can require a number of spin flips that is exponential in N (A. Haken et al. ca. 1989), and that one can in fact embed arbitrary computations in the dynamics (Orponen 1995). (More precisely, determining whether a given "output spin" is $+1$ or -1 in the local minimum reached from a given initial state is a "PSPACE-complete" problem.)

6.4 The NK Model

Introduced by Stuart Kauffman (ca. 1986) as a "tunable family of fitness landscapes".

A *fitness landscape* is a triple $\langle X, R, f\rangle$, where X is the *configuration* (or *state*) *space*, $R \subseteq X \times X$ is a *neighbourhood relation* on X, and $f : X \to \mathbb{R}$ is a *fitness* (or *objective*) *function*.

A point $x \in X$ is a *local optimum* (of f on X) if

$$f(y) \leq f(x) \quad \forall\, yRx$$

and a *global optimum* (*maximum*) if

$$f(y) \leq f(x) \quad \forall\, y \in X$$

Questions of the "ruggedness" of landscapes (correlation structure), number and height of local optima, sizes of "attraction basins" of local optima with respect to "hill-climbing" algorithms etc. are of great interest for natural landscapes.

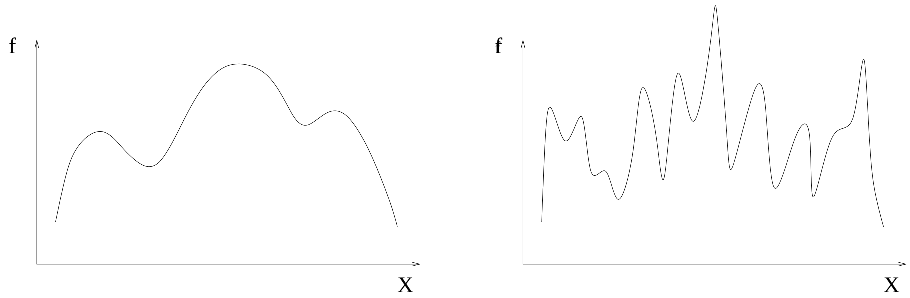
In Kauffman's NK models, $X = A^N$ (usually just $X = \{0,1\}^N$) and K is a tunable neighbourhood size parameter that influences the landscape characteristics, especially its ruggedness (cf. Figure 4).

The model can be seen as a toy model of "epigenetic interactions in chromosomes" — or also a generalisation of the spin glass model.

In Kauffman's model, a *chromosome* is a N-vector of *loci* (*genes*, "positions"), each of which has a value from a set of *allelles* A (usually just $A = \{0,1\}$). A "filled-in" chromosome $\alpha \in A^N$ is called a *genotype*.

The fitness of each gene $i \in \{1,\dots,N\}$ in a genotype $\alpha = (a_1,\dots,a_N) \in A^N$ depends on the allele a_i and K other alleles $a^i_1,\dots,a^i_K$ via some local fitness function $f^i(\alpha) = f^i(a_i; a^i_1,\dots,a^i_K)$, usually normalised so that $f^i(\alpha) \in [0,1]$. The total fitness of a genotype $\alpha \in A^N$ is the normalised sum of its genes' local fitnesses:

$$f(\alpha) = \frac{1}{N}\sum_{i=1}^{N} f^i(a_i; a^i_1,\dots,a^i_k) \quad \in [0,1].$$



(a) $K \ll N$ (b) $K \approx N$

Figure 4: A smooth (a) and a rugged (b) NK fitness landscape.

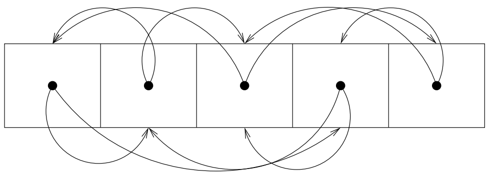


Figure 5: An NK interaction network with $N = 5$, $K = 2$.

Figure 5 illustrates an NK network with five loci and two “epigenetic interactions” per locus.

In Kauffman’s versions of the model, the K loci affecting locus i can either be systematically selected as e.g. $i+1,\dots,i+K(\mathrm{mod}N)$, or the chromosome can be simply “randomly wired”. The f^i are usually determined as randomly generated 2^{K+1}-element “interaction tables”.

