“The finding that real networks are rapidly evolving dynamical systems has catapulted the study of complex networks into the arms of physicists as well.”

Albert-László Barabási - „Linked“, p. 102 (2003)

Network Theory

An executive summary

Lukas Voss | Advanced Interdisciplinary Statistical Methods | PD Dr. Stockburger

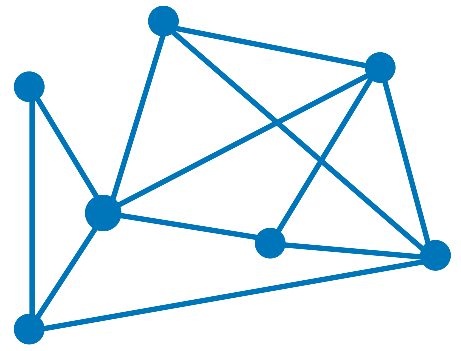
29th January, 2022

ulm university universität ulm

Agenda

Why physicists are curious about nodes and links

FUNDAMENTALS



- What is a network?
  - ➡ Nodes, links and distances
- History of network theory
  - ➡ Random networks (Paul Erdős and Alfréd Rényi)
  - ➡ Scale-free networks by Albert-László Barabási

detailed in the following

OBSERVATIONS & APPLICATION

- Why physicists are interested in networks
- Different topologies of networks
  - ➡ Power-law
- Directed networks
  - ➡ *Islands* emerge

CONCLUSION

- When network theory is a helpful method
  - ➡ Applications and limits of the theory
- Summary
  - ➡ Take home messages

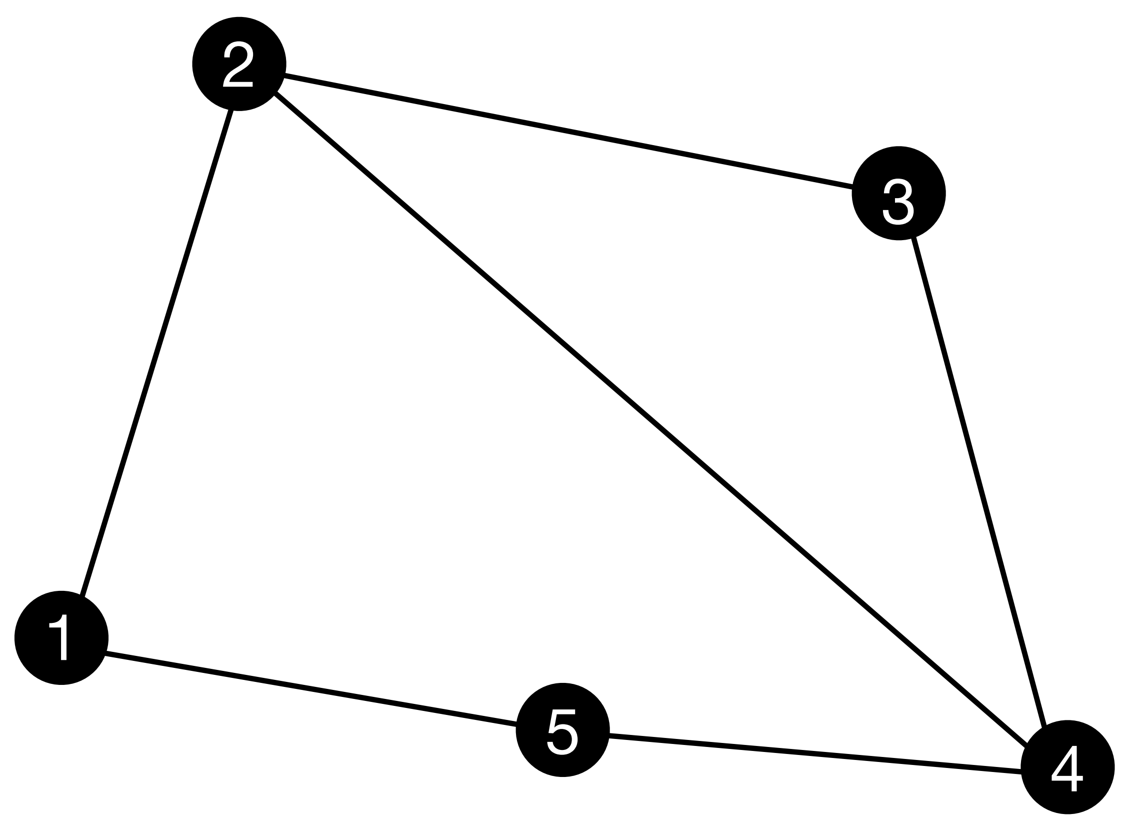
What is a network?

An overview

Graph

A graph G is characterized by the set of nodes P and the set of links/edges E that connect nodes in the graph

$$G = \{P, E\}$$



Characteristics of the upper graph

- $N = 5$
- $P = \{1,2,3,4,5\}$
- $E = \{(1,2), (1,5), (2,3), (2,4), (3,4), (4,5)\}$

History

Leonhard Euler

- First small graphs
- High degree of regularity

Paul Erdős and Alfréd Rényi

- Random graphs
- (More) complex topology
- Unknown organization principles

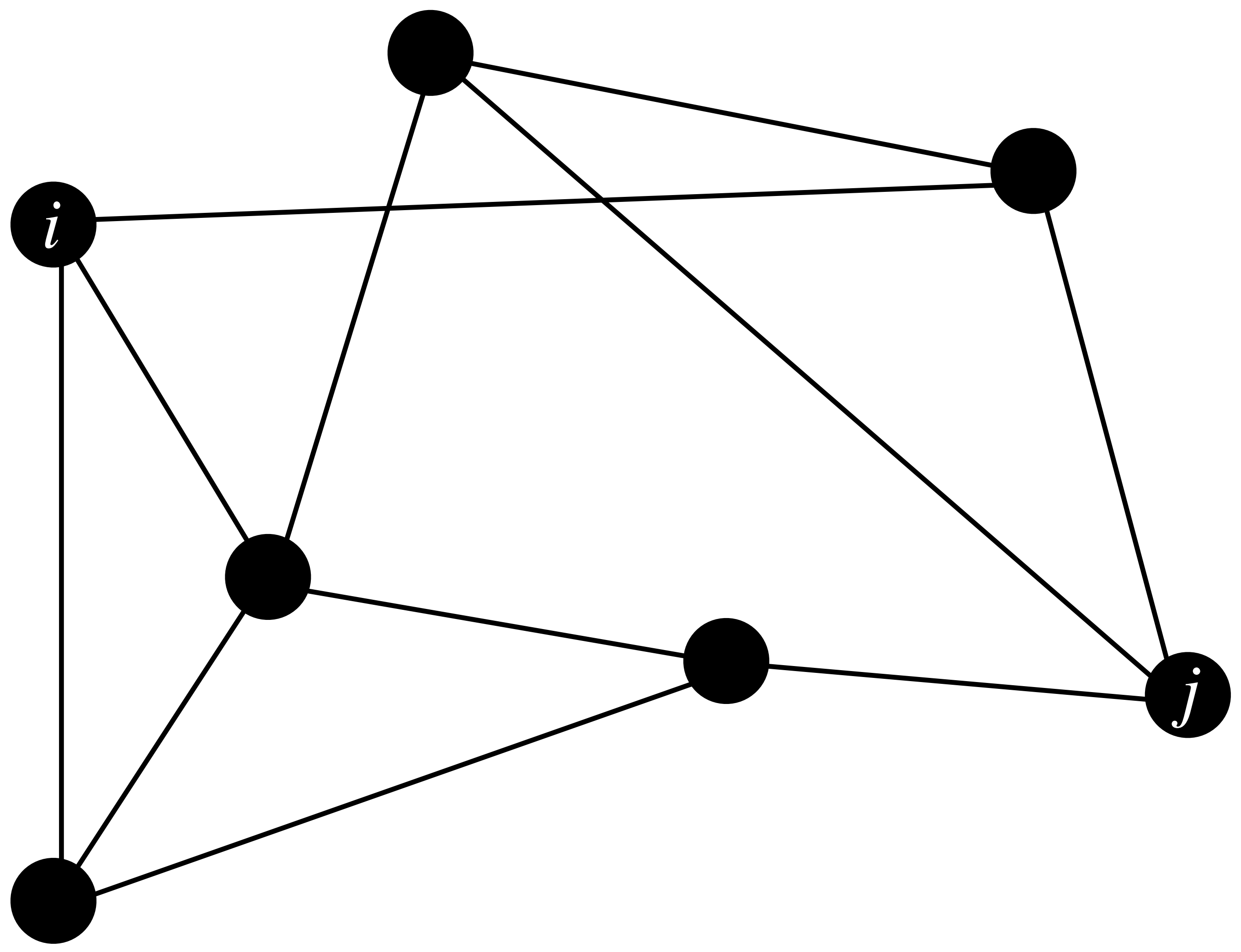
What is a network?

