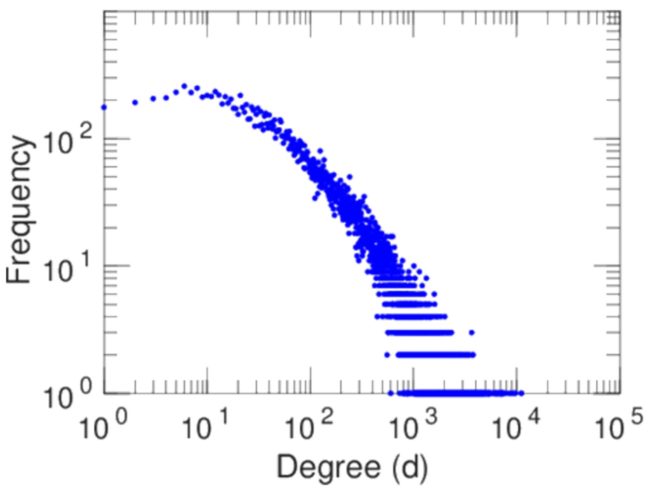
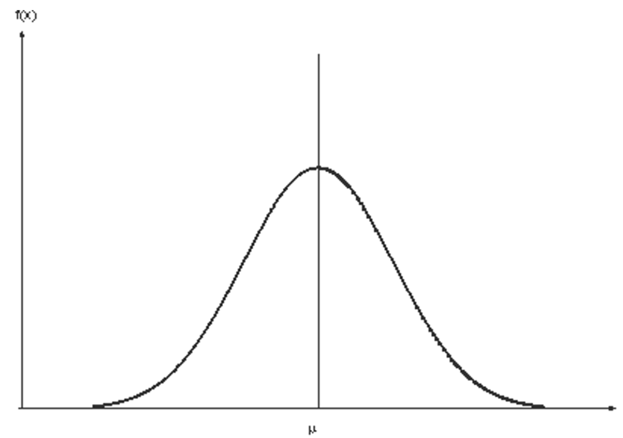
CS4973
network analysis

