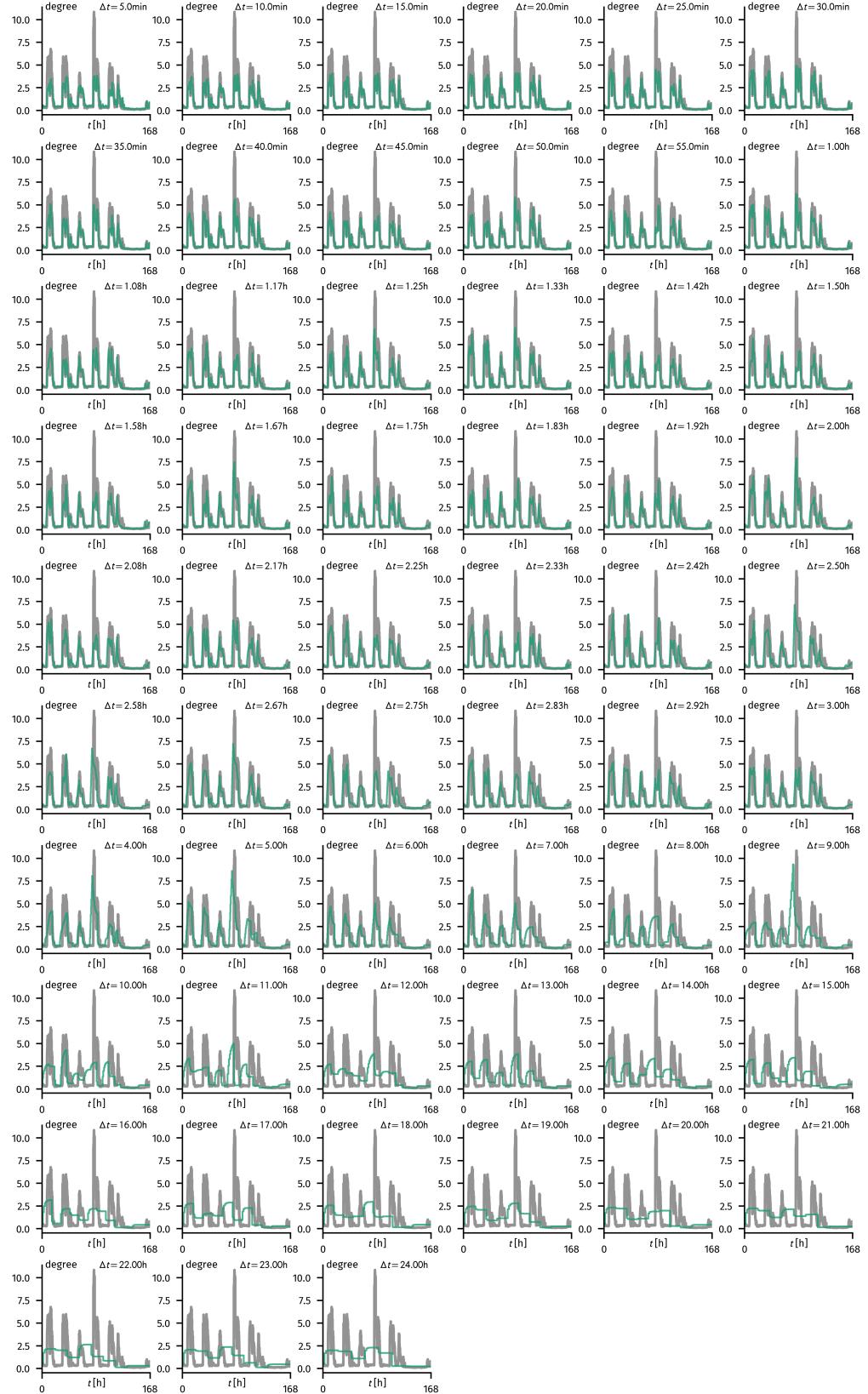


Figure B.4: Inference of the expected mean degree k_0 from the DTU data for increasing inference parameter Δt and comparison to the actual mean degree $\langle k \rangle(t)$.

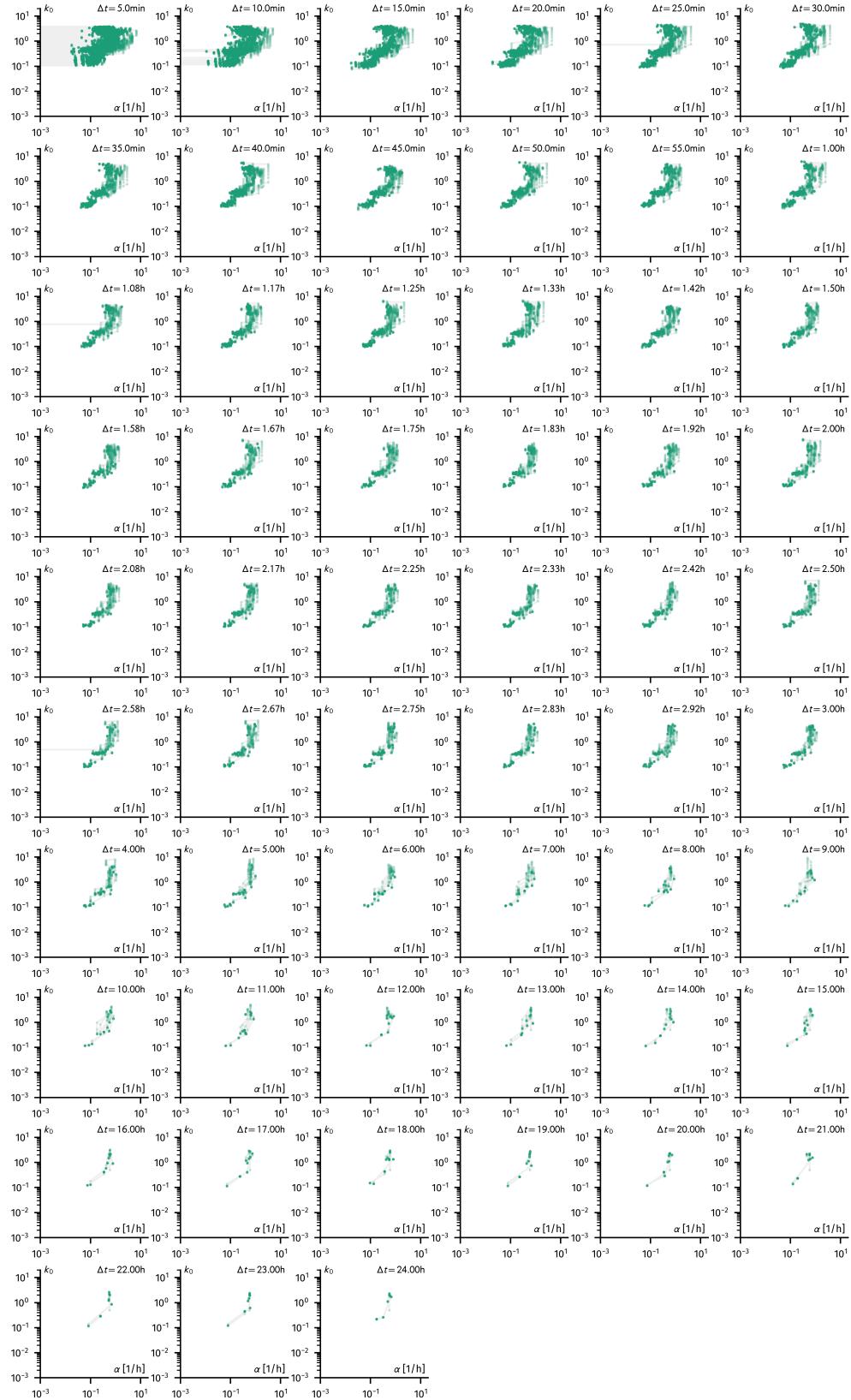


Figure B.5: Inferred trajectory in the (α, k_0) -plane from the DTU data for increasing inference parameter Δt

