Graph Reconstruction and Structure

Jane Tan
St John's College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2023

In memory of Ciaran

Acknowledgements

To Alex, for your kindness and generosity,
for giving me the most wonderful combinatorics family,
for keeping me excited about mathematics,
and for signing so many forms;

To Carla, the big sister I always wanted
for all the walks, chats and check-ins;

To Tom and Candy, for being there for me from the start;

To Gal, Little Alex, Dániel, Freddie and Antonio,
for your friendship and advice on all things;

To Emil and Michael of 007 and Youri and Lukas of 008,
for the antics, the adventures and always being close by;

To Youri again, for the figures, the food,
and maybe a few other things;

To Alice, Luci, Filippos, Christoph×2, Lukas and Jan,
for adopting me into your group when it really mattered;

To Ruben and Minying, for finding me in that first MCR event;

To Stuart, for your steady support and understanding;

To Adele, Chris and Caroline, for moving hemispheres with me;

To Brendan, for introducing me to research and to graphs;

and to my family, for always being on the phone.

Thank you all!

Abstract

This thesis explores the difficulties of capturing graph structure, and the recent approaches that have been able to do so effectively in some settings.

The famous Graph Reconstruction Conjecture (Kelly-Ulam, 1941) states that every graph on at least 3 vertices is uniquely determined by its deck of vertex-deleted subgraphs. Despite sustained efforts, it remains wide open. In Part I, we break new ground on several variants of this reconstruction problem. Two such settings are even harder than the classical case; we reconstruct graphs or their properties from decks of k-vertex deleted subgraphs as well as from incomplete decks where ℓ subgraphs are missing. While most general results in literature only apply when k and ℓ are small constants, in both cases we develop techniques that facilitate significant leaps to even handle deficits that are linear in the number of vertices. For the third variant, we introduce axioms to define a family of reconstruction problems in which decks are defined by 'switching' at a vertex instead of deletions. This encompasses a well-known problem of Stanley from 1985. Our key result is a new sufficient condition for switching-reconstructibility that applies under this generic framework.

Part II examines the problem of formulating approximate characterisations of graphs in terms of simpler building blocks. A recent focal point has been on graph product structure theory, which allows many graph classes to be described using graphs of bounded treewidth. Having first appeared in 2019, this theory is still actively under development, but has already facilitated progress on longstanding problems. As a new application in the intersection of structural and geometric graph theory, we use it to prove that all proper minor-closed classes of graphs have touching representations by comparable boxes in bounded dimensions. We also show that product structure can, in turn, describe the bounded treewidth building blocks from the original theorems using even simpler parts.

Statement of Originality

In Part I, the results of Sections 4.1.3, 4.2 and 5.1 are based on joint work with Carla Groenland, Tom Johnston and Alex Scott [70]. Section 4.3 is based on joint work with Carla Groenland, Tom Johnston, Andrey Kupavskii, Kitty Meeks and Alex Scott [69]. An earlier version appears in the DPhil thesis of Carla Groenland. The formulation in Section 3.4.2 and results in Section 5.2 are based on unpublished work with Brendan McKay and Béata Faller.

In Part II, Chapter 8 is based on joint work with Zdeněk Dvořák, Daniel Gonçalves, Abhiruk Lahiri and Torsten Ueckerdt [48]. Chapter 9 is based on work from a paper with Rutger Campbell, Katie Clinch, Marc Distel, J. Pascal Gollin, Kevin Hendrey, Robert Hickingbotham, Tony Huynh, Freddie Illingworth, Youri Tamitegama and David Wood [29].

Contents

List of Symbols

$\mathcal{P}(S)$	Power set of S	9
$\mathbb{N}$	$\{1, 2, \ldots\}$	9
$\mathbb{N}_0$	$\{0, 1, 2, \ldots\}$	9
$[n]$	$\{1, \ldots, n\}$	9
$\lvert V(G)\rvert, \lvert G\rvert$	Order of a graph	9
$G^c, \overline{G}$	Graph complement	15
$\omega(G)$	Clique number	17
$\alpha(G)$	Independence number	17
$\chi(G)$	Chromatic number	17
$\sqcup$	Disjoint union (of graphs or sets)	24, 63
$n_H(G)$	# vertex sets in G inducing a copy of H	14, 39
$\widetilde{n}_H(G)$	# instances of H in G	14, 39
H_e	H-extension	44
$B_d(H, G)$	The (closed) d-ball of H in G	44
$m_d(H_e, G)$	# copies of H in G with d-ball isomorphic to H_e	44
$n(H_e, H'_e)$	# sub-H-extensions of H'_e isomorphic to H_e	45
$T_1 \frown T_2$	Tree grafting operation	54
$[a, b]$	Set of integers between a and b inclusive	59
$d_t(G), d_{<t}(H)$	# vertices in G with degree equal to/less than t	59
$P^n_k(x), \alpha_k(x, n)$	Binary Krawtchouk polynomials	73, 78
$\operatorname{tw}(G)$	Treewidth	96
$\square$	Graph Cartesian product	98
$\boxtimes$	Graph strong product	100
$\sqsubseteq$	Comparability relation for boxes	106
$\dim_{cb}$	Comparable box dimension	106
$\chi_a(G)$	Acyclic chromatic number	108
$\chi_s(G)$	Star chromatic number	108
$\operatorname{tpw}_c(G)$	c-tree-partition-width	124

Chapter 1

Introduction

In what ways can we capture the structure of a graph using collections of associated graphs as descriptors? More pertinently, how much information does such a characterisation require, and how simple can those descriptors be? Good answers naturally vary depending on what exactly we wish to preserve, be it the entire structure of a graph, a collection of distinguishing properties of a family of graphs, or one specific feature. These questions underpin an important and extensive body of research which touches many areas within graph theory. Indeed, a basic strategy for studying complex structures and their properties is to try to extract a simpler characterisation using objects that we can get our hands on, and then reassemble these in some way to infer what we need about the original structure.

This thesis comprises two parts tied together by the underlying theme of understanding particular graph characterisations. Part I focuses on graph reconstruction, where there is some immediate bad news: it is not even known whether the full collection of vertex-deleted induced subgraphs of an unknown graph G is enough to determine G fully up to isomorphism. Nonetheless, it is enough to recover graphs in certain classes and to deduce many graph parameters. In fact, some of these can be recovered from much less data. Part II lies within the area of structural graph theory, where some remarkably general characterisations have been developed. Our focus is on graph product structure theory, which provides a suite of tools particularly suited to studying sparse graphs. This theory is relatively young but has seen rapid progress and rise to significance, making for an exciting outlook.

While the parts are essentially self-contained, they have a similar structure interleaving the literature with new results consisting of both concrete theorems and more open-ended theoretical contributions. We now provide more detailed introductions highlighting the new results that appear in each.

1.1 Contributions in Part I

Broadly speaking, reconstruction problems ask whether a combinatorial object is uniquely determined (up to some notion of equivalence) by a collection of its subobjects or associated data. Questions of this type are interesting not only mathematically, due to their fundamental nature, but also because they arise naturally in various diverse fields. A standard example is the reconstruction of a genome from sequenced DNA fragments in bioinformatics.

Returning to graphs, the key information that we work with in these problems is the *deck* of a graph G, denoted $\mathcal{D}(G)$, which in the classical setting can be defined as the multiset consisting of all vertex-deleted subgraphs of G, i.e. $\mathcal{D}(G) = \{G - v : v \in V(G)\}$. The elements of the deck are called *cards*. For example, if $G \cong C_n$ (the cycle on n vertices), then its deck is the multiset $\mathcal{D}(G) = \{P_{n-1}, P_{n-1} \ldots, P_{n-1}\}$ consisting of n cards where each one is a path on $n-1$ vertices. In this case, it is not difficult to show that the deck by itself is enough to recover the original graph.

On the other hand, the two graphs of order two (K_2 and $K_1 \sqcup K_1$) have the same deck since each card is just K_1. In 1941, Kelly and Ulam [85] made the striking conjecture that these are the only non-reconstructible graphs. At first glance, this conjecture seems quite believable since even one card already contains a lot of information about our graph. However, this belies the complexities at play, such as difficulties relating to graph isomorphism. Despite a steady flow of results in the intervening years, the Kelly-Ulam Reconstruction Conjecture remains very much open and continues to fascinate. Indeed, Harary [76] called it a 'graphical disease' due to the widespread interest it has piqued, and it is the first entry in Bondy's list of 'Beautiful conjectures in graph theory' [18].

By now, this problem has blossomed into an active area of research with work in many directions, including directly toward the conjecture as well as on related variants. Partial results in the former category usually state that the conjecture is true for specific classes of graphs, or that certain graph parameters are reconstructible from the deck so we can learn something about our unknown graph even if we do not know how to recover it fully. Instances of the latter are obtained by varying the definition of a deck, so as to investigate how far we can reduce the information of the deck and still reconstruct certain classes or properties, or whether differently obtained graphs determine the full structure instead.

The purpose of Chapter 2 is to give an overview of the status of work on the classical problem that we have defined above. Chapter 3 then does the same for a collection of variants that we have identified as the most promising or well-studied. Our account is not intended to be comprehensive, but rather present a fairly broad introduction paying special attention to aspects that have seen relatively little treatment in the many surveys that exist, or those that are relevant to new, recent, and potential future results. Chapter 6 discusses some of these future directions. Our new contributions are split between Chapter 3, where we contextualise our work, and Chapters 4 and 5, which contain the proofs of theorems that use direct counting and algebraic methods respectively.

One of the first nontrivial results in reconstruction was that all trees are reconstructible [85, 86]. In the first paper where this appears, Kelly [86] also quietly introduced the idea of considering the collection of k-vertex deleted subgraphs for $k > 1$: the *ℓ-deck* of a graph G is the multiset of all subgraphs induced by subsets of $V(G)$ with $\ell = n - k$ elements. For example, the 2-deck $\mathcal{D}_2(C_4 + e)$ consists of five copies of K_2 and one copy of $K_1 \sqcup K_1$. Note that if $|V(G)| = n$, then $\mathcal{D}(G) = \mathcal{D}_{n-1}(G)$. As before, for any fixed ℓ, this raises the question of whether the deck $\mathcal{D}_\ell(G)$ uniquely determines G or some of its properties. This 'small cards' variant is discussed in greater detail in Section 3.2.

An easy counting argument proves that $\mathcal{D}_\ell(G)$ determines $\mathcal{D}_{\ell-1}(G)$. Thus, for any class of graphs that have been shown to be reconstructible from the $(n-1)$-deck, we can ask: what is the smallest ℓ for which graphs in that class can be reconstructed from the ℓ-deck? There has been little progress on this question, even for simple classes like trees (see Section 3.2). In 1990, Nýdl [130] conjectured that trees are reconstructible from their ℓ-decks whenever $\ell \geq \lfloor n/2 \rfloor + 1$. The best bound in the literature comes from an earlier result proved by Giles [64] in 1976 when put together with one of Manvel [118] from 1974. These imply that trees with at least 6 vertices are reconstructible from their $(n-2)$-decks.

We make a significant improvement on this bound allowing a linear number of vertices to be removed:

Theorem 1.1. *Any n-vertex tree T can be reconstructed from $\mathcal{D}_\ell(T)$ whenever $\ell > \frac{8n}{9} + \frac{4}{9}\sqrt{8n+5} + 1$.*

The proof of this theorem is split into several parts that occupy Section 4.2 and includes a recognition statement that was independently developed. A key ingredient is a new counting lemma that allows us to count certain subgraphs in which we can keep track of a marked structure. This counting tool is of independent interest, and is the topic of Section 4.1.3.

There is more to be said when we consider the normally more approachable problem of reconstructing properties or parameters in the context of small cards. In general, we say that a property is *reconstructible* if when G satisfies the property and $\mathcal{D}(G) = \mathcal{D}(G')$, then G' also satisfies the same property. That is, one can determine from $\mathcal{D}(G)$ whether or not G satisfies the property. Simple examples in the classical case include the number of edges, degree sequence and connectedness. These have also been studied for decks of smaller cards but it quickly becomes very challenging, as can be seen from the survey in Section 3.2. For example, after it was shown that the degree sequence of G is reconstructible from $\mathcal{D}_{n-2}(G)$ [30], it took more than 35 years for this to be improved to $n-3$ [95]. In Section 5.1.1, we do much more.

Theorem 1.2. *The degree sequence of an n-vertex graph G can reconstructed from $\mathcal{D}_\ell(G)$ for any $\ell \geq \sqrt{2n\log(2n)}$.*

In the same section, we also address the question of determining if a graph is connected from a deck of small cards. Again, the bound we obtain is a marked improvement on the previous best.

Theorem 1.3. *The connectedness of an n-vertex graph G can be recognised from $\mathcal{D}_\ell(G)$ provided $\ell \geq 9n/10$.*

The proofs of the two preceding theorems rely on a tightened version of a pre-existing algebraic tool, which we discuss in Section 5.1. Our work on this variant is joint with Carla Groenland, Tom Johnston and Alex Scott.

The second variant for which we have new results comes from the observation that the procedures by which we reconstruct some basic graph parameters do not require all of the cards (here, we return to 1-vertex-deleted subgraphs). We call this the 'missing cards' variant, where the main question is to determine the threshold on the number of cards that can be missing from the deck before we can no longer reconstruct a specific parameter. It is also possible to ask about reconstructing classes, but progress on this is limited by how difficult it already is for the most basic parameters.

For general graphs, Groenland, Guggiari and Scott [68] proved that the number of edges can be reconstructed from any $n - \frac{1}{20}\sqrt{n}$ cards. This improves on the work of Myrvold [126] and the work of Brown and Fenner [28], who proved the bounds $n-1$ and $n-2$ respectively. In addition, Myrvold [124, 126] showed that the degree sequence is reconstructible from any $n-1$ cards. A more complete discussion of work within and surrounding this variant can be found in Section 3.3; the best bounds here are far from tight.

Our contribution, made in collaboration with Carla Groenland, Tom Johnston, Andrey Kupavskii, Kitty Meeks and Alex Scott, is to show that better bounds can be achieved for graphs with bounded average degree. The following theorem is proved in Section 4.3 using direct but approximate counting techniques, and is the only result of this kind for reconstructing degree sequence that goes beyond one missing card.

Theorem 1.4. *Let G have $n \geq 3$ vertices and average degree bounded by some $d \in \mathbb{N}$. Then the degree sequence of G can be reconstructed from its deck when at most $\frac{n}{10^4 d^3}$ of the cards are missing.*

Under the same bounded average degree assumption, we prove the following two theorems in Section 4.3.1.

Theorem 1.5. *Let G be a graph on $n \geq 3$ vertices with average degree at most d for an unknown $d \in \mathbb{N}$. Then the number of edges in G can be reconstructed from any deck missing at most $\frac{n}{4d+6} - d - 5$ cards.*

The preceding statement is a special case of the next one, although with a slighter better bound.

Theorem 1.6. *Let $d, r \in \mathbb{N}$. For any graph G on $n \geq 3$ vertices with average degree at most d, the number of cliques of size r in G can be reconstructed from any deck missing at most $\left(1 + \binom{2(d+1)}{r-1}\right)^{-1} \left(\frac{n}{2} - 1\right) - d - 5$ cards.*

It is noteworthy that all of these results cover graphs that can be embedded on a fixed surface, such as planar graphs. This class is of special interest since the fact that planar graphs are not yet known to be reconstructible in the classical setting is seen as a serious obstruction to settling the Kelly-Ulam Reconstruction Conjecture.

To round off our adventures in reconstruction, we examine the ‘switching reconstruction’ variant surveyed in Section 3.4. The term *vertex-switching* was first used by Stanley [162] in 1985 to describe a reconstruction variant in which each card of the deck is obtained by a switching operation at a vertex that deletes every edge incident to that vertex in the graph, and inserts all possible edges incident to the vertex that are not in the graph.

Switching reconstruction should seem easier than classical reconstruction. In particular, since we do not delete the vertex once we have switched at it, this means that we can take any card in the switching-deck of a graph G, switch on every vertex again to form the switching-deck of this card, and then our unknown G is guaranteed to be among these n graphs. To some extent this intuition is correct; for example, it is known that all graphs whose order is not divisible by 4 are switching-reconstructible, which is much stronger than anything we can say for the original reconstruction problem. Nonetheless, Stanley conjectured in 1985 that every graph of order at least 5 vertices is switching-reconstructible, and this is also still open.

In addition to Stanley's version, there are numerous operations on graphs and related combinatorial structures that have also been termed switchings, and the corresponding reconstruction problems in these settings have also been studied. A brief list of examples can be found in Section 3.4.2 where, together with Brendan McKay and building on work of Beáta Faller, we also give a generic framework for switching reconstruction that unifies all of these separate definitions. The purpose of this is that it allows us to standardise some existing work and prove results for all these settings simultaneously.

An interesting aspect of switching reconstruction is that it has well-established direct algebraic connections to certain polynomials that we introduce and discuss in Section 5.2. With Brendan McKay, we explored this relationship in the generic setting, and use it in a novel way to derive many identities involving taking iterated decks in Section 5.2.2. Our main contribution for this setting is then contained in Section 5.2.3, where we use some of those identities to prove a new sufficient condition for switching-reconstructibility. These results all apply to Stanley's conjecture, but unfortunately they do not yet break new ground in this case. There is, however, good scope to push these methods much further.

1.2 Contributions in Part II

Graph product structure theory was developed in 2019 by Dujmović, Joret, Micek, Morin, Ueckerdt and Wood [40] in the course of their work on queue number. The main idea is to describe a class of graphs in terms of a highly structured supergraph. Specifically, we want the bigger graph to be the strong product of the simplest possible building blocks, where simplicity is primarily measured by treewidth.

Results of this type have roots in Robertson and Seymour's influential work on graph minors, which is documented in a series of 23 papers published between 1983 and 2010. One of the many important results that came out of this is that minor-closed classes of graphs can be built up from graphs embedded on surfaces using some fairly simple operations on embedded graphs. In turn, graphs embeddable on surfaces can be described using graph product structure in terms of, for example, a graph with small treewidth, a path (which also has small treewidth), and a small complete graph.

In this way, the theory is capable of providing finer characterisations. More of this story is told in Chapter 7, which culminates in the statements of the main product structure theorems.

Being relatively recent, graph product structure theory is in a stage where new theory is still being rapidly developed. Not all of this is done with a specific application in mind, so in tandem, applications of the existing results are also being sought out in wide contexts. We make new contributions in both of these directions.

Chapter 8 presents a new application of product structure that lies in the intersection of geometric and structural graph theory. Famously, Koebe [93] proved that a graph is planar if and only if it is a touching graph of balls in $\mathbb{R}^2$. This result has motivated numerous strengthenings and variations; for instance, Felsner and Francis [55] showed that every planar graph is a touching graph of cubes in $\mathbb{R}^3$. On one hand, such characterisations demonstrate that graphs which admit certain geometric representations may form 'nice' classes, in the sense that they satisfy desirable structural properties as is the case for planar graphs. On the other, if we can find a representation so that the class of graphs which are representable coincide with one that we wish to study, this opens up the possibility of using geometric tools to work with graphs in the target class.

We can gain different structural insight when particular geometric object are chosen – this may be seen for instance by comparing packing balls versus tiling boxes, which immediately restricts what a neighbourhood of a vertex can look like. A *touching representation by comparable boxes* of a graph G is a collection of interior-disjoint boxes in $\mathbb{R}^2$ such that:

- there is one axis-aligned box for each vertex $v \in V(G)$,
- two boxes touch (intersect) if and only if their corresponding vertices are adjacent in G, and
- for every $u, v \in V(G)$, the boxes corresponding to u and v are comparable in the sense that one fits inside a translate of the other.

The *comparable box dimension* $\dim_{cb}(G)$ of a graph G is the smallest integer d such that G has a touching representation by comparable boxes in $\mathbb{R}^d$. This representation is specially geared toward studying sparse graphs, as they preclude complete and complete bipartite graphs from being representable in few dimensions. We use product structure theorems to prove the following main theorem.

Theorem 1.7. *Every proper minor-closed class of graphs has finite comparable box dimension.*

The same proof outline gives an explicit bound for graph on surfaces.

Theorem 1.8. *For every graph G of genus g,* $\dim_{cb}(G) \leq 3 \cdot 81^7 \cdot \max(2g, 3)^{\log_2 81}$.

To facilitate the use of product structure theorems, we also prove a host of new results that relate comparable box dimension to numerous graph parameters such as chromatic number and treewidth, and operations such as strong products, taking subgraphs and clique sums. This is based on joint work with Zdeněk Dvořák, Daniel Gonçalves, Abhiruk Lahiri and Torsten Ueckerdt.

On the theoretical side, Chapter 9 goes one layer deeper on the scale of structural characterisations by showing that many graphs with bounded treewidth can be described as subgraphs of the strong product of a graph with even smaller treewidth and a bounded-size complete graph. To measure how effectively this can be done for graphs of different families, we introduce a new parameter. Let us define the *underlying treewidth* of a graph class $\mathcal{G}$ to be the minimum non-negative integer c such that, for some function f, for every graph $G \in \mathcal{G}$ there is a graph H with $\mathrm{tw}(H) \leq c$ such that G is isomorphic to a subgraph of $H \boxtimes K_{f(\mathrm{tw}(G))}$. The smaller the underlying treewidth, the simpler the main building blocks for product structure can be. It turns out that this is related to several existing parameters such as tree-partition-width and clustered colouring number.

Our main contribution is to introduce the notion of disjointed coverings of graphs and show they determine the underlying treewidth of any graph class. Disjointed coverings are easier to work with than the original definition and provide a practical way to compute the underlying treewidth of many typical classes of interest. For example, we use it to show that the class of graphs embeddable on any fixed surface has underlying treewidth 3, among other applications. This is based on a project in collaboration with Rutger Campbell, Katie Clinch, Marc Distel, J. Pascal Gollin, Kevin Hendrey, Robert Hickingbotham, Tony Huynh, Freddie Illingworth, Youri Tamitegama and David Wood.

1.3 Conventions

We establish some notational conventions that apply throughout this thesis. For a finite set S, let $|S|$ be its cardinality and let $\mathcal{P}(S)$ be its power set. Let $\mathbb{N} \coloneqq \{1, 2, \dots\}$, $\mathbb{N}_0 \coloneqq \{0, 1, \dots\}$, and $[n] = \{1, \dots, n\}$.

The objects under consideration in this thesis are primarily simple, finite and undirected graphs, although Part I briefly touches on other related structures. We use $V(G)$ and $E(G)$ to denote the vertex set and edge set of a graph G respectively. We reserve $n = |V(G)| = |G|$ for the number of vertices (or *order*) of G, and $m = |E(G)|$ for the number of edges. Both uv and $\{u, v\}$ are used to denote elements of $E(G)$, where $u, v \in V(G)$ and $u \neq v$. For a vertex $v \in V(G)$, we denote its neighbourhood in G by $\Gamma_G(v)$ (or $\Gamma(v)$ if the ambient graph is clear). Let $\Delta(G)$ be the maximum degree and $\delta(G)$ be the minimum degree of G.

When we say that a graph is *unique* or *uniquely determined*, we always mean up to graph isomorphism, which is denoted by '$\cong$'. A *class of graphs* is a collection of graphs that is closed under isomorphism, that is, if $\mathcal{C}$ is a class of graphs containing G and $H \cong G$, then $H \in \mathcal{C}$. If H' is a subgraph of G (so that $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$), and $H' \cong H$, then we allow ourselves to say that H is a subgraph of G as well. There are times where we will want to keep track of actual occurrences of a particular subgraph within a graph for counting purposes, but separate terminology will be introduced for this later.

Lastly, we mention some standard constructions and parameters. We will write $G - v$ for the graph formed by deleting the vertex v and its incident edges, and $G - e$ for the graph formed by deleting the edge e (for set difference we use either $\setminus$ or $-$ depending on context). Let P_n be the n-vertex path, which has length $n - 1$. The complete graph on r vertices (or *r-clique*) is denoted by K_r. Let $\omega(G)$ be the size of the largest clique in a graph G, and $\chi(G)$ be the chromatic number of G. The *(Euler) genus* of a surface with h handles and c cross-caps is $2h + c$. The *(Euler) genus* of a graph G is the minimum $g \in \mathbb{N}_0$ such that there is an embedding of G in a surface of genus g (see [123]).

For graph theoretic notation and definitions that do not appear here, we follow Diestel [34]. A list of recurring symbols is included on page iii.

Part I

Reconstruction

Chapter 2

First look

We begin by describing the classical setting for graph reconstruction. The *deck* of a graph G, denoted $\mathcal{D}(G)$, is the multiset consisting of all (1-)vertex-deleted subgraphs of G, i.e. $\mathcal{D}(G) = \{G - v : v \in V(G)\}$. In keeping with this terminology, elements of the deck are referred to as *cards*. We will say that two decks are *equal* or two graphs have the *same deck* (and write $\mathcal{D}(G) = \mathcal{D}(G')$ accordingly) if there is a 1-1 correspondence between those multisets pairing up isomorphic cards[1]. Given a fixed deck $\mathcal{D}$, any graph G for which $\mathcal{D}(G) = \mathcal{D}$ is called a *reconstruction* of $\mathcal{D}$. Now comes the key property that we wish to study: we say that a graph is *reconstructible* if it is uniquely determined (up to isomorphism) by its deck, that is, if every reconstruction of $\mathcal{D}(G)$ is isomorphic to G. This leads to the natural question: which graphs are reconstructible?

An example of a deck is shown in Figure 2.1. While it is convenient to use the notation $G - v$ for cards, it is important to note that the cards do not come labelled (else reconstruction would be easy) by which we mean that we are only given their isomorphism type and multiplicity. Thus, for the graph depicted, all we would know given the deck is that there are four cards consisting of two copies of C_3 and two copies of P_3, so $\mathcal{D}(C_4 + e) = \{C_3, C_3, P_3, P_3\}$.

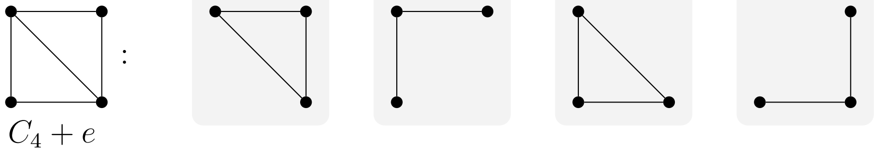


Figure 2.1: $\mathcal{D}(C_4 + e)$

---

[1] In older literature, the term *hypomorphic* is often used to describe two graphs with the same deck. However, this term seems to have fallen out of common usage lately.

In the above example, it is easy to recover the complete structure of the original graph given only the deck even in a naive fashion: the fact that there are four cards means that our unknown graph G has four vertices, and one can verify directly by enumerating all 11 such graphs that they have different decks exactly one of which is $\{C_3, C_3, P_3, P_3\}$. Of course, this exhaustive strategy is limited. The current best computer search due to McKay [121] (see also [120]) has verified that all graphs on between 3 and 13 vertices[2] are reconstructible from their decks.

On the other hand, we have already mentioned that both graphs on 2 vertices are not reconstructible since their decks trivially consist of two cards, each with one vertex and hence no edges. In fact, these are the only known graphs that are not reconstructible. The long-standing Reconstruction Conjecture posed by Kelly and Ulam, formulated in 1941[3] and appearing in the PhD thesis of the former, asserts that they are the only ones.

Conjecture 2.1 (Kelly-Ulam Reconstruction Conjecture [85, 173])**.** *Every graph with order at least 3 is reconstructible.*

Despite its ostensible simplicity, the Reconstruction Conjecture remains wide open and looks very unlikely to be solved in the near future. Indeed, direct progress on the conjecture has been more or less at a standstill in recent times. Consequently, there is great interest in any partial results as well as parallel progress on the many variants of the classical reconstruction problem, which arise by asking whether we can reconstruct graphs and sometimes other combinatorial objects after tweaking or replacing the definition of a deck. All of this constitutes an extensive and ever-growing body of related work.

Partial reconstruction results, both for the classical problem and variants, typically fall into one of two broad categories. The first are those that involve reconstructing parameters. We say that a graph parameter is *reconstructible* if for any graph G with, say, value p for the parameter, every reconstruction of G also has value p for that parameter. That is, the parameter is uniquely determined by the deck. The following observation provides some easy examples of reconstructible parameters.

Observation 2.2. *The number of vertices, number of edges and degree sequence of a graph are all reconstructible.*

---

[2] There are 50,502,031,367,952 isomorphism classes of graphs on 13 vertices, so even a naive exhaustive approach checking all pairs of graphs won't do here. The technique employed in [121] shows that many cases can be quickly eliminated without a full check.

[3] This is the widely cited year, although it is believed that the problem may have been known to Ulam already in 1929 (see [77]). These earlier occurrences were stated in terms of metric set systems, but the graph language quickly became much more prevalent.

The number of vertices is clear from the definition of the deck, this being precisely the number of cards. To reconstruct the number of edges m, we take the total number of edges present in all cards and divide by $n-2$ where n is the number of vertices. Then, the degree sequence can be obtained by looking at each card $G-v$ and noting that the degree of v is given by subtracting the number of edges in $G-v$ from m.

The second direction is to show that particular classes of graphs are reconstructible, meaning that every graph in the class is reconstructible. A class $\mathcal{C}$ of graphs is *recognisable* if, for any $G \in \mathcal{C}$, every reconstruction of G is also in $\mathcal{C}$. Recognition results go hand-in-hand with reconstructing properties, since recognising a class of graphs amounts to ensuring that we can reconstruct the defining property of a given class. Independent of recognisability, one can still consider whether a graph class is *weakly reconstructible*, meaning that no two graphs in the class have the same deck. This is equivalent to being given the initial data of the deck together with the additional assumption that the deck has a reconstruction within the particular class. There are many families of graphs for which this is the sticking point, even when recognition is known.

Proving that a family of graphs is reconstructible is therefore a two-step process: given the deck, we first to need to be able to tell that the graph it came from is in our family (recognition), and then we actually need to go through the process of uniquely reconstructing the graph (weak reconstruction). We give two elementary examples.

Theorem 2.3. *Regular graphs and Eulerian graphs are reconstructible.*

Proof. Suppose that G is our unknown graph, and we are given its deck $\mathcal{D}(G)$. By Observation 2.2, we know that every reconstruction of $\mathcal{D}(G)$ is regular, say with degree d. Now to reconstruct G, take any card H and add one new vertex joined to each vertex in H of degree $d-1$ (there are necessarily exactly d such vertices). Due to the definition of a deck, all such d-regular reconstructions are isomorphic so G is uniquely determined.

