Network Optimization

Learning Through Examples

Zoltán Toroczkai

Center for Nonlinear Studies, Theoretical Division, Los Alamos National Laboratory, New Mexico, USA

Department of Physics, University of Notre Dame, IN, USA

(after June 1st, 2006)
What are networks?

Collection of discrete entities [nodes], which might be connected via links [edges] representing interactions or associations between the connected elements.

Mathematical term for these objects: **Graph**

Typical notation: $G(V, E)$, where $V=$\{1,2,...,N\} is the set of nodes (vertices, sites) and $E$ is the set of edges.

An edge typically connects a pair of vertices $x$ and $y$, however it can also connect more than two vertices, called hyperedges and this case the resulting graph is called a **Hypergraph**. For now we exclusively deal with simple graphs, where $E \subseteq V \times V$.

Typical notations for an edge: $e = \{x, y\} \equiv (x, y) \equiv xy$
If there are several edges between two nodes, the graph is called a **multigraph**.

If the interaction or association is unidirectional, then this fact is resolved by making \( xy \neq yx \)

Such an edge \( \overrightarrow{e} = xy \) is called a *directed edge* and the corresponding graph a *directed graph*, or *digraph* for short.

Note: \( xy \in E \nRightarrow yx \in E \)

Both nodes and edges can have associated a number of properties, parameters, called *weights*.

Graphs and weights can be time dependent.

Typical real-world graphs are the result of complex processes with stochastic components makes sense to talk about *Graph Ensembles and* probabilistic descriptions.
Representations:

Visual, geometric:

Abstract:

- e.g. with the adjacency matrix: \( A = \{a_{ij}\}_{N \times N} \) where

\[
  a_{ij} = \begin{cases} 
  1 & \text{if } ij \in E \\
  0 & \text{if } ij \notin E 
  \end{cases}
\]

- “expensive” representation, requires \( O(N^2) \) resources

- it is hard to simply recover patterns/clusters from.

- sometimes advantageous for analytical calculations
How do we describe and study networks?

The party problem

What is the minimum nr. of people $R$, one should invite to a party that would surely have $k$ people who all know each other, or $k$ who do not know each other (at all)?

For $k=3$, \[ R(k) = 6 \]

For $k=4$:

\[ R(k) = 18 \text{ (hard proof)} \]
For $k=5$: $R(k)=…$ NOT KNOWN!

Only the bounds are known: $43 \leq R(5) \leq 49$.

**Come on, use a computer!**

We are looking for complete graphs with $n$ nodes that have a monochromatic complete subgraph of $k$ nodes ($k$-clique). (Here $k=5$.)

There are $\frac{n(n-1)}{2}$ edges in a complete graph. There are $2^{n(n-1)/2}$ such graphs whose edges are either blue or red.

Since for $k=3$, $R(3)=6$, an $n=6$ node complete graph would have a monochromatic triangle.

$n=6$: $2^{n(n-1)/2} = 2^{15} = 32,768$

$n=18$: $2^{n(n-1)/2} = 2^{153} \approx 1.46 \times 10^{46}$

$43 \leq n \leq 49$: $2^{903} - 2^{1176}$ graphs.
Operating at the physical limits of computation (as determined by the Planck constant, the speed of light and the gravitational constant) the 1kg laptop of Set Lloyd performs

\[ f = 5.4218 \times 10^{50} \text{ operations per second} \]

To check all graphs for monochromatic complete subgraphs takes at least

\[ \frac{2^{n(n-1)/2}}{f} \text{ seconds} = \frac{2^{n(n-1)/2-193.44}}{f} \text{ seconds} = 2^{n(n-1)/2-193.44} \text{ years} \]

Or, for \( k=5 \) it would take at least \( 2.693 \times 10^{213} \) years!

The age of the universe is estimated to be:

1.1-2 \times 10^{10} \text{ yrs!}

Probabilistic ensemble approach.
Optimization Problem Types

1. **Structural optimization**: find a graph \( G(V,E) \) which maximizes (minimizes) a topology functional \( R(G,\{\alpha\}) \).

2. **Dynamic optimization on static graphs**: for a given graph \( G(V,E) \) and a given Dynamical System on \( G \):

\[
\partial_t \Phi(x, \dot{x}, ..., \{\alpha\}, t) = 0
\]  

(1)

Find values of the parameters \( \{\alpha\} \) which maximize a global functional \( F \) on the trajectory of \( \Phi \).

3. **Structural optimization for dynamics**: given a set of parameters \( \{\alpha\} \) and the dynamical system (1), find a graph \( G(V,E) \) for which a global functional \( F \) on the trajectory of \( \Phi \) is maximized.

4. **Dynamics driven network optimization**: if \( G(V,E) = G(V,E,t) \) through coupling to flow (1) find the values of parameters \( \{\alpha\} \) for which a global functional \( F \) on the trajectory of \( \Phi \) AND over \( G(V,E,t) \) is maximized.
Structural Optimization

An algorithm for optimizing Randić indices on simple graphs

With: Péter L. Erdős (Alfréd Rényi Inst. Budapest)
      László Székely (University of South Carolina)
Structural optimization: find a graph $G(V,E)$ which maximizes (minimizes) a topology functional $R(G,\{\alpha\})$.

