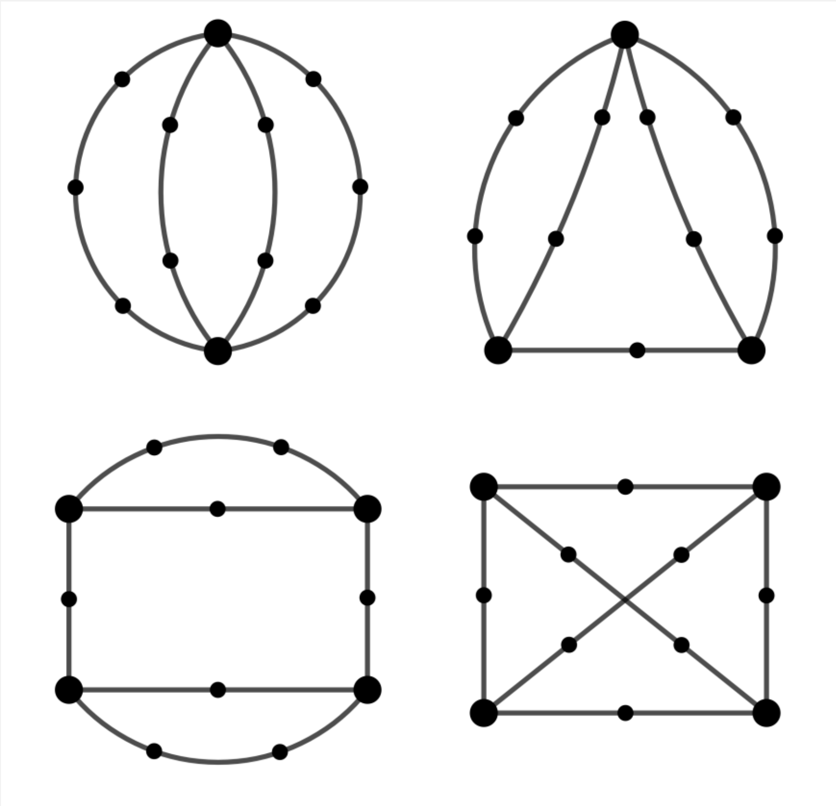
CHALMERS
UNIVERSITY OF TECHNOLOGY



Optimal Graphs for Connectedness under Random Edge Deletion

Master's thesis in mathematics

LORENTS LANDGREN

DEPARTMENT OF MATHEMATICAL SCIENCES

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2021
www.chalmers.se

Optimal Graphs for Connectedness under Random Edge Deletion

Lorents Landgren

CHALMERS
UNIVERSITY OF TECHNOLOGY

A master's thesis from the department of Mathematical Sciences,
Chalmers University of Technology,
Gothenburg, Sweden, 2021.

Optimal Graphs for Connectedness under Random Edge Deletion,
Lorents Landgren

Supervisor and examiner: Jeffrey Steif, professor at the dept. of Mathematical Sciences.

Master Theses 2021
Department of Mathematical Sciences
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Four types of graphs having two more edges than vertices. For a given number of vertices, a graph in one of these categories will prove to be "optimal".

Typeset in LaTeX.

Chalmers Reproservice
Gothenburg, Sweden, 2021

Optimal Graphs for Connectedness under Random Edge Deletion
Lorents Landgren
Mathematical Sciences,
Chalmers University of Technology

Abstract

Performing percolation on a finite graph G means independently keeping or discarding each edge according to a probability parameter p. The focus is the probability $P_c(G, p)$ that a percolation outcome turns out to be a connected graph. Specifically, if we fix p, the number of vertices n and the number of edges m, we try to find the graph(s) with the highest such probability. We call such graphs the most stable or the (n, m, p)-optimal graphs. It is shown that any (n, m, p)-optimal graph consists of a single so-called *block*.

For $m = n$, $m = n + 1$ and $m = n + 2$ respectively, we show the existence of a unique optimal graph, which is actually independent of p. However, in general, the relative stability of two (n, m)-graphs is p-dependent. We make some investigations into when this is the case.

Keywords: random graph theory, percolation on finite graphs, connected outcome, combinatorics and graph theory, Erdős–Rényi model, stability ordering, optimal graph

Acknowledgements

The undersigned author wishes to express gratitude to his supervisor, professor Jeffrey Steif, for availability and support which went far beyond the call of duty. The original idea belongs to him, and the presentation would not be of recognizable form without his feedback at every stage and level. The remaining mathematical and linguistic flaws are my own.

Lorents Landgren,
Mariestad, June 2021

Contents

1 Introduction **1**

2 Preparations **2**
2.1 Graph theory . . . 2
2.2 Percolation on finite graphs . . . 4

3 First cases: Trees and $m = n$ **7**

4 A few general results **8**
4.1 Counting connected graphs . . . 8
4.2 An (n, m, p)-optimal graph consists of one block . . . 9

5 Chains, multigraphs and $m = n + 1$ **13**
5.1 Chains and multigraphs . . . 13
5.2 The p-optimal $(n, n + 1)$-graph . . . 16

6 A study of the case $m = n + 2$ **19**
6.1 The four D-graph designs . . . 19
6.2 Calculating and maximizing c_2 . . . 21
6.3 Calculating c_3 for graphs of type D1, D2 and D4 . . . 23
6.4 A D1- or D2-graph does not maximize c_3 . . . 24
6.5 A D3-graph does not maximize c_3 . . . 26
6.6 Finding the p-optimal $D4_m$-graph . . . 27

7 The ordering depends upon p **32**
7.1 The relative stability of two graphs . . . 32
7.2 Handcuff graphs . . . 33
7.3 Graphs with branches . . . 34
7.4 Optimality for large and small p . . . 36
7.5 Ordering the D2-graphs . . . 37

1 Introduction

Graph theory is a fundamental area of mathematics. Its study lays the groundwork for network modeling, with applications across the board of sciences. Percolation theory concerns the random removal of edges (or vertices) from graphs, often infinite grids. When applied to the complete graph on n vertices, one has the famous $G(n, p)$-model introduced by Edgar Gilbert, known as the Erdős–Rényi model. Random graph theory and percolation theory are active areas of mathematical research.

In spite of all the activity in these areas, not much research has been done on percolation on general finite graphs. Some investigations have been made into the critical probability for the emergence of a giant component [1]. In this paper we are concerned simply about the probability that a graph is connected after performing percolation on it. Among other things, we find graphs that maximize this probability (for given numbers of vertices and edges and other constraints). As far as we know, this has not been studied elsewhere.

2 | Preparations

2.1 Graph theory

The content of this section belongs to the common goods of graph theory. Much practical notation has been borrowed from Diestel [2], here and throughout. Diestel has also been followed in cases where there are conflicting definitions of a concept in current use.

Formally, a graph is an ordered pair $G = (V, E)$, where V is the vertex set and E is the edge set, which contains a collection of unordered pairs of vertices. The vertex set of G can also be written as $V(G)$ and the edge set as $E(G)$. We generally consider graphs to be *unlabeled*, that is, we identify graphs that are isomorphic. (However, an exception is introduced in Section 2.2.)

We will allow ourselves a few practical shorthands when there is no risk of confusion. For example, an edge $\{u, v\}$ is usually written uv. If e is an edge and u is a vertex of G, we will often write $e \in G$ and $u \in G$ instead of the formally correct $e \in E(G)$ and $u \in V(G)$.

The letter n will generally be used for the number of vertices, or *order*, of G, and m for the number of edges, or *size*, of G. Let us define, as shorthands, $|G| = n$ for the order of G, and $\|G\| = m$ for the size of G.

A graph of order $n \geq 1$ is called an n-graph. A graph of order n and size m is called an (n, m)-graph. More generally, if $\mathcal{K}$ is a collection of graphs, a graph in $\mathcal{K}$ is called a $\mathcal{K}$-graph.

Take an edge $e = uv$; we then say that e is *incident* to u and v, which are themselves incident to e. The vertices u and v are said to be *adjacent* or *neighbors*; likewise two edges are adjacent if they have a vertex in common.

Definition 1 (Vertex degree)**.** The degree of a vertex v, denoted $\deg(v)$, is the number of edges incident to v. A vertex of degree k is a k-vertex.

Definition 2 (Degree sequence)**.** The degree sequence of G is the (weakly) decreasing sequence of vertex degrees of G.

The following well known formula is proven by counting the number of vertex-edge incidences in two different ways.

Lemma 1 (Degree sum formula)**.** *The sum of the vertex degrees of $G = (V, E)$ is twice* *the number of edges:*

$$\sum_{v \in V} \deg(v) = 2|E|\,. \tag{2.1}$$

Definition 3 (Path)**.** A path Π of a graph G is a sequence of *distinct* vertices in G, in which every vertex is adjacent to the next, and the sequence of edges effecting these adjacencies. The length of a path is its number of edges.

A path can be specified by its vertex sequence. We will occasionally use the following notation: If Π is the path $x_0x_1 \ldots x_k$ in G, and $0 \leq i \leq k$, then Πx_i is the path $x_0 \ldots x_i$. A union denotes concatenation of paths; for example, $\Pi x_i \cup x_i x'$ is the path $x_0 \ldots x_i x'$, and more generally, $\Pi_1 \cup \Pi_2$ is the path Π_1 followed by Π_2 (which presupposes that the endvertex of Π_1 matches the starting point of Π_2).

Definition 4 (Cycle)**.** A cycle C of a graph is a sequence of at least four (three distinct) sequentially adjacent vertices, and the sequence of edges effecting these adjacencies. Except for the first and the last vertex being identical, the vertices are distinct.

Cycles with the same vertex sets are identified, so that the specific starting point and the orientation does not matter. The length of a cycle is its number of edges. A cycle of length k is called a k-cycle. A graph without a cycle is called *acyclic*.

Many graphs have their own names and notation. Some useful special graphs are defined as follows. A graph which is itself a path is a *path graph*, denoted P^k, where k is the number of edges. A k-cycle, taken by itself, is the *cycle graph* C^k. An n-graph which contains every possible edge is the *complete graph* on n vertices, denoted K^n. The complete graph on n vertices has $\binom{n}{2}$ edges.

Definition 5 (Connectedness)**.** A graph G is connected if, for every pair of vertices u and v in G, there exists a path from u to v; otherwise it is disconnected.

There are several equivalent definitions that characterize a (finite) *tree*. We will have use of these three:

Definition 6a (Tree)**.** A connected acyclic graph is called a tree.

Definition 6b. A tree is a connected graph that becomes disconnected if any one of its edges is removed.

Definition 6c. A tree is a connected n-graph with $n - 1$ edges, for some $n \geq 1$.

Definition 7 (Leaf)**.** A leaf is a vertex of degree 1.

Lemma 2. *If G is an n-graph with $m < n - 1$ edges, then G is disconnected.*

Proof. Let G be an n-graph with $m < n - 1$ edges and suppose that G is connected. Add $n - 1 - m$ edges, and we have a tree. Remove the last added edge, and the graph disconnects. This is absurd, so G must be disconnected. □

Definition 8 (Spanning tree)**.** T spans G if T is a tree which is a subgraph of G, and $|T| = |G|$.

As is well known, every connected graph has a spanning tree. *Proof:* Let G be a connected graph. If the graph is not a tree, pick a cycle and remove one edge from it. Repeat until you have a tree, which spans G.

Lemma 3. *If G is an (n, m)-graph, it is possible to remove $m - n + 1$ edges from G without disconnecting the graph, and this number is maximal.*

Proof. Let G be a connected (n, m)-graph, and let T be a spanning tree. Since $\|T\| = n-1$, we have $\|G\| - \|T\| = m - n + 1$. It is therefore possible to remove $m - n + 1$ edges from G without disconnecting G. Since by Lemma 3 no n-graph with less than $n - 1$ edges is connected, this number cannot be increased. □

We will generally assume that G stands for a connected graph with at least two vertices. Since G is assumed connected, and since no n-graph can have more edges than K^n, the number of edges m of G is in the integer interval $[n - 1 \mathrel{..} \binom{n}{2}]$.

Definition 9 (Subgraph, $G' \subseteq G$)**.** A graph $G' = (V', E')$ is a subgraph of $G = (V, E)$ if $V' \subseteq V$ and $E' \subseteq E$. This is written $G' \subseteq G$.

Definition 10 (Induced subgraph)**.** If $G' \subseteq G$ and E' contains every edge $xy \in E$ such that $x, y \in V'$, then G' is an induced subgraph of G.

By $G - e$ we mean the subgraph of G in which the edge e is removed; formally, $G - e = (V(G), E(G) \setminus \{e\})$. A little more care is needed to define $G - v$. Vertex removal is defined as follows:

Definition 11 (Vertex deletion)**.** When a vertex v is *removed* or *deleted* from a graph G, any edges incident to v are also removed, so that deleting a vertex yields an induced subgraph of G, denoted $G - v$.

If X and Y are subsets of $V(G)$, then $E(X, Y)$ is the set of edges $xy \in G$ such that $x \in X$ and $y \in Y$.

Definition 12 (Cut)**.** If $\{V_1, V_2\}$ is a non-trivial partition of the vertex set of G, then the set of edges $E(V_1, V_2)$ is called a *cut* in G.

2.2 Percolation on finite graphs

Exploring the concept of percolation on finite graphs, we need a probability parameter p. We will also use $q \coloneqq 1 - p$. To avoid trivialities, we will always assume p to be in the open interval $(0, 1)$.

Definition 13 (Percolation, G_p)**.** Fix $p \in (0, 1)$. Performing *percolation* on a graph G means independently keeping or removing each edge of G with probability p and $1 - p$, respectively. A subgraph of G thus obtained is called a percolation realization of G, denoted by G_p.

When performing percolation on a graph G, the possible outcomes should be considered labeled graphs, so that isomorphic outcomes are not considered "the same". For example, the unique $(3,2)$-graph P^2 has four possible percolation realizations (two of which are isomorphic), as the reader should verify.

Definition 14 (Open/closed edges and paths)**.** Suppose G_p is a percolation realization of G. The edges retained in the percolation process will be referred to as *open* in G_p and the ones discarded as *closed* in G_p. A path in G is open in G_p if all its edges are open; otherwise it is closed.

Definition 15 (Outcome sets, $\mathcal{G}$, $\mathcal{G}_i$)**.** The set of percolation realizations of G is denoted $\mathcal{G}$. The subset of $\mathcal{G}$ where exactly i of the edges of G are closed is denoted $\mathcal{G}_i$. G_i will typically denote an element of $\mathcal{G}_i$.

Suppose G and G' are two (n,m)-graphs, where G' has been obtained by a relatively simple modification of G, for example by somehow "moving" an edge. There is then an obvious bijection between the vertices and edges of G and those of G'. The bijection induces a natural one-one correspondence between the outcome sets $\mathcal{G}$ and $\mathcal{G}'$, which will be useful later on.

Performing percolation on G gives rise to an obvious probability space $(2^E, \mathcal{P}(2^E), P_p)$. We are interested in the probability that a percolation realization is a connected graph, and define $P_c(G,p) := P_p\{\omega \in 2^E : (V,\omega) \text{ is connected}\}$. (The subscript c here stands for "connected".) The probability that an outcome is disconnected is denoted P_d, and obviously $P_d(G,p) = 1 - P_c(G,p)$.

For fixed n, m and p, the association of every (n,m)-graph G with the number $P_c(G,p)$ induces a *weak ordering* on the (n,m)-graphs (meaning an ordering where ties are allowed). An interesting question is whether, or how and when, this ordering depends upon p. We will use the following terminology when comparing the relative placement of graphs in the ordering:

Definition 16 (Stability ordering, $\prec_p$)**.** Fix n, m and p and consider only (n,m)-graphs. Define the weak ordering $\preceq_p$ on these graphs by $G \preceq_p G' :\Leftrightarrow P_c(G,p) \le P_c(G',p)$, and define the associated strict weak ordering $\prec_p$ similarly. ($G \prec_p G'$ is read "G strictly precedes G' with respect to p" and $G' \succ_p G$ is read "G' strictly succeeds G with respect to p".)

We call $\preceq_p$ the *stability ordering*, and normally use the following more suggestive terminology instead of "precedes" and "succeeds":

- If $G \succ_p G'$, then G is strictly more stable than G' with respect to p.
- If $G \succ_p G'$ for all p, then G is strictly more stable than G'.

We are especially interested in graphs which, for some p, are maximal (have no successors) in their stability ordering; in other words, the graphs which maximize $P_c(G,p)$ for given n, m and p.