Similarly, note that G is connected if and only if at least two of its cards are connected. We can therefore decide whether G is Eulerian from its connectedness and reconstructed degree sequence. Then, to reconstruct G, we take any card and join one additional vertex to every vertex in the card that has odd degree. ☐

In this part of the thesis, we paint a broad picture of the area of graph reconstruction with the goal of motivating and contextualising new contributions, while also setting up for continued work in the future. This will not be a comprehensive survey; indeed, after so many years of sustained work, it is no longer really practical to write a truly comprehensive survey on this topic. There is, however, a series of staple references. We suggest that the most useful introduction comes from Bondy's *Graph Reconstructor's Manual* from 1991 [17] (marking 50 years of the Reconstruction Conjecture), which is organised around recurring techniques, together with Lauri's chapter in the 2nd edition of the Handbook of Graph Theory from 2014 [72, Chapter 2.3], which is the most recent entry.

Organisation. Within this chapter, Section 2.1 presents a simple outline of developments in classical graph reconstruction, based in part on an amalgamation of the aforementioned surveys. This gives us the opportunity to get a flavour for lines of research, working our way up to discussing the most recent key results. Unfortunately, we do not make direct progress on the Reconstruction Conjecture in this thesis, although the work that we do is geared toward improving our understanding of what might be within reach and developing tools in parallel settings that could have transferrable ideas. With a view to this, the remainder of this part shifts away from the classical vertex reconstruction problem to different variants where the majority of current activity lies. Each section in Chapter 3 introduces one such variant. We then present some new contributions which use two families of tools that have deep associations with reconstruction. The first, appearing in Chapter 4 are those primarily derived from direct counting methods. The second, in Chapter 5, use algebraic methods, especially those involving studying polynomials. We conclude this part with an outline of potential next steps for the area in Chapter 6.

2.1 Key results in vertex reconstruction

This section summarises some standard and significant partial results toward the Reconstruction Conjecture. Where a specific citation is not provided, we refer to [110]. To highlight how certain results have grown out of each other, or out of common lines of enquiry, we will organise them around themes.

Basics. We begin with some more basic facts that come from directly examining cards, starting from a common foundation on which many results in reconstruction are built. The following simple yet central counting lemma arises naturally when we attempt to generalise the formula for reconstructing edge counts in Observation 2.2. Suppose that G and H are graphs. Let $\widetilde{n}_H(G)$ and $n_H(G)$ denote the number of subgraphs and induced subgraphs of G isomorphic to H respectively. In other words, these parameters count the number of copies or instances of H in G, which might be zero. Importantly, these are reconstructible parameters.

Lemma 2.4 (Kelly's Lemma [86]). *For any graphs G and H with $|H| < |G|$, both $\widetilde{n}_H(G)$ and $n_H(G)$ are reconstructible from $\mathcal{D}(G)$.*

The proof easily follows from the formula $\widetilde{n}_H(G) = \frac{1}{n-|H|} \sum_{C \in \mathcal{D}(G)} \widetilde{n}_H(C)$ (and similarly for $n_H(G)$) by noting that all of the summands can be read directly from the given deck. Several important tools that fundamentally rely on Kelly's lemma have since been developed; these are the topic of Section 4.1, and require a bit more work. A simple application generalising the procedure to reconstruct degree sequence gets us a little more mileage. If v is any vertex of the graph, the number of edges $|E(\Gamma(v))|$ within the neighbourhood is the number of triangles containing v in G. This is given by $n_{K_3}(G) - n_{K_3}(G-v)$, so using Kelly's Lemma (to determine $n_{K_3}(G)$) we can reconstruct $|E(\Gamma(v))|$ for all v. As a small note, we can also reconstruct the degrees of all vertices in $\Gamma(v)$ for each v by comparing the degree sequence of a card $G - v$ with the reconstructed degree sequence of G; the deletion of v reduces the degrees of its neighbours by exactly 1 and does not affect anything else.

Connectivity. A graph G is connected if and only at least two cards in $\mathcal{D}(G)$ are connected. This gives an easy way to recognise disconnected graphs with at least 3 vertices, which are in fact a reconstructible class. Kelly [85] gave a direct proof of this by reconstructing the collection of components, but it is also the easy consequence of a tool that we discuss in Section 4.1.2. We can also reconstruct the (vertex-)connectivity of a graph using the observation that, for $k \geq 2$, a graph is k-connected if and only if every card is $(k-1)$-connected. This means that we can, for instance, recognise the class of separable graphs. However, reconstruction of graphs with higher connectivity is not known in general. For a separable graph G, Bondy [16] showed that the blocks of G are reconstructible. Although this does not determine G uniquely, Bondy used it to prove that the class of non-separable graphs without pendant vertices[4] is reconstructible.

Another easy observation is that if we take the multiset of complements of every card in a deck $\mathcal{D}(G)$, we obtain precisely the deck of G^c, the complement of G.

Observation 2.5. $\mathcal{D}(G^c) = \{D^c : D \in \mathcal{D}(G)\}$.

---

[4] A *pendant vertex* is a vertex of degree 1.

It follows that G is reconstructible if and only if its complement G^c is reconstructible. This is pertinent to disconnected graphs in particular, since the complementary class contains no disconnected graphs. Using this, Yang [181] showed that the Reconstruction Conjecture is true if it is true for 2-connected graphs.

From trees to planarity. Along with regular graphs, Eulerian graphs and disconnected graphs, one of the earliest graphs classes known to be reconstructible is the class of trees. This was proved by Kelly [86]. While recognition is straightforward using either the edge count or that they are minimally connected, the procedure for reconstruction involves some fairly substantial casework. This proof has since been revisited many times, resulting in simpler arguments (see, for instance, [77, 19, 17]). Following on from this, Manvel [115] showed that the class of unicyclic graphs is also reconstructible. As both trees and cycles are special cases of cacti[5], Geller and Manvel [60] took the natural step of reconstructing this broader class.

The class of cacti is in turn a subclass of the class of outerplanar graphs, making this the next target. Manvel [117] made progress toward this in 1972, showing that maximal outerplanar graphs are reconstructible. Giles [62] was then able to extend this to general outerplanar graphs in 1974. Going the other way, Fiorini and Manvel [58] showed that the class of maximal planar graphs with minimum degree at least 4 is reconstructible. The degree condition was not straightforward to remove here, and even recognition becomes more difficult without it. Nonetheless, Fiorini and Lauri [56] subsequently showed that maximal planar graphs are recognisable, after which Lauri [57] was able to confirm reconstructibility in 1981. The sharply increasing difficulty in this sequence correctly signifies the challenge that general planar graphs would pose. No concrete progress was made until 2007.

Theorem 2.6 (Bilinski, Kwan and Yu [9])**.** *Planar graphs are recognisable.*

This theorem is still the most recent breakthrough toward the Reconstruction Conjecture. Bilinski, Kwan and Yu were also able to show that certain 5-connected planar graphs are reconstructible as a byproduct of the methods used to prove Theorem 2.6, but the general reconstructibility of planar graphs is a glaring omission from the list of graph classes known to be reconstructible.

---

[5] A *cactus* is a graph in which every pair of cycle has at most one vertex in common.

Polynomials and perfect graphs. In 1979, Tutte [172] published a groundbreaking paper providing a general framework that facilitates the reconstruction of a host of parameters using multiplicative recursion. These include the characteristic polynomial (and hence eigenvalues) and chromatic polynomial (and hence chromatic number), as well as the number of spanning subgraphs that are disconnected (such as perfect matchings) or separable (such as Hamilton cycles). The methods were later picked up by Kocay [88] in 1981, who was able to express them in more elementary terms. We discuss this in greater detail in Section 4.1.1.

An easy application of Kelly's Lemma allows us to reconstruct the clique number $\omega(G)$ of a graph G. Putting this together with the fact that we can reconstruct chromatic number immediately implies that the class of perfect[6] graphs is recognisable. Von Rimscha [176] made this observation and went on to prove that many subclasses of perfect graphs are recognisable; namely, triangulated graphs, interval graphs, comparability graphs and split graphs. Further restricting these classes, he also showed that certain split graphs (threshold graphs, those for which $\omega(G)+\alpha(G) = |V(G)|+1$, and those that are also comparability graphs) as well as the class of unit interval graphs are reconstructible. Not much further work has appeared in this direction. We remark that these results came much earlier than the proof of the Strong Perfect Graph Theorem (see [72, Chapter 2.5.1]), but the advent of that theorem has not so far translated into progress on reconstructing perfect graphs.

---

[6] A graph G is *perfect* if, for every induced subgraph $H \subseteq G$, we have $\chi(G) = \omega(G)$.

Chapter 3

Variants

We have so far only considered the original reconstruction problem (which we will also refer to as *vertex reconstruction*, or *classical reconstruction*). Given the difficulty in making any substantial progress on it though, work in the area quickly branched out into many reconstruction-type problems with a similar spirit. Some of these lie parallel, whilst others are deeply intertwined with the classical setting so that results and techniques in one can inform lines of enquiry in the other. Importantly, these are often also independently interesting problems.[1]

It does not require too much imagination to see how variants can be formulated. The essential components of a reconstruction problem are a structure to be reconstructed, and the data provided to do so. Breaking this down reveals several moving parts; the (combinatorial) objects used, a notion of equivalence, the way in which a single card is defined, and the way that the whole deck is defined. If we change any of these from graphs, isomorphism, single vertex-deletion, and one card per vertex, we produce a variant in which we can ask the same question; does the given data determine this combinatorial structure up to equivalence?

One way to view reconstruction variants is to organise them into three types. The first consists of those that are meant to be more tractable than the classical version. These are typically obtained by modifying the definition of a card so that each one carries more information. Studying questions of this type, which are (at least ostensibly) easier, provide an opportunity to develop and test new strategies with the possibility that some of the ideas might transfer back to the original question. In reality, these variants have still proved to be very difficult. We will discuss two problems of this type in Sections 3.1 and 3.4.

---

[1] Indeed, apart from the classical Reconstruction Conjecture, two (edge and switching reconstruction) variants also appear in Bondy's list of 'Beautiful conjectures in graph theory' [18].

The next category on our list consists of extensions of vertex-reconstruction, in the sense that they contain classical reconstruction as a special case. These are usually obtained by modifying the definition of a card to reduce the amount of information seen, and/or limiting the deck in some way. We know that the Reconstruction Conjecture is difficult and these problems are strictly even more so, but the idea is that they can help measure the difficulty of vertex-reconstruction by closing in on the minimum information needed to reconstruct certain graphs or properties. We will look at several ways in which these questions arise in Sections 3.2 and 3.3.

The third type of variant are those that arise from transferring the core structure and nature of the classical problem to different settings, such as by changing the combinatorial objects involved from graphs to something richer, or changing the notion of equivalence. We discuss a few selected examples in Section 3.5.

3.1 Edge reconstruction

The *edge deck* of a graph G is the multiset $\{G - e : e \in E(G)\}$ of edge-deleted subgraphs of G. We say that a graph is *edge-reconstructible* if it is uniquely determined by its edge deck. This is a natural counterpart to vertex reconstruction, with part of its appeal coming from the fact that it feels easier since many cards contain an even greater proportion of the information of the whole graph as compared to vertex-deletion. In 1964, Harary posed the following conjecture.

Conjecture 3.1 (Harary [75]). *Every graph with at least 4 edges is edge-reconstructible.*

All known counterexamples on 3 edges (and as many vertices as we like) are given by considering $K_3 \sqcup cK_1$ and $K_{1,3} \sqcup (c-1)K_1$, where cK_1 denotes c isolated vertices. While the Edge Reconstruction Conjecture is also wide open, there is a qualitative sense in which this variant is easier than the vertex setting. The first such comparison comes from a result of Hemminger, which says that a graph (apart from K_3) is edge-reconstructible if and only if its line graph is reconstructible, so this is a special case of vertex reconstruction. This becomes clearer with the following theorem.

Theorem 3.2 (Greenwell [65]). *Let G be a graph without isolated vertices. Then $\mathcal{D}(G)$ is uniquely determined by the edge deck of G. Therefore, if G is reconstructible, then it is also edge-reconstructible.*

The preceding theorem means whenever we prove that a class or parameter is vertex reconstructible, then we also obtain that it is edge reconstructible. Many tools, such as Kelly's Lemma or other direct counting arguments based on directly inspecting the deck, also automatically apply. There are still statements about vertex-reconstruction that do not transfer, however. For instance, it is not true that the complement of any edge-reconstructible graph is edge-reconstructible since Observation 2.5 fails. On the other hand, there is much more that we can say for edge-reconstruction that is not known in the vertex case, and the real interest in this variant comes from results obtained independent of the vertex-deck. These provide a guide for what might be possible in classical reconstruction, and a sandbox for the development of new techniques.

Within the edge-reconstruction literature, there is a particularly important collection of results and tools lauded by Lauri as "[arguably] the deepest and most general results obtained in reconstruction". A short note of Lovász [111] from 1972 proved that if $|E(G)| > \binom{n}{2}/2$, then G is edge-reconstructible. The proof is based on a surprisingly simple inclusion-exclusion argument, and was later tightened by Müller to show that $2^{|E(G)|-1} > n!$ is sufficient. Taking an algebraic approach, Nash-Williams gave a strikingly general 'mapping counting' lemma which can be used to prove both of these results but also has much further applications. For graphs G and H with the same vertex set V, the basic idea is to count the number of homomorphisms[2] from $G \to H$, denoted $[H]_G$. Given a subset $X \subseteq E(G)$, let $[H]_{G\setminus X}$ denote the number of homomorphisms forbidding X, meaning that all edges in $E(G) \setminus X$ map to edges in H and all edges in X map to non-edges in H.

Lemma 3.3 (Nash-Williams [158])**.** *Let G and H be graphs on the same vertex set and with the same edge-decks. Then $[H]_G = |\mathrm{Aut}(G)| + (-1)^{|X|}([H]_{G\setminus X} - [G]_{G\setminus X})$ for any $X \subseteq E(G)$.*

Apart from the two results mentioned above, an especially noteworthy application of this lemma due to Pyber [137] showed that any Hamiltonian graph with sufficiently many vertices is edge-reconstructible. What was particularly novel about Nash-Williams' methods, however, was that the formulation allows for the reconstruction of other objects with the following setup: take the structure to be a triple (D, Γ, E) consisting of:

D: a finite set,
Γ: a group of permutations acting on D,
E: a subset of D.

---

[2] A *homomorphism from G to H is a permutation of their common vertex set such that every edge in G is mapped to an edge in H.*

The edge-reconstruction problem that arises from this data asks whether the subsets $E - x$ uniquely determine E up to translation by an element of Γ. To obtain the graph variant, we can set D to be the set of all $\binom{n}{2}$ edges on a vertex set $[n]$, Γ to be the symmetric group on $[n]$, and E to be the edges that define the graph we are looking to reconstruct. For a toy example outside of graphs, one can consider setting D to be the set of cells in a 3×3 square grid, Γ to be the dihedral group acting on the square, and we ask which sets E can be reconstructed. This setting is used as a running example in [110, Chapter 11], to which refer for a more complete account.

While edge reconstruction is not going to be a focus of this thesis, we take away the precedent set here for algebraic formulations of counting arguments, and of generic formulations of problems that abstract away from graphs. These will both be present in the approach we take concerning the switching variant introduced in Section 3.4.

3.2 Reconstruction from small cards

The *ℓ-deck* of G, denoted $\mathcal{D}_\ell(G)$, is the multiset of all induced subgraphs of G on ℓ vertices. This is also the collection of k-vertex-deleted subgraphs of G, where $k = n - \ell$. It is common to use both k and ℓ, and we will do this interchangeably. In this notation, $\mathcal{D}(G) = \mathcal{D}_{n-1}(G)$, but ℓ can be potentially much smaller than $n - 1$. We say that a graph, class or parameter is *reconstructible from the ℓ-deck* if it is determined by the ℓ-deck. Intuitively, individual cards that are smaller carry less information. Indeed, the $(\ell - 1)$-deck is determined by the ℓ-deck for each ℓ, which can be shown by a simple counting argument (see Section 4.1). Thus, if a graph is reconstructible from its ℓ'-deck then it is reconstructible from its ℓ-deck for all $\ell \geq \ell'$. The main question is then to determine the threshold; that is, the smallest ℓ for which a given class of graphs or property is reconstructible from the ℓ-deck.

Reconstruction from small cards is generally attributed to Kelly, although the strengthening of the Reconstruction Conjecture that follows seems to be formulated by Manvel (who calls it “Kelly’s Conjecture”).

Conjecture 3.4 (The k-Reconstruction Conjecture [86, 118])**.** *For every $k \in \mathbb{N}$, there is an integer N_k such that every graph with at least N_k vertices is reconstructible from its $(n - k)$-deck.*

Kelly and Ulam’s Graph Reconstruction Conjecture posits that $N_1 = 3$. This stronger conjecture did not receive much attention until 1974 when it was studied by Manvel [118], who showed that several classes of graphs, such as connected graphs, trees, regular graphs and bipartite graphs, can be recognised from the $(n-2)$-deck where $n \geq 6$ is the number of vertices. Since then, a lot more progress has been made.

We begin by looking at two parameters. In 1982, Chernyak [30] showed that the degree sequence of an n-vertex graph can be reconstructed from its $(n-2)$-deck for $n \geq 6$. This was later extended by Kostochka, Nahvi, West, and Zirlin [95] to the $(n-3)$-deck for $n \geq 7$. The best known asymptotic result is due to Taylor [169], who proved that the degree sequence of a graph G on n vertices can be reconstructed from $\mathcal{D}_\ell(G)$ where $\ell \sim (1-1/e)n$. In Section 5.1 of this thesis, we strengthen aspects of Taylor's proof to show that $\ell \geq \sqrt{2n\log(2n)}$ suffices (Theorem 1.2).

The same method can be adapted to recognise whether a graph is connected. The story in the literature here is similar to that for degree sequence. We have already mentioned Manvel's result that the class of connected graphs is recognisable from the $(n-2)$-deck for $n \geq 6$. Extending this, Kostochka, Nahvi, West, and Zirlin [95] showed that the connectedness of a graph on $n \geq 7$ vertices is determined by $\mathcal{D}_{n-3}(G)$. Spinoza and West [157] proved that connectedness can be recognised from $\mathcal{D}_\ell(G)$ provided $n - \ell \leq (1+o(1))\sqrt{\frac{2\log n}{\log(\log n)}}$. Our method from Section 5.1 gives a significant improvement to allow a linear gap between n and ℓ. Specifically, we show that the connectedness of an n-vertex graph G can be recognised from $\mathcal{D}_\ell(G)$ provided $\ell \geq 9n/10$ (Theorem 1.3).

Spinoza and West [157] also gave some telling lower bounds concerning connectedness. If we take $G_1 = P_n$ (the path on n vertices) and $G_2 = C_{\lceil n/2\rceil+1} \sqcup P_{\lfloor n/2\rfloor-1}$ the disjoint union of a cycle and a path, we find $\mathcal{D}_\ell(G_1) = \mathcal{D}_\ell(G_2)$ for all $\ell \leq \lfloor n/2\rfloor$. However, G_1 is connected and G_2 is not. In light of this construction, they believe that for $n \geq 6$ and $\ell \geq \lfloor n/2\rfloor + 1$, the connectedness of an n-vertex graph G is determined by $\mathcal{D}_\ell(G)$, which would be sharp.

We now move to reconstructing graph classes from decks of smaller cards. For general graphs, it is not possible to guarantee reconstructibility from the $(n-k)$-deck unless $k = o(n)$, as shown by the following theorem of Nýdl.

Theorem 3.5 (Nýdl [131])**.** *For any integer n_0 and $0 < \alpha < 1$, there exists an integer $n > n_0$ such that there are two non-isomorphic graphs on n vertices which share the same multiset of subgraphs of order at most αn.*

However, Nýdl's theorem does not say what happens if we limit to specific families of graphs. This has been demonstrated in various recent results. It is known that n-vertex graphs with maximum degree 2 are reconstructible from their $\lceil n/2\rceil$-decks [97], 3-regular graphs are reconstructible from their $(n-2)$-decks [94], disconnected

Figure 3.1: Non-isomorphic trees on 13 vertices that have the same 7-deck.

graphs for which every component has at most ℓ vertices is reconstructible from their ℓ-decks [97], and every complete r-partite graph is reconstructible from its $(r+1)$-deck [97]. Most statements involve either very restrictive classes or k-decks for k only just under $n-2$. In general, there is a delicate balance here that reflects the difficulty of obtaining these kinds of results.

An interesting exception is the case of trees, which have been studied for much longer. Giles [64] proved in 1976 that no two non-isomorphic n-vertex trees (with $n \geq 5$) have the same $(n-2)$-deck. In 1990, Nýdl then made the bold conjecture that no two non-isomorphic trees have the same ℓ-deck when ℓ is slightly larger than $n/2$.

Conjecture 3.6 (Nýdl [130]). *For any $n \geq 4$ and $\ell \geq \lfloor n/2 \rfloor + 1$, any two trees on n vertices with the same ℓ-deck are isomorphic.*

The conjectured bound would be sharp: Nýdl [130] presented trees for which $\ell \geq \lfloor n/2 \rfloor + 1$ is necessary (it is non-trivial to verify that two graphs have the same ℓ-decks generally, but a short proof has been found in [97] in this case). There was no progress on Conjecture 3.6 until recently. We first note that it is false as stated; Figure 3.1 shows two trees on 13 vertices with the same 7-deck. This was found by a computer search of our coauthor, Tom Johnston, which also verified the conjecture for $4 \leq n \leq 25$. On the other hand, we will show that for $\ell \geq (2n+4)/3$, the class of trees on n vertices is recognisable from the ℓ-deck (Theorem 4.8). Then, taking a big step beyond Giles' theorem and matching Nýdl's linear conjecture, we prove that any n-vertex tree T can be reconstructed from $\mathcal{D}_{n-k}(T)$ when $k < \frac{n}{9} - \frac{4}{9}\sqrt{8n+5} - 1$ (Theorem 1.1). The proofs of these theorems are detailed in Section 4.2, and rely on a new counting result that we introduce in Section 4.1.3.

At the same time as our result appeared, Kostochka, Nahvi, West and Zirlin [96] independently proved that one can recognise if a graph is acyclic from the ℓ-deck when $\ell \geq \lfloor n/2 \rfloor + 1$, which also verifies the believed bound for reconstructing connectedness in the special case of forests. This has the particularly nice consequence that trees can be recognised from their ℓ-deck, and so Conjecture 3.6 (which is only for weak reconstruction) is equivalent to the reconstruction of trees amongst general graphs. Finally, while Nýdl's conjectured lower bound for reconstructing trees is false for

$n = 13$, it is still the best known for all other values of n. It may still be case that the conjecture is asymptotically true, or even true exactly for large enough n.

Problem 1. *Is there a function* $\ell(n) = (1/2 + o(1))n$ *such that all* n*-vertex trees can be reconstructed from their* $\ell(n)$*-deck?*

3.3 Reconstruction from subdecks

On examining the proofs of basic reconstructibility results in the classical setting (returning again to 1-vertex-deleted subgraphs), we quickly see that not all cards are used equally and some are not really needed at all. There is a family of variants of reconstruction highlighting this that arise from restricting the initial information we are given to a subset of the deck.

This observation was investigated early on in relation to trees. Specifically, suppose we consider what subsets of cards are enough to distinguish trees, assuming that we know the graph is a tree (so we are just left to verify weak reconstructibility) and we also know something about how the subset of cards was picked. Harary and Palmer [78] showed how to recover a tree using only the cards which are subtrees (i.e. connected, so these are the cards corresponding to leaf deletions), Bondy [15] showed that only the cards where peripheral vertices have been removed are needed and Manvel [116] showed that the set (as opposed to the multiset) of cards which are trees suffices. It has also been shown that trees with at least 3 cutvertices can be reconstructed (amongst all graphs) from the cards corresponding to removing a cutvertex [109], and that only three carefully chosen cards are needed to reconstruct a tree when $n \geq 5$ [125]. The reconstruction of other classes of graphs from structurally significant cards has also been studied (such as in [61]).

Manvel's result also falls under a variant introduced by Harary in 1964, who considered reconstruction from the set (rather than multiset) of 1-vertex-deleted subgraphs. The loss of multiplicity data does make a difference; the decks of the graphs P_3 and $P_2 \sqcup v$ (disjoint union of an edge with an isolated vertex) are different as multisets, but both consist only of P_2 and the empty graph on two vertices so they are the same as sets. On the other hand, we don't have any bigger examples where this is the case, leading to the Set Reconstruction Conjecture. Let us say a graph is *set reconstructible* if it is uniquely determined by its set of 1-vertex-deleted subgraphs.

Conjecture 3.7 (Harary [75]). *Any graph of order at least 4 can be reconstructed uniquely from its set of non-isomorphic vertex-deleted subgraphs.*

Much of what we know about Set Reconstruction is due to Manvel [119], who showed that the number of vertices, number of edges, minimum degree, degree sequence for certain classes, and connectivity are set reconstructible. Manvel used these to prove that, apart from trees, the disconnected graphs, separable graphs without pendants, and maximal outerplanar graphs are also set reconstructible. These results closely follow the early trajectory in classical reconstruction. Continuing with this as a guide, the Set Reconstruction Conjecture has also been verified for 2-connected outerplanar graphs [63] and unicyclic graphs [6]. Additionally, McKay's computer verification that graphs of up 13 vertices are reconstructible only use the set of cards.

As a measure of how difficult it is to reconstruct certain graphs, Harary and Plantholt [80] introduced the definition of the *reconstruction number* of a graph G as the smallest number of cards in $\mathcal{D}(G)$ which guarantee that we can uniquely determine G. They formulated another strengthened conjecture in terms of this parameter.

Conjecture 3.8 (Harary and Plantholt [80])**.** *Any graph of order at least 3 has reconstruction number at most* $\frac{n}{2}+2$.

The same quantity was also defined by Myrvold [124] under the term *ally reconstruction number*, with the idea that we have an ally who has seen the unknown graph we are trying to reconstruct and they give us a well-chosen selection of cards. The aforementioned result of Myrvold implies that trees on at least 5 vertices have reconstruction number 3, and another result of Myrvold [125] states that disconnected graphs for which not all components are isomorphic have reconstruction number 3. This is complemented by a more recent result that any disconnected graph with reconstruction number at least $c+1$ must consist of copies of K_c [7]. While there are relatively few results about reconstruction numbers in general, Bollobás proved the following result by probabilistic means.

Theorem 3.9 (Bollobás [13])**.** *Almost every graph has reconstruction number 3.*

Here, 'almost every' is meant in the random graph sense[3], and can be understood as saying that the proportion of labeled graphs on n vertices satisfying this property tends to 1 as $n \to \infty$. We have already mentioned that it is not possible to reconstruct any graph from only two cards, so in that sense this result is the best possible. In fact, Bollobás' proof clearly identifies this barrier as it follows from a stronger result which says that almost every graph is determined by any two cards up to the

---

[3] More precisely, the statement says that this is true for Erdős-Rényi graphs $G_{n,1/2}$ with probability tending to 1 as $n \to \infty$. It also works for any fixed probability p.

edge joining the two deleted vertices. There is also a deterministic way to phrase this result. A graph G is said to have property A_k if, whenever A and B are distinct k-sets of vertices of G, the graphs $G[V(G) \setminus A]$ and $G[V(G) \setminus B]$ are not isomorphic. The preceding result can then be broken into two statements; firstly, that every graph with property A_3 (assuming this is known *a priori*) has reconstruction number 3, and secondly, that almost all graphs have property A_3.

Missing cards. Bollobás' result is even stronger than Theorem 3.9 (which is the standard statement) suggests, because not only do three well-chosen cards suffice, but in fact any three cards are enough for reconstruction. All of the subdeck variants that we have discussed so far either allowed us to carefully select a subdeck of cards, or at least tell us how those cards were chosen. Another subdeck variant comes from supposing that we are missing a part of our deck, so we have no control over (nor knowledge of) which cards are not present. This is equivalent to a parameter of Myrvold called the *adversary reconstruction number* of G, which is the size of the largest collection of cards that is also a subdeck of a graph not isomorphic to G, as a counterpart to the previous ally reconstruction number. The former is a more difficult problem since we must assume the worst case scenario at every step, and it follows that the results here are much stronger and accordingly even fewer.

For upper bounds on the number of missing cards, let $cc(G, H)$ denote the number of cards that G and H have in common, and let $cc(n) := \max\{cc(G, H) : G, H \text{ distinct graphs on } n \text{ vertices}\}$. The graph reconstruction conjecture states that $cc(n) \leq n - 1$ for $n \geq 3$. Bowler, Brown and Fenner [24] have constructed infinitely many families of pairs of graphs with $2\lfloor (n-1)/3 \rfloor$ cards in common, so $cc(n) \geq (\frac{2}{3} + o(1))n$. They conjecture that this may be the threshold.

Conjecture 3.10 (Bowler, Brown and Fenner [24]). *For large enough n every graph G is determined up to isomorphism by any $2\lfloor (n-1)/3 \rfloor + 1$ cards from $\mathcal{D}(G)$.*

A good approach would be to determine whether $cc(n) \geq (1 - o(1))n$. A positive answer would disprove Conjecture 3.10, whereas a negative answer would prove the Reconstruction Conjecture in a strong form.

We are nowhere near verifying Conjecture 3.10, even for simple parameters and classes. The main difficulty when cards are missing is that we lose important counting results including Kelly's Lemma. This makes it a challenge already to reconstruct the number of edges, which has been a major pursuit in the area. Myrvold [126] showed that this can be done from any $n - 1$ cards from the deck (i.e. with 1 card missing) for graphs with at least 7 vertices. Monikandan and Balakumar gave an approximate result, showing that the number of edges can be determined within 1 by any $n-2$ cards, and Woodall gave a much stronger version showing that for any $p \geq 3$ and n sufficiently large, the difference in number of edges between two graphs with $n-p$ cards in common is at most $p-2$. Brown and Fenner [28] then made the next concrete improvement, showing that the number of edges can be reconstructed when 2 cards are missing for graphs with at least 29 vertices, and Groenland, Guggiari and Scott [68] made the jump to show that this can still be done with $\frac{1}{20}\sqrt{n}$ missing cards.