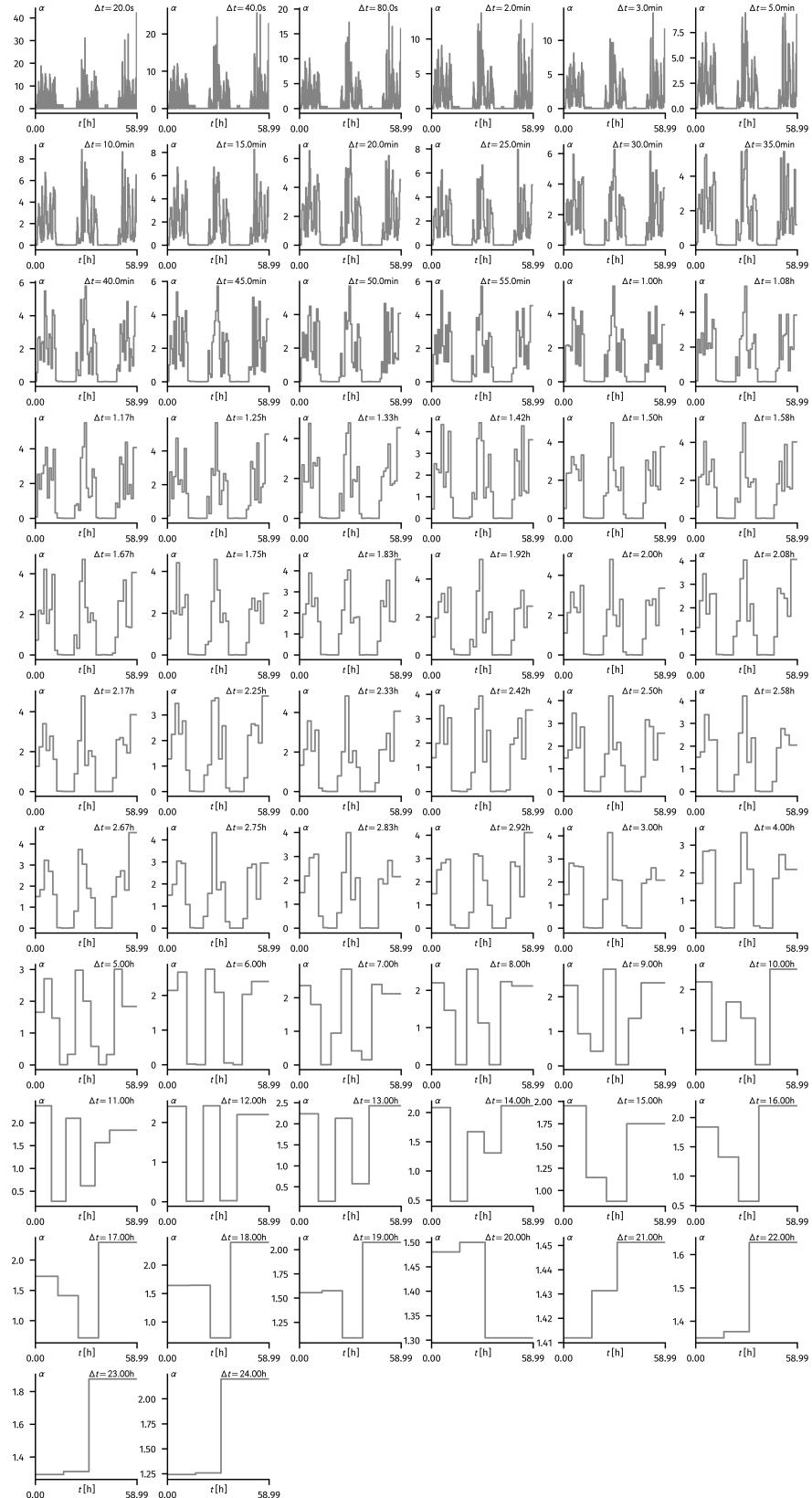


Figure B.6: Inference of the reconnection rate α from the HT09 data for increasing inference parameter Δt .