Definition 17 (Optimal graphs)**.** Fix n, m and p and consider only (n, m)-graphs. Let $\mathcal{K}$ be a subset of these graphs.

- G is an (n, m, p)-optimal graph if $G \succeq_p G'$ for every G'.
- $G \in \mathcal{K}$ is a p-optimal $\mathcal{K}$-graph if $G \succeq_p G'$ for every $G' \in \mathcal{K}$.

For every n, m and p, at least one (n, m, p)-optimal graph clearly exists. Pertinent questions, for given n, m and p, are:

- How can an (n, m, p)-optimal graph be characterized?
- Is there one unique (n, m, p)-optimal graph, or several?
- Does an (n, m, p)-optimal graph depend upon p for its optimality?

3 | First cases: Trees and $m = n$

Consider the case $m = n - 1$. If G is a connected $(n, n-1)$-graph, G is by definition a tree, and needs all of its edges to stay connected. Therefore $P_c(G, p) = p^m$, independently of G. Any n-tree is therefore $(n, n-1, p)$-optimal for all p. We note that for every $n \geq 4$ there are several nonisomorphic trees, so that the $(n, n-1, p)$-optimal graphs are not unique.

From this trivial case, we move on to $m = n$; $n \geq 3$. A connected graph G with as many edges as vertices has one edge too many to be a tree. Since a tree is a connected, acyclic graph, G has at least one cycle; in fact, a moment of thought makes clear that G has exactly one cycle, which we call C. (Proof: Pick a cycle and remove one edge, say uv. Since there is a path from u to v, the graph is still connected. Since it has $n - 1$ edges, it is a tree and therefore acyclic. The unique path from u to v, together with the removed edge uv, forms the only cycle, C.)

Let $\ell_1 \leq m$ be the number of edges in C. A percolation outcome G_p is connected if all edges remain or if exactly one of the edges in C is closed; otherwise G_p is disconnected. Since there are ℓ_1 different ways to remove one of the edges of C, and each of these outcomes has probability $p^{m-1}q$, the probability of connectedness is

$$P_c(G, p) = p^m + \ell_1 p^{m-1} q \,. \tag{3.1}$$

For any fixed p, $P_c(G, p)$ is determined only by the length of C, so that the (n, n)-graphs are ordered according to the lengths of their cycles. The stability ordering is therefore independent of p. The probability $P_c(G, p)$ is strictly maximized when $\ell_1 = n$, i.e. when G is the cycle C^n, which is therefore the unique (n, n, p)-optimal graph for all p.

By a 2^{nd}-(n, m, p)-optimal graph we mean a graph G which, for given p, has an immediate successor that is (n, m, p)-optimal. Similarly a 3^{rd}-(n, m, p)-optimal graph has an immediate successor that is 2^{nd}-(n, m, p)-optimal. Now, in the $n = m$ case, for $n \geq 4$, the graph C^{n-1} with a leaf attached is uniquely 2^{nd}-(n, n, p)-optimal for all p. For $n \geq 5$, any graph containing C^{n-2} is 3^{rd}-(n, n, p)-optimal for all p. An (n, n)-graph with such a cycle has either two leafs or a single leaf two edges away from C^{n-2}. We note that the 3^{rd}-(n, n, p)-optimal graphs are therefore not unique.

4 | A few general results

4.1 Counting connected graphs

Performing percolation on any connected (n, m)-graph G, an outcome with i edges removed is obtained with probability $p^{m-i}q^i$. There are $\binom{m}{i}$ such outcomes; say that c_i of these outcomes are connected. In principle, then, $P_c(G, p)$ can be calculated by counting connected outcomes and summing up their probabilities. By Lemma 3, the largest i for which c_i is nonzero is $i = m - (n - 1)$.

$$P_c(G, p) = \sum_{i=0}^{m-n+1} c_i \, p^{m-i} q^i \, . \tag{4.1}$$

We might also want to make a direct calculation of the probability that G_p is *disconnected*. Let d_i be the number of disconnected outcomes with i edges removed. Then

$$P_d(G, p) = \sum_{i=1}^{m} d_i \, p^{m-i} q^i \, . \tag{4.2}$$

Remark. Since $c_i + d_i = \binom{m}{i}$, the binomial expansion of $1 = (p + q)^m$ confirms that $P_c + P_d = 1$.

To emphasize that the coefficients c_i and d_i vary for different (n, m)-graphs G we will sometimes write $c_i(G)$ and $d_i(G)$. When comparing two graphs G and G', we will often write simply c_3 for $c_3(G)$ and c'_3 for $c_3(G')$.

The following lemma follows almost immediately from the definitions above.

Lemma 4. *If G and G' are (n, m)-graphs, then the following condition on the coefficients c_i (or d_i) and c'_i (or d'_i) is sufficient for G' to be strictly more stable than G (for any p): For each $i \in [0, m - n + 1]$, $c'_i \geq c_i$ (or equivalently $d'_i \leq d_i$), with a strict inequality for at least one i.*

Proof. Let G and G' be two (n, m)-graphs, and suppose that the condition of the lemma is true. Inserting the coefficients into (4.1) gives $P_c(G', p) > P_c(G, p)$ for any p, which means by definition that G' is strictly more stable than G. □

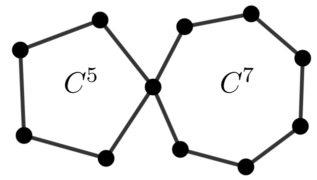


Figure 4.1: A central cutvertex is common to the two blocks C^5 and C^7.

4.2 An (n, m, p)-optimal graph consists of one block

In this section we disregard the trivial case of G being a tree, and thus assume that $m \geq n$. We introduce three useful concepts, and arbitrarily fix $p \in (0, 1)$, which makes the subsequent results valid for any p in this interval.

Definition 18 (Bridge)**.** A *bridge* is an edge that, if removed, separates a graph into two components.

It is not hard to see that the bridges of a graph are exactly those edges that are not part of a cycle. We give two equivalent definitions of the related concept of *cutvertex*.

Definition 19a (Cutvertex)**.** A *cutvertex* is a vertex that, if removed, separates a graph into more than one component.

Definition 19b. A vertex $u \in G$ is a cutvertex if there exist other vertices v and w in G such that every vw-path contains u.

Example. The central vertex in Figure 4.1 is a cutvertex.

A *block* can loosely be described as what you get if you start with a non-cutvertex and add every piece that is not "beyond" a cutvertex. More precisely, we have:

Definition 20 (Block)**.** A block is a maximal connected (induced) subgraph that has no cutvertex of its own.

Example. The upper graph of Figure 5.1 has six blocks: the cycle starting at A, the 4-cycle (with a chord) starting at B, and the four paths of length one between A and B.

See [2, p. 60–62] for further discussion of blocks.

Proposition 1. *An (n, m, p)-optimal graph has no bridges.*

Proof. Let G be a connected graph with a bridge. G contains a cycle by the assumption that $m \geq n$. Pick two adjacent edges $e_b = uv$ and $e_c = vw$, such that e_b is a bridge and e_c is in a cycle. This can be done since every edge either is a bridge or belongs to a cycle.

We construct a modified graph G' from G by removing the edge e_c and adding an edge $e'_c = uw$, as illustrated in Figure 4.2. (The illustrated graph is not completely general: there could be more vertices and edges on the right hand side.) We will presently

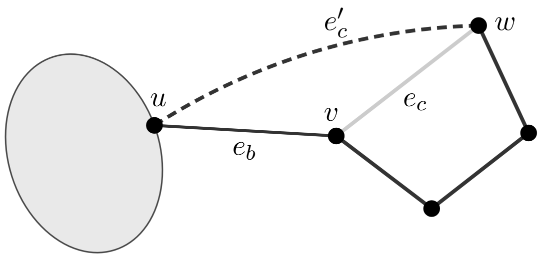


Figure 4.2: When one of the incidences of e_c is changed from vertex v to u, the edge e_b ceases to be a bridge; the result is a strictly more stable graph. (The gray area represents an unknown part of the graph. There could also be other vertices and edges strictly to the right of the symbol e_b.)

show that none of the coefficients c_i in (4.1) have decreased, and that in fact at least one (namely c_1) has strictly increased, which makes G' strictly more stable than G according to Lemma 4.

Fix $i \in [1\,..\,m]$ and let $\mathcal{G}_i$ and $\mathcal{G}'_i$ be outcome sets according to Definition 15. The natural one-to-one correspondence between the edges of G and G' induces a bijection between $\mathcal{G}_i$ and $\mathcal{G}'_i$. Given $G_i \in \mathcal{G}_i$, let the corresponding graph in $\mathcal{G}'_i$ be denoted by G'_i. Suppose that G_i is connected. In particular this means that $e_b \in G_i$: otherwise there would be no open path from u to v.

Pick an arbitrary vertex x and an open path Π in G_i from x to v. If $e_c \notin \Pi$, the same path is open in G'_i. If $e_c \in \Pi$, adapt the path to G'_i by exchanging $e_c = wv$ for both e'_c and e_b in sequence (wuv, vertex-wise). Since the result is an open path from x to v in G'_i, and x was arbitrary, G'_i is connected. This means that the connected graphs in $\mathcal{G}'_i$ are at least as many as in $\mathcal{G}_i$, so that $c'_i \geq c_i$ in (4.1) for all i.

Now consider the particular $G_1 \in \mathcal{G}_1$ where exactly the edge $e_b = uv$ has been closed. The graph is disconnected, since e_b is a bridge of G. However, in G'_1 there is a path connecting u and v: Consider that $e_c = vw$ is part of a cycle in G, so that there is an alternate path between v and w. This path together with the edge $e'_c = wu$ forms an open path in G'_1 from v to u, which makes up for the closed e_b. Therefore, G'_1 is connected, and we can conclude that $c'_1 \geq c_1 + 1$. Now an application of Lemma 4 gives that G' is strictly more stable than G, so that G is not (n, m, p)-optimal. ☐

Remark. Since the edge incident to a leaf is a bridge, the preceding Proposition 1 implies that an (n, m, p)-optimal graph has no leaves. Any vertex incident to a bridge is either a leaf or a cutvertex, but there can be a cutvertex without there being a bridge, as shown in Figure 4.1. Therefore, the following proposition strengthens Proposition 1.

Proposition 2. *An (n, m, p)-optimal graph has no cutvertices.*

Proof. Let G be a connected graph with a cutvertex. If G has a bridge, use Proposition 1 to conclude that G is not (n, m, p)-optimal.

Otherwise, pick a cutvertex u. Choose an edge incident to u and call it $e_1 = uv$. Since e_1 is part of a cycle C, there is another edge e_2 incident to u, which ends the cycle. Since u is a cutvertex, it has another incident edge, $e_3 = uw$, such that every path from w to v passes u. Since e_3 is part of a cycle (which cannot include e_1 or e_2), u must have a fourth incident edge, e_4, which ends the cycle. Note that e_1 and e_3 belong to different blocks, which we call block A and block B, so that $v \in A$, $w \in B$ and $A \cap B = u$. Figure 4.3 is a minimalistic illustration of this; there could be any number of other blocks containing u, blocks sharing some other vertex with A or with B (but not both), and other blocks further away, depending on the number of cutvertices of G.

Let us construct a graph G' in a way analogous to the construction of G' in Proposition 1: We remove the edge e_1 from G, and add the edge $e_1' = vw$, as illustrated in Figure 4.3.

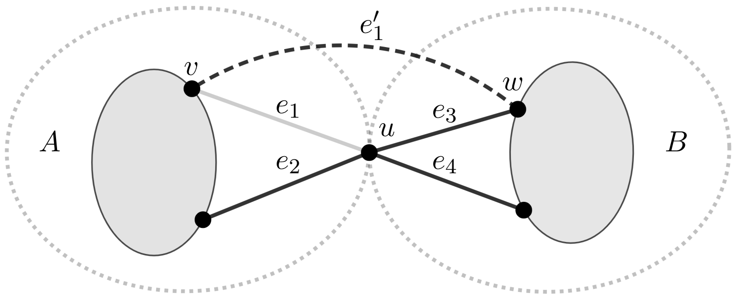


Figure 4.3: When the edge e_1 is moved from the vertex u to vertex w, the blocks A and B become a single block, and the resulting graph is strictly more stable.

Fix $i \in [1\,..\,m]$, let $\mathcal{G}_i$ and $\mathcal{G}_i'$ be outcome sets according to Definition 15, and let $G_i \in \mathcal{G}_i$ and $G_i' \in \mathcal{G}_i'$ be corresponding percolation outcomes. Suppose that G_i is connected. From an arbitrary vertex x there is an open path $\Pi \in G_i$ to u. If $e_1 \notin \Pi$ then Π is open in G_i' also. If $e_1 \in \Pi$ (which implies that Π goes through A) then in G_i' we have the open path $\Pi' = \Pi v \cup vw$ from x to w. We wish to concatenate Π' with an open path in G_i' from w to u. Such a path exists in G_i, since G_i is connected, and it has to be contained in block B. Therefore, the same wu-path is open in G_i'. The open path $\Pi' \cup w \ldots u$ from x to u in G_i' shows that G_i' is connected. Thus, $c_i' \geq c_i$, for any i.

Consider the particular subgraph of G where exactly e_3, e_4 and any other edge incident to u and contained in B has been removed. If there are j such edges, this is a graph $G_j \in \mathcal{G}_j$. Since there is no open path in G_j from any vertex in $B - u$ to u, G_j is disconnected. However, in G_j', there there is a path from any vertex in $B - u$ to u, which means that G_j' is connected.

To see this, let x be a vertex in $B - u$. Choose a path from x to $w \in B$ (by the definition of block, $B - u$ is connected). Then, from w, use the edge e_1' to $v \in A$, and at last follow the path $C - e_1$ from v to u.

Since G_j' is connected even though G_j is not, $c_j' \geq c_j + 1$. According to Lemma 4, G' is strictly more stable than G, so G' is not (n, m, p)-optimal. $\square$

Corollary 1. *An (n, m, p)-optimal graph consists of a single block.*

Proof. A connected graph without a cutvertex is a single block by definition. □

Given a graph G, one can iterate the procedures of Proposition 1 and Proposition 2, each time producing a more stable graph. Since the number of (n, m)-graphs is finite, the process terminates, and one has a single-block (n, m)-graph which we can call B.

However, for the given G, the terminating graph B is not uniquely determined, and it is not true in general that it will be (n, m, p)-optimal. For example, consider the graph of two cycles sharing a vertex in Figure 4.1; a bridgeless $(11, 12)$-graph with a 5-cycle, a 7-cycle and a single cutvertex. With this graph as G in the proof of Proposition 2, an edge e_1 should be "loosened" from the cutvertex and attached to one of its neighbors in the other cycle. The resulting graph will have two chordless cycles sharing an edge, but the length of the cycles depends on which edge was chosen as e_1. By the end of Section 5.2, it will become clear that one of these is strictly more stable than the other for all p, but that they are both distinct from the unique $(11, 12)$-optimal graph (which exists, and will turn out to look very similar to the graph in Figure 5.2).

5 | Chains, multigraphs and $m = n + 1$

5.1 Chains and multigraphs

In this section we assume that $m > n$. For ordinary graphs (but not for multigraphs) this implies that n is at least 4. It also implies that there is at least one vertex of degree more than 2, by the degree sum formula.

Definition 21 (Multigraph)**.** A multigraph is a generalized graph, which allows for multiple edges between any pair of vertices, and also for a vertex to have *loops*. A loop is an edge incident to the same vertex in both ends, and increases the degree sum of the vertex by 2. A k-multigraph is a multigraph with k vertices.

To study the possible ways to construct an (n, m)-graph, we can consider its possible degree sequences. Since we are interested in finding optimal, connected graphs, we rule out any degree sequence containing a 0 or a 1, that is, we consider only connected graphs without leaves (in accordance with Proposition 1 and the subsequent remark).

Consider the $(n, n + 1)$-graphs. The degree sum is $2n + 2$, and there are two possible degree sequences:

$$\begin{aligned} 1\colon\ &(4, 2, \dots, 2) \\ 2\colon\ &(3, 3, 2, \dots, 2) \end{aligned} \tag{5.1}$$

To refer to a graph by its degree sequence, we introduce the following notation:

Definition 22 ($\langle d_1, d_2, \dots, d_n\rangle$-graph)**.** A graph or multigraph with degree sequence $(d_1, d_2, \dots, d_n)$ is referred to as a $\langle d_1, d_2, \dots, d_n\rangle$-graph or -multigraph. For repeated degrees, the number of repetitions are put in preceding superscript, so that an $\langle {}^2a, {}^3b\rangle$-graph means an $\langle a, a, b, b, b\rangle$-graph.

