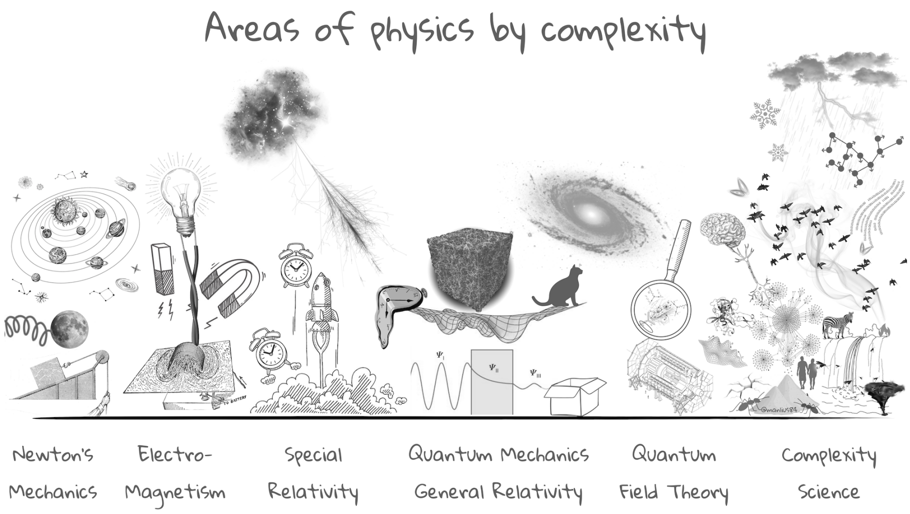
Final Report

Physics of Complex Networks: Structure and Dynamics

UNIVERSITÀ DEGLI STUDI DI PADOVA



PoCN Final Project

Conte, William

Last update: September 22, 2025

Contents

1 | Task #30: Axelrod model for dissemination of culture

1.1 | Introduction

The Axelrod model is an agent-based model used to study how social influence shapes cultural traits in individuals. In this context, the term "culture" refers to a set of traits, beliefs, or characteristics that an agent possesses and that may be shared or modified through interactions with neighboring agents [1]. The more similar these traits are, the more likely two neighboring agents will interact.

The objective of this model is to investigate how, and under what conditions, interactions between individual agents can create regions with a shared culture (consensus) as opposed to regions with different cultures (polarization) [4]. Agent-based modeling provides a bottom-up approach in which global effects emerge from local interactions governed by a precise set of rules [1].

The Axelrod model considers N agents as nodes of a network (or lattice). Each agent i has a vector of F cultural features, $i = (i_1, i_2, \ldots, i_F)$, where each feature i_f can take one of q integer values $(1, \ldots, q)$, initially assigned independently with equal probability $1/q$. The system evolves in discrete time steps as follows [4]:

1. A connected pair of agents (i, j) is selected at random.
2. The overlap l_{ij}, i.e., the number of shared features, is computed.
3. If $0 < l_{ij} < F$, the link is **active**, and i and j interact with probability l_{ij}/F. In that case, a feature g for which $i_g \neq j_g$ is selected randomly, and the corresponding value of i is updated: $i_g = j_g$.

The model has q^F possible cultural configurations. Consensus is reached when a single configuration occupies the entire network; otherwise, the system is polarized. To quantify the system's state, we measure the size of the largest cultural domain, $S_{\max}$, often normalized by N. Here, $S_{\max}$ is defined as the size of the largest connected component of agents sharing the same culture.