If we know the number of edges m in a graph G, deducing its degree sequence from the complete deck was a simple matter of subtracting from m the number of edges seen in each card. Losing one card G_i does not pose a problem as the missing degree is given by $d_G(v_i) = 2m - \sum_{j \neq i} d_G(v_j)$. However, it is already an open problem as to whether we can reconstruct the degree sequence when 2 cards are missing. In Section 4.3 of this thesis, we show that much more can be done if we add an average degree assumption. Specifically, for any graph G with $n \geq 3$ vertices and average degree bounded by some $d \in \mathbb{N}$, we prove that:

- the number of edges can be reconstructed from any deck missing at most $\frac{n}{4d+6} - d - 5$ cards (Theorem 1.5);
- the number of copies of K_r can be reconstructed from any deck missing at most $\left(1 + \binom{2(d+1)}{r-1}\right)^{-1} \left(\frac{n}{2} - 1\right) - d - 5$ cards (Theorem 1.6);
- the degree sequence of G can be reconstructed from any deck missing at most $\frac{n}{10^4 d^3}$ cards (Theorem 1.4).

Since graphs that can be embedded on a fixed surface have bounded average degree, the following corollary is immediate.

Corollary 3.11. *For any surface S, there is an $\varepsilon > 0$ such that for any n-vertex graph G embeddable on S, the degree sequence of G can be reconstructed from any collection of at least $(1-\varepsilon)n$ cards.*

In particular, it is possible to reconstruct the degree sequence of planar graphs with a linear number of missing cards. This is tight up to a constant. For example, consider the graphs

$$G_1 = K_{1,p+1} \sqcup K_{1,p+1} \sqcup K_{1,p-1} \text{ and } G_2 = K_{1,p+1} \sqcup K_{1,p} \sqcup K_{1,p}$$

formed by the disjoint union of three stars. For both graphs, roughly two thirds of their cards are equal to $K_{1,p+1} \sqcup K_{1,p} \sqcup K_{1,p-1}$, and we might be unable to distinguish the two graphs even with nearly two thirds of the deck. Yet G_1 has two vertices of degree $p+1$, whereas G_2 has only one such vertex. These graphs do have the same number of edges, but we can find examples with a linear number of common cards and a different number of edges. For example, $K_{2,p} \sqcup K_{1,p}$ and $K_{2,p+1} \sqcup K_{1,p-1}$ share approximately half their cards yet have a different number of edges. These examples can be generalised to graph classes with a larger (constant) average degree d as well. Indeed, consider adding disjoint copies of the same $(3p+4)$-vertex graph H to both G_1 and G_2. The resulting graphs will still have about $\frac{2}{3} \times \frac{1}{2} = \frac{1}{3}$ of their cards in common, and we can create the desired average degree by choosing the density of H.

Groenland, Guggiari and Scott [68] conjectured that the degree sequence of a graph can be reconstructed from a deck of cards with a constant number k of missing cards (for n sufficiently large). Our results imply that this conjecture holds for graphs where the average degree is at most $c_k n^{\frac{1}{3}}$ (for some c_k depending only on k), but it is not yet known to hold for general graphs.

Conjecture 3.12 (Groenland, Guggiari and Scott [68])**.** *Fix $k \in \mathbb{N}$ and let n be sufficiently large. For any graph G on n vertices, the degree sequence of G is reconstructible from any $n-k$ cards.*

We also mention a result of Brown, Bowler, Fenner and Myrvold [25] that proves that the connectedness of a graph can be recognised from any $\lfloor n/2 \rfloor + 2$ cards. There does not appear to be much more related work on reconstructing other parameters, and the difficulty that this has posed has so far made it impractical to look at reconstructing classes with any significant number of cards missing.

3.4 Switching problems

In 1985, Stanley [162] introduced a variant of reconstruction that arises by replacing vertex- or edge-deletion with the following operation. For a graph G, *vertex-switching* at a particular vertex means to delete all the edges of G incident to that vertex, and inserting all possible non-edges incident to the same vertex. The remaining definitions follow the same pattern as in other variants. The *vertex-switching deck* of G consists of the multiset of graphs obtained by switching G at each of its vertices. An example for C_4+e is shown in Figure 3.2, where the switched vertex is marked only for illustrative purposes. We say that G is *vertex-switching-reconstructible* (or *VSR*) if it is uniquely determined by its switching deck. Although this deck is very much different from that in our classical problem, we will mirror the notation of the classical case and write

$\mathcal{D}(G)$ for the deck of G. In general throughout this thesis, is should always be clear which deck we are referring to since the definitions are not mixed within sections.

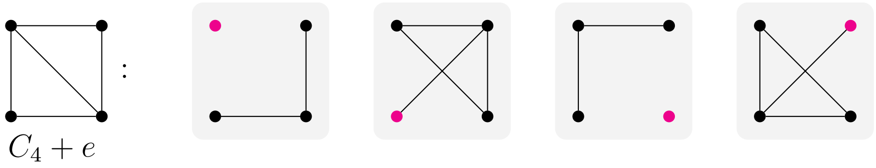


Figure 3.2: The vertex-switching deck of $C_4 + e$

As in the classical case, we can now ask the question: for two graphs G and G', does $\mathcal{D}(G) = \mathcal{D}(G')$ imply that $G \cong G'$? The answer is negative in general, since for example $C_4 + e$, P_4, and $K_2 \sqcup 2K_1$ (an edge with two isolated vertices) all have the same vertex-switching decks. Other known examples exist; the 11 non-isomorphic simple graphs on 4 vertices have just 6 different vertex-switching decks in total. However, these constitute all known examples of non-switching-reconstructible graphs. A search by Royle (see [104]) has shown that all graphs on 5 to 8 vertices are switching-reconstructible, and a search by Niesink [128] using an algorithm by McKay extends this to graphs on up to 12 vertices[4]. Accordingly, the main conjecture in this version of reconstruction is the following.

Conjecture 3.13 (Stanley [162])**.** *Every graph of order at least 5 vertices is switching-reconstructible.*

A key difference between switching reconstruction and vertex (or edge) reconstruction is that we can switch on the same vertex multiple times. This leads to quite a different flavour of arguments, and in particular lends itself to a host of enumerative and algebraic techniques that we will elaborate on in Chapter 5. To make use of this property though, we require some standard notation that will allow us to encode multiply switched decks. We can think of our switching operation as a function $\phi : \mathcal{G} \times V(G) \to \mathcal{G}$, and abbreviate $\phi(G, z)$ to G_z, which represents the graph G switched at the vertex z. This notation is sensible since we have an implied labelling of the vertices of G induced by the cards; given a deck, if we arbitrarily label the cards, then we know that each of those labels corresponds to some vertex in G. Of course, we cannot identify this labelling with any predetermined labelling of G. For

---

[4] Actually, the contribution from these computations are checking graphs on exactly 8 and 12 vertices, since we will shortly see that the remaining cases have been settled with an algebraic or counting argument.

multiple subsequent switches, we will write $(\ldots((G_{z_1})_{z_2})\ldots)_{z_k} = G_{z_1 z_2 \ldots z_k}$. It is easy to verify directly that the order of the switchings is not important.

Many of the first results that we saw for classical vertex reconstruction can be adapted to vertex-switching reconstruction. It is clear that the order of a graph is reconstructible since the switching operation does not change the number of vertices. We can also reconstruct the number of edges and degree sequence of a graph in a similar fashion to before.

Theorem 3.14 (Stanley [162]). *If G has order $n \neq 4$, the number of edges and degree sequence of G are VSR.*

Proof. Each edge of G will appear in $n-2$ of the cards in $\mathcal{D}(G)$, specifically all but the cards obtained by switching an endpoint of the edge. Meanwhile, every non-edge will appear in exactly two cards in $\mathcal{D}(G)$. Hence, if G has q edges then the total number of edges in all the graphs of $\mathcal{D}(G)$ is given by $q(n-2)+2\left(\binom{n}{2}-q\right) = n^2-n+(n-4)q$ which can be solved for q as long as $n \neq 4$. Then, as G_z is obtained by deleting $d(z)$ edges and adding the $n-d(z)-1$ non-edges adjacent to z, then G_z has $q+n-1-2d(z)$ edges which can be solved for $d(z)$. $\square$

It is really necessary that $n \neq 4$ in the preceding statement since we have mentioned that $C_4 + e$ and P_4 have different numbers of edges but the same vertex-switching decks. This begins to hint at the general phenomenon that the order of the graph is pertinent to the difficulty of reconstructibility. Knowing the degree sequence now leads us to our first VSR class.

Theorem 3.15 (Ellingham and Royle [52]). *Regular graphs of order $n \neq 4$ are VSR.*

The idea of the proof is that in an r-regular graph G, each card G_v has one vertex of degree $n-1-r$ (that being v), r vertices of degree $r-1$ (neighbours of v) and $n-1-r$ vertices of degree $r+1$. If all of these degrees are distinct, it is easy to recover v and then use the property $G_{vv} = G$. In cases where they are not all distinct, it turns out that it is still possible to find a vertex u such that $G_{vu} \cong G$. Slightly more sophisticated partitionings of the vertex set based on adjacency have successfully been used to prove that disconnected graphs [98] and triangle-free graphs [52] of order $n \neq 4$ are VSR. Alongside the latter result, Ellingham and Royle also proved that we can reconstruct the number of triangles in any graph on $n \neq 4$ vertices. This is a special case of an analogue of Kelly's Lemma for vertex-switching reconstruction.

Theorem 3.16 (Ellingham and Royle [52])**.** *Given a graph G on n vertices and a graph S on $k < \frac{n}{2}$ vertices, the number of induced subgraphs (or just subgraphs) of G isomorphic to S can be reconstructed from $D(G)$.*

As a sanity check, we note that the condition in this statement does exclude the possibility of counting edges (which involve two vertices) when $n = 4$.

The next class of VSR graphs that we consider sets this problem apart from all other variants we have discussed in terms of approachability. Suppose we are given a deck $\mathcal{D}(G)$, and form the multiset $\mathcal{D}(\mathcal{D}(G))$ by performing every possible vertex-switching in every card of $\mathcal{D}(G)$. Then $\mathcal{D}(\mathcal{D}(G))$ provides a finite list of graphs among which G is guaranteed to be present. We can be more specific about the counts here. Since vertex-switchings are commutative, there are exactly n cards of the form $G_{ii} = G$ in $\mathcal{D}(\mathcal{D}(G))$ and all other cards G_{ij} $(i \neq j)$ are equal[5] to G_{ji}. Now if n is odd, then the only graph appearing an odd number of times in $\mathcal{D}(\mathcal{D}(G))$ must be G. This proves that all graphs with an odd number of vertices are VSR.

The preceding argument was observed by Stanley at the time that he introduced this variant. With some more work, Stanley showed further that any graph with order not divisible by 4 is switching-reconstructible.

Theorem 3.17 (Stanley [162])**.** *If G has order n, and $n \not\equiv 0 \pmod 4$, then G is VSR.*

Stanley's original proof involves a moderately involved argument using linear algebra which we will briefly touch on in Chapter 5. Several other proofs exist of this result using different methods [1, 52, 102]. Among these is a beautiful argument attributed to Alon and Coppersmith. To describe it, however, we will need the more general setup described in Section 3.4.1 so we will return to this later.

Finally, we mention one more result that is typical in switching-reconstruction but does not fall into the usual categories of reconstructible properties or (natural and easily described) classes. In [98], Krasikov determines a bound for the ratio of numbers of induced subgraphs of G, and then bounds these in terms of degrees of vertices in the graph. This leads to a sufficient condition for vertex-switching reconstructibility.

Theorem 3.18 (Krasikov [98])**.** *If $\min\left(n\binom{n-1}{\Delta}, n\binom{n-1}{\delta}\right) < 2^{\frac{n}{2}-3}$, then G is VSR.*

This can be useful when applied in conjunction with the reconstructibility of the degree sequence. For example, it implies that a graph with a vertex of degree 1 is reconstructible if it has order at least 28 which might make casework manageable. We will see related sufficient conditions in other settings.

---

[5] They are really equal as graphs with labellings inherited from G, and not just isomorphic.

3.4.1 k-switchings

Just as the classical vertex reconstruction problem had a natural extension to ℓ-decks, this is also the case for switchings. When we discussed iterated switchings, we noted that the order in which switches are performed did not matter. In light of this, we can extend the definition of our switching operation ϕ in a natural way; namely, given a set $W = \{z_1, z_2 \dots, z_k\} \subseteq G$ where the labelling of vertices is arbitrary, then $G_W = G_{z_1 z_2 \dots z_k}$. That is, switching on each vertex in the set is the same as switching 'simultaneously' on a set. The *k-switching-deck* (or *k-deck*) of G is the multiset $\mathcal{D}_k(G) = \{[G_W] : W \subseteq Z,\ |W| = k\}$ where $[G]$ is the isomorphism class of graphs containing G. It is convenient to write this as a formal sum

$$\mathcal{D}_k(G) = \sum_{W \subseteq Z \,:\, |W|=k} [G_W].$$

Note that $\mathcal{D}_0(G) = [G_\varnothing] = [G]$, $G_{\{z\}} = G_z$ and the 1-deck of G is $\mathcal{D}_1(G) = \mathcal{D}(G)$.

Confirming what we would intuitively expect of k-switchings, the following is a summary of some basic properties most of which are proved in [128]. Note that these statements use equality of graphs, and not just isomorphism.

Lemma 3.19. *Let $G, H \in \mathcal{G}$ have the same labelled vertex set Z, $u, v \in Z$ and $W, X, Y \subseteq Z$. Then*

1. *$G_{vv} = G$ and more generally, $(G_X)_Y = G_{X \triangle Y}$.*
2. *$G_{vu} = G_{uv}$ and more generally $G_{z_1 \dots z_k} = G_{z_{\rho(1)} \dots z_{\rho(k)}}$ where $\rho \in S_k$ is any permutation. In particular, this means that $\phi' : \mathcal{G} \times \mathcal{P}(Z) \to \mathcal{G}$, where $\phi'(G, W) = G_W$ is given by switching G at the vertices in W in any order, is well-defined.*
3. *If $G_X = G_Y$ then either $X = Y$ or $X = Z - Y$.*
4. *$G_W = G_{Z-W} = G_{\overline{W}}$, and in particular $G_Z = G$.*
5. *For any $v \in Z$, we have $G_v = H_v \iff G = H$.*
6. *If $|Z| \geq 3$, then $G_v = G_w \implies v = w$.*

Studying k-decks is an important technique for proving reconstructibility results from the 1-deck. This is a point where switching reconstruction really deviates from the classical setting; whereas taking a k-deck before meant deleting more information and therefore obtaining a harder problem, switching on k-subsets forms larger decks that actually provide more data via algebraic relationships. One example of this is in the proof of Theorem 3.17, which follows the ideas of Alon and Coppersmith. We record it since it is quite elusive in literature despite being alluded to in several places (mentioned in [162]).

Proof of Theorem 3.17 (following Alon and Coppersmith). Suppose G and G' are both graphs of order n with $n \not\equiv 0 \pmod 4$, and have the same deck $\mathcal{D}_1(G) = \mathcal{D}_1(G')$. We shall prove that G and G' are isomorphic. For $0 \le k \le n$, consider $\mathcal{D}_k(G)$ and observe that $\mathcal{D}_0(G) = \{G\}$ and $\mathcal{D}_i(G) = \mathcal{D}_{n-i}(G)$. Moreover, it is clear from the definition that $\mathcal{D}_1(G) = \mathcal{D}_1(G')$ implies $\mathcal{D}_k(G) = \mathcal{D}_k(G')$. So consider $\mathcal{D}_k(\mathcal{D}_1(G))$, which is the multiset formed by switching about every set of k vertices for each card in $\mathcal{D}_1$. This contains $n\binom{n}{k}$ graphs, and in particular it can be seen that

$$\mathcal{D}_k(\mathcal{D}_1(G)) = (n-k+1)\mathcal{D}_{k-1}(G) + (k+1)\mathcal{D}_{k+1}(G) \tag{3.4.1}$$

and similarly with G replaced by G'. Since $\mathcal{D}_1(G) = \mathcal{D}_1(G')$, we have $\mathcal{D}_k(\mathcal{D}_1(G)) = \mathcal{D}_k(\mathcal{D}_1(G'))$ for all k, so

$$(n-k+1)\mathcal{D}_{k-1}(G)+(k+1)\mathcal{D}_{k+1}(G) = (n-k+1)\mathcal{D}_{k-1}(G')+(k+1)\mathcal{D}_{k+1}(G'). \tag{3.4.2}$$

If we let $k = 2j-1$, this becomes

$$(n-2j+2)\mathcal{D}_{2j-2}(G) + 2j\mathcal{D}_{2j}(G) = (n-2j+2)\mathcal{D}_{2j-2}(G') + 2j\mathcal{D}_{2j}(G'). \tag{3.4.3}$$

Setting $j = 1$ in (3.4.3) gives $n\mathcal{D}_0(G) + 2\mathcal{D}_2(G) = n\mathcal{D}_0(G') + 2\mathcal{D}_2(G')$. There is only one term on either side of the equation that has an odd coefficient, so we must have $\mathcal{D}_0(G) = \mathcal{D}_0(G')$ which then implies $G \cong G'$. The remaining case is when $n \equiv 2 \pmod 4$, so without loss of generality we can write $n = 4i+2$. Letting $k = 2i+1$ in (3.4.2) and noting that $\mathcal{D}_{2i+2} = \mathcal{D}_{n-(2i+2)} = \mathcal{D}_{4i+2-2i-2} = \mathcal{D}_{2i}$, we have

$$\begin{aligned}(n+2)\mathcal{D}_{2i}(G) &= (n-2i)\mathcal{D}_{2i}(G) + (2i+2)\mathcal{D}_{2i+2}(G)\\ &= (n-2i)\mathcal{D}_{2i}(G') + (2i+2)\mathcal{D}_{2i+2}(G')\\ &= (n+2)\mathcal{D}_{2i}(G')\end{aligned}$$

so $\mathcal{D}_{2i}(G) = \mathcal{D}_{2i}(G')$. From (3.4.3), if $\mathcal{D}_{2j}(G) = \mathcal{D}_{2j}(G')$ then $\mathcal{D}_{2j-2}(G) = \mathcal{D}_{2j-2}(G')$. Hence $\mathcal{D}_0(G) = \mathcal{D}_0(G')$, which implies that $G \cong G'$. □

Considering k-decks also leads to another reconstruction problem: is a graph G uniquely determined by its k-switching-deck? We will say G is k-VSR when the answer is positive. One might hope that results from the previous section will generalise to this closely related problem, but the answer is mixed. It is true that a graph is k-VSR if and only if its complement is k-VSR, and the proof is largely the same. On the other hand, the counting and partitioning arguments that dominated proofs in the previous section are no longer tractable in general, so we do not even know whether degree sequences are reconstructible for any $k > 1$. In better news, it is known that the number of edges is k-VSR when $k < n/2$ with $n \neq a^2$ and $n \neq \binom{a}{2}$ for any a [103], complete bipartite graphs on at least 4 vertices are k-VSR for all k [104]. In addition, we have the following (partial) analogue of Theorem 3.16.

Theorem 3.20 (Krasikov and Roditty [103])**.** *Let G and H be graphs with n and n' vertices respectively. The number of subgraphs of G isomorphic to H is reconstructible from $\mathcal{D}_k(G)$ provided*

$$\binom{n-n'}{k} + \binom{n-n'}{k-n'} > \frac{1}{2}\binom{n}{k}.$$

Krasikov and Roditty [103] also gave a stronger version of the preceding lemma which allows us to obtain edge counts in more cases. It says that if $|V(H)| = 2, 3$, then the number of subgraphs of G isomorphic to H is reconstructible from $\mathcal{D}_k(G)$ except possibly when $k = \binom{t}{2}$ and $n = t^2$, or $n = (t-1)^2$ for an integer $t \geq 2$. Furthermore, they proved that for each $k \geq 4$ there exists an integer N_k such that a graph is k-VSR whenever $n > N_k$ for k even, or $n > N_k$ and $n \not\equiv 0 \pmod 4$ when k is odd. This is one of multiple extensions of Theorem 3.17 for the k-switching setting (others can be found in [162, 103, 104]). We also have another sufficient condition in terms of degrees that generalises Theorem 3.18 [99]. As a disclaimer, we mention that most of these statements involve some rather mysterious conditions that come from studying the roots of a certain family of polynomials called the Krawtchouk polynomials. We will discuss how this comes about in Section 5.2 en route to proving new results that further exploit this algebraic connection.

3.4.2 Other settings and a generic formulation

There are myriad reconstruction problems in literature concerning slightly different objects that also call themselves ‘switching reconstruction problems’. After Stanley’s vertex switchings, the most widely investigated is a digraph switching problem introduced by Bondy and Mercier [20, 122] in 2011. If D is an oriented graph, then switching at a vertex z forms a card D_z which is the oriented graph obtained by reversing the orientation of every arc adjacent to z in D. Here, we work up to isomorphism of oriented graphs. A different problem arises if we replace oriented graphs with general digraphs.

Bondy and Mercier [20] also introduced switching reconstruction on edge-coloured graphs. Consider the set of (improperly) 2-edge-coloured simple graphs, where switching at a vertex swaps the colours of the edges incident to that vertex. This operation does not change the underlying graph, so the problem here is to recover the colouring. The same proof as that for Theorem 3.17 carries over to this problem, and so we have again that a 2-edge-coloured graph is reconstructible (up to automorphism of the underlying graph) if it has order $n \not\equiv 0 \pmod 4$. On the other hand, there are examples of non-reconstructible two-edge-coloured graphs on 8 vertices. In fact, this is a special case of Stanley's problem since if we consider only complete graphs and let the two colours be edge and non-edge, then we have a bijective correspondence between 2-coloured complete graphs and simple graphs on the same vertex set where the switching operations coincide.

Interpolating between the digraph and coloured edge problems, Bondy and Mercier also proposed working with edge-coloured hypergraphs with a digraph defined on the vertices of each hyperedge. Switching at a vertex v would entail permuting the colours of each hyperedge that contains v, and also reversing the orientation of all arcs incident to v within each hyperedge. There are again some results given in [20] for particular permutations, and when we have two colours.

In [1], Abatangelo and Dragomir considered a weakening of Stanley's problem involving the combinatorial Laplacian spectrum[6]. The weakened problem is stated as follows. Let $n \geq 4$ and $\mathcal{G}$ be the set of graphs on a vertex set Z such that for any $G \in \mathcal{G}$, the degree of each vertex x of G satisfies $2 \leq d(x) < n - 1$. We keep the switching operation as before, so a card G_x is obtained by exchanging the adjacencies and non-adjacencies at x. Now given $G, H \in \mathcal{G}$, does $\sigma(G_x) = \sigma(H_x)$ for each $x \in V$ imply that $\sigma(G) = \sigma(H)$?

All of these problems have natural k-switching versions as well by simply extending the switching operation to sets of vertices, although it is a bit more involved to verify properties equivalent to those in Lemma 3.19 for some operations than others. On the other hand, the fact that some of the results we have mentioned can be directly transferred to other switching problems by simply translating the proofs suggests that it might be possible to treat these problems together. This leads to the question; what characteristics do all of these problems have in common? In partial answer to this, we now develop a generic formulation for switching problems.

---

[6] The Laplacian matrix of a graph, denoted δ is defined as $\delta = D - A$ where D is the degree matrix and A is the adjacency matrix of the graph. The spectrum $\sigma(G)$ of G is the set of eigenvalues of δ.

Let $\mathcal{G}$ be a set and Z a finite set. Since switching does not change our vertex set in all the instances of our problem, we should think of $\mathcal{G}$ as being graphs (or digraphs, or edge-coloured graphs...) on a fixed vertex set Z. The examples we have seen only specify switching operations that are carried out on a single vertex. Given such an elementary switching $(G, z) \mapsto G_z$ for $G \in \mathcal{G}, z \in Z$ satisfying $(G_z)_z = G$ and $(G_y)_z = (G_z)_y$, we can extend to all of $\mathcal{G} \times \mathcal{P}(Z)$ by setting $G_\emptyset = G$ and $G_{\{w_1,\ldots,w_k\}} = G_{w_1,\ldots,w_k}$. Thus, we may as well consider a function $\phi : \mathcal{G} \times \mathcal{P}(Z) \to \mathcal{G}$. For $G \in \mathcal{G}$, $z \in Z$ and $W \subseteq Z$, we will write G_W for $\phi(G, W)$ and G_z for $\phi(G, \{z\})$. Let "$\sim$" be an equivalence relation on $\mathcal{G}$, and $\mathcal{G}/\sim$ the set of equivalence classes. For $G \in \mathcal{G}$, the equivalence class containing G is denoted by $[G]$. In summary, a switching reconstruction problem consists of:

$\mathcal{G}$: the set of objects under consideration,
Z: index set of switching operations,
ϕ: a function $\mathcal{G} \times \mathcal{P}(Z) \to \mathcal{G}$ that defines the switching operations,
$\sim$: a notion of equivalence between objects.

Our next task is to distil the behaviour that characterises switching operations, and distinguishes it from one of the deletion operations in vertex- or edge- reconstruction. We will say that $(\phi, \sim)$ is a *switching operation* if it satisfies these properties:

S1. For $G \in \mathcal{G}$, $[G_\emptyset] = [G]$. That is, $\mathcal{D}_0(G) = [G]$.
S2. For $G, G' \in \mathcal{G}$, $G \sim G' \Rightarrow \mathcal{D}_1(G) = \mathcal{D}_1(G')$.
S3. For $X, Y \subseteq Z$ and $G \in \mathcal{G}$, $(G_X)_Y \sim G_{X \triangle Y}$.

This formulation was developed with Brendan McKay and Béata Faller. As far as we are aware, all 'switching reconstruction problems' in literature have an operation satisfying these conditions. The goal, then, is to unify some switching reconstruction arguments to prove generic statements that can then be applied to specific instances. This is the framework in which we will work in Section 5.2. For the moment, we mention a few basic but indispensable consequences of our chosen axioms.

Lemma 3.21. *For $G, G' \in \mathcal{G}$ and any switching operation $(\phi, \sim)$, the following hold.*

(a) For each $W \subseteq Z$ with $W = \{z_1, z_2, \ldots, z_k\}$, $G_{z_1,z_2,\ldots,z_k} \sim G_W$. In particular, this is independent of the ordering.

(b) For any $W \subset Z$, $(G_W)_W \sim G$

(c) If $G \sim G'$ then $\mathcal{D}_k(G) = \mathcal{D}_k(G')$ for all $k \geq 0$.

We will freely apply these facts in later sections, as well as add some more interesting properties to the list.

3.5 Non-reconstructible structures

The fact that each variant of graph reconstruction we have seen so far has come with a plausible open conjecture is not merely due to cherry-picking; it seems genuinely to be the case that while reconstruction-type problems on graphs are difficult, subgraphs (and switching decks) do carry sufficient information to suggest a positive answer. For edge and switching reconstruction, we have seen some variants on structures other than finite undirected graphs. It should not come as a surprise that similar questions have also been considered in the classical reconstruction setting. What may be surprising, however, is that several of the natural corresponding reconstruction conjectures fail.

In 1964, Harary [75] asked whether all sufficiently large tournaments are reconstructible from their multiset of vertex-deleted subtournaments. This was widely believed to be true until Stockmeyer [163] gave a family of arbitrarily large counterexamples in 1977.[7] In light of this, work in digraph reconstruction has been geared toward finding a characterisation of the reconstructible digraphs. By refining his earlier methods, Stockmeyer also found many more examples of non-reconstructible digraphs including ones that are not tournaments (see [164], as well as an unpublished Part II [166] written circa 1986). One positive result is that tournaments on at least five vertices which are not strongly connected are reconstructible [79]. Another direction, initiated by Ramachandran [138], has been on reconstructing digraphs from a deck where each card $D-v$ is augmented with the degree data $(d^+(v), d^-(v))$. In what is known as the 'New Digraph Reconstruction Conjecture' (or the 'N-Reconstruction Conjecture for Digraphs'), Ramachandran conjectured that all digraphs are reconstructible from this information and there are no known counterexamples.

It is less clear when looking at reconstruction problem on hypergraphs how to generalise our standard definitions. Kocay [90] gave the following formulation in the k-uniform setting. Given a k-uniform hypergraph G with vertex set $V(G)$ and edge set $\mathcal{E}(G) \subseteq V(G)^{(k)}$, and a set $X \subset V(G)$, the induced subhypergraph $G[X]$ has vertex set X and edge set $\{e \in \mathcal{E}(G) : e \in X^{(k)}\}$. For a vertex $u \in V(G)$, the vertex-deleted hypergraph $G - u$ is then defined to be $G[V(G) \setminus \{u\}]$ and this allows us to state the directly analogous reconstruction problem. When $k = 2$, this is simply the classical vertex reconstruction problem, and just as we had a simple example of a small non-reconstructible pair in that case, for general k the empty hypergraph and the graph with a single edge $[k]$, both on the vertex set $[k]$, have the same deck.

---

[7] The original paper contained a couple of errors which are corrected in [89, 165].

For $k = 3$, Kocay [90] showed that the corresponding reconstruction conjecture is false and expressed confidence that it should also be false for higher k, although this has not been verified and is expected to be very involved. Much later, Kocay [92] gave a catalogue of further examples of non-reconstructible 3-uniform hypergraphs on few vertices. It turns out that this question is not totally independent from that for tournaments; Conilh [31] demonstrated a connection between Kocay's early hypergraph counterexamples and Stockmeyer's tournaments.

For infinite graphs, it is easy to see that the direct analogue of the Reconstruction Conjecture is false; let T be the tree in which every vertex has degree $\aleph_0$, and $T \sqcup T$ be the disjoint union of two copies of T. Then each of the $\aleph_0$ cards in $\mathcal{D}(G)$ consists of countably many disjoint copies of T, but the same is true for $T \sqcup T$ which is not isomorphic to T. This example, found by Fisher, Graham and Harary [59], shows that connectedness is already not a reconstructible property when it comes to infinite graphs. Moreover, the complements $\overline{T}$ and $\overline{T \sqcup T}$ provide examples of connected graphs that are not reconstructible, and Andrae [3] has given further pairs of non-reconstructible infinite graphs in which both the graphs and their complements are connected among a number of other examples.

