1

Complex Network DYNAMICS

Complex → somewhere in between regular and random

Network → A structure of units (nodes) and their interaction (links)

Dynamics → something changing over time, we'll get to that later

Motivation

- networks are everywhere!
- Game: Find a topic that is not a network and we'll try to prove you wrong!
- There is a lot of fun historical and pop-sci motivation on youtube and we have a bunch of links ... somewhere

---

- Networks sort of make the opposite simplification to almost everywhere else in physics:

Physics: Figure out how a thing behaves and how 2 things interact

Put them together and explain everything but Hypergraphs

Network science: Never mind what the things and their interactions are, can we already say something by the structure of interactions

Show my funny plot proving that Networks are still growing and are still cool

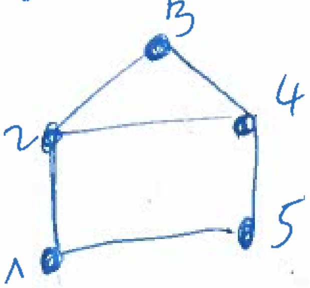
Reading Material

- Newmann Book or review paper
- Statistical Mechanics of complex networks Barabasi
- Spatial Networks Barthélemy

③

WALKS, PATHS & CONNECTIVITY

walk : sequence of edges



i.E.

1 - 2 - 4 - 3 - 2 - 1 - 2

closed walk → starts & ends on the same node

Path : walk without repeats of vertices or edges

5 - 1 - 2 - 4 - 3

cycle : closed path

4 - 3 - 2 - 4

Trail : vertices may repeat but edges not

4 - 2 - 3 - 4 - 5 - 1

Just so you know the lingo, mostly we'll stick to paths and leave the other terms to the Graph theorists.

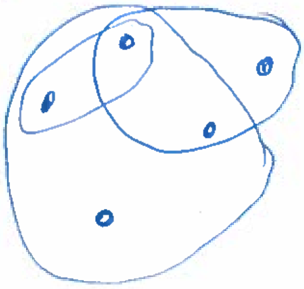
Connectivity

an undirected Graph G is said to be CONNECTED if every pair u, v of vertices is connected, that is if G contains a path from u to v

"disconnected" not connected

connected

Hypergraph: Graph where edges "hyper-edges" are tuples of several vertices

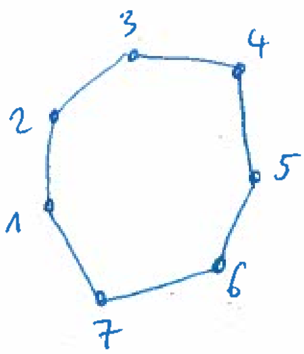
↳ 

haven't really found a use for this, but they get brought up occasionally

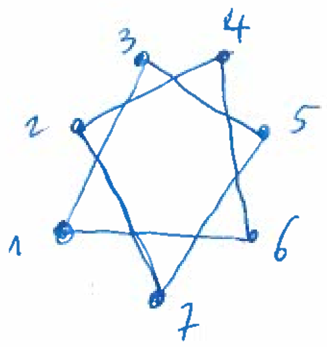
---

Network MEASURES

- Comparing several different networks can ~~be~~ quickly get out of hand as the network grows ⇒ we need aggregated quantities for the quantitative understanding



N



N'

First idea: count the differences

Hamming distance:

$$H(N,N') = \sum_{i,j} |A^{N}_{ij} - A^{N'}_{ij}| = 14$$

⟹ However, often we don't even care about the numbering of the nodes

Quantitatively understanding Networks

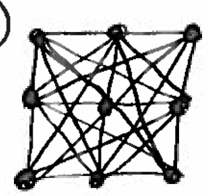
How can we describe and compare networks more simply?

What is the difference between?

a) $N = 9$

and

b) $N = 9$



$\Longrightarrow$

> \# links M
> or link density
> $P = \frac{M}{M_{max}} = \frac{M}{N \cdot (N-1)/2}$

$M = 0$

$P = 0$

$\langle k \rangle = 0$

$M = \frac{N \cdot (N-1)}{2} = 9 \cdot 4 = 36$

$P = 1$

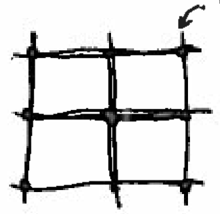
$\langle k \rangle = 8$

$\langle k \rangle = \frac{2M}{N}$

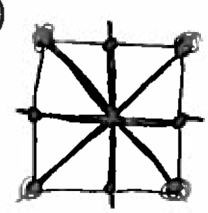
What is the difference between?