Nodes, links and distances

- A network consists of nodes that are connected by links
- Links can but do no have to have a direction

- Nodes can be interpreted as involved entities in a given system
  ➡ Humans, computers, power stations, etc.

- Distance between nodes is a function of the total number of nodes N and the average number of links per node k



Distance in random networks

$$k^d = N \quad \Longleftrightarrow \quad d = \frac{\log(N)}{\log(k)}$$

Number of links along the shortest path from node i to node j

What is a random network?

The Erdős-Rényi model

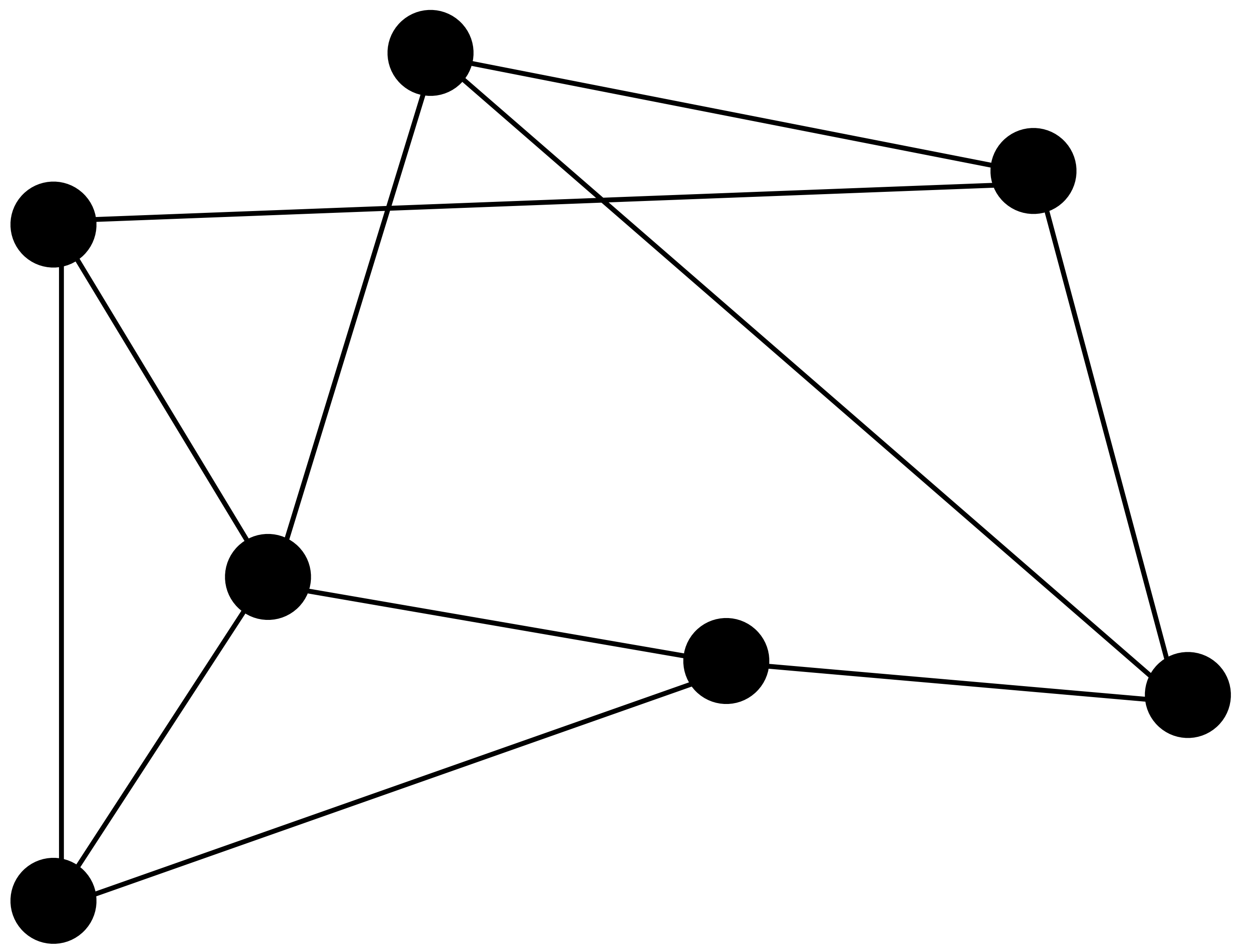
- A random network emerges by randomly adding links to the nodes with probability p
  ➡ At the end almost all nodes have the same k

Degree distribution

$$P(k) = \binom{N-1}{k} \cdot p^k (1-p)^{N-1-k}$$

Probability distribution over the number of links k from one node to other nodes in the network

- Random network theory has dominated scientific thinking since its introduction in 1959
- It requires only one link to stay connected



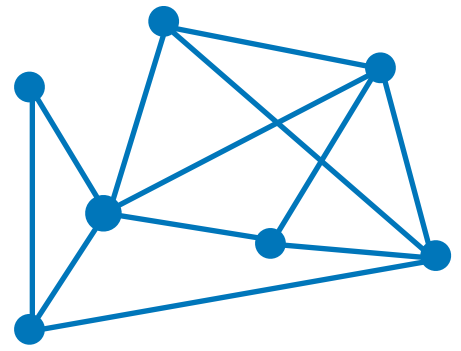
Important property

- For $k > 1$, the number of nodes left out of the giant clusters decreases exponentially
- The more links we add, the harder it is to find a node that remains isolated

Agenda

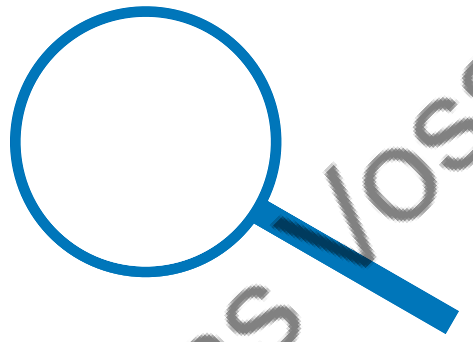
Why physicists are curious about nodes and links