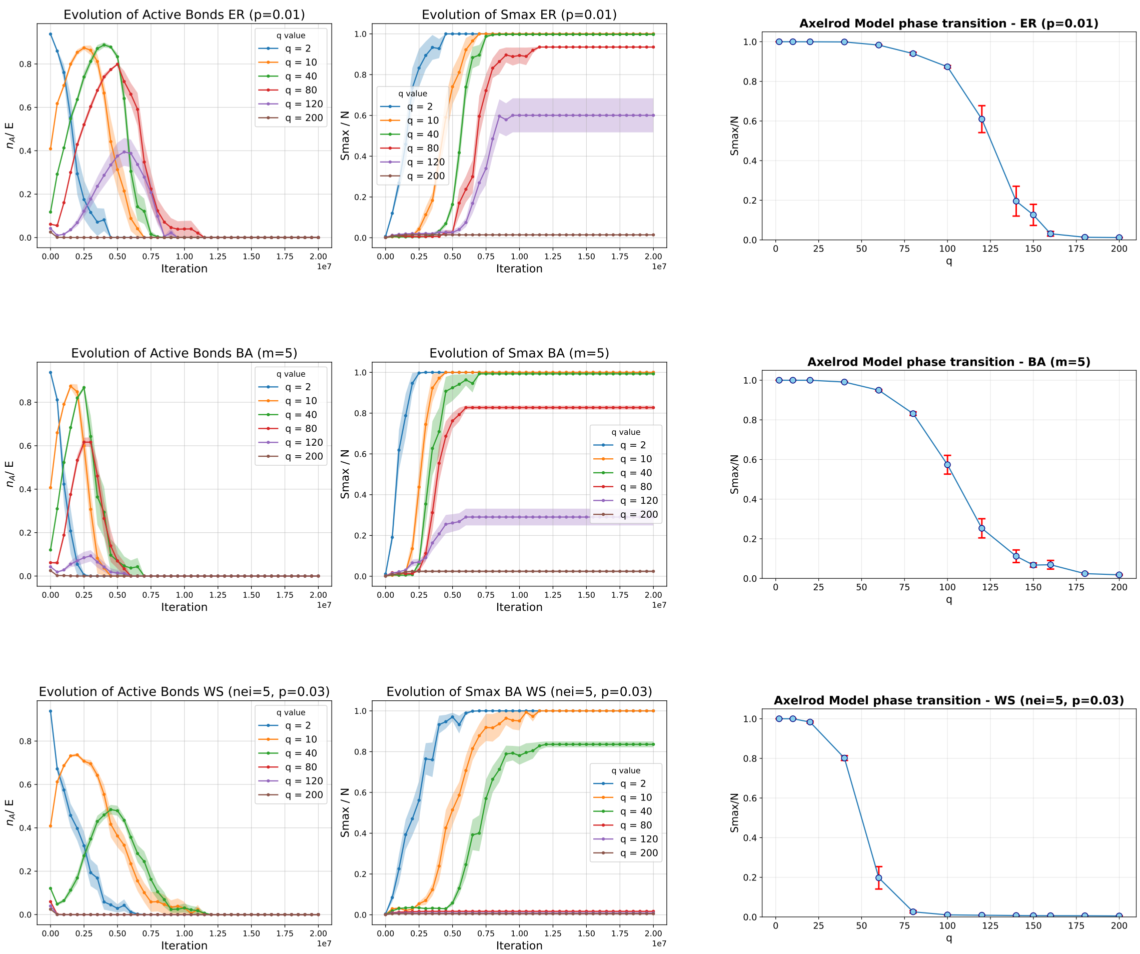
This model is notable because, for a fixed F, there exists a critical value of q, denoted q_c, at which the system undergoes a phase transition from consensus to polarization, with $S_{\max}/N$ as the order parameter and q as the control parameter [2].

1.2 Numerical simulations

Given the goal of this model, the dynamics have been studied under different synthetic and real topologies (see Supplementary Material for this last one), with a fixed number of cultural features $F = 5$. In particular, regarding the synthetic networks (N denotes the number of nodes):

- **Erdős-Rényi network:** $N = 1000$ and $p = 0.01$;
- **Barabási-Albert network:** $N = 1000$ and $m = 5$;
- **Watts-Strogatz network:** $N = 1000$, $nei = 5$, and $p_{ws} = 0.03$ (where nei is the number of neighbors for each node and p_{ws} is the rewiring probability).

For 10 runs of 2×10^7 iterations each, we tracked the number of active links (referred to as "bonds" in the plots), n_A, and the size of the largest cultural domain, $S_{\max}$. On the left, we show the evolution of these quantities throughout the dynamics, while on the right, a phase transition plot is presented.