Nonetheless, there are more challenging questions regarding infinite graphs, many of which appear in Nash-Williams [160] survey from 1991. In particular, Nash-Williams asked whether locally finite connected infinite graphs are reconstructible, as well as whether infinite trees are weakly reconstructible[8]. The former is related to a conjecture of Harary, Schwenk and Scott [81] that every locally finite tree is reconstructible, which had been verified in several cases. A question of Halin also appears in [160], which asks whether for any two infinite graphs G and H with the same decks, it is true that there exist embeddings $G \hookrightarrow H$ and $H \hookrightarrow G$. All of these were settled negatively by Bowler, Erde, Heinig, Lehner and Pitz [26] in 2017, when they gave a recursive construction of two infinite trees with maximum degree three, the same deck, and neither tree embeds into the other. A follow-up paper by the same authors [27] presents a pair of non-reconstructible locally finite connected graphs with one and countably many ends[9]. This completes a picture started by Nash-Williams [159, 161] who showed that locally finite infinite connected graphs with finitely many (at least two) ends are reconstructible.

---

[8] The example given in the preceding paragraph shows that they are not reconstructible in general, but one of these graphs is a forest so weak reconstruction within the class of trees was still possible.

[9] Informally, an *end* of an infinite graph is a direction in which it extends out infinitely.

Chapter 4

Direct counting

While many direct arguments in reconstruction rely on specific properties of a set-up or graph class, the beauty of counting techniques is that they can be very general. We have already alluded to the fact that Kelly's Lemma, our most basic counting tool, is of central importance in reconstruction. The purpose of this section is to discuss and develop further methods that rely on direct counting for the small cards variant, where we can draw generously from the substantial literature on counting in classical reconstruction, and the missing cards variants, where we are in a considerably earlier stage of development.

In Section 4.1, we state the generalisation of Kelly's Lemma to small cards and use this to derive three more counting tools. The first two of these, due to Kocay and Greenwell-Hemminger, are well-established, whilst the third is a new technique for counting 'extensions' of subgraphs. These tools are central to current work: Kocay's Lemma was an important ingredient in Bilinksi, Kwon and Yu's [9] breakthrough on recognising planar graphs, and we use Greenwell and Hemminger's Lemma together with our extension counting theorem in Section 4.2 to prove Theorem 1.1 on reconstructing trees from small cards. We then use approximate counting methods to show that the degree sequence can be reconstructed from a deck with linearly many cards missing in Section 4.3, proving Theorems 1.4, 1.5 and 1.6.

4.1 Beyond Kelly's Lemma

We begin by saying a bit more about Kelly's Lemma. We will reserve the word *copy* of H for an induced subgraph isomorphic to H, and say an *instance* of H to mean a not necessarily induced subgraph. Let $n_H(G)$ be the number of vertex sets in G inducing a copy of H, and let $\widetilde{n}_H(G)$ denote the number of instances of H in G. In the classical graph reconstruction problem, Kelly's Lemma (Lemma 2.4) states that we can reconstruct $n_H(G)$ and $\widetilde{n}_H(G)$ provided $|V(H)| < |V(G)|$, and there are many variants of the lemma for other reconstruction problems (see [17]). In particular, there is a very natural extension using exactly the same proof for reconstruction from smaller cards.

Lemma 4.1. *Let $\ell \in \mathbb{N}$ and let H be a graph on at most ℓ vertices. For any graph G, the multiset of ℓ-vertex induced subgraphs of G determines both the number of subgraphs of G that are isomorphic to H and the number of induced subgraphs that are isomorphic to H.*

Proof. Suppose we count the number of copies of H in each of the ℓ-cards of G, and take the sum over all cards. Each copy of H in G will be counted exactly $\binom{n-|V(H)|}{\ell-|V(H)|}$ times toward this total. Hence, we can reconstruct the number $n_H(G)$ of copies of H in G from the ℓ-deck as

$$n_H(G) = \binom{n-|V(H)|}{\ell-|V(H)|}^{-1} \sum_{C \in \mathcal{D}_\ell(G)} n_H(C).$$

The same argument applies with instances rather than copies. □

In particular, Kelly's Lemma implies that $\mathcal{D}_{\ell'}(G)$ can be reconstructed from $\mathcal{D}_\ell(G)$ for all $\ell' \leq \ell$. Foreshadowing later usage in Section 4.2, we remark that Kelly's Lemma only requires the subset of the deck consisting of the cards which contain at least one copy of the fixed graph H.

The remainder of this section is devoted to discussing three more tools (which can be equivalently viewed as reconstructible parameters, although they are not standard parameters from a graph theory point of view) that build off Kelly's Lemma in an essential way. The common idea is to reconstruct certain parameters by expressing them, via standard tools such as double counting or inclusion-exclusion, in terms of subgraphs that are small enough to be covered by Kelly's lemma. We will give the more general ℓ-deck versions in each case. These follow directly from the original arguments, although we have not seen these exact statements in literature.

4.1.1 Spanning subgraphs

In 1979, Tutte [172] proved that it is possible to reconstruct counts of certain spanning subgraphs of a graph from its deck. These results are especially nice because, in some sense, they overcome the reasonable necessity in Kelly's Lemma that the subgraph we are counting fits within a single card. While Tutte's original proof was algebraic, soon after in 1981, Kocay [88] gave an alternative proof by elementary counting methods. We describe the lemma that he formulated for this, which has turned out to be quite powerful in the later development of more counting results.

Let $\vec{\mathcal{F}} = (F_1, F_2, \ldots, F_s)$ be a finite sequence of graphs which do not necessarily need to be distinct. For a graph G, a *cover* of G by $\vec{\mathcal{F}}$ is a sequence $(G_1, G_2, \ldots, G_s)$ of instances of subgraphs of G such that $G_i \cong F_i$ for each $1 \leq i \leq s$ and $\bigcup_{i=1}^{s} G_i = G$. We will write $c(\vec{\mathcal{F}}, G)$ for the number of distinct covers of G by $\vec{\mathcal{F}}$. In this count, it is really important to distinguish between different instances of a subgraph, and the union of subgraphs is still a subgraph.

Lemma 4.2 (Kocay [88])**.** *Let G be any graph and $\vec{\mathcal{F}} = (F_1, F_2, \ldots, F_s)$ be any of sequence of graphs such that $|V(F_i)| \leq \ell$ for each $1 \leq i \leq s$. Then we can reconstruct from the ℓ-deck of G the parameter*

$$\kappa_\ell(\vec{\mathcal{F}}, G) := \sum_X c(\vec{\mathcal{F}}, X)\widetilde{n}_X(G)$$

where the sum is taken over all isomorphisms classes of graphs X with order in $\{\ell + 1, \ldots, |V(G)|\}$.

Proof. By double counting the number of sequences of subgraphs $(G_1, G_2, \ldots, G_s)$ of G such that $G_i \cong F_i$ for each i, we get the expression

$$\prod_{i=1}^{s} \widetilde{n}_{F_i}(G) = \sum_X c(\vec{\mathcal{F}}, X)\widetilde{n}_X(G)$$

where the sum is currently taken over all isomorphism classes of graphs X. Since $|V(F_i)| \leq \ell$ for each i, Kelly's Lemma implies that $\widetilde{n}_{F_i}(G)$ is reconstructible. Therefore, the product on the left hand side is reconstructible, and so too is the right hand side of the equation. Note that for X with $|V(X)| > |V(G)|$, we have $\widetilde{n}_X(G) = 0$ and hence these do not contribute to the sum. On the other hand, if $|V(X)| \leq \ell$ then we can reconstruct $\widetilde{n}_X(G)$ using Kelly's Lemma and $c(\vec{\mathcal{F}}, X)$ directly (since both $\vec{\mathcal{F}}$ and X consist of explicitly known graphs). Thus, we can restrict this summation to X with order in the claimed range. □

To see how Kocay's lemma in the $\ell = n - 1$ case can be used to count spanning structures, we use the example of disconnected spanning subgraphs of a graph G on n vertices, such as perfect matchings. Suppose we have a disjoint graph F with components $\vec{\mathcal{F}} = (F_1, \ldots, F_s)$ satisfying $s \geq 2$ and $\sum_{i=1}^{s} |V(F_i)| = n$. By Kocay's lemma, we can reconstruct $\sum_X c(\vec{\mathcal{F}}, X)\widetilde{n}_X(G)$, and we may as well restrict the sum to be taken over X with n vertices and for which $c(\vec{\mathcal{F}}, X) > 0$. However, with the conditions $|X| = |V(F)| = n$ and $c(\vec{\mathcal{F}}, X) > 0$, the only such X must be isomorphic to F. Thus, the summation becomes $c(\vec{\mathcal{F}}, F)\widetilde{n}_F(G)$, which allows us to deduce $\widetilde{n}_F(G)$.

There is also an induced version of Kocay's Lemma, where an *induced vertex cover* is a sequence of induced subgraphs of G and we replace $\widetilde{n}_X(G)$ with $n_X(G)$. This is the version that is used in an important recent application to recognising planar graphs from the classical deck. Specifically, Bilinksi, Kwon and Yu [9] apply Kocay's Lemma to count the number of non-separating cycles in a graph, which they combine with a lemma of Tutte that states that a 3-connected graph G is planar if and only if every edge of G is contained in exactly two induced non-separating cycles. This is a key part of their argument allowing them to handle graphs of connectivity at least 3, and the remaining 2-connected case is then addressed directly. We believe that this lemma is quite likely to play a role in upcoming developments in vertex reconstruction as well.

4.1.2 Maximal subgraphs

Given a class of graphs $\mathcal{F}$, a subgraph F' of some graph G is said to be an *$\mathcal{F}$-subgraph* if F' is isomorphic to some $F \in \mathcal{F}$, and is a *maximal $\mathcal{F}$-subgraph* if the subgraph F' cannot be extended to a larger $\mathcal{F}$-subgraph, that is, there does not exist an $\mathcal{F}$-subgraph F'' of G such that $V(F') \subsetneq V(F'')$. Let $m(F, G)$ denote the number of maximal $\mathcal{F}$-subgraphs in G which are isomorphic to F. We give a slight variation of a classical "Counting Theorem" due to Bondy and Hemminger [19] (see also the statement of Greenwell and Hemminger [66]) which allows us to reconstruct $m(F, G)$ from the ℓ-deck.

Lemma 4.3. *Let $n \in \mathbb{N}$, let $\ell \in [n-1]$ and let $\mathcal{G}$ be a class of n-vertex graphs. Let $\mathcal{F}$ be a class of graphs such that for any $G \in \mathcal{G}$ and for any $\mathcal{F}$-subgraph F of G,*

(i) $|V(F)| \le \ell$;

(ii) F is contained in a unique maximal $\mathcal{F}$-subgraph of G.

Then for all $F \in \mathcal{F}$ and $G \in \mathcal{G}$, we can reconstruct $m(F, G)$ from the collection of cards in the ℓ-deck of G that contain an $\mathcal{F}$-subgraph.

The following proof is essentially that of Bondy and Hemminger [19], only with a few additional observations used to accommodate slight changes to the assumptions.

Proof. Define an *(F,G)-chain* of length k to be a sequence $(X_0,\dots,X_k)$ of $\mathcal{F}$-subgraphs of G such that

$$F \cong X_0 \subsetneq X_1 \subsetneq \cdots \subsetneq X_k \subsetneq G.$$

The *rank* of F in G is the length of a longest (F,G)-chain, and two chains are called *isomorphic* if they have the same length and the corresponding terms are isomorphic. Following Bondy and Hemminger's argument, we first show that

$$m(F,G) = \sum_{k=0}^{\operatorname{rank} F} \sum (-1)^k \widetilde{n}_F(X_1)\widetilde{n}_{X_1}(X_2)\cdots\widetilde{n}_{X_{k-1}}(X_k)\widetilde{n}_{X_k}(G) \tag{4.1.1}$$

where the second summation is over all non-isomorphic (F,G)-chains of length k. When $\operatorname{rank} F = 0$, we have $m(F,G) = \widetilde{n}_F(G)$. Let $\operatorname{rank} F = r$, and suppose that (4.1.1) holds for all graphs $F \in \mathcal{F}$ with rank less than r. The second assumption states that every copy of F has a unique maximal extension X, which implies that

$$\widetilde{n}_F(G) = \sum_X \widetilde{n}_F(X)m(X,G),$$

where the sum is over all non-isomorphic $\mathcal{F}$-subgraphs X of G. This gives the expression

$$m(F,G) = \widetilde{n}_F(G) - \sum_{X \ncong F} \widetilde{n}_F(X)m(X,G).$$

In the summation, we can restrict to X for which $\widetilde{n}_F(X) > 0$. Such a graph X has rank at most $r-1$, so we may apply the induction hypothesis to rewrite each $m(X,G)$-term into a double sum. The resulting triple sum can be simplified to obtain (4.1.1).

It now suffices to show that the RHS of (4.1.1) is reconstructible. To see this, we note that the inner summation is over (F,G)-chains for which X_k has size at most ℓ (since X_k is an $\mathcal{F}$-subgraph and by condition (i)), and so all such chains can be seen on cards. The remaining terms can be reconstructed by Kelly's Lemma (again using (i)), and this only requires the cards from $\mathcal{D}_\ell(G)$ that contain an $\mathcal{F}$-subgraph. □

Kelly's Lemma is a special case of the above; to count the number of copies of H in G, we can set $\mathcal{F}$ to contain just H and let $\mathcal{G}$ be the class of all graphs with same deck as $\mathcal{D}(G)$. As observed in [17], Greenwell-Hemminger also gives a quick proof that any disconnected graph G on n vertices is reconstructible from the classical deck. It is clear that they are recognisable. For weak reconstructibility, we can take $\mathcal{F}$ to be the class of connected graphs on up to $n-1$ vertices, and $\mathcal{G}$ the class of disconnected graphs on n vertices. Condition (ii) from the statement holds as each F can be extended to a unique component, and the counts $m(F, G)$ give us the isomorphism classes of the components of G.

4.1.3 Marked extensions

We now introduce a new technique akin to those of the preceding sections, which first appeared in [70] joint with Groenland, Johnston and Scott. The purpose of this strategy is to count copies of a graph within G while keeping track of a specific distinguished subgraph. Intuitively, this is useful for reconstruction because it can preserve crucial information for determining how to glue pieces of a graph back together.

Given a graph H, we define an H*-extension* to be a pair $H_e = (H^+, A)$ where H^+ is a graph and $A \subseteq V(H^+)$ is a subset of vertices with $H^+[A] \cong H$. The idea is that H^+ may contain multiple copies of H as a subgraph, so we are picking out one in particular. The *order* of $H_e = (H^+, A)$ is $|H_e| = |V(H^+)|$.

We will usually work with H-extensions in a setting where H is an induced subgraph of an ambient graph G, and in this case a natural family of H-extensions can be obtained by considering neighbourhoods. Specifically, for $d \in \mathbb{N}$, the *(closed) d-ball* of an induced subgraph H of a graph G is

$$B_d(H, G) = G[\{v \in V(G) : d_G(v, H) \leq d\}],$$

the subgraph induced by the set of vertices of distance at most d from H including the vertices of H itself. It is useful to view the d-ball of H as the H-extension $(B_d(H, G), V(H))$. Two H-extensions (G_1, A_1) and (G_2, A_2) are *isomorphic* if there is a graph isomorphism $\varphi : G_1 \to G_2$ with $\varphi(A_1) = A_2$. Let $m_d(H_e, G)$ be the number of copies of H in G whose d-ball is isomorphic (as an H-extension) to H_e. In addition, we say that an H-extension (H^+, A) is a *sub-*H*-extension* of (H^{++}, B) if H^+ is an induced subgraph of H^{++} and $A = B$.

Our key counting result for extensions states that it is possible to reconstruct $m_d(H_e, G)$ from the ℓ-deck provided the d-balls of all copies of H are small enough to appear on a card.

Lemma 4.4. *Let $\ell, d \in \mathbb{N}$ and let G be a graph on at least $\ell + 1$ vertices. For any graph H on at most $\ell - 1$ vertices, at least one of the following conditions must hold:*

1. *There is a copy of H in G whose d-ball in G has at least ℓ vertices.*
2. *For any H-extension H_e, we can reconstruct $m_d(H_e, G)$ from the ℓ-deck of G.*

Proof. Let $\mathcal{H}$ denote the set of graphs H^+ such that $|V(H^+)| \leq \ell$, and there is a copy H' of H in H^+ such that all vertices of H^+ are at distance (in H^+) at most d from H'. These represent all possible d-neighbourhoods of H with at most ℓ vertices, and in particular, $\mathcal{H}$ contains all actual d-balls of copies of H in G with at most ℓ vertices. Note that it is not necessary (nor guaranteed) that all copies of H in H^+ satisfy the above distance condition, rather only that there is at least one such copy.

For any $H^+ \in \mathcal{H}$, we can reconstruct $n_{H^+}(G)$ from the ℓ-deck using Lemma 4.1. We can also recognise (from the ℓ-deck) whether Condition 1 of Lemma 4.4 holds. Suppose that it does not hold, meaning no copy of H has a d-ball containing more than $\ell - 1$ vertices. Then set

$$k = \max\{|V(H^+)| : H^+ \in \mathcal{H},\ n_{H^+}(G) > 0\}.$$

For a fixed $H^+ \in \mathcal{H}$ with $|V(H^+)| = k$, we observe that every copy H' of H for which $B_d(H', H^+) \cong H^+$ also satisfies $B_d(H', G) \cong H^+$ by the maximality of k and the definition of $\mathcal{H}$.

Let $\mathcal{H}_e$ denote the set of isomorphism classes of H-extensions (H^+, A) with $H^+ \in \mathcal{H}$. By the preceding observation, if $H_e = (H^+, A) \in \mathcal{H}_e$ with $|H^+| = k$, then

$$m_d(H_e, G) = n_{H^+}(G) m_d(H_e, H^+), \tag{4.1.2}$$

the number of copies of H^+ in G times the number of copies of H in H^+ whose d-ball (within H^+) is isomorphic to H_e (as H-extensions). Both of these quantities are reconstructible from the ℓ-deck, so we are done in this case.

If $|V(H^+)| < k$, the d-ball of H may be strictly larger than H^+ and the formula (4.1.2) does not apply. This can be corrected by subtracting the number of $H \subseteq H^+$ for which H^+ is not the d-neighbourhood of that copy of H in G. To count these, we select each 'maximal' d-neighbourhood in turn, and subtract one from the relevant count for each strictly smaller H^+ that it contains. Any leftover H^+ that have not been accounted for must then be maximal.

For $H'_e \in \mathcal{H}_e$ distinct from H_e, let $n(H_e, H'_e)$ give the number of sub-H-extensions of H'_e isomorphic to H_e. We claim that

$$m_d(H_e, G) = n_{H^+}(G) m_d(H_e, H^+) - \sum_{\substack{H'_e \in \mathcal{H}_e \\ |H'_e| > |H_e|}} n(H_e, H'_e) m_d(H'_e, G).$$

Note that when $|H_e| = k$, this formula agrees with (4.1.2). The terms $m_d(H_e, H^+)$ $n(H_e, H'_e)$ and the domain of the summation are already known to us, and we can reconstruct $n_{H^+}(G)$ for all $H^+ \in \mathcal{H}$ using Kelly's Lemma. Moreover, we may assume that we have reconstructed the terms $m_d(H'_e, H^+)$ for $|H'_e| > |H_e|$ by induction with base case $|H_e| = k$, so verifying the formula will complete the proof.

The first term of the formula $n_{H^+}(G)m_d(H_e, H^+)$ counts the number of pairs $(A, B) \subseteq V(G) \times V(G)$ such that

- $G[B]$ is a copy of H^+ and corresponds to one object counted by n_{H^+},
- $A \subseteq B$,
- $G[A]$ is a copy of H and is counted by $m_d(H_e, H^+)$ for a fixed copy of H^+ (determined by B),
- B is a subset of the d-ball around A (i.e. $B \subseteq B_d(G[A], G)$) due to the definition of H-extension.

Compared to $m_d(H_e, G)$, we are overcounting whenever $B \subsetneq B_d(G[A], G)$. Thus, it just remains to verify that there are $\sum_{|H'_e|>|H_e|} n(H_e, H'_e)m_d(H'_e, G)$ pairs for which $B \neq B_d(G[A], G)$. To see this, we can think of the correction term as counting triples (A, B, C) with $A \subseteq B \subsetneq C \subseteq V(G)$ such that

- $G[A]$ is a copy of H,
- $G[B]$ is a copy of H^+
- $G[C] \cong B_d(G[A], G)$.

The first two conditions follow from the definition when H_e is a sub-H-extension of H'_e, and the latter follows from the definition of $m_d(H'_e, G)$. The fact that $B \subsetneq C$ follows from the strict inequality $|H'_e| > |H_e|$. Each pair (A, B) with $B \neq B_d(G[A], G)$ is in a unique such triple, namely with $C = V(B_d(G[A], G))$; if $B = B_d(G[A], G)$ then no suitable C with $B \subsetneq C$ can be found. □

As a first application, by setting $d = 1$ and considering the H-extension $(H, V(H))$ in Lemma 4.4, one can count the number of components isomorphic to H.

Corollary 4.5. *Let H and G be graphs with $|V(H)| \leq \ell - 1$ and $n = |V(G)|$. If there is no copy of H in G for which $|B_1(H, G)| \geq \ell$, then we can reconstruct the number of components of G isomorphic to H from $\mathcal{D}_\ell(G)$.*

4.2 Reconstructing trees from small cards

The counting tools of the last section enable us to show that trees are reconstructible even from the deck consisting of cards where a linear proportion of the vertices have been deleted. We recall this statement below, which comes from joint work with Groenland, Johnston and Scott.

Theorem 1.1. *Any n-vertex tree T can be reconstructed from $\mathcal{D}_\ell(T)$ whenever $\ell > \frac{8n}{9} + \frac{4}{9}\sqrt{8n+5} + 1$.*

The proof of Theorem 1.1 is organised into three parts. In Section 4.2.1, we address the recognition problem (Theorem 4.8), which also serves as another application of extension counting. The other two parts contain the proof of reconstruction, which is split into cases depending on whether or not the tree T contains a path that is long relative to the size of the graph n and the size of each card ℓ.

Let the *length* of a path P be the number of edges in P, or equivalently $|V(P)|-1$. The *diameter* of a graph G is the maximum distance between two vertices in G, and for a tree T this is the same as the length of a longest path. We will refer to the aforementioned cases as the 'high diameter' and 'low diameter' cases.

Assuming we have already determined that T is a tree, the high diameter case is handled by the following lemma which we prove in Section 4.2.2. This main ingredient here is Lemma 4.4.

Lemma 4.6. *Let $\ell, k \in [n]$ with $k > 4\sqrt{\ell} + 2(n-\ell)$. If T is an n-vertex tree with diameter $k-1$, then T can be reconstructed from its ℓ-deck provided $\ell \geq \frac{2n}{3} + \frac{4}{9}\sqrt{6n+7} + \frac{11}{9}$.*

If T has low diameter, then we instead leverage the fact that we can identify longest paths on individual cards. This provides an anchor that we can use to reconstruct pieces of the graph that we obtain using Lemma 4.3. Following this outline, the next lemma is proved in Section 4.2.3.

Lemma 4.7. *Suppose that T is an n-vertex tree with diameter $k-1$. Then T can be reconstructed from its ℓ-deck for any $\ell \in [n]$ such that $n - \ell < \frac{n-3k+1}{3}$ if k is odd or $n - \ell < \frac{n-3k-1}{3}$ if k is even.*

With these lemmas in hand, the proof of Theorem 1.1 then amounts to verifying that the assumptions are sufficient for recognition, and that our definitions of high and low diameter together cover the full range.

Proof of Theorem 1.1. Let k be the number of vertices in the longest path in T. The conditions on ℓ and n imply that $\ell \geq \frac{2n}{3} + \frac{4}{9}\sqrt{6n+7} + \frac{11}{9}$. This allows us to recognise that T is a tree by Theorem 4.8, and moreover that T is reconstructible by Lemma 4.6 when $k > 4\sqrt{\ell} + 2(n-\ell)$. For the remaining k, we show that the conditions of Lemma 4.7 are then satisfied. It suffices to verify that $n - \ell < \frac{n-3k-1}{3}$. The right hand side is decreasing in k, and now $k \leq 4\sqrt{\ell} + 2(n-\ell)$, so Lemma 4.7 applies provided

$$n - \ell < \frac{n - 12\sqrt{\ell} - 6(n-\ell) - 1}{3}$$

which is equivalent to our assumed condition

$$\ell > \frac{8n}{9} + \frac{4}{9}\sqrt{8n+5} + 1. \qquad \square$$

4.2.1 Tree recognition

This section proves that we can recognise trees as an application of the extension-counting result established in Section 4.1.3.

Theorem 4.8. *For $\ell \geq (2n+4)/3$, the class of trees on n vertices is recognisable from the ℓ-deck.*

Proof. Let G be a graph and suppose we are given $\mathcal{D}_\ell(G)$. By Kelly's Lemma (Lemma 4.1), we can reconstruct the number of edges m provided $\ell \geq 2$. We may suppose that $m = n - 1$, otherwise we can already conclude that G is not a tree. It suffices to show that we can determine whether G contains a cycle, or equivalently to determine whether G is connected.

If G has a cycle of length at most ℓ, then the entire cycle will appear on a card and we can conclude that G is not a tree. We may therefore assume that every cycle in G has length greater than ℓ. By inspecting the ℓ-deck, we can easily determine whether all components of G have order at most $\ell - 1$, and if so conclude that G is not a tree. We may therefore assume that the largest component in G, say A, has at least $\ell \geq (2n+4)/3$ vertices, and all other components have at most $\ell - 1$ vertices.

Let $d = \lceil \ell - n/2 - 1 \rceil$. For a vertex $x \in V(G)$, denote the d-ball around x by $B_d(x) := B_d(\{x\}, G)$. Using Lemma 4.4 with H being the graph consisting of a single vertex, we find that either there is an $x \in V(G)$ with d-ball of order at least ℓ or we can reconstruct the collection of d-balls (with 'distinguished' centres).

Suppose firstly that there exists $x \in V(G)$ such that $|B_d(x)| \geq \ell$. We claim then that G is a tree. Assume towards a contradiction that there is a cycle in G. Since every cycle has length at least $\ell+1$, any cycle in G must be contained in the largest component A. Let C be a shortest cycle in A. Note that $x \in A$, since otherwise the d-ball around x cannot have ℓ vertices (the smaller components have order at most $\ell-1$). If $|B_d(x) \cap C| \leq 2d+1$, then

$$|B_d(x)| \leq n - |C \setminus B_d(x)| \leq n - (\ell+1) + (2d+1) \leq \ell - 1$$

by our choice of d. Thus, $B_d(x) \cap C$ contains at least $2d+2$ vertices. We can choose two vertices $c_1, c_2 \in B_d(x) \cap C$ such that there is a subpath C' of C between c_1 and c_2 that does not contain any other vertices of $B_d(x) \cap C$, allowing the possibility that C' is a single edge. Let C'' be the other path from c_1 to c_2 in C. This must contain at least $2d$ other vertices of $B_d(x) \cap C$, so C'' is a path of length at least $2d+1$. However, there is also a path P from c_1 to c_2 in the d-ball around x of length at most $2d$, and this intersects C' only at the endpoints c_1 and c_2. Replacing the path C'' with the path P forms a cycle which is strictly shorter than C, giving a contradiction. Hence, G cannot have any cycles and must be a tree.

We may now assume that we can reconstruct the collection of d-balls and will show how to recognise whether the graph is connected in this case. In any component of order at most $n-\ell$, there must be some vertex x such that the distance from x to any vertex in the same component is at most $(n-\ell)/2$. By our choice of ℓ and d,

$$\frac{n-\ell}{2} \leq \ell - \frac{n}{2} - 2 \leq d - 1.$$

Thus, if there is a component of order at most $n-\ell$ (which happens if and only if G is not a tree), then there must be a d-ball with radius at most $d-1$. Conversely, if we discover such a d-ball, then we know that the graph is disconnected since the d-ball must form a component due to its radius, yet has at most $\ell-1$ vertices. Hence, G is a tree if and only if all d-balls have radius d. This shows that we can recognise connectedness and completes the proof. □

4.2.2 High diameter

The main result in this section is Lemma 4.6, which states that a tree T is reconstructible from its ℓ-deck provided it contains a sufficiently long path.

Consider the collection of components of $T-e$ as e varies over all edges, viewed as induced subgraphs. Each element R in this collection has a natural counterpart $R^c := T[V(G) - V(R)]$. Our goal is to recognise a pair (R, R^c) for which we can also deduce which vertex in each subgraph was incident to e (assuming this is the pair of components in $T - e$). With this information, one can easily obtain T by gluing via one extra edge between the 'indicated vertices'.

Instead of working just with copies of R (and R^c), we are specifically interested in copies which connect to the rest of the graph by a single edge. For a graph H let a *leaf H-extension* be a pair $H_e = (H^+, A)$ where

- H^+ is obtained by adding a single vertex connected by a single edge to a vertex of H, and
- $A \subset V(H^+)$ is such that $H^+[A] \cong H$.

This is a special case of the extensions defined in Section 4.1.3. Note that the 1-ball of our special component R of T gives a leaf R-extension, but there may be multiple (non-isomorphic) leaf R-extensions in T.

The extra edge in a leaf extension indicates where to glue, so we would be done if we could identify two leaf extensions $C = (C^+, V_C)$ and $D = (D^+, V_D)$ for which the vertex set of G is the disjoint union of $V(C)$ and $V(D)$. We demonstrate in Lemma 4.9 that this can be done using counts of the relevant leaf extensions in cases where these can be obtained by Lemma 4.4. The final step to proving Lemma 4.6 consists of showing that there exist suitable R and R^c in T for which Lemma 4.9 applies, and it is only then that we use the assumption of high diameter.