FUNDAMENTALS



- What is a network?
  - ➡ Nodes, links and distances
- History of network theory
  - ➡ Random networks (Paul Erdős and Alfréd Rényi)
  - ➡ Scale-free networks by Albert-László Barabási

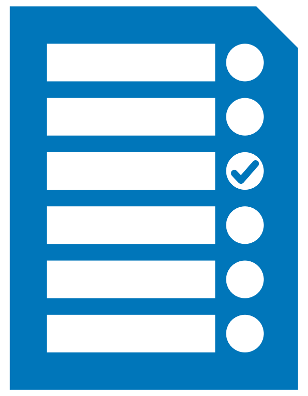
OBSERVATIONS & APPLICATION



- Why physicists are interested in networks
- Different topologies of networks
  - ➡ Power-law
- Directed networks
  - ➡ *Islands* emerge

detailed in the following

CONCLUSION



- When network theory is a helpful method
  - ➡ Applications and limits of the theory
- Summary
  - ➡ Take home messages

Six degrees of separation

How we are actually closer to knowing the American president than we think

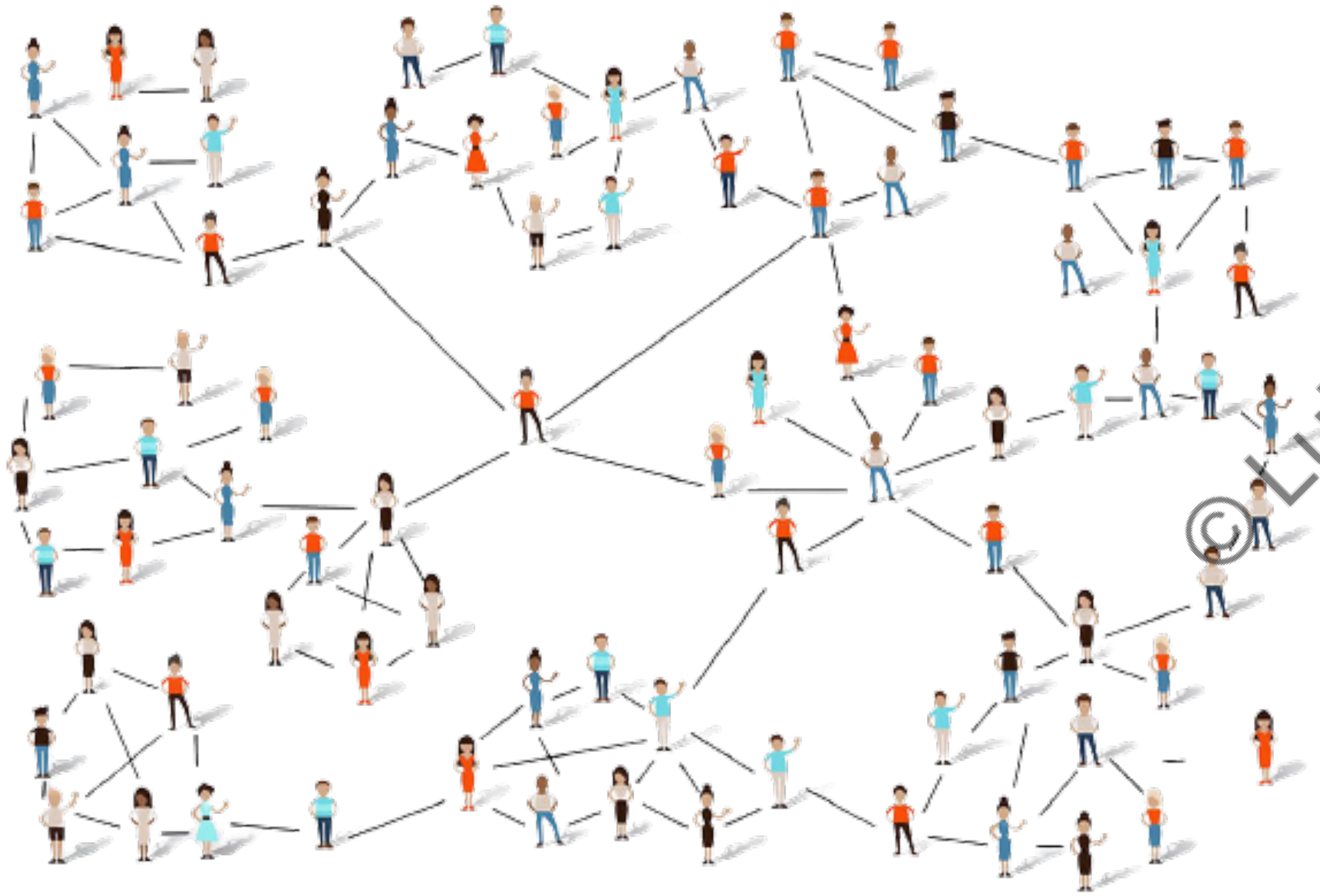


Fig 1 A sketch of a personal network. [8]

Small world problem

- A study by Stanley Milgrim (1963) has investigated the probability that two randomly selected people know each other
  - ➡ Find the average path lengths in the social network of the United States
  - ➡ Results indicate that the average path length is 5.5 to 6 links

On average, any person in the world is only a maximum of 6 links away from you!

- Exception: Isolated tribes (e.g. Sentinelese)

Scale-free networks

How growth and preferential attachment leads to a complex topology with hubs

Growth

Real network grow over time, start with m_0 nodes

- Numerous new website join the WWW every day
- People get to know more people as they get older

➡ Let m be the number of links that get added to the new node

Preferential attachment

The observed probability of new nodes attaching to existing ones is **not** equally distributed

- *Older* nodes have a big advantage

➡ Probability Π of a node connecting to another node i depends on the degree k_i with

$$\Pi(k_i) = \frac{k_i}{\sum_l k_l}$$

➡ Contradiction to the random network theory!

