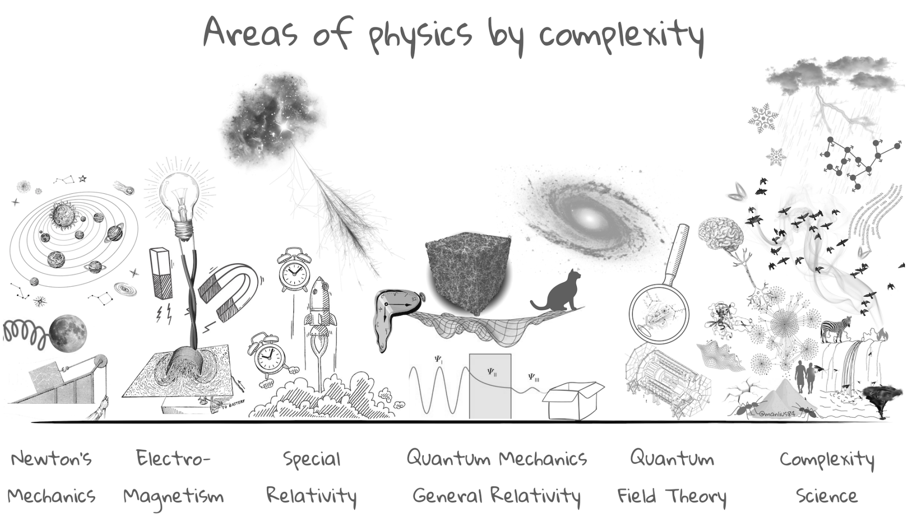
Final Report

Physics of Complex Networks: Structure and Dynamics

Università degli Studi di Padova



Projects Report: 15, 44

Tuscano Alessio

Last update: February 11, 2026

Contents

1 | Task 15: Self-Organized Criticality on Networks

1.1 | Framework (BTW sandpile on a graph)

We simulate the Bak–Tang–Wiesenfeld (BTW) sandpile on a network with N nodes. Each node i carries an integer load $z_i \in \{0, 1, 2, \dots\}$ and a threshold $z_c(i)$. At each driving step one grain is added to a uniformly random node; then the system relaxes via topplings. On a generic graph we take $z_c(i) = k_i$ (degree threshold), so one toppling at i sends one grain to each neighbor. If $z_i \geq z_c(i)$, a toppling performs

$$z_i \leftarrow z_i - k_i, \qquad z_j \leftarrow z_j + 1 \;\; \forall j \in \partial i. \tag{1.1}$$

To reach a stationary regime on finite networks we include dissipation with small probability f (per-toppling) so that avalanches remain finite. An *avalanche* is the full relaxation triggered by one grain addition. We measure standard observables [2, 6]: size S (total topplings), duration T (number of update waves), and area A (distinct toppled nodes).

1.2 | Results

From SOC on networks to SOC with interdependence. Power-law avalanche statistics on networks are a standard signature of self-organized criticality in sandpile models [2, 6]. Here we focus on a complementary question: how *interdependence* between two networked systems reshapes the probability of extreme events, following Brummitt et al. [3].

Coupled sandpiles and large cascades. We consider two modules A and B and add sparse interconnections controlled by a coupling parameter p. Here a *module* simply denotes one of the two subnetworks (A or B) in the coupled system [3]. In this section, N denotes the number of nodes *per module* (so the coupled system has $2N$ nodes). Because the BTW threshold is degree-based ($z_c(i) = k_i$), adding bridges modifies both (i) pathways for load to propagate between modules and (ii) node capacities/total load that the system can hold [3]. Following [3], we classify events by whether a *large* cascade occurs in module A, using a fixed cutoff on the avalanche size in that module: $S_A > C$ with $C = \frac{N}{2} = 1000$. We also track *global* cascades using $S > C_g$ with $C_g = N = 2000$, following [3] like before. Dissipation is implemented in per-toppling mode with probability $f = 0.01$. The "regular" modules are random regular graphs $R(z_a)$ and $R(z_b)$ (here $z_a = z_b = 3$), whereas the scale-free modules are generated by a