The simulation results show that network topology strongly affects the critical threshold for cultural polarization, with scale-free networks promoting consensus more effectively than random or small-world networks. Low q values lead to rapid equilibration and consensus (active links decay quickly to zero), while high q values ($q > q_c$)

produce persistent cultural boundaries, with active links remaining non-zero for extended periods. A natural extension is to repeat the simulations on larger networks to reduce finite-size effects and to explore how varying model parameters influences the results.

2 | Task #40: Subways II

2.1 | Introduction

The objective of this data task was to extract node and edge lists from the raw data of subway networks across multiple cities worldwide. To achieve this, three types of scripts were developed, depending on the available information:

1. When the topology for a given year was provided as an adjacency matrix, it was used as the primary source to determine the edges.
2. In the absence of an adjacency matrix, the adjacency list for each year was utilized.
3. If neither format was available (as was the case for the Chicago subway), edges were inferred from the "lines" data.

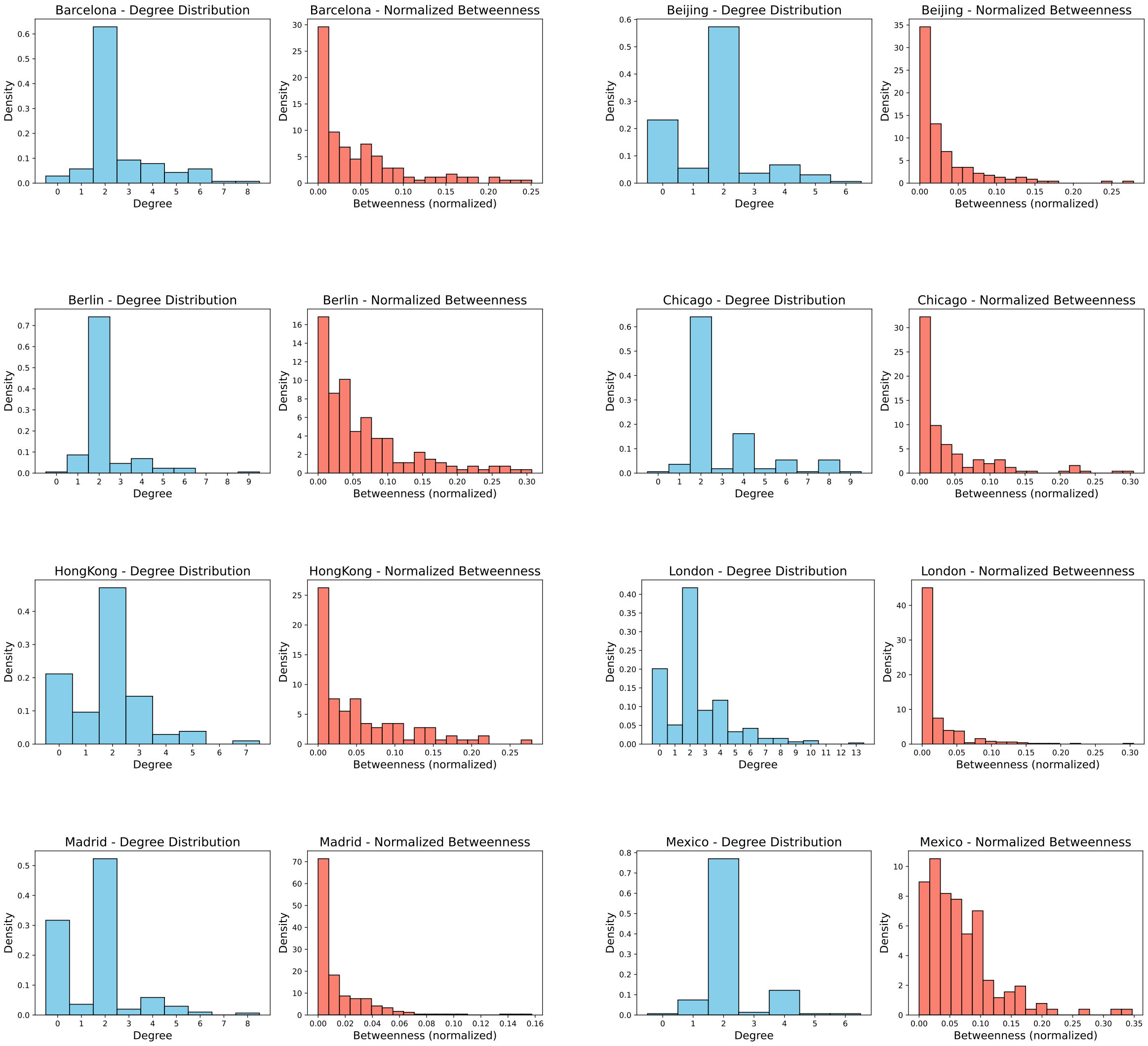
In instances where specific information was missing from the raw files, the corresponding entries were filled with the placeholder "unknown" to preserve the network structure and avoid the removal of any nodes or edges. Data were only removed in cases where there was no correspondence between stations in the files used to infer edges/node IDs and those used to define the nodes.