Starting with a degree sequence, and considering how the possible graphs are structured, we can choose to regard a vertex of degree 2 as "only a stop on the way" between "junctions" (higher degree vertices). The 2-vertices do not from this perspective add complexity to the graph. We want to temporarily disregard the 2-vertices by *suppressing* them, which in general yields a multigraph. One can imagine suppression of a 2-vertex as letting the vertex melt into its two incident edges, making them one.

Definition 23 (Suppression)**.** Suppressing a 2-vertex means deleting it and adding an edge between its neighbors. This might yield a multigraph.

Suppose that $v_1 \dots v_{\ell-1}$ is a maximal path of 2-vertices in an ordinary graph. For reasons to become apparent, we want to include the edges incident to the path at its ends and call the alternating sequence $e_1 v_1 e_2 \dots v_{\ell-1} e_\ell$ a *chain*.

Definition 24 (Chain)**.** A maximal alternating sequence of distinct edges and vertices in G, where every object is incident to the next one, and in which every vertex has degree 2, is called a *chain* of G. The *length* of a chain is its number of edges. An ℓ-chain is a chain of length ℓ.

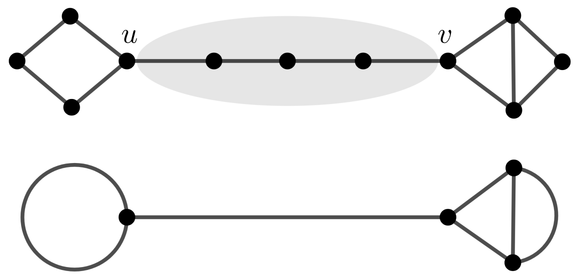


Figure 5.1: In the first graph, the part enclosed in light gray is a 4-chain adjacent to the vertices u and v. The graph below shows the associated multigraph, formed by collapsing every chain into an edge.

In the first graph of Figure 5.1, a chain of length four is highlighted. A few remarks are in order. (1) A chain need not contain a vertex - a 1-chain is an edge connecting two vertices of degree 3 or higher. (2) A chain is *not* a subgraph of G, since it has "loose ends". (3) We will often pragmatically use the same symbol (a subscripted ℓ) for both the length of a chain and the chain itself.

Definition 25 (Chain length sequence)**.** Referring to a graph G (or possibly a collection of chains in G), the chain length sequence is the monotonically (weakly) decreasing sequence of chain lengths.

Due to the maximality, every 2-vertex belongs to one and only one chain. It should also be clear that suppressing all the vertices on a chain, in any order, yields a single edge, incident to higher degree vertices. Therefore, the order of suppression does not matter, and the suppression of all 2-vertices in a graph G yields a well-defined result, which we call the *associated multigraph*.

Definition 26 (Associated multigraph)**.** The multigraph obtained by suppressing all 2-vertices of a graph G is called the multigraph associated with G, denoted by G_M.

In effect, the multigraph associated with G is obtained by exchanging every chain for an edge (as shown in Figure 5.1) and conversely, every edge in G_M corresponds to a chain in G. If the two ends of a chain are incident to the same vertex, the corresponding edge in G_M is a loop. Since suppressing a vertex does not change the degree of any of the other vertices, every vertex remaining in G_M has the same degree as it had in G. The reader can verify an example of this by comparing the vertex degrees of the graph and associated multigraph in Figure 5.1.

Definition 27 (Chain incidence and adjacency)**.** A chain and a vertex of degree ≥ 3 are said to be incident in G if the corresponding edge and vertex are incident in G_M. Likewise two chains are adjacent if the corresponding edges are adjacent in G_M.

Since the chains yield a partition of the edges of G, the sum of their lengths (the *chain sum*, for short) equals m. Fix m and fix a multigraph G_M. The coefficients $c_i(G)$ and $d_i(G)$ (where we let G vary over the possible graphs, i.e. graphs associated to G_M and with size m) will often attain extreme values when some or all chains are *as equal as possible* in length. We therefore define

Definition 28 ($\simeq$, AEAP, as equal as possible)**.** A collection of chains, $(\ell_1, \ell_2, \ldots \ell_k)$, whose sum is fixed, are *as equal as possible* (AEAP) when the pairwise difference of their lengths is at most one. If this is true, we write $\simeq(\ell_1, \ell_2, \ldots \ell_k)$.

Remark. If a collection of chains have a fixed sum, the condition that the chains are AEAP determines the chain length sequence.

We return our focus to the concept of the multigraph of a graph, and prove a useful lemma. The notions of *bridge*, *cutvertex* and *block* for multigraphs are the same as for ordinary graphs; the definitions need not be repeated here. The key in the proof of the following lemma is that G has a bridge if G_M has a bridge and G has a cutvertex if G_M has a cutvertex or a loop. (The converses are also true.)

Lemma 5. *If G_M has a bridge, a cutvertex or a loop, G is not (n, m, p)-optimal for any p.*

Proof. Firstly, suppose G_M has a bridge. Then G has a chain which separates G, and every edge in such a chain is a bridge. Therefore, according to Proposition 1, G is not (n, m, p)-optimal for any p.

Secondly, suppose G_M has a cutvertex u. Let A and B be two distinct blocks in G_M containing u, and let $v \in A$ and $w \in B$ be other vertices in G_M. Since u is a cutvertex, every vw-path in G_M passes u. G_M is obtained from G by suppressing 2-vertices, and vertex suppression clearly does not block or remove paths between other vertices; it can only make them shorter. Therefore, every vw-path passes u also in G, which implies that u is a cutvertex in G. According to Proposition 2, G is not (n, m, p)-optimal for any p.

Lastly, suppose that G_M has a vertex u with a loop ℓ_1. We prove that u is a cutvertex in G, which is somewhat obvious, geometrically. Since an associated multigraph has no vertices of degree two, the degree of u is at least 3, so G_M contains at least one vertex w distinct from u or a loop ℓ_2 distinct from ℓ_1. A loop in G_M corresponds to a chain in G that contains at least two vertices (which are distinct from u). Therefore, in G, let

v be a vertex in the chain ℓ_1, let w' be a vertex not in ℓ_1 and distinct from u and let Π be a path in G from v to w'. Since the chain ℓ_1 is adjacent to u in both ends, Π passes (contains) u, which makes u a cutvertex in G. Therefore, according to Proposition 2, G is not (n, m, p)-optimal for any p. $\square$

Remark. If G_M has a bridge, it automatically has a cutvertex. The statement of the lemma and the structure of the proof are therefore logically slightly redundant.

5.2 The p-optimal $(n, n+1)$-graph

We can now further refine the categorization of $(n, n+1)$-graphs that we began in (5.1). Suppressing the 2-vertices, every $\langle 4, {}^x2\rangle$-graph corresponds to some $\langle 4\rangle$-multigraph, and every $\langle {}^23, {}^x2\rangle$-graph corresponds to some $\langle {}^23\rangle$-multigraph. Fortunately, the number of possible $\langle 4\rangle$- or $\langle {}^23\rangle$-multigraphs is very limited.

Case 1: The $\langle 4, {}^x2\rangle$-graphs all look essentially like the graph illustrated in Figure 4.1, but with the lengths of the cycles varying. Any such graph G corresponds to a $\langle 4\rangle$-multigraph; that is, G_M is a vertex with two loops. According to Lemma 5, G is not $(n, n+1, p)$-optimal, for any p.

Case 2: G_M is a $\langle {}^23\rangle$-multigraph. An edge in G_M is either a loop or it connects the two 3-vertices. If G_M has a loop, G is not $(n, n+1, p)$-optimal for any p according to Lemma 5. If G_M does not have a loop, it has three edges in parallel, which means that G has three chains in parallel between two 3-vertices; see Figure 5.2 for an example. Let us call an $(n, n+1)$-graph that consists of three parallel chains (all incident in each end to one of two 3-vertices) a *3-parallel-chains graph*, and we have proven the following:

Lemma 6. *An $(n, n+1, p)$-optimal graph, for any p, must be a 3-parallel-chains graph.*

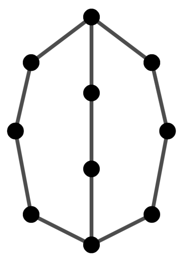


Figure 5.2: A $(10, 11)$-graph with three parallel chains of length 4, 4 and 3. This graph is shown to be $(10, 11, p)$-optimal for all p.

Take a 3-parallel-chains (n, m)-graph G and call the lengths of the chains ℓ_1, ℓ_2 and ℓ_3. Consider a percolation realization G_p. It is easy to see that G_p is disconnected if and only if some chain has at least two edges missing or at least one edge is missing from each chain. In other words, G_p is connected if and only if no edge is closed, exactly one edge is closed, or exactly two edges, which belong to different chains, are closed. Referring to

(4.1), this means that $c_0 = 1$, $c_1 = m$ and $c_2 = \ell_1\ell_2 + \ell_1\ell_3 + \ell_2\ell_3$, and the formula for $P_c(G,p)$ when G is a 3-parallel-chains graph is:

$$P_c(G,p) = p^m + mp^{m-1}q + (\ell_1\ell_2 + \ell_1\ell_3 + \ell_2\ell_3)\, p^{m-2}q^2\,. \tag{5.2}$$

Proposition 3. *Fix $n \geq 4$. There exists a graph which is uniquely $(n, n+1, p)$-optimal for all p, namely the 3-parallel-chains graph of order n with AEAP chains.*

Proof. First note that for given n, there is a unique 3-parallel-chains graph with AEAP chains. We can therefore prove the contrapositive statement that if a 3-parallel-chains graph G does not have AEAP chains, then G is not $(n, n+1, p)$-optimal for any p. Since no other graph can be $(n, n+1, p)$-optimal by Lemma 6, the conclusion will follow.

Let G be an $(n, n+1)$-graph of 3 parallel chains, and suppose that the chains are not as equal as possible. The expression for P_c given in (5.2) is strictly maximized when the last coefficient c_2, which is a function of (ℓ_1, ℓ_2, ℓ_3), is.

Let WLOG $\ell_1 > \ell_2 + 1$. Construct G' by moving one edge (and vertex) from ℓ_1 to ℓ_2, so that the first chain becomes shorter and the second longer. Substitution of the new chain lengths $\ell_1 - 1$ and $\ell_2 + 1$ into the formula for c_2 gives $c_2' := (\ell_1 - 1)(\ell_2 + 1) + (\ell_1 - 1)\ell_3 + (\ell_2 + 1)\ell_3 \;=\; \ell_1 - (\ell_2 + 1) + c_2 > c_2$.

According to Lemma 4 G' is strictly more stable than G, so G is not $(n, n+1, p)$-optimal, for any p. ☐

Example. The graph in Figure 5.2 is the unique 3-parallel-chains $(10, 11)$-graph G which has chains of lengths 4, 4 and 3. By Proposition 3, G is the unique $(10, 11, p)$-optimal graph for all p.

The following lemma generalizes an important idea used in the proof of Proposition 3.

Lemma 7. *Consider a coefficient $c_k(G)$, where G is a variable graph of fixed size and order and associated to a fixed multigraph G_M. Let two of the chains of G be ℓ_i and ℓ_j and consider the lengths of the other chains to be fixed. Suppose that*

$$c_k(G) = c_i(\ell_i, \ell_j) = a\ell_i\ell_j + b_1(\ell_i + \ell_j) + b_2\,, \tag{5.3}$$

where $a > 0$ and $b_{1,2} \geq 0$ are constants that depend upon the lengths of the other chains. Under these constraints, $c_k(G)$ will strictly increase as the difference $|\ell_i - \ell_j|$ decreases. In particular, c_k can only attain a maximum if $\simeq(\ell_i, \ell_j)$.

Proof. Suppose that $c_k(G)$ and G are as in the statement of the lemma. Since we have to keep m and the other chains fixed, $\ell_i + \ell_j$ is also fixed, because $\ell_i + \ell_j$ equals m minus the sum of the other chains. Therefore (5.3) can be simplified as

$$c_i = a\ell_i\ell_j + b\,, \tag{5.4}$$

where b is the constant $b_1(\ell_i + \ell_j) + b_2$.

Choose a particular graph G and suppose WLOG that $\ell_i \geq \ell_j$ in G. If $\ell_i = \ell_j$ or $\ell_i = \ell_j + 1$, it is not possible to make $|\ell_i - \ell_j|$ any smaller. Therefore suppose that

$\ell_i \geq \ell_j + 2$. Construct a modified graph G' by moving one edge from ℓ_i to ℓ_j. The difference between the chains is then decreased by 2, which is the smallest possible change. By (5.4)

$$
\begin{aligned}
c_k(G') - c_k(G) &= a(\ell_i - 1)(\ell_j + 1) + b - (a\ell_i\ell_j + b) \\
&= a\,((\ell_i - 1)(\ell_j + 1) - \ell_i\ell_j) = a\,(\ell_i - (\ell_j + 1)) > 0\,.
\end{aligned}
$$

The lemma follows, since the chain sum and the associated multigraph are obviously the same for G and G'. □

Remark. In essence, the above lemma is an algebraic statement, that need not have anything to do with graphs, c-coefficients or chains. We have phrased it with our typical application in mind.

6 | A study of the case $m = n + 2$

6.1 The four D-graph designs

In this section we prove that there is a unique $(n, n+2, p)$-optimal graph for all p. We follow essentially the same technique as for the $(n, n+1)$-case. Since there is only one $(4, 6)$-graph (the complete graph on four vertices) we assume that $n \geq 5$. With the degree sum being $2m = 2n + 4$, there are *a priori* five possible degree sequences (excluding graphs with leaves):

$$\begin{aligned} &1\colon\ (6, 2, \ldots, 2) \\ &2\colon\ (5, 3, 2, \ldots, 2) \\ &3\colon\ (4, 4, 2, \ldots, 2) \\ &4\colon\ (4, 3, 3, 2, \ldots, 2) \\ &5\colon\ (3, 3, 3, 3, 2, \ldots, 2) \end{aligned} \tag{6.1}$$

The first and second cases are only possible for $n \geq 7$ and $n \geq 6$, respectively. Anyway, they can be ruled out without much ado. Fix p.

Case 1: G is a $\langle 6, {}^{x}2\rangle$-graph, so that G_M a $\langle 6\rangle$-multigraph; i.e. G_M is a single vertex with three loops. According to Lemma 5, G is not (n, m, p)-optimal.

Case 2: G is a $\langle 5, 3, {}^{x}2\rangle$-graph and so G_M is a $\langle 5, 3\rangle$-multigraph. An edge in G_M is either a loop or it connects the two vertices. Since there are at most 3 edges between the vertices, the 5-vertex has at least one loop. According to Lemma 5, G is not $(n, n+2, p)$-optimal.

Case 3: G is a $\langle {}^{2}4, {}^{x}2\rangle$-graph and so G_M is a $\langle {}^{2}4\rangle$-multigraph. Suppose G is (n, m, p)-optimal. Then G_M cannot have a loop, so all the four edges are incident to both vertices, as illustrated by the first graph in Figure 6.1, labeled D1.

Case 4: G is a $\langle 4, {}^{2}3, {}^{x}2\rangle$-graph and so G_M is a $\langle 4, {}^{2}3\rangle$-multigraph. The degree sum of G_M is $4 + 3 + 3 = 10$, so G_M has 5 edges. Suppose G is (n, m, p)-optimal. For G_M not to have a loop, every edge incident to the 4-vertex must go to one of the 3-vertices, 2 to each of them, and the fifth edge has to be incident to both 3-vertices. G_M therefore is isomorphic to the second multigraph shown in Figure 6.1, labeled D2.

