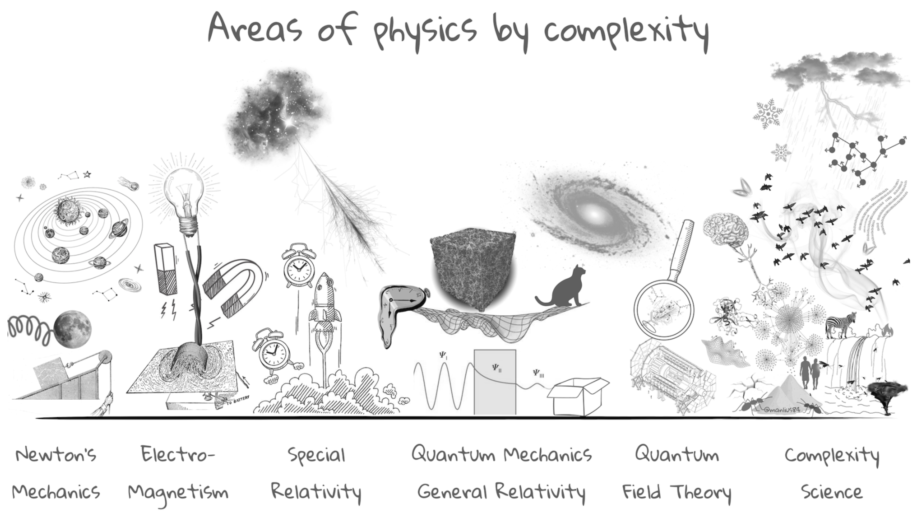
Università
degli Studi
di Padova

Final Report
Physics of Complex Networks: Structure and Dynamics



Final Report

Carotenuto, Jacopo

Last update: September 15, 2024

Contents

1 | Task # 15 - Cascade Failures: SOC Sandpile Model

Task leader(s): *Jacopo Carotenuto*

1.1 | Introduction

In this task, the Bak-Tal-Wiesenfeld Sandpile Model[11] is studied on different types of networks. This model is a type of dynamic process that can be simulated on a network that exhibits self-organized criticality (SOC). The model has been studied in different kinds of networks, and some interesting relations have emerged, especially about the distribution of the avalanche size and duration.

1.2 | Model Description

The Bak-Tal-Wiesenfeld Sandpile Model[11] is a type of dynamic process that can be simulated on a network that exhibits self-organized criticality (SOC). The mechanism of this model is rather simple:

1. The network is initialized with every node having zero height h and having threshold $z_i = k_i$ where k_i is the degree of the node
2. At every time step, one randomly chosen node i height is incremented by 1.
3. If the height exceeds the threshold, the node sheds some of its height to its neighbor in this manner: $h_i \to h_i - k_i$ and $h_j \to h_j + 1$ where j is the neighbors of node i
4. A small fraction f of the height of node i is lost. This acts as a sink to prevent the overloading of the system. f is called the "shedding probability."

Numerous works have studied this type of dynamic, and some interesting relations have emerged, especially about the distribution of the avalanche size and duration. Both quantities seem to have a power-law distribution of this type:

$$p(x) \sim x^{-\tau_x} \exp(-x/c_x)$$

Where x is either the avalanche size or the avalanche duration, τ_x is the associated exponent, and c_x is the characteristic value of the quantity. [10]

These values vary based on the type of network and parameters used for the simulations, but the underlying power-law distribution holds for all types of networks. In [10] it's found that, for the avalanche size, $\tau = \frac{\gamma}{\gamma-1}$ in scale-free networks with power-law degree distribution with exponent γ (in range $2 < \gamma < 3$).

Other works [8] analyzed different types of networks with different degree distributions and found similar power-law distributions. In this task, the model is simulated on eight different types of networks (Gaussian Degree Distribution, Uniform Degree Distribution, Static Scale Free Network, Erdos-Renyi Network, and Barabasi-Albert Network), and the corresponding distributions are calculated.