From the spin glass perspective, e.g. a 1-D Ising model with N spins can be seen as an $N2$ network where $f^i(S_i;S_{i-1},S_{i+1}) = \frac{J}{2}(S_{i-1}S_i + S_iS_{i+1})$, and an SK spin glass with coefficients J_{ij} and local fields h_i as an $N(N-1)$ network where

$$f^i(S_i;\sigma\setminus\{S_i\}) = \frac{1}{2}\sum_{\langle ij\rangle} J_{ij}S_iS_j + h_iS_i.$$

Basic properties of the NK model, for binary alleles $A = \{0,1\}$ and varying values of K, include the following:

$K = 0$**:**

If $f^i(0) \neq f^i(1)\ \forall i = 1,\dots,N$, then there is a unique global optimum, which is easily found by e.g. the obvious 1-locus mutation “hill-climbing” algorithm.

Expected length of the hill-climbing path is $N/2$. (Half of the alleles are "right" in the beginning, after that one allele gets fixed at each step.)

Neighbouring genotypes α, α' are always highly correlated, as necessarily $|f(\alpha) - f(\alpha')| \leq 1/N$.

$1 \leq K < N-1$**:**

For $K = 1$, a global optimum can still be found in polynomial time. For $K \geq 2$, global optimisation is NP-complete. However, for adjacent affecting loci $(i \curvearrowleft i+1,\ldots,i+K)$, the problem can be solved in time $\mathcal{O}(2^K N)$ (Weinberger).

$K = N-1$**:**

Neighbouring genotypes are totally uncorrelated.

$\Rightarrow$ Probability that a given genotype α is a local optimum is equal to the probability that α has the highest rank within its 1-mutant neighbourhood. This probability is equal to $1/(N+1)$.

$\Rightarrow$ The expected number of local optima is $2^N/(N+1)$.

The expected number of improvement steps for 1-mutant hill-climbing to hit a local optimum is proportional to $\log_2 N$ (each improvement step typically halves the rank of the genotype within the neighbourhood).

The expected waiting time for finding an improvement step is proportional to N.

7 Random Graphs

7.1 The Erdős-Rényi Model(s)

Two closely related "uniform" random graph models introduced in 1959 by P. Erdős & A. Rényi and E. N. Gilbert.

Consider the family $\mathcal{G}_n$ of all (labelled, undirected) graphs on n nodes. Denote $N = \binom{n}{2}$; then $|\mathcal{G}_n| = 2^N$.

Define the following two probability spaces

[Erdős & Rényi:] $\mathcal{G}(n,M)$ = all $G \in \mathcal{G}_n$ with exactly $M \leq N$ edges, taken with uniform probability, i.e.

$$\Pr(G_M = H) = \begin{cases} \binom{N}{M}^{-1}, & \text{if } H \text{ has } M \text{ edges} \\ 0; & \text{otherwise.} \end{cases}$$

[Gilbert:] $\mathcal{G}(n,p)$ = all $G \in \mathcal{G}_n$, taken so that each edge has occurrence probability p, $0 \leq p \leq 1$, independently of the other edges, i.e.

$$\Pr(G_p = H) = p^M \underbrace{(1-p)}_{q}^{N-M}, \text{ if } H \text{ has } M \text{ edges.}$$

These spaces are in a precise sense "close" if $M \sim pN$, and are often both referred to (unfairly to Gilbert) as the "Erdős-Rényi random graph model", or alternatively as the $\mathcal{G}(n,M)$ and $\mathcal{G}(n,p)$ random graph models.

Let $\Omega_n, n = 0,1,2,\ldots$ be a sequence of probability spaces of n-node graphs. Say that *almost every* (*a.e.*) graph in Ω_n has property Q if

$$\Pr(G \in \Omega_n \text{ has } Q) \to 1, \text{ as } n \to \infty.$$

Conversely, *almost no* graph in Ω_n has property Q if a.e. graph in Ω_n has property $\neg Q$, i.e.

$$\Pr(G \in \Omega_n \text{ has } Q) \to 0, \text{ as } n \to \infty.$$

Theorem 7.1 *Let H be a fixed graph and p a constant, $0 < p < 1$. Then a.e. $G \in \mathcal{G}(n,p)$ contains an induced copy of H.*

Remark: an "induced copy" means here a subset of nodes whose induced subgraphs is isomorphic to H.