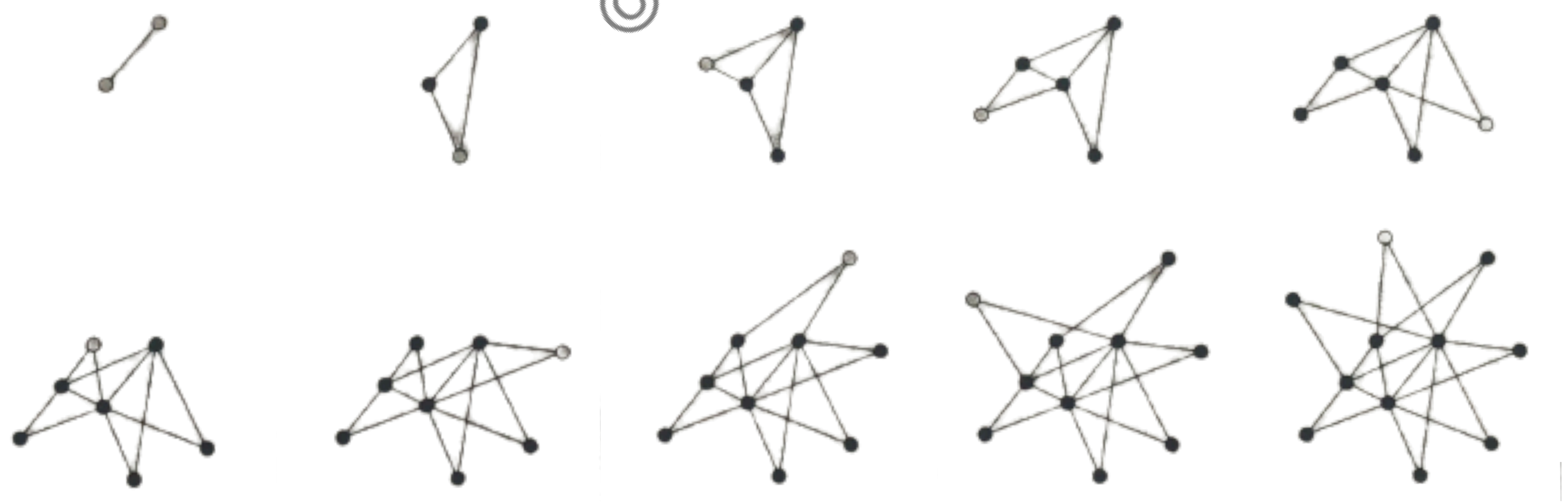


Fig 2 At every step a new nodes joins (lighter circle). Due to growth and preferential attachment, a few highly connected hubs emerge. [7]

Random vs. Scale-free networks

A comparison of degree distributions

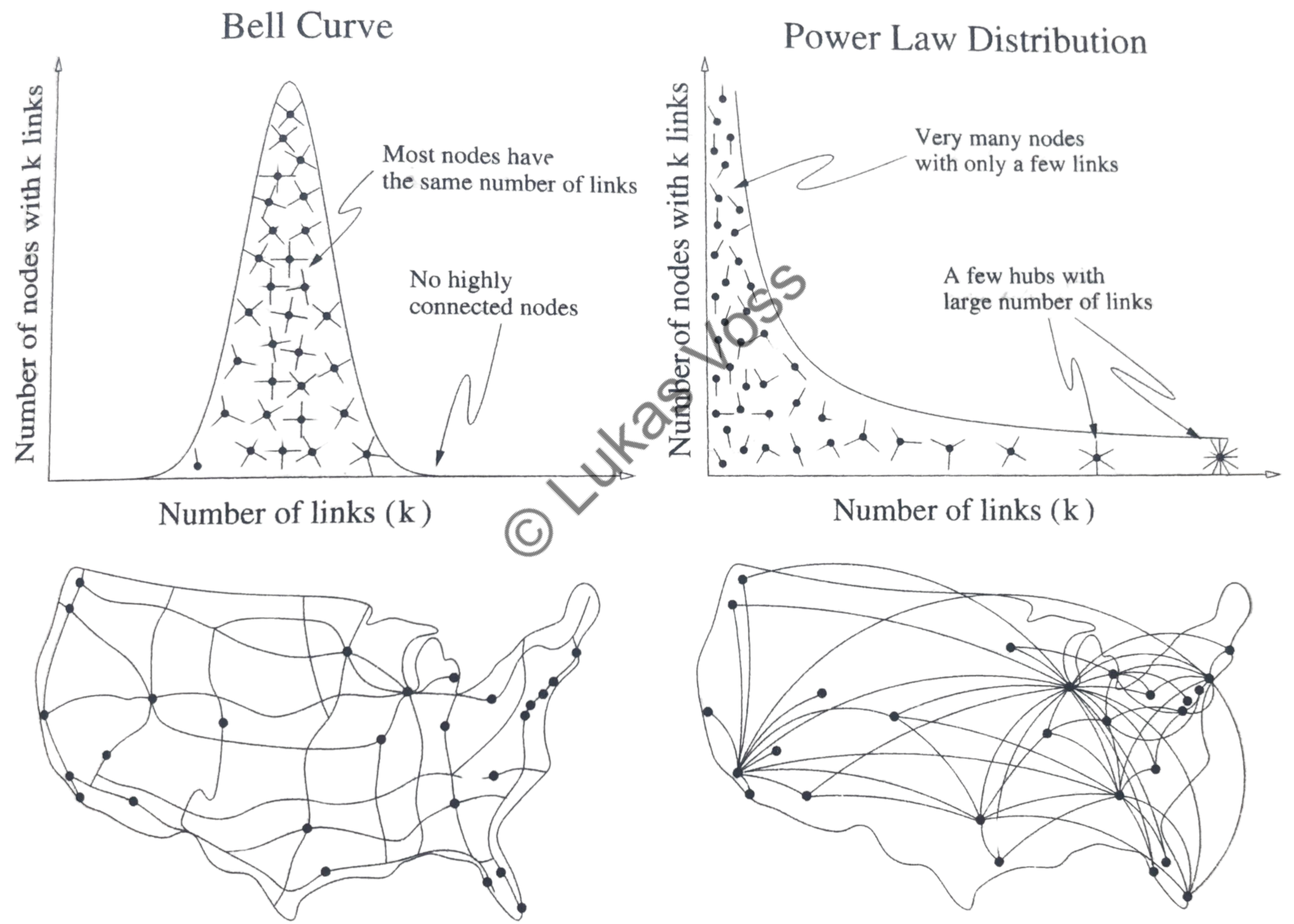


Fig 3 The degree distribution $P(k)$ of a random network follows a bell curve (upper left). The American highway network serves as an example (lower left). Instead $P(k)$ of a scale-free network follows a power law (upper right). Here, the American air traffic system is being presented as an example (lower right). [7]

Power law observations

Examples

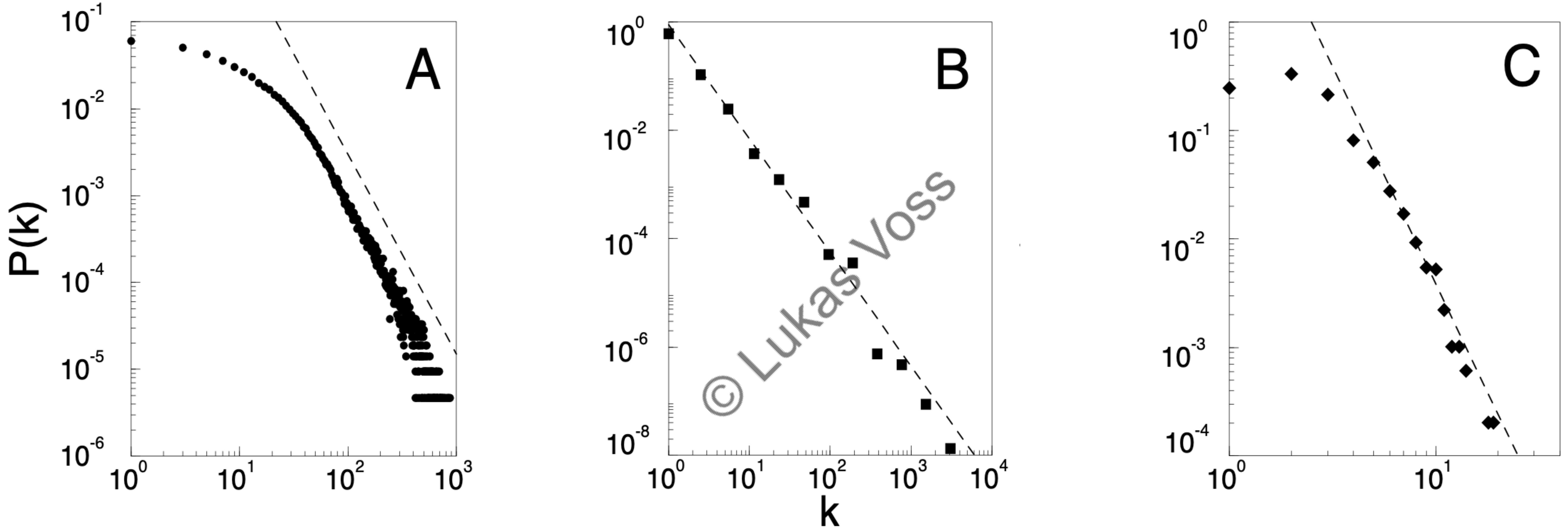


Fig 4 Distribution function of connections for large networks in a log-log plot.