Periodische randbedingungen

a)



b)

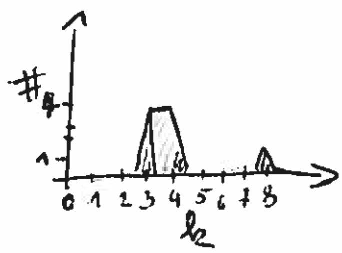


$2M = N \cdot \langle k \rangle$
$= 4 \cdot 9 = 36$

$M = 18$

$M = 18$

$\Longrightarrow$ Degree distribution



> these are
> Global network measures Describing the network as a hole in a Scalar or a Vector
> (the adjacency matrix can be viewed as a matrix global network measure)

Network Ensembles

Reasons why just one adjacency matrix doesn't tend to be of much use:

SYSTEMS IN Nature tend to be
- fluctuating / evolving
- completely or partially unknown
- probabilistic / random in some way
- exist in many realizations

⟹ often we're not actually interested in the one true Network but rather in a TYPE OF NETWORKS

Examples:
- Power Grid - like networks
- friendship - like networks
- protein - like networks

DEF : Network Ensemble

A set of possible networks together with a probability distribution of their occurrance

---

ERDŐS - Renyi - RANDOM Graphs

widely used as Null - Model, we will be computing a bunch of things about it analytically today.

a) a graph with N nodes and M edges is uniformly chosen from all such graphs

$G(N,M)$

b) each edge is realized with probability p

$G(N,p)$

Degree Distribution

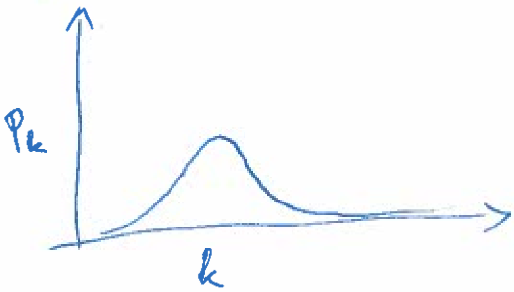
the probability for a node to be connected to a specific set of k other nodes is

$$p^k (1-p)^{n-1-k}$$

ways to select a set of k other nodes:

$$\binom{n-1}{k} \quad \Rightarrow \quad \boxed{P_k = \binom{n-k}{k} p^k (1-p)^{n-1-k}}$$

Binomial degree distribution



Clustering Coefficient

$C \hat{=}$ probability that two neighbours of a vertex are also neighbours of each other

$\Rightarrow$ $C = p$ all node pairs have this very same connection probability

---

other features

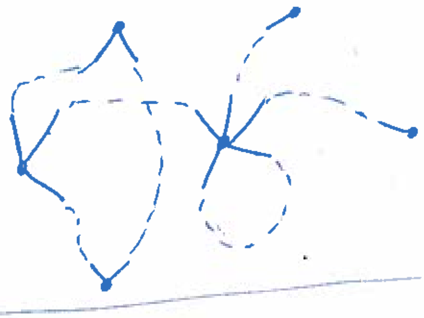
many other things can be analytically derived about ER-networks

- Size of the Giant component
  - ↳ percolation seminar topic
- Component size distribution
  - ↳ see Newman & some new works from 2019

The Configuration model

For some purposes you may not want everything about your network to be random a common thing to fix is the degree distribution or degree sequence.

- given a bunch of nodes with "stubs" or "half-edges", combine them randomly to obtain a realization of the configuration model
  (# stubs must be even)



$$\boxed{\text{Beware! this model has self-edges \& multi-edges}}$$

Link probability between nodes i and j

$$\boxed{P_{ij} = \frac{k_i k_j}{2m-1}}$$

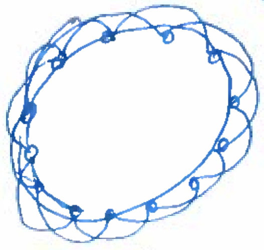
Say we start from node i. The probability that a specific one of the stubs of node i is connected to node j is $\frac{\#\text{ stubs of } j}{\#\text{ all other stubs}} = \frac{k_j}{2m-1}$

and we get k_i chances.