Maximizing the generalized Randić index:

*Problem*: Given a graphic sequence of integers $d_1 \geq d_2 \geq \ldots \geq d_n$, find a simple graph $G(V,E)$ realizing $\{d\}$ for its degree sequence and maximizing:

$$R(\alpha,G) = \sum_{(v_i,v_j) \in E} (d_id_j)^\alpha$$

The case $\alpha = -1/2$ is the classical Randić index.

Here we present the case $\alpha = 1$. 
Randić index, molecular branching index, connectivity index -- describes the degree of branching of the molecular skeleton.
It strongly correlates with physical properties such as "boiling points of hydrocarbons and the retention volumes and retention times obtained from chromatographic studies".


Figure 1. Boiling points of alkane isomers having from two to seven carbon atoms plotted against the topological branching index. (Experimental data are taken from ref 24.)
Figure 3. The experimental values of enthalpies of formation of alkane isomers having from two to seven carbon atoms plotted against the branching index. Each group of isomers has a separate correlation line approximately equally displaced and of the same slope.

Figure 4. Correlations between empirical parameters \( A, B, \) and \( C \) based on the Antoine equation for nine isomers of \( C_7H_{15} \) and the topological branching index. (Data taken from ref 24.)

Figure 5. A correlation between theoretically calculated total surface area of selected saturated acyclic hydrocarbons based on the model of Hermann and topological branching indices.

Figure 1. Representation of the external accessibility of an atom as a function of its valence (vertex degree).
The case $\alpha = 1$.

$$R(G) = \sum_{(v_i,v_j)\in E} d_id_j = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}d_id_j$$

It can be simply related to the number of paths of length 3, $N_3$:

$$N_3 = \sum_{(v_i,v_j)\in E} (d_i - 1)(d_j - 1) = R(G) - \sum_{i=1}^{n} (d_i)^2$$

From a social network application point of view: the optimal graph has maximum assortative mixing property (M.E.J. Newman).

The later statement is obvious from the observation:
Thus, maximizing \( R(G) \) is equivalent to minimizing \( F(G) - \text{DS index} \).

Representation using stubs:

\[
0 \leq F(G) = \sum_{(v_i,v_j) \in E} (d_i - d_j)^2 = \sum_{i=1}^{n} (d_i)^3 - 2R(G)
\]

\( d_1 = d_2 = \ldots = d_{k_1} = D_1, \quad d_{k_1+1} = \ldots = d_{k_1+k_2} = D_2, \quad \ldots \)

\( d_{n-k_l+1} = \ldots = d_n = D_l \).
Theorem 1 (Erdős, Gallai, 1960): Let $d_1 \geq d_2 \geq \ldots \geq d_n$ be a sequence of integers. Then they are the degree sequence (graphic sequence) of a simple graph if and only if:

1) $d_1 + d_2 + \ldots + d_n$ is even,

2) for all $1 \leq k \leq n-1$,

$$\sum_{i=1}^{k} d_i \leq 2 \binom{k}{2} + \sum_{i=k+1}^{n} \min\{k, d_i\}.$$ 

For a $D$-regular graph: $d_1 = \ldots = d_n = D$ condition 2) is:

$$n \geq D + 1$$

Theorem 2 (Largest R-index): if for all $1 \leq i \leq l$, $D_i k_i$ are even and $k_i \geq D_i + 1$, then the disjoint union of $D_i$-regular graphs on all vertices $i$ will maximize the R-index (and the DS-index will be identically zero, the smallest possible).

What if the conditions in Thm 2 are not fulfilled?
Observation 1:

Let \( \{v_j, v_k, v_l, v_m\} \in V \), such that \( j < k < l < m \)

1) If \((v_j, v_m) \in E\) and \((v_k, v_l) \in E\) and they can exchange for \((v_j, v_k) \in E\), \((v_l, v_m) \in E\), the DS-index decreases.

2) If \((v_j, v_m) \in E\) and \((v_k, v_l) \in E\) and they CANNOT exchange for \((v_j, v_k) \in E\), \((v_l, v_m) \in E\), but can exchange for \((v_j, v_l) \in E\), \((v_k, v_m) \in E\), the DS-index decreases.

3) If \((v_j, v_l) \in E\), \((v_k, v_m) \in E\) and can exchange for \((v_j, v_k) \in E\), \((v_l, v_m) \in E\), the DS-index decreases.
Lemma 1: Assume that in the setup from Observation 1, one of the intermediate nodes ($v_k$ or $v_l$) is not connected to $v_j$ and the exchanges in the previous observation that reduce the DS-index cannot be performed. Then the other intermediate ($v_l$ or $v_k$) node must be connected to $v_m$.

Non-reducible configurations
Lemma 2: Assume G is optimal and \((v_1, v_n) \in E\). Then:

(i) For every vertex \(v \in V\) with \(d(v) > d(v_n)\), \((v_1, v) \in E\)

(ii) \(D_1 \geq n - k_1\)

Proof. Assume on the contrary that there is \(v \neq v_1\) with \(d(v) > d(v_n)\), 
\((v_1, v) \not\in E\).