Some slides adapted from ECE442, Gonzalo Mateos, Univ. of Rochester

Real Networks are not Poisson



n=28,093
m=3,148,447
CC =0.28
Diameter = 9
Mean distance = 2.82

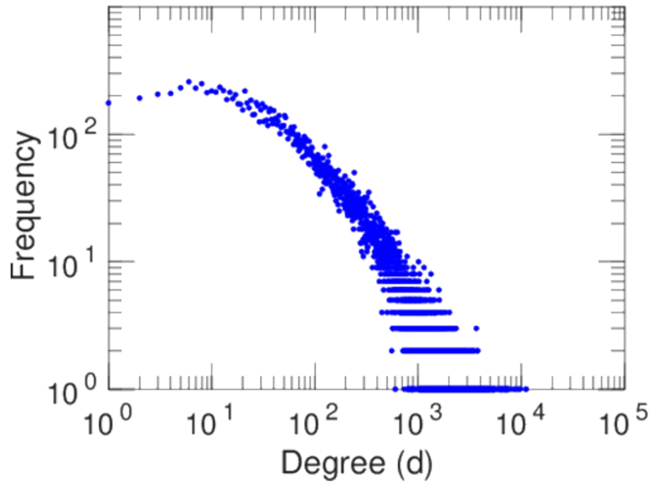


WS Model

k=120

Std. Dev. changes with p

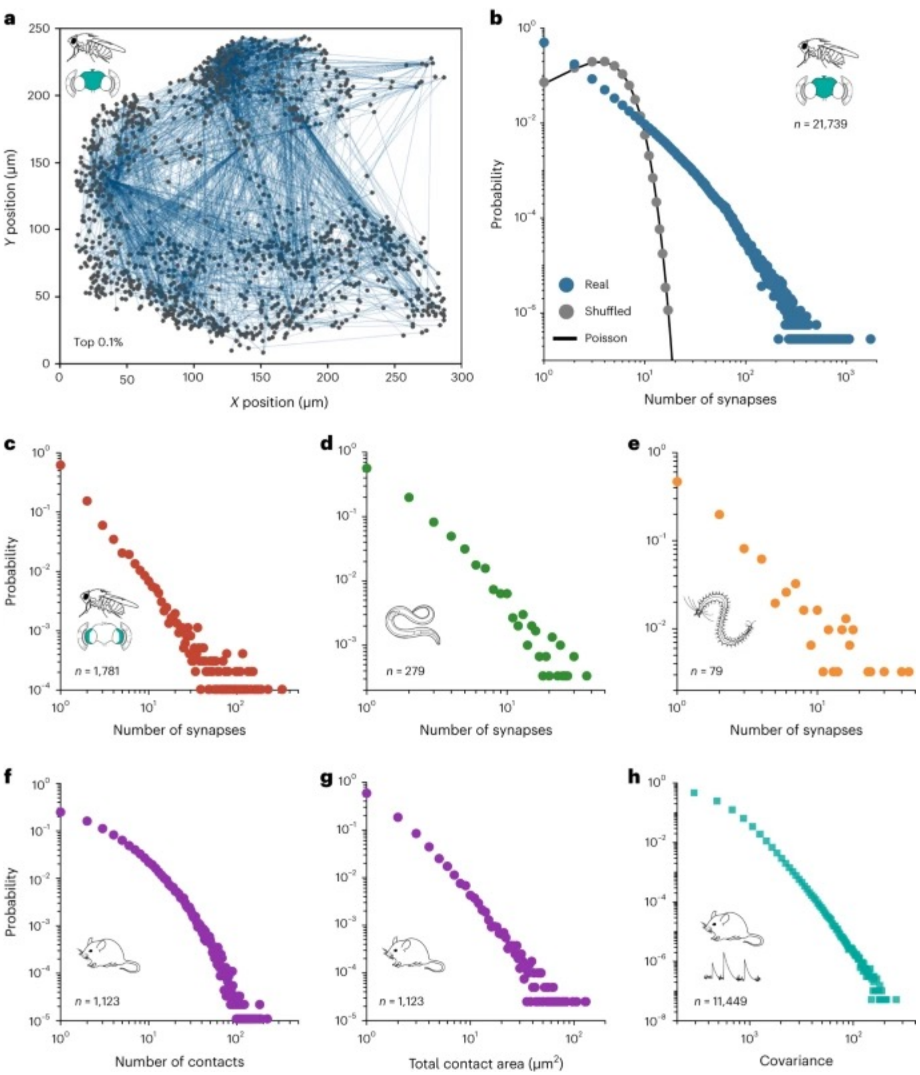
Real Networks are not Poisson



n=28,093
m=3,148,447
CC =0.28
Diameter = 9
Mean distance = 2.82

- Small number of nodes with very high degree.
- Large number of nodes with very low degree.

From 2 weeks ago!



- Connections in networks of neurons are heavy-tailed!
- Pruned at random, rearrange using preferential attachment

Heavy-tailed neuronal connectivity arises from Hebbian self-organization
Christopher W. Lynn, Caroline M. Holmes & Stephanie E. Palmer, ***Nature Physics*** **20**, pages 484–491 (2024)

The Power Law

- Forms a straight-line on a log-log plot.
- $\log P(k) = C - \alpha \log k \Rightarrow P(k) \propto k^{-\alpha}$
- Power law exponent (negative slope) typically in range $[2,3]$.
- Normalization constant C – not super important – highly network dependent.

Scale Free Property

- Scale Free function $f(ax) = bf(x)$, for $a, b \in \mathbb{R}$.
- Power law functions are scale free:
    - $f(ax) = (ax)^{-\alpha} = a^{-\alpha} f(x) = bf(x)$
- Functional form of the distribution is scale-invariant:
    - (Informally,) zooming in on the distribution, shape looks similar. (log-log plots in case of network degree dist)

Discrete Formulation

- As node degrees are positive integers, $k = 0, 1, 2 \ldots$

  What is the probability that a node has exactly k edges?

$$P(k) = p_k = Ck^{-\alpha}$$

- For constant C, use normalization condition:

$$\sum_{k=1}^{\infty} p_k = 1 \Rightarrow C \sum_{k=1}^{\infty} k^{-\alpha} = 1 \Rightarrow C = \frac{1}{\sum_{k=1}^{\infty} k^{-\alpha}} = \frac{1}{\zeta(\alpha)}$$

- Discrete power law dist:

$$p_k = \frac{k^{-\alpha}}{\zeta(\alpha)}$$

Continuous Formulation

- For weighted networks, degrees can be arbitrary (positive) real numbers.

  What is the probability that a node's adjacent edge weights add up to k?