(A) Actor collaboration graph with $N = 212\,250$ links with an average connectivity $\langle k \rangle = 22{,}78$ leading to $\gamma_A = 2{,}3$.

(B) World Wide Web graph with $N = 325\,729$ links with an average connectivity $\langle k \rangle = 5{,}46$ leading to $\gamma_B = 2{,}1$.

(C) Graph of an American power grid with $N = 4\,941$ links with an average connectivity $\langle k \rangle = 2067$ leading to $\gamma_C = 4{,}0$.

Slope of the dashed lines correspond to the respective γ. [1]

ulm university universität uulm

Scale-free networks

How growth and preferential attachment leads to a complex topology with hubs [1]

- Because of preferential attachment, a node that acquired more links than other nodes will increase its connectivity at a higher rate

➡ An initial difference in the connectivity between nodes will increase further as the network grows

- The rate at which a node acquires new links holds for

$$\frac{\partial k_i}{\partial t} = \frac{k_i}{2t} \implies k_i(t) = m\sqrt{\frac{t}{t_i}}$$

t_i : time at which the node i was added to the graph

m : number of links that get added to the new node

➡ Older nodes (smaller t_i) increase their connectivity at the expense of younger nodes (larger t_i)

The winner takes all / The rich get richer phenomenon

➡ Some nodes are highly connected and form a hub/cluster

The fitness of a node

How well does a node acquire new links?

- The fitness η of a node describes its ability to acquire new links compared to the others

  ➡ In the scale-free model a node's fitness is determined by its amount of links

- The probability to attach to a node with k links and fitness η then holds for $\Pi(k) = \frac{k\,\eta}{\sum_l k_l\,\eta_l}$

- Nodes still acquire links following a power law t^{β_i}, but the dynamic exponent is different for each node $\beta_i \propto \eta_i$

The age of a node is less important, now it's the node's fitness η_i that is in the driver's seat

Bianconi-Barabási model

How the fitness of a node helps understanding the Bose-Einstein Condensate

- Assign an energy level ϵ to each node $\epsilon = -\beta^{-1}\log(\eta), \ \beta \propto T^{-1}$ ➡ The larger η, the smaller the corresponding energy level

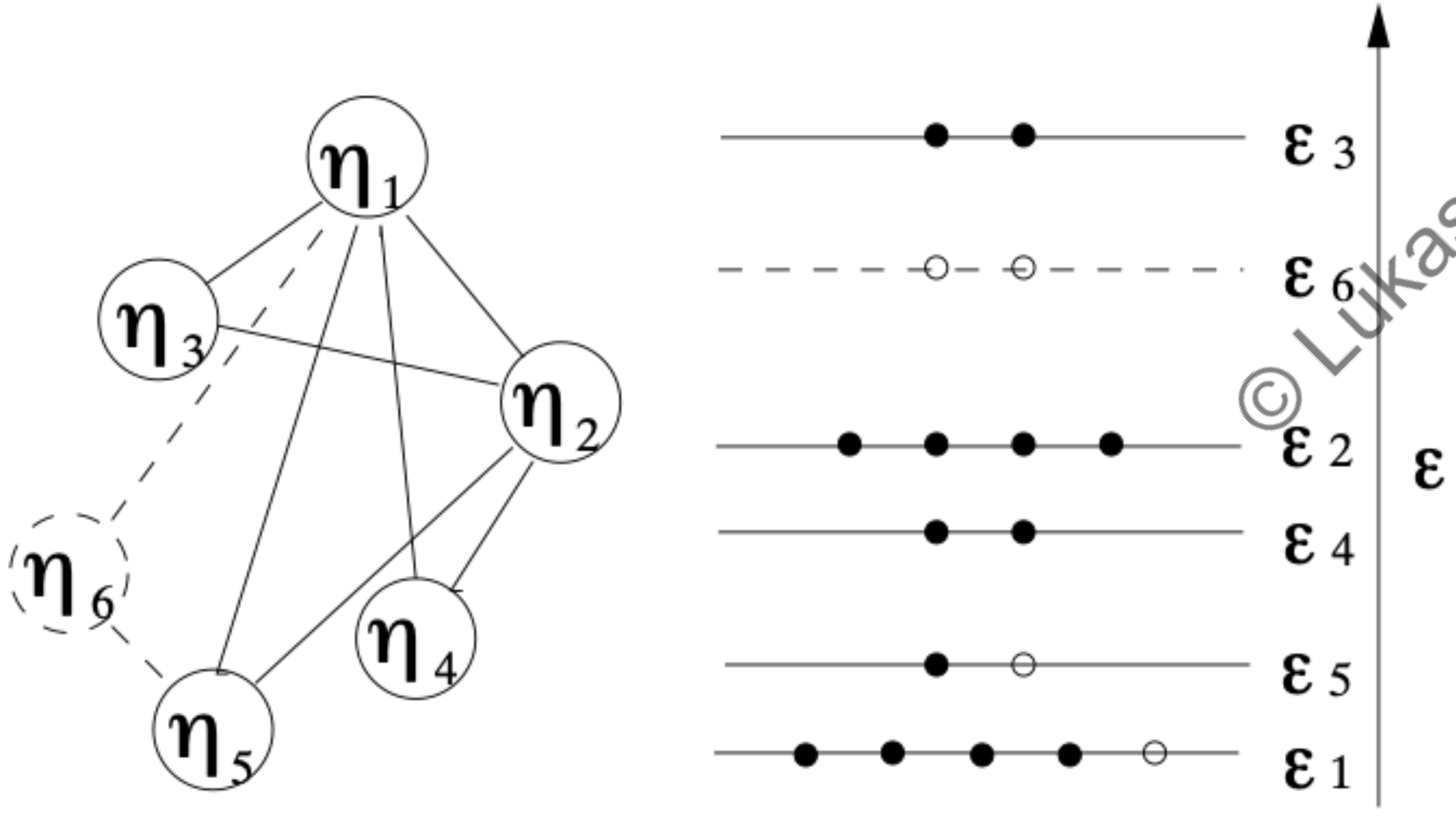


Fig 5 Schematic illustration of the mapping between a network model and a Bose-Einstein Condensate. [9]

➡ Every time step a new node is added that connects to $m = 2$ particles

➡ A link from node i to node j corresponds to a particle at level ϵ_i and one at ϵ_j

