

# Random Lifts of Graphs

Thesis submitted for the degree

**Doctor of Philosophy**

by

Alon Amit

Submitted to the Senate of the Hebrew University

July 2000

This work was carried under the supervision of  
Professor Nathan Linial

# Contents

|                                               |           |
|-----------------------------------------------|-----------|
| <b>Abstract</b>                               | <b>v</b>  |
| <b>1 Basic Definitions and Connectivity</b>   | <b>1</b>  |
| 1.1 Introduction . . . . .                    | 1         |
| 1.1.1 Lifts . . . . .                         | 2         |
| 1.1.2 The Model . . . . .                     | 5         |
| 1.1.3 The Role of Labels . . . . .            | 6         |
| 1.2 Connectivity . . . . .                    | 8         |
| 1.2.1 Preliminaries . . . . .                 | 9         |
| 1.2.2 Lifts of $K_2(\alpha)$ . . . . .        | 11        |
| 1.2.3 Thick Sets . . . . .                    | 14        |
| 1.2.4 Thin Sets . . . . .                     | 18        |
| 1.2.5 Conclusion . . . . .                    | 21        |
| 1.3 Isomorphisms and Labeling . . . . .       | 22        |
| 1.3.1 The Number of Unlabeled Lifts . . . . . | 23        |
| 1.3.2 Asymptotics . . . . .                   | 25        |
| 1.3.3 Equivalence of the Models . . . . .     | 27        |
| 1.4 Variations . . . . .                      | 29        |
| <b>2 Edge Expansion of Random Lifts</b>       | <b>31</b> |
| 2.1 Introduction . . . . .                    | 31        |

|          |                                                                |           |
|----------|----------------------------------------------------------------|-----------|
| 2.2      | General Graphs . . . . .                                       | 33        |
| 2.2.1    | The Second Eigenvalue . . . . .                                | 37        |
| 2.3      | Random lifts of $B_l$ . . . . .                                | 38        |
| 2.4      | An Improved Bound Through $\epsilon$ -nets . . . . .           | 45        |
| 2.4.1    | Random Regular Graphs . . . . .                                | 45        |
| 2.4.2    | $\epsilon$ -nets . . . . .                                     | 46        |
| <b>3</b> | <b>Independence and Chromatic Number</b>                       | <b>51</b> |
| 3.1      | Introduction . . . . .                                         | 51        |
| 3.2      | The Independence Number . . . . .                              | 52        |
| 3.2.1    | A Greedy Lower Bound . . . . .                                 | 55        |
| 3.2.2    | The Independence Number of Lifts of Complete Graphs .          | 57        |
| 3.3      | The Chromatic Number . . . . .                                 | 59        |
| 3.3.1    | Lower Bounds . . . . .                                         | 60        |
| 3.3.2    | $\tilde{\chi}_l(G)$ and Variations on Graph Coloring . . . . . | 64        |
| 3.3.3    | An Upper Bound . . . . .                                       | 66        |
| 3.3.4    | Persistence of the Chromatic Number . . . . .                  | 68        |
| 3.4      | Conclusion . . . . .                                           | 70        |
| <b>A</b> | <b>The Chromatic Number of <math>D_s</math></b>                | <b>73</b> |

# Abstract

The notion of covering maps between graphs is essentially a restriction to the case of graphs (as, say, one dimensional simplicial complexes) of the general topological notion of covering map. It is described in purely combinatorial terms as a mapping of graphs that maps the neighbours of a vertex one-to-one onto the neighbours of its image vertex (We will later refine this slightly to properly account for multiple edges and loops).

Covering maps have received considerable attention from several different points of view. For example, Leighton’s remarkable theorem on common finite coverings [17] was inspired by Angluin’s application of coverings to problems in multi-processor networks [4], and in turn was found to have connections with deep algebraic questions related to groups acting on trees (Bass and Kulkarni [6]). Graph coverings are naturally related to subgroups of free groups; see Stillwell [26] and Stallings [25], for example. Finally, as purely graph theoretic objects, they were studied by Gross and Tucker [11], [12], Negami [22] and Archdeacon and Richter [5], among others.

The main theme of this work is the introduction of a probabilistic structure on the set of graphs that cover a fixed base graph. We call such graphs *lifts* of the base graph, to distinguish from other uses of the term “cover” in graph theory. This point of view enables one to apply the powerful probabilistic method to questions concerning lifts. We may investigate what a *typical* lift looks like, and

for instance prove the existence of lifts having properties that may be hard to construct explicitly. Some results of this kind appear in chapter 2.

Moreover, we can “forget” that the random graphs we generate cover some base graph, and view our model as some sort of random model for (a special class of) graphs. Inasmuch as the probabilistic method is used to prove existence theorems of graphs with extremal properties, a new model for random graphs (complete with “knobs” we can fiddle with) may be a useful addition to our arsenal.

In a different direction, we may hope that the model can be used to “randomize” a given graph, in the following sense. Suppose we wish to calculate, or approximate, some invariant of a graph  $G$ . It is often the case that the invariant is hard to determine efficiently, but an efficient algorithm is known to work on a *random* graph. We can now run the algorithm on a random lift  $\tilde{G}$  of  $G$ , hoping that it is “sufficiently random” so that the algorithm runs reliably, but still “sufficiently similar” to  $G$  itself so the result is close to the original invariant. We hope to pursue this idea in the future. Some promising first steps are the results on the chromatic number in chapter 3.

In chapter 1 we introduce the model and prove the first property of random lifts: a random lift of a graph  $G$  is almost surely  $\delta(G)$ -connected, provided  $\delta(G) \geq 3$ . Here  $\delta(G)$  is the minimal degree of  $G$ . In addition, we investigate the relation between our random model which actually generates *labeled* random lifts, and the more naturally defined *unlabeled* lifts. This is similar to the results on labeled and unlabeled random graphs.

In chapter 2 we turn to edge expansion properties. We prove that random lifts of a fixed base graph have edge expansion bounded away from 0. We then obtain numerical bounds on the edge expansion of lifts of the bouquet graphs  $B_l$ . Using the method of  $\epsilon$ -nets, we improve the naive union bound usually used in such investigations and exhibit better bounds for the edge expansion of regular graphs than were available before.

In the final chapter we study the independence number and chromatic number of random lifts. For the independence number, upper and lower bounds are obtained as solutions to certain optimization problems on the base graph. For a graph  $G$  with chromatic number  $\chi$  and fractional chromatic number  $\chi_f$ , we show that the chromatic number of typical lifts is bounded from below by  $\text{const} \cdot \sqrt{\chi/\log \chi}$  and also by  $\text{const} \cdot \chi_f/\log^2 \chi_f$ . We have examples of graphs where the chromatic number of the lift equals  $\chi$  almost surely, and others where it is a.s.  $O(\chi/\log \chi)$ .



# Chapter 1

## Basic Definitions and Connectivity

We begin by describing our simple model for random graphs that have an  $n$ -fold covering map onto a fixed finite base graph. Roughly, given a base graph  $G$  and an integer  $n$ , we form a random graph by replacing each vertex of  $G$  by a set of  $n$  vertices, and joining these sets by random matchings whenever the corresponding vertices are adjacent in  $G$ . The resulting graph covers (is a lift of) the original graph in the sense that the two are locally isomorphic. The main specific result that we prove in this chapter (Theorem 1) is that if  $\delta \geq 3$  is the smallest vertex degree in  $G$ , then almost all  $n$ -lifts of  $G$  are  $\delta$ -connected. We also discuss the important connection between the labeled and unlabeled random models.

### 1.1 Introduction

In this section we describe the model and its basic properties. Section 1.2 contains our first application of the model, showing that almost every lift of

a graph with minimal degree  $\delta$  is  $\delta$ -connected. In section 1.3 we return to the model itself, showing the equivalence between the models of labeled and unlabeled lifts. Finally in Section 1.4 we present variations and generalizations of the model, some of which we will utilize in forthcoming papers.

### 1.1.1 Lifts

Let  $G$  and  $\tilde{G}$  be graphs (possibly with multiple edges and loops), and let  $\pi : \tilde{G} \rightarrow G$  be a graph homomorphism, namely a pair of maps  $V(\tilde{G}) \rightarrow V(G)$  and  $E(\tilde{G}) \rightarrow E(G)$  that preserve adjacency.  $\pi$  is a *covering* if the star of each vertex  $\tilde{v} \in \tilde{G}$  is mapped bijectively to the star of its image  $\pi\tilde{v}$ , where a star is the collection of edges emanating from a vertex. A loop is considered as *two* edges in the star.

We call  $G$  the *base graph* and the inverse image  $\pi^{-1}(v)$  is called a *fiber*, denoted  $\tilde{G}_v$ . We use suggestive terms such as “ $\tilde{v}$  lies above  $v$ ” when  $\pi\tilde{v} = v$  etc. Indeed it is best to visualize  $\tilde{G}_v$  as a vertical stack of vertices above  $v$ . It is easy to see that the inverse image of an edge  $[u, v]$  sets a bijective correspondence between  $\tilde{G}_u$  and  $\tilde{G}_v$ , since each  $\tilde{u}$  in  $\tilde{G}_u$  is connected to precisely one  $\tilde{v}$  in  $\tilde{G}_v$ .

In particular all the fibers have the same cardinality (when the base graph is connected). This common cardinality is called the *degree* of the covering; if this is finite and equal to  $n$  we call  $\pi$  an  $n$ -covering. Every path in  $G$  is covered by  $n$  disjoint paths in  $\tilde{G}$ , called the *liftings* of the path. This is the well-known *unique path-lifting* property of coverings.

We emphasize that a covering, as an object, consists of the graph  $\tilde{G}$  *together* with the covering projection  $\pi : \tilde{G} \rightarrow G$ . Two coverings  $\pi : \tilde{G} \rightarrow G$  and  $\pi' : \tilde{G}' \rightarrow G$  are considered isomorphic if there is an isomorphism  $f : \tilde{G} \rightarrow \tilde{G}'$  that respects the coverings, namely  $\pi = \pi' \circ f$ . An automorphism is of course an isomorphism from a covering to itself, and a covering is *regular* (or *normal*) if its automorphism group acts transitively on the fibers.

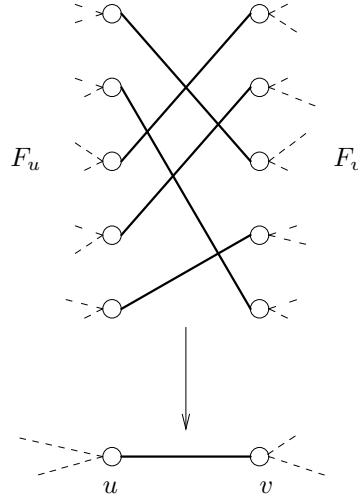
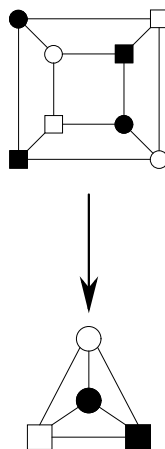


Figure 1.1: Typical adjacent fibers

To avoid confusion with other notions of covers in graph theory, we call a graph  $\tilde{G}$  a *lift* of  $G$  if there is a covering map  $\tilde{G} \rightarrow G$ . Figure 1.1 shows part of  $G$  and the part of a typical lift  $\tilde{G}$ .

Here are some examples of lifts:

- The graph  $Q_3$  of the three dimensional cube is a lift of  $K_4$ . In Figure 1.2, the vertices of each fiber are equally colored.
- The cycle  $C_n$  is a lift of  $C_m$  iff  $m|n$ . This holds even for  $m = 1$ , since we count loops twice.
- The infinite path  $P_\infty$  (The integers  $\mathbb{Z}$  with edges  $[n, n + 1]$ ) covers every cycle  $C_n$ .
- More generally, the  $d$ -regular tree covers every  $d$ -regular graph. This is the *universal cover* of a  $d$ -regular graph. A universal cover exists for every graph, and it is always a tree (unless the base graph is not connected).
- Every regular graph  $G$  of even degree  $2m$  covers the bouquet  $B_m$  of  $m$

Figure 1.2:  $Q_3$  covers  $K_4$ 

circles, which is the graph with one vertex and  $m$  loops. To see this, take a 2-factorization of  $G$  (which exists by a theorem of Petersen. See, for example, [7] ); there are  $m$  factors, and we map all the edges in a factor to the same loop in  $B_m$ . Incidentally, this proves that every regular graph of even degree is a Schreier coset graph (See [13])

- Figure 1.3 is an example of an *irregular*, or *non-normal*, covering. Irregularity refers to the fact that the automorphism group of this covering does not act transitively on the fibers. In topology and algebra, regular coverings (and normal field extensions) are more well-behaved, but one should keep in mind that for random coverings this is the exception rather than the rule: most coverings have no automorphisms, and are thus irregular. See 1.3.
- An interesting example of an infinite covering is the covering of  $B_2$  by the 2-dimensional grid graph. In terms of the fundamental group, this covering realizes the commutator subgroup of  $F_2$ , and furnishes a proof that this group is not finitely generated (see [26]).

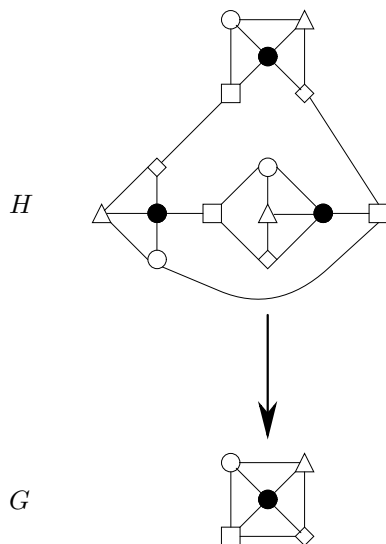


Figure 1.3: An Irregular Covering

### 1.1.2 The Model

Given the base graph  $G$  and the degree  $n$ , we can construct a lift  $\tilde{G}$  by putting  $n$  vertices  $v_1, \dots, v_n$  above every vertex  $v$  of  $G$ . For every edge  $e = [u, v]$ , we need to decide which  $u_i$  is connected to which  $v_j$ . This is determined by a single permutation on  $n$  elements: Given a permutation  $\sigma = \sigma(e)$  we connect  $u_i$  to  $v_{\sigma(i)}$ , and this is done independently for every edge. For example, if we number the vertices in Figure 1.1 from the bottom up, the corresponding permutation has cycle presentation  $(1\ 2\ 4)(3\ 5)$ .

Notice that we need to assign an orientation to  $e$  so that we know which end is  $u$  and which is  $v$ . This choice, however, is arbitrary and has no real effect on the possible outcomes, i.e. reversing the edge and inverting the permutation yield the same lift. Having made this choice, we see that a lift is defined by attaching permutations to the edges of  $G$ . All  $n$ -lifts of  $G$  are obtained in this manner.

It is now fairly obvious how to define a *random*  $n$ -lift: simply choose a

permutation  $\sigma(e) \in S_n$  uniformly and independently for every edge  $e$  in  $G$ , and form the lift  $\tilde{G}$  as above. This is our model for a random *labeled*  $n$ -lift of  $G$  (see 1.1.3). Already we see several ways we can tweak it to suit our needs: we may choose the permutations non-uniformly in  $S_n$ , and possibly not independently. Unless otherwise stated, however, we will work with this standard model. We summarize it in the following definition:

**Definition 1** *Given a graph  $G$ , a random labeled  $n$ -lift of  $G$  is obtained by arbitrarily orienting the edges of  $G$ , choosing a permutation  $\sigma_e$  in  $S_n$  for each edge  $e$  uniformly and independently, and constructing the graph  $\tilde{G}$  with  $n$  vertices  $u_1, \dots, u_n$  for each vertex  $u$  of  $G$  and edges  $e_i = (u_i, v_{\sigma_e(i)})$  whenever  $e = (u, v)$  is an oriented edge. A covering  $\pi : \tilde{G} \rightarrow G$  is defined by  $\pi(u_i) = u$  and  $\pi(e_i) = e$ .*

The random model is well-defined in the context of pseudographs as well. Different permutations can be assigned to parallel edges, and a single permutation is attached to each loop.

### 1.1.3 The Role of Labels

We note that the model actually gives a little more than what we required, since the resulting graph  $\tilde{G}$  is equipped with a labelling  $\{1, \dots, n\}$  of the vertices in each fiber. A graph  $\tilde{G}$  covering  $G$  whose vertices are so labeled is called a *labeled* lift, and it is these objects for which we have a random model. The situation is analogous to that of random graphs, where the standard model is defined for labeled graphs instead of (isomorphism classes of) abstract graphs.

Let  $L_n(G)$  be the set of labeled  $n$ -lifts of  $G$ . The same notation is used also for the probability space we have just defined. Clearly, the probability distribution is uniform. Let  $C_n(G)$  be the set of isomorphism classes of unlabeled  $n$ -lifts of  $G$ . Obviously  $|L_n(G)| = (n!)^e$  where  $e = |E(G)|$  is the number of edges in  $G$ . In Section 1.3 we prove that, when  $v = |V(G)| < e$ ,

$$|C_n(G)| = (1 + o(1))(n!)^{e-v}$$

(For the remaining cases  $e - v = -1, 0$  see Corollary 2). This implies that almost every lift has a trivial automorphism group, and more importantly, that properties of lifts have the same asymptotic distribution in the labeled and unlabeled models. (This is reminiscent of the situation for random graphs  $G(n, p)$ , compare with [8], chapter IX). We may therefore adopt the standard policy of proving results in the labeled model, and viewing them as valid statements in the unlabeled case as well.

In a certain sense, our model may be considered somewhat redundant, since it generates  $n!^e$  labeled lifts that correspond to roughly  $n!^{e-v}$  unlabeled ones. The ratio of  $n!^v$  indicates the number of ways to label a (typical) unlabeled graph. In the standard theory of random graphs, this factor is  $n!$ . We can eliminate this redundancy, as follows.

Suppose a labeled lift  $\tilde{G}$  of  $G$  is given by specifying the permutations  $\sigma_e$  for the edges of  $G$ . Fix an edge  $e = (u, v)$  of  $G$ . We may relabel the vertices in  $\tilde{G}_v$  in such a way that  $\sigma_e$  becomes the identity. We may now relabel another fiber, trivializing or “flattening” another permutation. Continuing in this manner, it is easy to see that we can flatten all the permutations belonging to a spanning tree of  $G$ . Of course we cannot hope to simultaneously flatten edges that form a cycle, so this is best possible.

This process of flattening shows that a completely equivalent model to the one defined above is obtained if one chooses in advance a spanning tree  $T$  of  $G$ , and randomly selects permutations for edges  $e$  *outside*  $T$ , fixing the permutations of edges in  $T$  as the identity. The new model is called the  $T$ -flat model, and the sample space now consists of  $n!^{e-v+1}$  lifts, which is only  $n!$  away from the number of unlabeled lifts.

Any property of unlabeled lifts has the same distribution in the original and the  $T$ -flat model, so we may work with either one. We will not discuss the  $T$ -flat model more formally in this work since we do not use it here, but see the remarks in 1.2.5.

## 1.2 Connectivity

For a fixed base graph  $G$ , what is the typical degree of connectivity of a random (labeled) lift of  $G$ ? If  $\delta = \delta(G)$  is the minimal degree of  $G$ , it is also the minimal degree in every lift of  $G$ . Therefore no lift can have connectivity higher than  $\delta$ . We prove that a random lift is indeed almost surely  $\delta$ -connected when  $\delta \geq 3$ .

**Theorem 1** *Let  $G$  be a connected simple graph with minimal degree  $\delta \geq 3$ . Then with probability  $1 - o(1)$ , a random labeled  $n$ -lift of  $G$  is  $\delta$ -connected.*

The  $o(1)$  term is taken with respect to  $n$ , of course. Here, and throughout, we use the term “almost surely” instead of “with probability  $1 - o(1)$  as  $n$  tends to infinity”, so the theorem says that a random labeled lift is almost surely  $\delta$ -connected. As remarked above (and proved in Section 1.3), the same statement also applies to random unlabeled lifts.

We point out that for some base graphs  $G$ , it is not even true that a random lift is likely to be connected. A trivial example is that of a base graph that is a tree  $T$ : Every lift of  $T$  is disconnected, except for  $T$  itself. Another case is when  $G$  is unicyclic: a tree  $T$  with one extra edge (i.e., a graph with zero Euler characteristic). In this case, a random lift is essentially determined by a single permutation  $\sigma$  (see the discussion of flattening in Section 1.1.3), and it is connected iff  $\sigma$  is cyclic, which happens only with probability  $1/n$ . There is no contradiction, though, since both these examples fail to satisfy  $\delta \geq 3$ . See also the discussion of  $K_2(2)$  and  $K_2(3)$  in Section 1.2.2.

Let us briefly indicate the structure of the proof. After some preliminaries we begin in Section 1.2.2 by considering connectivity properties of lifts of  $K_2(\alpha)$ , the graph with two vertices and  $\alpha$  edges. Usually, such lifts are highly connected: every set of vertices has many neighbours. This is essentially just a property of random independent permutations. The “usually” here refers to the fact that for  $\alpha = 1, 2$  we don’t get high connectivity at all, and even when  $\alpha = 3$ , very

small sets may fail to have enough neighbours. It is this difficulty with seemingly minor cases that makes the proof a bit messy.