Random d-regular graph

configuration model where all nodes have degree d

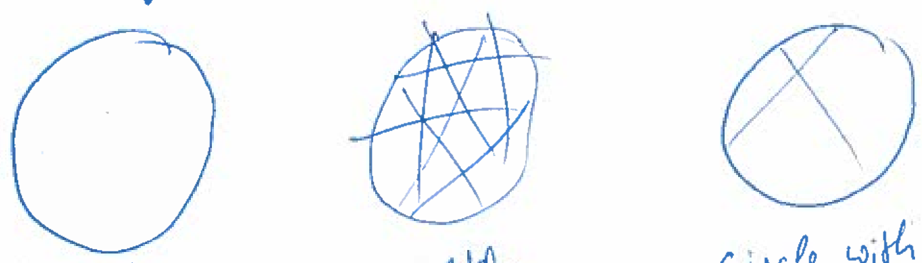
Rewiring or adding links



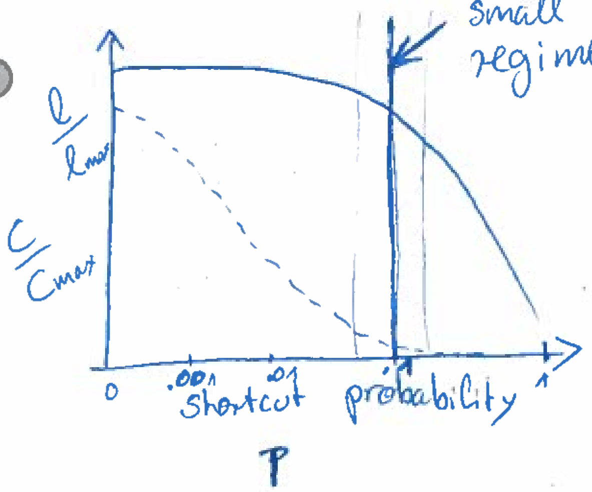
go through all edges, remove each with prob. p and add a random one instead — rewiring — preserves m

just add random edges with probability p for each existing edge. $m \cdot p$ shortcuts are made $\Rightarrow$ $\langle k \rangle p$ per node — shortcuts — easier but increases m

$P=0$ $\qquad$ $P=1$ $\qquad$ $0 < P < 1$



original circle $\qquad$ completely random $\qquad$ circle with some shortcuts



Degree distribution

at $P=0$ it's regular, all nodes have degree $\langle k \rangle$
as we add shortcuts it becomes $\langle k \rangle + \#\text{shortcuts}_i$
(Binomial distribution)

The ensemble Zoo

More on Watts-Strogatz Networks

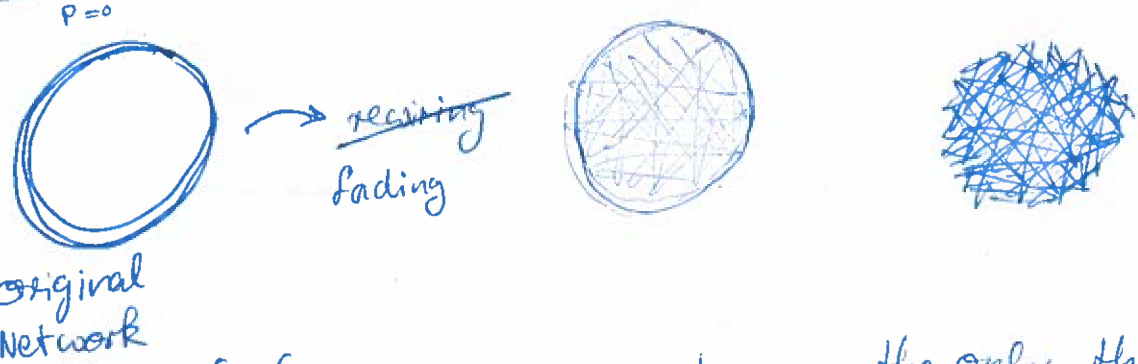
Graph Laplacian

$$\boxed{L = -D + A} \qquad -\begin{pmatrix} k_1 & & \\ & \ddots & \\ & & k_n \end{pmatrix} + \begin{pmatrix} 0 & 1 & & 0 \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ 0 & & 1 & 0 \end{pmatrix}$$

→ more on this in 2 to 4 weeks

- Diffusion, synchronization etc.
  ↳ predict Eigenvalue spectrum of WS networks

- first (and for ~~our~~ considerations main) step: constructing the mean field Laplacian



- $$L^{MF} = \begin{pmatrix} c_0 & c_1 & c_2 & \cdots & c_{N-1} \\ c_{N-1} & \ddots & & & \vdots \\ \vdots & & \ddots & & c_2 \\ & & & \ddots & c_1 \\ & & & & c_0 \\ c_1 & \cdots & & & c_{N-1} \end{pmatrix}$$

because the only thing that determines the entries of A^{MF} and thus L^{MF} is the distance of two nodes along the circle

this structure is called "Circulant Laplacian"