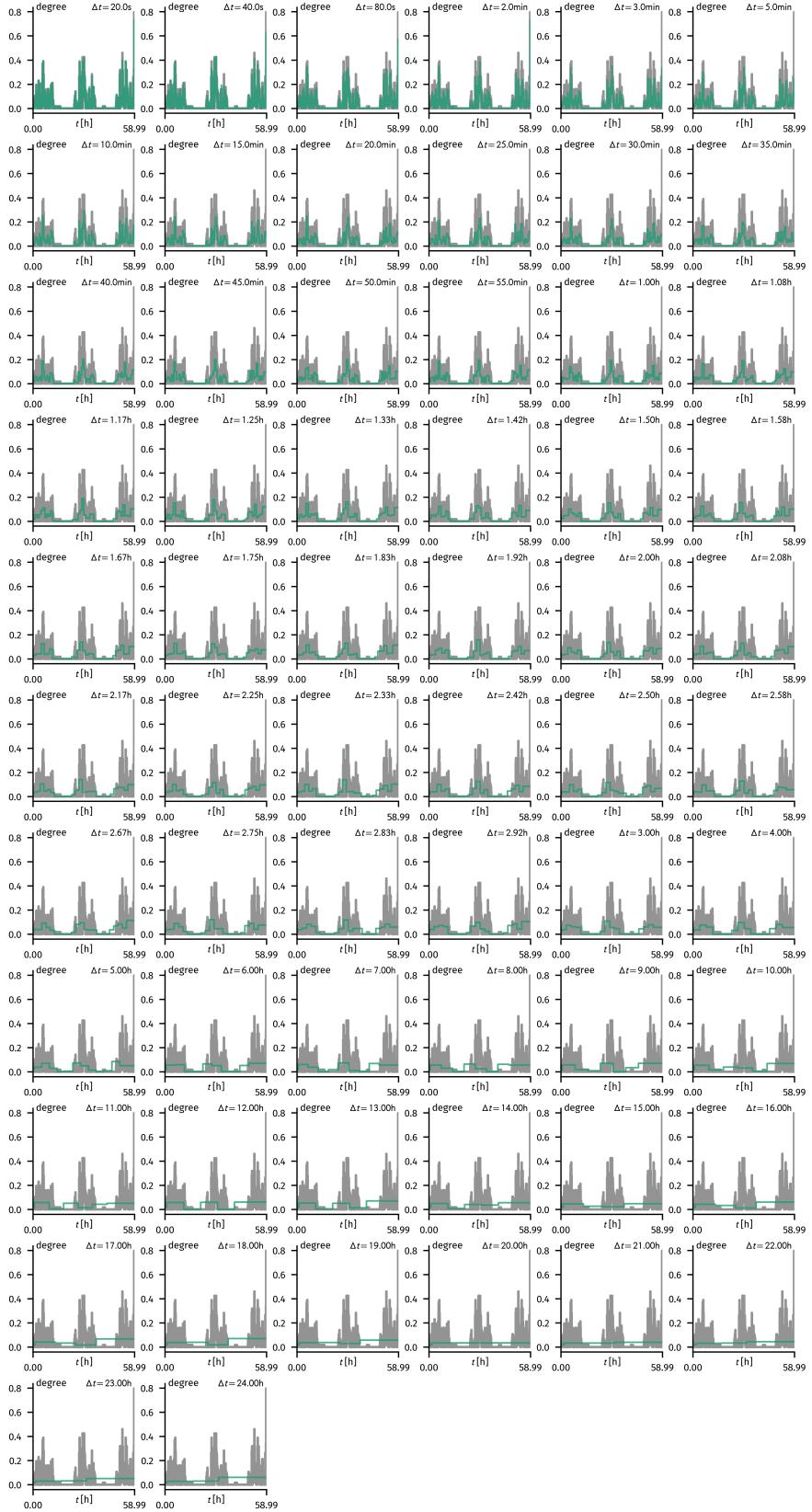


Figure B.7: Inference of the expected mean degree k_0 from the HT09 data for increasing inference parameter Δt and comparison to the actual mean degree $\langle k \rangle(t)$.

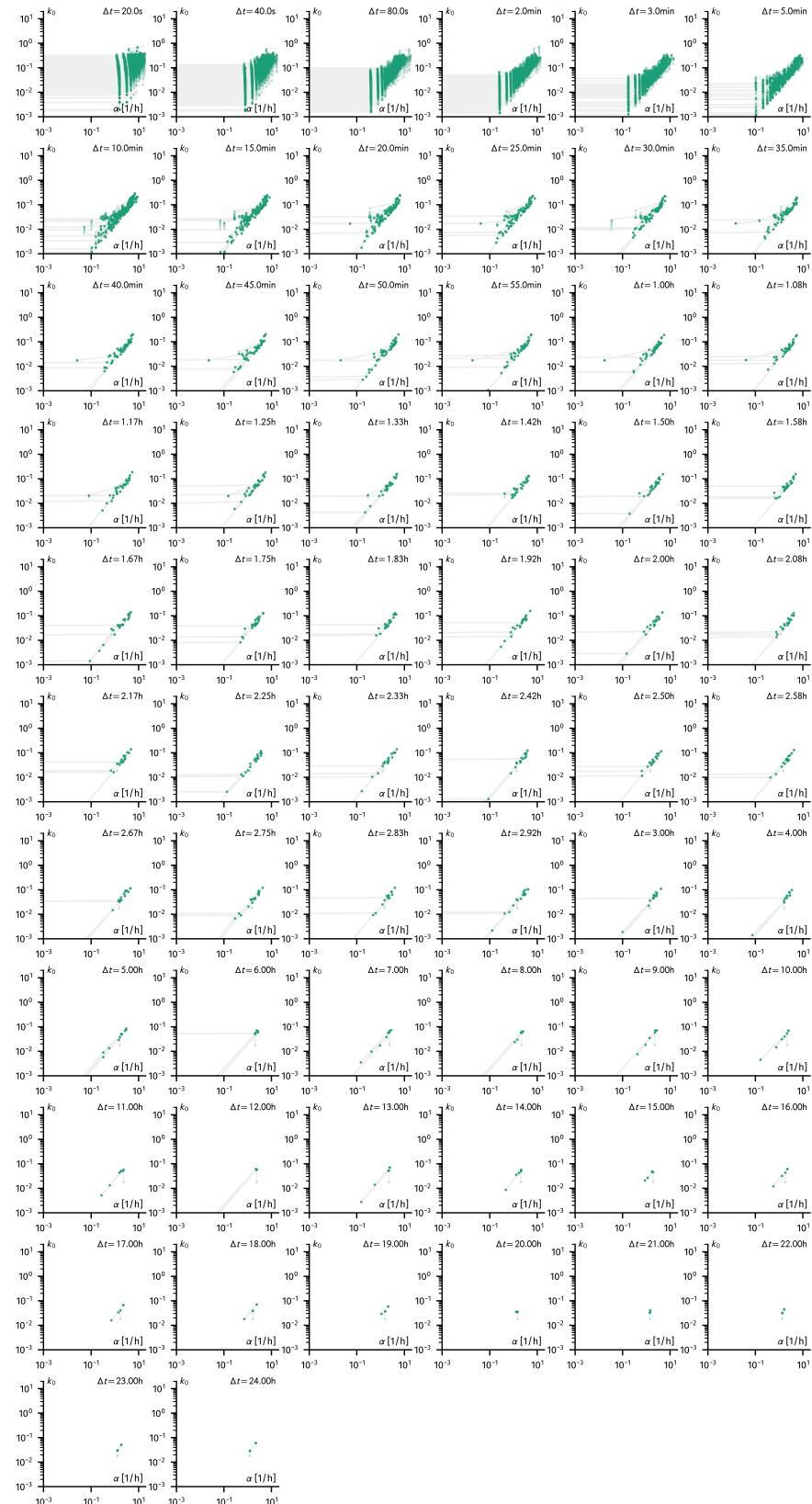


Figure B.8: Inferred trajectory in the (α, k_0) -plane from the HTog data for increasing inference parameter Δt