Given a general base graph  $G$ , we consider a set  $X$  of vertices in the lift  $\tilde{G}$ , and try to show that it has a large vertex boundary. We look for a topological  $K_2(\alpha)$  inside  $G$  over which  $X$  has enough vertices to apply the previous analysis. This works well for most sets  $X$ , namely those that have at least two vertices in some fiber. The special cases  $\alpha = 1, 2$ , however, require some additional effort. This is handled in Section 1.2.3. Section 1.2.4 deals with “thin” sets, that have at most one vertex in each fiber.

### 1.2.1 Preliminaries

To show that a lift  $\tilde{G}$  is  $\delta$ -connected, we need to show that for every set  $X$  of vertices with  $|X| < |\tilde{G}|/2$ ,  $|\partial X| \geq \delta$  where  $\partial X$  is the set of vertices outside of  $X$  that are adjacent to some vertex in  $X$ . An important characteristic of sets  $X$  in  $\tilde{G}$  is the way they are distributed across the fibers. Let  $X_v = X \cap \tilde{G}_v$  be the set of vertices of  $X$  that lie above a vertex  $v \in V(G)$ , and  $x_v = |X_v|$  its size. We begin with a simple observation concerning these numbers.

Suppose that for some  $u, v$  in  $G$ ,  $x_u - x_v \geq \delta$ . Since  $G$  is connected, there is a  $u$ - $v$  path  $P$  in  $G$ . Consider the liftings of  $P$  to paths in  $\tilde{G}$  that start at a vertex in  $X_u$ . They are disjoint, so at least  $\delta$  of them end in a vertex *outside*  $X_v$ . Therefore each such path contains a vertex in  $\partial X$  (possibly more than one), so  $X$  satisfies  $|\partial X| \geq \delta$  automatically, regardless of the actual lift. In Figure 1.4,  $|\partial X|$  is even larger than  $x_u - x_v$ , since the third (from the top) lifted path contributes more than one vertex to  $\partial X$ .

Thus, in proving  $\delta$ -connectivity, we may restrict to sets  $X$  that are rather evenly distributed across the fibers, in the sense that

$$|x_u - x_v| \leq \delta - 1 \tag{1.1}$$

for every  $u, v \in G$ .

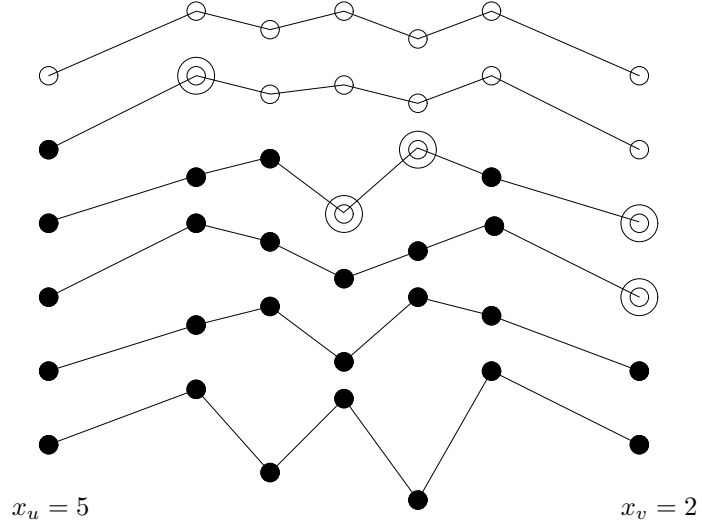


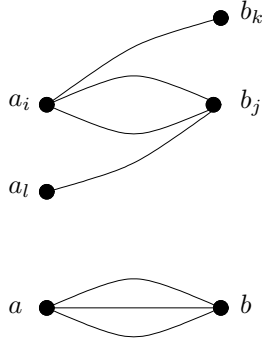
Figure 1.4: An Example with  $x_u - x_v = 3$ . The black vertices are in  $X$ , the white ones outside  $X$ , and the circled white are in  $\partial X$ .

Let  $m = \max\{x_v \mid v \in G\}$ . A set  $X$  for which  $m = 1$  is called *thin*. Thus a thin set is a set of vertices of  $\tilde{G}$  which contains at most one vertex from each fiber.  $X$  is called *thick* if it is not thin.

$\kappa(a, b) = \kappa(a, b; G)$  For distinct vertices  $a, b$  in a graph  $G$ , define their local connectivity as the maximal number of vertex disjoint paths in  $G$  between them. This is denoted  $\kappa(a, b) = \kappa(a, b; G)$ . The maximum local connectivity, i.e. the maximum of  $\kappa(a, b; G)$  over all pairs of vertices of  $G$ , is denoted  $\bar{\kappa}(G)$ . Note that the analogous *minimum* is just the vertex connectivity of  $G$ .

We need the following result of Mader:

**Theorem 2 (Mader [20])** *In every finite graph  $G$  there is an edge  $[a, b]$  such that  $\kappa(a, b) = \min(\deg(a), \deg(b))$ .*

Figure 1.5: A bad set in a lift of  $K_2(3)$ 

### 1.2.2 Lifts of $K_2(\alpha)$

**Lemma 1** *Let  $\tilde{G}$  be a random  $n$ -lift of  $G = K_2(\alpha)$ , the graph with two vertices and  $\alpha$  edges, where  $\alpha \geq 3$ . Then a.s. every subset  $X \subset V(\tilde{G})$  such that  $3 \leq |X| < 2|\tilde{G}|/3$  satisfies  $|\partial X| \geq \alpha$ .*

It is interesting to see why the restrictions in the proposition are necessary. For  $\alpha = 1$ ,  $K_2(1) = K_2$  is a tree with one edge, and every lift is just a matching, so there are many sets  $X$  with  $|\partial X| = 0$ .

For  $\alpha = 2$ , a lift is determined by two permutations  $\sigma_1$  and  $\sigma_2$ , given that an orientation has been assigned for the two edges, say from  $a$  to  $b$  where  $a, b$  are the vertices of  $K_2(2)$ . Suppose that  $i$  is a fixed point of  $\sigma_2^{-1}\sigma_1$ , so that  $\sigma_1(i) = \sigma_2(i) = j$ , say. Then the set  $X = \{a_i, b_j\}$  spans a  $K_2(2)$  in  $\tilde{G}$ , and is disconnected from the rest of the graph. For every constant  $k$  the probability that there are  $k$  such fixed points is bounded away from zero, and so there are likely to be several such small disconnected sets.

When  $\alpha = 3$ , a typical lift is *nearly* 3-connected, but not quite. The lift is determined by three permutations  $\sigma_1, \sigma_2, \sigma_3$ . Consider again an index  $i$  for which  $\sigma_1(i) = \sigma_2(i) = j$ . Then  $a_i$  is connected to  $b_j$  by two parallel edges, and is further connected to  $b_k$  where  $k = \sigma_3(i)$ . Also  $b_j$  is connected to some  $a_l$  where

$l = \sigma_3^{-1}(j)$ . This yields a set  $X = \{a_i, b_j\}$  which only has 2 neighbours, as in Figure 1.5. This shows why  $2 < |X|$  is necessary in Proposition 1. Incidentally, this also demonstrates why simplicity is required in Theorem 1.

*Proof:* (of the Lemma): Let  $\{a, b\}$  be the vertices of  $G = K_2(\alpha)$ . A random lift of  $G$  is determined by  $\alpha$  random permutations  $\sigma_1, \dots, \sigma_\alpha$  in  $S_n$ , by connecting each vertex  $a_i$  to the  $\alpha$  (not necessarily distinct) vertices  $a_{\sigma_p(i)}$ ,  $1 \leq p \leq \alpha$ .

For a subset  $A \subset \{1, \dots, n\}$ , let  $A_p = \sigma_p(A)$  and  $A_* = \cup_p A_p$ . The letter  $m$  always denotes the size of  $A$ . Call a set  $A$  *bad* if  $2 \leq m \leq 3n/4$  and  $|A_*| < m + \alpha - \epsilon$  where  $\epsilon$  is 0 unless  $m = 2$  and  $\alpha = 3$ , in which case it is 1. (The role of  $\epsilon$  will soon be revealed). We claim that, if the  $\sigma_p$ s are chosen uniformly and independently in  $S_n$ , then a.s. there are no bad sets.

*Proof of claim:* We start with small sets  $A$ , those for which  $m \leq 10\alpha$ , say. Given  $A$ , the probability that  $|A_*| \leq t$  for some  $t$  is not greater than the probability that  $s = \alpha m$  balls that are thrown into  $n$  cells occupy no more than  $t$  cells, which is bounded by

$$\binom{n}{t} \left(\frac{t}{n}\right)^s.$$

Since there are  $\binom{n}{m}$  sets  $A$  with  $|A| = m$ , to show that no set of size  $m$  is bad a.s. it suffices to verify that

$$\binom{n}{m} \binom{n}{t} \left(\frac{t}{n}\right)^s \rightarrow 0$$

as  $n \rightarrow \infty$ , where  $s = \alpha m$  and  $t = \alpha + m - \epsilon - 1$ . By our assumption that  $m \leq 10\alpha$ ,  $s$  and  $t$  are bounded by constants so the expression is  $\Theta(n^{m+t-s})$ . Now  $m + t - s = (2 - \alpha)m + \alpha - \epsilon - 1$  which is always negative. Note that, when  $m = 2, \alpha = 3$ , it is not zero only thanks to the  $\epsilon$ . It is indeed true that given three permutations there is likely to exist a set of size two whose image contains just  $4 < 2 + 3$  elements.

For large sets  $A$ ,  $10\alpha < m < 3n/4$ .  $|A_*|$  is the union of  $\alpha$  randomly chosen sets of size  $m$ , and we need to bound the probability that this union has size  $m + \alpha$  or less. A union bound over the possible sets of size  $m + \alpha$  gives (again

recalling that we have  $\binom{n}{m}$  possible candidates for  $A$  itself)

$$B(m) = \binom{n}{m} \binom{n}{m+\alpha} \left( \binom{m+\alpha}{m} / \binom{n}{m} \right)^\alpha$$

and we must show that this tends to 0 as  $n \rightarrow \infty$ . The easy inequality

$$\frac{\binom{n}{m+\alpha}}{\binom{n}{m}} \leq n^\alpha$$

implies

$$B(m) \leq n^\alpha \binom{n}{m}^{2-\alpha} \binom{m+\alpha}{\alpha}^\alpha.$$

We rewrite the upper bound as the product of two terms

$$B(m) \leq n^\alpha \binom{n}{m}^{2-3\alpha/4} \cdot \left( \frac{\binom{m+\alpha}{\alpha}}{\binom{n}{m}^{1/4}} \right)^\alpha$$

and we turn to show that each of these two terms tend to 0. In the first one, the exponent is negative since  $\alpha \geq 3$ , and so it is maximized when  $m$  is minimized, namely  $m = 10\alpha$ . For this  $m$  it is bounded by

$$(ne)^{\alpha+10\alpha(2-3\alpha/4)} \rightarrow 0$$

as required. The second expression is certainly negligible in the range  $3n/4 > m \geq n/10$  where the numerator is polynomial in  $n$  whereas the denominator is exponential in  $n$ . For  $m \leq n/10$  one easily sees that the maximum is obtained again when  $m = 10\alpha$ , where the numerator is bounded by a constant and the denominator grows like a polynomial. This proves our claim. ■

We now translate what all this means for the lift  $\tilde{G}$  of  $G$ . Every set  $X$  of vertices in  $\tilde{G}$  is the union of two disjoint sets  $X_a$  and  $X_b$  lying above  $a$  and  $b$  respectively, with  $|X_a| = x_a$  and  $|X_b| = x_b$ . Assume w.l.o.g that  $x_a \geq x_b$ . If  $x_a$  and  $x_b$  differ by more than  $\alpha$ , then  $\partial X$  automatically contains  $\alpha$  vertices, as we have seen. This, and our assumptions on  $|X|$ , implies that  $2 \leq x_a \leq 3n/4$ .

We may think of  $X_a$  as a subset  $A$  of  $\{1, \dots, n\}$ , since the vertices of  $\tilde{G}_a$  are indexed by this set. By the construction of the lift,  $A_*$  corresponds to  $\partial X_a$ , as

a subset of  $\tilde{G}_b$ . Thus we may assume that  $|\partial X_a| \geq x_a + \alpha - \epsilon_1$ . If  $x_b \geq 2$  then similarly  $|\partial X_b| \geq x_b + \alpha - \epsilon_2$ . Since  $\partial X = \partial X_a \cup \partial X_b \setminus (X_a \cup X_b)$ , its size is at least

$$|\partial X| \geq x_a + x_b + 2\alpha - (\epsilon_1 + \epsilon_2) - (x_a + x_b) = 2\alpha - (\epsilon_1 + \epsilon_2) \geq \alpha$$

or, if  $x_b \leq 1$ , we just ignore  $\partial X_b$  and

$$|\partial X| \geq x_a + \alpha - \epsilon_1 - x_b \geq \alpha$$

as required. ■

This result can be naturally extended to topological  $K_2(\alpha)$ s as well. Let  $G$  be a topological  $K_2(\alpha)$ , namely a graph with two vertices  $a, b$  and  $\alpha$  vertex-disjoint paths between them.

**Definition 2** *An  $n$ -lift  $\tilde{G}$  of  $G = K_2(\alpha)$  is called good if every set  $X \subset V(\tilde{G})$  with  $3 \leq x_a + x_b < 2 \cdot 2n/3$  has  $|\partial X| \geq \alpha$ .*

Notice that the numbers  $x_v$  for  $v \neq a, b$  are not important. The upper bound  $2 \cdot 2n/3$  is actually the same as in Lemma 1, since  $|\tilde{G}_a| + |\tilde{G}_b| = 2n$ . An immediate corollary of Lemma 1 is

**Corollary 1** *A random lift of a topological  $K_2(\alpha)$  with  $\alpha \geq 3$  is a.s. good.*

This follows at once since a lift of a topological  $K_2(\alpha)$  is a subdivision of a lift of  $K_2(\alpha)$ . ■

### 1.2.3 Thick Sets

Consider a finite base graph  $G$  and a subgraph  $H$  which is a topological  $K_2(\alpha)$  for  $\alpha \geq 3$ , namely  $H$  consists of  $\alpha$  internally disjoint paths between two vertices  $a$  and  $b$  of  $G$ . Given a lift  $\tilde{G}$  of  $G$ , we can look at its restriction  $\tilde{H}$  lift  $H$ , i.e., the

subgraph of  $\tilde{G}$  comprising of vertices and edges that are mapped to  $H$  by the covering projection. It is clear that if  $\tilde{G}$  is a random lift, then  $\tilde{H}$  is a random lift of  $H$ . For every  $H$ , this lift is a.s. good, and since there are only a finite number of possible such subgraphs  $H$ , a.s. every restricted lift  $\tilde{H} \rightarrow H$  is good.

In this section we prove that for such lifts, thick sets have at least  $\delta$  neighbours.

**Proposition 1** *Let  $G$  be a finite graph with  $\delta = \delta(G) \geq 3$ , and let  $\tilde{G} \rightarrow G$  be a lift such that the restriction  $\tilde{H} \rightarrow H$  is good for every  $H$  that is a topological  $K_2(\alpha)$  with  $\alpha \geq 3$ . Then for every thick set  $X \subset V(\tilde{G})$ ,  $|\partial X| \geq \delta$ .*

We fix the base graph  $G$  and an  $n$ -lift  $\tilde{G}$  satisfying the condition of the Proposition.  $X$  denotes a thick set of vertices in  $\tilde{G}$  with  $|X| < \tilde{G}/2$ , and as usual,  $X_v = X \cap \tilde{G}_v$ ,  $x_v = |X_v|$  and  $m = \max x_v$  so  $m \geq 2$ . Notice that (1.1) implies that  $m < 2n/3$ . We assume that  $X$  is such that  $|\partial X| < \delta$ , and arrive at a contradiction.

Notice that if  $a, b \in G$  are such that  $x_a = x_b = m$ , and there are  $\alpha \geq 3$  disjoint paths between  $a$  and  $b$ , then  $\partial X$  contains at least  $\alpha$  vertices which all lie above these paths. This is an immediate consequence of the definition of good lift and the bounds on  $m$ .

Let  $M'$  be the subgraph of  $G$  spanned by the vertices  $v$  for which  $x_v = m$ , and let  $M$  be any connected component of  $M'$ . For a vertex  $a \in M$ , set  $d(a) = \deg(a; M)$  and  $e(a) = \deg(a; G) - d(a)$ , so  $e(a)$  is the number of vertices outside  $M$  that are adjacent to  $a$ . Let  $E(a)$  denote this set of vertices.

We will make frequent use of the following observation: If a vertex  $v \in M$  is adjacent to a vertex  $z \notin M$ , then  $\partial X$  contains at least one vertex in  $\tilde{G}_z$ . (This is because the edges covering  $[v, z]$  form a matching between  $\tilde{G}_v$  and  $\tilde{G}_z$ , and  $x_z < x_v$  so some vertex in  $X_v$  is matched to one outside  $X_z$ ). It follows that every  $a \in M$  implies the existence of at least  $e(a)$  vertices in  $\partial X$ , which lie above  $E(a)$ . We also see that  $|\partial M| < \delta$ , and in particular  $M$  cannot be a single

vertex.

Let  $d = \delta(M)$  be the minimal degree of  $M$ , and  $k = \bar{\kappa}(M)$  be its maximum local connectivity. We will gradually restrict the structure of  $M$ , until all possibilities are exhausted.

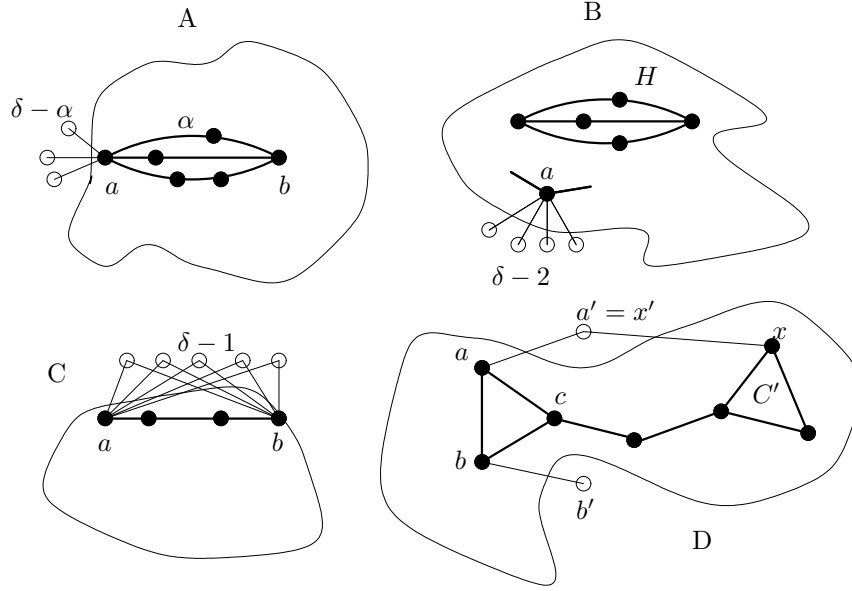
First, assume that  $d \geq 3$ . We use Mader's Theorem 2 to find vertices  $a, b$  in  $M$  with  $d(a) = \alpha \leq \beta = d(b)$  having  $\alpha$  disjoint paths between them. Since  $2 \leq m \leq 2n/3$  and  $\alpha \geq d \geq 3$ , we find at least  $\alpha$  vertices of  $\partial X$  lying above  $M$ . But  $e(a) \geq \delta - \alpha$ , and so  $\partial X$  contains also  $\delta - \alpha$  vertices that do not lie over  $M$ , giving a total of  $\delta$  vertices in  $\partial X$ , a contradiction. Therefore  $d \leq 2$ . See Figure 1.6A.

Now suppose that  $k \geq 3$ . Let  $a$  be a vertex of  $M$  with  $d(a) = \alpha \leq 2$ , and let  $H$  be a topological  $K_2(3)$  in  $M$ .  $e(a)$  is at least  $\delta - 2$ , and there are at least 3 vertices in  $\partial X$  above  $H$ . This gives at least  $\delta + 1$  vertices in  $\partial X$ , a contradiction. Note that it makes no difference if  $a$  belongs to  $H$  or not. This case is depicted in Figure 1.6B.

Now we know that  $k \leq 2$ . Suppose that  $K$  is a block of  $M$  which is not an edge. Then  $K$  is 2-connected, and we claim that it is a cycle. Otherwise, let  $C$  be a cycle in  $K$  and  $[c, x]$  an edge outside  $E(C)$ , with  $c \in C$ . If  $x \in C$  then  $\kappa(x, c; M) \geq 3$  which is impossible. Otherwise, since  $c$  is not a cutvertex of  $K$ , there is a path which avoids  $c$  from  $x$  to some vertex  $z$  on  $C$ . Thus  $\kappa(c, z; M) \geq 3$  which is again impossible. We obtain that every block of  $M$  is either an edge or a cycle.

If  $M$  has at least two leaves, we proceed as follows. Let  $a$  and  $b$  be leaves. Since  $e(a) \geq \delta - 1$ , every vertex in  $E(b) \setminus E(a)$  would give an additional vertex in  $\partial X$  which cannot happen. Therefore,  $E(b) = E(a)$ , and in particular there are at least  $\delta$  disjoint  $a$ - $b$  paths in  $G$  (one in  $M$  and  $\delta - 1$  paths of length 2; See Figure 1.6C). But this is also impossible, since it implies  $\delta$  vertices of  $\partial X$  above these paths.