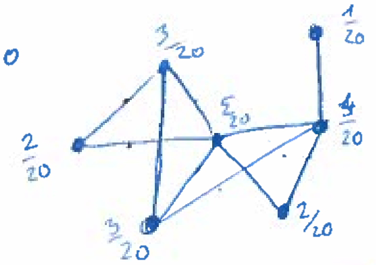
$$c_i^{p=0} = \begin{cases} -k & \text{if } i = 0 \\ 1 & \text{if } i \in \{1 \dots \frac{k}{2}, N-\frac{k}{2} \dots N-1\} := S_1 \quad \text{(ring)} \\ 0 & \text{if } i \in \{\frac{k}{2}+1, \dots, N-\frac{k}{2}-1\} := S_2 \quad \text{(off-ring)} \end{cases}$$

average total weight $p \cdot k \cdot N/2 \mathrel{\hat{=}}$ # expected rewired links is subtracted UNIFORMLY from existing edges of the Ring and then distributed uniformly on ALL possible edges (on and off ring)

BARABÁSI-ALBERT NETWORKS

Growth + preferential attachment (Earlier lesser known model by Price for citation growth)

① Start from some network m_0



② Add nodes with degree $k = m$

→

③ connect stumps to nodes in the existing network with

$$P_{k_i} = \frac{k_i}{\sum_j k_j} = \frac{k_i}{20}$$

degree distribution of B-A-Networks

1+2+2+3+3+4+5

Probability a new vertex attaches to a node of degree k:

$$\frac{k P_k}{\sum_k k P_k} = \frac{k P_k}{2m}$$

⟹ expected # degree k nodes to gain a link

$$m \cdot \frac{k P_k}{2m}$$

⟹ use this to formulate master equation

$$\boxed{\underbrace{(n+1)P_{k,n+1}}_{\text{expected \# nodes with degree } k \text{ at step } (n+1)} = \underbrace{n P_{k,n}}_{\text{expected \# nodes with degree } k \text{ at step } n} - \underbrace{\tfrac{1}{2} k P_{k,n}}_{\text{expected \# deg } k \text{ nodes gaining link}} + \underbrace{\tfrac{1}{2}(k-1) P_{k-1,n}}_{\text{expected \# deg } (k-1) \text{ nodes gaining a link}}}$$

for $k > m$

or

$$\boxed{(n+1) P_{m,n+1} = n P_{m,n} - \tfrac{1}{2} m P_{m,n} + 1}$$

$\langle k\rangle = Np \Rightarrow$ we can work out $K(\langle k\rangle)$ or the radius for which we get a desired target mean degree

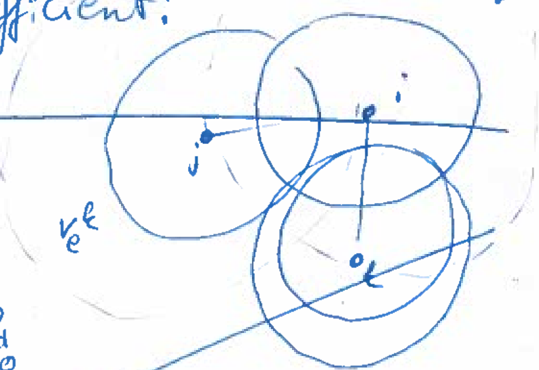
$p = \frac{\langle k\rangle}{N}$ ↓ equate to above expression & solve for R

$$R = \frac{1}{\sqrt{\pi}}\left(\frac{\langle k\rangle}{N}\,\Gamma\left(\frac{d+2}{2}\right)\right)^{1/d}$$

Random Geometric Graphs

- Null-model for embedded networks in space
- "randomized" alternative to grids
- many things can be computed about them
- model for peer to peer networks of devices based on short-range connections such as wifi or blue tooth "ad-hoc-networks"

we can still analytically compute things about this! namely the clustering coefficient!



$$\langle C_d\rangle = \underbrace{\frac{1}{V_e}\int_{V_e}}_{\text{which is also in the excluded volume of j}} \rho_d(r)\, dV$$