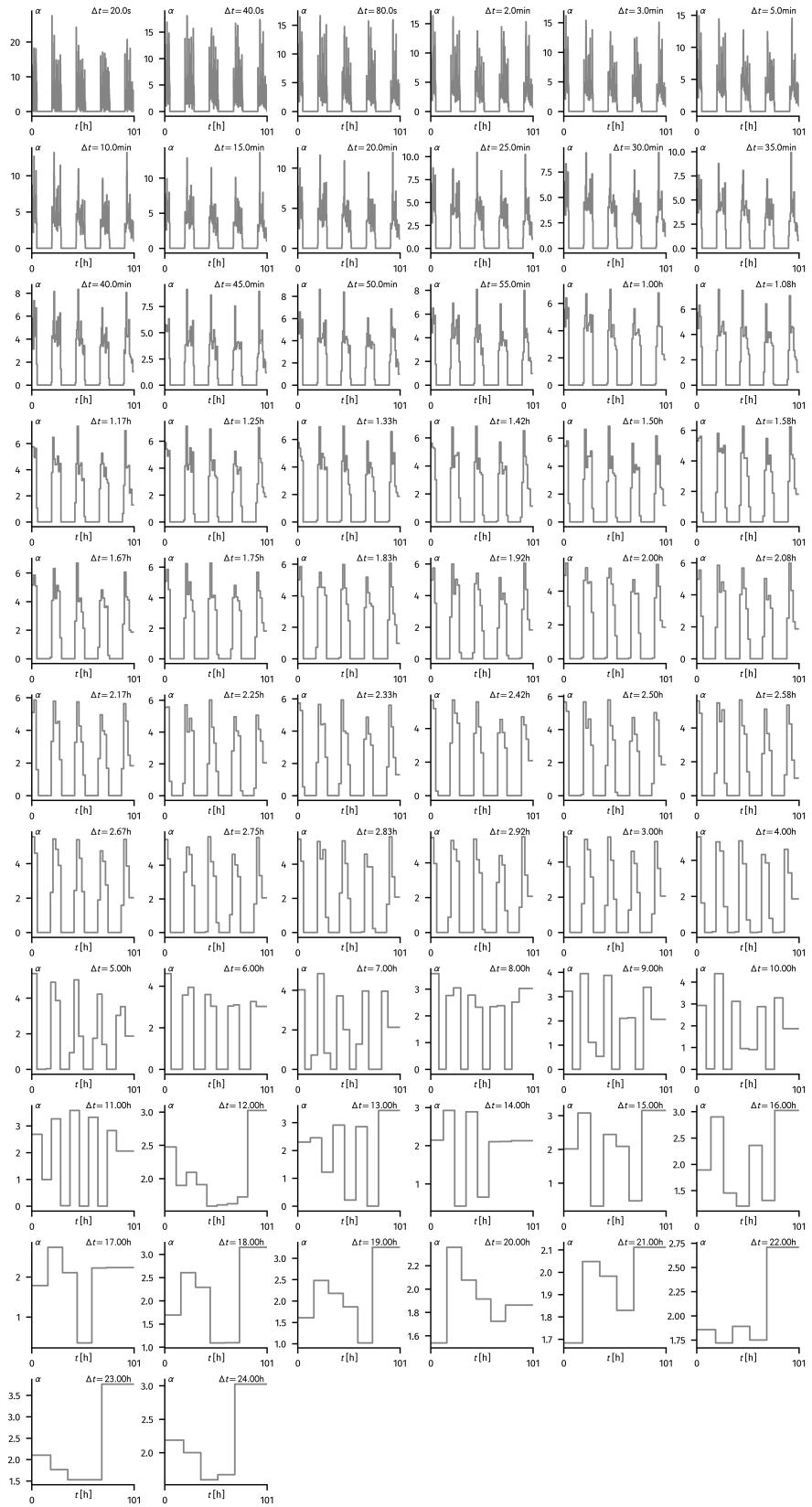


Figure B.9: Inference of the reconnection rate α from the HS13 data for increasing inference parameter Δt .

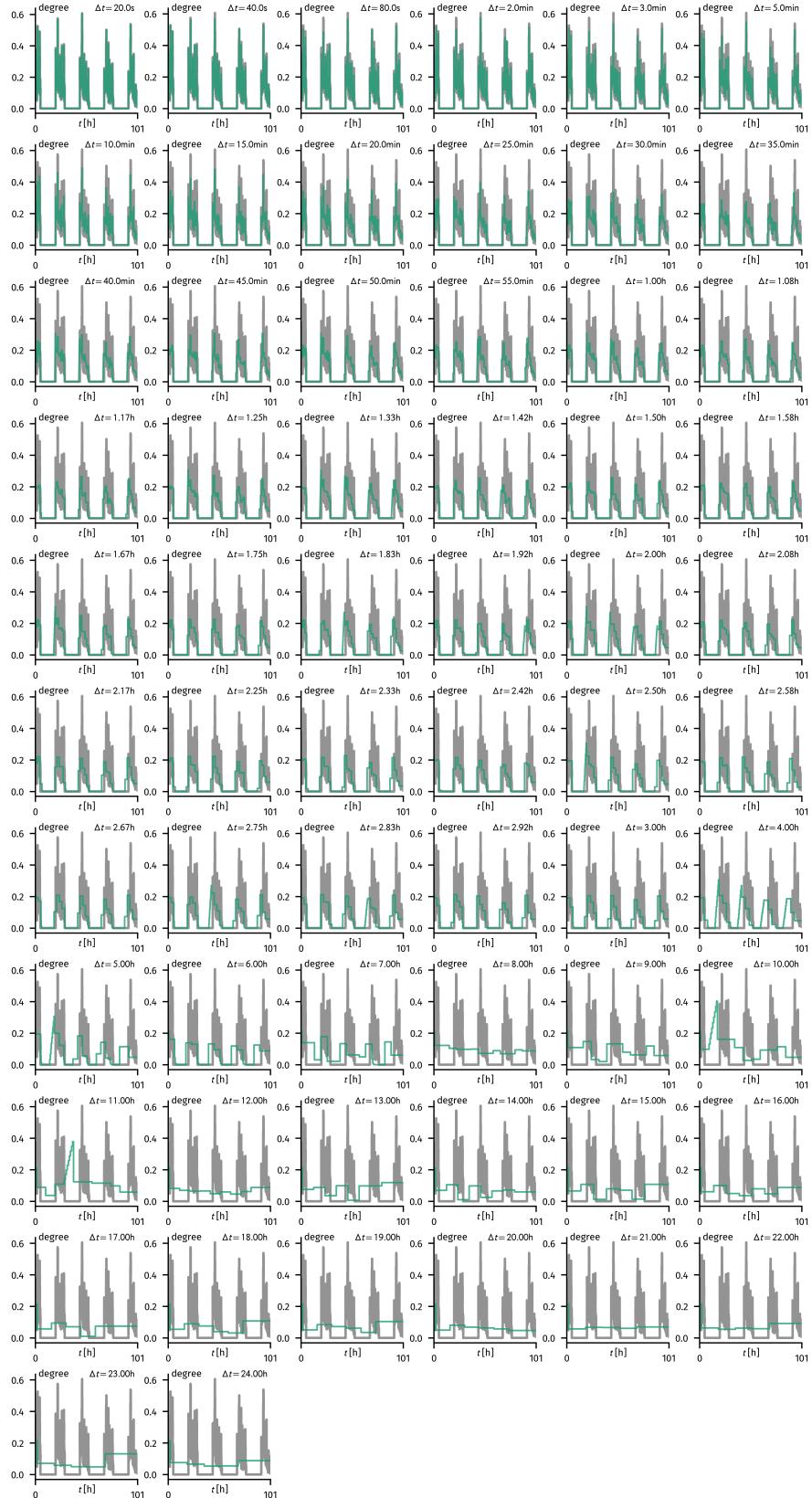


Figure B.10: Inference of the expected mean degree k_0 from the HS13 data for increasing inference parameter Δt and comparison to the actual mean degree $\langle k \rangle(t)$.