Since  $M$  is not a single vertex and contains at most one leaf, it has an

Figure 1.6: Various cases for  $M$ .

endblock  $C$  which is a cycle. Let  $a$  and  $b$  be two vertices in  $C$  with  $d(a) = 2$ . We claim that  $E(a)$  is disjoint from  $E(b)$ . Otherwise, there are 3 disjoint  $a$ - $b$  paths in  $G$ , two of which are in  $M$  and another one outside it.  $E(a)$  still contains at least  $\delta - 3$  vertices disjoint from these paths, which together yield  $\delta$  elements of  $\partial X$ , as usual.

It follows that  $C$  cannot contain 3 vertices  $a, b, c$  with  $d(a) = d(b) = d(c) = 2$ , since then  $E(a)$ ,  $E(b)$  and  $E(c)$  are three pairwise disjoint sets of size  $\geq \delta - 2$  each, and  $3(\delta - 2) \geq \delta$ . In particular  $M$  is not a single cycle and must contain more than one block. Since  $C$  is an endblock, it must contain at least two vertices of degree 2, so the only remaining possibility is the one that appears in Figure 1.6D:  $C = \{a, b, c\}$  is a cycle of length 3 and one of its vertices, say  $c$ , belongs to another block.  $E(a)$  and  $E(b)$  together contain  $2(\delta - 2)$  vertices, so  $\delta = 3$  and  $e(a) = e(b) = 1$ . Set  $E(a) = \{a'\}$  and  $E(b) = \{b'\}$ .

Let  $C'$  be another endblock of  $M$ . There is a path from  $c$  to  $C'$  that avoids

$C$ .  $C'$  can be an edge or a cycle, but in either case it contains a vertex  $x$  with  $d(x) \leq 2$ , so  $e(x) \geq 1$ . Let  $x' \in E(x)$  be some vertex. Then  $x' = a'$  or  $x' = b'$ , since otherwise we get 3 distinct vertices in  $\partial M$ . Say  $x' = a'$ . There are now 3 disjoint paths from  $a$  to  $c$ , two in  $M$  (through  $C$ ) and one outside, through  $a'$  and  $x$ . Once more this implies that  $|\partial X| \geq 3$ , which contradicts  $\delta = 3$ .

This final contradiction establishes the Proposition. ■

### 1.2.4 Thin Sets

In this section we show that a.s. every *thin* set in  $\tilde{G}$  has more than  $\delta$  neighbours. Unlike our analysis of thick sets, here we work with certain subgraphs of  $\tilde{G}$ , the lift itself.

**Definition 3** *A subgraph  $H$  of a lift  $\tilde{G}$  is called edge-thin if it does not contain a pair of parallel edges, namely edges covering the same edge of the base graph  $G$ .*

An edge-thin subgraph contains at most one edge in each edge fiber, in contrast with a thin set which contains at most one vertex in each vertex fiber. Notice that a thin set spans an edge-thin subgraph, but the vertex set of an edge-thin subgraph is not necessarily thin. Also, a pair of parallel edges cannot be incident with the same vertex, by the definition of a lift.

**Lemma 2** *Let  $G$  be a finite base graph and  $\tilde{G}$  a random  $n$ -lift. Then a.s. in every edge-thin subgraph of  $\tilde{G}$ , every connected component is a tree or is unicyclic.*

*Proof:* Let  $\Delta = \Delta(G)$  be the maximum degree of  $G$ . An edge-thin subgraph  $H$  of  $\tilde{G}$  contains at most  $\Delta$  vertices in each fiber, since if  $\tilde{G}_v$  has  $\Delta + 1$  vertices in  $H$ , there is an edge in  $G$  incident with  $v$  that is covered twice by  $H$ . In

particular, the number of vertices of an edge-thin subgraph is bounded by a constant  $K = \Delta|V(G)|$ .

It suffices to show that for each  $1 \leq k \leq K$ , a.s. no edge-thin subgraph  $H$  with  $k$  vertices has more than  $k$  edges. The number of possible vertex-sets  $V(H)$  is  $O(\binom{n|V(G)|}{k}) = O(n^k)$ .

Given a vertex set  $V$ ,  $|V| = k$ , we bound the probability that there is an edge-thin subgraph  $H$  with  $V = V(H)$  and  $|E(H)| = j > k$  as follows. First we choose a collection of  $j$  putative edges  $[u, v]$ ,  $u, v \in V$ , no two of which cover the same edge of  $G$ . For each pair, the probability that  $[u, v]$  is indeed an edge of  $\tilde{G}$  is  $1/n$ , and these events are independent by the assumption that no edge of  $G$  is covered twice. Therefore the probability that all putative edges exist is  $n^{-j}$ , and the probability that *some* choice of putative edges succeeds is still  $O(n^{-j})$  because the possible number of choices is bounded by, say,  $2^{K^2}$ .

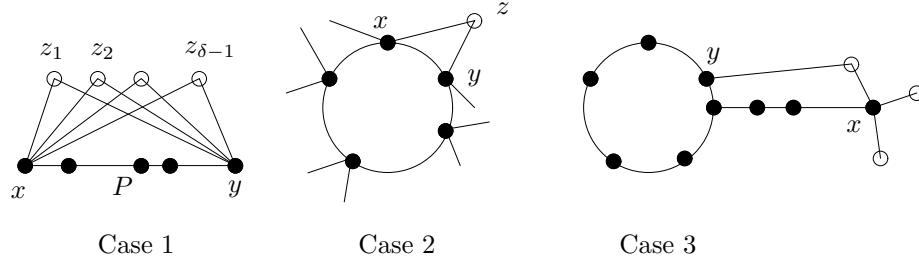
A union bound over all possible  $V$ s gives  $n^k n^{-j}$  which is  $O(1/n)$  since  $j > k$ . Therefore, a.s. there are no edge-thin subgraphs with  $k$  vertices, as required. ■

We now prove

**Proposition 2** *Let  $G$  be a finite graph with  $\delta = \delta(G) \geq 3$ , and let  $\tilde{G} \rightarrow G$  be a lift such that every edge-thin subgraph  $H$  of  $\tilde{G}$  satisfies  $|E(H)| \leq |V(H)|$ . Then for every thin set  $X \subset V(\tilde{G})$ ,  $|\partial X| \geq \delta$ .*

Fix a thin set  $X$ ,  $|X| = k$ , and let  $H$  be the subgraph of  $\tilde{G}$  spanned by  $X$ . Since  $H$  is edge-thin, it has at most  $k$  edges. It follows that  $H$  has a connected component which is a tree or a tree with an extra edge. Let  $H'$  be that connected component. We proceed by considering various possibilities for  $H'$ . See Figure 1.7.

**Case 1.** Suppose that  $H'$  has two leaves, say  $x, y$ . Namely,  $\deg(x; H) = \deg(y; H) = 1$ . Since  $\deg(x; \tilde{G}) \geq \delta$ ,  $x$  is adjacent to at least  $\delta - 1$  vertices  $z_1, \dots, z_{\delta-1}$  outside  $H$ , namely outside  $X$ . If  $y$  has a neighbour outside  $H$

Figure 1.7: Cases for  $H'$ 

which is different from the  $z_i$ 's, we have  $\delta$  vertices in  $\partial X$  and we are done. If not, then each  $z_i$  is adjacent to both  $x$  and  $y$ .

We construct a certain subgraph  $Z$  of  $\tilde{G}$ : Take a simple  $x$ - $y$  path  $P$  in  $H'$ , along with the vertices  $z_i$  and all the edges  $[x, z_i]$  and  $[z_i, y]$ . Clearly  $Z$  is edge-thin. Since  $\delta \geq 3$ ,  $Z$  is a connected edge-thin subgraph with two distinct cycles, contrary to our assumption. Henceforth,  $H'$  has at most a single leaf.

**Case 2.**  $H'$  is a cycle. Since  $\tilde{G}$  is simple,  $H'$  contains at least 3 vertices, each of which has at least  $\delta - 2$  neighbours outside  $H'$ . Suppose that two vertices  $x, y$  of  $H'$  share a common neighbour  $z$  outside  $H'$ . Then the graph consisting of  $H'$ ,  $z$  and the edges  $[x, z]$  and  $[y, z]$  is edge-thin and has more edges than vertices, which is impossible by assumption. Therefore every vertex outside  $H'$  has at most a single neighbour in  $H'$ . Consequently  $|\partial X| \geq 3(\delta - 2) \geq \delta$ .

**Case 3.** The only remaining possibility is that  $H'$  is a lollipop: a cycle  $C$  with an attached path  $P$ . Let  $x$  be the endpoint of the path; it has  $\delta - 1$  neighbours outside  $H'$ . Let  $y$  be a vertex of degree 2 on  $C$ . This  $y$  has at least one neighbour outside  $H'$ . If it is not one of  $x$ 's neighbours, then we are done. Otherwise the graph formed by adding this common neighbour to  $H'$  along with the two edges to  $x$  and  $y$  is edge-thin and has more edges than vertices, which is impossible. ■

### 1.2.5 Conclusion

The proof of Theorem 1 is now complete. A random lift  $\tilde{G}$  almost surely satisfies the conditions of both Proposition 1 and Proposition 2, as shown in the beginning of Section 1.2.3 and Lemma 2. Therefore both thick and thin sets have  $\delta$  neighbours outside them, so  $\tilde{G}$  is  $\delta$ -connected. ■

Let us make some further comments concerning the proof. It is not hard to extend Lemma 2 as follows: Fix a constant  $B$ . Then a.s. every subgraph  $H$  of  $\tilde{G}$  with at most  $B$  parallel edges in each edge fiber satisfies  $|E(H)| \leq |V(H)|$ . It follows that topological  $K_2(\delta)$ s in  $\tilde{G}$ , which exist in abundance, must contain an unbounded amount of parallel edges, and in particular they are quite large. In a typical random lift, every two vertices are connected by  $\delta$  disjoint paths, but these paths are necessarily quite long and utilize the lifts of some particular edges many times.

Theorem 1 is an asymptotic result. It will be interesting to estimate the probability that a random  $n$ -lift fails to be  $\delta$ -connected in terms of  $n$ . We expect this probability to become small already for moderate values of  $n$ .

A different proof of Theorem 1 can be given using the model of  $T$ -flat lifts (see 1.1.3). We give a sketch of it now. In a  $T$ -flat labeled lift  $\tilde{G}$ , the subgraphs spanned by the vertices with label  $i$  all contain a copy of the spanning tree  $T$  of  $G$ . Call these subgraphs *layers*. Define the *layer graph*  $\Lambda$  to be the graph on  $[n]$  that has an edge  $[i, j]$  whenever some edge in the lift connects a vertex in layer  $i$  to one in layer  $j$ .  $\Lambda$  is a random graph obtained by superposing  $|E(G)| - |V(G)| + 1$  random permutations, and in particular it is highly connected.

Given two vertices  $u_i$  and  $v_j$  in  $\tilde{G}$ , a path in  $\Lambda$  from  $i$  to  $j$  can be used to construct a path from  $u_i$  to  $v_j$ : we move between the layers using the edges that connect them (these correspond to edges of  $\Lambda$ ), and within a layer using the copy of  $T$ . But we need to exhibit not one but  $\delta$  disjoint paths from  $u_i$  to  $v_j$ . The high connectivity of  $\Lambda$  implies that there are indeed many disjoint paths

between  $i$  and  $j$  in  $\Lambda$ , but trying to realize those in  $\tilde{G}$ , we run into a difficulty: the paths may collide in the layers  $i$  and  $j$ , as they travel along the single copy of  $T$  in each of these layers.

To overcome this, we need to show that each vertex  $u_i$  of  $\tilde{G}$  admits a  $\delta$ -fan: a collection of  $\delta$  disjoint paths emanating from  $u_i$ , each of which is contained in layer  $i$  except for its final edge that “jumps” from this layer to some other layer, and these terminal layers are all distinct. Given those, we can find  $\delta$  disjoint paths from  $u_i$  to  $v_j$  by starting with the fans and gluing them using paths realized from  $\Lambda$ . It turns out, however, that not every vertex admits such a fan. In order to guarantee that a.s. every vertex has a fan we need to allow slightly more complicated fans, in which some paths meet more than two layers. We omit the details.

We return to our original proof. Following the steps, we find that for most sets  $X$ ,  $|\partial X|$  is much larger than  $\delta$ . However it is not true that a.s. only singletons of  $\tilde{G}$  can be disconnected by removing  $\delta$  vertices. For example, a random lift of  $K_4$  is likely to contain a 3-cycle with only 3 neighbours. Yet the proof suggests that random lifts are good expanders, as large sets tend to have large boundaries. We will investigate expansion properties of random lifts in the next chapter.

### 1.3 Isomorphisms and Labeling

In this section we prove the claim made in 1.1.3, that the random labeled model is applicable to unlabeled lifts as well. The basis of this claim is an asymptotic estimate on the number of unlabeled lifts of a fixed base graph  $G$ , which we proceed to derive. Recall that, for two labeled lifts to be considered isomorphic, there must be a covering-preserving isomorphism between them. The question of counting the unlabeled graphs covering a fixed base graph up to *isomorphism of graphs*, regardless of the coverings, is more difficult and will not be addressed

here.

### 1.3.1 The Number of Unlabeled Lifts

First, let us determine exactly when two labeled lifts given by permutations are isomorphic as unlabeled lifts. We assume that the edges of the base graph  $G$  have been oriented, and  $e = (u, v)$  refers to an edge from  $u$  to  $v$ . Let  $\sigma : E(G) \rightarrow S_n$  and  $\sigma' : E(G) \rightarrow S_n$  be two assignments of permutations to each edge of  $G$ . They determine lifts  $\pi : \tilde{G} \rightarrow G$  and  $\pi' : \tilde{G}' \rightarrow G$ , as in Definition 1.

Suppose  $f : \tilde{G} \rightarrow \tilde{G}'$  is an isomorphism of graphs that respects the coverings.  $f$  must map a vertex in  $\tilde{G}_u$  to  $\tilde{G}'_u$ , so for each vertex  $u \in V(G)$ , The restriction of  $f$  to the fiber  $\tilde{G}_u$  is given by a permutation  $\phi(u)$  on  $[n]$ , such that  $f(u_i) = u'_{\phi(u)i}$ .

In other words,  $f$  is a renaming function: it relabels the vertices in each fiber so that the edges in  $\tilde{G}$  become those of  $\tilde{G}'$ . In  $\tilde{G}$ , an edge  $e_i$  connects  $u_i$  to  $v_{\sigma(e)i}$ . Now  $f(u_i) = u'_{\phi(u)i}$  and  $f(v_{\sigma(e)i}) = v'_{\phi(v)\sigma(e)i}$ . Since  $f(e_i)$  connects these two vertices, we must have:

$$\phi(v)\sigma(e)i = \sigma'(e)(\phi(u)i)$$

for every  $i \in [n]$ , or

$$\sigma'(e) = \phi(v)\sigma(e)\phi(u)^{-1}. \quad (1.2)$$

Conversely, if there exists a relabeling scheme  $\phi : V(G) \rightarrow S_n$  such that (1.2) holds for every  $e$ , then the lifts determined by  $\sigma$  and  $\sigma'$  are isomorphic.

Put differently, the group  $K = S_n^{V(G)}$  acts on  $L_n(G)$  by  $\phi\sigma = \sigma'$  where  $\sigma'$  is given by (1.2) (it is easy to see that this is indeed an action). The orbits under this action correspond to isomorphism classes of labeled lifts, namely, unlabeled lifts.

We are now ready to calculate the cardinality of  $C_n(G)$ , the collection of

unlabeled lifts of  $G$ . To this end we apply Burnside's lemma

$$|C_n(G)| = \frac{1}{n!|V(G)|} \sum_{\phi \in K} |\text{Fix}(\phi)|$$

where  $\text{Fix}(\phi)$  denotes the collection of permutation assignments  $\sigma$  such that  $\phi\sigma = \sigma$ . Such a  $\sigma$  satisfies for every edge  $e = (u, v)$

$$\sigma(e) = \phi(v)\sigma(e)\phi(u)^{-1}$$

or

$$\phi(v) = \sigma(e)\phi(u)\sigma(e)^{-1}.$$

Therefore, the number of  $\sigma$ -s that are invariant under  $\phi$  is

$$|\text{Fix}(\phi)| = \prod_{e=(u,v)} |\{\sigma \in S_n \mid \phi(v) = \sigma\phi(u)\sigma^{-1}\}|. \quad (1.3)$$

This number is 0 unless all the permutations  $\phi(u)$ ,  $u \in V(G)$  are conjugate in  $S_n$ . Let  $\mathcal{C}$  be a conjugacy class in  $S_n$ , of size  $|\mathcal{C}|$ . Let  $s(\mathcal{C})$  be the size of the centralizer of any permutation in  $\mathcal{C}$ , which is, clearly, independent of the particular permutation. For any conjugacy class  $\mathcal{C}$  in  $S_n$ , the number of elements  $\phi \in K$  for which  $\phi(u) \in \mathcal{C}$  for every  $u$  is  $|\mathcal{C}|^{|V(G)|}$ . For such a  $\phi$ ,  $|\text{Fix}(\phi)|$  is simply  $s(\mathcal{C})^{|E(G)|}$  by (1.3). Therefore

$$|C_n(G)| = \frac{1}{n!|V(G)|} \sum_{\mathcal{C}} |\mathcal{C}|^{|V(G)|} s(\mathcal{C})^{|E(G)|}$$

where the sum is extended over all conjugacy classes in  $S_n$ . Now for every  $\mathcal{C}$ , the product  $|\mathcal{C}|s(\mathcal{C}) = n!$ , so finally:

$$|C_n(G)| = \sum_{\mathcal{C}} s(\mathcal{C})^{|E(G)| - |V(G)|} \quad (1.4)$$

For example, suppose that  $G$  is a graph with  $|E(G)| = |V(G)| + 1$ . The number of unlabeled 3-lifts is  $6 + 2 + 3 = 11$ . The reader may find it instructive to list them for some specific choice of  $G$ .

### 1.3.2 Asymptotics

Let us consider some special cases of this formula. When  $G$  is a tree,  $|E(G)| - |V(G)| = -1$ ,  $s(\mathcal{C})^{-1} = |\mathcal{C}|/n!$  and so  $|C_n(G)| = 1$ . This is indeed the case, since a tree is simply connected and the only  $n$ -sheeted covering of a tree is a disjoint union of  $n$  copies of that tree. When  $G$  is unicyclic,  $|E(G)| - |V(G)| = 0$  and  $|C_n(G)|$  is just the number of conjugacy classes in  $S_n$ , namely the number  $p(n)$  of partitions of  $n$ . Indeed, lifts of a cycle are disjoint unions of cycles of various lengths, corresponding to the cycle structure of the permutation formed by multiplying all the edge permutations  $\sigma(e)$  together.

Another special case is  $n = 2$ . There are two conjugacy classes, both of which have  $s(\mathcal{C}) = 2$ , so

$$|C_2(G)| = 2^{|E(G)| - |V(G)| + 1}$$

which is just the number of  $T$ -flat lifts.

These examples are the exception rather than the rule. When  $n$  is large and  $|E(G)| > |V(G)|$ , the dominant term in (1.4) is  $n!^{|E(G)| - |V(G)|}$ , coming from the trivial conjugacy class  $1 \in S_n$ .

**Lemma 3** *Let  $k > 1$  be a fixed integer. Let  $c_n = \sum_{\mathcal{C}} s(\mathcal{C})^k$  where  $s(\mathcal{C})$  is the size of the centralizer of an element in  $\mathcal{C}$  and the sum is taken over all conjugacy classes in  $S_n$ . Then as  $n \rightarrow \infty$ ,*

$$c_n = (1 + o(1))n!^k$$

*Proof:* For every class  $\mathcal{C}$  we have  $s(\mathcal{C}) \leq n!$ , and for the class of the identity  $s(1) = n!$ . Therefore

$$1 \leq \sum_{\mathcal{C}} \left( \frac{s(\mathcal{C})}{n!} \right)^k \leq \sum_{\mathcal{C}} \frac{s(\mathcal{C})}{n!}$$

and so it suffices to prove that  $1/n! \sum_{\mathcal{C}} s(\mathcal{C}) \rightarrow 1$ , the case  $k = 1$  of the Lemma.

A conjugacy class  $\mathcal{C}$  is determined by a partition of  $n$ , namely the cycle structure of the permutations in  $\mathcal{C}$ . Let  $k_j$  be the number of cycles of length  $j$ .