Lemma 4.9. *Let G be a connected graph with a bridge e, and $R, R^c \subseteq G$ be the connected components of $G - e$. If G has no induced subgraph H isomorphic to R or R^c with $|V(B_1(H, G))| \geq \ell$, then G is reconstructible from $\mathcal{D}_\ell(G)$.*

Proof. Given any connected graph H on at most $\ell - 1$ vertices and $\mathcal{D}_\ell(G)$, we can check whether there is a copy of H in G with $|V(B_1(H, G))| \geq \ell$. Suppose H is a connected graph for which no such copy exists. For every leaf H-extension H_e of H, apply Lemma 4.4 to reconstruct $m_1(H_e, G)$. Recall that this is the number of copies of H in G whose 1-ball in G is obtained by adding an edge at a specified vertex. Let $\mathcal{H}$ denote the set of connected graphs H for which we have now reconstructed that $m_1(H_e, G) > 0$ for at least one leaf H-extension H_e.

We may assume that $|V(R^c)| \geq |V(R)|$. Note that $\mathcal{H}$ is reconstructible from $\mathcal{D}_\ell(G)$, and by assumption R and R^c are elements of $\mathcal{H}$. Consider all pairs (C, D) of elements in $\mathcal{H}$ for which $|V(C)| + |V(D)| = n$ and $|V(C)| \leq |V(D)|$. Given a leaf C-extension $C_e = (C^+, V_C)$ with $m_1(C_e, G) > 0$ and also a leaf D-extension $D_e = (D^+, V_D)$ with $m_1(D_e, G) > 0$, we will show that it is possible to determine whether $C \cong D^c$ and gluing C_e and D_e on their additional edge gives G. This will complete the proof that G is reconstructible from its ℓ-deck since the existence of such a good pair is guaranteed by the fact that (R, R^c) is necessarily among the pairs considered, and suitable leaf extensions R_e and R^c_e exist.

Fix any pair (C, D) together with leaf extensions C_e and D_e. Since D_e is a leaf extension, we know that G is obtained by adding a single edge between D and D^c where $|V(D^c)| = |V(C)|$. Let D^c_e be the other extension in this gluing. The fact that G is connected implies that D and D^c are connected. In particular, if $D' \subset D$ is a non-spanning subgraph, then $B_1(D')$ contains a vertex of $V(D) - V(D')$. The same holds if we replace D with D^c. This implies that any copy of C containing vertices from both D^c and D satisfies $|V(B_1(C))| \geq |V(C)| + 2$, and so does not contribute to $m_1(C_e, G)$. Since $|V(C)| = |V(D^c)|$, a copy of C cannot cover some of D^c and none of D. Hence, the only way a copy of C which contributes to $m_1(C_e, G)$ can contain vertices from D^c is if it covers all of D^c, and since $|V(C)| = |V(D^c)|$, this implies that $C \cong D^c$. There is therefore at most one leaf C-extension for which the copy of C contains vertices from D^c and it exists if and only if C_e and D_e glue on the indicated edge to give G.

Any other contributing copy of C must be contained in D, and the corresponding copy of C^+ is contained in D^+ (possibly using the extra edge). Now let $N(C_e, D^+)$ be the number of leaf C-extensions $(C^{+\prime}, V'_C)$ in D^+ isomorphic to C_e with $V'_C \subseteq V_D$, which can be calculated directly for our fixed C_e and D_e. By the preceding discussion, either $m_1(C_e, G) = N(C_e, D^+)$ or $m_1(C_e, G) = N(C_e, D^+) + 1$. In the latter case, the additional leaf C-extension contains all vertices of D^c (and hence $C = D^c$), and exists if and only if $C_e \cong D^c_e$. Since $m_1(C_e, G)$ and $N(C_e, D^+)$ are known, we can hence recognise whether C_e and D_e glue on the indicated edge to give G. □

We are now ready to prove the main result in this section. It remains to show that Lemma 4.9 applies to trees with large enough diameter (depending on both n and ℓ). For this, we need to find subtrees R and S for which T has no copy of R or S with a large 1-ball; informally, we would like T to not be too star-like, and this is the case when T has a long path.

Proof of Lemma 4.6. Let $k, \ell \in [n]$ with $k > 4\sqrt{\ell} + 2(n-\ell)$ and $\ell \geq \frac{2n}{3} + \frac{4}{9}\sqrt{6n+7} + \frac{11}{9}$. We assume that we have already determined that T is a tree, and note that we can recognise from the ℓ-deck whether a longest path contains more than $4\sqrt{\ell} + 2(n - \ell)$ vertices. Indeed, our choice of ℓ guarantees that $\ell \geq 4\sqrt{\ell} + 2(n - \ell) + 1 > \lceil 4\sqrt{\ell} + 2(n - \ell) \rceil$.

Fix a longest path in T with k vertices. Create two rooted subtrees R and S by removing the central edge of the path if k is even, or one of the two central edges if k is odd (and rooting the subtrees at the vertex which had an incident edge removed). By Lemma 4.9, if T has no induced subgraph H isomorphic to R or S with $|V(B_1(H,T))| \geq \ell$, then T is reconstructible from $\mathcal{D}_\ell(T)$. We assume, in order to derive a contradiction, that T contains a copy S' of S with $|V(B_1(S',T))| \geq \ell$. The same argument can then be used to show that T also has no copy R' of R with $|V(B_1(R',T))| \geq \ell$.

Set $r = n - \ell$. Let $\varphi : S \to S'$ be an isomorphism, and let P_0 be a path in R containing at least $(k-1)/2$ vertices which starts at the root of R. Consider the intersection of S' with the path P_0. Since $V(S') \neq V(S)$, this intersection must be non-empty, and it must be connected since both T and S are trees, so S' and P_0 intersect on a subpath Q_0. Moreover, the intersection of $B_1(S',T)$ and P_0 must also be a path with at most $|V(Q_0)| + 2$ vertices. This means that there are at least $|V(P_0)| - |V(Q_0)| - 2$ vertices on P_0 which are not in $B_1(S',T)$. By assumption we have $|V(B_1(S',T))| \geq n - r$ meaning T has at most r vertices not in $B_1(S',T)$, so it follows that $|V(Q_0)| \geq |V(P_0)| - r - 2$.

Now let P_1 be the path $\varphi^{-1}(V(Q_0))$ in S and note that P_1 is vertex-disjoint from P_0 as P_0 is contained in R. Define Q_1 to be the intersection of S' with P_1, which is again a path. Furthermore, the intersection of $B_1(S',T)$ and P_1 is also a path, this time with at most $|V(Q_1)|+2$ vertices. The number of vertices of P_1 and P_0 which are not in $B_1(S',T)$ is at least $|V(P_0)|+|V(P_1)|-|V(Q_0)|-|V(Q_1)|-4$, which gives the inequality $|V(Q_0)| + |V(Q_1)| \geq |V(P_0)| + |V(P_1)| - r - 4$. Since $|V(Q_0)| = |V(P_1)|$, this becomes $|V(Q_1)| \geq |V(P_0)| - r - 4$.

We now continue iteratively to build two sequences of paths: given P_i, define $Q_i := S' \cap P_i$ which is a subpath of P_i and then let $P_{i+1} := \varphi^{-1}(V(Q_i))$ which is a path in S. Note that P_{i+1} is disjoint from $P_0, \ldots, P_i$. Since P_0 is contained in R, P_{i+1} cannot intersect P_0. If P_{i+1} intersects a path P_j, then Q_i must intersect Q_{j-1} which in turn implies P_i intersects P_{j-1}. Hence, the paths are disjoint by induction. By the finiteness of T, we must eventually reach a j such that $|V(Q_{j-1})| = |V(P_j)| = 0$. At this point, we have disjoint paths $P_1, \ldots, P_j$ in S that satisfy $|V(P_i)| = |V(Q_{i-1})| \geq |V(P_0)| - r - 2i$ for all $i = 1, \ldots, j$. In particular, setting $i = j$ to use the fact that $|V(P_j)| = 0$ shows that $j \geq (|V(P_0)| - r)/2$.

Let us use $|P_0|$ as shorthand for $|V(P_0)|$. We may then calculate

$$\begin{aligned}
|V(S)| &\geq |V(P_1)| + \cdots + |V(P_j)| \\
&\geq \sum_{i=1}^{\lfloor (|P_0|-r)/2 \rfloor} (|P_0| - r - 2i) \\
&= (|P_0| - r) \left\lfloor \frac{|P_0| - r}{2} \right\rfloor - 2 \binom{\lfloor (|P_0| - r)/2 \rfloor + 1}{2} \\
&= \left\lfloor \frac{|P_0| - r}{2} \right\rfloor \left\lceil \frac{|P_0| - r - 2}{2} \right\rceil \\
&\geq \frac{(|P_0| - r)(|P_0| - r - 2)}{4}.
\end{aligned}$$

Since $|V(S)| \leq n - |V(P_0)|$, we must have $|V(P_0)| \leq \sqrt{4n - 4r + 1} + r - 1$ and $k \leq 2|V(P_0)| + 1 \leq 2\sqrt{4n - 4r + 1} + 2r - 1$. Finally, note that $2\sqrt{x+1} - 1 \leq 2\sqrt{x}$ for all $x \geq 1$ to find $k \leq 4\sqrt{\ell} + 2r$, a contradiction. Repeating the argument for R, by Lemma 4.9 we can reconstruct T from $\mathcal{D}_\ell(T)$. □

4.2.3 Low diameter

Throughout this section, we will assume that T has n vertices and $r := n - \ell < \frac{n-3k+1}{3}$, where k is the number of vertices in a longest path in T. This means that $k + 1 \leq \ell$ so we can reconstruct k from the ℓ-deck.

If k is odd, the *centre* of T is the vertex in the middle of each longest path, and if k is even, the centre consists of the two middle vertices. The centre is unique, so in particular it does not depend on the choice of longest path. Let us assume for now that T has a unique central vertex, leaving the even case to be handled later by subdividing the central edge. Given a vertex $u \in T$ with neighbours $v_1, v_2, \ldots, v_a$, let the *branches* at u be the rooted subtrees $b_1, b_2, \ldots, b_a$ where b_i is the component of $T - u$ that contains v_i, rooted at v_i.

An *end-rooted path* is a path rooted at an endvertex. In this section, all longest paths P_k will be rooted at the central vertex c, and are hence not end-rooted, whilst all of the shorter paths mentioned will be end-rooted. Given two rooted trees T_1 and T_2 with roots u and v respectively, let $T_1 \frown T_2$ denote the (unrooted) tree given by adding an edge between u and v (see Figure 4.1).

By restricting our attention to the cards that have diameter $k - 1$, we may assume that we can always identify the centre of the graph. Our basic strategy is to reconstruct the branches at the centre separately, knowing that we can later join them together using the centre as a common point of reference. This can be done

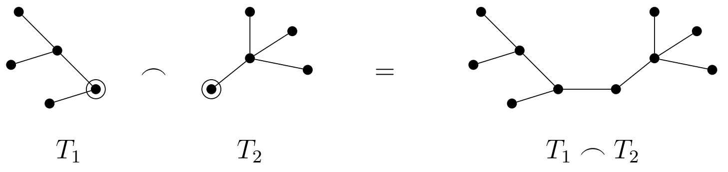


Figure 4.1: An example of the tree grafting operation $T_1 \frown T_2$.

via a counting argument when all branches at the centre have at most $\ell - k$ vertices, but when one branch is 'heavy' and contains many (at least $\ell - k$) of the vertices a slightly more finicky argument is required to reconstruct the heavy branch which cannot be seen on a single card. It is possible to recognise these cases from the ℓ-deck, a statement which we prove as part of Lemma 4.7. As such, we simply state Lemma 4.10 and Lemma 4.11 with the assumption that we know *a priori* whether or not all branches have less than $\ell - k$ vertices.

Lemma 4.10. *Let T be a tree with even diameter $k - 1$. Suppose it is known that every branch from the centre has fewer than $\ell - k$ vertices. Then T is reconstructible from the subset of the ℓ-deck consisting only of cards that contain a copy of P_k.*

Proof. Let c be the central vertex of T and $B = \{b_1, \dots, b_a\}$ be the branches at c that we wish to reconstruct. We first reconstruct all branches that are not end-rooted paths. For any fixed b which is a rooted tree but not an end-rooted path, we will use Lemma 4.3 to count each branch at c isomorphic to b once for every P_k in T. Dividing this number, denoted N_b, by the number $n_{P_k}(T)$ of copies of P_k in T then tells us the multiplicity of b in T (which may be zero). Note that $n_{P_k}(T)$ can be determined by Kelly's Lemma as $k < \ell$ (recall that we are assuming $n - \ell < \frac{n-3k+1}{3}$), so it suffices to reconstruct N_b.

Fix b to be any rooted tree that is not an end-rooted path. We will actually determine N_b in two parts. Let π_b be the number of pairs consisting of one copy b' of b that is a branch at c, and one copy P'_k of a longest path that is disjoint from b'. Similarly, let τ_b count pairs (b', P'_k) where the copy P'_k intersects b'. It is clear that $N_b = \pi_b + \tau_b$.

We begin with π_b. Let $\mathcal{G}$ be the family of all n-vertex trees with diameter $k - 1$ and where all branches from the centre have fewer than $\ell - k$ vertices, and let $\mathcal{F}$ be the family of graphs of the form $P_k \frown S$, where S is a rooted tree that is not an end-rooted path and P_k is rooted at its central vertex (see Figure 4.2). Fix $G \in \mathcal{G}$ and consider some $F \in \mathcal{F}$. If $F' = P'_k \frown S'$ is a copy of F in G, then it is contained in a unique maximal $\mathcal{F}$-subgraph, namely P'_k together with the unique branch b'

containing S'. Also, since every branch has fewer than $\ell - k$ vertices by assumption, these maximal elements have fewer than ℓ vertices. Thus, by Lemma 4.3 we can reconstruct the number of $\mathcal{F}$-maximal copies of each F in G from $\mathcal{D}_\ell(G)$. If this is non-zero for $F = P_k \frown S$ then G has a branch isomorphic to S (but the converse may not hold).

Now let $F = P_k \frown b$. Since $T \in \mathcal{G}$ and $F \in \mathcal{F}$, we may reconstruct the number of $\mathcal{F}$-maximal copies of F in T as above. This is precisely π_b. To see this, consider a particular copy b' of b that occurs as a branch and observe that F occurs as a maximal $\mathcal{F}$-subgraph with this b' as the copy of b once for every longest path in the tree which avoids b' (see Figure 4.3).

There is a similar argument to determine τ_b. Keeping $\mathcal{G}$ as before, let $\mathcal{F}'$ be the family of graphs of the form $P_{(k-1)/2+1} \frown S$ where S is a rooted tree which contains an end-rooted $P_{(k-1)/2}$, but is not itself an end-rooted path. Again, an element $F = P_{(k-1)/2+1} \frown S$ is $\mathcal{F}'$-maximal when S is an entire branch, and for any $G \in \mathcal{G}$ and $F \in \mathcal{F}'$ we can reconstruct the number of $\mathcal{F}'$-maximal copies of each F in G by Lemma 4.3. This time there is at least one $\mathcal{F}'$-maximal copy of $F = P_{(k-1)/2+1} \frown S$ if and only if G has a branch isomorphic to S (although we do not need to use both directions explicitly).

Let $m_{F'}$ be the number of $\mathcal{F}'$-maximal copies of $F' = P_{(k-1)/2+1} \frown b$ in T, which we can reconstruct as argued above. A particular copy b' of b that occurs as a branch contributes one to $m_{F'}$ for each copy of $P_{(k-1)/2+1}$ that starts at the central vertex c and is disjoint from b'. Thus, letting $n_{P^\bullet}(b)$ be the number of end-rooted copies of $P_{(k-1)/2+1}$ in b' with roots coinciding (this is the same for any copy of b and does not depend on the deck), one can construct all of the copies of longest paths that intersect b' by gluing together one $P_{(k-1)/2+1}$ from inside b' and one that is disjoint from it. Doing so for every copy of b shows that we can reconstruct $\tau_b = m_{F'} \cdot n_{P^\bullet}(b)$. The number of copies of b that occur as a branch can then be reconstructed as

$$\frac{N_b}{n_{P_k}(T)} = \frac{\pi_b + \tau_b}{n_{P_k}(T)}.$$

It remains to determine the number of branches isomorphic to an end-rooted path P_i, which we do using the fact that we know all of the other branches not of this form. Starting with $j = (k-1)/2$, this being the maximum possible length of a path branch, we compare the number of copies of $P_{(k-1)/2+j+1}$ in T to the number of copies in the graph $\widetilde{T}$ obtained by gluing all of the known branches at a single vertex c. The former count can be obtained by Kelly's Lemma, and the latter directly from inspecting $\widetilde{T}$. If there are more copies in T than in the current $\widetilde{T}$, then there must be

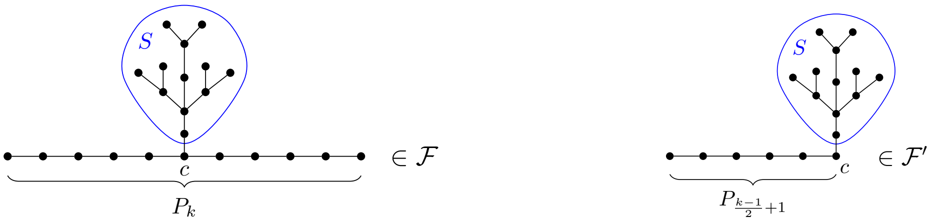


Figure 4.2: Elements of $\mathcal{F}$ and $\mathcal{F}'$.

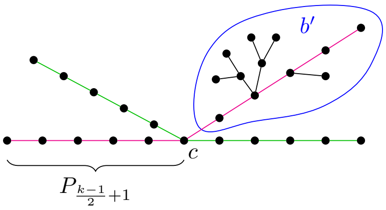


Figure 4.3: A longest path that avoids b' (green) contributes one to the number of times $P_k \frown b$ occurs as a maximal subgraph. A longest path that uses b' (magenta) consists of a $P_{(k-1)/2+1}$ outside b' and a $P_{(k-1)/2}$ inside.

at least one more end-rooted P_j as a branch so we add one copy to our list of known branches. We then repeat this step with j fixed but $\widetilde{T}$ updated to include this new path branch. If the counts match, meaning all copies of $P_{(k-1)/2+j+1}$ in G are already present in $\widetilde{T}$, then reduce j by 1 and continue iteratively until $j = 0$. Note that it is important that we handle the different path lengths in this order. At this point, we have reconstructed all branches and the final $\widetilde{T}$ is exactly T. $\square$

We now consider the case where one of the branches contains a lot of vertices. Since this branch contains so many vertices, we can find a card showing all the other branches in their entirety. This reduces the problem to reconstructing the large branch. We will move the "centre" one step inside the branch and continue doing this until no branch is too big. At this point we can apply the proof of the previous lemma (with some minor modifications) to reconstruct the branches off the new "centre". The condition that T has small diameter ensures that we do not have to take too many steps away from the true centre in this process. Recall that $r = n - \ell < \frac{n-3k+1}{3}$ by assumption.

Lemma 4.11. *Let T be a tree of even diameter $k-1$. Suppose it is known that T has a branch with at least $\ell - k$ vertices. Then T is reconstructible from the subset of the ℓ-deck consisting only of cards which contain a copy of P_k.*

Proof. Let c be the central vertex of T. First, note that $\ell - k = n - r - k > 2n/3$, so there can be at most one branch with at least $\ell - k$ vertices. Call this the *heavy* *branch*. The total number of vertices in the remaining branches is at most $k+r$. Taking a connected card in which the maximum number of vertices in any branch is as small as possible among those which contain a copy of P_k, we see that the heavy branch must still have at least $n-2r-k$ vertices visible on the card whilst each of the other branches have at most $k+r$ vertices. Thus, on any card containing P_k we can directly identify which is the heavy branch and the entirety of all of the smaller branches.

Set $c_0 := c$. To reconstruct the heavy branch, we construct a sequence of vertices $c_0, c_1, c_2, \ldots$ to act as new "centres" until the branches at some c_j are all small enough for us to apply Lemma 4.3. Let c_1 be the vertex in the heavy branch adjacent to c_0. If the branches at c_1, which we call *1-branches*, all have less than $\ell-k-1$ vertices, then we terminate with $j=1$. Otherwise, take a connected card containing a copy of P_k in which the heaviest 1-branch is as small as possible. The heaviest branch has at least $\ell-k-r-1$ vertices, which is greater than the maximum number of vertices in any other branch (now $k+r+1$). This ensures that the heaviest 1-branch is contained in the original heavy branch, and we can identify all smaller 1-branches. Now set c_2 to be the vertex in the heaviest 1-branch adjacent to c_1 and repeat the argument. In the ith step, we terminate if every branch from c_i has weight less than $\ell-k-i$, and otherwise completely determine all but the heaviest i-branch and proceed by taking a step into this i-branch. To do this we only require $\ell-k-r-i > k+r+i$, which holds for $i \leq (k-1)/2$ by our choice of r. Suppose the process terminates at the jth step. Since the longest path in T contains k vertices, the longest path in the heavy branch with one endvertex at the root contains at most $(k+1)/2$ vertices and hence $j \leq (k-1)/2$.

The remainder of the argument closely follows the proof of Lemma 4.10. Let $\mathcal{G}$ be the family of trees of n-vertex trees with diameter $k-1$, and $\mathcal{F}$ be the family of graphs that can be constructed as follows. Let $i \in \{0, \ldots, j-1\}$, let $v_1, \ldots, v_k$ be the vertices in a P_k and let $u_1, \ldots, u_{j-i}$ be the vertices in a (disjoint) P_{j-i}. A graph in $\mathcal{F}$ is formed by adding an edge from u_1 to $v_{\frac{k+1}{2}+i}$, and then attaching a rooted tree S which is not an end-rooted path to the vertex u_{j-i}. The condition that the attached tree is not a path ensures that it is easy to identify P_k and the added tree in any $\mathcal{F}$-graph. An example is given in Figure 4.4.

Each $\mathcal{F}$-subgraph of $G \in \mathcal{G}$ is contained in a unique maximal $\mathcal{F}$-subgraph, given by extending the tree attachment to the whole of the relevant branch at u_{j-i}. Applying Lemma 4.3 allows us to determine the number of occurrences of each maximal $\mathcal{F}$-subgraph, as we did in the proof of Lemma 4.10.

Figure 4.4: Potential elements of $\mathcal{F}$ along with their 'moving centres'.

At this point, each branch b' has contributed one to the relevant count for each copy of P_k which does not use b', so we again need to determine the number of P_k which use b'. This can be done using an identical argument to that in Lemma 4.10 except replacing c with c_j, replacing $P_{(k-1)/2+1} \frown S$ with $P_{(k-1)/2+j+1} \frown S$ and suitably adjusting S.

We have now identified the total number of branches of each isomorphism class (except those which are end-rooted paths), although we do not know they are all branches at c_j. However, we have already reconstructed all of the tree except for the branches at c_j, so we can subtract the counts of all the appropriate branches not at c_j from the total: the remainder must be attached at c_j.

Finally, the end-rooted paths attached at c_j can be reconstructed using the argument from the end proof of Lemma 4.10. □

Proof of Lemma 4.7. Let T be an n-vertex tree with diameter $k-1$, and hence k vertices in any longest path. First assume that k is odd, so there is a central vertex c of T. If one of the branches at c has at least $\ell - k$ vertices, then there must be a card containing a longest path with a branch of at least $\ell - k$ vertices (the branch and the path need not be disjoint, but their union contains at most ℓ vertices). Thus we can recognise from the ℓ-deck whether there is a branch with at least $\ell - k$ vertices. If there is no such branch then we are done by Lemma 4.10, and if there is one, we are done by Lemma 4.11.

When k is even, we reduce to the odd case by instead considering the tree T' obtained by subdividing the central edge. Since we can recognise the central vertex in T' and hence recover T, it suffices to reconstruct T'. Moreover, we can obtain the cards in the $(\ell+1)$-deck of T' that contain a longest path by taking the cards in the ℓ-deck of T that contain a longest path and subdividing the central edge. These are the only cards required by Lemma 4.10 and Lemma 4.11. □

4.3 Approximate counts

We now move to the setting of reconstruction from an incomplete $(n-1)$-deck where we do not have control over which cards are missing. The arguments in this variant have a different flavour to those of the preceding section because we no longer have Kelly's Lemma, nor any of the tools derived from it. Instead, we employ a general strategy of bounding away errors on counts that arises from the missing information. While we previously counted different structures and subgraphs, the theory in this case is much less developed so we focus on the smaller task of counting vertices with particular constraints.

This section is based on joint work with Groenland, Johnston, Kupavskii, Meeks and Scott. Throughout, we shall let G be the original graph on n vertices to be reconstructed. We write $G_i = G - v_i$ and order the vertices as $v_1, \ldots, v_n$ such that $G_1, \ldots, G_{n-k}$ are the cards that have been given from the deck $\mathcal{D}(G)$. For any graph H, let $d_t(H)$ be the number of vertices of degree t in H. That is, $d_t(H) = |\{v \in V(H) : d_H(v) = t\}|$. Similarly, let $d_{<t}(H) = |\{v \in V(H) : d_H(v) < t\}|$ be the number of vertices of degree less than t in H. We write $[a, b]$ to denote the set of integers between a and b inclusive.

Our goal is to prove that the number of edges, number of cliques, and degree sequence of graphs with bounded average degree can be reconstructed from the $(n-1)$-deck even with a linear number of missing cards. The added assumption of having bounded average degree is critical to overcoming the error resulting from the missing cards. The following simple fact gives an idea of how it helps.

Observation 4.12. *Every graph with average degree bounded above by $d \in \mathbb{N}$ has at least $\frac{n}{2}$ vertices of degree at most $2d$ and at least $\frac{n}{d+1}$ vertices of degree at most d.*

Given enough cards, this will allow us to assume that the card with the most edges corresponds to a low-degree vertex, and therefore has small "error". Another important feature is that the property of having small average degree is recognisable from the partial deck in the following sense.

Lemma 4.13. *Let G be a graph on $n \geq 8$ vertices with average degree d^*. From any deck of G missing $k \leq \frac{n}{4}$ cards, we can reconstruct a quantity $\widetilde{d}$ that satisfies $0 \leq \widetilde{d} - d^* < 1$.*

Proof. By [68, Lemma 2.1], we can calculate from the cards $G_1, \ldots, G_{n-k}$ an estimate $\widetilde{m}$ for the number of edges m that satisfies

$$0 \leq \widetilde{m} - m \leq \frac{k(n-1)}{n-2-k}.$$

Since $k \leq \frac{1}{4}n$, we find

$$\frac{k(n-1)}{n-2-k} < \frac{n^2}{3n-8} \leq \frac{n}{2}$$

for $n \geq 8$. Hence $\widetilde{d} = \frac{2\widetilde{m}}{n}$ satisfies the claimed inequality. □

Lemma 4.13 allows us to assume that the average degree d^* is known up to a potential error of 1. This approximation will be used to prove that we can reconstruct the number of edges (Theorem 1.5 below), but once we have that result and enough cards for it to apply, the average degree can then be computed exactly since the number of vertices is known. Thus, by the time we come to reconstructing the degree sequence afterwards, we will be able to assume that d^* is known exactly.

4.3.1 Edge and clique counts

We now begin in earnest by reconstructing the number of edges.

Theorem 1.5. *Let G be a graph on $n \geq 3$ vertices with average degree at most d for an unknown $d \in \mathbb{N}$. Then the number of edges in G can be reconstructed from any deck missing at most $\frac{n}{4d+6} - d - 5$ cards.*

Proof. Let k be the number of missing cards, where $k \leq n/(4d+6) - d - 5$ and let the partial deck of cards consist of $G_1, \dots, G_{n-k}$ with $G_i = G - v_i$. After possibly reordering the cards, we may assume that $|E(G_1)| \geq |E(G_2)| \geq \cdots \geq |E(G_{n-k})|$ or, equivalently, $d_G(v_1) \leq \cdots \leq d_G(v_{n-k})$. We may assume that $n \geq 2(d+6)(2d+3) \geq 36$, else there would be no cards missing. So by Lemma 4.13, it follows that any graph H with this partial deck has average degree at most $d+1$. Then Observation 4.12 and the conditions on k together imply that any such graph has at least $k+1$ vertices of degree at most $d+1$ and in particular the vertex v_1 has degree at most $d+1$. Therefore, the corresponding card G_1 must satisfy

$$d_{<t}(G_1) \in [d_{<t}(G) - 1, d_{<t}(G) + d + 1] \tag{4.3.1}$$

for each $t \in [0, n-1]$. The upper bound comes from the observation that v_1 has at most $d+1$ neighbours. Only these vertices can drop degree from t to $t-1$ when v_1 is deleted, and hence be counted in $d_{<t}(G_1)$ but not in $d_{<t}(G)$. For the lower bound, since the degree of a vertex in G_1 is at most the degree of the corresponding vertex in G, the only possible loss comes from the possibility that the deleted vertex v_1 may have had degree less than t. Applying Observation 4.12 first and then (4.3.1), we find

$$\frac{1}{2}n \leq \sum_{t=0}^{2(d+1)} d_t(G) \leq 1 + \sum_{t=0}^{2(d+1)} d_t(G_1).$$

It follows that there must be some $t \in [0, 2(d+1)]$ such that

$$d_t(G_1) \geq \frac{1}{2d+3}\left(\frac{1}{2}n - 1\right) \geq k + d + 4, \tag{4.3.2}$$

where the last inequality holds by our assumptions on k. Let us choose t to be the smallest integer satisfying (4.3.2), noting that this is determined by G_1. Our next goal is to find a card corresponding to a vertex with degree exactly t.