- $p(k) = p_k = Ck^{-\alpha}$

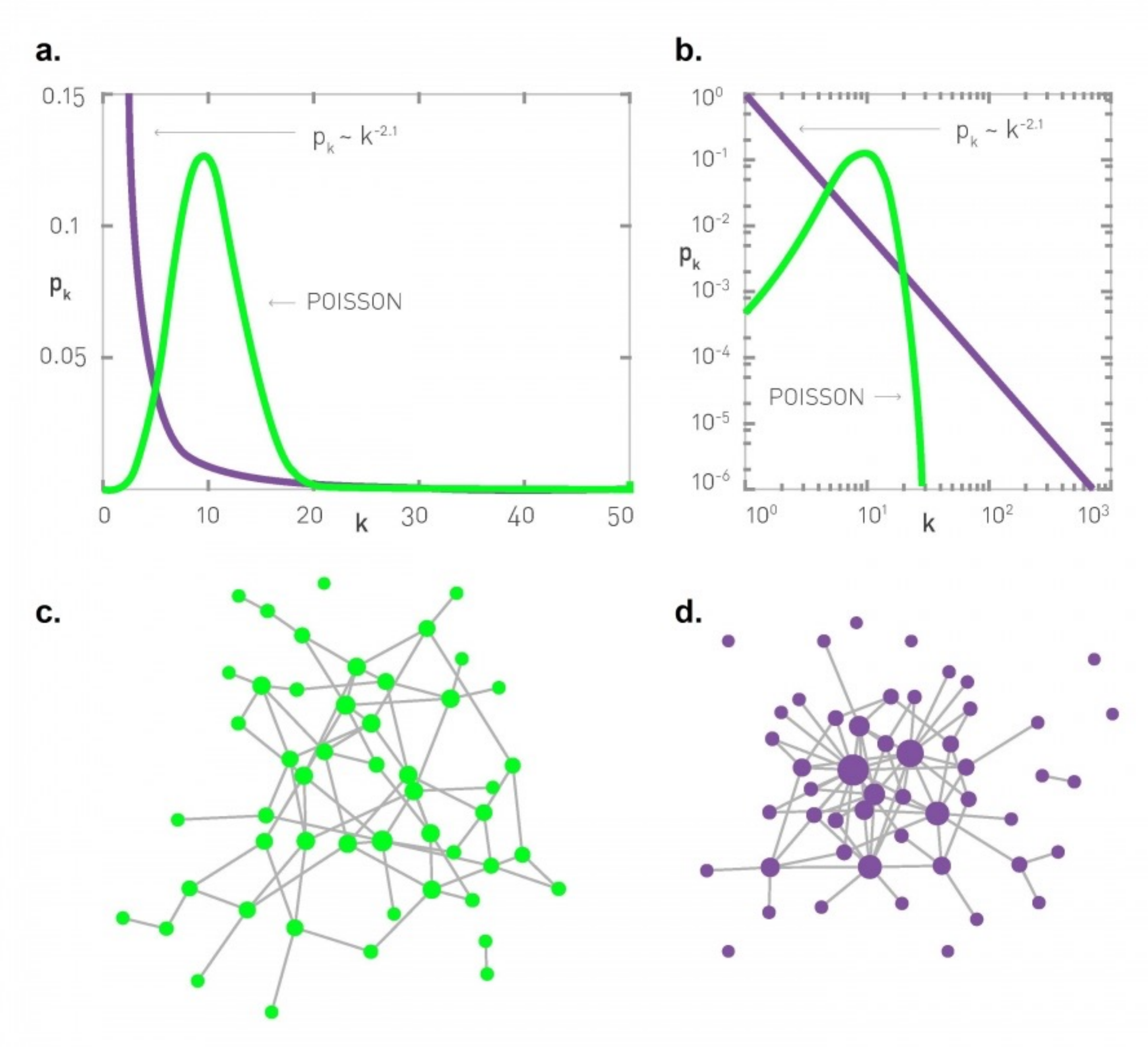
- Normalization condition in a continuous domain:

$$\int_{k=k_{min}}^{\infty} p(k)\, dk = 1 \Rightarrow \int_{k=k_{min}}^{\infty} Ck^{-\alpha} = 1 \Rightarrow C = \frac{1}{\int_{k=k_{min}}^{\infty} k^{-\alpha} dk} = (\alpha - 1)k_{min}^{\alpha-1}$$

$$p_k = (\alpha - 1)k_{min}^{\alpha-1}\, k^{-\alpha}$$

$\int_{k1}^{k2} p(k)dk$ is the probability that a random node has degree between $k1$ and $k2$.

Probability of Observing Hubs



https://networksciencebook.com/chapter/4#hubs

Probability of Observing Hubs

- Consider Wikipedia
  - $n = 6{,}948{,}615$
  - Assume $\bar{k} \approx 15$
- The article on African Humid Period has 965 links!
- What are the probabilities of observing such a node under a random $G(n, p)$ model, and the power law distribution?
- What are the probabilities of observing a node with $k \geq 100$ under a random $G(n, p)$ model, and the power law distribution?