Since: \(d(v) > d(v_n) \geq 1 \rightarrow d(v) \geq 2 \rightarrow \exists w \in V \setminus \{v_1, v_n\}\) s.t. \((v, w) \in E\) Then, from Lemma 1: \((w, v_n) \in E\)

Thus, all neighbors of \(v\) are also neighbors of \(v_n\). That, however, contradicts \(d(v) > d(v_n)\).
“A mathematician is a machine turning coffee into theorems…” - Erdős quoting Rényi. Turán’s extension: “Weak coffee is only good for Lemma’s”
Now on to more “dynamical” problems…

Given a static graph, will assess global performance of the dynamics and then ask how to change the graph such as to improve that performance.
What is Parallel Computing?

Distributed programming, distributed architectures.

Parallel architecture: # of PE-s, L > 1.

Need for coordination among the PE-s (for interacting particle systems)

Agent-based systems: simulate practical problems in engineering, physics, social systems.
Examples for

Interacting “agent/particle” systems:

• Condensed matter: strongly correlated electron sys., magnetic sys.,
surface growth, etc.

• Chemical reactions (activated processes)

• Complex network dynamics (biological networks, www, internet,
communication)

• Traffic models, spatial epidemic models (social/political sys.)

• Models of economy/multi-player game theory (social/political sys.)

• Internet traffic routing/queing problems

• Computationally complex problems: optimization, scheduling,
hardest (NP) problems related to a phase transition (K-SAT), etc.
Some parallel devices

- 9472-node ASCI Red @ Sandia
- 12288-node QCDSP Teraflop Machine @ Brookhaven
- Connection Machine CM-2 with 65536 PE-s
- IBM Blue L, P.

BRAIN

- Weight: ~ approx. 1kg.
- Volume: ~ approx. 1400 cm$^3$
- #PE-s: ~ approx. $10^{11}$ neurons

Hedgehog

- 1 neuron connects to approx. 10000 other through synapses → $10^{15}$ synapses

If On-Off states: $2^{10^{11}}$
A fundamental question of II computing: Scalability

Utilization (or efficiency) of the algorithm:

\[ u(t, L) = \frac{\# \text{ of non-idling PE-s}}{\text{total } \# \text{ of PE-s}} \]

Scalability:

Scalable iff:

\[ \lim_{L \to \infty, t \to \infty} u(t, L) = u_\infty > 0 \]

Note: one also has to increase the size of the problem accordingly.
Massively Parallel Algorithms


Applicable to continuous time stochastic cellular automata with local dynamics and Poisson distributed events.

Example I: The Kinetic Ising Model, Glauber's Continuous-time Interpretation

- the spin flips are Poisson distributed.

$$\mathcal{H} = -J \sum_{<ij>} s_i s_j - H \sum_i s_i$$
Thus:

Each PE simulates the (physical) time of its next update attempt using a Poisson stream. Then the update will take place iff the simulated time is a *local minimum* among the simulated times of its neighbors.
• assures causality, conservative approach ("wait and update")
• there is no global synchronizing clock (would not be scalable if there was one)
• it is scalable (to be shown below)

Example II: Wireless cellular communications

- call arrivals obey Poisson statistics

1 cell / processor

Mapping to a surface model

Simulated time horizon = Non-equilibrium fluctuating surface

$PE_i \leftrightarrow \text{site}$

$i \leftrightarrow \text{# of layers}$

$\tau_i(t) \leftrightarrow h_i(t)$, the height of the interface

Dynamics:

Increments are Poisson arrivals

Deposit “material” of height $\Delta h$ with probability:

$$p(\Delta h) = \frac{1}{\alpha} e^{-\frac{1}{\alpha} \Delta h}$$

$\alpha > 0$
Where to deposit:  *only in local minima*!

(Parallel update)
This surface growth model is independent on the object of simulation, it corresponds to the massively parallel algorithm.

\[
\text{utilization} = \frac{\# \text{ of active sites}}{\text{total \# of sites (L)}} = \text{density of minima}
\]

Notation: \( M = M(t) \); \( M' = M(t+1) \).

\[
\tau'_i = \tau_i + \theta (\tau_{i-1} - \tau_i) \theta (\tau_{i+1} - \tau_i) \eta_i
\]

\( \eta_i \) is independent of \( t, i, \) and \( \{\tau_i\}_i \).
**Slope variables:**

\[ \varphi_i \equiv \tau_i - \tau_{i-1} \]

\[ \varphi'_i - \varphi_i = \theta(-\varphi_i)\theta(\varphi_{i+1})\eta_i - \theta(-\varphi_{i-1})\theta(\varphi_i)\eta_{i-1} \]

**PBC:**

\[ \sum_{i=1}^{L} \varphi_i = 0 \]

**Biased diffusion**

\[ \langle \varphi'_i \rangle - \langle \varphi_i \rangle = -[\langle j_i \rangle - \langle j_{i-1} \rangle] \]

**Continuity equation**

\[ \langle j_i \rangle = -\langle \theta(-\varphi_i)\theta(\varphi_{i+1}) \rangle \]

**Average current**

\[ \langle u \rangle = \frac{1}{L} \sum_{i=1}^{L} \langle \theta(-\varphi_i)\theta(\varphi_{i+1}) \rangle \]