Set $j := d_{<t}(G_1)$. We claim that $d_G(v_{j+2}) = t$. From (4.3.1), we see that $d_G(v_{j+2}) \geq t$ since $j + 2 = d_{<t}(G_1) + 2 > d_{<t}(G)$. Moreover,

$$\begin{aligned} j + 2 &= d_{<t+1}(G_1) - d_t(G_1) + 2 \\ &\leq d_{<t+1}(G) + d + 1 - (k + d + 4) + 2 < d_{<t+1}(G) - k. \end{aligned}$$

by the bounds in (4.3.1) and (4.3.2). Since we are missing at most k cards from our deck, the $(j+2)$nd card is certainly within the first $j + k + 2$ cards in the whole deck. Hence, we also have the reverse inequality $d_G(v_{j+2}) \leq t$. This proves our claim that $d_G(v_{j+2}) = t$. Since we can compute j and t, and we have the card G_{j+2}, the number of edges may now be reconstructed by the formula $|E(G)| = |E(G_{j+2})| + t$. □

The preceding proof extends easily to reconstructing clique counts from an incomplete deck by replacing $d_t(H)$ and $d_{<t}(H)$ with analogous notions in terms of "clique degree". Namely, for each fixed $r \in \mathbb{N}$, let $c(v)$ be the number of r-cliques which contain the vertex v. Then let the number of vertices v for which $c(v) = t$ (similarly, $c(v) < t$) be denoted by $c_t(H)$ ($c_{<t}(H)$). Any vertex of degree at most $2(d+1)$ is in at most $\binom{2(d+1)}{r-1}$ cliques, and therefore

$$\frac{n}{2} \leq 1 + \sum_{t=0}^{2(d+1)} d_t(G_1) \leq 1 + \sum_{t=0}^{\binom{2(d+1)}{r-1}} c_t(G_1).$$

By the pigeonhole principle, there is some $t \in \{0, \ldots, 2(d+1)\}$ such that

$$c_t(G_1) \geq \frac{1}{\binom{2(d+1)}{r-1} + 1}\left(\frac{n}{2} - 1\right).$$

Again, only the neighbours of v_1 may appear in fewer cliques in G_1 than in G, and so $c_{<t}(G_1) \in [c_{<t}(G) - 1, c_{<t}(G) + d_G(v_1)]$. Since the number of r-cliques in G_{j+2} is the number of r-cliques in G minus $c(v_{j+2})$, it suffices to choose k to guarantee $c(v_{j+2}) = t$ for $j = c_{<t}(G_1)$. This proves:

Theorem 1.6. *Let $d, r \in \mathbb{N}$. For any graph G on $n \geq 3$ vertices with average degree at most d, the number of cliques of size r in G can be reconstructed from any deck missing at most $\left(1 + \binom{2(d+1)}{r-1}\right)^{-1}\left(\frac{n}{2} - 1\right) - d - 5$ cards.*

4.3.2 Degree sequence reconstruction

We now prove our main result on reconstructing degree sequence.

Theorem 1.4. *Let* G *have* $n \geq 3$ *vertices and average degree bounded by some* $d \in \mathbb{N}$. *Then the degree sequence of* G *can be reconstructed from its deck when at most* $\frac{n}{10^4 d^3}$ *of the cards are missing.*

Note that we have $\frac{n}{10^4 d^3} \leq \frac{n}{4d+6} - d - 5$ when $d \geq 1$ and $n \geq 10^4 d^3$. Both of these assumptions can be made for free, the former being clear and the latter since we know how to reconstruct the degree sequence if no cards are missing. Applying Theorem 1.5 therefore allows us to reconstruct the number of edges in G. This means that we can determine $d_G(v_i)$ for all vertices corresponding to cards in our partial deck, as well as the actual average degree of G.

The total number of occurrences of degree t vertices across all of the cards is $\sum_{i=1}^n d_t(G_i)$. At the same time, each vertex v of degree t in G still has degree t in $n-(t+1)$ cards, namely all those G_i for which $v_i \notin N_G(v) \cup \{v\}$. A vertex v of degree $t+1$ has degree t in G_i if and only if $v_i \in N_G(v)$. Hence

$$\sum_{i=1}^{n} d_t(G_i) = (n-1-t)d_t(G) + (t+1)d_{t+1}(G). \tag{4.3.3}$$

In order to guess $d_t(G)$ from $d_{t+1}(G)$ or vice versa, we first obtain a good estimate on the left-hand side of the equation above.

Lemma 4.14. *Suppose we know the number of edges of* G *and that the average degree of* G *is at most* $d \in \mathbb{N}$. *Moreover, assume that* $k \leq \frac{n}{1100d^2}$ *is the number of missing cards and* $n \geq 10^4 d^3$ *is the number of vertices. Then for every* $t \in [0, n]$, *we can reconstruct an estimate* $\widetilde{s}_t$ *from the given cards such that* $|\widetilde{s}_t - \sum_{i=1}^n d_t(G_i)| < \frac{n}{8}$.

Proof. Fix any $t \in [0, n]$. We again label the given cards $G_1, \ldots, G_{n-k}$ so that $|E(G_1)| \geq \cdots \geq |E(G_{n-k})|$. Let the missing cards be $G_{n-k+1}, \ldots, G_n$ ordered arbitrarily.

To estimate $\sum_{i=1}^n d_t(G_i)$, we partition $[n]$ into three sets I_1, I_2, and I_3 defined as follows:

$$\begin{aligned} I_1 &= \{i \in [2, n] : d_{G_1}(v_i) > 100d^2\}, \\ I_2 &= \{i \in [n-k] : d_G(v_i) \leq 100d^2\}, \\ I_3 &= [n] - (I_1 \cup I_2). \end{aligned}$$

We assume that the vertex numbering of G_1 is inherited from G for the sake of the argument, but we do not exploit that these labels are present on our given card. In particular, we do not access the set I_1, only the multiset $\{d_{G_1}(v_i) : i \in [2,n] \text{ with } d_{G_1}(v_i) > 100d^2\}$. Note that $I_1 \cap I_2 = \emptyset$ as $d_{G_1}(w) \leq d_G(w)$ for all $w \in V(G_1)$, so we can write $[n] = I_1 \sqcup I_2 \sqcup I_3$ as a disjoint union. Moreover, note that $1 \in I_2$ by Observation 4.12.

For each $j = 1, 2, 3$, we estimate $\sum_{i \in I_j} d_t(G_i)$. Recall that we know the number of edges of G and hence the degrees of $v_1, \dots, v_{n-k}$. This is enough to reconstruct the set I_2, and to read off $d_t(G_i)$ for each $i \in I_2$ by examining the relevant card. Therefore, we can determine $\sum_{i \in I_2} d_t(G_i)$ exactly.

We estimate $\sum_{i \in I_1} d_t(G_i)$ by $\sum_{i \in I_1} d_t(G_1 - v_i)$, and we now bound the error of this estimation. The vertex v_1 has degree at most d in G by Observation 4.12, and so, for each $i \in [2,n]$, the vertex v_1 has degree at most d in the graph G_i. It follows that for all $i \in [2,n]$,

$$|d_t(G_i) - d_t(G_1 - v_i)| = |d_t(G_i) - d_t(G_i - v_1)| \leq d + 1.$$

Since there are at most $\frac{n}{100d}$ vertices with degree greater than $100d^2$ in G and hence also at most $\frac{n}{100d}$ such vertices in G_1, we find that

$$\sum_{i \in I_1} |d_t(G_i) - d_t(G_1 - v_i)| \leq (d+1) \cdot |I_1| \leq (d+1) \cdot \frac{n}{100d} \leq \frac{2n}{100}. \tag{4.3.4}$$

Finally, we can express I_3 as the union

$$\{i \geq n-k+1 : d_G(v_i) \leq 100d^2\} \cup \{i > 1 : d_{G_1}(v_i) \leq 100d^2 \text{ and } d_G(v_i) > 100d^2\}.$$

In this form, we see that all vertices v_i with $i \in I_3$ have degree at most $100d^2 + 1$ in G. The first set in the union has cardinality at most k and the second has cardinality at most $d_G(v_1) \leq d$ (since all such vertices must be adjacent to v_1). Thus $|I_3| \leq k + d$. Moreover, observe that $|d_t(G_i) - d_t(G_j)| \leq d_G(v_i) + d_G(v_j) + 1$. This implies that

$$\sum_{i \in I_3} |d_t(G_i) - d_t(G_1)| \leq (100d^2 + d + 2)|I_3| \leq (100d^2 + d + 2)(k + d) \tag{4.3.5}$$

so we can estimate $\sum_{i \in I_3} d_t(G_i)$ by $|I_3| d_t(G_1)$. Note that we can reconstruct $|I_3| = n - |I_1| - |I_2|$ from the cards.

We now estimate $\sum_{i=1}^{n} d_t(G_i)$ by

$$\widetilde{s}_t = \sum_{i \in I_1} d_t(G_1 - v_i) + \sum_{i \in I_2} d_t(G_i) + |I_3| d_t(G_1),$$

which is reconstructible from our partial deck.

Using (4.3.4) and (4.3.5), the margin of error $|\sum_{i=1}^n d_t(G_i) - \widetilde{s_t}|$ is then given by

$$\begin{aligned}
&\left|\sum_{i\in I_1} d_t(G_i) + \sum_{i\in I_3} d_t(G_i) - \sum_{i\in I_1} d_t(G_1 - v_i) - |I_3| d_t(G_1)\right| \\
&\leq \sum_{i\in I_1} |d_t(G_i) - d_t(G_1 - v_i)| + \sum_{i\in I_3} |d_t(G_i) - d_t(G_1)| \\
&\leq \frac{2n}{100} + (100d^2 + d + 2)(k + d)
\end{aligned}$$

and this is less than $\frac{n}{8}$ for $k \leq \frac{n}{1100d^2}$ and $n \geq 10^4 d^3$. □

We now deduce the main result.

Proof of Theorem 1.4. Following the discussion at the start of this section, we may assume that $n \geq 10^4 d^3$, that we have already reconstructed the number of edges in G, and that we have therefore determined the best possible upper bound $d \in \mathbb{N}$ on the average degree. In particular, for every $t \in [0, n]$, Lemma 4.14 provides an estimate $\widetilde{s_t}$ for $\sum_{i=1}^n d_t(G_i)$ with $|\widetilde{s_t} - \sum_{i=1}^n d_t(G_i)| < n/8$.

Rewriting (4.3.3), we obtain

$$d_t(G) = \frac{1}{(n-1-t)}\left(\sum_{i=1}^n d_t(G_i) - (t+1)d_{t+1}(G)\right)$$

and, estimating $\sum_{i=1}^n d_t(G_i)$ by $\widetilde{s_t}$, we obtain the following estimate for $d_t(G)$

$$\widetilde{d_t} = \frac{1}{(n-1-t)}\left(\widetilde{s_t} - (t+1)d_{t+1}(G)\right).$$

If $t + 1 \leq \frac{3n}{4}$, then $\frac{n}{8} \leq \frac{1}{2}(n-1-t)$ and hence

$$\left|d_t(G) - \widetilde{d_t}\right| = \frac{1}{(n-1-t)}\left|\sum_{i=1}^n d_t(G_i) - \widetilde{s_t}\right| < \frac{1}{2}. \tag{4.3.6}$$

If d_{t+1} is known exactly, this means that we can reconstruct d_t exactly by rounding $\frac{1}{(n-1-t)}\left(\widetilde{s_t} - (t+1)d_{t+1}(G)\right)$ to the nearest integer. A symmetric argument, obtained by solving (4.3.3) for $d_{t+1}(G)$ and using the same estimate $\widetilde{s_t}$, shows that we can also reconstruct d_{t+1} given d_t and the partial deck when $t \geq \frac{1}{4}n$.

We now show that there is a $t \in \left[\frac{1}{4}n, \frac{3}{4}n\right]$ for which we can reconstruct that $d_t(G) = 0$. Observe that if there are two cards in our partial deck with no vertices of degree t or $t-1$, then $d_t(G) = 0$. Moreover, if there is a t such that $d_{t-1}(G)$, $d_t(G)$ and $d_{t+1}(G)$ are all 0, then no vertices of degree $t-1$ or t will appear on any card, and we can reconstruct that $d_t(G) = 0$. We claim that this is the case for some $t \in \left[\frac{1}{4}n, \frac{3}{4}n\right]$. Suppose for a contradiction that $d_{t-1}(G) + d_t(G) + d_{t+1}(G) \geq 1$ for every $t \in \left[\frac{1}{4}n, \frac{3}{4}n\right]$. Then

$$\sum_{t\in\left[\frac{1}{4}n,\frac{3}{4}n\right]} d_t(G) \geq \frac{1}{3} \sum_{t\in\left[\frac{1}{4}n+1,\frac{3}{4}n-1\right]} (d_{t-1}(G) + d_t(G) + d_{t+1}(G)) \geq \frac{1}{3} \cdot 1 \cdot \left(\frac{n}{2} - 4\right)$$

which implies that

$$dn \geq \sum_{t\in\left[\frac{1}{4}n,\frac{3}{4}n\right]} td_t(G) \geq \frac{1}{4 \cdot 6} n(n-8).$$

This contradicts the assumption that $n \geq 10^4 d^3$.

Fix a $t \in \left[\frac{1}{4}n, \frac{3}{4}n\right]$ such that there are no vertices of degree t or $t-1$ on any card as found above. Since $d_t(G)$ is known exactly (to be 0), we may now reconstruct the estimate for d_{t-1} given in (4.3.6) and round to determine d_{t-1} exactly. This process allows us to iteratively reconstruct $d_{t-1}(G), \ldots, d_0(G)$. Returning to $d_t(G)$, we can also 'push' in the other direction using the symmetric estimate to determine $d_{t+1}(G), d_{t+2}(G), \ldots, d_{n-1}(G)$ in order as well. □

There are two tricks in this proof that we wish to highlight as they seem to be becoming staples in this variant (see [68]). The first is surprisingly powerful for its simplicity; when we are estimating counts for some parameter, we should pick values to aim for an error of strictly less than 1/2 so that we can recover the exact value by rounding to the nearest natural number. This gets us quite far, at least up to a certain limit on the number of missing cards. The second, currently specific to degree sequence, is the method of pushing and pulling that consists of establishing $d_t(G)$ exactly for one value, and using this to step up and down until the whole range is covered. This capitalises on the fact that when comparing any two cards, the degree of any specific vertex can differ by at most 1. It would be interesting to consider whether a similar approach could be adapted for other parameters.

Chapter 5

Algebraic methods

Algebraic techniques abound in reconstruction, and they have long been regarded among the most promising ways forward[1]. Anecdotally, when algebraic connections have been established, they have often led to elegant proofs, sometimes created a shift in the way the problems are viewed, and still enjoy the generality of more elementary methods. Interestingly, the connections occasionally also work the other way, where relationships derived for graphs can be translated back to algebraic results.

The relevance of algebraic structures is not very surprising. For example, given that reconstruction can be done easily for labelled graphs, it is expected that automorphism groups are relevant to reconstruction. This has been studied in vertex reconstruction by the likes of Kocay [91], and to an even greater extent in edge reconstruction (see the work of Lóvasz [112], Nash-Williams [158], and Krasikov, Lev and Thatte [100], or the overview from Bondy [17]). Another motive to look at algebraic formulations is that exact expressions from direct counting can quite quickly become complicated and unwieldy. The use of polynomials for enumeration by Tutte [172] and linear algebraic formulations stemming from Kocay's Lemma by Oliveira and Thatte [132] represent a long timeline of attempts to address this issue in vertex reconstruction, and similar techniques have been employed effectively by Stanley in switching reconstruction.

The two applications that we develop in this chapter are of new polynomial methods for reconstruction from small cards (Section 5.1) and generic switching reconstruction (Section 5.2). In both cases, the basic algebraic framework for the methods we use had already been established. We are able to go further by applying different results from outside standard reconstruction literature to strengthen these tools.

---

[1] In [17], Bondy writes: "it is likely, if not inevitable, that group-theoretic methods will play a decisive role in the resolution of the Vertex and Edge Reconstruction conjectures".

5.1 In vertex reconstruction

In 1989, Taylor showed that, for any fixed k, the degree sequence of every sufficiently large graph is determined by its deck of k-vertex-deleted subgraphs, where the required size is given by the function

$$f(k) = (k + \log k + 1)\big(e + \frac{e \log k + e + 1}{(k-1)\log k - 1}\big) + 1.$$

Theorem 5.1 (Taylor [169]). *Let $k \geq 3$ be an integer. Then the degree sequence of all graphs on at least $f(k)$ vertices are reconstructible from their $(n-k)$-decks.*

This implies that it is enough to have cards with size about $(1 - 1/e)n$, where n is the order of the graph as usual. The particularly impressive part of this result is its novel proof. The basic observation that Taylor used was that, for a graph G, the star subgraph count $n_{K_{1,j}}(G)$ is equal to $\sum_{v \in V(G)} \binom{d(v)}{j}$. This corresponds to the fact that all stars $K_{1,j}$ with centre v in G are obtained by choosing any j neighbours of v, and gives us a reconstructible quantity to work with by Kelly's Lemma. So, if we cannot reconstruct the degree sequence of a graph, then this means that there are two graphs G_1 and G_2 with different degree sequences for which

$$\sum_{v \in V(G_1)} \binom{d(v)}{j} = \sum_{v \in V(G_2)} \binom{d(v)}{j}.$$

The next simple yet significant observation is that this has little to do with the graphs at all; we simply have two distinct sequences of n natural numbers in $\{0, \ldots, n-1\}$ for which the above condition holds. A fairly involved calculation then shows that no such sequences can exist. Getting to the heart of the calculation, Taylor noted that his proof produces a more general result which can be formulated in number theoretic terms.

Lemma 5.2 (Taylor [169]). *Let $k \geq 3$ be an integer and $n \geq f(k)$ where f is the function defined above. Let $(a_i)_{i=1}^{n}$ and $(a'_i)_{i=1}^{n}$ be two non-increasing sequences of integers taking values in $\{0, \ldots, n-1\}$ where*

$$\sum_{i=1}^{n} (a_i)^r = \sum_{i=1}^{n} (a'_i)^r \quad \textit{for all } r \in \{1, \ldots, n-k-1\}.$$

Then $a_i = a'_i$ for all $1 \leq i \leq n$.

Improving on Taylor's method, we will now show that it is possible to obtain tighter bounds by replacing Lemma 5.2 with a stronger result on shared moments of sequences. We will then apply this to not only degree sequence but also reconstructing the connectedness of a graph from small cards in Section 5.1.1. This is based on joint work with Groenland, Johnston and Scott.

Let us begin with the algebraic aspect. For this, we will need a bound on the maximum number of shared moments that two sequences $\alpha, \beta \in \{0, \ldots, n\}^m$ can have. This follows from a theorem on the number of positive real roots of a polynomial.

Theorem 5.3 (Theorem A in [22]). *Suppose that the complex polynomial*

$$p(z) := \sum_{j=0}^{n} a_j z^j$$

has k positive real roots. Then

$$k^2 \leq 2n \log \left(\frac{|a_0| + |a_1| + \cdots + |a_n|}{\sqrt{|a_0 a_n|}} \right).$$

The history of this theorem is a little hazy, although it does predate Taylor's result by some 50 years. This theorem is commonly attributed to Schmidt, but the first published proof is due to Schur and a series of simplifications have followed (see [22]). The specific transformation that we require was given by Borwein and Ingalls [23, Proposition 1]. Applying this, we shall use the following formulation which is tailored to our purposes.

Lemma 5.4. *Let $\alpha, \beta \in \{0, \ldots, n\}^m$ be two sequences that are not related to each other by a permutation. If*

$$\binom{\alpha_1}{j} + \cdots + \binom{\alpha_m}{j} = \binom{\beta_1}{j} + \cdots + \binom{\beta_m}{j} \quad \text{for all } j \in \{0, \ldots, \ell\}, \tag{5.1.1}$$

then $\ell + 1 \leq \sqrt{2n \log(2m)}$.

Proof. Since $\alpha_i, \beta_j \in \{0, \ldots, n\}$ for all $i, j \in [m]$,

$$p_{\alpha,\beta}(x) := \sum_{i=1}^{m} x^{\alpha_i} - \sum_{i=1}^{m} x^{\beta_i} \tag{5.1.2}$$

is a polynomial of degree at most n. For $c \in \mathbb{C}$, let $\mathrm{mult}_c(p_{\alpha,\beta})$ denote the multiplicity of the root at c, or 0 if c is not a root of $p_{\alpha,\beta}$. We will show that $\ell + 1 \leq \mathrm{mult}_1(p_{\alpha,\beta}) \leq \sqrt{2n \log(2m)}$.

Since α and β are not related by a permutation, the polynomial $p_{\alpha,\beta}$ is non-zero. We may write (with $r = \mathrm{mult}_0(p_{\alpha,\beta})$)

$$p_{\alpha,\beta}(x) = x^r \left(\sum_{j=0}^{n'} a_j x^j \right)$$

where a_0 and $a_{n'}$ are non-zero and $n' \leq n$. The coefficients are all integral, so $\sqrt{|a_0 a_{n'}|} \geq 1$. Moreover, from the definition of the polynomial in (5.1.2) there are at most $2m$ contributions of ± 1 to the coefficients, so we have $\sum_{i=0}^{n'} |a_i| \leq 2m$.

By Theorem 5.3, the number of positive real zeros of $\sum_{j=0}^{n'} a_j x^j$ is at most

$$\sqrt{2n' \log \left(\frac{|a_0| + |a_1| + \cdots + |a_{n'}|}{\sqrt{|a_0 a_{n'}|}} \right)} \leq \sqrt{2n \log(2m)}$$

and in particular, $\mathrm{mult}_1(p_{\alpha,\beta}) \leq \sqrt{2n\log(2m)}$. On the other hand, for all $j \in \{0, \ldots, \ell\}$, equation (5.1.1) shows that

$$\left| \left(\frac{d}{dx^j} \left[\sum_{i=1}^{m} x^{\alpha_i} - \sum_{i=1}^{m} x^{\beta_i} \right] \right) \right|_{x=1} = \sum_{i=1}^{m} j! \binom{\alpha_i}{j} - \sum_{i=1}^{m} j! \binom{\beta_i}{j} = 0.$$

Hence, $\ell + 1 \leq \mathrm{mult}_1(p_{\alpha,\beta})$, and $\ell + 1 \leq \sqrt{2n\log(2m)}$ as desired. □

Condition (5.1.1) is equivalent to the condition that the first ℓ moments of α and β agree. To see this, observe that $\{x^i : i \in \{0, \ldots, \ell\}\}$ and $\{\binom{x}{i} : i \in \{0, \ldots, \ell\}\}$ both form a basis for the polynomials of degree at most ℓ.

5.1.1 Two applications

Lemma 5.4 will allow us to prove that the degree sequence of an n-vertex graph G can be reconstructed from the ℓ-deck of G whenever $\ell \geq \sqrt{2n\log(2n)}$. This proof is essentially identical to that given by Taylor [169], except for the use of the better bounds.

Theorem 1.2. *The degree sequence of an n-vertex graph G can reconstructed from $\mathcal{D}_\ell(G)$ for any $\ell \geq \sqrt{2n\log(2n)}$.*

Proof. Let G be an n-vertex graph with vertices $v_1, \ldots, v_n$, and let $\ell \geq \sqrt{2n\log(2n)}$ be an integer. By Lemma 4.1, we can reconstruct the number of subgraphs of G

isomorphic to the star $K_{1,j}$ for all $j \in \{2, \dots, \ell - 1\}$. Since vertex v lies at the centre of $\binom{d(v)}{j}$ copies of $K_{1,j}$, we can compute the quantity

$$\widetilde{n}_{K_{1,j}}(G) = \sum_{v \in V(G)} \binom{d(v)}{j}$$

from the ℓ-deck. We can also reconstruct

$$\sum_{v \in V(G)} \binom{d(v)}{0} = n \text{ and } \sum_{v \in V(G)} \binom{d(v)}{1} = 2 \cdot e(G)$$

from the 2-deck. Write $\alpha_i = d(v_i)$ for $i \in [n]$ where we may assume $d(v_1) \leq \cdots \leq d(v_n)$. Suppose, for a contradiction, that a different degree sequence $\beta_1 \leq \cdots \leq \beta_n$ gives the same counts. Thus, for $j \in \{0, \dots, \ell - 1\}$,

$$\sum_{i=1}^{n} \binom{\alpha_i}{j} = \sum_{i=1}^{n} \binom{\beta_i}{j}.$$

Since $\alpha, \beta \in \{0, \dots, n-1\}^n$ are not permutations of each other, Lemma 5.4 applies to show $\ell \leq \sqrt{2(n-1)\log(2n)}$, giving the desired contradiction. $\square$

We now turn to reconstructing connectedness from the ℓ-deck, where we will employ a slightly more sophisticated version of the last proof. Recall that a copy H' of H in a graph G refers to an induced subgraph of G that is isomorphic to H where in particular we wish to keep track of the vertices involved.

Theorem 1.3. *The connectedness of an n-vertex graph G can be recognised from $\mathcal{D}_\ell(G)$ provided $\ell \geq 9n/10$.*

Proof. Let G be an n-vertex graph and let $\varepsilon = 1/10$. Suppose ℓ is an integer such that $\ell \geq 9n/10 = (1 - \varepsilon)n$. We wish to recognise whether G is connected from the ℓ-deck.

We begin by making some additional assumptions, which can be done for free using existing results. First, it was shown by Kostochka, Nahvi, West, and Zirlin [95] that the connectedness of a graph can be recognised from the $(n-3)$-deck for $n \geq 7$, so we can assume that $n \geq 39$. Second, Kostochka and West [97] showed that, for $\ell > n/2$, if every connected subgraph of G has at most ℓ vertices, then G is reconstructible from the ℓ-deck (this is an application of Lemma 4.3). We may therefore assume that the largest component of G has order at least ℓ. In particular, if G is not connected then it has a component of order at most $n - \ell$.

We will reconstruct all components that have at most $n - \ell$ vertices from the ℓ-deck. Let H be a connected graph with h vertices, where $1 \leq h \leq \varepsilon n$. Since $h \leq \ell$, we may compute $n_H(G)$ from the ℓ-deck by Lemma 4.1. Suppose $m = n_H(G) > 0$. Write $H_1, \dots, H_m$ for the induced copies of H in G, and define the *neighbourhood* of H_i by

$$\Gamma(H_i) = \{v \in V(G) \setminus V(H_i) : vu \in E(G) \text{ for some } u \in H_i\}.$$

Define the *degree* of H_i to be $|\Gamma(H_i)|$, and denote it by α_i. Note that G has a component isomorphic to H if and only if $\alpha_i = 0$ for some $i \in [m]$. Thus, reconstructing the sequence $(\alpha_1, \dots, \alpha_m) \in \{0, \dots, n-h\}^m$ determines the number of components isomorphic to H.

We now show that we can reconstruct $(\alpha_1, \dots, \alpha_m)$ up to permutation. Since $1 \leq h \leq \varepsilon n$ and $m \leq \binom{n}{h} \leq \left(\frac{en}{h}\right)^h$ we have

$$\begin{aligned}\sqrt{2(n-h)\log(2m)} &\leq \sqrt{2(n-h)h\log(en/h) + 2n\log(2)} \\ &\leq n\sqrt{2(1-\varepsilon)\varepsilon\log(e/\varepsilon) + 2\log(2)/n}\end{aligned}$$

using that $(n-h)h\log(en/h)$ increases with h within the given range. Hence by Lemma 5.4, it suffices to show that we can reconstruct

$$\sum_{i=1}^{m} \binom{\alpha_i}{j} \text{ for all integers } 0 \leq j \leq N, \tag{5.1.3}$$

where $N = n\sqrt{2(1-\varepsilon)\varepsilon\log(e/\varepsilon) + 2\log(2)/n}$.

Let P denote the set of pairs of vertex sets (A, B) where $A \subseteq B \subseteq V(G)$, $G[A] \cong H$, $|B| = |A| + j$ and A is *dominating* in $G[B]$ – that is, each vertex in $B \setminus A$ is adjacent to some vertex in A. Each $(A, B) \in P$ has some $i \in [m]$ for which $G[A] \cong H_i$ and B is contained in the neighbourhood of H_i, so $|P| = \sum_{i=1}^{m} \binom{\alpha_i}{j}$.

For $j \geq 0$, let $\mathcal{H}_j$ denote the set of $(h+j)$-vertex graphs that consist of H along with j additional vertices, all of which are adjacent to at least one vertex in the copy of H (we include each isomorphism type once). If $(A, B) \in P$, then B corresponds to some $H' \in \mathcal{H}_j$. By definition, there are $n_{H'}(G)$ vertex sets $B \subseteq V(G)$ with $G[B] \cong H'$. Since $\mathcal{H}_j$ and H are known to us, for each $H' \in \mathcal{H}_j$ we can calculate the number $n(H, H')$ of dominating copies of H in H'. Since

$$\sum_{H' \in \mathcal{H}_j} n(H, H') n_{H'}(G) = |P| = \sum_{i=1}^{m} \binom{\alpha_i}{j},$$

it only remains to show that we can determine $n_{H'}(G)$ from the ℓ-deck.

We may use Lemma 4.1 to reconstruct $n_{H'}(G)$ if $|H'| = h + j \leq \ell$. For $j \leq N$ and $n \geq 39$, we find that

$$h + j \leq \varepsilon n + N \leq n - \varepsilon n \leq \ell,$$

where the middle inequality follows from the fact that

$$\sqrt{2(1-\varepsilon)\varepsilon \log(e/\varepsilon) + 2\log(2)/39} \leq 1 - 2\varepsilon$$

for $\varepsilon = 1/10$.

This shows that we can reconstruct (5.1.3), and hence the number of components isomorphic to H. In particular, doing so for every graph H with at most $n - \ell$ vertices allows us to determine whether any component of G has at most $n - \ell$ vertices, which we saw would hold if and only if G is disconnected. □

The constant 9/10 can be improved slightly in the proof above provided n is large enough. Indeed, the proof holds for any n and ε such that

$$\sqrt{2(1-\varepsilon)\varepsilon \log(e/\varepsilon) + 2\log(2)/n} \leq 1 - 2\varepsilon,$$

and, for large enough n, we can take $\varepsilon \approx 0.1069$.

5.2 In switching reconstruction

The intimate connection between switching reconstruction problems and algebraic structures has been investigated right from beginning in Stanley's vertex-switching variant. The goal of this section is to shed light on how these structures come about as well as how they can they be used to draw conclusions on generic reconstructibility.

To this end, we first demonstrate how the Krawtchouk polynomials arose in early results for the vertex-switching problem (Section 5.2.1). We will then show that this algebraic structure still appears under the generic framework for switchings that we introduced in Section 3.4.2, and use it in a novel way to derive a host of identities relating different notions of extended decks to one another (Section 5.2.2). It follows that certain extended decks are (switching-)reconstructible, which allows us to derive a new sufficient condition for generic reconstruction problems (Section 5.2.3). Unfortunately, these results narrowly miss out on improving on existing work for Stanley's problem. Nonetheless, further development of these methods could lead to such a leap. These last two sections are based on joint work with Brendan McKay.