1.3 Simulations

The Sandpile model was implemented in the Julia programming language [7]. Each network was generated with $100k$ nodes, and 2 million grains were simulated. The shedding probability was 10^{-4} for all networks. The Cascade Size is defined as the number of nodes (with the possibility of repeating nodes) that the cascade affects. The Cascade Duration is the number of toppling events until relaxation. The following networks were used:

- Gaussian Degree Distribution ($\mu = 20, \sigma = 6$)
- Uniform Degree Distribution ($[1, 40]$)
- Static Scale Free Network ($\gamma = 2, 2.5, 3, 4$) [9]
- Erdos-Renyi Network ($p = 2 \cdot 10^{-4}$)
- Barabasi-Albert Network ($k = 10$)

Here, we can see some examples of distributions, where in blue is the simulation data and in red is the fitted power-law. For the size-duration relation, a simple exponential relation was hypothesized according to [10], where the exponent z is called "Dynamic Exponent". In both distributions, the tail was cut to exclude noise from finite-size effects.

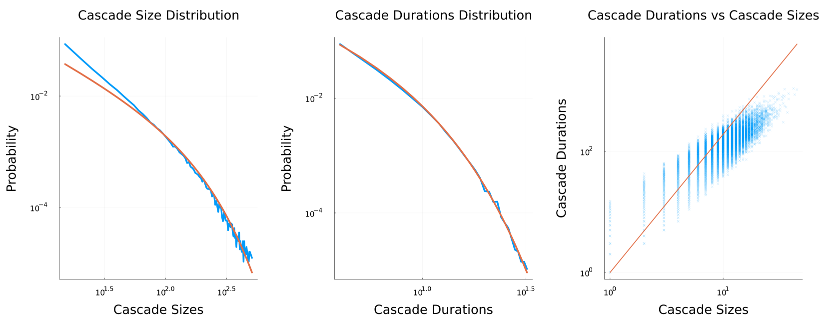


Figure 1.1: Static Scale-Free Network, $\gamma = 2.5$

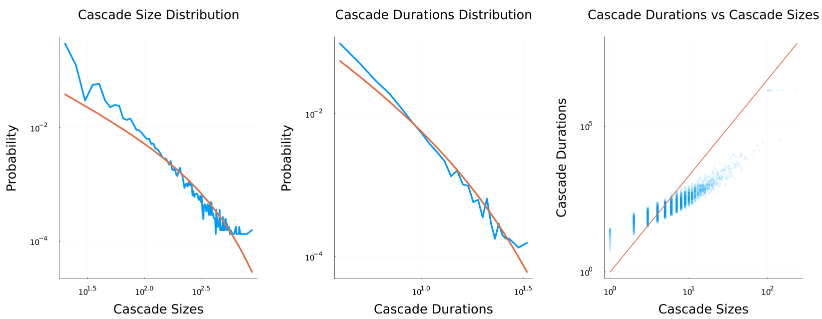


Figure 1.2: Erdos Renyi Network, $p = 2 * 10^{-4}$

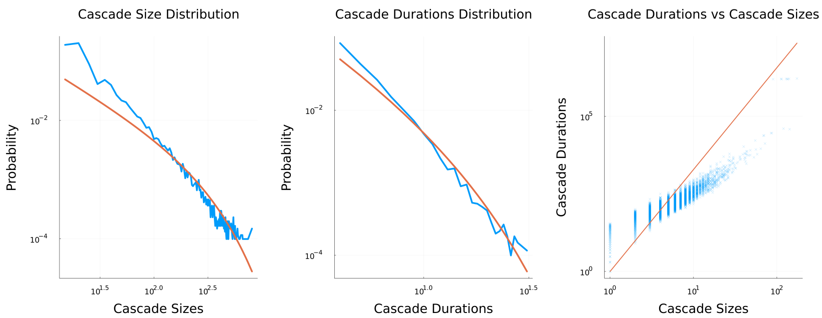


Figure 1.3: Gaussian Degree Distribution, $\mu = 20, \sigma = 6$

For the scale-free networks, the τ_{size} were not exactly following the value predicted by [10], but it was near enough to hypothesize that the difference was caused by the low number of simulations performed here. Here are the results for the τ_{size} and their expected values (slight abuse as with $\gamma = 4$ we could not use the formula provided by [10])