➡ Adding the new node with η_6 leads to adding the new energy level ϵ_6

Spread of information

How information propagates through a network

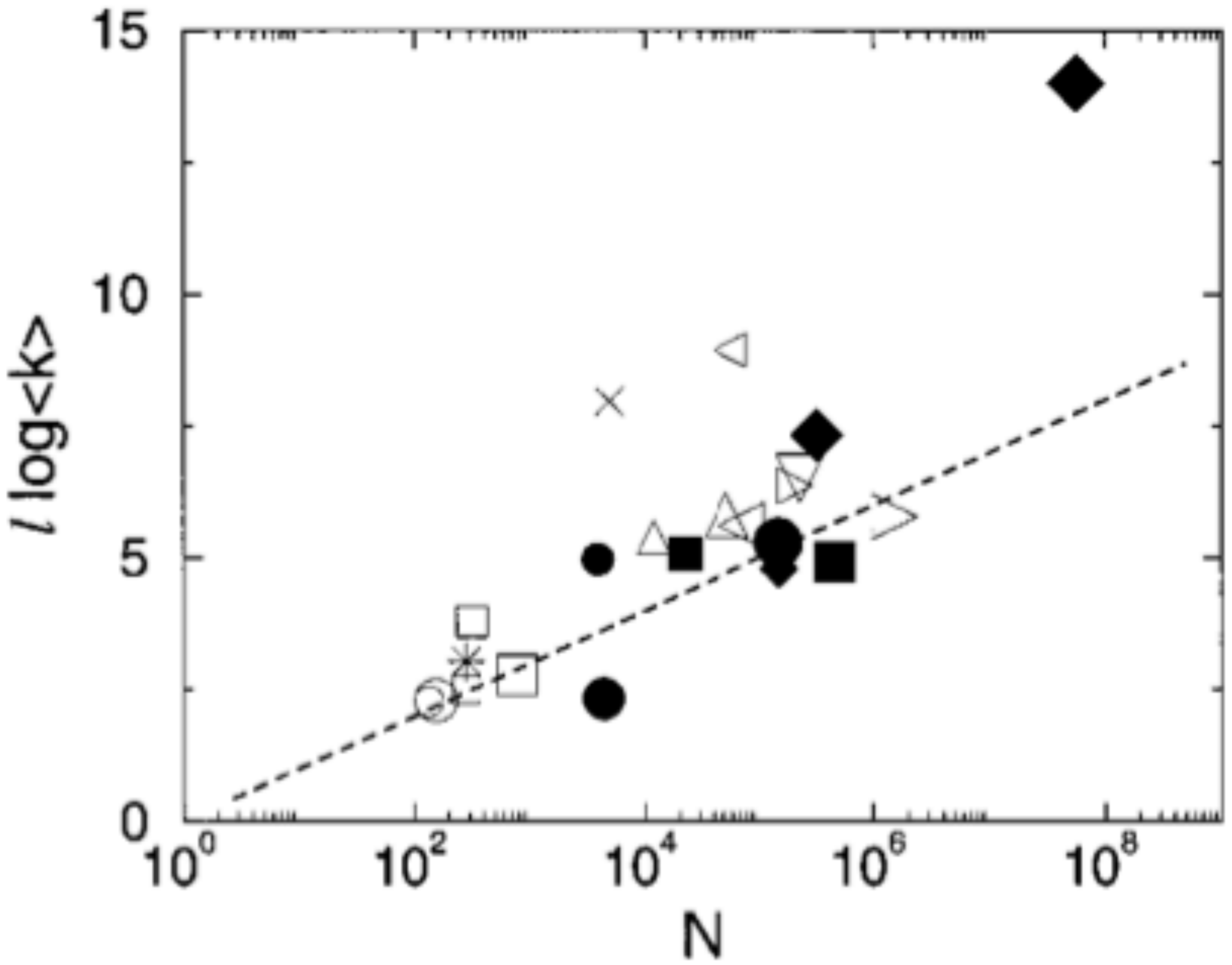


Fig 6 Comparison between the average path lengths of real networks and the prediction (dashed line) of random network theory. [3]

Legend

- □ Metabolic, *E. coli*
- ▷ Medline, co-authorship
- ● Internet, router (1999)
- ◆ WWW (2000)

- Average distance d in a random network

$$d = \frac{\log(N)}{\log(k)}$$

➡ Random network theory is a good estimate for the average distance between a pair of nodes

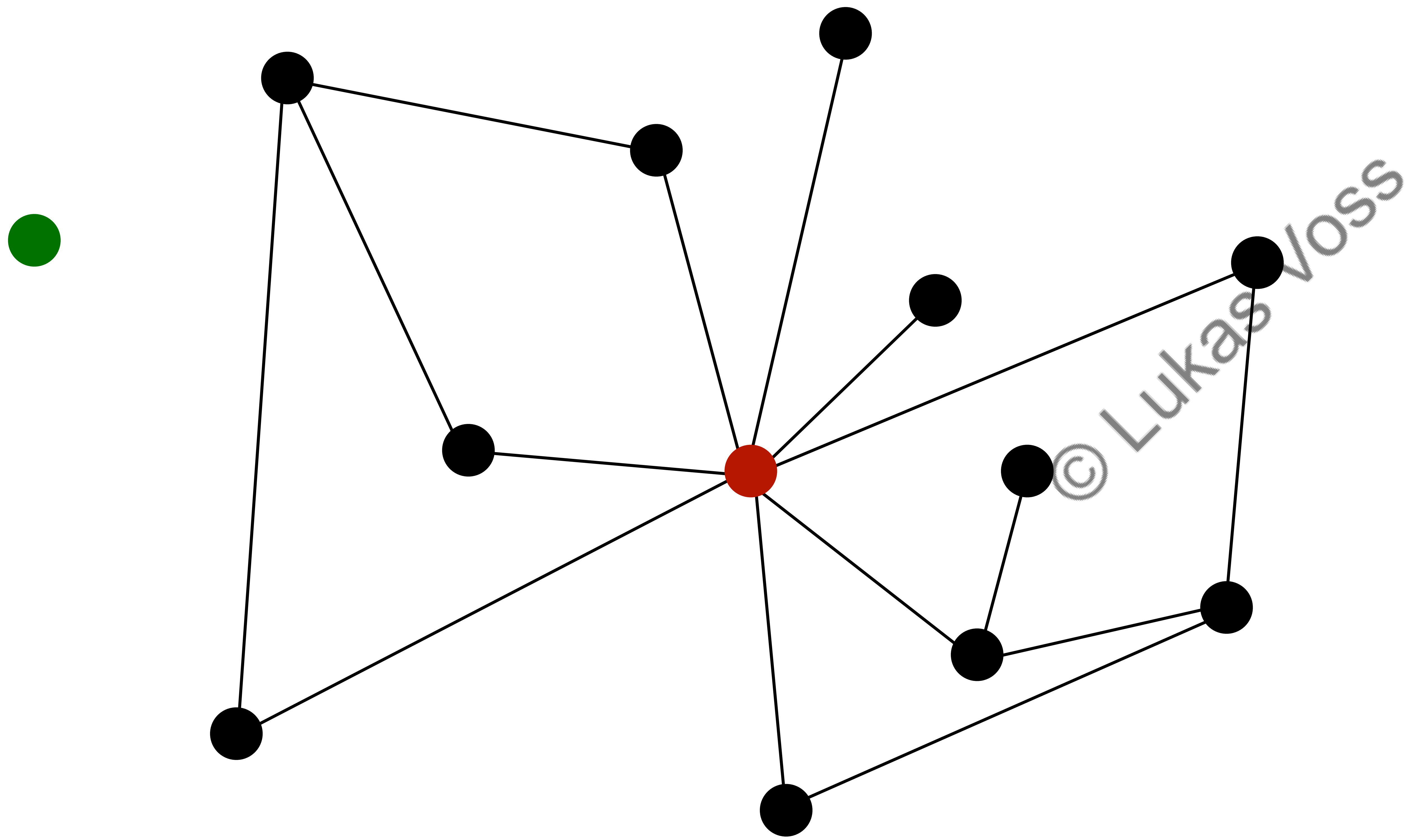
- Each information has a unique spreading rate representing the likelihood that it will be adopted by another node introduced to it