The line associated with each edge in the edge list was inferred from the "line" files, where available, or from station information files when such data were absent. Mismatches between station names in the node and edge files were resolved using a function that performs a fuzzy string match, accepting matches with at least 70% character similarity to account for differences in case or whitespace. Node metadata, including latitude, longitude, and year, were extracted from the "positions-years" files. The resulting node and edge lists were subsequently organized into the folder "nodes_edges".

Having constructed the node and edge lists for the subway networks, we next perform a basic network analysis to characterize their structural properties.

2.2 Basic network analysis

To understand the structural properties of the network, we analyze the distributions of node degrees and betweenness centrality, which provide insights into node connectivity and their roles in facilitating information flow.



We further investigate the network's structural characteristics by analyzing the average degree, assortativity, and shortest path length across nodes, providing insights into centrality patterns, degree correlations, and network compactness.

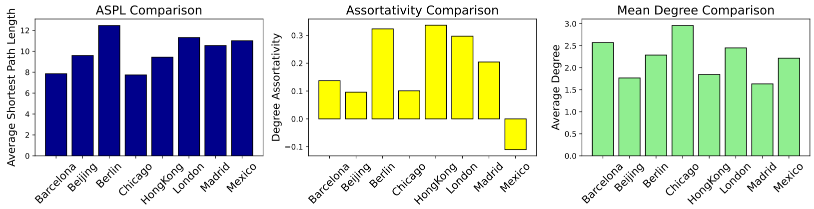


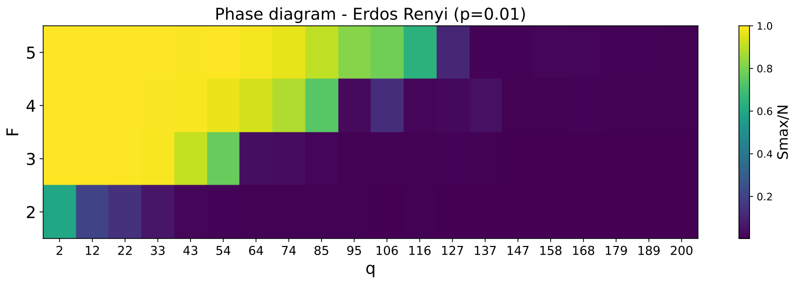
Figure 2.1: Comparison plot of some network's characteristics.