Proof. Let $k = |H| =$ number of nodes in H. Then a graph G with $n = |G| \geq k$ nodes can be partitioned into $\lfloor n/k \rfloor$ disjoint sets of k nodes (with some left over). For each of these sets, the probability that it forms an induced copy of H is $r > 0$. (Precisely, $r = \frac{k!}{|\mathrm{Aut}(H)|} p^{e(H)} q^{\binom{k}{2} - e(H)}$.)

Thus, the probability that none of these sets forms an induced copy of H is

$$(1-r)^{\lfloor n/k \rfloor} \to 0, \text{ as } n \to \infty. \square$$

Let $k, l \in \mathbb{N}$. Say that a graph $G = (V,E)$ has property Q_{kl} if $\forall\, U, W, |U| \leq k, |W| \leq l, U \cap W = \varnothing$, G contains a node $v \in V$ such that v is adjacent to all $u \in U$ and no $w \in W$ (cf. Figure 6).

Lemma 7.2 *For every constant p, $0 < p < 1$, and all $k, l \in \mathbb{N}$, a.e. $G \in \mathcal{G}(n,p)$ has property Q_{kl}.*

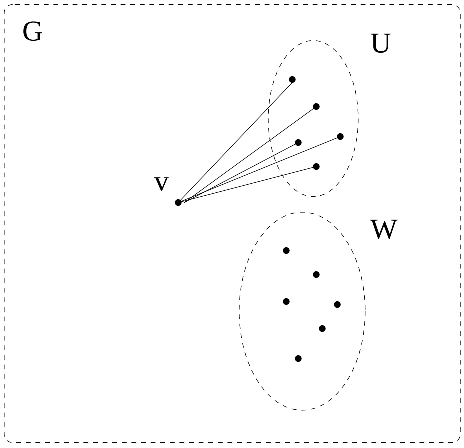


Figure 6: Property Q_{kl}.

Proof. For a fixed $U, W, v \in V \setminus (U \cup W)$, the probability that the condition is satisfied is

$$p^{|U|}q^{|W|} \geq p^k q^l$$

The events are independent for different v, so the probability that no appropriate v exists is

$$\left(1 - p^{|U|}q^{|W|}\right)^{n-|U|-|W|} \leq \left(1 - p^k q^l\right)^{n-k-l}.$$

There are at most n^{k+l} (U,W)-pairs to be considered, so the probability that some pair has no good v is bounded by

$$n^{k+l}(\underbrace{1 - p^k q^l}_{<1})^{n-k-l} \rightarrow 0, \text{ as } n \rightarrow \infty.$$

Thus in a.e. $G \in \mathcal{G}(n, p)$ all (U,W)-pairs have some appropriate v. □

Corollary 7.3 *Let p, $0 < p < 1$, be a constant. Then (i) a.e. $G \in \mathcal{G}(n, p)$ has minimum degree $\geq k$, for given constant k (ii) a.e. $G \in \mathcal{G}(n, p)$ has diameter 2 (iii) a.e. $G \in \mathcal{G}(n, p)$ is k-connected for given constant k.*

Proof. (i) and (ii) are immediate.

(iii) In a.e. $G \in \mathcal{G}(n, p)$, no two nodes u_1, u_2 can be separated by a cutset of size $k-1$, because we may choose in Lemma 7.2 $U = u_1, u_2$, $W = w_1, \ldots, w_{k-1}$ for arbitrary $w_1, \ldots, w_{k-1}$, and obtain a path u_1—v—u_2 connecting u_1, u_2 and avoiding $w_1, \ldots, w_{k-1}$. □

Corollary 7.4 *Let ϕ be any first-order sentence about graphs (i.e. quantification over nodes, relations $E(u,v)$ + identity). Then either $G \models \phi$ or $G \models \neg\phi$ for a.e. $G \in \mathcal{G}(n,p)$.*

Proof. Induction on the structure of ϕ, using Lemma 7.2 to eliminate quantifiers. □

Thus, all the first-order properties of $\mathcal{G}(n,p)$ for fixed p are easily captured. Things are more interesting when the number of nodes discussed and/or the probability p depends on n.

Given graph G, denote:

independence number $\alpha(G)$	=	size of the largest independent set in G,
clique number $\omega(G)$	=	size of the largest clique in G,
chromatic number $\chi(G)$	=	smallest number of colours needeed for colouring nodes in G so that no two adjacent nodes get the same colour.