➡ R-value for Covid-19: How many other nodes (people) will be infected (on average) by one node having the information (virus)

$$R = \begin{cases} > 1, & \text{exponential increase in infections} \\ = 1, & \text{linear increase in infections} \\ < 1, & \text{exponential decrease in infections} \end{cases}$$

Application of network theory

German physicist Viola Priesemann: Modeling Covid-19



Arbitrary personal network of physics students

Risk Classification

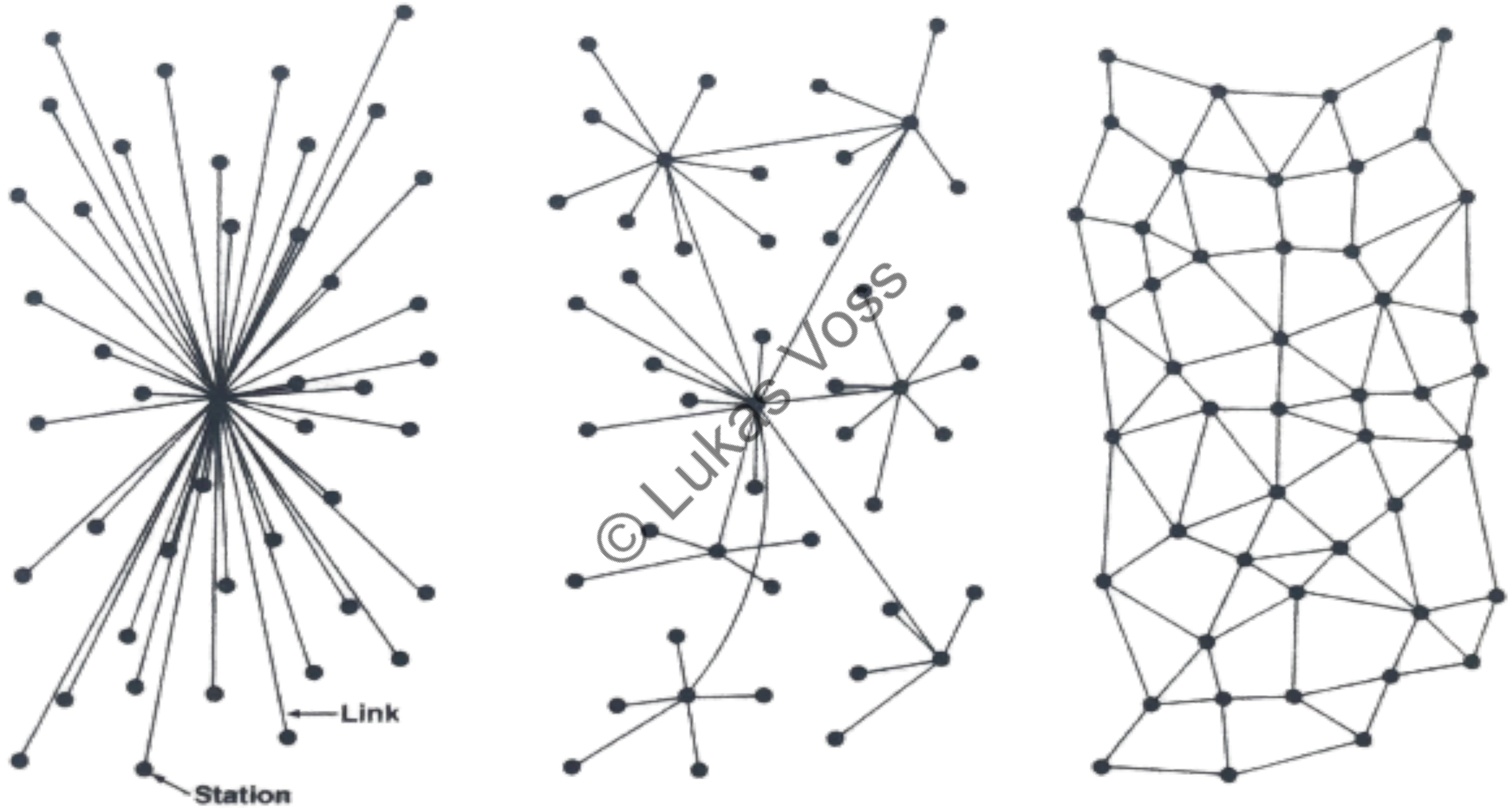
Super spreader $\implies R \gg 1$

Isolated person $\implies R \ll 1$

➡ For the whole network an average R value will be calculated

Topology of networks

Why their architecture matters a lot



Centralized
(A)

Decentralized
(B)

Distributed
(C)

Fig 7 Paul Baran was thinking about the ideal structure of the Internet back in the sixties. At that time the structures A, B were dominating the architecture of the communication system - but they are vulnerable. He therefore proposed a distributed architecture as shown in C. [7] following Paul Baran (1964)