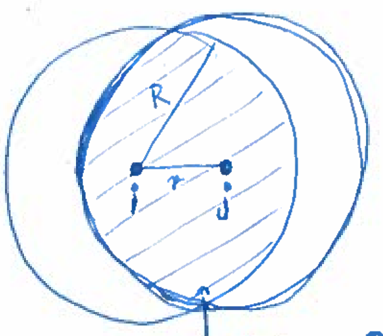
fraction of the excluded volume of i

overlap ρ_d of the two spheres

In 1d:

$$\rho_1(r) = \frac{(2R-r)}{2R} = 1 - \frac{r}{2R}$$

$$\langle C_1\rangle = \frac{1}{2R}\cdot\int_0^{2R}\left(1-\frac{r}{2R}\right)dr = \frac{1}{2R}\cdot R$$



$$C_d(r) = \frac{\rho(r)}{V_e}$$

fractional volume overlap $= \rho_d(r) = \frac{\text{absolute overlap}}{\text{whole volume}}$

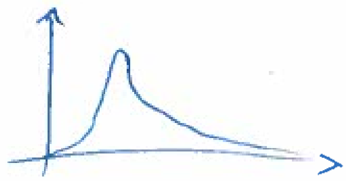
Intuition behind this:

Small q: "you never have time for me"
lower-degree side feels neglected.
when chosing a new connection however, they still want a popular friend linearly more

large q: "you're not a sufficiently valuable connection"
higher-degree side thinks they can "do better"

Results:

→ this process is repeated until an equilibrium in the degree distribution is reached



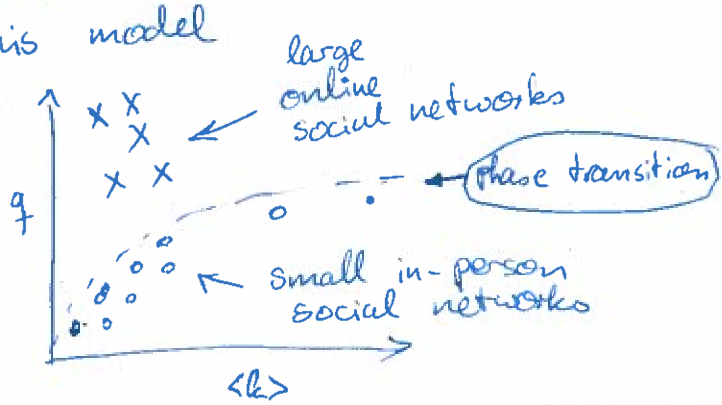
Small q: peaked degree distribution

large q: power law-ish

very large q: two peaks
$q > 0.98$

↑
sort of leader-follower network

real networks fitted with this model



NON-overlapping CHain

Proteins are chains of amino acids. Amino acids are physical objects that can't overlap indefinitely

1/

Exponential Random Graphs
OR Canonical Network Ensembles

Defining a network ensemble through its network measure.

→ sort of the opposite of what we did before where we defined it through the formation mechanisms and then computed expected network measures

→ more of a "statistical mechanics approach"

- fix the desired average value of a network measure
- draw from an ensemble that tends to be close to that value

⇒ this approach first came from social sciences, where you had measurements)

Let $\mathcal{G}$ be the set of all graphs with n nodes

and $\sum_{G \in \mathcal{G}} P(G) = 1$

and let $x_1, x_2, \dots$ be network measures

$$\Rightarrow \langle x_i \rangle = \sum_{G \in \mathcal{G}} P(G)\, x_i(G)$$

↑ now we fix this

Now under these Constraints we want to MAXIMIZE ENTROPY

So we are looking for the most generic ensemble satisfying the constraints.

⇒ $\boxed{\text{Use Lagrange Multipliers to maximize Gibbs Entropy under network measure constraints}}$

Lagrange multipliers:

$$\mathcal{L}(x, \lambda) = f(x) - \lambda g(x)$$

← constraint $g(x) = 0$: $g(x) = \left(\sum_{G \in \mathcal{G}} P(G) - 1\right)$

and $g(x) = \left(\sum_{G \in \mathcal{G}} P(G)\, x_i(G) - \langle x_i \rangle\right)$

↑ Gibbs Entropy $S = -\sum_{G \in \mathcal{G}} P(G) \ln(P(G))$

Simplest Example:

Erdös - Renyi again (sort of)