Lemma 7.5 *Given $n \geq k \geq 2$, random $G \in \mathcal{G}(n,p)$:*

$$\Pr(\alpha(G) \geq k) \leq \binom{n}{k} q^{\binom{k}{2}}.$$

Proof. Probability that given k-set of nodes in G is independent is $q^{\binom{k}{2}}$. Total number of k-sets is $\binom{n}{k}$. □

Theorem 7.6 *Let $p, 0 < p < 1$ and $\varepsilon > 0$ be constant. Then for a.e. $G \in \mathcal{G}(n,p)$:*

$$\chi(G) > \frac{\ln(1/q)}{2+\varepsilon} \cdot \frac{n}{\ln(n)} = \Omega\left(\frac{n}{\ln(n)}\right) = \textit{large!}$$

Proof. By Lemma 7.5, for any fixed $n \geq k \geq 2$:

$$\begin{aligned}
\Pr(\alpha(G) \geq k) &\leq \binom{n}{k} q^{\binom{k}{2}} \leq n^k q^{\binom{k}{2}} \\
&= q^{k\frac{\ln n}{\ln q} + \frac{1}{2}k(k-1)} \\
&= q^{\frac{k}{2}[-\frac{2\ln(n)}{\ln(1/q)} + k - 1]} \\
&\to 0 \text{ for } k \text{ large,}
\end{aligned}$$

when

$$\frac{k}{2}\left[-\frac{2\ln(n)}{\ln(1/q)}+k-1\right] \to \infty \text{ for } k > (2+\varepsilon)\frac{\ln(n)}{\ln(1/q)}.$$

Thus there cannot be a.s. k nodes coloured with same colour for large k
$\Rightarrow$ More than $\frac{n}{k} = \frac{\ln 1/q}{2+\varepsilon} \cdot \frac{n}{\ln n}$ colours needed. □

Theorem 7.7 *Let p, $0 < p < 1$ be constant. Then for a.e. $G \in \mathcal{G}(n,p)$:*

$$\omega(G) \in \{d, d+1\},$$

where $d = d(n,p)$ is the largest integer such that

$$\binom{n}{d} p^{\binom{d}{2}} \geq \ln\, n.$$

(This implies $d = 2\log_{1/p}(n) + O(\log\log(n).)$ □

A *graph property* Q is an isomorphism-closed family of graphs, i.e. if $G \in Q$ (or "G has Q") and $G \approx G'$, then also $G' \in Q$.

A graph property is *monotone* if it is preserved under addition of edges, i.e. if $G = (V,E)$ and $G' = (V,E')$ are graphs such that $E \subseteq E'$ and G has Q, then also G' has Q.

A *threshold function* for the graph property Q is a function $t : \mathbb{N} \to \mathbb{R}$ such that

$$\Pr(G \in \mathcal{G}(n, p(n)) \text{ has } Q) \xrightarrow[n\to\infty]{} \begin{cases} 1, \text{ if } p \succ t \\ 0, \text{ if } p \prec t. \end{cases}$$

Notation:

$$p \succ t \Leftrightarrow \lim_{n\to\infty} \frac{p(n)}{t(n)} = \infty,$$
$$p \prec t \Leftrightarrow \lim_{n\to\infty} \frac{p(n)}{t(n)} = 0,$$
$$p \sim t \Leftrightarrow \lim_{n\to\infty} \frac{p(n)}{t(n)} = 1,$$
$$p \approx t \Leftrightarrow p(n) = \Theta(t(n)).$$

Denote: $P_n^Q(p) = \Pr(G \in \mathcal{G}(n,p) \text{ has } Q)$.

Then for monotone Q: $p_1 \leq p_2 \Rightarrow P_n^Q(p_1) \leq P_n^Q(p_2)\ \forall n$.

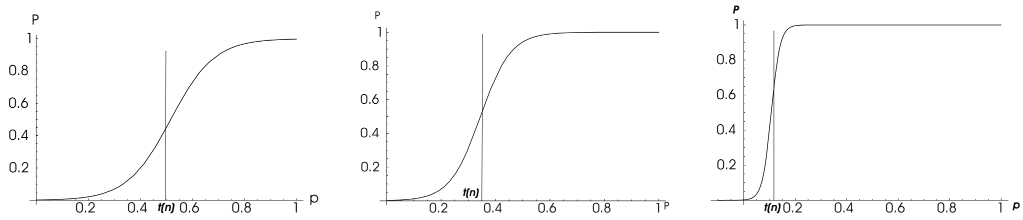


Figure 7: $P_n^Q(p)$ for (a) small, (b) intermediate and (c) large n.