γ	2	2.5	3	4
Simulated	1.98	1.16	1.49	1.27
Expected	2	1.66	1.5	1.33

We can compare all the size distributions by plotting them all together:

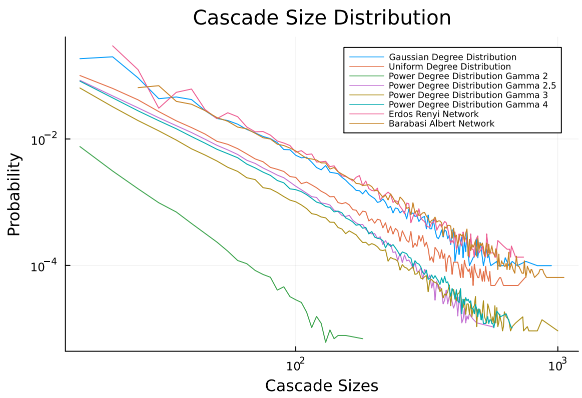


Figure 1.4: All Networks Size Distributions

2 | Task # 34 - Sociophysics: Game Theory on networks

Task leader(s): *Jacopo Carotenuto*

2.1 | Introduction

In this task, the evolution of strategies will be studied according to different "survival" rules. We will consider the "Ultimatum Game" on different networks with different types of players and strategies.

2.2 | Playing the Ultimatum Game on Networks

The setup of [12] will be followed. The Ultimatum Game is a two-player game where one player, the proposer, is given a sum of money (here 1) and proposes how to divide it between himself and the other player. The other player, the responder, can either accept or reject the proposal. If the responder accepts, the money is divided as proposed. If the responder rejects, both players receive nothing. The game is played twice, with the roles of proposer and responder reversed in the second round.

Different network types will be considered, where each node represents a player, and each edge represents a connection between players. The players will play only with their neighbors. Each player will have two numbers associated with it (from 0 to 1): p, the division offered when playing as the proposer, and q, the threshold for accepting the division when playing as the responder. There will be three types of players:

(A) **Fair Player**: The player will have $p = q$.

(B) **Pragmatic Player**: The player will have $p = 1 - q$. This player wants the same reward both as proposer and responder.

(C) **Random Player**: The player will have independent p and q.

At each cycle, each player will play the game with each of its neighbors, and the score will be updated (the score being the sum of the money received). Two updating rules will be considered:

1. **Natural Selection**: In this framework each player i in the network selects at random one neighbor j and compares its payoff Π_i with that of j, Π_j. If $\Pi_j > \Pi_i$, player i adopts the strategy of $j, (p_j, q_j)$, for the next round of the UG with a probability proportional to the payoff difference:

$$P_{ij} = \frac{\Pi_j - \Pi_i}{2 \max \{k_i, k_j\}}$$

where k_i and k_j are the degrees of i and j respectively. However, if $\Pi_i \leq \Pi_j, i$ keeps its strategy for the following round.

2. **Social Penalty**: The player with the lowest payoff in the whole population and its neighbors, no matter how wealthy they are, are removed. These agents are replaced in their nodes by new players with random strategies (so that they only inherit their contacts).

2.3 Simulations

Three types of networks were considered: Erdos-Renyi, Barabasi-Albert, and Scale-Free. The networks were generated with $10k$ nodes and an average degree of 4. On all three networks, each type of player was simulated with each type of updating rule. The distribution of p and q among the players was recorded at predetermined intervals. The following figures show the distribution of p for the different types of players and networks.