**Utilization:**
Due to translational invariance:

\[
\text{mean velocity of the surface} = |\langle j \rangle| = \langle u \rangle
\]

Naïve coarse-graining:

\[
\theta(\varphi) = \lim_{\kappa \to 0} \theta(\varphi) = \lim_{\kappa \to 0} \frac{1}{2} \left[ 1 + \text{th} \left( \frac{\varphi}{\kappa} \right) \right]
\]

To leading order in \( \varphi / \kappa \):

\[
\langle \varphi'_i \rangle - \langle \varphi_i \rangle = \frac{1}{4\kappa} \langle \varphi_{i+1} - 2\varphi_i + \varphi_{i-1} \rangle - \frac{1}{4\kappa^2} \langle \varphi_i (\varphi_{i+1} - \varphi_{i-1}) \rangle + \ldots
\]

In the continuum limit:

Burger’s equation for the coarse-grained field \( \hat{\varphi} \):

\[
\frac{\partial \hat{\varphi}}{\partial t} = \frac{\partial^2 \hat{\varphi}}{\partial x^2} - \lambda \frac{\partial}{\partial x} (\hat{\varphi}^2)
\]
The fluctuations are captured by adding noise.

In 1-D, $t \to \infty$, the steady state is governed by the Edwards-Wilkinson Hamiltonian:

$$\mathcal{H}_{EW} = \int dx \left( \frac{\partial \hat{\tau}}{\partial x} \right)^2$$

A simple random-walk surface. The slopes are independent!
Utilization is not a universal quantity, it depends on the details of the model. However, it cannot vanish for our model: a zero density of local minima as $L \to \infty$ would imply that it is zero on all length-scales, which would contradict the fact that the coarse-grained level is described by the EW Hamiltonian.

The massively parallel algorithm is scalable.
Numerical support

\[ < w^2 (L, t) >= \frac{1}{L} \langle \sum_{i=1}^{L} [\tau_i(t) - \bar{\tau}(t)]^2 \rangle \]

\[ < w^2 (L, t) > \sim t^{2\beta} \]

\[ \beta \approx 0.326 \pm 0.005 \]

KPZ exact is: \( \beta = 1/3 \)

\[ < w^2 (L, \infty) > \sim L^{2\alpha} \]

\[ \alpha \approx 0.49 \pm 0.01 \] (roughness exponent)

KPZ exact is: \( \alpha = 1/2 \)

**Dynamic scaling:** \( z = \alpha/\beta \)

\[ < w^2 (L, t) >= L^{2\alpha} f(t / L^z) \]
Finite system size effects on the efficiency

\[ \langle u \rangle_L - \langle u \rangle_\infty \propto L^{-1} \]

\[ \sigma_L \equiv \sqrt{\langle u^2 \rangle - \langle u \rangle^2} \propto L^{-1/2} \]

(from 'roof-top')

\[ u \text{ is self-averaging } \]

\[ \tilde{u} \equiv \frac{u - \langle u \rangle_L}{\sigma_L} \]

\[ \tilde{P}_L(\tilde{u}) = \frac{1}{\sqrt{2\pi}} e^{-\tilde{u}^2/2} \]

( the normal density )

\[ \langle u \rangle_\infty \approx 0.2464... \]
Another check for the EW universality class

\[ \Phi(x) = \langle w^2 \rangle P(w^2) \]

\[ \Phi(x) = \frac{\pi^2}{3} \sum_{n=1}^{\infty} (-1)^{n-1} n e^{-\frac{\pi^2}{6} n^2 x} \]

Microscopic features in the steady state

$$P^{st}(\varphi_{i+1}, \varphi_i)$$

L=10^4

L=3

Up-down symmetry is broken
Higher-$d$ simulations (one site per PE)

$d=1$
\[ N_{PE} = L^d \]
\[ \langle u \rangle_\infty \approx 0.246... \]

$d=2$
\[ \langle u \rangle_\infty \approx 0.12... \]

$d=3$
\[ \langle u \rangle_\infty \approx 0.075... \]
Conclusions, so far

- The fundamental problem of scalability is solved via an exact mapping to a non-equilibrium growing surface.
- The evolution of this surface growth model (MPEU) is described by the KPZ equation.
- The scalability result is independent on the particular problem the massively parallel algorithm is simulating, it refers to the algorithm.
- Conservative schemes can be made scalable.
- The tools and machinery of non-equilibrium statistical physics (coarse-graining, finite-size scaling, universality, etc.) can be applied to scalability modeling and algorithmic engineering.
- Internet as a scalable supercomputer??
There is a problem, however!