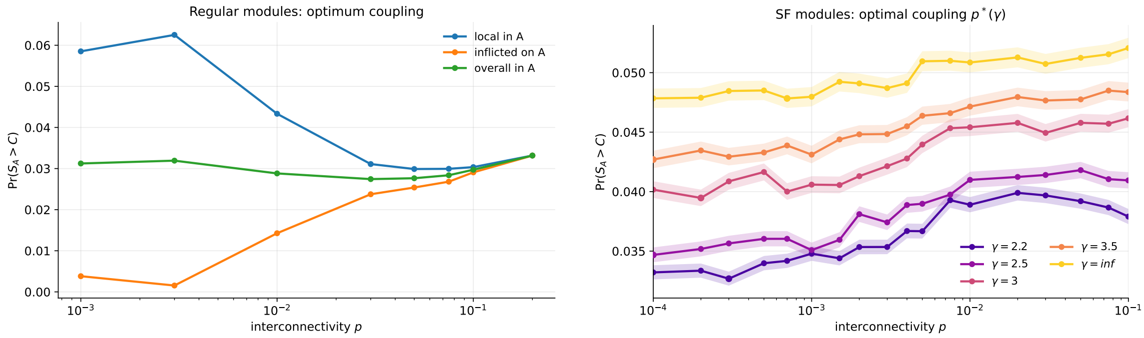


Figure 1.1: Coupled networks: probability of a large cascade in module A versus coupling p (regular vs scale-free modules). Scale-free modules show a stronger dependence on the generating exponent γ, while regular modules follow a smoother trend with p.

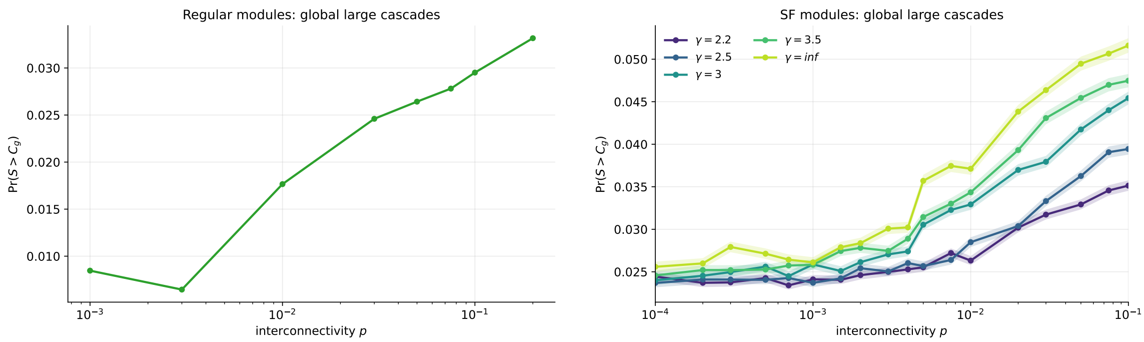


Figure 1.2: Robustness across event definitions: probability of a *global* large cascade ($S > C_g$, with $C_g = 2000$) versus coupling p. The qualitative non-monotonic dependence on p persists, supporting the interpretation of an "optimal coupling" rather than a cutoff artifact [3].

static model with tunable exponent γ (including $\gamma = \infty$, approaching a limit to where the degree distribution becomes uniform). In the scale-free panels, the shaded band shows a 95% Wilson score confidence interval (binomial) for the estimated probability from aggregated event counts. The three curves in Fig. 1.1 separate cascades that stay in A (local) from those triggered by activity arriving from the other module (inflicted), and their sum (any large in A). For small p, interconnections are beneficial because they suppress the largest within-module cascades; for large p, they become detrimental because they enable inflicted events and also increase capacities/total load. The trade-off can produce an intermediate optimum in p [3].