Having now alluded to the Krawtchouk polynomials several times though, we begin this section by defining them. The Krawtchouk polynomials were introduced by Mikhail Krawtchouk in [105] as a family of discrete orthogonal polynomials associated with multinomial distributions. Specifically, we are interested in the binary Krawtchouk polynomials defined by the following generating function;

$$\sum_{k=0}^{\infty} P_k^n(x) z^k = (1-z)^x (1+z)^{n-x}$$

There are several explicit expressions, including

$$\begin{aligned} P_k^n(x) &= \sum_{i=0}^{k} (-1)^i \binom{x}{i} \binom{n-x}{k-i} \\ &= \sum_{i=0}^{k} (-2)^i \binom{x}{i} \binom{n-i}{k-i} \\ &= \sum_{i=0}^{k} (-1)^i 2^{k-i} \binom{n-x}{k-i} \binom{n-k+i}{i}. \end{aligned}$$

We also have a third method of defining these polynomials, as among a number of nice properties, they satisfy the recurrence relation

$$(k+1)P_{k+1}^n(x) = (n-2x)P_k^n(x) - (n-k+1)P_{k-1}^n(x).$$

The Krawtchouk polynomials have many applications such as to coding theory (see [114, 167]). For reconstruction, we make heavy use of the various equivalent definitions to convert recurrence relations to explicit formulae expressing relationships between decks. Another important collection of results for reconstruction applications are those on the integral roots of Krawtchouk polynomials. From [99], all roots of P_s^n for fixed n are real, lie in the interval $(0, n)$, and are symmetric with respect to $\frac{n}{2}$. However, not much is known on integral roots. Most pertinent results are collected in [101, 51]. We will state specific results concerning Krawtchouk polynomials as we wish to apply them.

5.2.1 Emergence in Stanley's problem

Stanley's 1985 paper introducing vertex-switching reconstruction includes a proof of the fact that every graph on $n \not\equiv 0 \pmod 4$ vertices is vertex-switching reconstructible (Theorem 3.17). This was one of the first significant algebraic proofs in the area, and we outline it below. We shall use the notation introduced in Section 3.4.

Proof outline (following Stanley [162]). Let $\mathcal{G}_n$ be the set of all graphs with vertices $Z = \{z_1, \ldots, z_n\}$, and let $\mathcal{V}_N$ denote the real vector space of all formal linear combinations $\sum_{G \in \mathcal{G}_n} a_G G$, $a_G \in \mathbb{R}$. Define a linear map $T : \mathcal{V}_n \to \mathcal{V}_n$ by

$$T(G) = G_{z_1} + \cdots + G_{z_n}$$

where each G_{z_i} is a labelled card in $\mathcal{D}(G)$. We also require unlabelling map that maps a graph X to $[X]$, which is a formal sum of the graph vectors of all members of the isomorphism class of X. Then $[X] = [X']$ if and only if $X \cong X'$. This map is linear and commutes with T. The bulk of the proof lies in showing that T is invertible if and only if $n \not\equiv 0 \pmod 4$. This is done using a lemma by Diaconis and Graham (see [1, 128]). Then for any X and X' on the same vertex set with the same deck, that is $X_{z_i} \cong X'_{z_i}$ for all $1 \leq i \leq n$ which implies $[X_{z_i}] = [X'_{z_i}]$, combining these properties gives

$$T[X] = [TX] = [X_{z_1} + \ldots + X_{z_n}] = [X_{z_1}] + \ldots + [X_{z_n}] = [X'_{z_1}] + \ldots + [X'_{z_n}] = T[X'].$$

When $n \not\equiv 0 \pmod 4$, this means $[X] = [X']$ so $X \cong X'$. □

We have already mentioned that this proof has been superseded in simplicity, however its enduring importance is that it acted as a springboard for many later results. In particular, it introduces some useful ideas such as representing decks by formal sums and the 'deck-taking operator' T. These will be central to our work in the next section. It is also already interesting to express T as a linear map. For reconstruction, we are given $\mathcal{D}_1(G) = T[G]$ and wish to recover $[G]$. If T is invertible, then we can calculate $[G] = T^{-1}\mathcal{D}_1(G)$. A lot of work exists on this strategy, and hence on studying the eigenvalues of T. One weakness, however, is illustrated by the fact that the reconstructibility of $[G]$ from $\mathcal{D}_1(G)$ only requires the matrix form of T to have distinct columns—a much weaker condition than non-singularity.

Stanley also used a similar argument to prove a k-switching generalisation, via the following result which establishes a connection with Krawtchouk polynomials.

Theorem 5.5 (Stanley [162]). *Suppose that $P_k^n(x)$ has no even integer roots in the interval $[0, n]$. Then any graph with n vertices is k-switching reconstructible.*

The idea of this proof is to define $T_i(G) = \sum_{W \subseteq V(G), |W|=i} G_W$, and using the earlier argument, it turns out that T_i is invertible if and only if the coefficient of x^{n-i} in the polynomial $(1-x)^{o(G)}(1+x)^{n-o(G)}$ is nonzero, where $o(G)$ is the number of odd-degree vertices in G. This is precisely the generating function of the Krawtchouk polynomials, so we actually need $P_i^n(o(G)) \neq 0$. By suitable choice of G, $o(G)$ can achieve any even value in $[0, n]$, which gives the claimed condition. Combining this result with explicit forms of Krawtchouk polynomials for $k \leq 3$ and properties of the polynomials for larger k allows us to derive statements generalising Theorem 3.17. As an example, for $k = 2$ and setting $y = n - 2x$, we have $P_2^n(x) = \frac{1}{2}(y^2 - n) = 0$ and $n = y^2 \pmod 4$, so $y = 0, 1$. This corresponds to the sufficient condition that any graph on n vertices where n is not the square of an integer congruent to $0, 1 \pmod 4$ is reconstructible from $\mathcal{D}_k(G)$.

Similar matrix inversion methods were later used to count subgraphs in the vertex-switching setting. Recall that $n_S(G)$ is the number of induced subgraphs isomorphic to S contained in G. Let $\mu_1(G \to H)$ be the number of 1-switchings of G that are isomorphic to H. Then let $t_S(G)$ be the number of occurrences of S across all the cards in $\mathcal{D}(G)$, so $t_S(G) = \sum_{v \in V(G)} n_S(G_v)$ and this is clearly reconstructible. Finally, let $\mathcal{G}_k$ be the set of simple graphs on k vertices. If $|S| = s < n$, then we have

$$t_S(G) = (n - s)n_S(G) + \sum_{T \in \mathcal{G}_s} \mu_1(T \to S)n_T(G).$$

The first term comes from the fact that induced copies of S in G are unchanged when switching on vertices in $V(G) - V(S)$, and the second covers cases where an induced s-vertex graph yields a graph isomorphic to S after switching one of its vertices. To determine the number of induced subgraphs of G isomorphic to S, we consider the switching class $\{S = S_1, \ldots, S_m\}$ of all graphs that can be obtained by switching S in any number of vertices and write out the above equations to obtain an $m \times m$ system of linear equations in the variables $n_{S_j}(G)$. Let M be the matrix corresponding to this system. Since t is reconstructible, if M invertible (this requires work to establish) then we can solve for $n_{S_j}(G)$, which are hence reconstructible.

This is the method employed in [52] to show that the number of triangles is reconstructible when $n \neq 4$. We refer to that proof for a readable example illustrating the approach. A more involved version of this argument also leads to Krasikov and Roditty's proof of the k-vertex-switching analogue of Kelly's Lemma (Theorem 3.20).

5.2.2 Generic deck decompositions

We now address our goal of extending algebraic methods to the generic switching reconstruction problem that we introduced in Section 3.4.2. Let us first recall the notation and axioms from that section. Without specifying an instance, let $\mathcal{G}$ be the set of objects that we will reconstruct up to a notion of equivalence $\sim$, and let Z (with

$|Z| = N$) be the index set for our switching operation defined by $\phi : \mathcal{G} \times 2^Z \to \mathcal{G}$. We assume that the switching operation ϕ satisfies the following properties:

S1. For $G \in \mathcal{G}$, $[G_\emptyset] = [G]$. That is, $\mathcal{D}_0(G) = [G]$.
S2. For $G, G' \in \mathcal{G}$, $G \sim G' \Rightarrow \mathcal{D}_1(G) = \mathcal{D}_1(G')$.
S3. For $X, Y \subseteq Z$ and $G \in \mathcal{G}$, $(G_X)_Y \sim G_{X \triangle Y}$.

To work in this framework, we will do algebra with formal sums $\sum_{\mathcal{Y} \in \mathcal{G}/\sim} c_{\mathcal{Y}}\, \mathcal{Y}$, where the coefficients $c_{\mathcal{Y}}$ are rational numbers. We shall call these multisets even though negative and non-integer coefficients are allowed. The operations of addition, and multiplication or division by a scalar are defined in the obvious manner. Mapping our terminology for reconstruction to this notation, for an integer k and $G \in \mathcal{G}$, let the k-deck of G be

$$\mathcal{D}_k(G) = \sum_{W \subseteq Z \,:\, |W| = k} [G_W].$$

Note that $\mathcal{D}_k(G) = 0$ if $k < 0$ or $k > N$. The multiset $\mathcal{D}_1(G)$ is simply called the *deck* of G, and its elements are *cards*. We will say that a given function of G is *reconstructible* (for specified $(\phi, \sim)$) if it is determined uniquely within $\mathcal{G}$ up to equivalence by $\mathcal{D}_1(G)$. The most important question is whether $[G]$ is reconstructible, which we still write as "G is reconstructible".

An advantage of this generic formulation is that we can now unify some of the arguments we have seen in particular instances of switching problems. The particular type of argument that we are targeting are those that involve relationships between k-decks. To gain leverage from this strategy, we refine our notation $D_k(G)$ as follows. For $G \in \mathcal{G}$, $J \subseteq Z$, and $k, \ell \in \mathbb{Z}$, define an *extended k-deck* by

$$\mathcal{D}_k^{\ell,\, J}(G) = \sum_{W \subseteq Z \,:\, |W| = k, |W \cap J| = \ell} [G_W].$$

Note that $\mathcal{D}_k(G) = \mathcal{D}_k^{0,\, \emptyset}(G)$. We then have the following list of elementary properties.

Lemma 5.6. *Let $G \in \mathcal{G}$, $j, k, \ell \in \mathbb{Z}$, and $J \subseteq Z$. Then the following hold.*

(i) $\mathcal{D}_k^{\ell,\, J}(G) = 0$ *unless* $\max\{0, k + |J| - N\} \le \ell \le \min\{k, |J|\}$.
(ii) $\mathcal{D}_k(G) = \sum_{\ell=0}^{|J|} \mathcal{D}_k^{\ell,\, J}(G)$.
(iii) $\sum_{J \subseteq Z \,:\, |J| = j} \mathcal{D}_k^{\ell,\, J}(G) = \binom{k}{\ell}\binom{N-k}{j-\ell}\mathcal{D}_k(G)$.
(iv) $\mathcal{D}_0^{0,\, J}(G) = [G]$.
(v) $\mathcal{D}_k^{\ell,\, J}(G) = \mathcal{D}_k^{k-\ell,\, Z-J}(G)$.

The next property is perhaps less obvious, though it also follows easily from the extended definition of the deck.

Lemma 5.7. *If $G \in \mathcal{G}$, then for all $k, \ell \in \mathbb{Z}$ and $J \subseteq Z$, $\mathcal{D}_k^{\ell,\, J}(G) = \mathcal{D}_{N-k}^{|J|-\ell,\, J}(G_Z)$. In particular, $\mathcal{D}_{N-k}(G) = \mathcal{D}_k(G_Z)$.*

Proof. The definitions give immediately that

$$\begin{aligned} \mathcal{D}_k^{\ell,\, J}(G) &= \sum_{W \subseteq Z\,:\,|W|=k, |W \cap J|=\ell} [G_W], \\ \mathcal{D}_{N-k}^{|J|-\ell,\, J}(G) &= \sum_{W \subseteq Z\,:\,|W|=N-k, |W \cap J|=|J|-\ell} [(G_Z)_W]. \end{aligned}$$

For any W such that $[G_W]$ is a term in $\mathcal{D}_k^{\ell,\, J}(G)$, $[(G_Z)_{W^c}]$ is a term in $\mathcal{D}_{N-k}^{|J|-\ell,\, J}(G)$. By S3, $[(G_Z)_{W^c}] = [G_{Z-W^c}] = [G_W]$ and the equality follows. The second statement is a special case when $|J| = l = 0$. $\square$

With a view to capitalising on arguments that involve taking iterated decks, it is natural that we would want to have something analogous to taking the deck of an object in $\mathcal{G}$ for formal sums. Generalising the operator that Stanley studied in his proof of Theorem 3.17, we introduce a deck operator T defined as follows. If $\mathcal{Y} = \sum_{H \in \mathcal{G}} c_H[H]$ is a multiset of equivalence classes of $\sim$, define $T\mathcal{Y}$ to be the multiset obtained by replacing each element $[H]$ by $\mathcal{D}_1(H)$; that is, $T\mathcal{Y} = \sum_{H \in \mathcal{G}} c_H \mathcal{D}_1(H)$. We observe that T is a linear operator of dimension N, and so can be represented as a matrix. This operation is well-defined on account of property S2. For a single equivalence class, we have the following expression.

Lemma 5.8. *Let $G \in \mathcal{G}$ and $W \subseteq Z$. Then*

$$T[G_W] = \sum_{z \in W} [G_{W-\{z\}}] + \sum_{z \notin W} [G_{W \cup \{z\}}].$$

Proof. By definition of the T operator and S3, we have

$$T[G_W] = \mathcal{D}_1(G_W) = \sum_{z \in Z} [(G_W)_z] = \sum_{z \in W} [G_{W-\{z\}}] + \sum_{z \notin W} [G_{W \cup \{z\}}]. \qquad \square$$

The main reason to work with this operator T is that it preserves reconstructibility. That is, if some expression $f([G])$ is reconstructible then so is $Tf([G])$. This should be clear from the definition, where applying T is equivalent to performing every elementary switching on every object in our formal sum (where terms are assumed to be known). On the other hand, while it would be nice to also have the decks $D_k(G)$

or the more refined decks that we have defined, these are not directly reconstructible from $\mathcal{D}_1(G)$. Thus, in order to reconstruct a sum involving extended-k-decks, we can strive to express it using terms that only involve the T operator and objects that are known to be reconstructible. To carry this out effectively, we begin by establishing the following crucial lemma which describe precisely how decks behave under T and also establishes our connection to Krawtchouk polynomials.

Lemma 5.9. *Let $G \in \mathcal{G}$, $k, \ell \in \mathbb{Z}$ and $J \subseteq Z$. Then*

$$\begin{aligned} T\mathcal{D}_k^{\ell,\,J}(G) &= (|J| - \ell + 1)\mathcal{D}_{k-1}^{\ell-1,\,J}(G) + (N - |J| - k + \ell + 1)\mathcal{D}_{k-1}^{\ell,\,J}(G) \\ &\qquad + (k - \ell + 1)\mathcal{D}_{k+1}^{\ell,\,J}(G) + (\ell + 1)\mathcal{D}_{k+1}^{\ell+1,\,J}(G), \qquad (5.2.1) \\ T\mathcal{D}_k(G) &= (N - k + 1)\mathcal{D}_{k-1}(G) + (k + 1)\mathcal{D}_{k+1}(G). \qquad (5.2.2) \end{aligned}$$

Proof. In the cases when $\mathcal{D}_k^{\ell,\,J}(G) = 0$ by Lemma 5.6(i), the right side of (5.2.1) has only zero terms, so we restrict ourselves to $\max\{0, k + |J| - N\} \leq \ell \leq \min\{k, |J|\}$.

Applying the definition of $\mathcal{D}_k^{\ell,\,J}(G)$, and Lemma 5.8,

$$\begin{aligned} T\mathcal{D}_k^{\ell,\,J}(G) &= \sum_{W \subseteq Z\,:\,|W|=k, |W \cap J| = \ell} \Bigg(\sum_{z \in W} [G_{W-\{z\}}] + \sum_{z \notin W} [G_{W \cup \{z\}}] \Bigg) \\ &= \sum_{W \subseteq Z\,:\,|W|=k, |W \cap J| = \ell} \Bigg(\sum_{z \in W \cap J} [G_{W-\{z\}}] + \sum_{z \in W - J} [G_{W-\{z\}}] \\ &\qquad\qquad + \sum_{z \in Z - W - J} [G_{W \cup \{z\}}] + \sum_{z \in J - W} [G_{W \cup \{z\}}] \Bigg). \end{aligned}$$

Each of the four terms gives the corresponding term in (5.2.1). For example, in the first term $W - \{z\}$ has size $k-1$ and intersects J in $\ell - 1$ elements, with each such set being counted $|J| - \ell + 1$ times. Equation (5.2.2) is the special case $|J| = \ell = 0$. □

The recurrences in Lemma 5.9 take the form of the binary Krawtchouk polynomials with a change of variable $\alpha_k(x, n) = P_k^n\big((n - x)/2\big)$. Explicitly, $\alpha_k(x, n)$ are given by $\alpha_k(x, n) = 0$ if $k < 0$, $\alpha_0(x, n) = 1$, and

$$k\,\alpha_k(x, n) = x\,\alpha_{k-1}(x, n) - (n - k + 2)\alpha_{k-2}(x, n). \qquad (5.2.3)$$

For example, we have $\alpha_0(x, n) = 1$, $\alpha_1(x, n) = x$, $\alpha_2(x, n) = x^2/2 - n/2$, and $\alpha_3(x, n) = x^3/6 - \big(n/2 - 1/3\big)x$. Equation (5.2.3) applies for nonpositive k as well, so the following holds for all integers k.

Lemma 5.10. $\mathcal{D}_k(G) = \alpha_k(T, N)\,[G]$.

While previous applications of the Krawtchouk polynomials have been through linear algebraic properties and roots, our strategy to extract information from this relationship is to exploit known identities for the Krawtchouk polynomials and translate these to the language of decks. To facilitate this, we record some of properties of the Krawtchouk polynomials below, which can be found or follow from results in [101, 175]. Combining these with the switching axioms make it possible to derive many more identities relating decks and the T operator.

Lemma 5.11. *The following hold for all $k, n \in \mathbb{Z}$.*

(a) *$\alpha_k(x,n)$ is a polynomial in x. For $k \geq 0$, the leading term is $x^k/k!$. If k is odd, then $\alpha_k(x,n)$ is an odd polynomial, while if k is even then $\alpha_k(x,n)$ is an even polynomial with constant term*

$$(-1)^{k/2}\binom{n/2}{k/2}.$$

(b)

$$\alpha_k(x,n) = \sum_{j=0}^{k}(-2)^j\binom{n-j}{n-k}\binom{(n-x)/2}{j},$$

and, in particular, for all integers $q \geq 1$,

$$\alpha_{n+q}(x,n) = (-1)^{n+q}2^n\sum_{\ell=1}^{q}\binom{q-1}{\ell-1}\binom{(n-x)/2}{n+\ell}.$$

(c) *We have formally*

$$\sum_{k\geq 0}\alpha_k(x,n)\,y^k = \frac{(1+y)^{(x+n)/2}}{(1-y)^{(x-n)/2}}\,.$$

(d)

$$(n-x)\alpha_k(x{+}2,n) = 2(n-2k)\alpha_k(x,n) - (n+x)\alpha_k(x{-}2,n),$$
$$(n-k+1)\alpha_k(x{+}1,n{+}1) = (2n-2k+x+1)\alpha_k(x,n) - (n+x)\alpha_k(x{-}1,n{-}1).$$

(e) *For all $q \in \mathbb{Z}$,*

$$\alpha_k(x,n) = \sum_{\ell=0}^{\lfloor k/2\rfloor}(-1)^{\ell}\binom{q/2}{\ell}\alpha_{k-2\ell}(x,n{-}q),$$

and in particular (for $q = 2$)

$$\alpha_k(x,n) = \alpha_k(x,n{-}2) - \alpha_{k-2}(x,n{-}2).$$

(f) $\alpha_{n-k}(x,n) - \alpha_k(x,n) = 0$ *for* $x = n-2-4j$, *when* $0 \le j \le \lfloor (n+1)/2 \rfloor$, *and*
$\alpha_{n-k}(x,n) + \alpha_k(x,n) = 0$ *for* $x = n-4j$, *when* $0 \le j \le \lfloor n/2 \rfloor$.

Lemma 5.9 allows us to write the multisets $\mathcal{D}_k^{\ell,\,J}(G)$ in terms of a smaller basis using the T operator. One such basis consists of the multisets $\mathcal{D}_t^{t,\,J}(G)$ for $t \ge 0$[2]

Theorem 5.12. *For all* $G \in \mathcal{G}$, $J \subseteq Z$ *and* $k, \ell \in \mathbb{Z}$,

$$\mathcal{D}_k^{\ell,\,J}(G) = \sum_{t=0}^{|J|} \sum_{s=0}^{t} (-1)^{t+\ell} \binom{t}{s} \binom{|J|-t}{|J|-\ell-s} \alpha_{k+t-2s-2\ell}(T, N-2|J|)\, \mathcal{D}_t^{t,\,J}(G).$$

Proof. Note that the inner summand is 0 for many values of s. The actual range that gives non-trivial values for each particular t is

$$\max(0, t-\ell) \le s \le \min\bigl(|J|-\ell, t, \lfloor (k+t-2\ell)/2 \rfloor\bigr).$$

We prove the theorem by induction on $k-\ell$. When $k < \ell$, $k < 0$ or $\ell > |J|$ there are no non-zero terms in the sum so the correct value $\mathcal{D}_k^{\ell,\,J}(G) = 0$ is given. In the remaining cases of $k = \ell$, it is routine to see that the summation has a single non-zero term, for $t = k$ and $s = 0$, which again gives the correct value.

Since the coefficient of $\mathcal{D}_{k+1}^{\ell,\,J}$ in (5.2.1) is positive, the induction will be complete if we show that our summation satisfies (5.2.1). Substitute the summation into (5.2.1) and then on the left side replace $T\alpha_{k+t-2s-2\ell}(T, N-2|J|)$ by $(k+t-2s-2\ell)\alpha_{k+t-2s-2\ell+1}(T, N-2|J|) - (N-2|J|-k-t+2s+2\ell+1)\alpha_{k+t-2s-2\ell-1}(T, N-2|J|)$, in accordance with (5.2.3). For each t and q we can now verify that the coefficient of $\alpha_q(T, N-2|J|)\, \mathcal{D}_t^{t,\,J}$ on each side is the same. □

In the case that $J = \{z\}$, we shall use the abbreviations $\mathcal{D}_k^{0,\,\{z\}}(G) = \mathcal{D}_k^{-z}(G)$ and $\mathcal{D}_k^{1,\,\{z\}}(G) = \mathcal{D}_k^{+z}(G)$. The previous theorem then specialises to the following.

Corollary 5.13. *Suppose* $G \in \mathcal{G}$, $z \in Z$ *and* $k \in \mathbb{Z}$. *Then*
$\mathcal{D}_k^{-z}(G) = -\alpha_{k-1}(T, N-2)\,[G_z] + \alpha_k(T, N-2)\,[G]$ *and*
$\mathcal{D}_k^{+z}(G) = \alpha_{k-1}(T, N-2)\,[G_z] - \alpha_{k-2}(T, N-2)\,[G]$.

We shall also consider the polynomial $\bar{\alpha}_k(x,n) = x^{-1}(\alpha_k(x,n) - \alpha_k(0,n))$. Note that $\alpha_k(0,n)$ is the constant term of $\alpha_k(x,n)$, and hence $\bar{\alpha}_k(x,n)$ is polynomial in x. If k is odd then $\alpha_k(0,n) = 0$, so in that case we have simply that $\bar{\alpha}_k(x,n) = x^{-1}\alpha_k(x,n)$. This opens up further ways to decompose decks in terms of T.

---

[2] This expression, as well as others in this section, may be a little mysterious at first glance. They were conjectured by McKay with the aid of a computer algebra program.

Lemma 5.14. *Suppose $G \in \mathcal{G}$, $z \in Z$ and $k \in \mathbb{Z}$. Then*

(a) $\mathcal{D}_k(G) = \alpha_k(T,N)\,[G]$, and in particular $\mathcal{D}_k(G) = \bar{\alpha}_k(T,N)\,\mathcal{D}_1(G)$ if k is odd.

(b) More generally, for $J \subseteq Z$

$$\sum_{0 \le i \le \min\{|J|,k\}} \mathcal{D}^{|J|-i,\,J}_{k+|J|-2i}(G) = \alpha_k(T,N)[G_J]$$

and in particular $\mathcal{D}^{-z}_{k-1}(G) + \mathcal{D}^{+z}_{k+1}(G) = \alpha_k(T,N)\,[G_z]$.

(c) $(N-k+1)\mathcal{D}_{k-1}(G) + (k+1)\mathcal{D}_{k+1}(G) = \alpha_k(T,N)\,\mathcal{D}_1(G)$.

(d) If k is odd, then $\mathcal{D}^{-z}_k(G) = -\alpha_{k-1}(T,N-2)\,[G_z] + \bar{\alpha}_k(T,N-2)\,\mathcal{D}_1(G)$ and $\mathcal{D}^{+z}_k(G) = \alpha_{k-1}(T,N-2)\,[G_z] - \bar{\alpha}_{k-2}(T,N-2)\,\mathcal{D}_1(G)$.

(e) $\alpha_k(T,N-2)\mathcal{D}_1(G) = (k+1)\mathcal{D}_{k+1} + N\sum_{i=0}^{c-1}\mathcal{D}_{k-2i-1}$ where $c = \left\lceil \frac{k}{2} \right\rceil$.

Proof. (a) $\mathcal{D}_0(G) = T^0[G] = [G]$ by condition S1, and $\mathcal{D}_1(G) = T[G]$. Using the definition of α, we can rewrite these as $\mathcal{D}_0(G) = \alpha_0(T,N)[G]$ and $\mathcal{D}_1 = \alpha_1(T,N)[G]$. Then by comparing (5.2.2) and (5.2.3), we observe that $\mathcal{D}_k$ and α_k satisfy the same recurrence relation and the result follows by induction for $k > 1$. When k is odd, $\bar{\alpha}_k(T,N)\,\mathcal{D}_1(G) = T\bar{\alpha}_k(T,N)[G] = \alpha_k(T,N)[G] = \mathcal{D}_k$.

(b) Using (a) and S3, $\alpha_k(T,N)[G_J] = \mathcal{D}_k(G_J)$. Then

$$\begin{aligned} D_k(G_J) &= \sum_{X \subseteq Z\,:\,|X|=k} [(G_J)_X] \\ &= \sum_{0 \le i \le \min\{|J|,k\}} \left(\sum_{X \subseteq Z\,:\,|X|=k, |X \cap J| = i} [(G_J)_X] \right) \end{aligned}$$

Set $W = J \triangle X$, so $|W| = |J| + k - 2|X \cap J|$. Then by S3 the last expression becomes

$$\sum_{0 \le i \le \min\{|J|,k\}} \left(\sum_{\substack{W \subseteq Z\,:\,|W| = k+|J|-2i \\ |W \cap J| = |J|-i}} [G_W] \right) = \sum_{0 \le i \le \min\{|J|,k\}} \mathcal{D}^{|J|-i,\,J}_{k+|J|-2i}(G).$$

When $J = \{z\}$, this is

$$\alpha_k(T,N)\,[G_z] = \sum_{i=0,1} \mathcal{D}^{1-i,\,J}_{k+1-2i}(G) = \mathcal{D}^{1,\,J}_{k+1}(G) + \mathcal{D}^{0,\,J}_{k-1}(G) = \mathcal{D}^{-z}_{k-1}(G) + \mathcal{D}^{+z}_{k+1}(G).$$

(c) From (5.2.2) it is enough to show that $\alpha_k(T,N)\mathcal{D}_1(G) = T\mathcal{D}_k(G)$, which follows from (a) and the independence of ordering of switches.

(d) This follows from Lemma 5.13 and the definition of $\bar{\alpha}$.

(e) Applying Lemma 5.11(e) iteratively and noting that $\alpha_k(x, n-2) = \alpha_k(x, n)$ for $k = 0, 1$, we have $\alpha_k(x, n-2) = \sum_{i=0}^{m} \alpha_{k-i}(x, n)$ where $m = \lfloor \frac{k}{2} \rfloor$. Then using (c), we can write $\alpha_k(T, N-2)\mathcal{D}_1(G)$ as

$$\begin{aligned}
&\sum_{i=0}^{m} \alpha_{k-2i}(T, N)\, \mathcal{D}_1(G) \\
&= \sum_{i=0}^{m} ((k-2i+1)\mathcal{D}_{k-2i+1}(G) + (N-k+2i+1)\mathcal{D}_{k-2i-1}(G)) \\
&= (k+1)\mathcal{D}_{k+1}(G) + (N-k+2m+1)\mathcal{D}_{k-2m-1}(G) \\
&\quad + \sum_{i=0}^{m-1} ((N-k+2i+1)\mathcal{D}_{k-2i-1}(G) + (k-2i-1)\mathcal{D}_{k-2i-1}(G)) \\
&= (k+1)\mathcal{D}_{k+1}(G) + (N-k+2m+1)\mathcal{D}_{k-2m-1}(G) + N\sum_{i=0}^{m-1} \mathcal{D}_{k-2i-1}(G).
\end{aligned}$$

If k is even, then $k - 2m - 1 = -1$ so $(N - k + 2m + 1)\mathcal{D}_{k-2m-1}(G) = 0$, and if k is odd then $(N - k + 2m + 1)\mathcal{D}_{k-2m-1}(G) = N\mathcal{D}_0$. Thus, the last expression becomes $(k+1)\mathcal{D}_{k+1}(G) + N\sum_{i=0}^{c-1} \mathcal{D}_{k-2i-1}(G)$ where $c = \lceil \frac{k}{2} \rceil$. □

We now have the ingredients needed prove some reconstructibility results. Specifically, the reconstructibility of the following expressions involving decks is established by writing them as functions (polynomials) of T and $\mathcal{D}_1(G)$. We will demonstrate shortly why these expressions are of particular interest.