All combinations were averaged over ten different simulation runs for $2 \cdot 10^4$ cycles (Natural Selection) or 10^5 cycles (Social Penalty). The Social Penalty updating rules were simulated more because it was slower to reach a stable configuration.

The following figures show the distribution of p for the different types of players and networks.

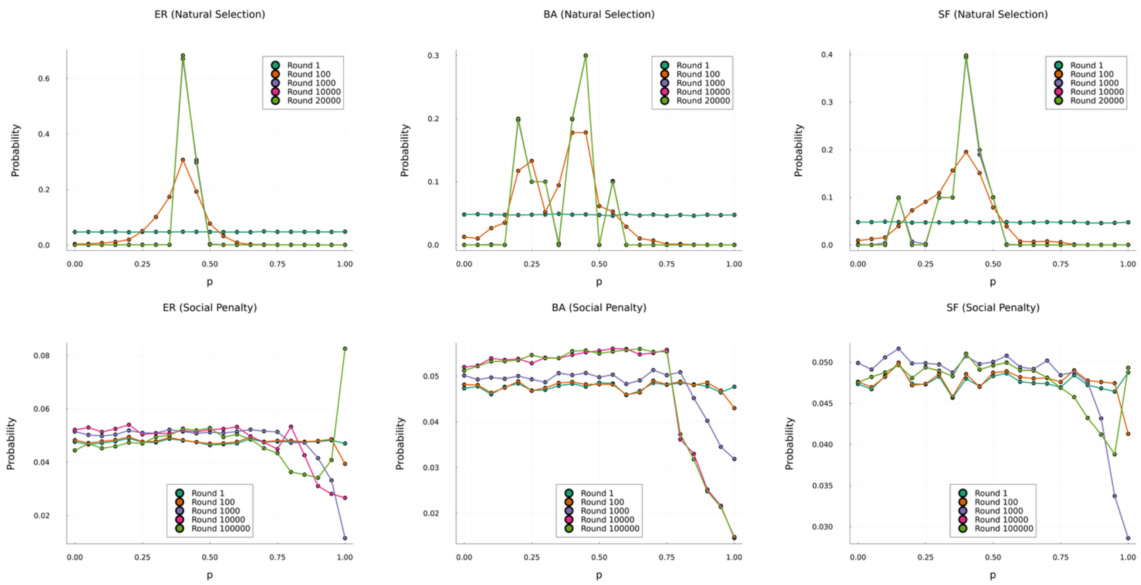


Figure 2.1: Networks with only type A players. The first row is with Natural Selection; the second row is with Social Penalty.