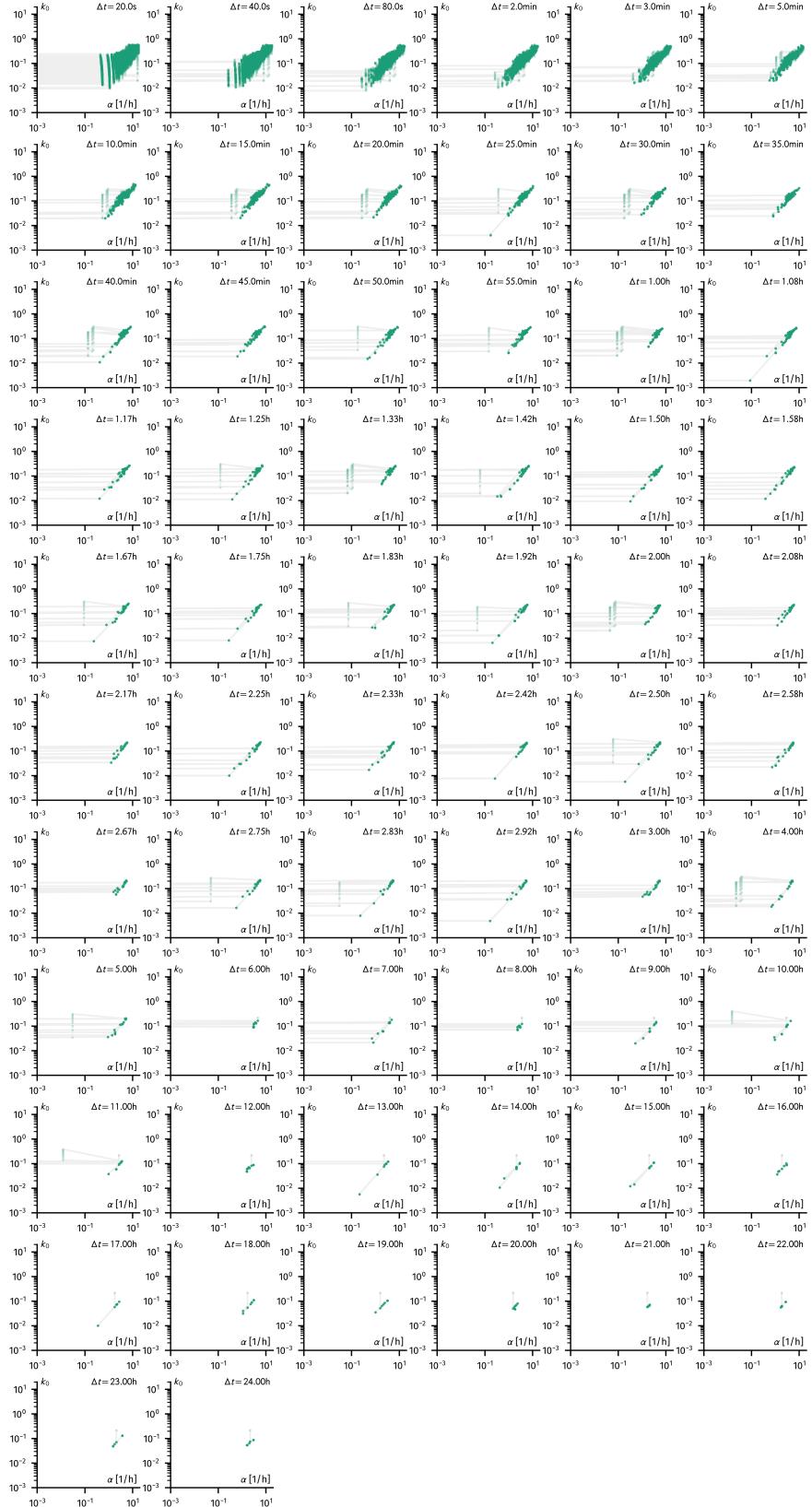


Figure B.11: Inferred trajectory in the (α, k_0) -plane from the HS13 data for increasing inference parameter Δt

C

Gillespie's Stochastic Simulation Algorithm

Gillespie's stochastic simulation algorithm (SSA) is a general algorithm to simulate the evolution of Poisson processes related to discrete entities (e.g. particles, individuals, etc.). It was originally developed to simulate chemical reactions as birth-death processes in well-mixed systems. This Appendix introduces sequentially how and why the algorithm works for general processes with a single constant rate, multiple constant rates (for multiple events), single time-varying rates and multiple time-varying rates. It will close with an outline on how to perform exact simulations for locally constant time-varying rates.

C.1 Homogeneous Poisson Processes

This Section introduces Gillespie's SSA for processes happening with constant rates. For all of those processes consider the initial time to be $t_0 = 0$.

C.1.1 A Single Event

The simplest application for a stochastic simulation algorithm is one where only a single event can take place. This event happens with constant rate λ and is therefore a Poisson process. A Poisson process is based on several assumption as stated in [?].

1. There is at maximum one event happening in an infinitesimal time interval of length dt .
2. The probability of one event happening in this time interval is λdt .
3. The process is Markovian in the sense that whether or not an event happens in this time interval does not depend on the events in any of the previous intervals.

Let $P_0(t)$ denote the probability that no event has taken place up to and including time t . According to assumptions 1 and 2, the proba-

bility that an event has happened in time dt is

$$P_1 = \lambda dt.$$

The probability that no event has taken place is consequently

$$P_0(dt) = 1 - P_1 = 1 - \lambda dt.$$

With Assumption 3 and due to the definition of P_0 , the probability that no event has happened up to and including time $t + dt$ is hence the probability that no event has happened up to time t and no event happened within the subsequent time interval and thus

$$P_0(t + dt) = P_0(t) \times P_0(dt) = P_0(t) - P_0(t)\lambda dt$$

which translates to the ordinary differential equation

$$\frac{dP_0}{dt} = -\lambda P_0$$

with solution

$$P[\text{"no event happened"}, t] = P_0(t) = \exp(-\lambda t).$$

The complementary probability that at least one event has happened up to and including time t is consequently given as

$$\begin{aligned} P[\text{"at least one event happened"}, t] &= 1 - P_0(t) \\ &= 1 - \exp(-\lambda t) \\ &\equiv P(t) \end{aligned}$$

and the corresponding probability density for the time of the first event evaluates to

$$p(t) = \partial_t P(t) = \lambda \exp(-\lambda t).$$

Since this is just an exponential distribution, the algorithm is simply as follows. Draw a time t from the exponential distribution $t \sim \mathcal{E}(\lambda)$ and update the global time. Let the event take place and update the rate λ according to the new circumstances. Then start over with $t_0 = t$.

C.1.2 Multiple Events

Instead of one single event channel which can be triggered, in the following it is assumed that there are M such events, each taking place with constant rate $\lambda^{(i)}$, where $i \in [1, M]$.