More on "Scale Free"

- The m^{th} moment of a RV, k, under power law distribution $p(k)$ is:

$$\mathbb{E}[k^m] = \int_{k_{min}}^{\infty} k^m \, p(k) \, dk = \frac{\alpha - 1}{k_{min}^{1-\alpha}} \left[\frac{k^{m+1-\alpha}}{m + 1 - \alpha}\right]_{k_{min}}^{\infty}$$

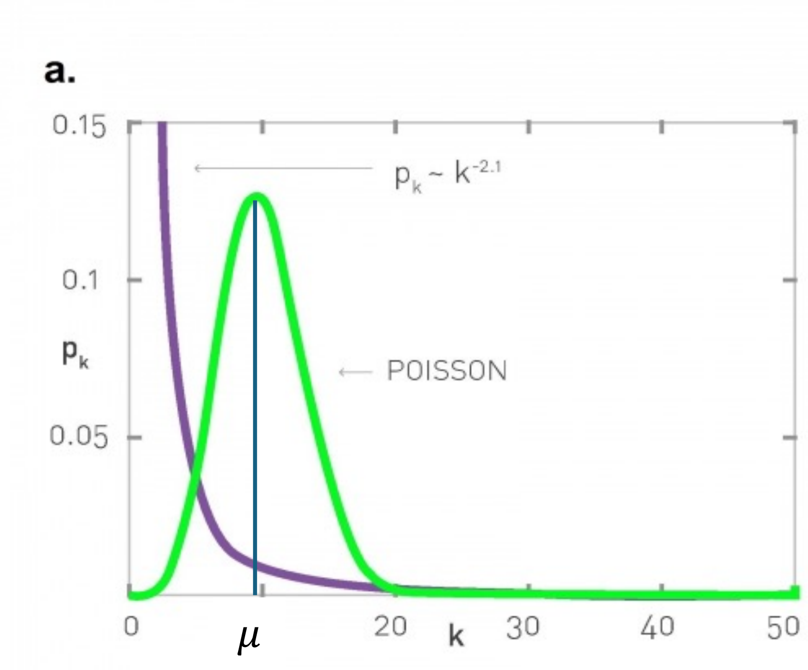
- Integral converges when $m + 1 < \alpha$.

- For real world networks, $\alpha \in (2,3)$, so

$$\mathbb{E}[k] = \left(\frac{\alpha - 1}{\alpha - 2}\right) k_{min} < \infty \text{ and } \mathbb{E}[k^m] = \infty, m \geq 2$$

- The second moment & variance are infinite!