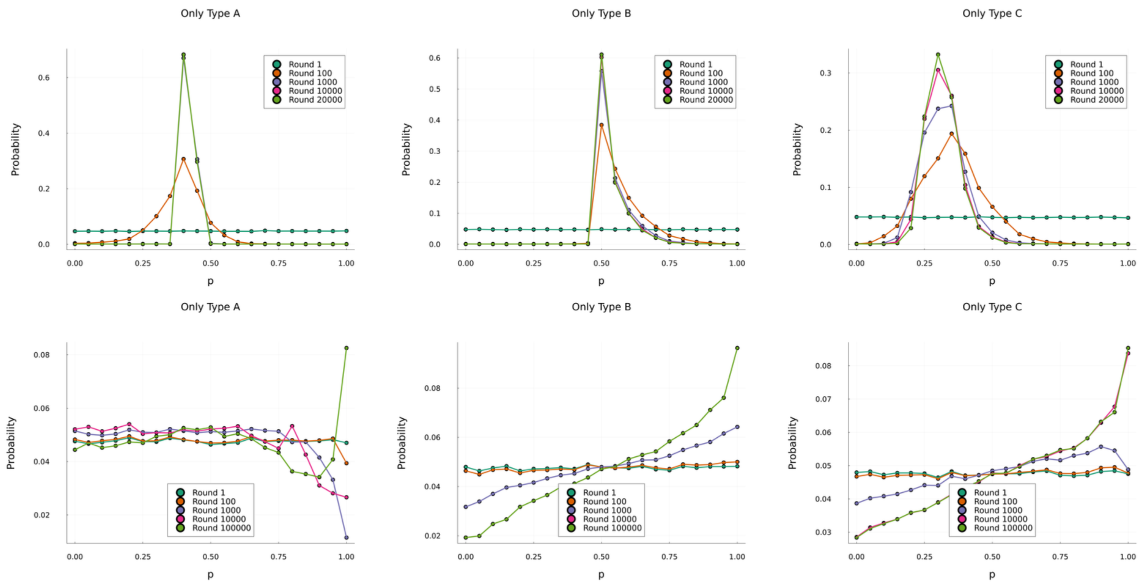


Figure 2.2: Networks with all types of players on Erdos-Renyi networks. The first row is with Natural Selection; the second row is with Social Penalty.

As stated in [12], the change in topology does not massively change the distribution of strategies, but the updating rule does. These results corroborate this conclusion. The individual distribution is different mainly because, in this task, a very low number of realizations was utilized.

3 | Task # 44 - Social Connectedness Index from Facebook

Task leader(s): *Jacopo Carotenuto*

3.1 | Introduction

In this task, the goal was to extract a network for each country present in the Facebook Scaled Social Connectedness Index data [1]. The data was collected by Facebook, and it is based on the number of friendships between people in different countries. From the documentation:

> The Social Connectedness Index uses an anonymized snapshot of active Facebook users and their friendship networks to measure the intensity of social connectedness between locations. Users are assigned to locations based on their information and activity on Facebook, including the stated city on their Facebook profile and device and connection information.

3.2 | Networks Extraction

3.2.1 Raw Data

The goal was to build two files:

1. `node_list.csv`: a file containing the node ID, longitude, latitude, and label for each node.
2. `edge_list.csv`: a file containing the source node ID, target node ID, weight (SCI), country, and country ISO code for each edge.

As the goal was to build a network for each country, edges between different countries were not considered. The data provided was in a ".tsv" file with the following columns: "`user_loc`," "`fr_loc`" and "`scaled_sci`" representing the User Location, Friend Location and Scaled Social Connectedness Index respectively. As written in the documentation, the SSCI was built to be between 1 and 10^9. All data manipulation and analysis were done in Julia[7].

3.2.2 Edge List

The first step was to extract all the unique locations present in the data and assign a unique ID to each of them. As the location was divided into different types of denominations (GADM[2], NUTS[3] and US Counties), the file "`gadm1_nuts3_counties_levels.csv`" was used to determine the type of denomination for each location. Then, each location was assigned the ISO3 code of the country it belonged to using the "Countries" package in Julia[5]. To each entry of the original file (so to each edge), the User ISO3 code and the Friend ISO3 code were added, and all the edges between different countries were removed. The final edge list was saved in a ".csv" file, saving only the source node ID, target node ID, weight, country, and country ISO code. For completeness, the edge list presents all the (symmetric) edges present in the source file.

3.2.3 Node List

The node list was more of a challenge as the longitude and latitude of each location were not provided. To assign to each location its coordinates, the geospatial data for each subdivision was found: for the GADM subdivisions, the data was found in the GADM website[2] (Version 2.8, world data), for the NUTS subdivisions, the data was found in the Eurostat website[3] (Year 2016, EPSG 4326) and for the US counties, the data was found in the Public Data website[4] (Year 2018).

The provided data was in the form of a ".shp" or ".geojson" file, so for the coordinates extraction, the "GeoStats.jl" package was used[6]. For all the administrative subdivisions, a polygon was provided, so the centroid of each polygon was used as the coordinates for the location. As the provided Facebook data contained different levels of administrative subdivisions, each level was analyzed separately. Then, as the Facebook area code used was different from all the other denominations, the area codes of each location were transformed into the Facebook format. Then, for each unique node, we associated the coordinates and the label (also found in the official data). The final node list was saved in a ".csv" file, saving the node ID, longitude, latitude, and label.