$H = \beta M$ as M is what we want to fix

$\Rightarrow P(G) = \frac{e^{\beta M}}{Z}$ where $Z = \sum_G e^{\beta M}$

Now all this summing over G can be done using adjacency matrices:

$$\boxed{M = \sum_{i<j} A_{ij}}$$

$$Z = \sum_{\{A_{ij}\}} e^{\beta \sum_{i<j} A_{ij}} = \sum_{\{A_{ij}\}} \prod_{i<j} e^{\beta A_{ij}} = \prod_{i<j} \sum_{A_{ij}=0,1} e^{\beta A_{ij}} = \prod_{i<j} (1+e^{\beta})$$

all allowed values of A_{ij}

$$= (1+e^{\beta})^{\binom{n}{2}}$$

$\Rightarrow$ compute free Energy $F = \ln(Z) = \binom{n}{2} \ln(1+e^{\beta})$

$$\langle M \rangle = \frac{\partial F}{\partial \beta} = \frac{e^{\beta}}{1+e^{\beta}} \binom{n}{2} = \binom{n}{2} \frac{1}{e^{-\beta}+1}$$

if we have a particular value $\langle M \rangle$ in mind we should fix β to

$$\beta = \ln\left(\frac{\langle M \rangle}{\binom{n}{2} - \langle M \rangle} \right)$$

probability for node-pair to be connected:

$$P_{i'j'} = \langle A_{i'j'} \rangle = \frac{1}{Z} \sum_{\{A_{ij}\}} A_{i'j'} e^{\beta \sum_{k<j} A_{ij}} = \frac{\sum_{A_{ij}=0,1} A_{i'j'} e^{(\beta A_{i'j'})}}{\sum_{A_{i'j'}=0,1} e^{\beta A_{i'j'}}}$$

$$= \frac{(0+e^{\beta})}{(1+e^{\beta})} = \frac{1}{1+e^{-\beta}} = \frac{\langle M \rangle}{\binom{n}{2}}$$ just like ER

RELATIVE CANONICAL NETWORK ENSEMBLES

so far we've maximized ENTROPY, however, is that reasonable?
why would Entropy always be maximal?

⇒ Real world networks may not actually be completely maximized in Entropy but relatively maximized to something we may know about the formation process, which is the

Background Ensemble

- Essentially the Background ensemble is what you get if you run the Ensemble for infinite temperature.
(It's practically very convenient to define it like this :)

$$\boxed{P^R_{\beta q} = \frac{e^{-\beta R(G)}}{Z_{R,q}(\beta)}\, q(G)}$$

R: Property function
β: inverse temperature/genericity
q: Background ensemble

RCNE's as a tool for analyzing network measures

- While Canonical network models are a good way of constructing network ensembles based on network measures and Background ensembles, their realism can be limited and the amount of things you can actually compute can be ~~[illegible]~~ constrained by computational complexity

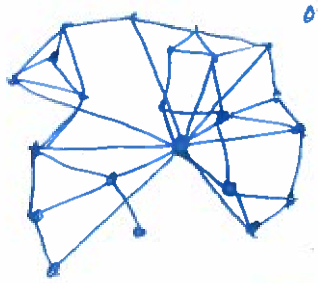
- But they can actually be used as a way to analyze the NETWORK MEASURES THEMSELVES what about another measure?

$$R_{WL} = W \cdot L$$

L ← Average shortest path length

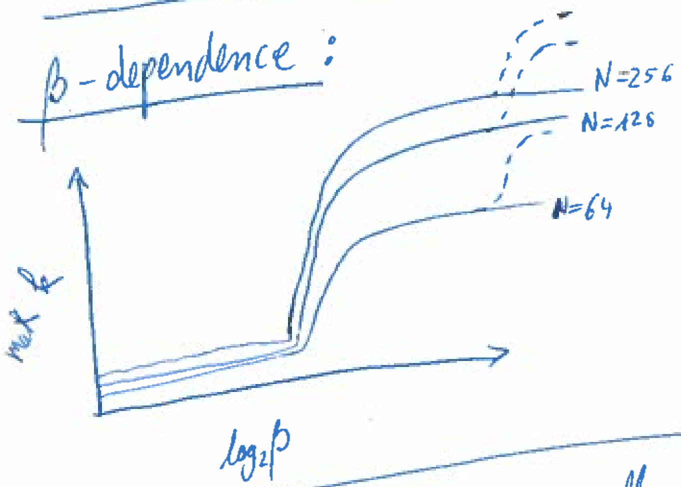
W ↑ total Euclidean Length of all Links

→ Better but still no



one Hub with long links, short links otherwise

β-dependence:



phase transition at some value of β (in all 4 cases)

→ sometimes actually 2 phase transitions one for clique and one for hub-formation

Problems:
15.3