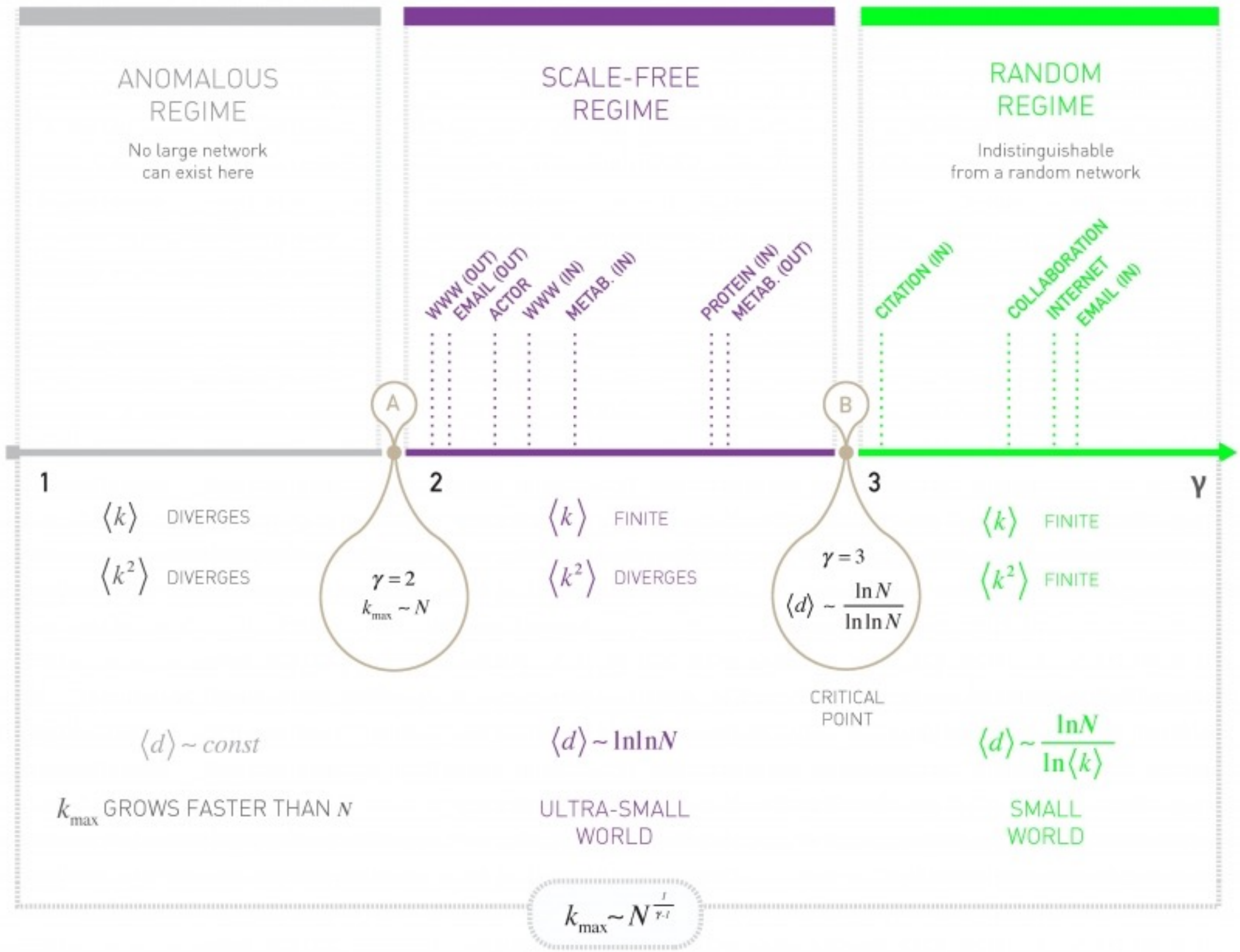
More on “Scale Free”



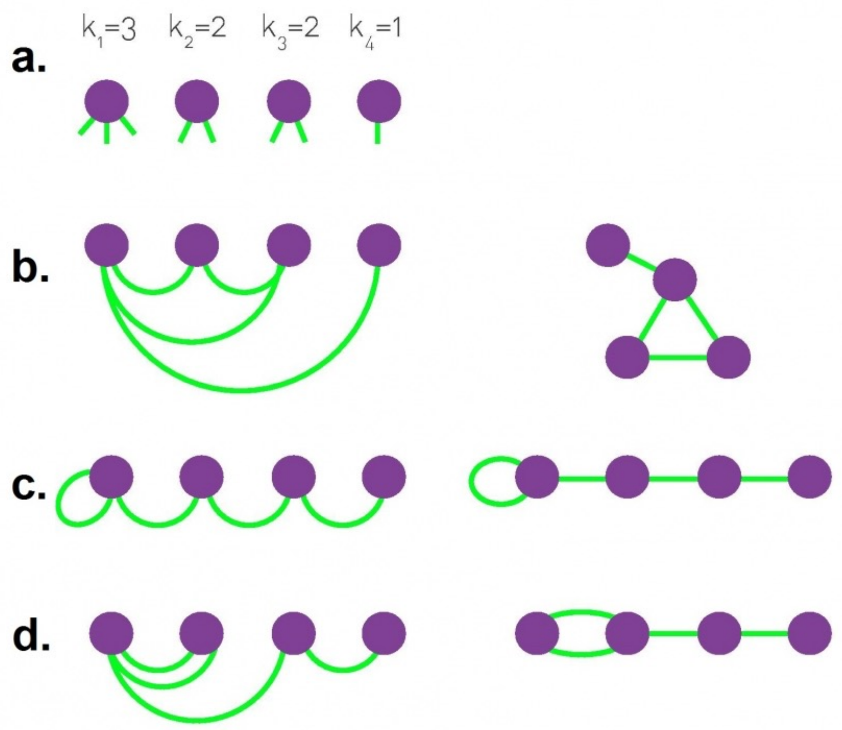
- In $G(n,p)$, prob. that node has degree k is $\mu \pm \sqrt{\mu}$.
  The scale is μ.

- In scale free network, prob. that node has degree k is $\mu \pm \infty$.
  There is no scale.

Dependence on α



Generating Networks with Arbitrary Degree Distributions



Configuration Model

- Assumes given discrete sequence of node degrees.
- Create half-links per given sequence.
- Randomly wire unpaired half-links.