Global large cascades (defined by $S > C_g$) are rarer but more systemic events than "large-in-A" cascades, as seen in Fig. 1.2. The non-monotonic dependence on coupling p persists: small p can suppress the largest within-module events, whereas large p enables inter-module propagation (inflicted cascades) and increases total capacity/load, which can fuel larger system-wide events. The similar qualitative shape across regular and scale-free modules supports the interpretation of an "optimal coupling" rather than a cutoff-specific artifact. For a detailed branching-process description and analytic estimates of the optimal coupling p^*, we refer to [3].

2 | Task 44: Social Connectedness Index II

2.1 | Dataset and goal

We use the Social Connectedness Index II (SCI) by Meta/HDX, which reports the intensity of Facebook friendship ties between pairs of administrative regions. Each record provides a pair of regions and a positive score (`scaled_sci`). This release covers 178 countries and corresponds to the reference period 2025-12-26 to 2026-01-25 (updated 2026-02-07, CC0) [7]. Following the task instructions, we exclude the USA from our outputs.

Concept of SCI. SCI is a normalized measure of friendship intensity between locations. Conceptually, for regions i, j one can think of

$$\mathrm{SCI}_{ij} \propto \frac{F_{ij}}{U_i\, U_j},$$

where F_{ij} is the number of friendships connecting the two regions and U_i, U_j are the corresponding Facebook-user populations; published values are then rescaled within each layer to $[1, 10^9]$ (`scaled_sci`) [7, 1]. In our analysis we treat `scaled_sci` as a *within-country* edge weight.

2.2 | Network construction

End-to-end pipeline. From the raw SCI layer CSVs, we:

- select the set of countries to include;
- set the resolution per country: EU countries use NUTS3 (higher resolution), while all other countries use GADM level 1;
- keep within-country rows only (`user_country = friend_country`), drop self-loops, store one edge per unordered pair, and sum duplicate `scaled_sci`;
- optionally add coordinates by joining a separate region-centroids table (built from boundary polygons), then export the submission CSVs and run sanity checks.

Nodes. Nodes represent subnational administrative regions at the best available resolution. We export a single `nodes.csv` with global consecutive `nodeID` and `nodeLabel`; labels are prefixed with ISO3 (e.g., `ITA:...`) to avoid collisions.

Edges. Edges are built *within country* only. We treat ties as undirected, keep one edge per unordered region pair, and aggregate duplicates by summing `scaled_sci`; self-loops are removed. The submission file `edges.csv` stores endpoints and country tags; `edges_weighted.csv` is produced for analysis/debug.

Geographic coordinates. The SCI layer CSVs do not include coordinates. We therefore build a region-to-coordinate table by downloading administrative boundary datasets and extracting one representative point per polygon (EPSG:4326), using:

- **GADM v4.1** level-1 polygons (downloaded as shapefiles; region code `GID_1`) [5].
- **EU NUTS 2024** level-3 polygons (GISCO GeoJSON; region code `NUTS_ID`) [4].

The resulting `(latitude, longitude)` are merged into `nodes.csv` when available; boundary files are large and are downloaded on demand and cached locally.

2.3 Sanity checks

For the selected top-100 countries (by available GADM1 regions, USA excluded), the global output contains $N = 3040$ nodes and $E = 134016$ edges. Coordinate coverage is high after including centroids ($\sim 98\%$ of nodes). Within-country graphs are often very dense (close to complete), so unweighted structure is less informative than weighted summaries. To mitigate size effects we inspect mean edge weight (total `scaled_sci` divided by E).