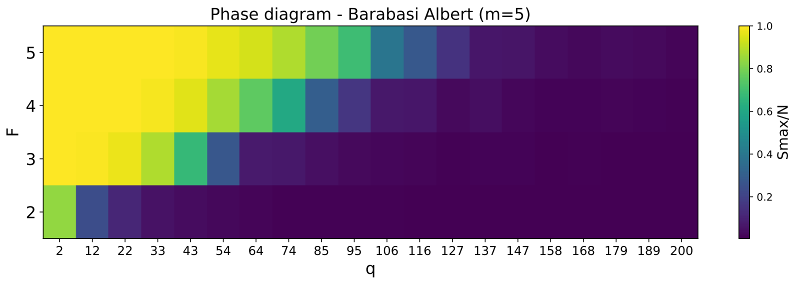
hyperref

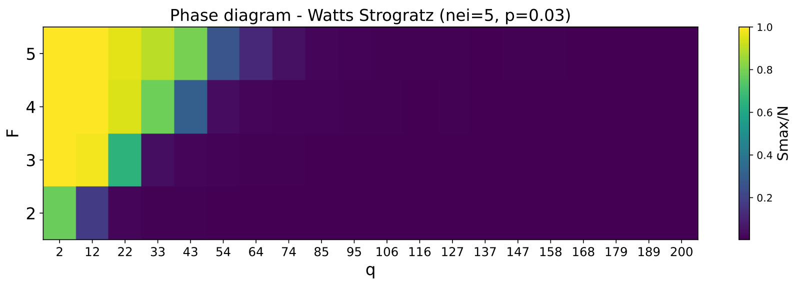
A | Supplementary material

A.1 | Axelrod model: phase diagrams

In addition to what is presented in the main section of the report, we also show how the critical value q_c varies with F for a given network topology. To this end, we repeated the same simulations as before, varying the value of F at each run. The network parameters and sizes were kept fixed, and each simulation was run once for 2×10^7 iterations. The goal was to reproduce a phase diagram of the Axelrod model for a given topology. Due to computational constraints, we could not explore higher values of F, although doing so would provide more complete results.







In addition to the observations reported in the main section, the phase diagrams show that increasing the value of F also increases q_c, as well as the number of iterations required for the system to reach an absorbing state. This behavior reflects the fact that higher cultural complexity (larger F) increases the number of possible configurations, thus raising the threshold q_c and slowing down convergence to consensus or polarization.

A.2 | Axelrod model: real network application

In this section, we perform a further analysis using a real network, namely the "B.F. Maier's FB friends network," available at this GitHub repository [3]. The dataset consists of an anonymized Facebook friendship network collected in fall 2014. Nodes represent Facebook profiles, and edges represent friendships between them. The node corresponding to the profile owner, which was connected to every other node, has been removed.

Some basic network statistics are as follows:

- Undirected: True
- Number of nodes: $N = 362$
- Number of edges: 1988
- Mean degree: 10.98
- Connected components: 20
- Nodes in largest component: 329
- Global clustering coefficient $C = 0.51$
- Average path length $= 3.58$

We also plotted the degree distribution, shown in Fig. A.1.

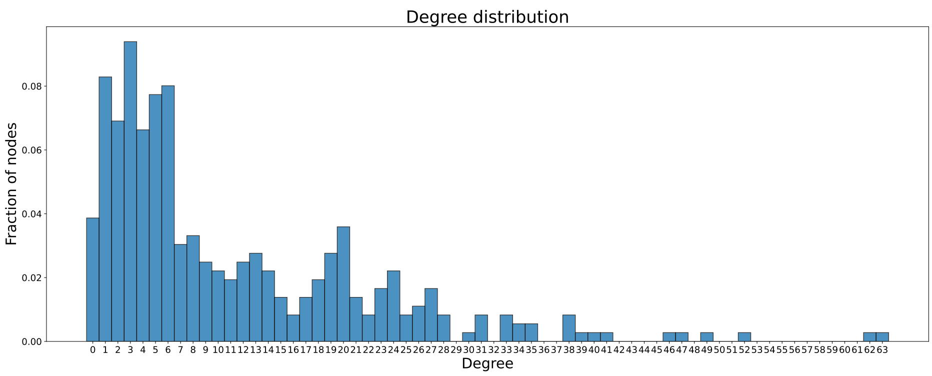


Figure A.1: Degree distribution of the B.F. Maier FB friends network.

We performed the same analysis as before and produced the same type of phase transition plot, comparing the real network with synthetic networks of the same size. The parameters of the synthetic networks were chosen to match the average degree of the real network ($p = 0.03$ for Erdős-Rényi, $m = 5$ for Barabási-Albert, and $nei = 5$, $p_{ws} = 0.03$ for Watts-Strogatz), in order to ensure comparability across topologies.