Simulation reaches steady state for (arbitrary $d$) $t >> L^z$

- **Computational phase:** scalable $\left\langle u \right\rangle_L \approx \left\langle u \right\rangle_\infty + \frac{\text{const.}}{L^{2(1-\alpha)}}$

  $\left\langle u \right\rangle_\infty$ asymptotic average growth rate (simulation speed or utilization) is non-zero

- **Measurement (data management) phase:** not scalable

**Measurement at $\tau_{\text{meas}}$**

$$\left\langle w^2 \right\rangle_L \sim L^{2\alpha}$$

**Solution: create a SW topology !**

$r$ is a randomly chosen site with prob. $p$

$$\tau_i \leq \min\{\tau_{nn}\} \quad \Rightarrow \quad \tau_i \leq \min\{\tau_{nn}, \tau_r\}$$

Does not alter causality!
It is a synchronization problem.

\[ \tau_i \leq \min \{ \tau_{nn} \} \]

Kardar-Parisi-Zhang:

\[ \partial_t \tau = \frac{\partial^2 \tau}{\partial x^2} - \lambda \left( \frac{\partial \tau}{\partial x} \right)^2 + \text{noise} \]

\[ d_i = \ln M(r) / \ln r \quad r = D, \quad M(D) = N \quad D \equiv a \ln N \]

\[ w \sim L^\alpha \quad (\alpha = 0.5) \]

\[ w = \text{const.} + O(L^{-1}) \]

\[ \tau_i \leq \min \{ \tau_{nn}, \tau_r \} \]

\[ \partial_r \tau = -\gamma(\tau(x,t) - \bar{\tau}(t)) + \frac{\partial^2 \tau}{\partial x^2} + \text{noise} \]

\[ d_i \equiv \ln N / \left[ \ln(\alpha \ln N) \right] \to \infty \quad \text{as} \quad N \to \infty \]

Steady-state “height” structure factors

\[ S(k) \propto \langle \tau(k) \tau(-k) \rangle \]

only short-range connections (KPZ)

\[ \partial_t \tau = \frac{\partial^2 \tau}{\partial x^2} - \lambda \left( \frac{\partial \tau}{\partial x} \right)^2 + \text{noise} \]

\((d=1)\)

\[ S(k) \sim \frac{1}{k^2} \]

+ random connections (relaxation)

\[ \partial_t \tau = -\gamma (\tau(x,t) - \bar{\tau}(t)) + \frac{\partial^2 \tau}{\partial x^2} + \text{noise} \]

\[ S(k) \sim \frac{1}{k^2 + \gamma} \]

Effective mass
Edwards-Wilkinson Process on a network

G. Korniss et al., cond-mat/0508056

Consider:

\[ \partial_t h_i = - \sum_{j=1}^{N} A_{ij} (h_i - h_j) + \eta_i(t) \]  

where \( h_i(t) \) is a scalar at a node (stochastic field variable such as virtual time)

\( \eta_i(t) \) is delta-correlated white noise with zero mean and variance

\[ \langle \eta_i(t) \eta_j(t') \rangle = 2 \delta_{ij} \delta(t - t') \]  

\( A_{ij} = A_{ji} \) is the effective coupling between nodes \( i \) and \( j \). \( A_{ii} = 0 \)

Defining the Network Laplacian:

\[ \Gamma_{ij} = \delta_{ij} \sum_{l=1}^{N} A_{il} - A_{ij} \]  

(1) becomes:

\[ \partial_t h_i = - \sum_{j=1}^{N} \Gamma_{ij} h_j + \eta_i(t) \]
The steady-state 2-point equal time correlation function is given by:

\[ G_{ij} = \langle (h_i - \overline{h})(h_j - \overline{h}) \rangle = \hat{\Gamma}_{ij}^{-1} = \sum_{k=1}^{N-1} \frac{1}{\lambda_k} \psi_{ki} \psi_{kj} \] (5)

where \( \overline{h} = \frac{1}{N} \sum_{i=1}^{N} h_i \) and \( \langle \cdot \cdot \cdot \rangle \) denotes averaging over noise.

\( \hat{\Gamma}^{-1} \) is the inverse of \( \Gamma \) in the space orthogonal to the zero mode.

\[ \{ \lambda_k, \{ \psi_{ki} \}_{i=1}^{N} \}, \quad k = 0, N - 1 \] are the \( k^{th} \) eigenvalues and normalized eigenvectors.

\( k = 0 \) represents the zero mode of the network where \( \lambda_0 = 0 \).

Thus

\[ \langle w^2 \rangle = \left\langle \frac{1}{N} \sum_{i=1}^{N} (h_i - \overline{h})^2 \right\rangle = \frac{1}{N} \sum_{i=1}^{N} G_{ii} = \frac{1}{N} \sum_{k=1}^{N-1} \frac{1}{\lambda_k} \] (6)

For large systems and quenched network disorder, typically we have self-averaging:

\[ \langle w^2 \rangle \sim \left[ \langle w^2 \rangle \right] \quad \text{calculate} \quad [G_{ii}] \quad \text{get} \quad N \rightarrow \infty \quad \text{limit.} \]
Resistor Networks

Stationary state, \( V, I \). Kirchoff’s and Ohm’s laws.

\[
\sum_n A_{mn}(V_m - V_n) = I_m
\]  \( (7) \)

Where \( A_{mn} \) is the conductance of the edge \( mn \) and \( I_m \) is the net source current into node \( m \).

\[
\sum_n \Gamma_{mn} V_n = I(\delta_{mi} - \delta_{mj})
\]  \( (8) \)

\( \Gamma_{mn} \) is the same network Laplacian as for EW!