Case 5: G is a $\langle {}^{4}3, {}^{x}2\rangle$-graph and G_M a $\langle {}^{4}3\rangle$-multigraph. The degree sum of G_M is 12, so G_M has 6 edges. Suppose G is (n, m, p)-optimal. Then G_M has no loops or cutvertices. There are two possible G_M that satisfy these requirements:

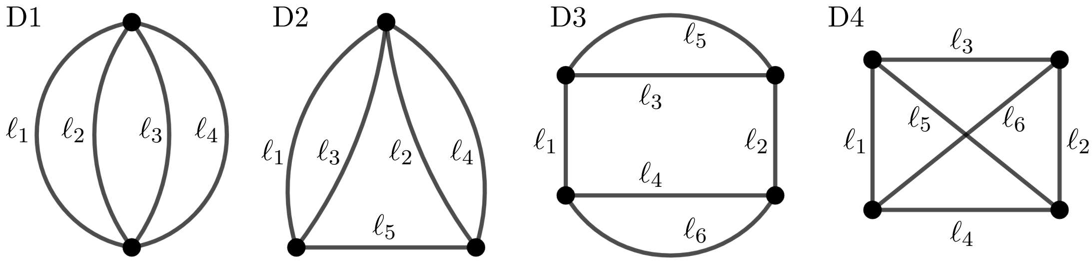


Figure 6.1: The four possible designs for a p-optimal $(n, n+2)$-graph.

To construct the first one, suppose that G_M has two vertices connected by two parallel edges, like the two uppermost vertices of D3 in Figure 6.1. A third edge between them would disconnect the graph, so the two vertices each have one edge that is incident to some other vertex. If these edges were incident to the same vertex, it would be a cutvertex, since every path to the fourth vertex would pass through it. So the two edges, labeled ℓ_1 and ℓ_2 in D3, connect to one each of the remaining two vertices, which then also must be doubly adjacent, as in D3.

For the second case, suppose that there are no multiple edges between any two vertices of G_M. This makes G_M an ordinary graph of 4 vertices and 6 edges. This is the complete 4-graph K^4, labeled D4 in Figure 6.1.

In conclusion: If G is an $(n, n+2, p)$-optimal graph, G_M has to be one of the four multigraphs in Figure 6.1, labeled D1, D2, D3 and D4, or, with a slight change of perspective, G has to be a *graph* of type D1, D2, D3 or D4. Call all such graphs collectively D-graphs, and we have proven the following lemma:

Lemma 8. *If G is an $(n, n+2, p)$-optimal graph for some p, then G is a D-graph.*

In Figure 6.1, we now let every ℓ_i represent a chain, so that the picture represents ordinary graphs. Let the set of all (n, m)-graphs represented by D1 be denoted $\mathrm{D1}^n_m$ (or equivalently, just $\mathrm{D1}_m$), and likewise for D2, D3, D4 and the entire collection D.

Let an ordered set of chain lengths be called a chain tuple, signified by a boldface $\boldsymbol{\ell}$. A chain tuple can be used to specify a graph G, given a multigraph G_M with the right number of edges and an edge labeling making clear how the chains should fit together. For example, referring to the chain labels in Figure 6.1, the D3-graph with chain tuple $\boldsymbol{\ell} = (\ell_1, \ell_2, \ell_3, \ell_4, \ell_5, \ell_6) = (2, 4, 1, 1, 3, 4)$ is a well defined graph, which is isomorphic to the D3-graph with chain tuple $(4, 2, 4, 3, 1, 1)$. There is, however, no D3-graph with chain tuple $(2, 2, 1, 3, 1, 3)$. (The reader could try to draw the graph to see why.) Note how the concept of chain tuple is different from that of the chain length sequence, which has been defined as the chain lengths of a graph in decreasing order.

Let G be any D-graph and perform percolation on it, obtaining G_p. The outcome graph is certainly connected if no or exactly one edge is closed, and disconnected if four or more edges are closed (since any $(n, n-2)$-graph is disconnected). If exactly two or three edges are closed, G_p could be either connected or disconnected. Our focus is therefore directed towards calculating, comparing and maximizing c_2 and c_3 of equation (4.1), (or equivalently, *minimizing* d_2 and d_3), for the four kinds of D-graphs.

6.2 Calculating and maximizing c_2

Consider again the graphs in Figure 6.1 viewed as multigraphs. Only one of them, D3, can be disconnected by removing two edges (ℓ_1 and ℓ_2). For this reason, D3-graphs will be treated somewhat separately.

Suppose G is a D1-, D2- or D4-graph. Calculating c_2 or d_2 is easy: A subgraph $G_2 \in \mathcal{G}_2$ with exactly two closed edges is counted by c_2 if the edges belong to different chains and to d_2 if they belong to the same chain. As previously noted, $c_2 + d_2 = \binom{m}{2} = |\mathcal{G}_2|$.

Thus, c_2 is a sum of the products of every possible pair of chains, and d_2 is the sum of all pairs of edges that belong to the same chain. Let $k \in \{4, 5, 6\}$ be the number of chains of G. Then

$$c_2 = \sum_{1 \le i < j \le k} \ell_i \ell_j \qquad\qquad d_2 = \sum_{i=1}^{k} \binom{\ell_i}{2}. \tag{6.2}$$

Lemma 9. *Among graphs of type* D1*,* D2 *or* D4*, the graphs with AEAP chains strictly minimize d_2 compared to other graphs of the same type and order.*

Proof. Let G be an (n, m)-graph of type D1, D2 or D4. Consider how d_2 changes if an edge is inserted into one of the chains of G, all else being equal. Call the chain ℓ_a, so that its length in the new graph G' is $\ell_a + 1$. Then $d_2' = d_2 + \ell_a$, which can be easily obtained from (6.2) above. Likewise, if $\ell_a \ge 2$ and an edge is collapsed in the chain, then d_2 decreases by $(\ell_a - 1)$.

Suppose now that G does not have AEAP chains. Then there are two chains, ℓ_i and ℓ_j, for which $\ell_j < \ell_i - 1$. Let us construct G' by first removing an edge from ℓ_i and then adding an edge to ℓ_j. By the above, $d_2' = d_2 - (\ell_i - 1) + \ell_j$, which is strictly less than d_2.

If, on the contrary, G is a graph with AEAP chains, this determines the chain length sequence, which determines d_2. Therefore d_2 is strictly minimized when the chains are AEAP. □

Remark. As an alternative to the proof above, one can observe that $c_2(G)$ in (6.2) satisfies the requirements of Lemma 7 for every pair of chains ℓ_i and ℓ_j. Therefore the chains of G have to be AEAP to maximize $c_2(G)$.

The next two lemmas help us compare d_2-values for two (n, m)-graphs belonging to different types.

Lemma 10. *Let G and G' be (n, m)-graphs of type* D1*,* D2 *or* D4 *such that G' has one or two more chains than G, and let a be the number of chains in G. If the i^{th} longest chain of G is at least as long as the i^{th} longest chain of G', for $i \in [1\,..\,a]$, then $d_2 > d_2'$.*

Proof. We prove the case where G is a D1-graph and G' is a D2-graph. The other two cases are proved analogously. Suppose that the chain length sequence (which is decreasing) is (k_1, k_2, k_3, k_4) for G and $(k'_1, k'_2, k'_3, k'_4, k'_5)$ for G'. As assumed, $k_i \geq k'_i$ for $i \in \{1, 2, 3, 4\}$.

By (6.2) we have

$$d_2 = \sum_{i=1}^{4} \binom{k_i}{2} \qquad\qquad d'_2 = \sum_{i=1}^{5} \binom{k'_i}{2}. \tag{6.3}$$

A binomial coefficient $\binom{k_i}{2}$ can be written as $0 + 1 + \cdots + (k_i - 1)$, which is a sum of k_i terms. Hence

$$d_2 = \sum_{i=1}^{4}(0 + 1 + \cdots + (k_i - 1)) \qquad d'_2 = \sum_{i=1}^{5}(0 + 1 + \cdots + (k'_i - 1)). \tag{6.4}$$

Counting every term in all partial sums once, both the expression for d_2 and the one for d'_2 has exactly m terms.

In the expression for d_2, since $k_i \geq k'_i$ for $i \in [1\,..\,4]$, we can split each partial sum into two parts, (where the second part might be empty): $0 + 1 + \cdots + (k_i - 1) = (0 + 1 + \cdots + (k'_i - 1)) + (k'_i + \cdots + k_i - 1)$. Note that each term of the second part is greater than or equal to k'_5.

Rearrange the sum for d_2 by moving the last part of each partial sum to the end of the whole sum. Since the total number of terms in the sums of d_2 and d'_2 are still equal, we can conclude that we have moved exactly k'_5 terms to the end of the d_2 sum:

$$d_2 = \sum_{i=1}^{4}(0 + 1 + \cdots + (k'_i - 1)) + [k'_5 \,\text{terms}] \quad d'_2 = \sum_{i=1}^{5}(0 + 1 + \cdots + (k'_i - 1)). \tag{6.5}$$

We have noted that each of the last k'_5 terms of d_2 is greater than or equal to k'_5. On the other hand, each of the last k'_5 terms in d'_2 is strictly less than k'_5. Since the preceding parts of the sums are equal, $d'_2 < d_2$. □

Lemma 11. *If G is a* D3*-graph and G' is a* D4*-graph with the same chain lengths, then* $d'_2 < d_2$ *.*

Proof. Fix n and m and let G and G' be as in the statement of the lemma. To calculate d_2, one can start with the formula for d'_2 in (6.2) and add the term $\ell_1\ell_2$ to account for the fact that the graph disconnects also when removing an edge from each of these two chains. (We include the formula for c_2 here as well, even if it is only used at a later point.)

$$c_2^{D3} = \sum_{1 \leq i < j \leq 6} \ell_i \ell_j - \ell_1 \ell_2 \qquad\qquad d_2^{\mathrm{D3}} = \sum_{i=1}^{6} \binom{\ell_i}{2} + \ell_1 \ell_2 \tag{6.6}$$

It is immediate that $d'_2 < d_2$. □

Remark. If one considers only $\mathrm{D3}_m$-graphs, it is possible to show that d_2 is minimized if and only if $\simeq(\ell_1+\ell_2, \ell_3, \ell_4, \ell_5, \ell_6)$.

Corollary 2. *Any* $\mathrm{D4}_m$*-graph whose chains are AEAP strictly minimizes* d_2 *compared to other* D_m*-graphs.*

Proof. Fix n and consider the D_m-graphs. By Lemma 11, no $\mathrm{D3}_m$-graph can minimize d_2.

By Lemma 9, we only need to compare d_2 for graphs whose chains are AEAP. Let $G \in \mathrm{D1}_m$, $G' \in \mathrm{D2}_m$ and $G'' \in \mathrm{D4}_m$ have AEAP chains. G'' has one more chain than G', which has one more chain than G. Since the m edges are distributed evenly over the chains, the longest chain in G is at least as long as the longest chain in G' which is at least as long as the longest chain in G'', and likewise for the second, third and fourth longest chains, (and the fifth longest chain in G' is at least as long as the fifth in G''). By Lemma 10, $d_2'' < d_2' < d_2$, so no graph in $\mathrm{D1}_m$ or $\mathrm{D2}_m$ minimizes d_2.

The only graphs remaining are any $\mathrm{D4}_m$-graphs with AEAP chains. Again by Lemma 9, any such graph equally minimizes d_2. □

Remark. For a given m, there might be several nonisomorphic $\mathrm{D4}_m$-graphs with AEAP chains.

6.3 Calculating c_3 for graphs of type D1, D2 and D4

Take a graph G of type D1 and consider a percolation outcome G_3 with exactly 3 edges closed. It is clear that G_3 is connected if and only if the three closed edges belong to three different chains. Therefore, c_3 is simply the number of ways to pick three edges from three different chains:

$$c_3^{\mathrm{D1}} = \ell_1\ell_2\ell_3 + \ell_1\ell_2\ell_4 + \ell_1\ell_3\ell_4 + \ell_2\ell_3\ell_4\,. \tag{6.7}$$

If G is of type D2, the situation is similar, except that closing three edges from three different chains will disconnect the graph if the three chains are incident to the same 3-vertex. There are $\binom{5}{3}-2=8$ combinations of three chains where closing an edge of each does not disconnect G. We get

$$c_3^{\mathrm{D2}} = \ell_1\ell_2\ell_3 + \ell_1\ell_2\ell_4 + \ell_1\ell_3\ell_4 + \ell_2\ell_3\ell_4 + (\ell_1\ell_2 + \ell_1\ell_4 + \ell_2\ell_3 + \ell_3\ell_4)\ell_5\,. \tag{6.8}$$

Lastly, if G is of type D4, a percolation outcome G_3 is connected if and only if the three closed edges belong to three different chains which are not incident to the same 3-vertex. There are $\binom{6}{3}-4=16$ such combinations of chains, and the formula for c_3 should therefore have 16 terms. It is easily checked that

$$\begin{aligned} c_3^{\mathrm{D4}} &= \ell_1\ell_2(\ell_3+\ell_4+\ell_5+\ell_6) + \ell_3\ell_4(\ell_1+\ell_2+\ell_5+\ell_6) \\ &\quad + \ell_5\ell_6(\ell_1+\ell_2+\ell_3+\ell_4) + \ell_1\ell_4\ell_5 + \ell_1\ell_3\ell_6 + \ell_2\ell_3\ell_5 + \ell_2\ell_4\ell_6\,. \end{aligned} \tag{6.9}$$

Remark. Notice how there is a natural pairing of the edges of the D4-multigraph where the edges of each pair are non-adjacent or "opposite" each other. The pairs are (ℓ_1, ℓ_2), (ℓ_3, ℓ_4) and (ℓ_5, ℓ_6).

6.4 A D1- or D2-graph does not maximize c_3

Lemma 12. *For every $G \in \mathrm{D1}_m$ there exists a graph $G' \in \mathrm{D2}_m$ with a strictly greater c_3-value (or, in case G has chains $(2,2,2,2)$, equal c_3-value) and which is also strictly more stable than G for all p. Furthermore, there exists a $\mathrm{D4}_8$-graph which has strictly greater c_3-value than the $D1_8$-graph with chains $(2,2,2,2)$.*

Proof. Let $G \in \mathrm{D1}_m$ with chain lengths $\boldsymbol{\ell} = (\ell_1, \ell_2, \ell_3, \ell_4)$. Suppose WLOG that $\ell_1 \geq \ell_{2,3,4}$ (by which we mean that ℓ_1 is at least as long as any of the other chains) and also WLOG that $\ell_3 \geq 2$ (since at most one chain can have length 1). Now modify G according to Figure 6.2: Let a be a 4-vertex, adjacent to $b \in \ell_1$ and $c \in \ell_3$. Exchange the edge ca for an edge cb, in effect disconnecting ℓ_3 in one of its ends and reconnecting it one vertex in along the first chain. (Notice the location of the chains ℓ_2 and ℓ_3.) This yields a $\mathrm{D2}_m$-graph which we call G', with chain lengths $\boldsymbol{\ell'} = (\ell_1 - 1, \ell_2, \ell_3, \ell_4, 1)$.

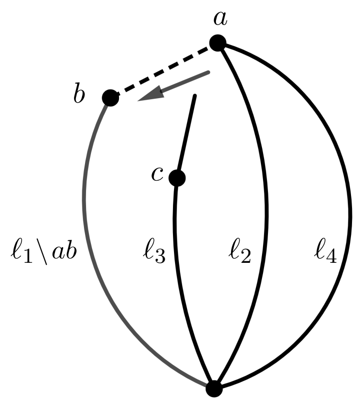


Figure 6.2: A graph G of type D1 is being transformed to type D2 by moving edge ca to cb. This makes the chain ℓ_1 one edge shorter and creates the 1-chain ab. (That the edge ab is dotted has significance only for the remark following Lemma 12.)

We obtain c_3' by substitution of the $\boldsymbol{\ell'}$-values into (6.8), and compare with $c_3 = c_3^{\mathrm{D1}}(G)$ from (6.7).

$$\begin{aligned} c_3' &= (\ell_1 - 1)(\ell_2\ell_3 + \ell_2\ell_4 + \ell_3\ell_4) + \ell_2\ell_3\ell_4 + ((\ell_1 - 1)(\ell_2 + \ell_4) + \ell_2\ell_3 + \ell_3\ell_4) \cdot 1 \\ &= c_3 - \ell_2\ell_3 - \ell_2\ell_4 - \ell_3\ell_4 + (\ell_1 - 1)(\ell_2 + \ell_4) + \ell_2\ell_3 + \ell_3\ell_4 \\ &= c_3 + (\ell_1 - 1)(\ell_2 + \ell_4) - \ell_2\ell_4 \,. \end{aligned} \tag{6.10}$$

To see if $c_3' >^? c_3$ we check the equivalent inequality $(\ell_1 - 1)(\ell_2 + \ell_4) >^? \ell_2\ell_4$. Since ℓ_1 is not shorter than the other chains, either $\ell_1 - 1 \geq \ell_2$ or $\ell_1 = \ell_2$.