Then  $\sum jk_j = n$  and

$$s(\mathbb{C}) = \prod_j k_j! j^{k_j}$$

since an element of  $S_n$  which centralizes a permutation with this cycle structure shuffles the cycles of equal length among themselves, and cyclically shifts each cycle. Hence we need to prove that

$$Q = \frac{1}{n!} \sum_{(k_j)} \prod_j k_j! j^{k_j} \rightarrow 1.$$

$$\sum jk_j = n$$

The number of summands here is  $p(n)$ . We split the sum according to the value of  $k_1$ . When  $k_1 = n$ , all the other  $k_j$ -s are zero, and this contributes 1 to  $Q$ . Now suppose  $n > k_1 > n/2$ . Since  $k_2 + \dots + k_n \leq n - k_1$ , we have

$$\prod_j k_j! \leq k_1! (n - k_1)!$$

and using  $j \leq 2^{j/2}$  we get

$$\prod_j j^{k_j} \leq 2^{\frac{n-k_1}{2}}.$$

Let  $t = n - k_1$ . The number of partitions in which  $t$  has a specified value is  $p(t) \leq e^{K\sqrt{t}}$  for some constant  $K$ . Letting  $Q'$  denote the contribution to  $Q$  from partitions satisfying  $n > k_1 > n/2$ , we get

$$\begin{aligned} Q' &\leq \sum_{2 \leq t < n/2} \frac{2^{t/2} e^{K\sqrt{t}}}{\binom{n}{t}} \\ &\leq \sum_{2 \leq t < n/2} \frac{(2e)^{t/2}}{n^t / 2^t} \\ &\leq \sum_{2 \leq t < n/2} \left( \frac{2^{3/2} e}{n} \right)^t \end{aligned}$$

and this obviously tends to 0 as  $n \rightarrow \infty$ .

Finally, let  $Q''$  denote the contribution to  $Q$  coming from partitions in which  $k_1 \leq n/2$ . Since  $2(k_2 + k_3 + \dots + k_n) \leq 2k_2 + 3k_3 + \dots + nk_n = n - k_1$ ,

$$k_2 + k_3 + \dots + k_n \leq \frac{n - k_1}{2}$$

hence

$$\sum_{j=1}^n k_j \leq k_1 + \frac{n - k_1}{2} \leq \frac{3n}{4}$$

and therefore

$$\prod k_j! \leq (\sum k_j)! \leq \left(\frac{3n}{4}\right)!$$

Once more  $\prod j^{k_j} \leq 2^{n/2}$  so

$$Q'' \leq \frac{p(n)\left(\frac{3n}{4}\right)!2^{n/2}}{n!}$$

which also tends to 0 by the same bound on  $p(n)$ . Adding, we get  $Q = 1 + Q' + Q'' \rightarrow 1$  as required.  $\blacksquare$

Combining this with (1.4) we obtain

**Corollary 2** *For a base graph  $G$  with  $v$  vertices and  $e$  edges, the number of unlabeled lifts is*

$$C_n(G) = \begin{cases} 1 & e - v = -1 \\ p(n) & e - v = 0 \\ (1 + o(1))(n!)^{e-v} & e - v > 0 \end{cases}$$

Since an unlabeled lift that admits a non-trivial automorphism can be labeled in no more than  $n!^v/2$  ways, this implies that for  $e > v$  a.e. unlabeled lift has no automorphisms, and the same holds of course for labeled lifts. In particular, almost every lift of  $G$  is irregular.

### 1.3.3 Equivalence of the Models

We now use Lemma 3 to prove the equivalence between the labeled and unlabeled models. Let  $P$  be a property of unlabeled lifts. For instance  $P$  may

be connectivity or 3-colorability of the lift, regularity of the covering, etc. Let  $\Pr_L(P)$  be the probability that  $P$  holds for a random labeled lift, and  $\Pr_C(P)$  the probability for unlabeled ones.

**Theorem 3** *Let  $G$  be a graph with  $|E(G)| = e > v = |V(G)|$ , and let  $P$  be a property of unlabeled lifts of  $G$ . Then almost every labeled lift has property  $P$  iff almost every unlabeled lift has it.*

*Proof:* This follows from the following Lemma:

**Lemma 4** *With  $G$  and  $P$  as above, as the degree  $n \rightarrow \infty$*

$$\Pr_L(P) \leq (1 + o(1))\Pr_C(P)$$

*Proof:* Fix a degree  $n$ . Let  $L(P)$  and  $C(P)$  denote, respectively, the collection of labeled and unlabeled lifts of order  $n$  of  $G$  that have property  $P$ , and let  $l(P) = |L(P)|$ ,  $c(P) = |C(P)|$ . Obviously  $l(P) \leq n!^v c(P)$ . Now

$$\begin{aligned} \Pr_L(P) &= l(P)/|L_n(G)| \\ &= l(P)/n!^e \\ &\leq c(P)/n!^{e-v} \\ &= (1 + o(1))c(P)/|C_n(G)| \\ &= (1 + o(1))\Pr_C(P) \end{aligned}$$

where the next to last inequality uses Lemma 3. ■

We return to the proof of Theorem 3. Suppose that almost every labeled lift has  $P$ , in the sense that  $\Pr_L(P) = 1 - o(1)$ . Then  $\Pr_C(P) = 1 - o(1)$  also, so almost every unlabeled lift has  $P$  as well. Conversely, if  $\Pr_C(P) = 1 - o(1)$ , then  $\Pr_C(P') = o(1)$  where  $P'$  is the negation of  $P$ , and so  $\Pr_L(P') = o(1)$  and  $\Pr_L(P) = 1 - o(1)$ . ■

The same argument also shows that  $\Pr_L(P) \rightarrow c$  for some constant  $c$  iff  $\Pr_C(P) \rightarrow c$  for the same constant.

## 1.4 Variations

In this chapter we considered the model of independent uniform permutations. Let us mention here some possible generalizations and variants of the model.

Given a group  $\Gamma$  acting on a set  $\Omega$ , we can attach an element  $g_e \in \Gamma$  to each edge  $e$  of  $G$  and view each fiber  $\tilde{G}_u$  of as a copy of  $\Omega$ ,  $\tilde{G}_u = \{u_\alpha\}_{\alpha \in \Omega}$ . The edges in  $\tilde{G}$  now connect  $u_\alpha$  to  $v_{g_e \alpha}$ , and  $\tilde{G}$  is a covering of degree  $|\Omega|$ . If it is possible to select elements of  $\Gamma$  at random, we obtain random lifts. The standard model is the special case where  $\Gamma = S_n$  and  $\Omega = [n]$  with the usual action.  $\Gamma$  and  $\Omega$  may be finite or infinite. (In fact, even  $G$  may be infinite as long as  $\Gamma^{E(G)}$  is equipped with a probability measure).

A special case of this is that of the group  $\Gamma$  acting on itself by left multiplication. The resulting lifts of  $G$  will then be regular in the usual topological sense. If  $G$  is the bouquet  $B_k$  with one vertex and  $k$  loops, the lift will be the Cayley graph of  $\Gamma$  with respect to the set  $S$  of the elements attached to the loops and their inverses, assuming this set generates  $\Gamma$ .

A variation in a different direction is obtained as follows. Let  $n = n_1 \cdot n_2 \cdots n_r$ . Starting from  $G$ , we form a random  $n_1$ -lift, then a random  $n_2$ -lift of the result, and so on. Since a composition of covering maps is itself a covering, the resulting graph is an  $n$ -degree cover of  $G$ , distributed differently than one formed by taking a random  $n$ -lift directly. This kind of model may be more suited for some applications.

A far-reaching generalization of the above would be a model for random lifts of higher-dimensional simplicial complexes. At present this seems difficult to obtain even for 2-dimensional complexes. However, we mention here an interesting model for random branched coverings of a 2-manifold.

Let  $M$  be a 2-dimensional manifold and let  $G$  be a graph embedded in  $M$  in such a way that the regions (connected components of  $M \setminus G$ ) are homeomorphic to open disks, and their boundaries form simple cycles on the graph (this is

sometimes called a *strong embedding*). For example, we can let  $G$  be the 1-skeleton of a triangulation of  $M$ .

Consider a covering  $\tilde{G}$  of  $G$ . Every cycle  $C$  in  $G$  that bounds a region in  $M$  is covered by the disjoint union of cycles  $\tilde{C}_j$  in  $\tilde{G}$ , and every edge of  $\tilde{G}$  belongs to exactly two such cycles. Let us attach a 2-cell to each  $\tilde{C}_j$ . We get a 2-complex  $\tilde{M}$  and it is easy to verify that it is, in fact, a manifold. There is also a natural map  $\tilde{M} \rightarrow M$  which maps the cell attached to  $\tilde{C}_j$  onto the region bounded by  $C$ , and restricts to  $\tilde{G} \rightarrow G$  on the 1-skeleton. This map will not in general be a covering of manifolds: it may be branched at the centers of the attached cells.

If the original covering  $\tilde{G} \rightarrow G$  is chosen randomly, this procedure yields a random branched covering  $\tilde{M}$  of  $M$ , although we do not get all coverings of  $M$  that way. It will be interesting to see if anything useful can be done with this model.

## Chapter 2

# Edge Expansion of Random Lifts

We continue the study of random lifts of graphs. Here we study the possibility of generating graphs with high edge expansion as random lifts. Along the way, we introduce the method of  $\epsilon$ -nets into the study of random structures. This enables us to improve (slightly) the known bounds for the edge expansion of regular graphs.

### 2.1 Introduction

In this chapter we try to utilize the random model to generate graphs with high edge expansion. We begin by showing the existence of a lower bound on the edge expansion of random lifts of a general base graph. We then calculate an explicit bound for the case of the bouquet  $B_l$  as base graph. Finally, using the method of  $\epsilon$ -nets, we show the existence of graphs whose edge expansion exceeds the known bounds.

Specifically, we are concerned with the following problem: Given a fixed

connected base graph  $G$ , what can be said about the edge expansion of a random lift  $\tilde{G}$  of  $G$ ? This should be compared with known results concerning edge expansion of random graphs (see [9] and the discussion closing Section 2.3). We prove:

**Theorem 4** *Let  $G = (V, E)$  be a connected graph with  $|E| > |V|$ . Then there is a positive constant  $\xi_0 = \xi_0(G)$  such that almost every lift of  $G$  has edge expansion at least  $\xi_0$ .*

For specific base graphs we can produce explicit bounds on the edge expansion of a random lift. Let  $B_l$  be the bouquet with one vertex and  $l$  loops.

**Theorem 5** *Almost every lift  $\tilde{G}$  of  $G = B_l$  has edge expansion  $\xi(\tilde{G}) \geq \xi_0$ , where  $\xi_0$  is the smaller root of  $H(\xi/d) = (d-2)/d$ ,  $d = 2l$ .*

It is possible to apply the same technique used for proving Theorem 2 for other base graphs, such as the complete graphs  $K_r$ , but the technical details are more tedious. Finally, we use the method of  $\epsilon$ -nets to improve the known bounds on the edge expansion of random  $d$ -regular graphs. The same method applies, mutatis mutandis, for random lifts.

**Theorem 6** *For every  $d \geq 3$  there is an  $\epsilon_0 = \epsilon_0(d)$  such that if  $0 \leq \epsilon \leq \epsilon_0$  then the following holds. Let  $\mu_0$  be the larger solution of*

$$\frac{2}{d}(1 - H(\epsilon)) = 1 - H(\mu_0)$$

*then the edge expansion of a  $d$ -regular graph is almost surely at least  $d(1 - \mu_0 - 2\epsilon)$ .*

Throughout these sections, we shall often use the entropy function

$$H(x) = -x \log_2(x) - (1-x) \log_2(1-x)$$

and the approximation

$$\binom{u}{v} = 2^{uH(v/u)(1-o_u(1))} \quad (2.1)$$

which is valid for arbitrary  $v$  and  $u \rightarrow \infty$ . The inequality

$$\binom{u}{v} \leq 2^{uH(v/u)} \quad (2.2)$$

is valid for all  $u, v$ .

## 2.2 General Graphs

We define the edge expansion of a graph as usual:

**Definition 4** *Given a set  $S \subset V(G)$  of vertices in a graph  $G$ , let  $E(S, \bar{S})$  be the set of edges with one vertex in  $S$  and one outside  $S$ . The edge expansion  $\xi(S)$  is defined to be  $|E(S, \bar{S})|/|S|$ , and the edge expansion of  $G$  is*

$$\xi(G) = \min\{\xi(S) \mid S \subset V(G), |S| \leq |V(G)|/2\}.$$

A lift  $\tilde{G}$  of  $G$  cannot have higher edge expansion than  $G$ : Given  $S \subset V(G)$  with some small  $\xi(S)$ , take  $\tilde{S}$  to be the union of the vertex fibers  $\tilde{G}_u = \{u\} \times [n]$  for  $u \in S$ . Then  $\xi(\tilde{S}) = \xi(S)$  and  $|\tilde{S}| \leq |V(\tilde{G})|/2$  iff  $|S| \leq |V(G)|/2$ .

The first natural question is whether the edge expansions of  $G$ 's lifts are almost surely bounded away from 0, that is whether there is a  $\xi_0 = \xi_0(G) > 0$  such that almost every lift  $\tilde{G}$  of  $G$  has edge expansion at least  $\xi_0(G)$ . We must leave out degenerate cases where  $G$  is a tree or unicyclic, since for such base graphs a random lift is not even a.s. connected. Having ruled out these graphs, the answer to our question is positive:

**Theorem 1** *Let  $G = (V, E)$  be a connected graph with  $|E| > |V|$ . Then there is a positive constant  $\xi_0 = \xi_0(G)$  such that almost every lift of  $G$  has edge expansion at least  $\xi_0$ .*

As in the  $\delta$ -connectivity theorem in chapter 1, the probabilistic core of the proof lies in the following simple lemma.

**Definition 5** Given two permutations  $\sigma_1, \sigma_2 \in S_n$  and an  $\eta > 0$ , a set  $A \subset [n]$  is called  $\eta$ -bad if  $0 < |A| < 2n/3$  and  $|A \cup \sigma_1(A) \cup \sigma_2(A)| \leq (1 + \eta)|A|$ .

**Lemma 5** There exists a positive  $\epsilon$  such that, for two uniformly chosen random permutations  $\sigma_1, \sigma_2 \in S_n$ , almost surely there are no  $\epsilon$ -bad sets.

Put differently, the probability that a bad set exists tends to 0 as  $n \rightarrow \infty$ . Notice that this is meaningful even for sets  $A$  that are singletons: almost surely, no element is a fixed point of both permutations.

*Proof:* For a fixed set  $A$  of size  $m$ , the probability that  $A$  is  $\epsilon$ -bad is bounded above by

$$\binom{n-m}{\lfloor \epsilon m \rfloor} \left( \frac{\binom{\lfloor (1+\epsilon)m \rfloor}{m}}{\binom{n}{m}} \right)^2$$

and since there are  $\binom{n}{m}$  possible such sets  $A$ , it suffices to show that

$$\sum_{m=1}^{2n/3} B(m) = o(1)$$

where

$$B(m) = \binom{n}{m} \binom{n-m}{\lfloor \epsilon m \rfloor} \left( \frac{\binom{\lfloor (1+\epsilon)m \rfloor}{m}}{\binom{n}{m}} \right)^2.$$

First, consider the range  $n/10 < m < 2n/3$ . Setting  $\mu = m/n$  and using (2.1), which is valid since  $m \rightarrow \infty$ , we obtain (in this range we can neglect the floor brackets):

$$\frac{1}{n} \log B(m) = (1 - \mu)H\left(\frac{\epsilon\mu}{1 - \mu}\right) + 2(1 + \epsilon)\mu H\left(\frac{1}{1 + \epsilon}\right) - H(\mu)$$

Since  $\mu/(1 - \mu) \leq 2$  and assuming  $\epsilon < 1/4$ , we have  $H(\epsilon\frac{\mu}{1 - \mu}) \leq H(2\epsilon)$ . Using the easy inequality  $H(\frac{1}{1 + \epsilon}) \leq H(\epsilon) \leq H(2\epsilon)$ ,

$$\frac{1}{n} \log B(m) \leq (1 - \mu)H(2\epsilon) + 3\mu H(2\epsilon) - H(\mu) = (1 + 2\mu)H(2\epsilon) - H(\mu)$$

which can certainly be made negative in the range  $1/10 \leq \mu \leq 2/3$  by choosing  $\epsilon$  small enough. The contribution of  $B(m)$  in this range to the sum is, therefore, negligible.

In the lower range,  $m \leq n/10$ , we can bound  $B(m)$  from above by

$$B(m) \leq \frac{\binom{n}{\lfloor \epsilon m \rfloor} 5^m}{\binom{n}{m}}$$

If  $\lfloor \epsilon m \rfloor = 0$ , this is  $O(\frac{1}{\binom{n}{m}})$  which is  $O(1/n)$  for  $m = 1$  but  $o(1/n)$  for  $m > 1$ .

Now assume  $\lfloor \epsilon m \rfloor \geq 1$ . Using  $\binom{n}{k} \leq (\frac{ne}{k})^k$  we get

$$\binom{n}{\lfloor \epsilon m \rfloor} \leq \left( \frac{ne}{\lfloor \epsilon m \rfloor} \right)^{\lfloor \epsilon m \rfloor} \leq \left( \frac{2ne}{\epsilon m} \right)^{\epsilon m}$$

For the denominator we use  $\binom{n}{m} \geq (\frac{n}{m})^m$ , so

$$B(m) \leq \left( \frac{5 \left( \frac{2ne}{\epsilon m} \right)^\epsilon}{\frac{n}{m}} \right)^m = \left( 5 \left( \frac{2e}{\epsilon} \right)^\epsilon \left( \frac{n}{m} \right)^{\epsilon-1} \right)^m \leq \left( 10 \left( \frac{n}{m} \right)^{\epsilon-1} \right)^m$$

Finally,  $(n/x)^{(\epsilon-1)x}$  is decreasing for  $x \leq n/e$ , so  $B(m) = o(1/n)$  throughout the range. ■

We now turn to the proof of Theorem 1.

*Proof:* Fix a vertex  $z \in V(G)$  and two closed walks  $P_1, P_2$  starting and ending at  $z$ . We assume that each  $P_i$  does not *backtrack*, i.e. it contains no step of the form  $x, y, x$ . Moreover, we require  $P_1$  and  $P_2$  to be *independent* in the sense that each  $P_i$  contains an edge not contained in the other walk. This is possible since  $G$  is not unicyclic (whence the cycle space of  $G$  is at least 2-dimensional).

A random lift  $\tilde{G}$  of  $G$  is determined by assigning permutations in  $S_n$  to the oriented edges of  $G$ , and we let  $\sigma_i$  be the product of these permutations along  $P_i$ , inverting the ones assigned to edges that are traversed against their orientation. An immediate consequence of the definition of  $\tilde{G}$  is that, for every  $k \in [n]$ , there is a unique path in  $\tilde{G}$  starting at  $(z, k)$  and ending in  $(z, \sigma_i(k))$ . Call this path  $W_i(k)$ . From the covering property, the paths  $W_i(k)$  for fixed  $i$  are disjoint.

We let  $\epsilon > 0$  be as in Lemma 5. Almost surely, the permutations  $\sigma_1$  and  $\sigma_2$  are such that there are no  $\epsilon$ -bad sets  $A \subset [n]$ . This follows from Lemma 5 and the fact that  $\sigma_1, \sigma_2$  are indeed uniformly distributed, independent permutations

(their independence follows from the independence of the paths  $P_i$ ). In the rest of the proof, we show that this property of  $\sigma_1$  and  $\sigma_2$  ensures that  $\xi(\tilde{G}) > \xi_0$  where  $\xi_0 = \frac{\epsilon(1-\epsilon)}{2|V(G)|}$ .

Let  $T \subset V(\tilde{G})$  be a set with  $0 < |T| \leq \frac{1}{2}|V(\tilde{G})|$ . We need to show that  $|E(T, \bar{T})| \geq \xi_0|T|$ . For a vertex  $v \in V(G)$ , let  $T_v = T \cap \tilde{G}_v$  be the part of  $T$  that lies above  $v$ , let  $t_v = |T_v|$ , and set  $m = \max_{v \in V(G)} t_v$ . Note that  $|T| \leq m|V(G)|$ .

Suppose that  $t_u < (1 - \epsilon)m$  for some  $u \in V(G)$ . Let  $v$  be any vertex for which  $t_v = m$ , and let  $Q$  be a  $v - u$  path in  $G$ . Using again the unique lifting property of paths, there are  $n$  disjoint paths in  $\tilde{G}$  starting at  $\tilde{G}_v$  and ending in  $\tilde{G}_u$ . at least  $\epsilon m$  of those connect a vertex in  $T_v$  to a vertex *outside*  $T_u$ . Therefore  $E(T, \bar{T})$  contains at least  $\epsilon m$  edges, so  $\xi(T) \geq \epsilon/|V(G)| \geq \xi_0$  and we are done.

We can now assume that  $t_u \geq (1 - \epsilon)m$  for every  $u \in V(G)$  (this means that  $T$  is quite evenly distributed across the vertex fibers). It follows that  $m \leq \frac{2}{3}n$  and  $t_u > 0$  for every  $u \in V(G)$ . Let  $A = \{k \in [n] \mid (z, k) \in T_z\}$  where  $z$  is the vertex chosen above. Since  $|A| = |T_z|$  we have  $0 < |A| < \frac{2}{3}n$ . It follows that  $|A \cup \sigma_1(A) \cup \sigma_2(A)| \geq (1 + \epsilon)|A|$ . Without loss of generality, there are at least  $\frac{\epsilon}{2}|A|$  indices  $k \in A$  for which  $\sigma_1(k) \notin A$ . For such indices, the path  $W_1(k)$  contains an edge in  $E(T, \bar{T})$ , so

$$|E(T, \bar{T})| \geq \frac{\epsilon}{2}|A| \geq \frac{\epsilon(1-\epsilon)}{2}m$$

and since  $|T| \leq m|V(G)|$  we have  $\xi(T) \geq \frac{\epsilon(1-\epsilon)}{2|V(G)|} = \xi_0$  as required.  $\blacksquare$

In what follows we seek more precise information on the edge expansion of random lifts. For example, let  $G$  be a connected  $d$ -regular base graph. We suspect that the behavior of edge expansion in random lifts of  $G$  resembles that of random  $d$ -regular graphs, in the following sense: Let

$$\xi^*(d) = \limsup\{\xi(G) \mid G \text{ a } d\text{-regular graph}\}.$$

We conjecture that for every connected  $d$ -regular  $G$ , almost every lift  $\tilde{G}$  of  $G$

has edge expansion

$$\xi(\tilde{G}) \geq \min(\xi(G), \xi^*(d)).$$

(Perhaps even equality holds). We cannot decide this question in general, but we were able to explicitly bound from below the edge expansion of almost every lift of the bouquet graph  $B_l$  with one vertex and  $l$  loops. Lifts of these graphs are essentially equivalent to random  $2l$ -regular graphs so, not surprisingly, we recover Bollobas' result [9] in this case; only our proof (and our model) are somewhat different. A novel argument using  $\epsilon$ -nets allows us to improve his bounds slightly.