Separate Evaluation

Since these processes are independent, one can draw a single potential leap time $t^{(i)}$ for each of the processes. Then, the process j where $t^{(j)} = \min\{t^{(i)}\}$ defines the new time $t \rightarrow t^{(j)}$ and the event which takes place is event j . Consequently, let the event take place, evaluate the new rates according to the new circumstances and start over with $t_0 = t$.

Joint Evaluation

Instead of evaluating each event channel on its own, one can evaluate the probability that none of the event channels have yet fired, which is

$$\begin{aligned} P["\text{no event happened}", t] &= P_0^{(1)}(t) \times \cdots \times P_0^{(M)}(t) \\ &= \prod_{i=1}^M \exp(-\lambda^{(i)} t) \\ &= \exp\left(-\sum_{i=1}^M \lambda^{(i)} t\right) = \exp(-\Lambda t). \end{aligned}$$

Consequently, the probability that at least one event has happened up to time t is given as

$$\begin{aligned} P["\text{at least one event happened}", t] &= 1 - P["\text{no event happened}", t] \\ &= 1 - \exp(-\Lambda t) \\ &\equiv P_\Lambda(t). \end{aligned}$$

The corresponding probability density for the time of the first event happening evaluates to

$$p_\Lambda(t) = \partial_t P_\Lambda(t) = \Lambda \exp(-\Lambda t).$$

Thus, the new time can simply be drawn from this exponential distribution as $t \sim \mathcal{E}(\Lambda)$. However, it is now necessary to decide which channel fired first. As argued above, the probability $\pi_j(t) \times p_\Lambda(t)dt$ that event j is the event which takes place *and* the time of the event is time t is equal to the probability that $t^{(j)} = \min\{t^{(i)}\}$ (the probability that it happens first). For $M = 2$ this is given as

$$\begin{aligned} \pi_1(t)p_\Lambda(t)dt &= P[t^{(1)} = \min\{t^{(i)}\} \text{ and } t^{(1)} = t] = P[t^{(1)} \text{ and } t^{(2)} \wedge t^{(1)} = t] \\ &= dt p^{(1)}(t) \int_t^\infty dt^{(2)} p^{(2)}(t^{(2)}) \\ &= dt p^{(1)}(t) [P^{(2)}(\infty) - P^{(2)}(t)] \\ &= dt \lambda^{(1)} \exp(-t\lambda^{(1)}) \exp(-t\lambda^{(2)}) \\ &= \frac{\lambda^{(1)}}{\lambda^{(1)} + \lambda^{(2)}} dt [\lambda^{(1)} + \lambda^{(2)}] \exp\left(-t[\lambda^{(1)} + \lambda^{(2)}]\right), \end{aligned}$$

such that at any time t , the probability that channel $j = 1$ fired first (given that it happened at time t) is

$$\pi_1(t) = \pi_1 = \frac{\lambda^{(1)}}{\lambda^{(1)} + \lambda^{(2)}}$$

and equivalently for the second event channel $\pi_2 = \lambda^{(2)}/[\lambda^{(1)} + \lambda^{(2)}]$. For $M = 3$ the probability that event channel $j = 1$ is the first one to fire given that the firing time is t is the probability that its

firing time $t^{(1)}$ fulfills $t^{(1)} < t^{(2)} < t^{(3)}$ or $t^{(1)} < t^{(3)} < t^{(2)}$ such that

$$\begin{aligned}
\pi_1(t)p_\Lambda(t)dt &= P[t^{(1)} < t^{(2)} \text{ and } t^{(1)} = t] \times P[t^{(2)} < t^{(3)}] + \\
&\quad + P[t^{(1)} < t^{(3)} \text{ and } t^{(1)} = t] \times P[t^{(3)} < t^{(2)}] \\
&= dt p^{(1)}(t) \left[\int_t^\infty dt^{(2)} p^{(2)}(t^{(2)}) \int_{t^{(2)}}^\infty dt^{(3)} p^{(3)}(t^{(3)}) + \right. \\
&\quad \left. + \int_t^\infty dt^{(3)} p^{(3)}(t^{(3)}) \int_{t^{(3)}}^\infty dt^{(2)} p^{(2)}(t^{(2)}) \right] \\
&= dt p^{(1)}(t) \left[\frac{\lambda^{(2)}}{\lambda^{(2)} + \lambda^{(3)}} \exp(-t[\lambda^{(2)} + \lambda^{(3)}]) + \right. \\
&\quad \left. + \frac{\lambda^{(3)}}{\lambda^{(2)} + \lambda^{(3)}} \exp(-t[\lambda^{(2)} + \lambda^{(3)}]) \right] \\
&= \frac{\lambda^{(1)}}{\lambda^{(1)} + \lambda^{(2)} + \lambda^{(3)}} \times \\
&\quad \times dt [\lambda^{(1)} + \lambda^{(2)} + \lambda^{(3)}] \exp(-t[\lambda^{(1)} + \lambda^{(2)} + \lambda^{(3)}])
\end{aligned}$$

where similar steps have been used as during the evaluation of the case $M = 2$. By considering that evaluations for $M > 3$ will take similar forms one sees that this result generalizes for any M such that

$$\pi_j = \frac{\lambda^{(j)}}{\sum_{k=1}^M \lambda^{(k)}} = \frac{\lambda^{(j)}}{\Lambda}.$$

To sum up, once the event time t is decided, the corresponding event channel j is found by simply drawing a random integer from the interval $[1, M]$, each with probability π_j . Consequently, the event takes place, the rates are updated accordingly and the process starts over with $t_0 = t$.

C.2 Inhomogeneous Poisson Processes

This Section introduces Gillespie SSAs for processes associated with time-varying rates. For all of those processes consider the initial time to be $t_0 = 0$.

C.2.1 A Single Event

As for the homogeneous case, consider a stochastic simulation algorithm where again only a single event can take place, whereas now, this event happens with a time-varying rate $\lambda(t)$ and is therefore associated with an inhomogeneous Poisson process. The assumptions made in Sec. C.1.1 have to be updated by replacing original assumption 2 with

2. The probability of one event happening in the time interval $[t, t + dt]$ is $\lambda(t)dt$.

Let $P_0|_t^{t+dt}$ denote the probability that no event has taken place within this time interval such that the probability that no event has taken place up to and including time t would be denoted as $P_0(t) = P_0|_0^t$. According to assumptions 1 and 2, the probability that an event has happened within the time interval $[t, t + dt]$ is

$$P_1|_t^{t+dt} = \lambda(t)dt,$$

such that the probability that no event has happened is

$$P_0|_t^{t+dt} = 1 - \lambda(t)dt.$$

The probability $P_0(t + dt)$ that no event has taken place up to time $t + dt$ is, again (by assumption 3)

$$P_0(t + dt) = P_0(t) \times P_0|_t^{t+dt} = P_0(t) - P_0(t)\lambda(t)dt$$

which evaluates to the differential equation

$$\frac{dP_0}{dt} = -\lambda(t)P_0$$

with general solution

$$P[\text{"no event happened"}, t] = P_0(t) = \exp \left[- \int_0^t d\tau \lambda(\tau) \right].$$

Consequently, the complementary probability that at least one event has happened is given as

$$\begin{aligned} P[\text{"at least one event happened"}, t] &= 1 - P_0(t) \\ &= 1 - \exp \left[- \int_0^t d\tau \lambda(\tau) \right] \\ &\equiv P(t) \end{aligned} \tag{C.1}$$

with corresponding probability density for the time of the first event happening

$$p(t) = \partial_t P(t) = \lambda(t) \exp \left[- \int_0^t d\tau \lambda(\tau) \right].$$

Considering the variable

$$\theta(t) = \int_0^t d\tau \lambda(\tau) \tag{C.2}$$

and comparing with Eq. (C.1) it becomes clear that θ is distributed according to an exponential distribution with unit mean $\theta \sim \mathcal{E}(1)$. Hence, to find the new time t , draw a random number θ and solve Eq. (C.2) for t (this can be done numerically). Subsequently, let the event take place, update the rate function $\lambda(t)$ accordingly and repeat the process with $t_0 = t$.