Suppose the former. Together with $(\ell_2 + \ell_4) > \ell_4$, this proves the strict inequality. Instead suppose $\ell_1 = \ell_2$. In this case our inequality becomes $(\ell_1 - 1)(\ell_1 + \ell_4) - \ell_1\ell_4 = (\ell_1 - 1)\ell_1 - \ell_4 >^? 0$. Since $\ell_1 \geq \ell_4$, we are done if $\ell_1 > 2$, and also if $\ell_1 = 2$ and $\ell_4 = 1$. The only remaining case is $\ell_1 = \ell_2 = \ell_4 = 2$. By our assumptions we must have also $\ell_3 = 2$. These values only give us an equality above, and we will return to it shortly.

For the additional conclusion that G' is strictly more stable than G for all p, consider the chain length sequences of G and G'. G' has the same chain lengths, except that one of the longest chains of G has become one edge shorter and that there is an extra chain of length one. G and G' therefore satisfy the conditions of Lemma 10, so that $d_2 > d_2'$. An application of Lemma 4 gives that G' is strictly more stable than G. (G is not (n, m, p)-optimal for any p.)

A D1-graph G with all chains of length 2 is a $(6, 8)$-graph. According to (6.7), it has c_3-value $4 \cdot 2^3 = 32$. Compare with the $(6, 8)$-graph H of type D4 which has chains $\boldsymbol{\ell'} = (2, 2, 1, 1, 1, 1)$. Insertion into (6.9) yields $c_3(H) = 4\cdot4+6+6+4\cdot2 = 36 > c_3(G)$. □

Remark. The content of (6.10), that $c_3' = c_3 + (\ell_1 - 1)(\ell_2 + \ell_4) - \ell_2\ell_4$, is possible to conclude by careful inspection of Figure 6.2. Consider the subgraphs of G and G' with exactly three edges removed (i.e. the graphs G_3 of $\mathcal{G}_3$ and G_3' of $\mathcal{G}_3'$). When does the transformation (the moving of the edge from ca to cb) change the connectedness of such a subgraph? (That is, when is the connectedness of G_3' different from that of G_3?) It is easy to see that ab has to be one of the closed edges, if the edge-moving is to make any difference for connectedness. By carefully considering where the other two closed edges might be located for the transformation to turn a disconnected G_3-graph into a connected G_3'-graph (there are $(\ell_1 - 1)(\ell_2 + \ell_4)$ possibilities) and then the other way around (there are $\ell_2\ell_4$ such possibilities), one can see that, certainly, $c_3' = c_3 + (\ell_1 - 1)(\ell_2 + \ell_4) - \ell_2\ell_4$.

Lemma 13. *For each $G \in \mathrm{D2}_m$ there exists a $G' \in \mathrm{D4}_m$ with a strictly greater c_3-value, which is also strictly more stable than G for all p.*

Proof. Let $G \in \mathrm{D2}_m$ have chain lengths $\boldsymbol{\ell} = (\ell_1, \ell_2, \ell_3, \ell_4, \ell_5)$ (see Figure 6.3) and suppose WLOG that $\ell_1 \geq \ell_{2,3,4}$. In particular, this means that $\ell_1 \geq 2$. Let a be the vertex where chains $\ell_{1,2,3,4}$ meet and let $d \in \ell_1$ be adjacent to a. Modify G to a $\mathrm{D4}_m$-graph G' by moving the edge in ℓ_4 incident to a so that it is instead incident to d, as in the middle graph of Figure 6.3. The sixth chain in this D4-graph is the edge ad.

Substitute $\boldsymbol{\ell}$ into (6.8) to obtain c_3 and $\boldsymbol{\ell'} = (\ell_1 - 1, \ell_2, \ell_3, \ell_4, \ell_5, 1)$ into (6.9) to obtain c_3', which we compare with c_3.

$$\begin{aligned}
c_3' = {} & (\ell_1 - 1)\ell_2(\ell_3 + \ell_4 + \ell_5 + 1) + \ell_3\ell_4(\ell_1 - 1 + \ell_2 + \ell_5 + 1) + \\
& (\ell_1 - 1 + \ell_2 + \ell_3 + \ell_4)\ell_5 + (\ell_1 - 1)\ell_3 + (\ell_1 - 1)\ell_4\ell_5 + \ell_2\ell_3\ell_5 + \ell_2\ell_4 \\
= {} & c_3 + \ell_1\ell_2 - \ell_2(\ell_3 + \ell_4 + \ell_5 + 1) + (\ell_1 + \ell_2 + \ell_3 + \ell_4 - 1)\ell_5 + \ell_1\ell_3 - \ell_3 - \ell_4\ell_5 + \ell_2\ell_4 \\
= {} & c_3 + (\ell_1 - 1)(\ell_2 + \ell_3 + \ell_5) + \ell_3\ell_5 - \ell_2\ell_3
\end{aligned}$$

We want to prove that $c_3' > c_3$ by confirming the inequality

$$(\ell_1 - 1)(\ell_2 + \ell_3 + \ell_5) + \ell_3\ell_5 \overset{?}{>} \ell_2\ell_3. \tag{6.11}$$

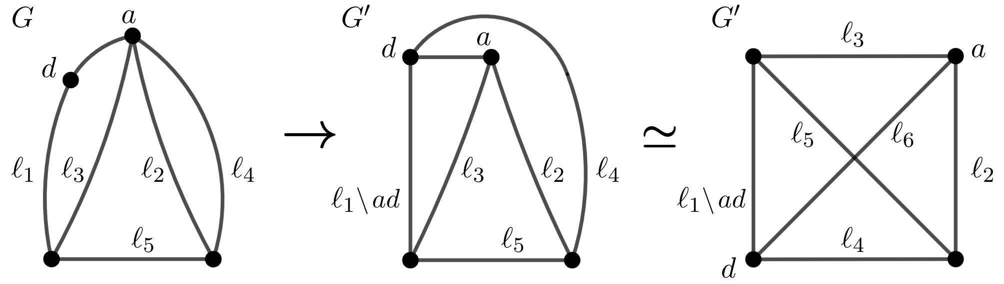


Figure 6.3: A graph of type D2 (left) is modified to one of type D4 (middle), by moving an edge of chain ℓ_4 from vertex a to the adjacent vertex d. The graph is then redrawn in standard D4-shape (right).

If $\ell_1 - 1 \geq \ell_3$, (6.11) is immediate. Suppose therefore that $\ell_1 = \ell_3$. Eliminating ℓ_3 and rearranging gives $(\ell_1 - 1)(\ell_1 + \ell_5) + \ell_1\ell_5 >^? \ell_2$. Since LHS $> (\ell_1 + \ell_5) > \ell_2$, we have confirmed (6.11).

Consider the chain length sequences of G and G': The latter has the same chain lengths, except that one chain is one edge shorter and that there is an extra chain of length one. G and G' therefore satisfy the conditions of Lemma 10, so that $d_2 > d_2'$. An application of Lemma 4 gives that G' is strictly more stable than G, so G is not (n, m, p)-optimal for any p. □

6.5 A D3-graph does not maximize c_3

For a graph G of type D3, c_3 is calculated in a similar way as for D4, but observing that a cut through ℓ_1 and ℓ_2 disconnects the graph. Going through the 20 possible combinations of three chains in D3, Figure 6.1, there are 12 that do not constitute a cut, which makes for 12 terms in c_3.

$$c_3^{\mathrm{D3}} = \ell_1(\ell_3\ell_4 + \ell_3\ell_6 + \ell_4\ell_5 + \ell_5\ell_6) + \ell_2(\ell_3\ell_4 + \ell_3\ell_6 + \ell_4\ell_5 + \ell_5\ell_6) + \\ \ell_3\ell_4\ell_5 + \ell_3\ell_4\ell_6 + \ell_3\ell_5\ell_6 + \ell_4\ell_5\ell_6\,. \tag{6.12}$$

The expression can be factorized further:

$$c_3^{\mathrm{D3}} = (\ell_1 + \ell_2)(\ell_3 + \ell_5)(\ell_4 + \ell_6) + \ell_3\ell_4\ell_5 + \ell_3\ell_4\ell_6 + \ell_3\ell_5\ell_6 + \ell_4\ell_5\ell_6\,. \tag{6.13}$$

Lemma 14. *A* $\mathrm{D3}_m$*-graph* G *in which* $\not\simeq(\ell_3, \ell_5)$ *has a strictly smaller* c_3*-value than the graph* G' *in which* $\not\simeq(\ell_3, \ell_5)$ *but the lengths of the other chains are unchanged. Furthermore,* G' *is strictly more stable than* G *for all* p*. These statements are equally true with* (ℓ_4, ℓ_6) *instead of* (ℓ_3, ℓ_5)*.*

Proof. In (6.13), c_3^{D3} can quite easily be rewritten as to satisfy (5.3) in Lemma 7 for ℓ_3 and ℓ_5. According to the lemma, $c_3(G') > c_3(G)$. Likewise, c_2^{D3} in (6.6) satisfies Lemma 7 for (ℓ_3, ℓ_5), so that $c_2(G') > c_2(G)$. Therefore, a graph in which these chains are not AEAP is not (n, m, p)-optimal for any p by Lemma 4. The case with (ℓ_4, ℓ_6) is entirely similar. $\square$

Lemma 15. *For every $G \in \mathrm{D3}_m$ with $\simeq(\ell_3, \ell_5)$ and $\simeq(\ell_4, \ell_6)$, there exists a $G' \in \mathrm{D4}_m$ with a strictly greater c_3-value, which is also strictly more stable than G for all p.*

Proof. Let G be a $\mathrm{D3}_m$-graph in which $\simeq(\ell_3, \ell_5)$ and $\simeq(\ell_4, \ell_6)$. Suppose WLOG that $\ell_3 \geq \ell_4, \ell_5, \ell_6$ and that $\ell_4 \geq \ell_6$.

Let G' be a $\mathrm{D4}_m$-graph with the same chain lengths. Let us consider $c_3' = c_3^{\mathrm{D4}}$ and $c_3 = c_3^{\mathrm{D3}}$ according to (6.9) and (6.12) respectively, and compute $c_3' - c_3$:

$$\begin{aligned} c_3' &= \ell_1\ell_2(\ell_3 + \ell_4 + \ell_5 + \ell_6) + \ell_1\ell_3\ell_4 + \ell_2\ell_3\ell_4 + \ell_3\ell_4\ell_5 + \ell_3\ell_4\ell_6 + \\ &\quad \ell_1\ell_5\ell_6 + \ell_2\ell_5\ell_6 + \ell_3\ell_5\ell_6 + \ell_4\ell_5\ell_6 + \ell_1\ell_4\ell_5 + \ell_1\ell_3\ell_6 + \ell_2\ell_3\ell_5 + \ell_2\ell_4\ell_6 \\ c_3 &= \ell_1\ell_3\ell_4 + \ell_1\ell_3\ell_6 + \ell_1\ell_4\ell_5 + \ell_1\ell_5\ell_6 + \ell_2\ell_3\ell_4 + \ell_2\ell_3\ell_6 + \ell_2\ell_4\ell_5 + \ell_2\ell_5\ell_6 + \\ &\quad \ell_3\ell_4\ell_5 + \ell_3\ell_4\ell_6 + \ell_3\ell_5\ell_6 + \ell_4\ell_5\ell_6 \\ c_3' - c_3 &= \ell_1\ell_2(\ell_3 + \ell_4 + \ell_5 + \ell_6) + \ell_2\ell_3\ell_5 + \ell_2\ell_4\ell_6 - \ell_2\ell_3\ell_6 - \ell_2\ell_4\ell_5 \\ &= \ell_1\ell_2(\ell_3 + \ell_4 + \ell_5 + \ell_6) + \ell_2(\ell_3 - \ell_4)(\ell_5 - \ell_6) \end{aligned}$$

We claim that this is strictly positive, which obviously hinges on the second (last) term. Consider the two possible relationships between ℓ_4 and ℓ_5. Case 1: $\ell_5 \geq \ell_4$. Then, by the initial assumptions, $\ell_3 \geq \ell_5 \geq \ell_4 \geq \ell_6$, which makes the last term nonnegative. Case 2: $\ell_4 > \ell_5$. Since $\simeq(\ell_3, \ell_5)$, this forces $\ell_3 = \ell_4$. Thus, the last term is 0. In either case, $c_3' > c_3$.

According to Lemma 11, $d_2' < d_2$, and by Lemma 4, G' is strictly more stable than G for all p. $\square$

Remark. One might wonder if it is possible to prove the preceding results more directly, by taking any D3-graph and cunningly rearrange the chains into a more stable D4-graph. There are cases where this doesn't work. For example, given fixed chain length sequence $(16, 4, 4, 1, 1, 1)$ the D3-graph with $\boldsymbol{\ell} = (1, 1, 16, 4, 1, 4)$ maximizes c_3 and is actually the most stable such graph for small values of p.

6.6 Finding the p-optimal $\mathrm{D4}_m$-graph

Lemma 16. *For every $G \in \mathrm{D4}_m$ in which the chains are not AEAP, there exists a $\mathrm{D4}_m$-graph G' with a strictly greater c_3-value, which is also strictly more stable than G for all p.*

Proof. Let $G \in \mathrm{D4}_m$ with $\not\simeq(\ell_1, \ldots, \ell_6)$. We construct a D4-graph G' of the same order by moving one edge from a longest chain to a shortest chain.

Pick two chains in G, say $\ell_i > \ell_j$, that has the largest possible pairwise difference, which is then at least 2. There are two different ways these chains can be situated relative to each other.

Case 1: The chains are adjacent. Suppose WLOG that $\ell_i = \ell_1$ and $\ell_j = \ell_4$. Construct G' from G by moving one edge from ℓ_1 to ℓ_4, so that $\ell'_1 = \ell_1 - 1$ and $\ell'_4 = \ell_4 + 1$. Substitution of ℓ' into (6.9) yields

$$\begin{aligned}
c'_3 &= (\ell_1 - 1)\ell_2(\ell_3 + \ell_4 + \ell_5 + \ell_6 + 1) + \ell_3(\ell_4 + 1)(\ell_1 + \ell_2 + \ell_5 + \ell_6 - 1) + \\
&\quad \ell_5\ell_6(\ell_1 + \ell_2 + \ell_3 + \ell_4) + (\ell_1 - 1)(\ell_4 + 1)\ell_5 + (\ell_1 - 1)\ell_3\ell_6 + \ell_2\ell_3\ell_5 + \ell_2(\ell_4 + 1)\ell_6 \\
&= c_3 + \ell_1\ell_2 - \ell_2(\ell_3 + \ell_4 + \ell_5 + \ell_6 + 1) - \ell_3\ell_4 + \ell_3(\ell_1 + \ell_2 + \ell_5 + \ell_6 - 1) \\
&\quad + \ell_1\ell_5 - \ell_4\ell_5 - \ell_5 - \ell_3\ell_6 - \ell_2\ell_6 \\
&= c_3 - \ell_2\ell_5 + \ell_3\ell_5 + (\ell_1 - \ell_4 - 1)(\ell_2 + \ell_3 + \ell_5)\,.
\end{aligned}$$

We need to show $c'_3 > c_3$. This is true if and only if

$$(\ell_1 - \ell_4 - 1)(\ell_2 + \ell_3 + \ell_5) > (\ell_2 - \ell_3)\ell_5\,.$$

If $\ell_1 - \ell_4 > \ell_2 - \ell_3$, the condition is trivial. Otherwise, by assumption, no pairwise difference can exceed that of $\ell_1 - \ell_4$, but equality is possible, i.e. $\ell_1 - \ell_4 = \ell_2 - \ell_3 =: k+1$, with $k \geq 1$. In that case we need to prove

$$k(\ell_2 + \ell_3 + \ell_5) > (k+1)\ell_5\,,$$

which is true if $\ell_2 + \ell_3 > \ell_5$. Suppose on the contrary that $\ell_5 \geq \ell_2 + \ell_3$, i.e. $\ell_5 - \ell_3 \geq \ell_2$. Since $\ell_1 - \ell_4 = \ell_2 - \ell_3$ is the largest possible difference, we have

$$\ell_2 - \ell_3 \geq \ell_5 - \ell_3 \geq \ell_2\,,$$

which is absurd, since chain lengths are positive.