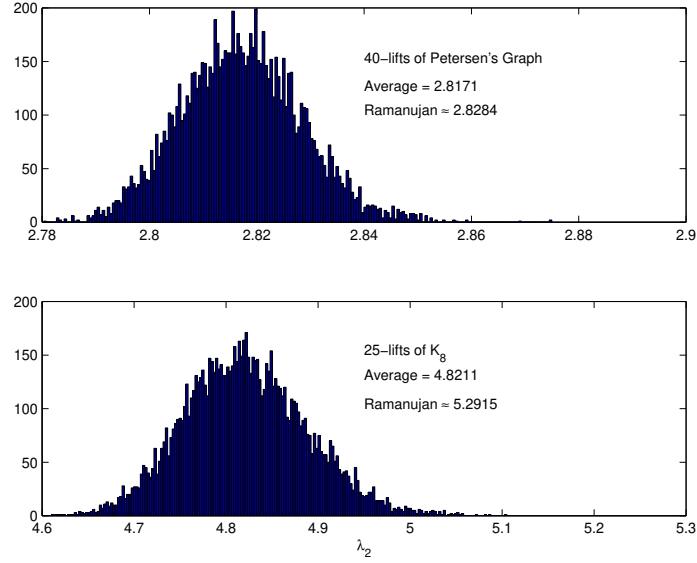
### 2.2.1 The Second Eigenvalue

It is well known (See, e.g., [2], [1]) that there are close relations between the expansion properties of a regular graph and the second-largest eigenvalue of its adjacency matrix. It is interesting to see what can be said about the spectrum of a lift, in particular random lifts. We denote the spectrum, as usual, by  $d = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_v$ .

A simple observation is that if  $\tilde{G}$  is a lift of  $G$ , then every eigenvalue of  $G$  is also an eigenvalue of  $\tilde{G}$ : any eigenvector  $f(v)$  of  $G$  can be lifted to  $\tilde{G}$  by setting  $\tilde{f}(\tilde{v}) = f(v)$  whenever  $\tilde{v} \in \tilde{G}_v$ . It follows that  $\lambda_2(\tilde{G}) \geq \lambda_2(G)$ , which should be compared with the fact that a lift cannot have better edge expansion than the base graph.

In Figure 2.1 we present empirical results for the distribution of the second eigenvalue of lifts of Petersen's graph and the complete graph  $K_8$ . In both cases,  $10^4$  random lifts were generated and the second eigenvalue was numerically calculated for each one. Even though the graphs are different, there is a striking similarity between the distributions.

The average value of  $\lambda_2$ , compared with the so-called Ramanujan bound  $2\sqrt{d-1}$ , (cf. [18], [19], [10]), are shown for the two cases. We see that the

Figure 2.1: Histograms of  $\lambda_2$  for Lifts of Two Graphs

(approximate) expectation of the second eigenvalue is below the Ramanujan bound. Recall that a  $d$ -regular graph is called *Ramanujan* if  $\lambda_2 \leq 2\sqrt{d-1}$ . A natural problem is to what extent random lifts satisfy the Ramanujan property. We hope to pursue this question, and similar ones concerning the spectra of lifts, in the future.

### 2.3 Random lifts of $B_l$

Let  $G = B_l$  be the pseudo-graph with one vertex and  $l$  loops. An  $n$ -lift  $\tilde{G}$  of  $G$  has a single vertex fiber  $V(\tilde{G}) = [n]$  and the edges are formed by choosing randomly  $l$  permutations  $\sigma_1, \dots, \sigma_l$  and connecting  $k$  to  $\sigma_i(k)$ , for every  $1 \leq i \leq l$  and  $k \in [n]$ . The result is, of course, a  $d = 2l$ -regular graph, possibly with loops and multiple edges. In the census of edges in  $E(S, \bar{S})$  for  $S \subset [n]$ , we count edges according to their multiplicity. Loops are never in  $E(S, \bar{S})$  so there is no

question of how to count them. In this section we prove

**Theorem 2** *Almost every lift  $\tilde{G}$  of  $G = B_l$  has edge expansion  $\xi(\tilde{G}) \geq \xi_0$ , where  $\xi_0$  is the smaller root of  $H(\xi/d) = (d-2)/d$ ,  $d = 2l$ .*

Let  $\xi_0(d)$  be the constant appearing in the Theorem. One can calculate  $\xi_0(4) \approx 0.440111$  and  $\xi_0(6) \approx 1.04371$ . As  $d$  grows,  $\xi_0(d) \rightarrow d/2$  since  $H^{-1}(1) = 1/2$ . This is quite easily seen to be tight (see [9]). Exactly the same formula for a lower bound on the edge-expansion appears in [9] for random regular graphs. As far as we know, however, the random model for  $\tilde{G}$  we used here is not known to be equivalent (technically, contiguous [14]) to the standard model for random regular graphs.

The proof has two main parts. The first (Proposition 3) is a bound on the probability that a random lift has edge expansion below a given  $\xi$ . The second, given in Propositions 4 and 5, is an analytical calculation which proves that for  $\xi < \xi_0$ , this bound tends to 0 as the order of the lift  $n \rightarrow \infty$ . We start with the bound, which is essentially a simple union bound:

**Proposition 3** *Let  $P(a, \xi)$  be the probability that a random lift of  $B_l$  contains a set  $S$  of cardinality  $|S| = a$  and  $\xi(S) \leq \xi$ . Then*

$$P(a, \xi) \leq \binom{n}{a}^{1-l} \sum_{\substack{0 \leq b_1, \dots, b_l \leq a \\ \sum b_i \geq a(l-\xi/2)}} \prod_i \binom{a}{b_i} \binom{n-a}{a-b_i} \quad (2.3)$$

*Proof:* Fix a subset  $S \subset [n]$  of the vertices of  $\tilde{G}$ . We can calculate the distribution of  $|E(S, \bar{S})|$  over the sample space of random lifts. Naturally, this random variable depends only on  $a$ , the size of  $S$ . For each permutation  $\sigma_i$ , let  $b_i = |\{k \in S | \sigma_i(k) \in S\}|$  be the number of elements of  $S$  that  $\sigma_i$  maps into  $S$ . The number of elements in  $S$  mapped *outside*  $S$  is  $a - b_i$ , and the number of elements *outside*  $S$  that  $\sigma_i$  maps into  $S$  is also  $a - b_i$ , so  $\sigma_i$  contributes  $2(a - b_i)$  to  $E(S, \bar{S})$  and

$$|E(S, \bar{S})| = 2la - 2 \sum_{i=1}^l b_i$$

whence

$$\xi(S) = 2l - \frac{2}{a} \sum_{i=1}^l b_i$$

so  $\xi(S) \leq \xi$  iff  $\sum b_i \geq a(l - \xi/2)$ .

The distribution of  $b_i$  is easy to calculate explicitly: the probability that  $b_i = b$  is

$$\frac{\binom{a}{b} \binom{n-a}{a-b}}{\binom{n}{a}}.$$

The bound in (2.3) is simply a union bound over all possible choices of such  $b_i$  and choices of  $S$  of size  $a$ . ■

The probability  $P(\xi)$  that the lift has edge expansion smaller than  $\xi$  is therefore bounded by

$$P(\xi) \leq \sum_{a=1}^{n/2} B(a)$$

where  $B(a)$  is the right-hand side of (2.3). Let us set

$$\lambda = 1 - \frac{\xi}{2l} = 1 - \frac{\xi}{d} \tag{2.4}$$

$$Q(a, b_1, \dots, b_l) = \binom{n}{a}^{1-l} \prod_i \binom{a}{b_i} \binom{n-a}{a-b_i}. \tag{2.5}$$

so

$$B(a) = \sum_{\substack{0 \leq b_1, \dots, b_l \leq a \\ \frac{1}{l} \sum b_i \geq \lambda a}} Q(a, b_1, \dots, b_l) \tag{2.6}$$

We seek the maximal  $\xi_0$  that guarantees  $P(\xi) \rightarrow 0$  as  $n \rightarrow \infty$  whenever  $\xi < \xi_0$ . As in the proof of Lemma 5, we need to treat separately the small sets  $a < \delta n$  and the large ones  $\delta n \leq a \leq n/2$ . Here  $\delta$  is some small constant that will be determined later (In fact,  $\delta = 1/10^7$  will suffice). For convenience, set

$$\begin{aligned} \sum_{a=1}^{n/2} B(a) &= \sum_{a=1}^{\delta n} B(a) + \sum_{a=\delta n}^{\frac{1}{2}n} B(a) = \\ &= B_1 + B_2 \end{aligned}$$

In the following two Propositions we prove  $B_2 \rightarrow 0$  and  $B_1 \rightarrow 0$  for the required  $\xi$ .

**Proposition 4** *If  $\xi < \xi_0$  then  $B_2 \rightarrow 0$ .*

*Proof:* Since the number of summands in  $B_2$  is  $O(n)$ , and the number of summands in  $B(a)$  is also polynomial in  $n$ , it suffices to prove an exponentially small upper bound on  $Q(a, b_1, \dots, b_l)$  in the range

1.  $\delta n < a \leq \frac{1}{2}n$
2.  $0 \leq b_i \leq a$
3.  $\frac{1}{l} \sum_{i=1}^l b_i \geq \lambda a$ .

Using (2.1) and (2.2) we obtain:

$$\frac{1}{n} \log_2 Q(a, b_1, \dots, b_l) \leq q(\alpha, \beta_1, \dots, \beta_l) \quad (2.7)$$

where  $\alpha = a/n$ ,  $\beta_i = b_i/n$  and

$$q(\alpha, \beta_1, \dots, \beta_l) = (1-l)H(\alpha) + \sum_{i=1}^l \alpha H\left(\frac{\beta_i}{\alpha}\right) + (1-\alpha)H\left(\frac{\alpha - \beta_i}{1-\alpha}\right). \quad (2.8)$$

If  $q$  is negative then  $Q \rightarrow 0$  exponentially as required. The following Lemma determines the points where  $q$  is maximized, which enables us to prove that  $q < 0$  in the required range, for the specified  $\xi$ .

**Lemma 6** *Let  $\Delta = \Delta_\lambda \subset \mathbb{R}^{l+1}$  be the region determined by*

1.  $0 \leq \alpha \leq 1/2$
2.  $0 \leq \beta_i \leq \alpha$
3.  $\frac{1}{l} \sum_{i=1}^l \beta_i \geq \lambda \alpha$ .

*with  $l \geq 2$  and  $1/2 < \lambda$ . The function  $q$  given by (2.8) attains its global maximum in  $\Delta_\lambda$  either when  $\alpha = 1/2$  and  $\beta_i = \lambda/2$  for all  $i$ , or when  $\alpha = 0$  and  $\beta_i = 0$  for all  $i$ .*

*Proof:* The concavity of  $H$  implies immediately that if  $\bar{\beta} = (\sum \beta_i)/l$  then

$$q(\alpha, \bar{\beta}, \dots, \bar{\beta}) \geq q(\alpha, \beta_1, \dots, \beta_l)$$

so we may assume that all  $\beta_i$  are equal,  $\beta_i = \beta$  where  $\beta \geq \lambda\alpha$ . Note that  $(\alpha, \beta_1, \dots, \beta_l) \in \Delta$  implies  $(\alpha, \bar{\beta}, \dots, \bar{\beta}) \in \Delta$ .

We fix  $\alpha$  and optimize over  $\alpha \geq \beta \geq \lambda\alpha$ . Now

$$\frac{\partial}{\partial \beta} q(\alpha, \beta, \dots, \beta) = \log\left(\frac{\alpha - \beta}{\beta}\right) - \log\left(\frac{1 - 2\alpha + \beta}{\alpha - \beta}\right) = \log\left(\frac{(\alpha - \beta)^2}{\beta(1 - 2\alpha + \beta)}\right).$$

This expression is always negative, because  $\beta \geq \lambda\alpha \geq \frac{1}{2}\alpha \geq \alpha^2$  since  $\lambda \geq 1/2 \geq \alpha$ . Therefore, to maximize  $q$ ,  $\beta$  needs to be as small as possible, namely  $\beta = \lambda\alpha$ . We are left with the univariate function  $f(\alpha) = q(\alpha, \lambda\alpha, \dots, \lambda\alpha)$ . Explicitly,  $f$  is given by

$$f(\alpha) = (1 - l)H(\alpha) + l\alpha H(\lambda) + l(1 - \alpha)H\left((1 - \lambda)\frac{\alpha}{1 - \alpha}\right).$$

and its derivatives are

$$\begin{aligned} f'(\alpha) &= [(1 - l)\log\frac{1 - \alpha}{\alpha} + lH(\lambda) - lH((1 - \lambda)\frac{\alpha}{1 - \alpha}) \\ &\quad + \frac{l}{1 - \alpha}\log\frac{1 - 2\alpha + \lambda\alpha}{(1 - \lambda)\alpha}] \end{aligned}$$

and

$$f''(\alpha) = \log(e) \frac{\lambda l - 1 - \alpha(l - 1)(2 - \lambda)}{\alpha(1 - \alpha)(1 - 2\alpha + \lambda\alpha)}.$$

Notice that  $f''(\alpha) = 0$  has at most one root, so  $f'$  has at most a single extremum. Further,

$$\lim_{\alpha \rightarrow 0^+} f'(\alpha) = -\infty$$

while

$$f'(\frac{1}{2}) = 2l \log \frac{\lambda}{1 - \lambda} > 0.$$

It follows that  $f'(\alpha) = 0$  has exactly one root in the range  $0 \leq \alpha \leq 1/2$ , so  $f(\alpha)$  has a unique minimum in that range, and the maximal value of  $f$  is achieved either at  $\alpha = 0$  or  $\alpha = 1/2$  as claimed. ■

Returning to the proof of Proposition 4, we set  $\lambda_0 = 1 - \frac{\xi_0}{2l}$ . Then  $\lambda_0 > 1/2$ , the Lemma applies and at the non-zero critical point we have

$$q(1/2, \lambda_0/2, \dots, \lambda_0/2) = (1-l) + lH(\lambda_0) = 0.$$

Since  $\lambda > \lambda_0$ ,  $q(1/2, \lambda/2, \dots, \lambda/2) < 0$ . We are assuming here  $a \geq \delta n$ , namely  $\alpha \geq \delta$ , so  $q$  is strictly negative throughout the whole range, and  $Q$  is exponentially small as required. ■

This proof cannot be applied to  $a = o(n)$  because  $q(0, 0, \dots, 0) = 0$ . For this reason we have to deal separately with  $B_1$ . In fact, small sets a.s. have larger edge expansion then that given by Theorem 2.

**Proposition 5** 1. If  $l = 2$  and  $\xi < 1$  then  $B_1 \rightarrow 0$ .

2. If  $l \geq 3$  and  $\xi < l$  then  $B_1 \rightarrow 0$ .

*Proof:* Fix  $a$  in the range  $1 \leq a < \delta n$ . Call  $\vec{b} = (b_1, \dots, b_l)$  *admissible* if  $0 \leq b_i \leq a$  and  $\frac{1}{l} \sum b_i \geq \lambda a$ . Suppose  $\vec{b} = (b_1 + 1, b_2, \dots, b_l)$  is admissible and  $b_1 \geq b_2 + 1$ . Then  $\vec{b}' = (b_1, b_2 + 1, \dots, b_l)$  is also admissible and since

$$\frac{Q(a, b_1, b_2 + 1, \dots, b_l)}{Q(a, b_1 + 1, b_2, \dots, b_l)} = \frac{b_1 + 1}{b_2 + 1} \cdot \frac{n - 2a + b_1 + 1}{n - 2a + b_2 + 1} \cdot \frac{(a - b_2)^2}{(a - b_1)^2}$$

we have  $Q(a, \vec{b}') > Q(a, \vec{b})$ . Starting from any  $\vec{b}$  and iterating this procedure we can find a  $\vec{b}'$  with  $Q(a, \vec{b}') > Q(a, \vec{b})$  and (up to rearrangements)  $\vec{b}' = (b_0, b_0, \dots, b_0 + 1, \dots, b_0 + 1)$  where  $b_0 + j/l \geq \lambda a$  for some  $0 \leq j < l$ .

Let  $\vec{b}_0 = (b_0, b_0, \dots, b_0)$ . A similar calculation shows that  $Q(a, \vec{b}_0) > Q(a, \vec{b}')$ . It is possible that  $\vec{b}_0$  is not admissible, but it is not far — in fact,

$$b_0 \geq \lambda a - \frac{l-1}{l} \geq \lambda a - 1 \quad (2.9)$$

We have shown that, given  $a$ , there exists an integer  $b_0$  satisfying (2.9) so that  $Q(a, b_1, \dots, b_l) < Q(a, b_0, \dots, b_0)$  for every admissible  $(b_1, \dots, b_l)$ . It follows that

$$B(a) \leq a^l Q(a, b_0, \dots, b_0) = a^l \binom{n}{a}^{1-l} \left( \binom{a}{b_0} \binom{n-a}{a-b_0} \right)^l = C(a, b_0).$$

Assume first that  $l \geq 3$ . Here  $\xi < l$  so  $\lambda > 1/2$ . Let us consider even values of  $a$ ; the odd case is similarly handled. Since  $b_0$  is an integer greater than  $\lambda a - 1$  we have  $b_0 \geq a/2$ . Decreasing  $b_0$  causes  $C(a, b_0)$  to increase, so we may take  $b_0 = a/2$ .

Set  $f(a) = C(a, a/2)$ . This is decreasing as a function of  $l$ , so it is enough to consider the case  $l = 3$ . We have

$$f(a) = \frac{a^3 \binom{a}{a/2}^3 \binom{n-a}{a/2}^3}{\binom{n}{a}^2}.$$

If  $\delta$  is small enough, then  $f(a) \geq f(a+2)$  for  $1 \leq a \leq \delta n$ . This follows from a straightforward calculation:

$$\begin{aligned} \frac{f(a+2)}{f(a)} &= \frac{(a+1)^5 (a+2)^8 (n - \frac{3}{2}a)^3 (n - \frac{3}{2}a - 1)^3 (n - \frac{3}{2}a - 2)^3}{a^3 (a/2 + 1)^9 (n-a)^5 (n-a-1)^5} \\ &\leq 2^9 (1 + \frac{2}{a})^{12} (a+2) \frac{(n - \frac{3}{2}a)^9}{(n-a-1)^{10}} \\ &\leq 2^{21} \frac{a+2}{n-a-1} \\ &\leq 2^{21} \frac{\delta}{1 - \delta - 1/n} \leq 1 \end{aligned}$$

Finally,  $f(2) = O(1/n)$  and  $f(4) = O(1/n^2)$  so  $\sum_1^{\delta n} B(a) = O(1/n)$ .

The case  $l = 2$  is done in the same manner. The difference is that  $C(a, a/2)$  is not decreasing in this case, but  $C(a, 3a/4)$  is. Since we just assumed  $\xi < 1$  for  $l = 2$ , then indeed  $\lambda > 3/4$  and  $b_0 \geq 3a/4 - 1/2$  and we proceed as before.

■

We close this section with some remarks concerning other base graphs. Let  $G = (V, E)$  be a connected graph. Trying to obtain tighter lower bounds on the edge expansion of lifts of  $G$  (tighter than the one obtained from the proof of Theorem 1), we can use a similar method as the one we used for  $B_l$ .

To do this, we need a bound on the probability that a set in the lift has many internal edges (i.e., small expansion). It is easy to write such a bound,

similar to equation (2.3), using variables  $a_v$  to parametrize the size of the set in each fiber  $\tilde{G}_v$  (instead of the single  $a$  in (2.3)) and  $b_{uv}$  for counting the edges connecting vertices in the fiber  $\tilde{G}_u$  to vertices in  $\tilde{G}_v$ . It is then necessary to find the critical value  $\xi_0$  of  $\xi$  for which the bound tends to 0.

Natural candidates for  $\xi_0$  are obtained by setting  $a_v = n/2$  and  $b_{uv} = b_0$  on the one hand, or else taking  $a_v = 0$  for all  $v \in V_0$  and  $a_v = 1$  for  $v \in V \setminus V_0$ . We conjecture that, in general, these are indeed the extremal cases. For regular graphs, this would imply the conjecture stated at the end of Section 2.2.

We have been able to carry out this program for the complete graphs  $K_r$  and the results are, once again, the same as those obtained for random regular graphs. It should be noted that, in many respects, a random lift of  $K_r$  is radically different from a random  $(r-1)$ -regular graph (for example, a random regular graph is not expected to cover a complete graph!).