C.2.2 Multiple Events

Instead of one single event channel that can be triggered, in the following it is assumed that there's M such events, each taking place with time-varying rate $\lambda^{(i)}(t)$, where $i \in [1, M]$.

Separate Evaluation

Again, since these processes are independent, one can compute a single potential leap time $t^{(i)}$ for each of the processes by drawing M exponentially distributed random numbers $\theta^{(i)} \sim \mathcal{E}(1)$ and solving Eq. (C.2) for $t^{(i)}$. Then, the process j where $t^{(j)} = \min\{t^{(i)}\}$ defines the new time $t \rightarrow t^{(j)}$ and the event which takes place is event j . Consequently, let the event take place, evaluate the new rate functions and start over with $t = t_0$.

Joint Evaluation

Similarly as for the homogeneous event, instead of evaluating each event channel on its own, one can evaluate the probability that none of the event channels have yet fired, which is

$$\begin{aligned} P["\text{no event happened}", t] &= P_0^{(1)}(t) \times \cdots \times P_0^{(M)}(t) \\ &= \prod_{i=1}^M \exp \left(- \int_0^t d\tau \lambda^{(i)}(\tau) \right) \\ &= \exp \left(- \int_0^t d\tau \sum_{i=1}^M \lambda^{(i)}(\tau) \right) \\ &= \exp \left(- \int_0^t d\tau \Lambda(\tau) \right). \end{aligned}$$

Consequently, the probability that at least one event has happened up to time t is given as

$$\begin{aligned} P["\text{at least one event happened}", t] &= 1 - P["\text{no event happened}", t] \\ &= 1 - \exp \left(- \int_0^t d\tau \Lambda(\tau) \right) \\ &\equiv P_\Lambda(t). \end{aligned}$$

Furthermore, the corresponding probability density for the time of the first event happening is

$$p_\Lambda(t) = \partial_t P_\Lambda(t) = \Lambda(t) \exp \left(- \int_0^t d\tau \Lambda(\tau) \right).$$

Again, defining the variable

$$\Theta = \int_0^t d\tau \Lambda(\tau) \tag{C.3}$$

the new time can be found by drawing a random variable $\Theta \sim \mathcal{E}(1)$ and solving for t . Now, given the time of the event t , an event channel to be triggered has to be chosen. The probability $\pi_j(t) \times p_\Lambda dt$ that

event j is the event which takes place (given that the time t which was drawn from p_Λ) is equal to the probability that $t^{(j)} = \min\{t^{(i)}\}$ (the probability that the event of channel j happens first). For $M = 2$ this is given as

$$\begin{aligned}\pi_1(t) \times p_\Lambda dt &= P[t^{(1)} = \min\{t^{(i)}\} \text{ and } t^{(1)} = t] = P[t^{(1)} < t^{(2)} \text{ and } t^{(1)} = t] \\ &= dt p^{(1)}(t) \int_t^\infty dt^{(2)} p^{(2)}(t^{(2)}) \\ &= dt p^{(1)}(t) [P^{(2)}(\infty) - P^{(2)}(t)] \\ &= dt \lambda^{(1)}(t) \exp \left[- \int_0^t d\tau \lambda^{(1)}(\tau) \right] \times \\ &\quad \times \exp \left[- \int_0^t d\tau \lambda^{(2)}(\tau) \right] \\ &= dt \lambda^{(1)}(t) \exp \left[- \int_0^t d\tau [\lambda^{(1)}(\tau) + \lambda^{(2)}(\tau)] \right].\end{aligned}$$

Given Eq. (C.3) one finds

$$d\Theta = \Lambda(t)dt$$

and thus

$$\pi_1(t(\Theta)) \exp(-\Theta) d\Theta = \frac{\lambda^{(1)}(t(\Theta))}{\Lambda(t(\Theta))} \exp(-\Theta) d\Theta.$$

Hence, given a random variable $\Theta \sim \mathcal{E}(1)$ and time $t(\Theta)$ as the (potentially numerical) solution of Eq. (C.3), the time-dependent probability that channel $j = 1$ fired at this time is

$$\pi_1(t) = \frac{\lambda^{(1)}(t)}{\Lambda(t)}.$$

This result is generalizable to $M > 2$. To see why this is the case, consider $M = 3$ where

$$\begin{aligned}
\pi_1(t) \times p_\Lambda dt &= P[t^{(1)} < t^{(2)} \text{ and } t^{(1)} = t] \times P[t^{(2)} < t^{(3)}] + \\
&\quad + P[t^{(1)} < t^{(3)} \text{ and } t^{(1)} = t] \times P[t^{(3)} < t^{(2)}] \\
&= dt p^{(1)}(t) \left[\int_t^\infty dt^{(2)} p^{(2)}(t^{(2)}) \int_{t^{(2)}}^\infty dt^{(3)} p^{(3)}(t^{(3)}) + \right. \\
&\quad \left. + \int_t^\infty dt^{(3)} p^{(3)}(t^{(3)}) \int_{t^{(3)}}^\infty dt^{(2)} p^{(2)}(t^{(2)}) \right] \\
&= dt p^{(1)}(t) \left[\int_t^\infty dt^{(2)} p^{(2)}(t^{(2)}) \exp \left[- \int_0^{t^{(2)}} dt \lambda^{(3)}(t) \right] + \right. \\
&\quad \left. + \int_t^\infty dt^{(3)} p^{(3)}(t^{(3)}) \exp \left[- \int_0^{t^{(3)}} dt \lambda^{(2)}(t) \right] \right] \\
&= dt p^{(1)}(t) \times \\
&\quad \times \left[\int_t^\infty d\tau \lambda^{(2)}(\tau) \exp \left[- \int_0^\tau dt' (\lambda^{(3)}(t') + \lambda^{(2)}(t')) \right] + \right. \\
&\quad \left. + \int_t^\infty d\tau \lambda^{(3)}(\tau) \exp \left[- \int_0^\tau dt' (\lambda^{(3)}(t') + \lambda^{(2)}(t')) \right] \right] \\
&= dt p^{(1)}(t) \exp \left(- \int_0^t d\tau (\lambda^{(3)}(\tau) + \lambda^{(2)}(\tau)) \right) \\
&= dt \lambda^{(1)}(t) \exp \left[- \int_0^t d\tau \lambda^{(1)}(\tau) \right] \times \\
&\quad \times \exp \left[- \int_0^t d\tau (\lambda^{(3)}(\tau) + \lambda^{(2)}(\tau)) \right] \\
&= dt \left(\frac{\lambda^{(1)}(t)}{\Lambda(t)} \right) \Lambda(t) \exp \left(- \int_0^t d\tau \Lambda(\tau) \right).
\end{aligned}$$

With increasing M one can see that the equations will take forms similar to the ones above such that one can apply recursive solutions and end with the same general equation. Hence, with $\Lambda(t) = \sum_{i=1}^M \lambda_i(t)$ and Eq. (C.3) one finds, again,

$$\pi_j(t) = \frac{\lambda^{(j)}(t)}{\Lambda(t)}. \quad (C.4)$$

Consequently, after finding the time leap t , choose a random integer j from the interval $[1, M]$ with corresponding probability $\pi_j(t)$.