Case 2: The two chains are not adjacent, but "opposite" each other. Suppose WLOG that $\ell_i = \ell_1$ and $\ell_j = \ell_2$, so that $\ell_1 > \ell_2 + 1$. As above, construct G' from G by moving one edge from ℓ_1 to ℓ_2 and substitute $\ell'_1 = \ell_1 - 1$ and $\ell'_2 = \ell_2 + 1$ into (6.9) to get

$$\begin{aligned}
c'_3 &= (\ell_1 - 1)(\ell_2 + 1)(\ell_3 + \ell_4 + \ell_5 + \ell_6) + \ell_3\ell_4(\ell_1 + \ell_2 + \ell_5 + \ell_6) + \ell_5\ell_6(\ell_1 + \ell_2 + \\
&\quad \ell_3 + \ell_4) + (\ell_1 - 1)\ell_3\ell_6 + (\ell_1 - 1)\ell_4\ell_5 + (\ell_2 + 1)\ell_3\ell_5 + (\ell_2 + 1)\ell_4\ell_6. \\
&= c_3 + (\ell_1 - \ell_2 - 1)(\ell_3 + \ell_4 + \ell_5 + \ell_6) - \ell_3\ell_6 - \ell_4\ell_5 + \ell_3\ell_5 + \ell_4\ell_6 \\
&= c_3 + (\ell_1 - \ell_2 - 1)(\ell_3 + \ell_4 + \ell_5 + \ell_6) + (\ell_3 - \ell_4)(\ell_5 - \ell_6)\,.
\end{aligned}$$

Now, as above, we want to show that $c'_3 > c_3$. This is true if and only if

$$(\ell_1 - \ell_2 - 1)(\ell_3 + \ell_4 + \ell_5 + \ell_6) > (\ell_4 - \ell_3)(\ell_5 - \ell_6)\,. \tag{6.14}$$

If the right hand side is non-positive, there is nothing left to prove. Suppose therefore that it is positive; we can assume WLOG that $\ell_4 > \ell_3$ and $\ell_5 > \ell_6$. Suppose that at least one of the differences $\ell_4 - \ell_3$ and $\ell_5 - \ell_6$ is strictly less than $\ell_1 - \ell_2$. WLOG, $\ell_1 - \ell_2 - 1 \geq \ell_4 - \ell_3$.

Then (6.14) is trivially true. On the other hand, if $\ell_1 - \ell_2 = \ell_4 - \ell_3 = \ell_5 - \ell_6 =: k+1 > 1$, then we need to prove the inequality on the next line (which is subsequently simplified):

$$
\begin{gathered}
k(\ell_3 + \ell_4 + \ell_5 + \ell_6) > (k+1)^2 \\
k((k+1+\ell_3) + \ell_3 + (k+1+\ell_6) + \ell_6) > (k+1)^2 \\
(2k)(k+1+\ell_3+\ell_6) > (k+1)^2 .
\end{gathered}
$$

The last line is true since $2k \geq k+1$ and $k+1+\ell_3+\ell_6 > k+1$.

In each case we have $c'_3 > c_3$. Since the expression for c_2 in (6.2) satisfies Lemma 7 for any pair of chains, we also have $d'_2 < d_2$. Therefore $G \prec_p G'$ for any p by Lemma 4. □

In Corollary 2 we proved that demanding AEAP chains is necessary and sufficient to maximize $c_2(G)$ for $G \in \mathrm{D4}_m$. When it comes to $c_3(G)$, having AEAP chains is necessary, but not sufficient. Suppose we have a D4-graph with $\simeq(\ell_1, \ldots, \ell_6)$. If the lengths of two or three chains differs from the others (by one edge), then we need to examine how the relative positions of these chains affect c_3.

Lemma 17. *Consider the* D_m*-graphs, and let* r *be the remainder when* m *is divided by 6. The following uniquely specifies a graph* $G \in \mathrm{D4}_m$*, with AEAP chains, which strictly maximizes* $c_3(G)$ *for* $G \in \mathrm{D}_m$*:*

- *If* r *equals* 0, 1 *or* 5*, there is only one such graph (or they are isomorphic).*
- *If* $r = 2$*, let the two longer chains be non-adjacent.*
- *If* $r = 3$ *let the three longer chains be arranged so that they form a (simple) path.*
- *If* $r = 4$*, let the two shorter chains be non-adjacent.*

Proof. By Lemma 12, Lemma 13, Lemma 14 and Lemma 15 no graph in $\mathrm{D1}_m$, $\mathrm{D2}_m$ or $\mathrm{D3}_m$ maximizes $c_3(G)$ for $G \in \mathrm{D}_m$. It is therefore enough to consider only the $\mathrm{D4}_m$-graphs. We will prove the equivalent result that the specified graph strictly minimizes d_3.

If $r = 0$, G is the D4-graph with chains of equal length.

If $r = 1$ or $r = 5$, one chain is longer or shorter than the others. Because of the symmetry of $G_M = K^4$, the possible graphs are isomorphic; it does not matter which chain differs in length.

If $r = 2$, two chains are one edge longer than the others. Suppose that the shorter chains have length ℓ and the longer length $\ell + 1$. The two longer chains might be either adjacent or non-adjacent ("opposite") each other. (See Figure 6.4.) Suppose that G is a graph of the former kind, in which $\ell_1 = \ell_3 = \ell + 1$ and G' is a graph of the latter, in which $\ell_1 = \ell_2 = \ell + 1$.

A percolation outcome $G_3 \in \mathcal{G}_3$ or $G'_3 \in \mathcal{G}'_3$ is disconnected if and only if either two edges are missing from one chain or if one edge is missing from each of three mutually adjacent chains (which disconnects a 3-vertex from the others). The number of disconnected outcomes of the first kind is obviously not affected by how the chains are

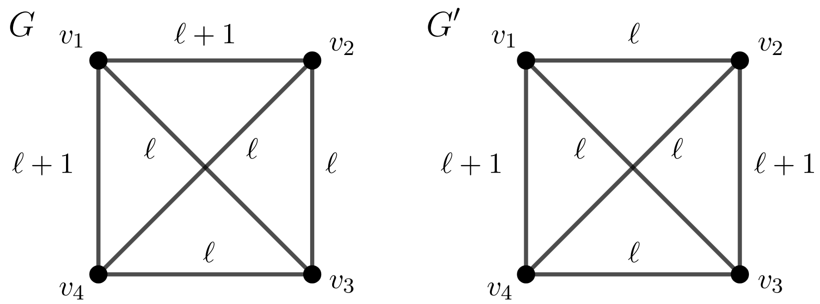


Figure 6.4: Two longer chains in a D4-graph can be either adjacent to or opposite each other.

arranged. As for the second kind, referring to Figure 6.4, there are $\ell(\ell+1)^2$ minimal cuts that disconnect v_1 from the other 3-vertices in G, and $\ell^2(\ell+1)$ in G'. For v_3 the numbers are ℓ^3 and $\ell^2(\ell+1)$, while for v_2 and v_4 there is no difference between G and G'. This implies that

$$d_3 - d_3' = \ell(\ell+1)^2 - \ell^2(\ell+1) + \ell^3 - \ell^2(\ell+1) = \ell\,. \tag{6.15}$$

If $r = 4$, two chains are one edge shorter than the others. The calculations are very similar to the case $r = 2$ (change $+1$ in the parentheses in (6.15) to -1), and the result is identical.

If $r = 3$, three chains are one edge longer than the other three. Consider an edge coloring of the D4-multigraph (K^4) where blue signifies that the corresponding chain should have length ℓ and red signifies length $\ell+1$. How many 2-edge-colorings of K^4 are there, with three edges of each color? The three possibilities are illustrated in Figure 6.5. In the first, the red edges form a 3-star and the blue a 3-cycle. In the second, the colors are reversed. In the third, each of the colors form a (simple) path. One might need a moment's thought to convince oneself that these are the only possibilities, especially that different-looking colorings of the last kind are isomorphic.

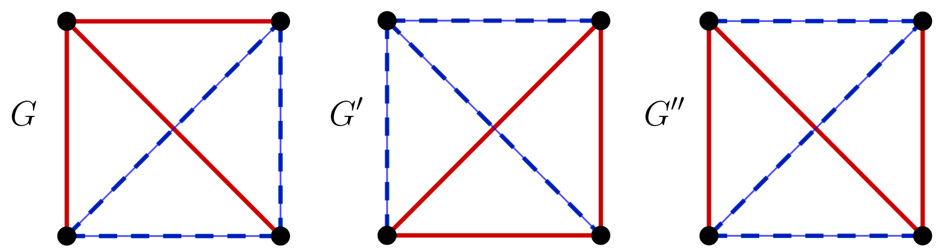


Figure 6.5: The three 2-edge-colorings of K^4 with three edges of each color. For our purposes, each edge should be interpreted as a chain, where red chains are one edge longer than blue chains. The blue edges are dashed for readability in black-and-white.

We now compare the values of d_3 for the three graphs G, G' and G'' that correspond to the three possibilities that the longer chains are arranged in a star, cycle and path, respectively.

Consider the outcome sets $\mathcal{G}_3$, $\mathcal{G}_3'$ and $\mathcal{G}_3''$. As previously noted, a graph in either of these sets is disconnected if and only if either two edges are closed in the same chain or one edge is missing from each of three mutually adjacent chains. The number of graphs that belong to the former category does not depend on the arrangement of the chains. Supposing that the number of such graphs is a, we can easily calculate d_3, d_3' and d_3'':

$$\begin{aligned} d_3 &= a + (\ell+1)^3 && + 3(\ell+1)\ell^2 \\ d_3' &= a \quad + 3(\ell+1)^2\ell && + \ell^3 \\ d_3'' &= a \quad + 2(\ell+1)^2\ell + 2(\ell+1)\ell^2 \,. \end{aligned} \tag{6.16}$$

Comparing them gives $d_3 = d_3' + 1$ and $d_3' = d_3'' + \ell$, i.e $d_3 > d_3' > d_3''$. □

Theorem 1. *For $n \geq 4$, there is a graph G which is uniquely $(n, n+2, p)$-optimal for all p. The graph is of type* D4 *with six AEAP chains. If two of the chains in G are longer or shorter than the others, they are non-adjacent. If three of the chains in G are longer than the others, they are arranged in a path.*

Proof. If $n = 4$, the only $(n, n+2)$-graph is K^4, which is of type D4 with equal chains.

Fix $n \geq 5$ and $p \in (0, 1)$. Lemma 8 says that an $(n, n+2, p)$-optimal graph, has to be a D-graph (D1-, D2-, D3- or D4-graph). By Lemma 17 the specified graph strictly maximizes $c_3(G)$ for $G \in \mathrm{D}_m$. According to Lemma 9, the graph also (non-strictly) maximizes $c_2(G)$ for $G \in \mathrm{D}_m$.

By Lemma 4, G is strictly more stable than the other D_m-graphs. Since p was arbitrary, G is the unique $(n, n+2, p)$-optimal graph for all p. □

7 | The ordering depends upon p

7.1 The relative stability of two graphs

Definition 29 (ℓ-edge-connected)**.** If $G = (V, E)$ is a graph of order at least 2, then G is ℓ-edge-connected if $G - F$ is connected for every edge set $F \subset E$ such that $|F| < \ell$.

Definition 30 (Edge connectivity, $\lambda(G)$)**.** The greatest integer ℓ such that G is ℓ-edge-connected is the edge connectivity of G, denoted $\lambda(G)$.

Define κ to be $m - n + 1$.

Lemma 18. *Let G and G' be (n,m)-graphs and let $\lambda = \min\{\lambda(G), \lambda(G')\}$. If $c_\lambda(G) > c_\lambda(G')$, then G is strictly more stable than G' for sufficiently large p. If $c_\kappa(G) < c_\kappa(G')$, then G is strictly less stable than G' for sufficiently small p.*

Proof. Let G, G' and λ be as above, so that λ is the smallest number of edges which can disconnect at least one of the graphs. Consider the formula (4.1) for $P_c(G,p)$. For $i < \lambda$, $c_i(G) = c_i(G') = \binom{m}{i}$, since neither graph can be disconnected by removing i edges, and c_λ is the first coefficient that is not equal to $\binom{m}{i}$ for both graphs.

The last nonzero coefficient is coefficient number $\kappa := m - n + 1$, by Lemma 3. By (4.1)

$$P_c(G,p) - P_c(G',p) = (c_\lambda - c'_\lambda)p^{m-\lambda}q^\lambda + \cdots + (c_\kappa - c'_\kappa)p^{m-\kappa}q^\kappa\,. \tag{7.1}$$

It is possible for the sum to be empty (consider C^n). However, if $\lambda \geq \kappa$, then by the above $c_i = c'_i$ for all i. Therefore, assume that $\lambda < \kappa$; the sum then contains at least two terms.

Suppose that $c_\lambda > c'_\lambda$, so that the first term of (7.1) is positive. By choosing a p close enough to 1 (so that q is close enough to zero), the other terms can be made sufficiently small, so that $P_c(G,p) - P_c(G',p)$ is positive. For large p, then, $G \succ_p G'$. Suppose also that $c_\kappa < c'_\kappa$. Then, if p is sufficiently small, the above difference is negative, so that $G \prec_p G'$. This pair of conditions is apparently sufficient for the relative stability of G and G' to be p-dependent. □

Corollary 3. *If G and G' are $(n, n+1)$-graphs with coefficients (c_1, c_2) and (c'_1, c'_2) respectively, then it is necessary and sufficient for their relative stability to depend upon p that either $c_1 > c'_1$ and $c_2 < c'_2$ (in which case $G' \prec_p G$ for large p and $G \prec_p G'$ for small p) or $c_1 < c'_1$ and $c_2 > c'_2$ (in which case $G \prec_p G'$ for large p and $G' \prec_p G$ for small p).*

Proof. Sufficiency follows from Lemma 18. (If $c_1(G) \neq c_1(G')$, then at least one of the graphs has edge connectivity $\lambda = 1$.) Necessity follows from considering the alternative: If $c_1 \geq c_1'$ and $c_2 \geq c_2'$, then G is strictly more stable than G' for all p by Lemma 4. If $c_1 \leq c_1'$ and $c_2 \leq c_2'$ then G' is strictly more stable than G for all p. □

7.2 Handcuff graphs

Since we proved Proposition 1 and Proposition 2, we have not been much concerned about graphs with bridges or cutvertices. Reexamining the stability ordering of the $(n, n+1)$-graphs with these graphs in mind gives us a lucid example of how the stability ordering can depend on p.

Let G be an arbitrary $(n, n+1)$-graph. To keep the analysis simple, we assume that G has no leaves. In Section 5.2, we analyzed the possible multigraphs G_M and discarded two multigraphs with loops: 1) the $\langle 4 \rangle$-multigraph; a vertex with two loops, and 2) the $\langle {}^2 3 \rangle$-multigraph in which each of the two vertices has a loop. In each of these cases, an associated graph G has exactly two cycles; in the second case it also has a chain which separates them.

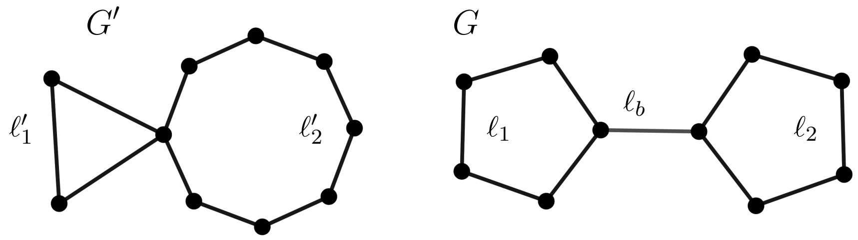


Figure 7.1: Two $(10, 11)$-handcuff graphs. G' is (strictly) more stable than G for large p, but G is more stable for small p.

Let us call graphs corresponding to the former multigraph figure-eight graphs. We can call graphs corresponding to the latter multigraph *proper handcuff graphs*. However, let graphs of both kinds together be called *handcuff graphs*; since we will handle them quite seamlessly together. One graph of each kind is illustrated in Figure 7.1.

If G is a proper handcuff graph, let ℓ_1 and ℓ_2 be the lengths of the cycles and ℓ_b be the length of the separating chain (where b is for bridge). However, if we allow for ℓ_b to be zero (when G is a figure-eight graph), we can use the same set of chains for all handcuff graphs. The chain sum formula, $\ell_1 + \ell_2 + \ell_b = m$, still holds.