## 2.4 An Improved Bound Through $\epsilon$ -nets

We now digress from the subject of random lifts to show how the method of  $\epsilon$ -nets can be applied to improve the lower bound on the edge expansion of regular graphs. We demonstrate the method in the standard model for random regular graphs, although the same can be done for random lifts. To prepare the ground, we recall briefly the proof of a lower bound on the edge expansion of random regular graphs from [9].

### 2.4.1 Random Regular Graphs

The standard probability space used to randomly generate regular graphs is the collection of perfect matchings among  $nd$  vertices, grouped into  $n$  clusters of size  $d$ . By shrinking each cluster into a single vertex, we get (with probability bounded away from zero for fixed  $d$  and growing  $n$ ) a random  $d$ -regular graph  $G$  on  $n$  vertices.

Let  $X$  be the set of  $nd$  vertices. For  $1 \leq \alpha \leq 1/2$  let  $\mathcal{S}_\alpha$  be the collection of subsets  $A \subset X$  of size  $\alpha nd$  that are unions of clusters. These are the ones that yield sets of size  $\alpha n$  in  $G$ . The probability that  $A \in \mathcal{S}_\alpha$  contains exactly  $\beta nd$  vertices that are paired to each other in the perfect matching is

$$p(\alpha, \beta) = \binom{\alpha nd}{\beta nd} \binom{(1-\alpha)nd}{(\alpha-\beta)nd} ((\alpha-\beta)nd)! \frac{N(\beta nd)N((1-2\alpha+\beta)nd)}{N(nd)}$$

where  $N(2v) = \frac{(2v)!}{2^v v!}$  is the number of perfect matchings of  $2v$  vertices. If  $\beta$  is so large so that the union bound  $|\mathcal{S}_\alpha|p(\alpha, \beta) = o(1)$ , then sets of size  $\alpha n$  with small expansion will not emerge in  $G$ . The relation between  $\beta$  and the edge expansion  $\xi$  is

$$\xi = d(1 - \frac{\beta}{\alpha}). \quad (2.10)$$

To calculate the union bound  $U = |\mathcal{S}_\alpha|p(\alpha, \beta)$  we note that  $|\mathcal{S}_\alpha| = \binom{n}{\alpha n}$ , and rewriting  $p(\alpha, \beta)$  we obtain <sup>1</sup>

$$U = \frac{\binom{n}{\alpha n}}{\binom{nd}{\alpha nd}} 2^{(\alpha-\beta)nd} \binom{nd/2}{\beta nd/2, 2(\alpha-\beta)nd/2}$$

and

$$\frac{1}{n} \log U = \left( H(\alpha)(1-d) + (\alpha-\beta)d + \frac{d}{2} H(\beta, 2(\alpha-\beta)) \right) (1 - o(1)).$$

It turns out that  $\alpha = 1/2$  is the extremal case (we will prove this in the next section), and so the critical  $\beta$  is determined by

$$H(2\beta) = \frac{d-2}{d}$$

which yields  $H(\xi/d) = \frac{d-2}{d}$  since  $\xi/d = 1 - 2\beta$  by (2.10). This is the same expression we obtained in Theorem 2.

## 2.4.2 $\epsilon$ -nets

At the heart of the above proof lies the union bound  $|\mathcal{S}_\alpha|p(\alpha, \beta)$ . This argument fails, however, to take into account that if a certain set  $A$  has high expansion

---

<sup>1</sup>We denote  $\frac{x!}{y!z!(x-y-z)!}$  as  $\binom{x}{y, z}$ . Correspondingly, we define  $H(x, y) = -x \log x - y \log y - (1-x-y) \log(1-x-y)$ .

(low  $\beta$ ), then other sets that have large intersection with  $A$  tend to expand, too. We wish somehow to exploit this fact. Let us first outline the method, in a general probability space.

**The general idea:** In the usual setting of the union bound approach, we are given a family  $\{A_i\}_{i \in I}$  of events and we wish to estimate  $\Pr(\cup A_i)$ . Suppose that  $\rho$  is some metric on  $I$  such that it is possible to estimate  $\Pr(A_i \triangle A_j)$  in terms of  $\rho(i, j)$ . It is then natural to define an  $\epsilon$ -net in  $I$ , namely a set  $J \subset I$  with the property that for every  $i \in I$  there is some  $j \in J$  such that  $\rho(i, j) \leq \epsilon$ .

Applying a union bound to the events  $\{A_j\}_{j \in J}$  may yield better results since there are fewer events in the  $\epsilon$ -net  $J$  than in the whole of  $I$ . It is then possible to apply these results to all events in  $I$  using the information on  $\Pr(A_i \triangle A_j)$ .

In our case, let us fix an  $\epsilon > 0$  and say that two sets  $A_1, A_2 \in \mathcal{S}_\alpha$  are  $\epsilon$ -close if

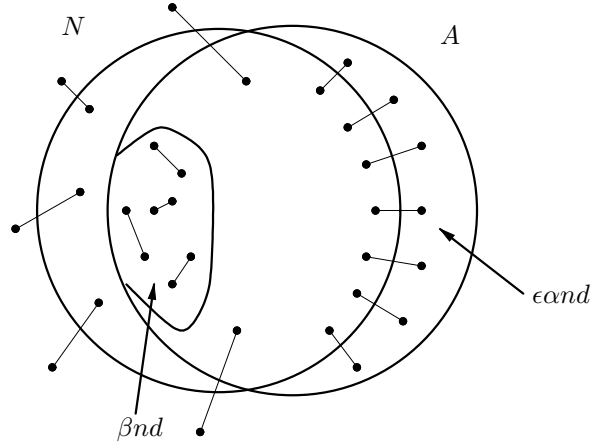
$$|A_1 \cap A_2| \geq (1 - \epsilon)\alpha n = (1 - \epsilon)|A_1|. \quad (2.11)$$

Now let  $\mathcal{N}_\alpha = \mathcal{N}_\alpha(\epsilon)$  be a subcollection of  $\mathcal{S}_\alpha$  with the property that for every  $A \in \mathcal{S}_\alpha$ , there is some  $N \in \mathcal{N}_\alpha$  such that  $N$  is close to  $A$ . This is the  $\epsilon$ -net. Since  $\mathcal{N}_\alpha$  is smaller than  $\mathcal{S}_\alpha$ , the union bound implies that the condition  $|\mathcal{N}_\alpha|p(\alpha, \beta') = o(1)$  already holds for some  $\beta' < \beta$ . That is, every  $N \in \mathcal{N}_\alpha$  contains at most  $\beta'nd$  vertices that are paired among themselves. This means that sets in the net have edge expansion  $\xi = d(1 - \beta'/\alpha)$ .

By the large intersection property (2.11), every  $A \in \mathcal{S}_\alpha$  has fewer than  $\beta''nd$  self-paired vertices, for some  $\beta''$  which is slightly larger than  $\beta'$  but still smaller than  $\beta$ . We are left with calculating  $\beta'$  and  $\beta''$ , and then optimizing  $\epsilon$  to get the best lower bound possible in this context.

We estimate  $\beta'$  as before, using the same formula for  $p(\alpha, \beta)$ . We only need to estimate the size of  $\mathcal{N}_\alpha(\epsilon)$ . For this we note that for every  $A \in \mathcal{S}_\alpha$ , the number of sets  $B$  that are  $\epsilon$ -close to  $A$  is at least

$$D = \binom{\alpha n}{\epsilon \alpha n} \binom{n(1 - \alpha)}{\epsilon \alpha n}.$$

Figure 2.2: The worst case when moving from  $N$  to  $A$ 

By a standard probabilistic argument (see, for example, [3], Theorem 2.2 in Chapter 1), we can find an  $\epsilon$ -net with size

$$|\mathcal{N}_\alpha| = |\mathcal{S}_\alpha| \frac{1 + \log(D+1)}{D+1}$$

so

$$\frac{1}{n} \log |\mathcal{N}_\alpha| \leq \left( H(\alpha) - (\alpha H(\epsilon) + (1-\alpha)H(\epsilon \frac{\alpha}{1-\alpha})) \right) (1 - o(1)).$$

A noticeable feature of this formula is that at  $\epsilon = 0$ , the derivative of  $\log |\mathcal{N}_\alpha|$  with respect to  $\epsilon$  is infinitely negative, so that the “immediate gain” from introducing the  $\epsilon$ -net is infinite.

Given the bound  $\beta'$  on self-paired vertices in members of  $\mathcal{N}_\alpha$ , we need to bound from above the number of such vertices in a set  $A \in \mathcal{S}_\alpha$ . Given  $A$ , let  $N \in \mathcal{N}_\alpha$  be a set  $\epsilon$ -close to  $A$ . Obviously, the worst case is when  $A$  contains all the self-paired vertices of  $N$ , and in addition all the vertices in  $A \setminus N$  are paired to ones in  $N$  (See Figure 2.2). This gives

$$\beta'' \leq \beta' + 2\alpha\epsilon$$

It is convenient to change variables and replace  $\beta$  by  $\mu = \beta/\alpha$ . To summarize, we have the following result.

**Proposition 6** *Let  $d \geq 3$  and  $\epsilon > 0$ . Define, for  $0 \leq \alpha \leq 1/2$  and  $0 \leq \mu \leq 1$ ,*

$$\begin{aligned} g_1(\alpha) &= H(\alpha) - \alpha H(\epsilon) - (1 - \alpha)H\left(\frac{\epsilon\alpha}{1 - \alpha}\right) \\ g_2(\alpha, \mu) &= -dH(\alpha) + \alpha(1 - \mu)d + \frac{d}{2}H(\alpha\mu, 2\alpha(1 - \mu)) \\ g(\alpha, \mu) &= g_1(\alpha) + g_2(\alpha, \mu), \end{aligned}$$

*and let  $\mu_0$  be such that  $g(\alpha, \mu) \leq 0$  for every  $\mu_0 \leq \mu \leq 1$  and  $0 \leq \alpha \leq 1/2$ . Then the edge expansion of a random  $d$ -regular graph is, almost surely, at least  $d(1 - \mu_0 - 2\epsilon)$ .*

*Proof:* For a given  $\epsilon$ ,  $g_1$  bounds the (logarithm of the) size of the  $\epsilon$ -net, and  $g_2$  bounds the probability that poorly expanding sets exist in the net. If their sum is negative, a random regular graph contains a net in which each set has edge expansion at least  $d(1 - \mu_0)$  by (2.10), and so every set in the graph has edge expansion at least  $d(1 - \mu_0 - 2\epsilon)$ . ■

We now determine  $\mu_0$  explicitly.

**Theorem 3** *For every  $d \geq 3$  there is an  $\epsilon_0 = \epsilon_0(d)$  such that if  $0 \leq \epsilon \leq \epsilon_0$  then the following holds. Let  $\mu_0$  be the larger solution of*

$$\frac{2}{d}(1 - H(\epsilon)) = 1 - H(\mu_0) \quad (2.12)$$

*then the edge expansion of a  $d$ -regular graph is almost surely at least  $d(1 - \mu_0 - 2\epsilon)$ .*

Notice that for  $\epsilon = 0$ , this gives the same edge expansion as Theorem 2. Also note that taking the larger solution implies  $\mu_0 \geq 1/2$ .

*Proof:* Define  $g_1$ ,  $g_2$  and  $g$  as in Proposition 6. Equation (2.12) is just the requirement that  $g(\frac{1}{2}, \mu_0) = 0$ . The point of the Theorem is that, indeed,  $\alpha = \frac{1}{2}$  is the critical case. We need to show that if  $\mu_0$  is defined by (2.12) then for any  $0 \leq \alpha \leq 1/2$  and  $\mu_0 \leq \mu \leq 1$  we have  $g(\alpha, \mu) \leq 0$ . First, for any fixed  $\alpha$ , the

function  $g(\alpha, \mu)$  has a unique maximum at  $\mu = \alpha$ . This follows from

$$\frac{\partial g}{\partial \mu} = \frac{\partial g_2}{\partial \mu} = -\alpha d + \frac{1}{2} \alpha d \log \frac{(2\alpha(1-\mu))^2}{\alpha\mu(1-2\alpha+\alpha\mu)}$$

which is easily seen to be negative for  $\mu \geq \alpha$  and positive otherwise. It therefore suffices to show that  $f(\alpha) = g(\alpha, \mu_0) \leq 0$ . We have  $f(0^+) = 0$  and  $f(1/2) = 0$  by definition. Also,  $f'(0^+) = -\infty$  and

$$f'(1/2) = -2\epsilon \log \frac{1-\epsilon}{\epsilon} + d(1-\mu) \log \frac{\mu}{1-\mu}$$

which is positive for  $\epsilon$  small enough. Therefore, if  $f$  is not negative throughout the interval  $0 \leq \alpha \leq 1/2$ , it must have at least three extremal points so  $f''$  must have at least two zeros. This is impossible for small  $\epsilon$  since for  $\epsilon = 0$ ,  $f''$  is a rational function with linear numerator. (In fact,  $f''$  is rational with quadratic numerator for any  $\epsilon$ ). ■

To find the best  $\epsilon$  we optimize  $2\epsilon + \mu$  under the constraint (2.12). Using Mathematica, we find that for  $d = 3$  the optimum is  $\epsilon \approx 0.00029082$  which gives  $\xi \approx 0.184682$ , compared with 0.184471 obtained without  $\epsilon$ -nets (as in [9]), an improvement of about 0.1%.

## Chapter 3

# Independence and Chromatic Number

In this chapter we present bounds on the chromatic number and on the independence number of typical random lifts, with  $G$  fixed and  $n \rightarrow \infty$ . For the independence number, upper and lower bounds are obtained as solutions to certain optimization problems on the base graph. For a base graph  $G$  with chromatic number  $\chi$  and fractional chromatic number  $\chi_f$ , we show that the chromatic number of typical lifts is bounded from below by  $\text{const} \cdot \sqrt{\chi/\log \chi}$  and also by  $\text{const} \cdot \chi_f/\log^2 \chi_f$  (trivially, it is bounded by  $\chi$  from above). We have examples of graphs where the chromatic number of the lift equals  $\chi$  almost surely, and others where it is a.s.  $O(\chi/\log \chi)$ . Many interesting problems remain open.

### 3.1 Introduction

In previous chapters we discussed the connectivity and expansion properties of random lifts. Here we investigate the independence number and chromatic

number.

### 3.2 The Independence Number

As usual, let  $\alpha(G)$  denote the maximal size of an independent set in a graph  $G$ . In this section we present bounds on the size and structure of independent sets in random lifts of  $G$ . Clearly, if  $\tilde{G}$  is an  $n$ -lift of  $G$  then  $\alpha(\tilde{G}) \geq n\alpha(G)$ , since if  $X$  is an independent set in  $G$  then  $\tilde{X} = \bigcup_{v \in X} \tilde{G}_v$  is an independent set in  $\tilde{G}$ . Note that if  $X$  is a maximal independent set in  $G$ , then  $\tilde{X}$  is maximal independent in  $\tilde{G}$ . However, there may be larger independent sets in  $\tilde{G}$  which are not a union of fibers.

For a set  $X \subset V(\tilde{G})$ , we let  $X_v = X \cap \tilde{G}_v$  be its intersection with the fiber over  $v \in V(G)$ , and we set  $x_v = |X_v|$ . A *profile* on  $G$  is a function  $f : V(G) \rightarrow [0, 1]$ . A set  $X$  determines a profile by  $f(v) = x_v/n$ , which represents the way  $X$  is distributed across the fibers. We present bounds on the probability that independent sets  $X$  exist, according to their profiles.

We need some definitions. Let

$$H(x_1, \dots, x_n) = - \sum_i x_i \log x_i - (1 - \sum_i x_i) \log(1 - \sum_i x_i)$$

be the entropy function (all logs are to the base 2), and set

$$I(x, y) = H(x) + H(y) - H(x, y)$$

letting  $I(x, y) = \infty$  if  $x + y > 1$ . For a profile  $f : V(G) \rightarrow [0, 1]$ , let

$$h(f) = \sum_{v \in V(G)} H(f(v)) - \sum_{[u, v] \in E(G)} I(f(u), f(v))$$

and

$$h_0(f) = \sum_{v \in V(G)} H(f(v)) - \log(e) \sum_{[u, v] \in E(G)} f(u)f(v).$$

It is easily seen that

$$I(x, y) \geq \log(e)xy$$

and therefore  $h(f) \leq h_0(f)$  for every profile  $f$ . Though  $h(f)$  yields tighter bounds, it is often easier to work with  $h_0$ . The fundamental result is the following:

**Lemma 7** *Let  $G$  be a graph, and let  $f : V(G) \rightarrow [0, 1]$  be a profile on  $G$ . The probability  $P$  that a random  $n$ -lift  $\tilde{G} \rightarrow G$  contains an independent set  $X$  with profile  $f$  satisfies  $P \leq 2^{nh(f)}$ .*

*Proof:* Fix a set  $X \subseteq V(\tilde{G})$  with  $x_v/n = f(v)$ . The probability that  $X$  is independent is

$$\prod_{[u,v] \in E(G)} \frac{\binom{n-x_v}{x_u}}{\binom{n}{x_u}}$$

since  $\binom{n-b}{a}/\binom{n}{a}$  is the probability that a random permutation on  $[n]$  takes a fixed  $a$ -set to a set disjoint from a fixed  $b$ -set. Therefore, the probability that *some* independent set  $X$  with this profile exists does not exceed

$$B = \prod_{v \in V(G)} \binom{n}{x_v} \prod_{[u,v] \in E(G)} \frac{\binom{n-x_v}{x_u}}{\binom{n}{x_u}}.$$

Using the bounds

$$\begin{aligned} \binom{n}{x_v} &\leq 2^{nH(x_v/n)} \\ \frac{\binom{n-x_v}{x_u}}{\binom{n}{x_u}} &\leq 2^{-nI(x_u/n, x_v/n)} \end{aligned}$$

we obtain

$$B \leq 2^{nh(f)}$$

as stated. ■

When  $h(f) < 0$ , this bound tends to 0, exponentially in  $n$ . In this case, almost surely no independent sets with profile  $f$  exist in  $\tilde{G}$ . The same holds, of course, when  $h_0(f) < 0$ . This leads us to define

$$\tilde{a}(G) = \max_f \left\{ \sum_v f(v) \mid h(f) \geq 0 \right\}$$

and

$$\tilde{a}_0(G) = \max_f \left\{ \sum_v f(v) \mid h_0(f) \geq 0 \right\}.$$

The following two results are immediate corollaries of Lemma 7.

**Theorem 4** *Almost every  $n$ -lift  $\tilde{G} \rightarrow G$  satisfies*

$$\alpha(\tilde{G}) \leq n\tilde{a}(G) \leq n\tilde{a}_0(G).$$

*Proof:* Given  $n$ , there are only  $(n+1)^{|V(G)|}$  profiles  $f$  such that  $nf(v)$  is an integer for every  $v$ . Therefore, the probability that there is an independent set  $X$  for which the profile  $f$  satisfies  $h(f) < 0$  is  $o(1)$ . In particular, this is the case for sets  $X$  with size larger than  $n\tilde{a}(G)$ , by the definition of  $\tilde{a}(G)$ . ■

Another useful bound concerns sets that have large intersection with each fiber.

**Corollary 3** *Let  $G$  be a graph with  $k$  vertices and  $\ell$  edges, and suppose that  $\beta$  satisfies  $kH(\beta) < \ell \cdot I(\beta, \beta)$ . A random  $n$ -lift  $\tilde{G} \rightarrow G$  almost surely contains no independent set  $X$  such that  $X \cap \tilde{G}_v \geq \beta n$  for every  $v$ .*

*Proof:* It is enough to show that a.s. there is no independent set with exactly  $n\beta$  vertices in each fiber. Let  $f_\beta$  be the constant profile  $f_\beta(v) = \beta$ . The probability that an independent set  $X$  exists with  $x_v = n\beta$  is smaller than  $2^{nh(f_\beta)}$ . Now

$$h(f_\beta) = kH(\beta) - \ell \cdot I(\beta, \beta) < 0$$

and so, almost surely, no such independent sets exist in  $\tilde{G}$ . ■

The *average degree* of a graph  $G$  with  $k$  vertices and  $\ell$  edges is  $\bar{d} = \bar{d}(G) = 2\ell/k$ .

**Corollary 4** *Let  $G$  be a graph with average degree  $\bar{d}$ , and suppose that  $\beta$  satisfies  $\bar{d}\beta/2 + \ln \beta \geq 1$ . A random  $n$ -lift  $\tilde{G} \rightarrow G$  almost surely contains no independent set  $X$  such that  $X \cap \tilde{G}_v \geq \beta n$  for every  $v$ .*

*Proof:* Since  $h(f_\beta) < h_0(f_\beta)$ , it suffices to show that  $h_0(f_\beta) < 0$  for  $\beta$  satisfying  $\bar{d}\beta/2 + \ln \beta \geq 1$ . Now  $h_0(f_\beta) = kH(\beta) - \log(e)\ell\beta^2$ . Using the easy inequality  $H(x) < -x \log(x) + \log(e)x$  for  $x > 0$ ,

$$\begin{aligned} h_0(f_\beta) &< -k\beta \log \beta + \log(e)k\beta - \log(e)\ell\beta^2 \\ &= -k\beta \log(e)(\ln \beta - 1 + \bar{d}\beta/2) \\ &\leq 0 \end{aligned}$$

as required. ■

Some remarks concerning the tightness of these bounds are in order. Let  $f$  be a profile, and for  $S \subset V(G)$  let

$$h(f, S) = \sum_{v \in S} H(f(v)) - \sum_{[u, v] \in E(G[S])} I(f(u), f(v)).$$

It is possible that  $h(f, S) < 0$  for some  $S$  even when  $h(f) \geq 0$ . Obviously, the proof of Theorem 4 shows that sets with such a profile almost surely do not exist — one just needs to restrict the proofs to the subgraph  $G[S]$  spanned by  $S$ . A natural conjecture is that the independence number of a random lift is, almost surely,  $(1 + o(1)) \sum f(v)$  where  $f$  is the profile which maximizes  $\sum f(v)$  while satisfying  $h(f, S) \geq 0$  for every  $S \subset V(G)$ .