Denote: $p_n^Q(\alpha)$ = the smallest p such that $P_n^Q(p) \geq \alpha$.

In fact $P_n^Q(p)$ is a continuous, strictly increasing function, so really $p_n^Q(\alpha) =$ unique p such that $P_n^Q(p) = \alpha$.

Figure 7 illustrates the evolution of the function P_n^Q, and a corresponding threshold function $t(n)$, for a monotone graph property Q from small to large values of n.

Lemma 7.8 *A function $t(n)$ is a threshold for monotone graph property Q if and only if*

$$t(n) \approx p_n^Q(\alpha)$$

for all $0 < \alpha < 1$.

Proof.
"⇒" Assume $t(n)$ is a threshold for Q. This means that if $p(n)/t(n) \to \infty$, then

$$P_n^Q(p(n)) \to 1 \qquad (*)$$

and if $p(n)/t(n) \to 0$, then

$$P_n^Q(p(n)) \to 0 \qquad (**)$$

Suppose then that

$$t(n) \not\approx p_n^Q(\alpha)$$

for some $0 < \alpha < 1$. This means that either there is a sequence $n_1, n_2, \ldots$ such that

$$p_{n_k}^Q(\alpha)/t(n_k) \to \infty,$$

contradicting (*), or there is a sequence $n_1, n_2, \ldots$ such that

$$p_{n_k}^Q(\alpha)/t(n_k) \to 0,$$

contradicting (**).

(Note that by definition, $P_n^Q(p_n^Q(\alpha)) = \alpha$!)

"$\Leftarrow$" Assume that $t(n)$ is *not* a threshold for Q. Then there is either a sequence $n_1, n_2, \ldots$ such that

$$p(n_k)/t(n_k) \to \infty,$$

but

$$P_n^Q(p(n_k)) \leq \alpha < 1$$

a.e., or a sequence $n_1, n_2, \ldots$ such that

$$p(n_k)/t(n_k) \to 0,$$

but

$$P_n^Q(p(n_k)) \geq \alpha > 0$$

a.e. In the former case,

$$t(n_k) \prec p(n_k) \leq p_{n_k}^Q(\alpha),$$

and in the latter case

$$t(n_k) \succ p(n_k) \geq p_{n_k}^Q(\alpha).$$

Thus in either case $t(n) \not\approx p_n^q(\alpha)$. $\square$

Theorem 7.9 *Every monotone graph property Q has a threshold function.*

Proof. For brevity, denote $p_n^Q(\alpha) = p(\alpha)$. Choose some arbitrary $0 < \alpha < \frac{1}{2}$. The goal is to prove that $p(\alpha) \approx p(1-\alpha)$, thus establishing e.g.

$$t(n) = p\left(\frac{1}{2}\right) = p_n^Q\left(\frac{1}{2}\right)$$

as a threshold function for Q. (Since $p(\alpha) \leq p(\frac{1}{2}) \leq p(1-\alpha)$.)

Let $m \in \mathbb{N}$ be such that $(1-\alpha)^m \leq \alpha$. Let $p = p_n(\alpha)$ and consider a sample of m independent graphs $G_1, \ldots, G_m$ from $\mathcal{G}(n,p)$. Then the graph $G_1 \cup \cdots \cup G_m \in \mathcal{G}(n,q)$, where $q = 1-(1-p)^m \leq mp$, and so

$$\Pr(G_1 \cup \cdots \cup G_m \text{ has } Q) \leq \Pr(G \in \mathcal{G}(n, mp_n(\alpha)) \text{ has } Q).$$

On the other hand, since Q is monotone, if any G_i has Q, then so does $G_1 \cup \cdots \cup G_m$. Thus,

$$\begin{aligned}\Pr(G_1 \cup \cdots \cup G_m \text{ does not have } Q) &\leq (1 - \Pr(G_i \text{ has } Q))^m \\ &= (1-\alpha)^m \leq \alpha.\end{aligned}$$

Hence,

$$\Pr_n^Q(mp_n(\alpha)) \geq \Pr(G_a \cup \cdots \cup G_m \text{ has } Q) \geq 1-\alpha,$$

and so

$$p_n(\alpha) \leq p_n(1-\alpha) \leq mp_n(\alpha),$$

i.e. $p(\alpha) \approx p(1-\alpha)$. (Since m depends only on α, not on n.) □

Consider a graph property Q defined as "G has Q" if $X(G) > 0$, where $X \geq 0$ is a random variable on $\mathcal{G}(n,p)$.