Now, let G be an (n, m)-handcuff graph. For G_p to be connected either: 1) all edges are open, 2) a single edge is closed in one cycle or 3) exactly one edge is closed in each of the cycles. Therefore

$$P_c(G, p) = p^m + (\ell_1 + \ell_2)p^{m-1}q + (\ell_1 \ell_2)p^{m-2}q^2 \,. \tag{7.2}$$

The only constraints upon ℓ_1 and ℓ_2 are $\ell_{1,2} \geq 3$ (by the definition of cycle) and $\ell_1 + \ell_2 \leq m$.

Therefore, to find two (n, m)-handcuff graphs G and G' whose relative stability depend upon p, we need

$$\ell_1 + \ell_2 < \ell_1' + \ell_2' \qquad \text{and} \qquad \ell_1\ell_2 > \ell_1'\ell_2' \,. \tag{7.3}$$

These conditions are necessary and sufficient by Corollary 3. It is not hard to find four positive integers $(\ell_1, \ell_2, \ell_1', \ell_2')$ which satisfy (7.3). Then one can pick $m \geq \max\{\ell_1 + \ell_2, \ell_1' + \ell_2'\}$, calculate ℓ_b and ℓ_b' and, *voilà*, one has two graphs where $G \prec_p G'$ for large p but $G' \prec_p G$ for small p.

We wish to find the smallest m for which this is possible, and therefore systematize the sums and products for small numbers in a table:

$\ell_1 + \ell_2 = c_1$	6	7	8	9	10	11
$\ell_1\ell_2 = c_2$	$3 \cdot 3 = 9$	$3 \cdot 4 = 12$	$3 \cdot 5 = 15$ $4 \cdot 4 = 16$	$3 \cdot 6 = 18$ $4 \cdot 5 = 20$	$3 \cdot 7 = 21$ $4 \cdot 6 = 24$ $5 \cdot 5 = 25$	$3 \cdot 8 = 24$ $4 \cdot 7 = 28$ $5 \cdot 6 = 30$

Table 1: Possible values of c_1 and c_2 for small handcuff graphs. For given m, all c_1-values up to and including m are allowed together with the c_2-values given below.

The first c_2-value which is lower than a value in a preceding column is $c_2 = 3 \cdot 8 = 24$ for $c_1 = 11$ compared to $c_2 = 5 \cdot 5 = 25$ for $c_1 = 10$. Because of the different growth rates of the top row compared to the lowest entry of each column, it is clear that such c_2-values will be found in every subsequent column.

Here, then, are our numbers: $(\ell_1, \ell_2) = (5, 5)$, $(\ell_1', \ell_2') = (3, 8)$ and $m = \max\{\ell_1 + \ell_2, \ell_1' + \ell_2'\} = 11$. The corresponding graphs are the ones shown in Figure 7.1: G is the proper handcuff graph with chains $(5, 5, 1)$ and G' is the figure-eight graph with chains $(8, 3)$. G is strictly more stable than G' for small p, but strictly less stable for large p. We have proven the following:

Proposition 4. *The smallest* (n, m) *for which the stability ordering of the* (n, m)*-handcuff graphs is* p*-dependent is* $(10, 11)$*. The* p*-dependency subsists for all* $(n, n + 1)$ *with* $n \geq 10$*.*

7.3 Graphs with branches

We now turn our focus to the 3-parallel-chains graphs (3pc-graphs), which are $(n, n + 1)$-graphs that were defined in Section 5.2. This time we add an additional feature.

If one connects a tree (which might be just a single vertex) to another graph G, by adding an edge between the graphs, then the number of vertices added to G equals the number of added edges. Therefore, a 3pc-graph with "leaves and branches" is also an $(n, n + 1)$-graph. Let us make the notion of *branch* precise:

Definition 31 (Branch)**.** Let G be a connected graph with at least one cycle, and let H be the maximal (induced) connected subgraph of G that has no leaves. A maximal connected subgraph B of G so that B and H does not share a vertex, *together* with the edge connecting B to H, is called a *branch* of G.

Let G be a 3pc-graph with branches allowed (a 3pc_b-graph). G needs every edge in every branch to stay connected. Let b be the number of edges belonging to branches. The edges which are not in a branch are precisely the edges in the three chains. We can therefore obtain $P_c(G,p)$ by a modification of (5.2):

$$P_c(G,p) = p^m + (\ell_1 + \ell_2 + \ell_3)p^{m-1}q + (\ell_1\ell_2 + \ell_1\ell_3 + \ell_2\ell_3)\, p^{m-2}q^2\,, \tag{7.4}$$

where $\ell_1 + \ell_2 + \ell_3 = m - b$.

Now, to obtain two 3pc_b-graphs G and G' whose relative stability depend on p, we need (ℓ_1, ℓ_2, ℓ_3) and $(\ell_1', \ell_2', \ell_3')$ such that

$$\ell_1 + \ell_2 + \ell_3 < \ell_1' + \ell_2' + \ell_3' \quad \text{and} \quad \ell_1\ell_2 + \ell_1\ell_3 + \ell_2\ell_3 > \ell_1'\ell_2' + \ell_1'\ell_3' + \ell_2'\ell_3'\,. \tag{7.5}$$

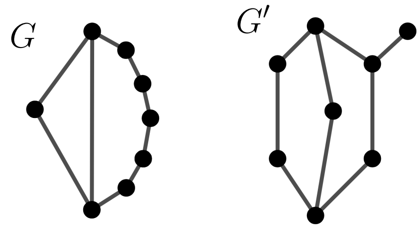


Figure 7.2: Two 3-parallel chains $(8, 9)$-graphs with branches allowed (the one to the right has a leaf). G is more stable than G' for large p but less stable for small p.

Proposition 5. *The stability ordering of* $(n, n+1)$*-3-parallel-chains graphs with branches is p-dependent when* $n \geq 8$*, but independent of* p *when* $n < 8$*.*

Proof. Let $\ell_1 + \ell_2 + \ell_3 =: k$. By Lemma 7, $c_2 = \ell_1\ell_2 + \ell_1\ell_3 + \ell_2\ell_3$ will, for fixed sum k, necessary attain its *minimum* value when $(\ell_1, \ell_2, \ell_3) = (k-3, 2, 1)$. This is because at most one chain can have length 1, all other possible (ℓ_1, ℓ_2, ℓ_3)-tuples can be obtained from $(k-3, 2, 1)$ by iteratively decreasing a difference between two chains. We can therefore modify the idea of the table in the last section by including, for each k, only the cases $(k-3, 2, 1)$ and $\simeq(\ell_1, \ell_2, \ell_3)$. In the former case we have $c_2 = 3(k-3) + 2 = 3k - 7$, so the first row increases by 3 for every column. Notice how the c_2-values in the lower row have a faster growth rate ($\sim k^2/3$).

$k = c_1$	5	6	7	8	9
(ℓ_1, ℓ_2, ℓ_3)	$(2,2,1)$	$(3,2,1)$	$(4,2,1)$	$(5,2,1)$	$(6,2,1)$
$\ell_1\ell_2+\ell_1\ell_3+\ell_2\ell_3$	4+2+2=8	11	14	17	20
(ℓ_1, ℓ_2, ℓ_3)	-	$(2,2,2)$	$(3,2,2)$	$(3,3,2)$	$(3,3,3)$
$\ell_1\ell_2+\ell_1\ell_3+\ell_2\ell_3$	-	4+4+4=12	6+6+4=16	9+6+6=21	9+9+9=27

Table 2: Possible values of c_1 and c_2 for small 3pc_b-graphs. For given m, all c_1-values up to and including m are allowed together with the possible c_2-values given below.

The first time $\ell_1\ell_2+\ell_1\ell_3+\ell_2\ell_3$ has decreased for an increasing sum k, is for $\boldsymbol{\ell} = (6, 2, 1)$ compared to $\boldsymbol{\ell}' = (3, 3, 2)$. The corresponding graphs are illustrated in Figure 7.2. For the $(8, 9)$-graph G with 3 parallel chains of length $(6, 2, 1)$ we have $(c_1, c_2) = (9, 20)$. For a $(8, 9)$-graph G' with 3 parallel chains of length $(3, 3, 2)$ and a leaf attached to any vertex, we have $(c_1, c_2) = (8, 21)$. By Corollary 3, $G' \prec_p G$ for large p but $G \prec_p G'$ for small p. □

Theorem 2. *The stability ordering of the* $(n, n + 1)$*-graphs is* p*-dependent when* $n \geq 7$*, but independent of* p *when* $n < 7$*.*

Proof. To see that the stability ordering of the $(n, n + 1)$-graphs is p-dependent when $(n, m) = (7, 8)$, but not for smaller (n, m), we can start by considering Table 1 and Table 2 together. Let us locate the first c_2-value which is strictly smaller than some c_2-value in a lower c_1-column. This is $c_2 = 3 \cdot 5 = 15$ for $c_1 = 8$ in Table 1, compared to $c_2 = 16$ for $c_1 = 7$ in Table 2.

Let G be a figure-eight graph with $(\ell_1, \ell_2) = (3, 5)$ and let G' be a 3pc_b-graph with chains (3,2,2) and one leaf, as illustrated in Figure 7.3. Then $(c_1, c_2) = (8, 15)$, $(c'_1, c'_2) = (7, 16)$ and by Corollary 3 $G' \prec_p G$ for large p but $G \prec_p G'$ for small p.

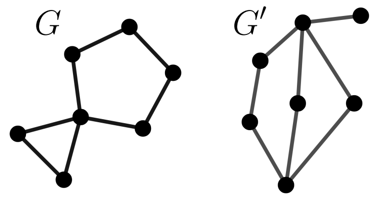


Figure 7.3: The smallest $(n, n+1)$-graphs $(n = 7)$ whose relative stability is p-dependent. G is more stable than G' for large p but less stable for small p. The position of the leaf on G' does not matter.

Before we conclude, we should consider whether we might have missed any relevant $(n, n + 1)$-graphs. Strip a general connected $(n, n + 1)$-graph of its branches, and what remains is a handcuff graph or a 3pc-graph. Therefore, the only graphs which are not included in the preceding analysis are handcuff graphs with with branches. However, for every n-handcuff graph H with branches, there is an n-handcuff graph H' without branches which is equal to H in the stability ordering. To construct H' from H, move all the edges and vertices belonging to branches of H to the ℓ_b-chain (ℓ_b might be zero in H). Therefore, the handcuff graphs with branches do not introduce any further complications in the stability ordering. □

7.4 Optimality for large and small p

Definition 32 (Λ)**.** The greatest edge connectivity among the set of (n, m)-graphs is denoted $\Lambda(n, m)$ (or simply Λ, with (n, m) implicit).

Definition 33 ($\mathcal{G}_\Lambda$)**.** The collection of (n,m)-graphs G for which $\lambda(G) = \Lambda(n,m)$ is denoted $\mathcal{G}_\Lambda$.

Lemma 19. *For p sufficiently large, the (n,m,p)-optimal graphs are contained in the set of graphs in $\mathcal{G}_\Lambda$ which maximize $c_\Lambda(G)$.*

Proof. Let G' be an (n,m)-graph not in $\mathcal{G}_\Lambda$ and let $G \in \mathcal{G}_\Lambda$. Then $\lambda(G') =: \lambda'$ is strictly less than $\lambda(G) = \Lambda$. Since G' can be disconnected by removing λ' edges, but G cannot, we have $c_{\lambda'}(G') < c_{\lambda'}(G)$. According to Lemma 18, G is strictly more stable than G' for large p.

Furthermore, if G^1 and G^2 are two graphs in $\mathcal{G}_\Lambda$, and $c_\Lambda(G^1) > c_\Lambda(G^2)$, then it also follows from Lemma 18 that G^1 is strictly more stable than G^2, for p sufficiently large. □

There is also a corresponding result for small p, which follows directly from Lemma 18:

Lemma 20. *For p sufficiently small, the (n,m,p)-optimal graphs are a subset of the (n,m)-graphs G that maximize $c_\kappa(G)$.*

7.5 Ordering the D2-graphs

In this section we reconsider the D-graphs (see Figure 6.1), particularly the D2-graphs. We describe the unique p-optimal D2_m-graph when p is sufficiently small and the unique p-optimal D2_m-graph for p sufficiently large. In general, these two graphs are not the same.

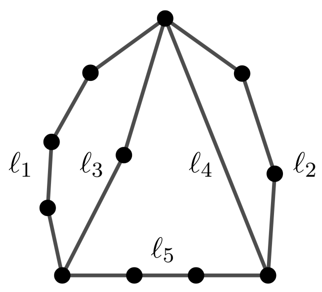


Figure 7.4: A D2_{13}-graph with chains labeled according to the conventions of this section.

Consider the formula (4.1) for $P_c(G,p)$. The D-graphs have edge connectivity 2 (all except K^4), which means that the first variable coefficient is c_2. By Lemma 3 the last nonzero coefficient is number $m - n + 1 = 3$, i.e c_3.

$$P_c(G,p) = p^m + mp^{m-1}q + c_2p^{m-2}q^2 + c_3p^{m-3}q^3 \,. \tag{7.6}$$

In the last chapter it became quite clear that for D1- and D4-graphs, both c_2 and c_3 are maximized by graphs with AEAP chains. When it comes to D2-graphs, this is true for c_2, but when c_3 is maximized remains uninvestigated. We reproduce (6.8) below (slightly more factorized), giving $c_3(G)$ for D2-graphs:

$$c_3 = \ell_2\ell_3\ell_4 + \ell_1\ell_3\ell_4 + \ell_1\ell_2\ell_4 + \ell_1\ell_2\ell_3 + \ell_5(\ell_1 + \ell_3)(\ell_2 + \ell_4)\,. \tag{7.7}$$

From now on we assume that all graphs of this section are D2-graphs. In labeling the chains, we will adopt the convention that $\ell_1 \geq \ell_3$, $\ell_2 \geq \ell_4$ and $\ell_1 + \ell_3 \geq \ell_2 + \ell_4$ (see Figure 7.4 for an example). This makes for a one-one correspondence between the D2-graphs and its set of permitted chain tuples. (Exception: When a graph G is modified into another graph G', for example by shortening the first chain, the convention might break for G'. We will not cause confusion by relabeling the chains.)

Lemma 21. *Consider all* $\mathrm{D2}_m$*-graphs with a fixed chain length* ℓ_5*. The coefficient* $c_3(G)$ *is strictly maximized for the unique graph* G *for which* $\simeq(\ell_1, \ell_2, \ell_3, \ell_4)$ *and* $\simeq(\ell_1+\ell_3, \ell_2+\ell_4)$*. Equivalently,* G *is the graph for which the first four chains are AEAP and satisfy* $\ell_1 \geq \ell_2 \geq \ell_3 \geq \ell_4$*.*

Proof. We start by observing that at least one chain in each chain pair (ℓ_1, ℓ_3) and (ℓ_2, ℓ_4) has to have length at least 2, $\ell_1 + \ell_2 + \ell_3 + \ell_4 \geq 6$. Therefore, it is necessary for a D2-graph that $m \geq 7$ and that $\ell_5 \in [1\,..\,m-6]$.

Since m and ℓ_5 are fixed, $(\ell_1+\ell_3)+(\ell_2+\ell_4)$ is also fixed. The last term of (7.7), call it s_2, is obviously strictly maximized when $\simeq(\ell_1 + \ell_3, \ell_2 + \ell_4)$. Let us call the expression for first four terms s_1. Algebraically, $s_1 = s_1(\ell_1, \ell_2, \ell_3, \ell_4)$ satisfies the conditions of Lemma 7 for each pair of chains. Therefore, s_1 is strictly maximized when $\simeq(\ell_1, \ell_2, \ell_3, \ell_4)$.

Since the condition for maximizing s_1 and the one for maximizing s_2 are not mutually exclusive, c_3 is maximized if and only if both are fulfilled.