If: \( \hat{V} = \frac{1}{N} \sum_{m=1}^{N} V_m \) and \( \hat{V}_m \equiv V_m - \hat{V} \) then

\[
\hat{V}_m = \sum_n \hat{\Gamma}^{-1}_{mn} I_n = \sum_n \hat{\Gamma}^{-1}_{mn} I(\delta_{ni} - \delta_{nj}) = I(G_{mi} - G_{mj})
\]  \( (9) \)

Where \( G \) is the same network propagator as in EW!
For the power source attachment points \( i \) and \( j \):

\[
V = V_i - V_j = \hat{V}_i - \hat{V}_j = I(G_{ii} + G_{jj} - 2G_{ij}) \tag{10}
\]

The total resistance between nodes \( i \) and \( j \):

\[
R_{ij} = \frac{V}{I} = G_{ii} + G_{jj} - 2G_{ij} = \sum_{k=1}^{N-1} \frac{1}{\lambda_k} \left( \psi_{ki}^2 + \psi_{kj}^2 - 2\psi_{ki}\psi_{kj} \right) \tag{11}
\]

If we calculate the steady-state \textit{height difference} correlation function for EW:

\[
\left\langle (h_i - h_j)^2 \right\rangle = G_{ii} + G_{jj} - 2G_{ij} = R_{ij} \tag{12}
\]

Namely it is the same as the network resistance between nodes \( i \) and \( j \)!

The height difference correlation function is a standard observable in surface growth:


Thus many results for resistor networks will fall out from the study of EW fluctuations on the same network. This connects \textit{fluctuations and flow/transport} properties.

The average resistance

\[
\overline{R} = \frac{1}{N(N-1)} \sum_{i \neq j} R_{ij} = 2 \left\langle w^2 \right\rangle \tag{13}
\]

The particular case of SW networks

Let $p$ be the shortcut link density, i.e., the link probability is $p/N$, and a node has in average $p$ shortcuts. The network is nothing but an ER random graph superimposed on a ring. We have:

$$A_{ij} = \delta_{i,j-1} + \delta_{i,j+1} + J_{ij}$$  \hspace{1cm} (14)

where $J_{ij}$ are quenched random variables $J_{ij} = 1(0)$ with probability $p/N (1 - p/N)$ (all conductances are unity on existing edges)

Network disorder average restores translational invariance:

$$[G_{ij}] = [G(|i - j|)]$$

These functions can be calculated using disorder-averaged self-consistent perturbation theory


$$[G(l)] \sim \frac{1}{2\sqrt{\Sigma}} e^{-\sqrt{\Sigma}l}$$  \hspace{1cm} (15)

where $\Sigma \sim p^2$ is an effective mass. Using (11):
\[ [R(l)] = 2 ([G(0)] - [G(l)]) \simeq \frac{1}{\sqrt{\Sigma}} \left( 1 - e^{-\sqrt{\Sigma}l} \right) \]

thus \[ \lim_{l \to \infty} [R(l)] = \Sigma^{-1/2} \sim p^{-1} \]

In contrast with the regular geometry (ring) where \( R(l) \sim l \to \infty \)

The average resistance is finite for \( 0 < p \ll 1 \)

\[ \bar{R} \simeq [R] = 2 \langle w^2 \rangle = 2 [G(0)] \simeq \frac{1}{\sqrt{\Sigma}} \sim p^{-1} \]

In contrast with the case for the ring: \( \bar{R} \simeq N/6 \)

Proving also the finiteness of the width for the synchronization problem when on SW.

Numerical support:
Fig. 1. Disorder-averaged (a) two-point function and (b) two-point resistance as a function of the separation \( l \) in simple SW networks for \( p = 0.10 \) and three system sizes. The solid line in (a) and (b) corresponds to the exponential decay and saturation given by Eqs. (14) and (15), respectively.

Fig. 2. (a) Average resistance vs the density of random links. The straight solid line indicate the asymptotic infinite system-size behavior [Eq. (16)]. (b) Average resistance vs the system size in simple SW networks. The straight solid line corresponds to the behavior of the one-dimensional regular network (ring), \([\bar{R}] \simeq N/6\).
**Gradient Networks**

*Gradients* of a scalar (temperature, concentration, potential, etc.) induce flows (heat, particles, currents, etc.).

Naturally, gradients will induce flows on *networks* as well.

Ex.: Load balancing in parallel computation and packet routing on the internet


References:


Setup:

Let $G=G(V,E)$ be an undirected graph, which we call the substrate network.