E.g. if $X(G)$ denotes the number of spanning trees of G, then property Q corresponds to connectedness.

A threshold function for property Q is a $t(n)$ such that

(i) $p(n) \prec t(n) \Rightarrow$ almost no $G \in G(n, p(n))$ has Q.
(ii) $p(n) \succ t(n) \Rightarrow$ almost all $G \in G(n, p(n))$ have Q.

If X is integral, then condition (i) can be verified by upper bounding $E[X]$; by Markov's inequality:

$$\Pr(X \geq 1) \leq E[X] \qquad \left(\begin{array}{l}\text{more generally, for } a > 0 \\ p(X \geq a) \leq E[X]/a\end{array}\right).$$

Condition (ii) is trickier, but can be approached by lower-bounding $E[X]$, and upper-bounding $\text{Var}[X]$. (So called "second-moment method".)

Denote $\mu = E[X]$, $\sigma^2 = \text{Var}[X] = E[(X-\mu)^2] = E[X^2] - \mu^2$.

Recall Chebyshev's inequality: for any $\lambda > 0$,

$$\Pr(|X-\mu| \geq \lambda) \leq \frac{\sigma^2}{\lambda^2}.$$

Lemma 7.10 *If $\mu > 0$ for n large, and $\frac{\sigma^2}{\mu^2} \to 0$ as $n \to \infty$, then $X(G) > 0$ for a.e. $G \in \mathcal{G}(n,p)$.*

Proof. If $X(G) = 0$, then $|X(G) - \mu| = \mu$. Hence

$$\Pr(X = 0) \le \Pr(|X - \mu| \ge \mu) \le \frac{\sigma^2}{\mu^2} \to 0 \text{ as } n \to \infty. \qquad \square$$

Denote the *density* of a graph G by $\delta(G) = \frac{e(G)}{|G|}$.

Say that a graph G is *balanced* if $\delta(G') \le \delta(G)$ for all subgraphs G' of G.

Theorem 7.11 *Let H be a balanced graph. Then the graph property "G has a subgraph isomorphic to H" has threshold function $n^{-1/\delta(H)}$.*

Proof. Let $X(G)$ =number of H-subgraphs of G. Let $k = |H|$, $l = e(H)$, so $\delta(H) = l/k$. Let us first bound $E[X]$ from above. Let $G \in \mathcal{G}(n,p)$, where $p = \gamma n^{-1/\delta(H)} = \gamma n^{-k/l}$ for some $\gamma = \gamma_n \to 0$, and denote

$$\mathcal{H} = \{\text{all copies of } H \text{ on vertex-set of } G\}.$$

Then $|\mathcal{H}| = \binom{n}{k} h \le \binom{n}{k} k! \le n^k$. Here h is the number of different arrangements of H on a set of k vertices, $h = k!/|\mathrm{Aut}(H)|$. Thus

$$\begin{aligned} E[X] &= \sum_{H' \subseteq \mathcal{H}} \Pr(H' \in G) = |\mathcal{H}| \cdot p^l \\ &\le n^k p^l = n^k (\gamma n^{-k/l})^l = \gamma^l \xrightarrow[\gamma \to 0]{} 0. \end{aligned}$$

Thus if

$$p(n) = \gamma_n \cdot n^{-1/\delta(H)} \prec n^{-1/\delta(H)},$$

then $E[X] \xrightarrow[n \to \infty]{} 0$.

By Markov's inequality this means that almost no $G \in \mathcal{G}(n,p)$ contains an H-subgraph for large n.