### 3.2.1 A Greedy Lower Bound

To find an independent set in a lift  $\tilde{G}$  of  $G$ , we may proceed as follows. We fix an ordering of  $G$ 's vertices  $v_1, \dots, v_r$ , and choose the desired profile  $\alpha_1, \dots, \alpha_r$  with the intention of finding  $\alpha_1 n$  independent vertices in  $\tilde{G}_{v_1}$ ,  $\alpha_2 n$  independent vertices in  $\tilde{G}_{v_2}$ , and so on. At the  $k$ -th step of this procedure, some of the vertices in  $\tilde{G}_{v_k}$  cannot be used since they are adjacent to vertices selected in some  $\tilde{G}_i$  with  $i < k$ . However, if

$$\alpha_k < \prod_{\substack{i < k \\ [v_i, v_k] \in E}} (1 - \alpha_i),$$

we are likely to succeed in finding  $\alpha_k n$  additional vertices in  $\tilde{G}_k$  that are not adjacent to the previously selected ones.

**Proposition 7** *Let  $V(G) = \{v_1, \dots, v_r\}$  and suppose that  $\{\alpha_1, \dots, \alpha_r\}$  satisfy for every  $1 \leq k \leq r$*

$$0 \leq \alpha_k \leq \prod_{\substack{i < k \\ [v_i, v_k] \in E(G)}} (1 - \alpha_i). \quad (3.1)$$

*Let  $S = \sum \alpha_i$ . For every  $\varepsilon > 0$ , a random lift  $\tilde{G}$  of  $G$  almost surely contains an independent set of size  $n(S - \varepsilon)$ .*

*Proof:* By replacing each positive  $\alpha_i$  by  $\alpha - \varepsilon/r$ , we may assume that for some positive  $\delta$ ,

$$0 \leq \alpha_k \leq \quad (3.2)$$

This reduces  $S$  by  $\varepsilon$ , and so it suffices to show there is almost surely an independent set in  $\tilde{G}$  with  $\alpha_i n$  vertices in  $\tilde{G}_{v_i}$  for every  $1 \leq i \leq r$ , where the  $\alpha_i$  satisfy (3.2).

To see this, it is useful to consider the lift as being generated as we proceed. Namely, the edges between  $v_k$  and  $v_1, \dots, v_{k-1}$  are being randomly lifted only at the  $k$ -th stage. This yields the usual random model for the lift  $\tilde{G}$ . It is now obvious that the vertices in  $\tilde{G}_{v_k}$  that cannot be used are the union of random subsets of size  $\alpha_i n$ , for  $1 \leq i < k$  such that  $[v_i, v_k] \in E(G)$ . The probability that the complement of this union is smaller than  $n \prod (1 - \alpha_i) - \delta$  is  $o(1)$ , and the claim follows. ■

To get the most out of this result we need to find an optimal ordering of the vertices and then optimize over the  $\alpha_i$ . For example, when  $G = K_3$ , it is easy to see that the optimum is  $\alpha_1 = \alpha_2 = \frac{1}{2}$  and  $\alpha_3 = \frac{1}{4}$ , which yields  $S = \frac{5}{4}$ . This is, however, not optimal, as a random lift of  $K_3$  always contains an independent set of size  $\frac{3}{2}n(1 + o(1))$ .

In fact, a random lift of a cycle always contains an independent set with  $\frac{1}{2}n(1+o(1))$  vertices in each fiber. This suggests an improvement to Proposition 7. Given  $G$  with  $\chi(G) \geq 3$ , start by choosing an odd cycle  $C$  in  $G$ . For vertices  $v_i \in C$  we can guarantee  $\alpha_i = \frac{1}{2}$ , and for the other vertices we choose  $\alpha_k$  so as to satisfy (3.1). For example, this implies that lifts of  $K_4$  a.s. have an independent set with  $\alpha_i = \frac{1}{2}$  for  $i = 1, 2, 3$  and  $\alpha_4 = \frac{1}{8}$ , yielding  $S = \frac{8}{3} = 1.625$ . This is the current best lower bound. An upper bound can be obtained from theorem 4: A random  $n$ -lift of  $K_4$  a.s. does not contain an independent set with more than  $1.8363n$  vertices.

Finally, let us consider the case in which all the  $\alpha_i$ 's are equal to the same number  $x$ . Condition (3.1) then becomes  $x \leq (1-x)^d$  where  $d$  is the degree of  $v_k$  in the subgraph spanned by  $v_1, \dots, v_k$ . To minimize  $d$ , we proceed as follows.

The *degeneracy* of a graph  $G$  is defined as  $\text{dgn}(G) = \max \delta(H)$  where the maximum is over all induced subgraphs of  $G$  (see, e.g., [15]). Choosing the ordering  $v_1, \dots, v_n$  backwards, taking  $v_n$  to be a vertex of degree  $\delta(G)$  and in general taking  $v_{k-1}$  to be a vertex of minimum degree in  $G \setminus \{v_k, \dots, v_n\}$ , we can make sure that the exponent  $d$  above is never larger than  $\text{dgn}(G)$ . For  $d$  large, the solution to  $x = (1-x)^d$  is  $(1+o(1))\frac{\ln d}{d}$ .

### 3.2.2 The Independence Number of Lifts of Complete Graphs

Combining the lower bound from the previous section with Theorem 4, we can determine the asymptotics of the independence number of random lifts of  $K_{r+1}$  (as  $r \rightarrow \infty$ ).

**Proposition 8** *The independence number of a random  $n$ -lift  $\widetilde{K_{r+1}}$  of a complete graph a.s. satisfies*

$$\alpha(\widetilde{K_{r+1}}) = \Theta_r(n \log r)$$

*Proof:* For a lower bound, set  $\alpha_i = C \frac{\log r}{r}$  for all  $i$  in Proposition 7, with

$C = 1/3$ , for example. Condition (3.1) is satisfied since

$$\left(1 - C \frac{\log r}{r}\right)^r \geq r^{-2C} \geq C \frac{\log r}{r}.$$

(The left inequality follows from  $e^{-2\varepsilon} \leq 1 - \varepsilon$  for  $\varepsilon < 1/2$ ). Therefore, there exists an independent set of size at least  $\Omega(n \log r)$ . Alternatively, this follows from corollary 5.

We now turn to the upper bound, using Theorem 4. Let  $V(K_{r+1}) = [r+1]$  and let  $f : [r+1] \rightarrow [0, 1]$  be a profile. Set  $x_1 = f(1)$  and  $x_2 = f(2)$ . Let  $f'$  be obtained from  $f$  by averaging the values of  $f(1)$  and  $f(2)$ ; that is  $f'(1) = f'(2) = \bar{x} := \frac{x_1 + x_2}{2}$  and  $f'(i) = f(i)$  for  $i > 2$ . We claim that  $h(f') \geq h(f)$ .

Indeed, consider the difference  $h(f') - h(f)$ . Noting that  $H(a) + H(b) - I(a, b)$  is just  $H(a, b)$ , we get

$$h(f') - h(f) = H(\bar{x}, \bar{x}) - H(x_1, x_2) + \sum_{i>2} I(x_1, x_i) + I(x_2, x_i) - 2I(\bar{x}, x_i)$$

which is positive since  $H(\bar{x}, \bar{x}) \geq H(x_1, x_2)$  (entropy is maximal for the uniform distribution) and  $I(x, c)$  is convex in  $x$  for constant  $c$ .

It follows that if  $f$  is a profile with  $h(f) \geq 0$  then also  $h(\bar{f}) \geq 0$  where  $\bar{f}$  the constant profile obtained by averaging all the values of  $f$ . To bound the independence number of  $\widetilde{K_{r+1}}$  from above it therefore suffices to consider constant profiles.

Let  $f$  be the constant profile with values  $p$  at all vertices. We need to estimate the critical value of  $p$  for which  $h(f) = 0$ . We will consider  $h_0(f) = 0$  instead since this saves some calculation and essentially does not change the end result. We have

$$(r+1)H(p) = \log(e) \frac{r(r+1)}{2} p^2$$

which can be restated as (the  $o(1)$  terms are with respect to  $r \rightarrow \infty$ ):

$$-\frac{2}{r}(1 + o(1))p \ln p = p^2(1 + o(1)).$$

This is achieved at

$$p = \frac{2 \ln r}{r} (1 + o(1)),$$

so  $\sum f(i) \leq O(\log r)$  and  $\alpha(\widetilde{K_{r+1}}) \leq O(n \log r)$ . ■

The upper bound on the independence number provides a lower bound on the chromatic number of typical lifts of  $K_r$ : For a.e. lift,  $\chi(K_r) \geq \Omega(r/\log r)$ , where  $\chi(G)$  denotes the chromatic number of a graph  $G$ . In Section 3.3.3 we will see that this is essentially tight.

### 3.3 The Chromatic Number

We now turn to investigate the chromatic number of random lifts. There is an obvious bound for the chromatic number of *any* lift  $\tilde{G}$  of  $G$ : a coloring of  $G$  can be “lifted” to a coloring of  $\tilde{G}$  by coloring all the vertices in each fiber with the color of the corresponding vertex in  $G$ . This means that  $\chi(\tilde{G}) \leq \chi(G)$ . We are interested in finding non-trivial upper and lower bounds for the chromatic number of *typical* lifts.

To this end, it is convenient to define:

**Definition 6** *Given a graph  $G$ , let*

$$\tilde{\chi}_h(G) = \min\{k \mid \chi(\tilde{G}) \leq k \text{ for a.e. lift } \tilde{G} \rightarrow G\}$$

$$\tilde{\chi}_l(G) = \max\{k \mid \chi(\tilde{G}) \geq k \text{ for a.e. lift } \tilde{G} \rightarrow G\}$$

These are the essential upper and lower bounds on the chromatic number of lifts of  $G$ . For non-trivial graphs  $G$  we have  $2 \leq \tilde{\chi}_l(G) \leq \tilde{\chi}_h(G) \leq \chi(G)$ . A natural conjecture is that the chromatic number of random lifts satisfies a zero/one law and is essentially single-valued:

**Conjecture 1** *For every graph  $G$ ,  $\tilde{\chi}_l(G) = \tilde{\chi}_h(G)$ .*

We will see some examples of graphs  $G$  for which this conjecture is true. A trivial case is for bipartite graphs: If  $G$  is bipartite, *every* lift of  $G$  is also bipartite and so  $\tilde{\chi}_l = \tilde{\chi}_h = 2$  in this case. The following is a very simple lower bound on  $\tilde{\chi}_l$  for non-bipartite graphs:

**Lemma 8** *If  $\chi(G) \geq 3$  then  $\tilde{\chi}_l(G) \geq 3$ .*

*Proof:* Since  $\chi(G) \geq 3$ ,  $G$  contains an odd cycle. A random lift of an odd cycle a.s. contains a component that is an odd cycle as well (since the cycle structure of a random permutation a.s. contains an odd cycle). ■

This is a lower bound on the chromatic number of *typical* lifts. For every graph  $G$  there do exist bipartite lifts of every even order: the 2-lift  $\tilde{G}_2$  which assigns to every edge the transposition  $(1\ 2)$  is bipartite, and so is every lift of  $\tilde{G}_2$ . Similarly, by assigning the permutation  $(1\ 2\ 3)(4\ 5)(6\ 7) \cdots (n-1\ n)$  to every edge, one obtains 3-chromatic lifts of any odd order  $n$ .

**Example:** Let  $G$  be a graph with maximal degree  $\Delta(G) \leq 3$ . By Brooks' Theorem, every lift of  $G$  is 3-colorable, and we have just seen that it is a.s. *not* 2-colorable, so conjecture 1 holds for these graphs with  $\tilde{\chi}_l(G) = \tilde{\chi}_h(G) = 3$ . The simplest case where we do not know whether a zero-one law holds is  $G = K_5$ . We have  $3 \leq \tilde{\chi}_l(K_5) \leq \tilde{\chi}_h(K_5) \leq 4$ , and it is possible (though, we believe, unlikely) that both 3-chromatic and 4-chromatic random lifts occur with positive probability.

### 3.3.1 Lower Bounds

We begin with a lower bound on the chromatic number of almost all lifts of  $G$ , given in terms of the chromatic number of  $G$  itself.

**Theorem 5** *For every graph  $G$ ,*

$$\tilde{\chi}_l(G) \geq \sqrt{\frac{\chi(G)}{3 \log \chi(G)}}$$

*Proof:* Suppose that  $\tilde{G}$  is colored with the colors  $1, \dots, s$ . For each color  $i$ , let

$$V_i = \{v \in V(G) \mid i \text{ is the most frequent color in } \tilde{G}_v\}$$

breaking ties arbitrarily. For  $v \in V_i$  let  $I_v \subset \tilde{G}_v$  be the vertices over  $v$  that are colored by the color  $i$ . Note that  $|I_v| \geq n/s$  for every  $v$ , and that  $\bigcup_{v \in V_i} I_v$  is an independent set in  $\tilde{G}$ . Set  $\beta = 1/s$ , so that  $|I_v| \geq \beta n$ .

Let  $r = \chi(G)$ . Since the sets  $V_i$  partition  $V(G)$ , there is some  $j \in [s]$  such that the chromatic number of  $G[V_j]$  is at least  $k = r/s$ . It follows that  $G[V_j]$  contains a subgraph  $H$  with minimal degree  $\delta(H) \geq k - 1$ , so  $\bar{d} = \bar{d}(H) \geq k - 1$ . The set  $\bigcup_{v \in V(H)} I_v$  is independent, so by the condition of Corollary 4 we must have

$$\bar{d}\beta/2 + \ln \beta < 1.$$

(Note that for any subgraph  $H$  of  $G$ , the subgraph of  $\tilde{G}$  induced by the fibers above  $H$  is a random lift of  $H$ ; moreover, since the number of subgraphs  $H$  is bounded while  $n \rightarrow \infty$ , a statement that is true almost surely for a random lift is true almost surely for the lifts of all  $H$  simultaneously.) Then

$$\begin{aligned} \frac{r-s}{2s^2} - \ln s &< 1 \\ r &< s + 2s^2(\ln s + 1) \\ r &< 5s^2 \ln s \end{aligned}$$

where the last inequality holds for  $s \geq 3$  (we may ignore  $s = 2$  by Lemma 8). From this it follows easily that  $s^2 \geq \frac{r}{3 \ln r}$ . ■

Currently, we have no example where  $\tilde{\chi}_l(G)$  would be  $o(\chi(G)/\log \chi(G))$ . The following problem seems quite interesting:

**Problem 1** *Is it true that for every graph  $G$ ,*

$$\tilde{\chi}_l(G) \geq \Omega(\chi(G)/\log \chi(G))?$$

In Section 3.3.3 we will see that for large complete graphs,  $\tilde{\chi}_l$  is indeed as low as  $\Theta(\chi/\log \chi)$ . We now prove a lower bound on the chromatic number of a random lift in terms of  $\chi_f(G)$ , the *fractional* chromatic number of  $G$ .

**Definition 7** For a graph  $G$ , the fractional chromatic number  $\chi_f(G)$  is defined as the solution  $Z$  of the following optimization problem: Let  $\mathcal{I}$  be the collection of all independent sets in  $G$ . We seek a mapping  $\phi : \mathcal{I} \rightarrow \mathbb{R}^+$  such that  $\sum_{I \in \mathcal{I}} \phi(I) = Z$  is minimized while for every  $x \in V(G)$ ,

$$\sum_{\substack{I \in \mathcal{I} \\ x \in I}} \phi(I) \geq 1. \quad (3.3)$$

Any mapping  $\phi$  satisfying these condition is called a *fractional coloring* of  $G$ . If  $G$  is properly colored with  $\chi(G)$  colors, we can define a fractional coloring by letting  $\phi(I) = 1$  if  $I$  is a color class and  $\phi(I) = 0$  otherwise, whence  $\chi_f(G) \leq \chi(G)$ .

**Theorem 6** For every graph  $G$ ,

$$\tilde{\chi}_l(G) \geq \Omega\left(\frac{\chi_f(G)}{\log^2 \chi_f(G)}\right).$$

*Proof:* Suppose  $\tilde{G}$  is colored with the colors  $1, \dots, s$ . As in the proof of Theorem 5, we assume that each restricted lift  $\tilde{H} \rightarrow H$  satisfies the conclusion of Corollary 4. Our plan is to slightly modify the coloring of  $\tilde{G}$ , use the color classes to obtain weighted independent sets in  $G$  which satisfy the requirement (3.3), and finally estimate the total weight to bound  $\chi_f(G)$ .

The new coloring of the vertices of  $\tilde{G}$  uses the set of colors  $[s] \times [k]$ , where  $k = O(\log s)$ . Let  $C_i$  be the class of vertices in  $\tilde{G}$  that received (old) color  $i$ . The new color of a vertex  $\tilde{v} \in C_i \cap \tilde{G}_v$  is  $(i, j)$ , where  $j$  is such that  $\frac{1}{n}|C_i \cap \tilde{G}_v| \in [2^{-j}, 2^{-j+1}]$ . More precisely, this holds for all vertices such that  $\frac{1}{n}|C_i \cap \tilde{G}_v| \geq \frac{1}{2s}$ , while the vertices with  $|C_i \cap \tilde{G}_v| < \frac{n}{2s}$  remain uncolored. We note that at least half the vertices in each fiber *are* colored, because there were  $s$  color classes

originally, and we have only ignored classes that occupied less than  $1/2s$  of the vertices in a fiber. The total number of colors is now  $O(s \log s)$ .

Letting  $C_{i,j}$  denote the new color classes, we see that the sets  $C_{i,j}$  are disjoint, they cover at least half of each fiber, and each one is an independent set in  $\tilde{G}$ . Furthermore, by definition, whenever  $C_{i,j}$  intersects a fiber, the intersection has size at least  $2^{-j}n$ .

Let  $A_{i,j}$  be the “shadow” of  $C_{i,j}$ , namely the collection of vertices  $v$  in the base graph  $G$  for which  $C_{i,j} \cap \tilde{G}_v$  is nonempty. For any subgraph  $H$  of  $G[A_{i,j}]$ , there is an independent set with occupancy at least  $2^{-j}$  in the lift of  $H$ , and by Corollary 4 we must have  $\bar{d}(H) \leq O(2^j j)$ . In other words,  $A_{i,j}$  has degeneracy  $O(2^j j)$ , and in particular it can be partitioned into  $O(2^j j)$  independent sets  $I_{ijq}$ .

These sets  $I_{ijq}$  support our fractional coloring of  $G$ . We assign the weight of  $\phi(I_{ijq}) = 2^{-j+1}$  to each  $I_{ijq}$ , and from the construction of  $C_{i,j}$  we see that the total weight at each vertex is at least 1, and so we have a fractional coloring. The total weight of this fractional coloring is at most

$$\sum_{i=1}^s \sum_{j=1}^{O(\log s)} \sum_{q=1}^{O(2^j j)} 2^{-j+1} = s \sum_{j=1}^{O(\log s)} O(j) = O(s \log^2 s).$$

We have shown that a.s.  $\chi_f(G) \leq \chi(\tilde{G}) \log^2 \chi(\tilde{G})$ , from which the desired inequality follows. ■

Theorem 6 shows that if there are examples of graphs with  $\tilde{\chi}_l(G)$  much smaller than  $\chi(G)/\log \chi(G)$ , they may not be easy to find, since only very few constructions of graphs with  $\chi_f$  much smaller than  $\chi$  are known. The most notable such construction are perhaps the Kneser graphs (see e.g. [15], [23]), but these do not seem to provide a good example for our problem, at least in their usual version.

### 3.3.2 $\tilde{\chi}_l(G)$ and Variations on Graph Coloring

In proving the bounds in the previous section, we have transformed a coloring of a lift into a coloring, or fractional coloring, of the base graph. The latter is natural since a coloring of a fiber  $\tilde{G}_v$  defines a “mixture” of colors at  $v$ . These considerations lead us to constructions on the base graph which are variations on the classical notion of coloring.