The vertex set: $V = \{x_0, x_1, \ldots, x_{N-1}\} \equiv \{0,1,2,\ldots, N-1\}$

The edge set: $E \subset V \times V$, $e \in E$, $e = x_i x_j = (i,j)$, $xx \notin E$ (no self-loops)

A simple representation of $E$ is via the $N \times N$ adjacency (or incidence) matrix $A$

$$A(x_i, x_j) = a_{ij} = \begin{cases} 1 & \text{if } (i,j) \in E \\ 0 & \text{if } (i,j) \notin E \end{cases} \quad (1)$$

Let us consider a scalar field $\{h\} : V \rightarrow \mathbb{R}$

Set of nearest neighbor nodes on $G$ of $i$ : $S_i^{(1)}$

$$h_m = \max \{h_i, h_a, h_b, h_c, h_d, h_m\}$$
**Definition 1**  The gradient $\nabla h(i)$ of the field $\{h\}$ in node $i$ is a directed edge:

$$\nabla h(i) = (i, \mu(i))$$

Which points from $i$ to that nearest neighbor $\mu \in S_i^{(1)} \cup \{i\}$ for $G$ for which the increase in the scalar is the largest, i.e.,:

$$\mu(i) = \arg \max_{j \in S_i^{(1)} \cup \{i\}} (h_j)$$

The weight associated with edge $(i, \mu)$ is given by:

$$|\nabla h(i)| = h_\mu - h_i$$

If $\mu(i) = i$ then $\nabla h(i) = (i, i) \equiv 0(i)$. The self-loop $0(i)$ is a loop through $i$ with zero weight.

**Definition 2**  The set $F$ of directed gradient edges on $G$ together with the vertex set $V$ forms the **gradient network**:

$$\nabla G = \nabla G(V, F)$$

If (3) admits more than one solution, than the gradient in $i$ is **degenerate**.
In the following we will only consider scalar fields with non-degenerate gradients. This means:

\[ \text{Prob.}\{h_i = h_j \text{ if } (i, j) \in E\} = 0 \]

**Theorem 1**  
Non-degenerate gradient networks form forests.

**Proof:**
Theorem 2  \[ \text{The number of trees in this forest} = \text{number of local maxima of } \{h\} \text{ on } G. \]
For Erdős - Rényi random graph substrates with i.i.d random numbers as scalars, the in-degree distribution is:

\[ R(l) = \frac{1}{N} \sum_{n=0}^{N-1} \binom{N-1-n}{l} [1 - p(1 - p)^n]^{N-1-n-l} [p(1 - p)^n]^l \]

\[ z = 100 \]

In the limit \( p \to 0, \ N \to \infty, \ z = Np = \text{const.}, \ z \gg 1 \):

\[ R_N(l) \approx \frac{1}{zl}, \quad 1 \leq l < z = Np, \]

- Red: \( N=1000, \ p=0.1, \ 10^4 \) averages, numerics
- Blue: Exact formula, \( N=1000, \ p=0.1 \)
In-degree distribution of the Gradient Network when $G = G_{N,p}$.

A combinatorial derivation

Version: Balazs Kozma (RPI)

Assume that the scalar values at the nodes are i.i.d, according to some distribution $\eta(h)$.

First, distribute the scalars on the node set $V$, then find those link configurations which contribute to $R(l)$ when building the $G_{N,p}$ graph.

Without restricting the generality, calculate $R(l)$ for node 0.

Consider the set of nodes with the property $h_j > h_0$

Let the number of elements in this set be $n$, and the set be denoted by $\tau[n]$.

The complementary set of $\tau[n]$ in $V\setminus\{0\}$ is: $C_{\tau[n]}$
In order to have exactly \( l \) nodes pointing their gradient edges into 0:

- they have to be connected to node 0 on the substrate
- they must NOT be connected to the set \( \tau[n] \)

For \( l \) nodes:

\[
\left[ p(1 - p)^n \right]^l
\]

Also need to require that no other nodes will be pointing their gradient directions into node 0:

(Obviously none of the \( \tau[n] \) will.)

\[
\left[ 1 - p(1 - p)^n \right]^{N-l-n}
\]

So, for a fixed \( h_0 \) and a specific set \( \tau[n] \):

\[
\binom{N-1-n}{l} \left[ p(1 - p)^n \right]^l \left[ 1 - p(1 - p)^n \right]^{N-l-n}
\]
The probability $Q_n$ for such an event for a given $n$ while letting $h$-s vary according to their distribution:

For one node to have its scalar larger than $h_0$: \[ \gamma(h_0) = \int_{h_0} dh \eta(h) \]

For exactly $n$ nodes:
\[ \left[ \gamma(h_0) \right]^n \left[ 1 - \gamma(h_0) \right]^{N-1-n} \]

Thus:
\[ Q_n = {N-1 \choose n} \int dh_0 \eta(h_0) \left[ \gamma(h_0) \right]^n \left[ 1 - \gamma(h_0) \right]^{N-1-n} = \frac{1}{N} \]

Combining:
\[ R_N(l) = \sum_{n=0}^{N-1} Q_n \binom{N-1-n}{l} \left[ p(1-p)^n \right]^l \left[ 1 - p(1-p)^n \right]^{N-1-l-n} \]

Finally:
\[ R_N(l) = \frac{1}{N} \sum_{n=0}^{N-1} \binom{N-1-n}{l} \left[ 1 - p(1-p)^n \right]^{N-1-n-l} \left[ p(1-p)^n \right]^l \]

Independent of $\eta$
What happens when the substrate is a scale-free network?
Gradient Networks and Transport Efficiency

- every node has exactly one out-link (one gradient direction) but it can have more than one in-link

- the gradient network has N-nodes and N out-links. So the number of “out-streams” is \( N_{\text{send}} = N \)

- the number of RECEIVERS is \( N_{\text{receive}} = \sum_{l \geq 1} N_{l}^{(in)} \)

\[
J = 1 - \left< \left< \frac{N_{\text{receive}}}{N_{\text{send}}} \right> \right>_G = 1 - \left< \left< \sum_{l \geq 1} \frac{N_{l}^{(in)}}{N} \right> \right>_G = \left< \left< \frac{N_{0}^{(in)}}{N} \right> \right>_G = R_N(0)
\]

- \( J \) is a congestion (pressure) characteristic.