Now, take a $\mathrm{D2}_m$-graph G for which $\simeq(\ell_1, \ell_2, \ell_3, \ell_4)$ and $\simeq(\ell_1 + \ell_3, \ell_2 + \ell_4)$. By our chain labeling convention, $\ell_1 \geq \ell_3$, $\ell_2 \geq \ell_4$ and $\ell_1 + \ell_3 \geq \ell_2 + \ell_4$. Let $k = \ell_4$ and consider two different cases: 1) $(\ell_2, \ell_4) = (k, k)$ and 2) $(\ell_2, \ell_4) = (k+1, k)$. The only ℓ_1- and ℓ_3-values consistent with the above requirements are either $(\ell_1, \ell_3) = (k, k)$ or $(\ell_1, \ell_3) = (k+1, k)$ in the first case, and $(\ell_1, \ell_3) = (k+1, k+1)$ or $(\ell_1, \ell_3) = (k+1, k)$ in the second. In all four cases, $\ell_1 \geq \ell_2 \geq \ell_3 \geq \ell_4$. The chain lengths can now be obtained as one of the four alternatives above by considering whether the fixed number $m - \ell_5$ equals $4k$, $4k+1$, $4k+2$ or $4k+3$. This proves that G is uniquely determined. □

The question is now, given m, which ℓ_5-length(s) is (are) optimal for maximizing c_3?

Lemma 22. *The following is necessary for* G *to maximize* $c_3(G)$ *(restricted to* $\mathrm{D2}_m$*): That the chains of* G *satisfy* $\simeq(\ell_1, \ell_2, \ell_3, \ell_4)$*,* $\simeq (\ell_1 + \ell_3, \ell_2 + \ell_4)$ *and that*

$$\ell_5 \in \left[\ell_1 - 1 - \frac{\ell_2\ell_4}{\ell_2+\ell_4},\ \ \ell_4 + 1 - \frac{\ell_1\ell_3}{\ell_1+\ell_3}\right]. \tag{7.8}$$

Proof. Let $\Gamma_m \subset \mathrm{D2}_m$ be the set of graphs satisfying $\simeq(\ell_1, \ell_2, \ell_3, \ell_4)$ and $\simeq (\ell_1+\ell_3, \ell_2+\ell_4)$. By Lemma 21, any graph that maximizes c_3 in $\mathrm{D2}_m$ has to be in Γ_m, and for any $G \in \Gamma_m$ we have $\ell_1 \geq \ell_2 \geq \ell_3 \geq \ell_4$. Furthermore, Γ_m contains exactly one graph for each possible length of ℓ_5. We refer to this graph as $\gamma_m(\ell_5)$.

We first compare $c_3(G)$ where $G = \gamma_m(\ell_5)$ with $c_3(G')$ where $G' = \gamma_m(\ell_5+1)$. To obtain G' from G, we move an edge to ℓ_5 from ℓ_1, which ensures that G' is in Γ_m. (We do not need to worry about relabeling the chains of G'.) Insertion into (7.7) yields

$$\begin{aligned} c_3(G') &= (\ell_1-1)(\ell_2\ell_3+\ell_2\ell_4+\ell_3\ell_4)+\ell_2\ell_3\ell_4+(\ell_5+1)(\ell_1+\ell_3-1)(\ell_2+\ell_4) \\ &= c_3(G)-(\ell_2\ell_3+\ell_2\ell_4+\ell_3\ell_4)+(\ell_1+\ell_3-1)(\ell_2+\ell_4)-\ell_5(\ell_2+\ell_4)\,, \end{aligned}$$

so that

$$c_3(G')-c_3(G) = (\ell_1-1)(\ell_2+\ell_4)-\ell_2\ell_4-\ell_5(\ell_2+\ell_4)\,. \tag{7.9}$$

We are only interested in the sign of $c_3(G')-c_3(G)$. Divide both sides of (7.9) by $(\ell_2+\ell_4)$ and define

$$f_1(G) = (\ell_1-1)-\frac{\ell_2\ell_4}{\ell_2+\ell_4} \tag{7.10}$$

so that $\mathrm{sgn}(c_3(G')-c_3(G)) = \mathrm{sgn}(f_1(G)-\ell_5)$. Now, if $\ell_5 < f_1(G)$, then $c_3(G') > c_3(G)$, which implies that G does not maximize c_3.

Secondly, we compare $c_3(G)$ with $c_3(G'')$ where $G'' = \gamma_m(\ell_5-1)$. To obtain G'' from G, we move an edge from ℓ_5 to ℓ_4, which ensures that G'' is in Γ_m. Insertion into (7.7) yields

$$c_3(G'') = (\ell_4+1)(\ell_1\ell_2+\ell_2\ell_3+\ell_1\ell_3)+\ell_1\ell_2\ell_3+(\ell_5-1)(\ell_1+\ell_3)(\ell_2+\ell_4+1)\,,$$

so that

$$\begin{aligned} c_3(G'')-c_3(G) &= (\ell_1\ell_2+\ell_2\ell_3+\ell_1\ell_3)-(\ell_1+\ell_3)(\ell_4+1)+\ell_5(\ell_1+\ell_3) \\ &= -(\ell_4+1)(\ell_1+\ell_3)+\ell_1\ell_3+\ell_5(\ell_1+\ell_3)\,. \end{aligned} \tag{7.11}$$

Now, $c_3(G'') > c_3(G)$ if and only if

$$(\ell_4+1)(\ell_1+\ell_3)-\ell_1\ell_3-\ell_5(\ell_1+\ell_3) < 0\,.$$

As above, define

$$f_2(G) = \ell_4+1-\frac{\ell_1\ell_3}{\ell_1+\ell_3}\,. \tag{7.12}$$

If $\ell_5 > f_2(G)$, then $c_3(G'') > c_3(G)$ and G does not maximize c_3. Therefore, it is necessary that $\ell_5 \in [f_1(G), f_2(G)]$ for c_3 to be maximal. □

The reader is reminded of Definition 17; a p-optimal $\mathrm{D2}_m$-graph is essentially "a graph which is (n,m,p)-optimal if only $\mathrm{D2}_m$-graphs are considered".

Theorem 3. *For each $m \geq 7$ there is a graph $G \in \mathrm{D2}_m$ which uniquely maximizes $c_3(G)$ in $\mathrm{D2}_m$, and therefore also is the unique p-optimal $\mathrm{D2}_m$-graph when p is sufficiently small. G is the graph whose chains satisfy $\simeq(\ell_1,\ell_2,\ell_3,\ell_4,2\ell_5)$ and $\simeq(\ell_1+\ell_3,\ell_2+\ell_4)$, except if $m = 9t+1$ for some $t \in \mathbb{Z}^+$, in which case G is the graph with chains $\boldsymbol{\ell} = (2t,2t,2t,2t,t+1)$.*

There is also, for each $m \geq 7$, a graph $H \in \mathrm{D2}_m$ which is the unique p-optimal $\mathrm{D2}_m$-graph when p is sufficiently large. H is the graph with AEAP chains satisfying $\simeq(\ell_1+\ell_3,\ell_2+\ell_4)$ and $\ell_5 \leq \ell_{1,2,3,4}$.

Proof of the first part. Let G be a D2$_m$-graph that maximizes c_3. According to Lemma 21, $\simeq(\ell_1, \ell_2, \ell_3, \ell_4)$ and $\ell_1 \geq \ell_2 \geq \ell_3 \geq \ell_4$. Let $k = \ell_4$. There are then four possibilities for $(\ell_1, \ell_2, \ell_3, \ell_4)$:

$$\begin{aligned} &1)\ \ (k, k, k, k) \\ &2)\ \ (k+1, k, k, k) \\ &3)\ \ (k+1, k+1, k, k) \\ &4)\ \ (k+1, k+1, k+1, k)\,. \end{aligned} \tag{7.13}$$

Correspondingly, there are four different formulas for the allowed interval for ℓ_5 according to Lemma 22. Call the interval $[f_1(k), f_2(k)]$. We also name the recurring expression $\epsilon_k = 1/(4 + 2/k)$, and note that $\epsilon_k \in [1/6, 1/4)$ for all k.

Case 1: With the first four chains equal to k, $f_1(k) = k - 1 - k^2/(2k) = k/2 - 1$, and $f_2(k) = k/2 + 1$. Obviously, ℓ_5 has to be an integer. The interval $[k/2 - 1, k/2 + 1]$ contains 3 integers if k is even and two if k is odd.

Case 2: The chain lengths $(k + 1, k, k, k)$ give $f_1(k) = k - k/2 = k/2$, and

$$\begin{aligned} f_2(k) &= k + 1 - \frac{(k+1)k}{2k+1} \\ &= \frac{k}{2} + 1 + \frac{k(2k+1)}{2(2k+1)} - \frac{2(k+1)k}{2(2k+1)} \\ &= k/2 + 1 - k/(4k+2) = k/2 + 1 - \epsilon_k\,. \end{aligned} \tag{7.14}$$

The interval $[k/2, k/2 + 1 - \epsilon_k]$ has length $1 - \epsilon_k$, and contains exactly one integer: $\lceil k/2 \rceil$.

Case 3: The chain lengths $(k + 1, k + 1, k, k)$ give

$$f_1(k) = k - \frac{(k+1)k}{2k+1} = k/2 - \epsilon_k\,, \tag{7.15}$$

and $f_2(k)$ is unchanged; $f_2(k) = k/2 - \epsilon_k$. The interval $[k/2 - \epsilon_k, k/2 + 1 - \epsilon_k]$ has length one, and contains exactly one integer: $\lceil k/2 \rceil$.

Case 4: The chain lengths $(k + 1, k + 1, k + 1, k)$ give the same f_1 as above; $f_1(k) = k/2 - \epsilon_k$, and

$$f_2(k) = k + 1 - \frac{(k+1)^2}{2k+2} = k + 1 - \frac{k+1}{2} = \frac{k}{2} + \frac{1}{2}\,. \tag{7.16}$$

The interval $[k/2 - \epsilon_k, k/2 + 1/2]$ contains only the integer $\lceil k/2 \rceil$.

Since the possible ℓ_5-values above depend upon whether k is even or odd, we separately set first $k = 2t$, then $k = 2t + 1$, and list all the possible chain tuples $\boldsymbol{\ell} = (\ell_1, \ell_2, \ell_3, \ell_4, \ell_5)$ that are compatible with the above calculations, in order, together with the resulting chain sums m.

#	$(\ell_1, \ell_2, \ell_3, \ell_4, \ell_5)$	m
~~1~~	$(2t, 2t, 2t, 2t, t-1)$	$9t-1$
2	$(2t, 2t, 2t, 2t, t)$	$9t$
3	$(2t, 2t, 2t, 2t, t+1)$	$9t+1$
~~4~~	$(2t+1, 2t, 2t, 2t, t)$	$9t+1$
5	$(2t+1, 2t+1, 2t, 2t, t)$	$9t+2$
6	$(2t+1, 2t+1, 2t+1, 2t, t)$	$9t+3$
7	$(2t+1, 2t+1, 2t+1, 2t+1, t)$	$9t+4$
8	$(2t+1, 2t+1, 2t+1, 2t+1, t+1)$	$9t+5$
9	$(2t+2, 2t+1, 2t+1, 2t+1, t+1)$	$9t+6$
10	$(2t+2, 2t+2, 2t+1, 2t+1, t+1)$	$9t+7$
11	$(2t+2, 2t+2, 2t+2, 2t+1, t+1)$	$9t+8$

The first thing to notice is that most m-values correspond to only one possible chain tuple. We have therefore proved, for example, that if m can be written as $9t+2$ for some integer t, the D2-graph that maximizes $c_3(G)$ is the graph with chains $\boldsymbol{\ell} = (2t+1, 2t+1, 2t, 2t, t)$. The second thing to notice is that the first and last line actually correspond to the same set of graph sizes. We might therefore want to rewrite the last line as:

$11'$	$(2t, 2t, 2t, 2t-1, t)$	$9t-1$

The third thing to note is that all lines except the first and third correspond to the unique graph (for fixed m) that satisfy $\simeq(\ell_1, \ell_2, \ell_3, \ell_4, 2\ell_5)$ in addition to $\simeq(\ell_1+\ell_3, \ell_2+\ell_4)$.

Now, when m is taken modulo 9, the two cases $m \equiv 1$ and $m \equiv -1$ each gives two alternative graphs, and we don't know if either one or both of these graphs maximize c_3. The straightforward way to find out is to substitute the chain lengths into (7.7) and compare the results:

Let G' be a graph with $\boldsymbol{\ell'} = (2t, 2t, 2t, 2t, t-1)$ (according to the first row of the table) and suppose that G has chains $\boldsymbol{\ell} = (2t, 2t, 2t, 2t-1, t)$ (according to row 11') for some positive t. (If $t=1$ only G exists). Insertion into (7.7) yields $c_3(G') = 3(2t)^3 + (t-1)(4t)(4t) = 40t^3 - 16t^2$ and $c_3(G) = (2t)^3 + 2\cdot(2t)^2(2t-1) + t\cdot 4t(4t-1) = 40t^3 - 12t^2$. Since $c_3(G) > c_3(G')$ we discard the first row.

Now suppose that G has chains $\boldsymbol{\ell} = (2t, 2t, 2t, 2t, t+1)$ (row 3) and let G' have $\boldsymbol{\ell} = (2t+1, 2t, 2t, 2t, t)$ (row 4). Glancing at the previous calculations we get $c_3(G) = 40t^3 + 16t^2$ and $c_3(G') = 40t^3 + 12t^2$. Since $c_3(G) > c_3(G')$ we keep row 3 and discard row 4 (surprisingly, perhaps, since row 3 breaks the pattern of $\simeq(\ell_1, \ell_2, \ell_3, \ell_4, 2\ell_5)$, compared to the other cases).

Thus, G, as specified in the theorem, is the unique $D2_m$ graph that maximizes $c_\kappa(G) = c_3(G)$. That G also is the unique p-optimal $D2_m$-graph for sufficiently small p follows from Lemma 20, when restricted to $D2_m$-graphs.

Proof of the second part. By Lemma 9, c_2 restricted to $D2_m$ is strictly maximized for graphs with AEAP chains. A p-optimal $D2_m$-graph must therefore have AEAP chains when p is large, according to Lemma 19.

Suppose that H is as specified in the theorem and that $H' \in \mathrm{D2}_m$ is another graph with AEAP chains, but distinct from H. Since $c_2(H) = c_2(H')$, we have

$$P_c(H,p) - P_c(H',p) = p^{m-3}q^3(c_3(H) - c_3(H'))\,.$$

If c_3 is maximized by H alone among the $\mathrm{D2}_m$-graphs with AEAP chains, then the above is positive and we are done.

Luckily, there are only two ways (not mutually exclusive) in which H' can differ from H. Let all chains refer to H'. Case 1: ℓ_5 is strictly longer than some other chain in H', so that by our conventions $\ell_5 = \ell_4 + 1$. Case 2: H' has $\not\simeq(\ell_1 + \ell_3, \ell_2 + \ell_4)$.

Suppose that the former holds. Obtain H'' by moving one edge from ℓ_5 to ℓ_4. We have done this before, in (7.11). From that equation we obtain $c_3(H'') - c_3(H') = (\ell_5 - (\ell_4 + 1))(\ell_1 + \ell_3) + \ell_1\ell_3$, which equals $\ell_1\ell_3$. In this case, H' clearly does not maximize c_3. Suppose therefore that ℓ_5 is the same in H' as in H, and consider case 2. By Lemma 21, $c_3(H) > c_3(H')$. □

Remark. According to the theorem, the p-optimal $\mathrm{D2}_m$-graph for large p has ℓ_5 AEAP to the other chains, but the p-optimal $\mathrm{D2}_m$-graph for small p has chain ℓ_5 equal to about half the length of the other chains (either $\simeq(2\ell_5, \ell_i)$ or $\ell_5 = \ell_i/2 + 1$, for $i \in [1\,..\,4]$). For reasonably large m, these numbers are not the same.

The smallest (n, m) for which the graphs differ is $(9, 11)$: the p-optimal $\mathrm{D2}^9_{11}$-graph has chains $(3, 2, 3, 2, 1)$ for small p but chains $(3, 2, 2, 2, 2)$ for large p.

The p-optimal graphs for small and large p are subsequently different for all $(n, n+2)$ greater than $(9, 11)$, except for the cases $(n, m) = (12, 14)$ and $(n, m) = (17, 19)$. In the latter case, the p-optimal $\mathrm{D2}_{19}$-graph is the graph with chains $(4, 4, 4, 4, 3)$, for all p.

Bibliography

[1] Noga Alon, Itai Benjamini, and Alan Stacey. "Percolation on Finite Graphs and Isoperimetric Inequalities." In: *Ann. Probab.* 32.3A (2004), pp. 1727–1745. DOI: https://doi.org/10.1214/009117904000000414.

[2] Reinhard Diestel. *Graph Theory*. Graduate Texts in Mathematics. Springer, 2005. ISBN: 9783662536216.