3.3 Networks Comparisons

The resulting countries' SCI-weighted undirected networks were all analyzed using Julia, and some important network properties were extracted. Some of them are contained in the following plots.

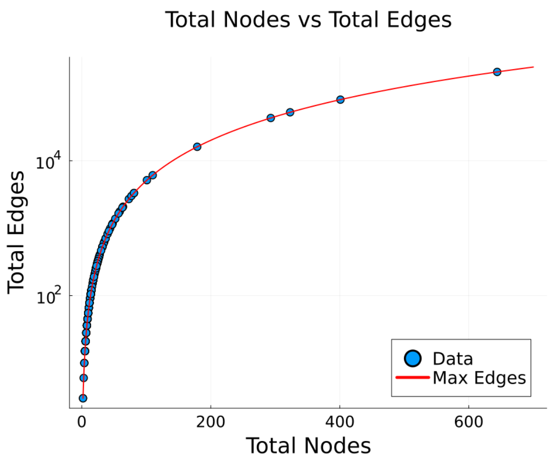


Figure 3.1:

From plot 3.1, we can see that the number of edges is proportional to the number of nodes (it's exactly the maximum number of undirected connections, allowing for self-loops), as expected, as all the networks are complete: every location has at least some people connected to every other location in the same country.

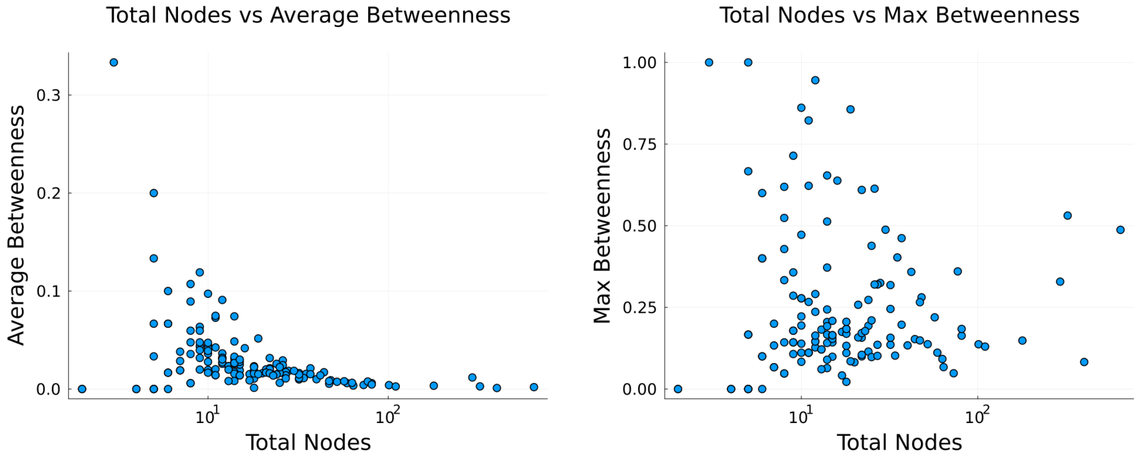


Figure 3.2:

From plot 3.2, we can see that the average betweenness is inversely proportional to the number of nodes in an exponential way, but the maximum betweenness is less correlated with the number of nodes.