For the other part, we need to bound from above

$$\frac{\sigma^2}{\mu^2} = \frac{1}{\mu^2}(E[X^2] - \mu^2).$$

Let us try to compute:

$$\begin{aligned} E[X^2] &= \sum_{H', H'' \in \mathcal{H}} \Pr(H' \cup H'' \subseteq G) \\ &= \sum_{H', H'' \in \mathcal{H}} p^{e(H') + e(H'') - e(H' \cap H'')} \\ &\le \sum_{H', H'' \in \mathcal{H}} p^{2l - i\delta(H)}, \end{aligned}$$

where $i = |H' \cap H''|$. (Note that $\delta(H' \cap H'') \leq \delta(H)$.)
Denote then $\mathcal{H}_i^2 = \{(H', H'') \in \mathcal{H}^2 : |H' \cap H''| = i\}$ and compute separately for each i the sum

$$A_i = \sum_{\mathcal{H}_i^2} \Pr(H' \cup H'' \subseteq G)$$

Case $i = 0$*:*

$$\begin{aligned}
A_0 &= \sum_{\mathcal{H}_0^2} \Pr(H' \cup H'' \subseteq G) \\
&= \sum_{\mathcal{H}_0^2} \Pr(H' \subseteq G) \cdot \Pr(H'' \subseteq G) \qquad H', H'' \text{independent} \\
&\leq \sum_{\mathcal{H}^2} \Pr(H' \subseteq G) \cdot \Pr(H'' \subseteq G) \\
&= \left(\sum_{\mathcal{H}} \Pr(H' \subseteq G) \right)^2 \\
&= \mu^2
\end{aligned}$$

Case $i \geq 1$*:*

$$\begin{aligned}
A_i &= \sum_{\mathcal{H}_i^2} \Pr(H' \cup H'' \subseteq G) \\
&= \sum_{H' \in \mathcal{H}} \sum_{\substack{H'': \\ |H' \cap H''| = i}} \Pr(H' \cup H'' \subseteq G) \\
&\leq |\mathcal{H}| \cdot \binom{k}{i} \binom{n-k}{k-i} h p^{2l} p^{-il/k} \qquad\qquad h = \frac{k!}{|\mathrm{Aut}(H)|} \\
&\leq |\mathcal{H}| \cdot c_1 n^{k-i} h p^{2l} (\gamma n^{-k/l})^{-il/k} \\
&= \mu \cdot c_1 n^{k-i} h p^l \gamma^{-il/k} n^i \\
&= \mu \cdot c_1 n^k h p^l \gamma^{-il/k} \\
&= \mu c_2 \underbrace{\binom{n}{k} h}_{|\mathcal{H}|} p^l \gamma^{-il/k} \\
&= \mu^2 \cdot c_2 \gamma^{-il/k} \\
&\leq \mu^2 \cdot c_2 \gamma^{-l/k}.
\end{aligned}$$

Thus, denoting $c_3 = kc_2$, we get the estimate

$$\frac{E[X^2]}{\mu^2} = \left(\frac{A_0}{\mu^2} + \frac{\sum_i A_i}{\mu^2}\right) \leq 1 + c_3\gamma^{-l/k}$$

and hence

$$\frac{\sigma^2}{\mu^2} = \frac{E[X^2] - \mu^2}{\mu^2} \leq c_3\gamma^{-l/k} \xrightarrow[\gamma\to\infty]{} 0.$$

Hence, if $p(n) = \gamma_n n^{-k/l}$ for $\gamma_n \to \infty$, then by Lemma 7.10 $X(G) > 0$ holds for almost every $G \in \mathcal{G}(n,p)$ for large n.□

Corollary 7.12 *For $k \geq 3$, the property of containing a k-cycle has threshold $t(n) = n^{-1}$. (Note: independent of k)*□

Corollary 7.13 *For $k \geq 2$, the property of containing a specific tree structure T on k nodes has threshold function $t(n) = n^{-k/(k-1)}$.* □

Corollary 7.14 *For $k \geq 2$, the property of containing a k-clique ($\approx K_k$) has threshold function $t(n) = n^{-2/(k-1)}$.* □

Denote $\delta^*(H) = \max\{\delta(H') | H' \text{ is subgraph of } H\}$.

Theorem 4.11' *The graph property "G has a subgraph isomorphic to H" has threshold function $n^{-1/\delta^*(H)}$.* □

Threshold functions for global graph properties

Also known as "the phase transition".

The "epochs of evolution": Consider the structure of random graphs $G \in \mathcal{G}(n,p)$, as $p = p(n)$ increases. The following results can be shown (note that np = average node degree):

0. If $p \prec n^{-2}$, then a.e. G is empty.
1. If $n^{-2} \prec p \prec n^{-1}$, then a.e. G is a forest (a collection of trees).
   - The threshold for the apperarance of any k-node tree structure is $p = n^{-k/(k-1)}$.