C.2.3 Application to Locally Constant Rates

A special case of an inhomogeneous Poisson process is one where event rates are time-varying but locally constant functions such that

$$\begin{aligned}
\lambda^{(i)}(t) &= \lambda_n^{(i)}, \quad t \in [t_n, t_{n+1}), \text{ and} \\
\Lambda(t) &= \Lambda_n, \quad t \in [t_n, t_{n+1}),
\end{aligned}$$

which is the case when the rates depend on changes of discrete variables such as contacts in time-varying networks or when using zero order approximations of continuous functions. Then, the integral Eq. (C.3) is exactly solvable. Considering that the process starts at time t_0 , define $\Theta_0 = 0$. Draw a random number $\tilde{\Theta} \sim \mathcal{E}(1)$. Now, iteratively solve the integral as

$$\Theta_{n+1} = \Theta_n + \Lambda_n(t_{n+1} - t_n)$$

until $\tilde{\Theta} \leq \Theta_{n+1}$. The time t corresponding to $\tilde{\Theta}$ consequently lies within the time interval $(t_n, t_{n+1}]$. Find the exact time of the event by solving the linear solution of this part of the integral such that

$$t = t_n + \frac{\tilde{\Theta} - \Theta_n}{\Lambda_n}.$$

Subsequently, choose an event channel by drawing a random integer j from the interval $[1, M]$, each with probability $\pi_j = \lambda_n^{(j)} / \Lambda_n$, let the event take place, evaluate the new intervals $[t_n, t_{n+1})$, the corresponding rates λ_n , and start over with $t_0 = t$.

D

SIS-Simulations on Temporal Network Models

Simulations of the SIS model on both the Flockwork model as well as the edge activity model were performed in the following way. The time scale was fixed using the active reconnection rate $\alpha = 1$ and the number of nodes was chosen as $N = 10^4$. The control parameters c_0 , k_0 and R_0 were varied.

In order to choose k_0 , the reconnection probability P was chosen as $P \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95\}$, then k_0 was calculated exactly from Eq. (8.7) for the Flockwork model. The corresponding parameter value for the edge activity model was consequently given as $p = k_0/(N - 1)$. For c_0 , the recovery rate ρ was given as $\rho \in \{0.01, 0.1, 0.5, 1, 2, 5, 10\}$ (in units of $\alpha = 1$). From k_0 and $\alpha = 1$, the Flockwork event rate $\gamma = \alpha/P$ was calculated such that simulations can be performed using the γ - P -formulation of the model. The corresponding parameter of link decay rate ω^- for the edge activity model was subsequently calculated using Eq. (8.14). Finally, the basic reproduction number was varied as $R_0 \in [10^{0.5}, 10^2]$ with 30 values evenly sized in logarithmic scale with base 10.

For each combination of those parameters, 40 simulations were performed for both the Flockwork model and the edge activity model to find the ratio I^* .

Concerning the Flockwork model, for each of those simulations, one Flockwork equilibrium configuration was sampled to be used as an initial configuration at $t = 0$ using Alg. 1 with the options “Poisson distribution” and “shuffled group sizes”. Then, 100 nodes were randomly chosen to be infected initially. Subsequently, a homogeneous joint Gillespie-simulation was performed with Flockwork event rate γ , infection rate per SI-contact η and recovery rate per infected ρ .

Considering the edge activity model, for each of those simulations of any control parameter combination, one equilibrium configuration was chosen as a sample from the Erdős–Rényi model with N nodes and connection probability $p = k_0/(N - 1)$. Then, 100 nodes were randomly chosen to be infected initially. Subsequently, a homogeneous joint Gillespie-simulation was performed with global link decay rate ω^- , link creation rate $\omega^+ = \omega^- p/(1 - p)$, infection rate per

Table D.1: Chosen values of the equilibration time t_{eq} for SIS simulations on both the Flockwork model as well as the edge activity model.

recovery rate $\rho [\alpha]$	repr. number R_0	eq. time $t_{\text{eq}} [(N\alpha)^{-1}]$
< 0.1	≥ 20	100
	$10 \leq R_0 < 20$	170
	$6 \leq R_0 < 10$	280
	$4 \leq R_0 < 6$	400
	$2.6 \leq R_0 < 4$	590
	$1.89 \leq R_0 < 2.6$	1000
	$1.36 \leq R_0 < 1.89$	2000
	$R_0 < 1.36$	4000
≥ 0.1	≥ 6	40
	$4 \leq R_0 < 6$	100
	$1.5 \leq R_0 < 4$	140
	$1 \leq R_0 < 1.5$	180 (if $k_0 \geq 1$) 500 (if $k_0 < 1$)
	$R_0 < 1$	500

SI-contact η and recovery rate per infected ρ .

Each system was equilibrated until a time $t = t_{\text{eq}}$ which is given in Tab. D.1. Afterwards, the measurement phase of the simulation was performed until $t = t_{\text{eq}} + 200(N\alpha)^{-1}$, sampling the value of $I^*(t)$ and $R_0(t) = \langle k \rangle(t)\eta/\rho$ after each $\Delta t = (N\alpha)^{-1}$. Consequently, I^* was calculated as an average over those 201 recorded values, as well as the corresponding value R_0 to account for possibly large fluctuations in the network structure during simulation. However, if the simulation ended with $I^* = 0$ at any time between $t = 0$ and $t = t_{\text{eq}} + 200(N\alpha)^{-1}$ then the average was not built over the values collected during the measurement. Instead, the simulation was counted as resulting in $I^* = 0$. For the edge activity model, simulations for $k_0 = k_0(P = 0.95)$ and $c_0 = 100$ were omitted because of the extensive runtime of the simulations. In a second analysis,

E

Fitting Distributions to Models

Acknowledgments

I want to thank my pet mosquito Guido.

Selbständigkeitserklärung

Hiermit erkläre ich, die vorliegende Arbeit selbständig ohne fremde Hilfe verfasst und nur die angegebene Literatur und Hilfsmittel verwendet zu haben.

Ort

Datum

Benjamin F. Maier