Our first example is this: The degeneracy of a set  $S \subset V(G)$  is defined as  $\text{dgn}(G[S])$ , the degeneracy of the subgraph spanned by  $S$  (see Section 3.2.2). Thus an independent set has degeneracy 0 and a cycle-free set of vertices has degeneracy  $\leq 1$ . Coloring a graph entails covering the vertex set by 0-degenerate subsets. We now allow for using non-disjoint, non-independent subsets, where the contribution of a dependent set to the solution is reduced. Specifically, a set  $S$  contributes  $\frac{1}{\text{dgn}(S)+1}$  to every vertex  $x$  it contains, and we seek a weighted cover, so the total contribution should be 1 everywhere. Formally:

**Definition 8** *Given a graph  $G$ , let*

$$L(G) = \min \left\{ k \mid \exists S_1, \dots, S_k \subseteq V(G) \text{ such that } \sum_{i: x \in S_i} \frac{1}{\text{dgn}(S_i)+1} \geq 1 \text{ for all } x \in V(G) \right\}. \quad (3.4)$$

Again, we see that  $L(G) \leq \chi(G)$  since we can take the color classes as the sets  $S_i$ . A simple adaptation of the proof of Theorem 6 shows that

**Proposition 9** *For every graph  $G$ ,*

$$\tilde{\chi}_l(G) \geq \Omega \left( \frac{L(G)}{\log^2(L(G))} \right).$$

■

The following lift-free result now provides an alternative proof of Theorem 5:

**Lemma 9** *For any graph  $G$ ,*

$$\chi(G) \geq L(G) \geq \sqrt{\chi(G)/2}.$$

*Proof:* The left inequality is trivial, as observed above.

Let  $S_1, \dots, S_k$  be sets satisfying the condition in (3.4). Let

$$J = \{j \mid \text{dgn}(S_j) \leq 2k - 1\}.$$

The sets  $S_j$ ,  $j \in J$  cover  $V(G)$  since for every vertex  $x$ ,

$$\sum_{\substack{j \in J \\ x \in S_j}} \frac{1}{\text{dgn}(S_j) + 1} \geq \frac{1}{2}.$$

The result follows easily, since each  $S_j$  for  $j \in J$  can be colored using at most  $2k$  colors, and using a different palette of colors for each  $j$  we have colored  $G$  by less than  $2k^2$  colors. ■

It is possible that the lower bound in Lemma 9 is not tight, and any significant improvement on that bound brings us closer to solving Problem 1. The question of improving Lemma 9 has nothing to do with lifts, and it seems quite interesting in its own right.

We have defined  $L(G)$  as a “covering version” of the degeneracy. More generally, for any real-valued graph parameter  $\Psi(G)$  attaining values in  $[1, \infty)$  for all nonempty graphs, we can define

$$\begin{aligned} \text{cover-}\Psi(G) &= \min \left\{ k \mid \exists S_1, \dots, S_k \subseteq V(G) \text{ such that} \right. \\ &\quad \left. \sum_{i: x \in S_i} \frac{1}{\Psi(G[S_i])} \geq 1 \text{ for all } x \in V(G) \right\}. \end{aligned}$$

With this notation,  $L(G) = \text{cover-dgn}_1(G)$ , where  $\text{dgn}_1(G) = \text{dgn}(G) + 1$ . These notions for various graph parameters  $\Psi$  seem to be of independent interest.

Let us note that if  $\Psi$  has value 1 for independent sets then always  $\text{cover-}\Psi(G) \leq \chi(G)$ . On the other hand, if  $\Psi(G) \geq \chi(G)$  for all  $G$ , then considerations as in the proof of Theorem 6 show that  $\text{cover-}\Psi(G) \geq \chi_f(G)$ . We also have

$\text{cover-}\Psi(G) \geq \Omega(\sqrt{\chi(G)})$ , by the same argument as in the proof of Lemma 9. So for various graph parameters  $\Psi$  bounding  $\chi(G)$  from above, we get new graph parameters  $\text{cover-}\Psi$  sandwiched between  $\chi_f$  and  $\chi$ .

For the smallest of such parameters, namely  $\text{cover-}\chi(G)$ , we have

$$\text{cover-}\chi(G) \leq L(G)$$

and so we can try to approach Problem 1 by improving the just mentioned lower bound  $\text{cover-}\chi(G) \geq \Omega(\sqrt{\chi(G)})$ . But here we can show that this bound is tight. An example showing this is the following Kneser-type graph, denoted by  $D_s = (V, E)$ , where

$$\begin{aligned} V &= \text{all vectors in } (\{*\} \cup [s])^{3s} \text{ in which exactly } 2s \text{ components are } *, \\ E &= \{[u, v] : \text{there is no } i \text{ with } u_i = v_i \neq *\}. \end{aligned}$$

Using topological methods (see Appendix A), it can be shown that  $\chi(D_s) \geq s^2$ . On the other hand,  $\text{cover-}\chi(D_s) \leq O(s)$ : consider the sets  $A_1, \dots, A_{3s}$  where  $A_i = \{u \mid u_i \neq *\}$ . Each  $A_i$  can be colored with  $s$  colors, corresponding to the values of  $u_i$ . In addition, every vertex belongs to exactly  $s$  of the  $A_i$ 's. It follows that  $\text{cover-}\chi(D_s) \leq 3s$ .

### 3.3.3 An Upper Bound

We use a result of Kim [16] to give an *upper* bound on the chromatic number of almost all lifts of  $G$ , in terms of the maximal degree  $\Delta(G)$ . First we need a lemma, which says all subgraphs of constant size in a random lift are trees or unicyclic.

**Lemma 10** *Let  $G$  be a graph and let  $M$  be any fixed integer. A random lift  $\tilde{G}$  of  $G$  has the following property almost surely: Every subgraph  $H \subset \tilde{G}$  with  $|V(H)| \leq M$  also satisfies  $|E(H)| \leq M$ .*

*Proof:* Let  $X$  be a collection of at most  $M$  vertices in  $V(\tilde{G})$ . The probability that  $X$  spans more than  $M$  edges is smaller than  $2^{M^2}(\frac{1}{n-M})^{M+1}$ : the number of possible sets of edges is bounded by  $2^{M^2}$ , and for a given potential edge, the probability that it exists in the random lift never exceeds  $\frac{1}{n-M}$ , even if we condition on the situation of other potential edges.

The number of possible sets of vertices  $X$  is less than  $(n|V(G)|)^M$ . It follows that the expected number of exceptions to our claim is  $o(1)$  and by Markov's inequality the property holds almost surely. ■

Subgraphs of larger size are, of course, not so degenerate, for example because almost every lift of  $G$  is  $\delta(G)$ -connected (see chapter 1).

**Theorem 7** *Let  $G$  be a graph with maximal degree  $\Delta = \Delta(G)$ . Then*

$$\tilde{\chi}_h(G) \leq \frac{\Delta}{\ln \Delta}(1 + o_\Delta(1))$$

It is important not to confuse the two asymptotic variables in this result: The maximum degree  $\Delta$  and the order of lift, hidden in the definition of  $\tilde{\chi}_h$ .

*Proof:* Let  $\tilde{G}$  be a lift of  $G$ . From the lemma, we may assume that the triangles and 4-cycles in  $\tilde{G}$  are disjoint, because adjacent triangles or 4-cycles are subgraphs with fewer than 8 vertices and more edges than vertices.

Let  $T$  be the set of vertices that are contained in triangles or 4-cycles. Removing  $T$  from  $\tilde{G}$  yields a graph with maximum degree  $\Delta$  and girth at least 5. Kim proves in [16] that such a graph can be colored with  $\frac{\Delta}{\ln \Delta}(1 + o(1))$  colors. The vertices in  $T$  can be colored with three extra colors, and the result follows. ■

As a corollary, we can determine the asymptotic behaviour of  $\chi(\tilde{K}_r)$ , the chromatic number of random lifts of complete graphs.

**Corollary 5** *There exist constants  $A > B > 0$  such that*

$$A \frac{r}{\log r} \geq \tilde{\chi}_h(K_r) \geq \tilde{\chi}_l(K_r) \geq B \frac{r}{\log r}.$$

Indeed, the upper bound follows immediately from Theorem 7, and the lower bound was obtained in Section 3.2.2. ■

It is interesting to see that the drop in chromatic number of lifts of  $K_r$  occurs, in a sense, already for lifts of order 2. A random 2-lift of  $K_r$  is a graph  $\tilde{H}$  with  $V(\tilde{H}) = V_1 \cup V_2$ . Each edge  $e = [i, j]$  of  $K_r$  is covered either by a “parallel” pair of edge  $[i_1, j_1]$  and  $[i_2, j_2]$  or by a “crossed” pair  $[i_1, j_2]$  and  $[i_2, j_1]$ , depending on whether the permutation assigned to  $e$  is trivial or the transposition  $(1\ 2)$ .

The subgraphs  $\tilde{H}[V_1]$  and  $\tilde{H}[V_2]$  induced by  $V_1$  and  $V_2$  respectively are isomorphic, and are in fact drawn uniformly from the random graph model  $G(r, 1/2)$ . The chromatic number of such graphs [8] is a.s.  $(1 + o(1))r / \ln(r)$ , and so the chromatic number of  $\tilde{H}$  is at most  $2r / \ln(r)$ .

### 3.3.4 Persistence of the Chromatic Number

In contrast to the situation with complete graphs, here we present examples of graphs for which the chromatic number of typical lifts equals that of the base graph.

For a graph  $G$ , let  $H$  denote the graph obtained by replacing each vertex of  $G$  by an independent set of size  $r$ ; formally  $V(H) = V(G) \times [r]$  and  $E(H) = \{[(u, i), (v, j)] \mid [u, v] \in E(G), i, j \in [r]\}$ .

**Proposition 10** *For any graph  $G$  with  $\chi(G) \geq 2$ , put  $r = 3\chi(G) \log \chi(G)$ , and let  $H$  be constructed from  $G$  as above. Then almost every lift  $\tilde{H}$  of  $H$  has chromatic number  $\chi(\tilde{H}) = \chi(H) = \chi(G)$ .*

*Proof:* Let  $\tilde{H}$  be a random lift of  $H$  and let  $\tilde{c} : V(\tilde{H}) \rightarrow [s]$  be a coloring of  $\tilde{H}$ ,  $s = \chi(\tilde{H})$ . We define the mapping  $c : V(G) \rightarrow [s]$  by letting  $c(v)$  be the color occurring most often in the union of the fibers  $\bigcup_{i \in [r]} \tilde{H}_{(v, i)}$ , breaking ties arbitrarily.

We show that, almost surely,  $c$  is a proper coloring of  $G$ . For  $v \in V(G)$ , let  $X_{v,i}$  be the set of vertices of the lift  $\tilde{H}$  in the fiber  $\tilde{H}_{(v,i)}$  colored with the color  $c(v)$ , and let  $X_v = \bigcup_{i \in [r]} X_{v,i}$ . It suffices to show that if  $[u, v] \in E(G)$ , then a.s. there is an edge of  $\tilde{H}$  connecting a vertex of  $X_u$  to a vertex of  $X_v$ .

Set  $\alpha_i = |X_{u,i}|/n$  and  $\beta_i = |X_{v,i}|/n$ . We have  $|X_v| \geq \frac{nr}{s}$  by the choice of  $c(v)$ , and so  $\sum_i \beta_i = \lambda$  with  $\lambda \geq \frac{r}{s} = 3 \log s$ . We may actually assume  $\lambda = \frac{r}{s}$  (if  $X_v$  has more vertices we discard some of them). Similarly, we assume  $\sum_j \beta_j = \lambda$ .

The calculation is now very similar to the one in the proof of Lemma 7. If the  $\alpha_i$  and the  $\beta_i$  are fixed, the number of ways of choosing the sets  $X_u$  and  $X_v$  is at most

$$2^{n(\sum_i H(\alpha_i) + \sum_j H(\beta_j))}.$$

The probability that there is no edge between fixed sets  $X_u$  and  $X_v$  with the given sizes of the intersections with the fibers is bounded by

$$2^{n(\sum_{i,j} I(\alpha_i, \beta_j))}.$$

In order to show that  $c$  is a proper coloring for a particular choice of the  $\alpha_i$ 's and  $\beta_j$ 's (with  $\sum_i \alpha_i = \sum_j \beta_j = \lambda = 3 \log s$ ), it is enough to verify that

$$\sum_i H(\alpha_i) + \sum_j H(\beta_j) < \sum_{i,j} I(\alpha_i, \beta_j). \quad (3.5)$$

Since  $I(\alpha_i, \beta_j) \geq \log(e) \alpha_i \beta_j$ , the right-hand side is at least  $\log(e) \lambda^2$ . By concavity of the function  $H(\cdot)$ , the left-hand side is at most  $2rH(\lambda/r) = 2rH(1/s)$ . So it suffices to check that  $2H(1/s) < 3 \log(e) \log(s)/s$  for all  $s \geq 2$ , and this is indeed true (elementary calculus).

There are infinitely many possible choices of the  $\alpha_i$ 's and  $\beta_j$ 's. But since the difference of the right-hand side and of the left-hand side in (3.5) is bounded away from 0 uniformly for all  $\alpha_i$  and  $\beta_j$  (for fixed  $s$ ), each application of the above argument with a given collection of  $\alpha_i$ 's and  $\beta_j$ 's actually covers a small range of  $\alpha_i$ 's and  $\beta_j$ 's, and the whole space of possible  $\alpha_i$ 's and  $\beta_j$ 's is covered

by a bounded number of such ranges. So the chromatic number of the random lift is a.s. equal to  $\chi(G)$ . ■

### 3.4 Conclusion

This chapter is an initial study of the independence number and chromatic number of random lifts. Many interesting questions remain open, and some of them may be quite challenging. Their formulation is, in our opinion, one of the main contributions of the present work. Here we repeat those which we like most.

- (Precise values of  $\alpha$ ) For a given graph  $G$ , determine the constant  $A = A(G)$  such that the independence number of a random  $n$ -lift of  $G$  is a.s.  $(1 + o(1))An$ . (Azuma's inequality or similar tools imply that the independence number of the random lift is indeed concentrated and so  $A(G)$  is well defined.) In particular, what is  $A(K_4)$ ? Our results show that  $1.62 \leq A \leq 1.837$ . Is the upper bound obtained from the “first moment method”, formulated at the end of the first part of Section 3.2, tight?
- The question about  $A(K_4)$  seems to be closely related to the independence number of random 4-regular graphs. Again, the independence number of a random 4-regular  $n$ -vertex graph is a.s.  $(1 + o(1))Bn$  for some constant  $B$ , but the value of  $B$  seems to be unknown.
- (Zero-one law) Is there a zero-one law for the chromatic number of random lifts? In particular, is the chromatic number of a random lift of  $K_5$  a.s. equal to a single number (which may be either 3 or 4)?
- (Gap between chromatic numbers) Are there graphs  $G$  such that the chromatic number of their random lift is a.s.  $o(\chi(G)/\log \chi(G))$ , or perhaps even close to  $\sqrt{\chi(G)}$ ?

- What about the “cover-degeneracy”  $L(G)$ : can it be much smaller than  $\chi(G)$ ? (Currently we even have no example with  $L(G) < \chi(G)$ .) Can one find some reasonable graph parameter  $\Psi(G)$  that upper-bounds  $\chi(G)$  and such that  $\text{cover-}\Psi(G)$  is always close to  $\chi(G)$ ? (For example, what about  $\Psi(G)$  being the maximum degree of  $G$  plus 1?)



## Appendix A

# The Chromatic Number of $D_s$

We prove here a lower bound on the chromatic number  $\chi(D_s)$  of the graph  $D_s$  defined in Section 3.3.2. First we need a slightly different representation of this graph: we represent it as the disjointness graph of some set system. Let  $X = [3s] \times [s]$  be a ground set and let  $L_0 \subseteq 2^X$  be the system consisting of all  $s$ -element subsets of  $X$  having the first components of their elements pairwise distinct. The graph  $D_s$  is isomorphic to the graph  $D'_s$  with vertex set  $L_0$  and edges being pairs of disjoint sets. The isomorphism is defined as follows: to a vector  $v \in V(D_s)$  assign the set

$$\{(i, v_i) \mid i \in [3s], v_i \neq *\} \in L_0 = V(D'_s)$$

We use Theorem 4.13 from [21], a slight generalization of theorems by Sarkaria [23], [24]. Let  $\sigma^N$  denote the  $N$ -dimensional simplex with vertex set  $[N + 1]$ . In what follows, simplicial complexes are abstract (i.e. hereditary set systems), and  $\|K\|$  denotes the polyhedron of a simplicial complex  $K$ .

**Theorem 8** *Write  $\Delta = \sigma^N$  and let  $K$  be a subcomplex of  $\Delta$ . Put  $L = \Delta \setminus K$*

and let  $L_0$  be the set of all inclusion-minimal simplices in  $L$ . Suppose that

$$\chi : L_0 \rightarrow 2^{[m]} \setminus \{\emptyset\}$$

is a coloring of the simplices of  $L_0$  by nonempty subsets of an  $m$ -element set, which satisfies the condition

$$\sigma_1 \cap \sigma_2 = \emptyset \Rightarrow \chi(\sigma_1) \cap \chi(\sigma_2) = \emptyset. \quad (\text{A.1})$$

Then for  $d \leq N - m - 1$  there is no embedding  $\|K\| \rightarrow \mathbb{R}^d$ .

In our case, the vertex set of  $\Delta$  is identified with  $X = [3s] \times [s]$ , so  $N = 3s^2 - 1$ . The set system  $L$  (the supersets of the sets in  $L_0$ ) consists of all subsets of  $X$  with at least  $s$  distinct first components. Thus,  $K$  consists of all sets with at most  $s - 1$  distinct first components. So  $K$  is a simplicial complex of dimension  $(s - 1)s - 1$ , and thus it can be embedded into  $\mathbb{R}^d$  with  $d = 2s(s - 1) - 1$ . Hence, by the theorem, there is no coloring of the sets in  $L_0$  by nonempty subsets of  $[m]$  for

$$m = N - d - 1 = 3s^2 - 1 - 2s(s - 1) + 1 - 1 \geq s^2.$$

In particular, there is no proper coloring of the graph  $D_s$  by at most  $s^2$  colors, so  $\chi(D_s) \geq s^2$ .

# Bibliography

- [1] Noga Alon. Tools from higher algebra. In R. L. Graham, M. Grötschel, and L. Lovász, editors, *Handbook of Combinatorics*, volume 2, page 1754. Elsevier Science, 1995.
- [2] Noga Alon and Vitali D. Milman.  $\lambda_1$ , isoperimetric inequalities for graphs and superconcentrators. *Journal of Combinatorial Theory Ser. B*, 38:78–88, 1985.
- [3] Noga Alon and Joel Spencer. *The Probabilistic Method*. John Wiley, 1992.
- [4] Dana Angluin. Local and global properties in networks of processors. In *Proceedings of the Twelfth Annual ACM Symposium on Theory of Computing*, Los Angeles, California, 1980.
- [5] Dan Archdeacon and R. Bruce Richter. On the parity of planar covers. *Journal of Graph Theory*, 14(2), 1990.
- [6] Hyman Bass and Ravi Kulkarni. Uniform tree lattices. *Journal of the American Mathematical Society*, 3(4):843–902, October 1990.
- [7] Claude Berge. *Graphs and Hypergraphs*. North-Holland, 1975.
- [8] Béla Bollobás. *Random Graphs*. Academic Press, 1985.
- [9] Béla Bollobás. The isoperimetric number of random regular graphs. *European Journal of Combinatorics*, 9:241–244, 1988.

- [10] Joel Friedman, Jeff Kahn, and Endre Szemerédi. On the second eigenvalue in random regular graphs. In *Proc. 21st ACM Symposium on the Theory of Computation*, pages 587–598, 1989.
- [11] J. L. Gross and T. W. Tucker. Generating all graph coverings by permutation voltage assignments. *Discrete Mathematics*, 18:273–283, 1977.
- [12] J. L. Gross and T. W. Tucker. *Topological Graph Theory*. Wiley-Interscience, 1987.
- [13] Jonathan L. Gross. Every connected regular graph of even degree is a schreier coset graph. *Journal of Combinatorial Theory Ser. B*, 22(2):227–232, 1977.
- [14] Svante Janson. Random regular graphs: Asymptotic distributions and contiguity. *Combinatorics, Probability and Computing*, 4:369–405, 1995.
- [15] Tommy R. Jensen and Bjarne Toft. *Graph Coloring Problems*. Wiley-Interscience, New York, NY, 1994.
- [16] Jeong Han Kim. On Brooks’ theorem for sparse graphs. *Combin. Probab. Comput.*, 4(2):97–132, 1995.
- [17] Frank Thomson Leighton. Finite common coverings of graphs. *Journal of Combinatorial Theory Ser. B*, 33:231–238, 1982.
- [18] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan conjectures and explicit construction of expanders. In *Proc. ACM Symposium on the Theory of Computation*, pages 240–246, 1986.
- [19] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8:261–277, 1988.
- [20] Wolfgang Mader. Grad und lokaler zusammenhang in endlichen graphen. *Math. Ann.*, 205:9–11, 1973.

- [21] Jiří Matoušek. Topological methods in combinatorics and geometry. *Lecture notes available at:*  
<http://www.ms.mff.cuni.cz/acad/kam/matousek/kt.ps.gz>, 1994.
- [22] Seiya Negami. The spherical genus and virtually planar graphs. *Discrete Mathematics*, 70:159–168, 1988.
- [23] Karanbir S. Sarkaria. A generalized Kneser conjecture. *J. Combinatorial Theory*, 49:236–240, 1990.
- [24] Karanbir S. Sarkaria. A generalized van Kampen-Flores theorem. *Proc. Amer. Math. Soc.*, 11:559–565, 1991.
- [25] John Stallings. Topology of finite graphs. *Inventiones Mathematicae*, 71:551–565, 1983.
- [26] John Stillwell. *Classical Topology and Combinatorial Group Theory*. Springer-Verlag, 1993.