- \( 0 \leq J \leq 1 \). \( \underline{J = 0} \): minimum congestion, \( \underline{J = 1} \): maximum congestion

\[
J^{G_{N,p}}(N, p) = \frac{1}{N} \sum_{n=1}^{N-1} \left[ 1 - p(1-p)^n \right]^{N-1-n}
\]
In the scaling limit \( p = \text{const.} \), \( N \to \infty \),

\[
J_{G_{N,p}}^N(N, p) = 1 - \frac{\ln N}{N \ln \left( \frac{1}{1-p} \right)} \left[ 1 + O \left( \frac{1}{N} \right) \right] \to 1
\]

- for large networks we get maximal congestion!

In the scaling limit \( p \to 0, \ N \to \infty, \ pN = z \),

\[
J_{G_{N,p}}^N(N, p) \geq \int_0^1 dx \ e^{-ze^{-x}} = \frac{1}{z} \left[ Ei(-z) - Ei(-ze^{-z}) \right]
\]

\[
J_{G_{N,p}}^N(N, p) \geq 1 - \frac{\ln z + C}{z} + \ldots \quad \text{as } z \to \infty
\]

- becomes congested for large average degree.
- For scale-free structures, the congestion factor becomes independent on the system (network) size!!

For LARGE and growing networks, where the conductance of edges is the same, and the flow is generated by gradients, **scale-free networks are more likely to be selected during network evolution than scaled structures.**
The Configuration model

A. Clauset, C. Moore, Z.T., E. Lopez, to be published.
Generating functions: \[ g(z) = \sum_i k_i z^k \]

\[ R(z) = \int_0^1 dx \ g \left( 1 - (1 - z) \frac{x g'(x)}{g'(1)} \right) \]

**K-th Power of a Ring**
\[ R^{(2K)}(l) = \begin{cases} 
\frac{4(3 + 9K + 4K^2 + 2 Kl)}{(2K + l)(2K + l + 1)(2K + l + 2)(2K + l + 3)}, & 1 \leq l \leq K - 1 \\
\frac{6(2 + 7K + 7K^2)}{3K(3K + 1)(3K + 2)(3K + 3)}, & l = K \\
\frac{4(2K + 1)}{(2K + l + 1)(2K + l + 2)(2K + l + 3)}, & K + 1 \leq l \leq 2K - 1 \\
\frac{1}{(4K + 1)}, & l = 2K 
\end{cases} \]
Power law with exponent $=-3$
Competition Games on Networks

Collaboration with:

- Marian Anghel (CCS-3)
- Kevin E. Bassler (U. Houston)
- György Korniss (Rensselaer)

References:


Resource limitations lead in human, and most biological populations to competitive dynamics.

The more severe the limitations, the more fierce the competition.

Amid competitive conditions certain agents may have better venues or strategies to reach the resources, which puts them into a distinguished class of the “few”, or elites.

Elites form a minority group. In spite of the minority character, the elites can considerably shape the structure of the whole society: since they are the most successful (in the given situation), the rest of the agents will tend to follow (imitate, interact with) the elites creating a social structure of leadership in the agent society.

Definition: a leader is an agent that has at least one follower at that moment. The influence of a leader is measured by the number of followers it has. Leaders can be following other leaders or themselves.

The non-leaders are coined “followers”.
The El Farol bar game on a social network
Agents communicate among themselves.

**Social network:** 2 components:

1) Aquintance (substrate) network: $G$ (non-directed, less dynamic)
2) Action network: $\mathcal{A}$ (directed and dynamic)
Emergence of scale-free leadership structure:

- Robust leadership hierarchy
  \[ 0 < k_i^{out} < k_i \]
  \[ N_k (N, m; p) \propto k^{-\beta} N_1 (N, m; p) \]
  \[ N_1 (N, m; p) = a(p) \]
  \[ N_k (N, m; p) = a(p) k^{-\beta} f_k (N, m; p) \]
  \[ f_k (N, m; p) \to 1, \quad \text{for } m \gg 1 \]

- RCG on the ER network produces the scale-free backbone of the leadership structure

✓ The influence is evenly distributed among all levels of the leadership hierarchy.
Structural unevenness appears in the leadership structure for low trait diversity.

The followers ("sheep") make up most of the population (over 90%) and their number scales linearly with the total number of agents.
A networked, low trait diversity system is more effective as a collective than a sophisticated group!

Can we find/evolve networks/strategies that achieve almost perfect volatility given a group and their strategies (or the social network on the group)?
In the limit $p \to 0, \ N \to \infty, \ z = Np = \text{const.}, \ z >> 1$:

$$R_N(l) \approx \frac{1}{zl}, \ 1 \leq l < z = Np,$$