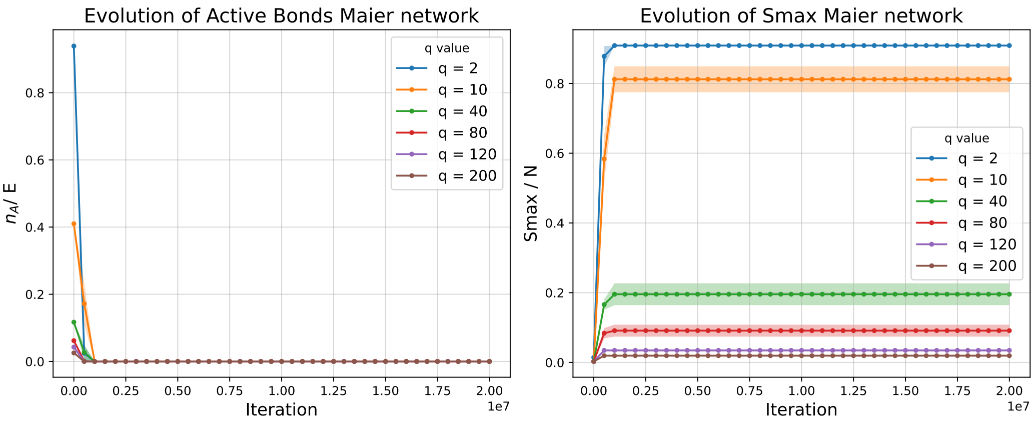


Figure A.2: Dynamics of the real network.

As shown in Fig. A.2, the number of active bonds drops rapidly to zero, indicating that the system reaches an absorbing state almost immediately. This behavior is likely due to the fragmented structure of the network and the low initial overlap between neighboring nodes.

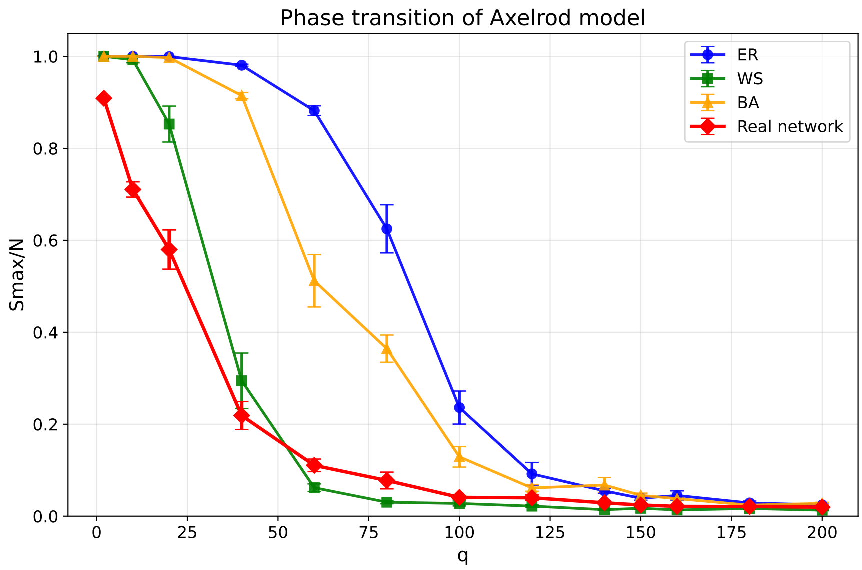


Figure A.3: Phase transition of the real network.

From Fig. A.3, we observe that the system behaves similarly to a small-world network, which is expected given its social network origin and the observed clustering and average path length. It is important to remark that this phase transition plot is qualitative, as the network contains fewer than 400 nodes and finite-size effects are dominant.

A.3 | Subways II: results discussion

Here a table with the main results is presented for visual and numerical clarity:

Network	Num nodes	Num edges	Assortativity	Avg short. path length	Avg degree	Avg betweenness
Barcelona	140	180	0.137	7.85	2.57	0.047
Beijing	164	145	0.096	9.59	1.77	0.031
Berlin	174	199	0.323	12.46	2.29	0.064
Chicago	167	247	0.101	7.74	2.96	0.040
HongKong	104	96	0.336	9.43	1.85	0.051
London	333	408	0.297	11.31	2.45	0.020
Madrid	306	250	0.204	10.55	1.63	0.015
Mexico	148	164	-0.110	11.01	2.22	0.068

Table A.1: Network statistics for the analyzed subway networks.