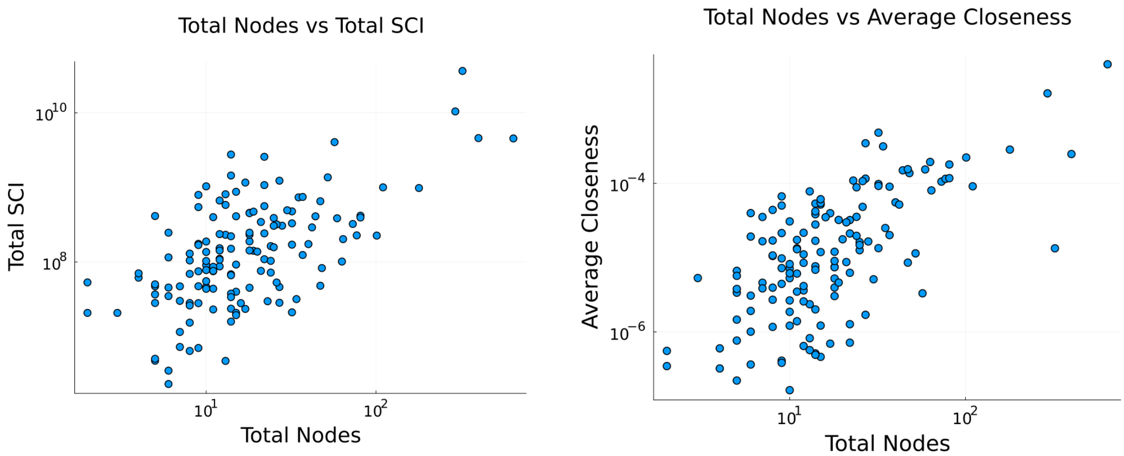


Figure 3.3:

From plot 3.3, it can be seen that the number of nodes is proportional to both the average closeness and the total SCI.

4 | Bibliography

[1] Facebook social connectedness index. https://data.humdata.org/dataset/social-connectedness-index. [Accessed 25 August 2024].

[2] Gadm. https://gadm.org/. [Accessed 25 August 2024].

[3] Nuts. https://ec.europa.eu/eurostat/web/gisco/geodata/statistical-units/territorial-units-statistics. [Accessed 25 August 2024].

[4] Us counties data. https://public.opendatasoft.com/explore/dataset/us-county-boundaries/export/?disjunctive.statefp&disjunctive.countyfp&disjunctive.name&disjunctive.namelsad&disjunctive.stusab&disjunctive.state_name. [Accessed 25 August 2024].

[5] Countries julia package. https://github.com/cjdoris/Countries.jl. [Accessed 25 August 2024].

[6] Geostats julia package. https://github.com/JuliaEarth/GeoStats.jl. [Accessed 25 August 2024].

[7] Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B Shah. Julia: A fresh approach to numerical computing. *SIAM review*, 59(1):65–98, 2017. URL https://doi.org/10.1137/141000671.

[8] Eric Bonabeau. Sandpile dynamics on random graphs. *Journal of the Physical Society of Japan*, 64(1):327–328, 1995. doi: 10.1143/JPSJ.64.327. URL https://doi.org/10.1143/JPSJ.64.327.

[9] K.-I. Goh, B. Kahng, and D. Kim. Universal behavior of load distribution in scale-free networks. *Phys. Rev. Lett.*, 87:278701, Dec 2001. doi: 10.1103/PhysRevLett.87.278701. URL https://link.aps.org/doi/10.1103/PhysRevLett.87.278701.

[10] K.-I. Goh, D.-S. Lee, B. Kahng, and D. Kim. Sandpile on scale-free networks. *Phys. Rev. Lett.*, 91:148701, Oct 2003. doi: 10.1103/PhysRevLett.91.148701. URL https://link.aps.org/doi/10.1103/PhysRevLett.91.148701.

[11] Bak-Tal-Wiesenfeld Sandpile Model. Self-organized criticality: An explanation of the 1/f noise. , 59(4):381–384, July 1987. doi: 10.1103/PhysRevLett.59.381.

[12] R Sinatra, J Iranzo, J Gómez-Gardeñes, L M Floría, V Latora, and Y Moreno. The ultimatum game in complex networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2009(09):P09012, sep 2009. doi: 10.1088/1742-5468/2009/09/P09012. URL https://dx.doi.org/10.1088/1742-5468/2009/09/P09012.