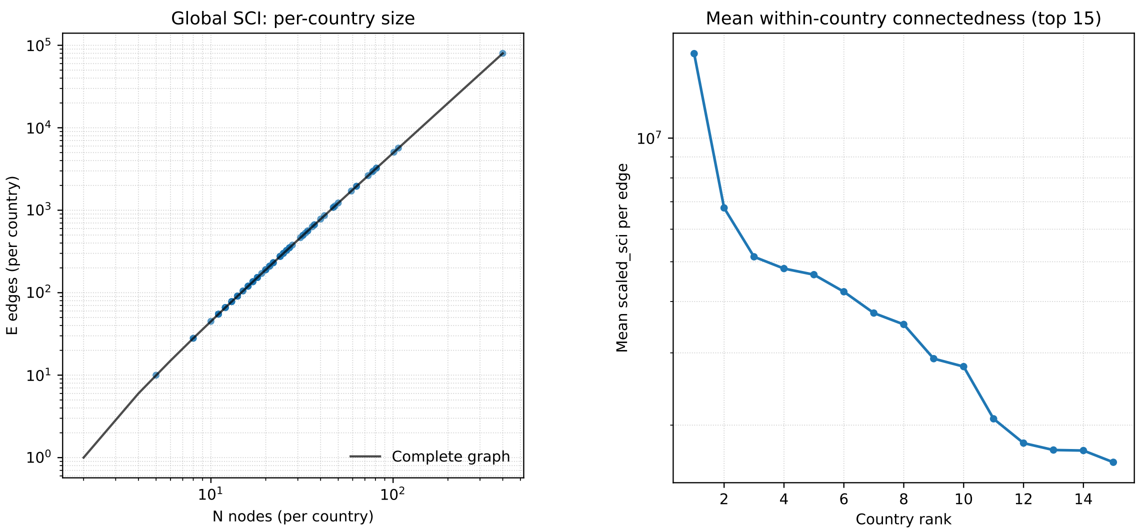


Figure 2.1: Left: number of edges versus number of nodes for each country network (log–log scale). The black curve is the complete-graph reference $E = N(N-1)/2$. Right: mean within-country connectedness (total `scaled_sci` divided by E) versus country rank (top 10).

3 | Bibliography

[1] AI and Data for Good at Meta. Social connectedness index. https://ai.meta.com/ai-for-good/datasets/social-connectedness-index/, 2026. [Accessed 04-Feb-2026]. Reference period: 2025-12-26 to 2026-01-25. Updated: 2026-01-29.

[2] Eric Bonabeau. Sandpile dynamics on random graphs. *Journal of the Physical Society of Japan*, 64(9):327–336, 1995. doi: 10.1143/JPSJ.64.327.

[3] Charles D. Brummitt, Raissa M. D'Souza, and E. A. Leicht. Suppressing cascades of load in interdependent networks. *Proceedings of the National Academy of Sciences*, 109(12):E680–E689, 2012. doi: 10.1073/pnas.1110586109.

[4] European Commission, Eurostat (GISCO). Nuts regions (2024) - gisco distribution (level 3, epsg:4326). https://gisco-services.ec.europa.eu/distribution/, 2024. [Accessed 04-Feb-2026]. GeoJSON endpoint used: NUTS_RG_01M_2024_4326_LEVL_3.

[5] GADM. Gadm database of global administrative areas (version 4.1). https://gadm.org/, 2025. [Accessed 04-Feb-2026]. Shapefile downloads mirrored at https://geodata.ucdavis.edu/gadm/.

[6] K.-I. Goh, D.-S. Lee, B. Kahng, and D. Kim. Sandpile on scale-free networks. *Physical Review Letters*, 91:148701, 2003. doi: 10.1103/PhysRevLett.91.148701.

[7] Humanitarian Data Exchange (HDX). Facebook social connectedness index. https://data.humdata.org/dataset/social-connectedness-index, 2026. [Accessed 04-Feb-2026]. Coverage: 178 countries. Reference period: 2025-12-26 to 2026-01-25. Updated: 2026-01-29. License: CC0.