Network sizes range from Hong Kong's compact 104 nodes to London's extensive 333-station system. Assortativity values span from Mexico's negative correlation (-0.110) to Hong Kong's positive correlation (0.336), indicating fundamentally different connectivity patterns where some networks favor hub-to-hub connections while others distribute load through hub-to-peripheral designs. Chicago demonstrates optimal connectivity-efficiency balance with the highest average degree (2.96) and short path lengths (7.74), while Madrid's sparse connectivity (1.63 average degree) results in longer traversal times. Betweenness centrality variations highlight structural vulnerabilities, with Mexico's high value (0.068) suggesting critical node dependencies (as we can see in the betweenness distribution plot), whereas London's distributed architecture (0.020) provides greater robustness despite its scale.

A.4 | Subways II: Average shortest path length vs degree assortativity

To further examine the relationship between network topology and structural efficiency, we analyze the correlation between assortativity and average shortest path length across the subway networks, with node degree represented by color intensity and point size indicating network scale.

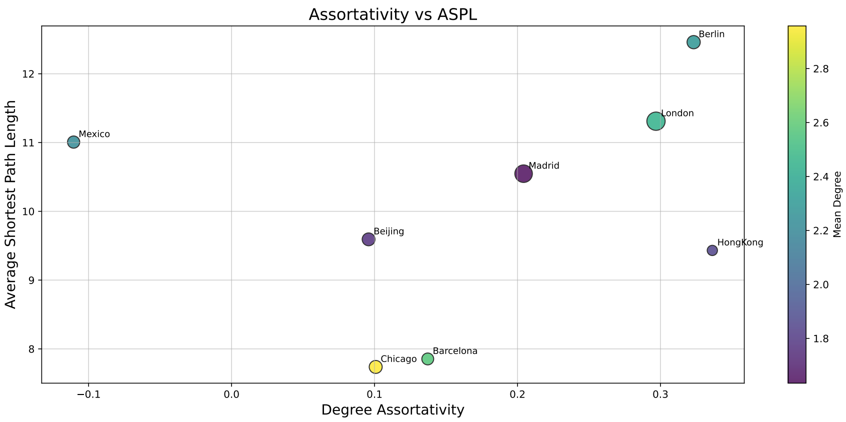


Figure A.4: Correlation plot between average shortest path length and degree assortativity

The scatter plot seems to indicate a positive correlation between assortativity and average shortest path length, suggesting that networks where high-degree nodes preferentially connect to other high-degree nodes tend to have longer average traversal distances. This finding indicates that hierarchical, hub-centric designs may actually reduce overall network efficiency by forcing paths to route through central connection points rather than providing more direct routes. Networks with lower or negative assortativity, such as Mexico and Beijing, achieve shorter average path lengths despite having fewer highly connected nodes, suggesting that these systems prioritize direct connectivity over hierarchical organization. The point sizes reveal that this relationship holds across different network scales, from compact systems like Hong Kong to bigger networks like London. Interestingly, larger networks tend toward higher assortativity but maintain relatively moderate path lengths, indicating that scale allows for more sophisticated hierarchical designs without completely sacrificing efficiency.

While this analysis suggests a positive relationship between assortativity and average shortest path length, the small sample size (n=8) and potential confounding factors (geography, historical development) limit the statistical significance of this observation.

B | Bibliography

[1] Robert Axelrod. The dissemination of culture: A model with local convergence and global polarization. *The Journal of Conflict Resolution*, 41(2):203–226, 1997. ISSN 00220027, 15528766. URL http://www.jstor.org/stable/174371.

[2] Claudio Castellano, Matteo Marsili, and Alessandro Vespignani. Nonequilibrium phase transition in a model for social influence. *Phys. Rev. Lett.*, 85:3536–3539, Oct 2000. doi: 10.1103/PhysRevLett.85.3536. URL https://link.aps.org/doi/10.1103/PhysRevLett.85.3536.

[3] Benjamin F. Maier and Dirk Brockmann. Cover time for random walks on arbitrary complex networks. *Physical Review E*, 96(4):042307, 2017. doi: 10.1103/PhysRevE.96.042307.

[4] M.S. Miguel, V.M. Eguiluz, R. Toral, and K. Klemm. Binary and multivariate stochastic models of consensus formation. *Computing in Science Engineering*, 7 (6):67–73, 2005. doi: 10.1109/MCSE.2005.114.