Directed networks

Different areas between nodes

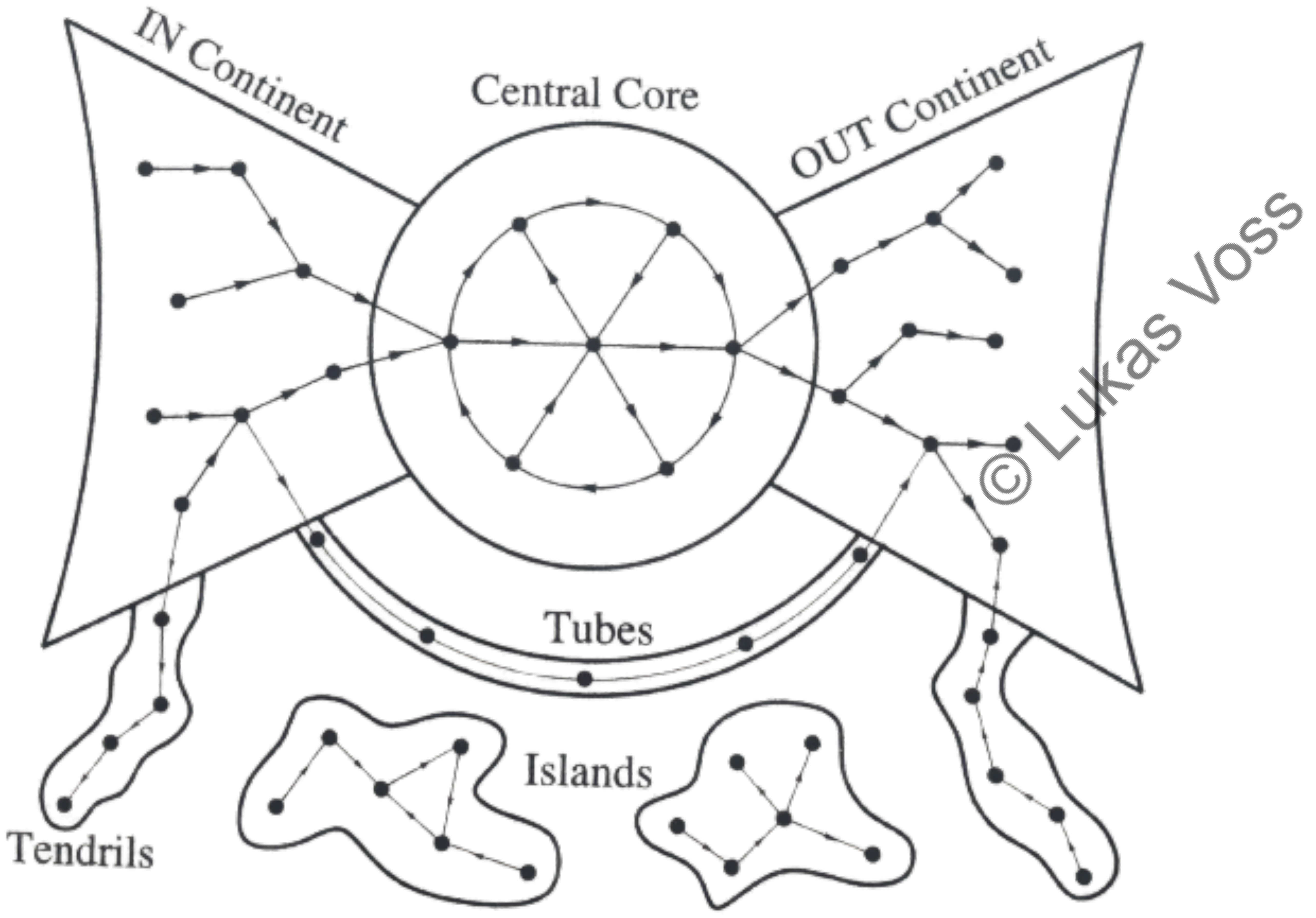


Fig 8 The Continents of a directed network [7]

Central Core

In the central core every node can be reached from every other node

IN Continent

Following the links will bring us to the central core

But starting from the central core the IN Continent in unreachable

OUT Continent

All nodes of the OUT continent can be reached from the central core

Once in the OUT continent there is no way back

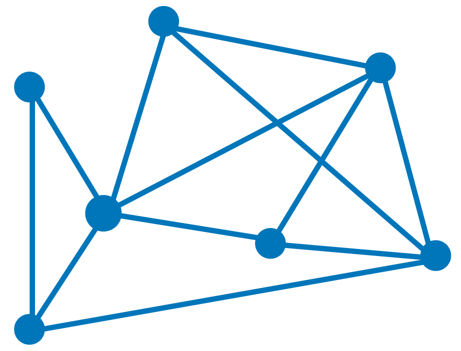
Islands

Isolated nodes that cannot be accessed from the rest of the nodes

Agenda

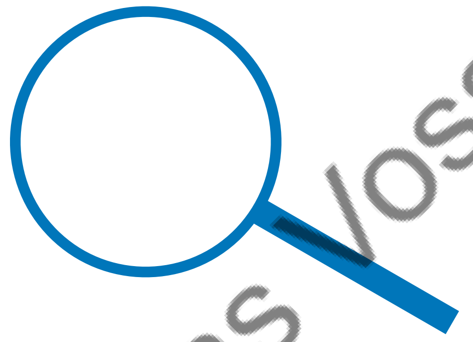
Why physicists are curious about nodes and links

FUNDAMENTALS



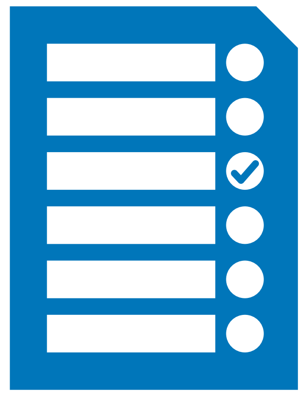
- What is a network?
  - ➡ Nodes, links and distances
- History of network theory
  - ➡ Random networks (Paul Erdős and Alfréd Rényi)
  - ➡ Scale-free networks by Albert-László Barabási

OBSERVATIONS & APPLICATION



- Why physicists are interested in networks
- Different topologies of networks
  - ➡ Power-law
- Directed networks
  - ➡ *Islands* emerge

CONCLUSION



- When network theory is a helpful method
  - ➡ Applications and limits of the theory
- Summary
  - ➡ Take home messages

detailed in the following

Applications in physics and beyond

Where network theory can be used effectively

- The correlation length of fluctuations in phase transitions is used as a rough measure of the cluster size [9,10]

  ➡ They follow a power law with a critical exponent ν

- Financial institutions are highly connected with each other due to lending each other money and sharing obligations

  ➡ The case of a default would also affect many others

- Corporations make use of network effects in their marketing strategy

  ➡ Paying highly connected nodes called *influencers* to promote their product or service is a common way nowadays

Summary

Take home messages

- First ambitions: Random networks
  - ➡ $P(k)$ with exponentially decaying tails
- Many observed network phenomena have a scale-free characteristic
  - ➡ $P(k)$ with a power law $k^{-\gamma}$

- There are interesting parallels in physics that can be explained by Network Theory
  - ➡ Bose-Einstein Condensate
  - ➡ Phase transitions

- Topology is key for the robustness of a network
  - ➡ Internet protocols use distributed architecture
- Network Theory gets applied ranging from marketing to modeling diseases

ulm university universität ulm

Remarks

What can be further considered

- Nodes of a real network often have a lifetime after which they disappear

➡ This can lead to an exponential cut-off for large k [3]

- In the Barabási-Albert model the preferential attachment was assumed to be linear $\Pi(k) \propto k$

➡ In general, $\Pi(k)$ could have an arbitrary non-linear form $\Pi(k) \propto k^{\alpha}$

- Accelerated growth instead of linear could be possible

- Links could have weights

ulm university universität uulm

Sources and literature

References

[1] „Emergence of Scaling in Random Networks“, Barabási and Albert, Science 286, 509 (1999)

[2] „On the evolution of random graphs“, Paul Erdős and Alfréd Rényi, Publ. Math. Inst. Hung. Acad. Sci 5.1 (1960): 17-60

[3] „Statistical mechanics of complex networks“, Albert, R., Barabási, A.-L., Rev. Mod. Phys., 74, 47-97 (2002)

[4] „Networks in motion“, Motter, A. E., Albert, R., Physics Today, 65, no. 4, 43-48 (2012)

[5] „Small Worlds: The Dynamics of Networks between Order and Randomness“, Watts, D. J., (Princeton University Press, Princeton, New Jersey) (1999)

[6] „Exploring complex networks“, Strogatz, S. H., 2001, Nature 410, 268

[7] „Linked“, Albert-László Barabási (2003)

[8] „The theory of Six Degrees of Separation“, Hackerearth-Website (2017)

[9] „Bose–Einstein condensation in complex networks“, Bianconi, Ginestra; Barabási, Albert-László, Physical Review Letters. 86 (24): 5632–5635 (2001)

[10] „Phase transitions in clusters“, Schmidt, M., & Haberland, H., Comptes Rendus Physique, 3(3), 327-340 (2002)