Lemma 5.15. *Suppose that $N = 2m$ is even. For $G \in \mathcal{G}$, $z \in Z$ and $j \geq 0$, define*

$$\begin{aligned}
\mathcal{X}_j^{-z}(G) &:= \mathcal{D}_{2j}^{-z}(G) - (-1)^j \binom{m-1}{j} [G], \\
\mathcal{X}_j^{+z}(G) &:= \mathcal{D}_{2j}^{+z}(G) - (-1)^j \binom{m-1}{j-1} [G], \\
\mathcal{X}_j(G) &:= \mathcal{D}_{2j}(G) - (-1)^j \binom{m}{j} [G].
\end{aligned}$$

Then we have the expressions:

(a) $\mathcal{X}_j^{-z}(G) = -\alpha_{2j-1}(T, N-2)\,[G_z] + \bar{\alpha}_{2j}(T, N-2)\,\mathcal{D}_1(G)$, *and*
$\mathcal{X}_j^{+z}(G) = \alpha_{2j-1}(T, N-2)\,[G_z] - \bar{\alpha}_{2j-2}(T, N-2)\mathcal{D}_1(G)$.

(b) $\mathcal{X}_j(G) = \bar{\alpha}_{2j}(T, N)\,\mathcal{D}_1(G)$.

Proof. We first prove (b). By Lemma 5.11(a), $\alpha_{2j}(0, N)[G] = (-1)^j \binom{m}{j}$. Together with Lemma 5.14(a) this gives $\mathcal{X}_j(G) = \mathcal{D}_{2j}(G) - (-1)^j \binom{m}{j} = \alpha_{2j}(T, N)[G] - \alpha_{2j}(0, N)[G] = T(\bar{\alpha}_{2j}(T, N))[G] = \bar{\alpha}_{2j}(T, N)\,\mathcal{D}_1(G)$. Similarly for (a), we have $\alpha_{2j}(0, N-2) = (-1)^j \binom{m-1}{j}$ and then using Lemma 5.13 gives

$$\begin{aligned}
\mathcal{X}_j^{-z}(G) &= \mathcal{D}_{2j}^{-z}(G) - (-1)^j \binom{m-1}{j} [G] \\
&= -\alpha_{2j-1}(T, N-2)\,[G_z] + \alpha_{2j}(T, N-2)\,[G] - \alpha_{2j}(0, N-2) \\
&= -\alpha_{2j-1}(T, N-2)\,[G_z] + T(\bar{\alpha}(T, N-2))\,[G] \\
&= -\alpha_{2j-1}(T, N-2)\,[G_z] + \bar{\alpha}_{2j}(T, N-2)\,\mathcal{D}_1(G).
\end{aligned}$$

For the second statement of (a), observe that $\mathcal{X}_j^{-z}G) + \mathcal{X}_j^{+z}(G) = \mathcal{X}_j(G)$. □

As a note of comparison, the earlier usage of algebraic methods in switching reconstruction (as outlined in Section 5.2.1) were centred around establishing invertibility of certain operators to then reconstruct desired quantities directly. While this method is powerful, in practice its applicability has been limited by the calculations it entails. The general method we have put forward in this section is an attempt to sidestep this issue. We try to reconstruct certain combinations of decks that come out of the algebra, and find structures in these that allows us to pick out our graphs of interest. This can be viewed as a broad generalisation of the Alon-Coppersmith argument.

5.2.3 Deriving generic reconstructibility conditions

The lemmas in the previous section now allow us to quite easily derive several sufficient conditions for generic switching reconstructibility.

Theorem 5.16. *Suppose $G \in \mathcal{G}$. Then the following hold.*

(a) If N is odd, then G is reconstructible.

(b) If $N \equiv 2 \pmod 4$, then $[G] \cup [G_Z]$ is reconstructible. In particular, if $G_Z \sim G$ then G is reconstructible.

(c) If $N \equiv 0 \pmod 4$, then G is reconstructible if $G_Z \not\sim G$.

Proof. (a) By Lemma 5.14(c) with $k = 1$, $\mathcal{D}_1(G)$ determines $N\mathcal{D}_0(G) + 2\mathcal{D}_2(G)$. Since N is odd and $\mathcal{D}_0(G) = [G]$, we can identify $[G]$ as the only element appearing in $N\mathcal{D}_0(G) + 2\mathcal{D}_2(G)$ with odd multiplicity. Another proof is that from Lemma 5.14(a) we can reconstruct $\mathcal{D}_N(G) = \{[G_Z]\}$ and then apply Lemma 5.7 to recover $[G]$.

(b) Suppose $N = 4t+2$. If $G_Z \sim G$, Lemma 5.7 shows that $\mathcal{D}_{2t}(G) = \mathcal{D}_{2t+2}(G_Z) = \mathcal{D}_{2t+2}(G)$. Therefore, by Lemma 5.14(c) with $k = 2t + 1$, $\mathcal{D}_1(G)$ determines $(N + 2)\mathcal{D}_{2t}(G)$ and thus $\mathcal{D}_{2t}(G)$. Since by Lemma 5.15(b) we can determine $\mathcal{X}_t(G) = \mathcal{D}_{2t}(G) - (-1)^t\binom{2t+1}{t}[G]$ from $\mathcal{D}_1(G)$, we can solve for $[G]$.

Now suppose $G_Z \not\sim G$. Define the relation "$\approx$" by $G \approx G'$ if $G \sim G'$ or $G_Z \sim G'$. By Lemma 5.14(a) with $k = N$, $G \sim G'$ implies $G_Z \sim G'_Z$, and by Lemma 5.7, $(G_Z)_Z \sim G$. Therefore $\approx$ is an equivalence relation. Since $G_Z \approx G$, it follows from the first part that $[G] \cup [G_Z]$ is reconstructible.

(c) In this case, Lemma 5.15(b) says that we can reconstruct $\mathcal{X}_{N/2} = \mathcal{D}_N(G) - \mathcal{D}_0(G) = \mathcal{D}_0(G_Z) - \mathcal{D}_0(G)$. If $G_Z \not\sim G$, $[G]$ is the sole member of $\mathcal{X}_{N/2}$ with negative multiplicity. □

With Stanley's switchings, we do have $G_Z \sim G$, so statements (a) and (b) of the preceding theorem essentially give an alternative derivation of the Alon-Coppersmith argument. Unfortunately though, this also means that the open cases of the Switching Reconstruction Conjecture are not handled by case (c).

Theorem 5.17. *Suppose that $N = 2m$ and $G \in \mathcal{G}$ is not reconstructible. Then, for each even j, there are at least $\binom{m}{j}$ sets $W \subseteq Z$ such that $|W| = 2j$ and $G_W \sim G$. Moreover, for any $z \in Z$, at least $\binom{m-1}{j-1}$ such sets include z and at least $\binom{m-1}{j}$ don't. In total, there are at least 2^{m-1} sets $W \subseteq Z$ such that $G_W \sim G$, including at least 2^{m-2} sets which include z and 2^{m-2} which don't.*

Proof. These claims follow from Lemma 5.15. For example, if there are fewer than $\binom{m}{j}$ sets W with $|W| = 2j$ and $G_w \sim G$, then $[G]$ is the unique element of $\mathcal{X}_j$ with negative multiplicity. For the totals, recall the identity $\sum_{k\geq 0}\binom{K}{2k} = 2^{K-1}$. □

Given a switching operation $(\phi, \sim)$, let us define the *redundancy* of $G \in \mathcal{G}$ to be $\rho(G) = \max_{G' \sim G}\left|\{W \subseteq Z : |W| \text{ is a multiple of 4 and } G_W = G'\}\right|$, that is, the maximum number of multiple-of-4 switchings producing the same (not just equivalent) result that is equivalent to G. This brings us to the following sufficient condition for switching reconstructibility.

Theorem 5.18. *$G \in \mathcal{G}$ is reconstructible if*

$$\left|\{W : W \subseteq Z, G_W \sim G, \textit{and } |W| \textit{ is divisible by 4}\}\right| < 2^{N/2-1}.$$

In particular, this is true if $\rho(G)\,|[G]| < 2^{N/2-1}$.

Proof. This follows from Theorem 5.16 if N is odd and Theorem 5.17 if N is even. □

We can obtain concrete conclusions from the previous results by applying them to specific instances of switching reconstruction problems. To illustrate this, we prove a sufficient condition for cycle reconstructibility in the 2-edge-coloured graphs problem which is specified by the following parameters:

- $\mathcal{G}$: All 2-edge-colourings of an underlying graph on n vertices.
- Z: The vertex set of the underlying graph.
- ϕ: For $G \in \mathcal{G}$ and $z \in Z$, let G_z be the graph obtained from G by flipping the colours of the edges incident to z. Extend this to $\mathcal{G} \times 2^Z$ in the usual way.
- $\sim$: Two elements of $\mathcal{G}$ are equivalent if they are related by an automorphism of the underlying graph.

The usual collection of nice basic properties hold here, in particular, that $(G_X)_Y = G_{X \triangle Y}$. This means if we have $W \subseteq Z$ such that $G_W = G'$ then $(G')_W = G$, so for each $V \subseteq Z$ with $G_V = G'$, we have $G_{V \triangle W} = (G_V)_W = (G')_W = G$. Hence, in the definition of redundancy we only need to consider $G' = G$.

Let G be an n-vertex cycle with a 2-edge-colouring. If $n \not\equiv 0 \pmod 4$ then it is reconstructible by Theorem 5.16, so we may assume $n \equiv 0 \pmod 4$. It is clear that $G_\varnothing = G$ and $G_Z = G$, and both of these sets have size divisible by 4 so $\rho(G) \geq 2$. In fact, if $\varnothing \subsetneq W \subsetneq Z$, then as G is connected, there is some edge with exactly one endpoint in W that changes colour in G_W, meaning $G_W \neq G$. Thus, $\rho(G) = 2$. By the orbit-stabiliser theorem $|[G]| = \frac{|\operatorname{Aut}(C_n)|}{|\operatorname{Aut}(G)|} = \frac{2n}{|\operatorname{Aut}(G)|}$ since the automorphism group of a cycle is the corresponding dihedral group. Theorem 5.18 now says that G is reconstructible if $2 \cdot 2n < 2^{N/2-1}|\operatorname{Aut}(G)|$ which is certainly true if $4n < 2^{n/2-1}$. This inequality holds whenever $n \geq 14$, and the $n = 13$ case is covered by Theorem 5.16.

Theorem 5.19. *Every 2-edge-colouring of a cycle of order at least 13 is 2-colour switching reconstructible.*

The same working can provide similar results for other classes of graphs and other problems, assuming the size of the automorphism groups are known or at least have good upper bounds. Another example for this problem is when the underlying graph is 3-connected and planar, in which case we find that any colouring is reconstructible under the colour switching operation provided the order of the underlying graph is at least 13.

There are certainly some loose ends in this chapter, in terms of results and identities in the previous section that we have not yet translated to reconstructibility results. Nonetheless, we are optimistic that this method shows potential.

Chapter 6

Next steps

The work in this section, together with the trajectory of recent work in the classical reconstruction problem, point to numerous possible next steps of varying levels of ambitiousness.

Our progress on reconstructing trees verified Nýdl's conjecture (Conjecture 3.5) up to the linear constant for sufficiently large n, which leaves the immediate question of what constant is correct asymptotically or for large enough graphs. On one hand, we feel that it would be unusual for the counterexample in the $n = 13$ case to be totally isolated, but we have not so far seen a straightforward way to generalise it to arbitrarily large counterexamples. At the same time, we are quite certain that the bound we have proved in Theorem 1.1 is not tight. A possible avenue for improvement could come from finding a way of using Lemma 4.4 on counting extensions effectively, as we current only use the very simple special case for leaf extensions. From the other side, given the simplicity of the current lower bound examples, it may be a worthwhile endeavour to try and improve these.

Most of the bounds on the size of cards that suffice to reconstruct certain parameters and classes recounted in Section 3.2 are far from tight, but some new ideas would be required to make anything more than incremental improvements (these may be possible, but are likely to take a lot of work). Deviating slightly from the past literature, it would be interesting to determine how large ℓ needs to be in order to recognise k-colourability of a graph on n vertices from its ℓ-deck. A first goal would be to pinpoint the threshold for recognising whether a graph is bipartite (2-colourable). A lower bound of $\lfloor n/2 \rfloor$ follows from an example of Spinoza and West [157] (consider a path and the disjoint union of an odd cycle and a path). Manvel [118] proved that the $(n-2)$-deck suffices, but it seems likely that it should be possible to determine whether a graph is bipartite when a linear number of vertices are deleted. For fixed

k, it may even be true that k-colourability is recognisable from the cn-deck for some $c = c(k) < 1$.

For reconstructing degree sequence (and other parameters) from missing cards, the current major goal is to verify Conjecture 3.12 for general graphs. To start with, one should certainly ameliorate the rather troubling fact that we still do not know how to reconstruct the degree sequence of a graph when 2 cards are missing from its deck. It seems that we could push this through with sufficient effort and casework, but it is not clear how much we could learn from doing so. The bounded average degree assumption that allowed us to obtain much better results in Section 4.3 are not readily done away with, so ideally we would develop some other efficient approach. We also note that since reconstructing the number of edges is a natural barrier to obtaining degree sequence, it may even be possible to go beyond Conjecture 3.12 and eventually show that the degree sequence is determined by any $n - \sqrt{n}$ cards.

Of all the major open reconstruction variant conjectures that we have listed, we feel strongly that Stanley's Switching Reconstruction Conjecture is the closest to being resolved. Our new methods have come tantalising close to making real progress toward this; for instance, Theorem 5.16 is a near-miss because we cannot distinguish between G and the graph G_Z switched at every vertex. However, the fact that we have unused algebraic structure may be evidence that they can be pushed further in this direction. In particular, we would really like to connect Theorem 5.12 to reconstructibility, as we are currently only using special cases given in Corollary 5.13.

Lastly, we circle back to the classical vertex reconstruction problem. The most pressing matter here is to somehow surpass the last major result of recognising planar graphs. It would be wonderful to able to reconstruct planar graphs, but this is a long-recognised sticking point for which we still have relatively few leads. A more feasible route could be toward recognising graphs embeddable on a fixed surface, or even general minor-closed classes. In light of work in structural graph theory where such classes have been intensively studied, an approachable related problem is to reconstruct the treewidth of a graph. This is with an eye to reconstructing partial k-trees for some small k, but knowing such parameters also seems potentially rewarding as it could open doors to applying an array of existing structural techniques. For a better idea of these concepts from structural graph theory, we refer to Part II.

Part II

Structure

Chapter 7

Background

The area of structural graph theory provides insight into the ways in which we can encapsulate the structure of particular classes of graphs. A candidate characterisation can be judged on its balance of two main criteria: first, that the ingredients it uses are as simple as possible to improve tractability, and second, that it still retains sufficient essential features of the graph class to prove results that are specific to it or to a slightly broader superclass (as opposed to general statements for all graphs).

Many structural theorems describe graph classes of interest by listing substructures that they either must all contain or cannot contain. The importance of this approach was established by extremely successful systematic studies, from which very general results of this type materialised. However, this is by no means the only way to proceed. Our main focus is on graph product structure theory, where results take a different form. Rather than looking within a graph for simpler pieces that reflect its properties, the idea is to zoom out and find highly structured supergraphs that can be constructed from simple building blocks. Despite being a fairly recent development, product structure theory has expanded rapidly and already proved to be extremely fruitful in terms of application to longstanding problems.

Organisation. To motivate this line of work, the remainder of this chapter is given to discussing the more classical graph minor structure theory (Section 7.1) and a key related structural invariant (Section 7.2). This leads to the statements of the main product structure theorems (Section 7.3). We then give a new application of product structure theory to studying the comparable box dimension of minor-closed classes of graphs in Chapter 8. Lastly, in Chapter 9 we push the existing theory a little further by introducing a new parameter that provides a way of describing graphs of bounded treewidth via a product structure.

7.1 Some minor motivation

In this part of the thesis, we will always work with simple, finite, undirected graphs. Let G be such a graph. Our starting point is to make precise several standard reduction operations. We have already worked with the notions of deletion; *deleting a vertex* $v \in V(G)$ means to take the induced subgraph $G - v := G[V(G) \setminus \{v\}]$ with vertex set $V(G) \setminus \{v\}$ and edge set $\{e \in E(G) : v \notin e\}$, while *deleting an edge* $e \in E(G)$ gives the spanning subgraph $G - e$ with vertex set $V(G)$ and edge set $E(G) \setminus \{e\}$. So far, we have generally understood 'substructure' to mean a subgraph or induced subgraph, where H is a *subgraph* of G if it can be obtained from G by vertex and edge deletions, and it is an *induced subgraph* of G if it can be obtained by only vertex deletions. Both of these relations define a partial order on graphs. A class of graphs that is closed under taking induced subgraphs is a *hereditary class*, and it is *monotone* if it is closed under taking subgraphs.

By allowing other reductions, we obtain more general relations on graphs. An *edge contraction* of $e = \{v, w\} \in E(G)$ forms the graph G/e with vertex set $V(G) \setminus \{w\}$ and edge set $E(G) \setminus \{e \in E(G) : w \in e\} \cup \{\{u, v\} : \{u, w\} \in E(G), v \neq u\}$. A contraction is *topological* if at least one of the endvertices v or w of the edge has degree 2 in G. This can be viewed as an inverse to taking a *subdivision at an edge* $e = \{v, w\} \in E(G)$, which creates the graph $G \cdot e$ with vertices $V(G \cdot e) = V(G) \cup \{u\}$ where u is a new vertex, and edges $E(G \cdot e) = E(G) \setminus \{e\} \cup \{\{v, u\}, \{u, w\}\}$.

We say that a graph is a *minor* of G if it can be obtained from a subgraph of G by a sequence of edge contractions. It is an *induced minor* if we use an induced subgraph, and *topological minor* if all of the contractions are topological. It is worth noting that the order of deletions and contractions do not affect the isomorphism class of the final graph. This last definition has a common alternative formulation. We say that a graph is a *subdivision* of G if it can be obtained from G by a sequence of edge subdivisions. Then H is a topological minor of G if and only if G contains a subdivision of H as a subgraph.

Both topological minors and (induced) subgraphs are minors. However, neither converse is true. For example, $K_{3,3}$ is a minor of the Petersen graph, but neither a topological minor nor a subgraph. A useful way to view graph minors is to think of performing contractions simultaneously so that if edges forming a connected subgraph are contracted, then in the minor we have one new supervertex connected to all of the neighbours of the endvertices of those original edges. The following lemma formalises this perspective.

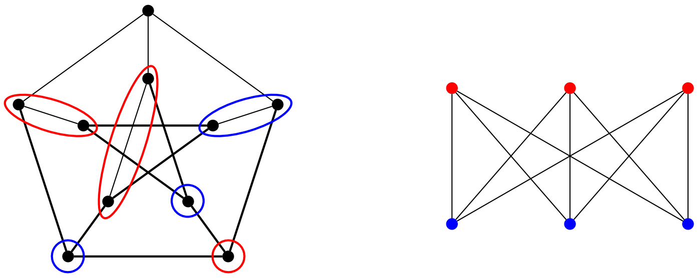


Figure 7.1: Minor model for $K_{3,3}$ in the Petersen graph

Lemma 7.1. *Let G and H be graphs. Then H is a minor of G if and only if there is a function $\mu : V(H) \to \mathcal{P}(V(G))$ such that*

- *$\mu(i)$ is nonempty and induces a connected subgraph $G[\mu(i)]$ for each i,*
- *$\mu(i) \cap \mu(j) = \varnothing$ for all $i \neq j$,*
- *for every edge $\{i, j\} \in E(H)$, there is an edge $\{x, y\} \in E(G)$ with $x \in \mu(i)$, $y \in \mu(j)$.*

The function in the preceding statement is often called a *model* of H in G. An example is given in Figure 7.1 which exhibits the non-planarity of the Petersen graph.

A class of graphs that is closed under taking minors is called a *minor-closed class*. The class of all graphs is clearly minor-closed, but it does not offer any nice structural properties so we usually restrict to *proper* minor-closed classes (that is, those for which at least one graph is missing from the class). Given a drawing of a graph in the plane without crossings, an easy observation is that the operations of deletion and edge contraction can all be performed without affecting the planarity of the embedding. It follows that planar graphs form a minor-closed class. If we restrict to topological contractions, it is even clearer that topological minors interact neatly with graph drawings since we are simply exchanging edges with induced paths. The fact that planar graphs can be characterised by topological minors cements this connection; in 1930, Kuratowski [108][1] proved that a graph is planar if and only if it does not contain a subdivision of K_5 or $K_{3,3}$. The same result with subdivisions replaced by minors was proved by Wagner not long after in 1937.

Theorem 7.2 (Wagner [177])**.** *A graph is planar if and only if it contains neither K_5 nor $K_{3,3}$ as a minor.*

---

[1] There are accounts of this theorem being independently proved by Frink and Smith, and Pontryagin all around the same time period. However, these other proofs were not published.

A result of this form, which characterises a class $\mathcal{G}$ of graphs via a finite list of graphs that cannot be contained as minors, is often called an excluded (or forbidden) minor characterisation of $\mathcal{G}$. Wagner's theorem is in some sense weaker than Kuratowski's theorem because topological minors are also minors, although it is not difficult to prove that these two statements are equivalent. At the same time, the greater generality of minors turn out to be more conducive than subdivisions when it comes to finding generalisations of these theorems. Indeed, Wagner's theorem was a precursor to the following fundamental conjecture[2],

Conjecture 7.3 (Wagner (see [113])). *Every minor-closed class of graphs can be characterised by a finite list of excluded minors.*

The natural generalisation of planar graphs is the class of graphs embeddable on a fixed surface, which provides a family of minor-closed classes. Indeed, the first significant step toward Conjecture 7.3 after its formulation was a theorem of Archdeacon and Huneke [5] from 1989 that for every non-orientable surface S, the class of graphs embeddable on S has an excluded minor characterisation. There is already a noteworthy difference between this result and the one for planar graphs, as the proof of the preceding theorem does not actually give the finite list of excluded minors explicitly, it only guarantees the existence of such a list. This is going to be a theme for the results to come.

In 1990, Robertson and Seymour [148] was able to extend this to orientable surfaces as well, proving that the class of graphs embeddable on any fixed surface has an excluded minor characterisation. This appears in Part VIII of their seminal series 'Graph Minors' (starting 1983), which together contain a wealth of results en route to one of their primary goals of verifying Conjecture 7.3. These papers continue to be extremely influential on the direction and form of modern structural results. In Part XX, published in 2004, they attained this goal in what is widely celebrated as one of the greatest achievements in graph theory. We record a common equivalent formulation.

Theorem 7.4 (Graph Minor Theorem [152]). *Every infinite set* $\{G_1, G_2, \ldots\}$ *of graphs contains* G_i *and* G_j *such that* G_i *is a minor of* G_j*.*

---

[2] This conjecture apparently only appears in a textbook of Wagner published in 1970 (see [113]), but was first formulated some time before this. According to Diestel [34, Chapter 12 Notes], Wagner denied making it.

In other words, this statement says that graphs are well-quasi-ordered[3] under the minor relation. From this perspective, the graph minor theorem is also a far-reaching generalisation of the result that finite trees are well-quasi-ordered with respect to the minor relation, which was independently proved in 1960 by Kruskal [106] and Tarkowski [168] after being conjectured as early as 1937 by Vázsonyi (see [107]). Another reason that this theorem is so significant is that it has major algorithmic implications. Expounding on this, Robertson and Seymour [150] also proved that, for a fixed graph H, there is a polynomial time algorithm to test whether a graph contains H as a minor, so it follows that any minor-closed property can be tested in polynomial time as well.

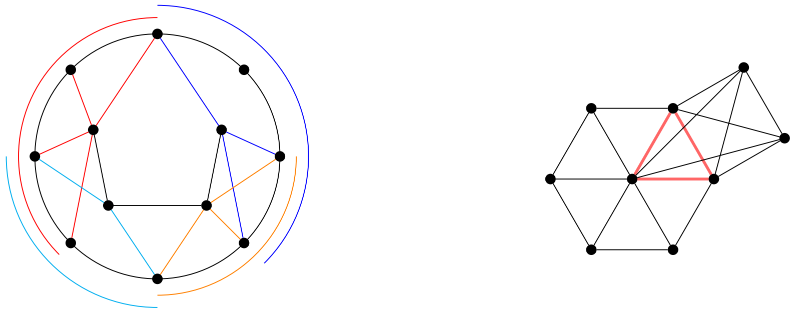
While it is difficult to overstate the importance of Theorem 7.4, it is not without limitations. One is that the proof is very involved and consequently opaque. A related point that we have already mentioned is the fact that it does not give an explicit list of excluded minors for any class. This tempers the discussion of complexity above, since it means that we do not automatically get an explicit algorithm at all. A more serious issue, in the sense that it cannot be overcome, is that even if there were a method to obtain the list of excluded minors they would be huge. Although two obstructions suffice for planar graphs, 35 are already needed for projective planar graphs [4] which is the simplest non-orientable surface, and no other exact lists have been obtained for any other surface. It is known, however, that there are more than 17500 for the torus [127], which hints at the scale for surfaces of higher genus. The size of the list of excluded minors for a class translates directly into the constant factor in the running time of algorithms, which again diminishes practicality.

In light of this last point that exact characterisations of graph classes are unavoidably complicated, we turn instead to approximate characterisations; that is, descriptions which cover all of the graphs in our class but are not 'if and only if' statements. The main result of Graph Minors XVII, which is one of the major stepping stones toward Theorem 7.4, provides a remarkable result of this type. In order to describe it, we need to define a few more graph operations.

First, starting with some graph G, *adding an apex vertex* means to add one new vertex together with edges to any arbitrary subset of $V(G)$. When we add multiple apices, the new vertices are also allowed to be joined to each other arbitrarily. Next, assuming that we have a 2-cell embedding of G (so that each face is homeomorphic

---

[3] A quasi-order is a reflexive and transitive binary relation. A well-quasi-order is a quasi-order for which every infinite sequence of elements contains a pair of comparable elements $x_i \leq x_j$ with $i < j$. It follows from a Ramsey argument that this is equivalent to there being no infinite antichains or infinite descending chains.



(a) Adding a vortex of width 2. (b) A full 3-clique sum of W_7 and K_5

Figure 7.2: Two operations

to an open disk) into a surface, we can *add a vortex of width at most k to a face F* by the following procedure (see Figure 7.2a):

(i) Take a family of arcs (connected subpaths of the boundary cycle of F) such that every vertex on the cycle is in at most k arcs.

(ii) For each arc A, create a vertex v_A inside F.

(iii) Arbitrarily connect v_A to any subset of vertices on A.

(iv) Arbitrarily include any subset of edges from $\{v_A v_B : A \cap B \neq \varnothing\}$.

The third operation is to take a *k-clique-sum* of two graphs G and H (Figure 7.2b):

(i) For some $m \leq k$, choose a copy of K_m in G and a copy of K_m in H.

(ii) Glue G and H by identifying the chosen copies of K_m via any bijective map.

(iii) Delete any (or none) of the edges in the resulting single copy of K_m.

When no edges are deleted in the final step, we call this a *full k-clique-sum*. Now let $\mathcal{L}_k$ be the family of graphs that can be constructed as follows:

(i) Take a graph G with 2-cell embedding into a surface of genus at most k.

(ii) Choose at most k faces, and to each one add a vortex of width at most k.

(iii) Add at most k apex vertices.

(iv) Repeatedly take clique-sums with the graph that we have so far and any graph that arises from steps (i)–(iii).

For a fixed graph H, let $\mathcal{N}_H$ be the class of graphs with no H minor. The Graph Minor Structure Theorem then says that all H-minor-free graphs can be constructed as above.

Theorem 7.5 (Graph Minor Structure Theorem [151])**.** *For every graph H, there exists an integer $k = k(H) > 0$ such that $\mathcal{N}_H \subseteq \mathcal{L}_k$. In addition, for every integer $k > 0$ there exists a graph H such that $\mathcal{L}_k \subseteq \mathcal{N}_H$.*

There are many variants of this result for specific cases where we can skip some of the steps in creating $\mathcal{L}_k$ so that it is replaced by a smaller class, or start with a surface where H does not necessarily embed. As with Theorem 7.4, this result is also very general and has the advantage that the operations we have used are relatively easy to work with. However, we now have a different limitation. The building blocks involved are graphs on surfaces, which are still complicated structures. By comparison, an earlier approximate characterisation for planar graphs from Graph Minors V sets the standard for simplicity.

Theorem 7.6 (Robertson and Seymour [147])**.** *For every planar graph H, there is an integer $k = k(H) > 0$ such that every graph in $\mathcal{N}_H$ has treewidth at most k. In addition, for every $k > 0$ there is a planar graph H such that every graph with treewidth at most k is in $\mathcal{N}_H$.*

To convey why this provides a simpler structural description, we now shift to a discussion of treewidth.

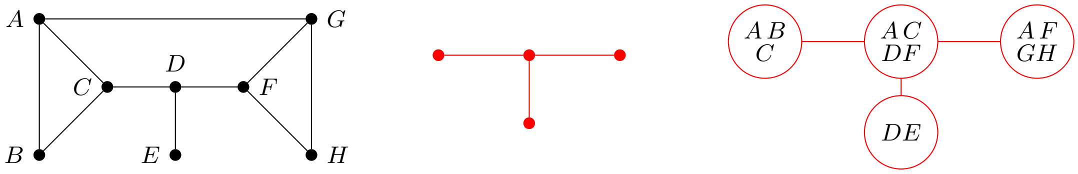
7.2 Treewidth

In this section, we define the treewidth parameter, discuss some of its basic properties, and look at what kinds of graphs have high treewidth to simultaneously indicate why graphs of small treewidth are considered nice. A good deal of the basic knowledge surrounding treewidth was developed within the series of papers of Robertson and Seymour highlighted in the previous section, particularly Graph Minors III [145]. Indeed, treewidth played a crucial role in the development of Theorem 7.5. We refer to [11, 141, 82] for detailed surveys of the history of treewidth and its applications.

There are several different ways to define treewidth. We begin with the most direct. Given a tree T, a *T-decomposition* of a graph G is a collection $\mathcal{W} = (W_x \colon x \in V(T))$ of subsets of $V(G)$ indexed by the vertices of T such that

(i) for every edge $vw \in E(G)$, there exists a vertex $x \in V(T)$ with $v, w \in W_x$; and

(ii) for every vertex $v \in V(G)$, the set $\{x \in V(T) \colon v \in W_x\}$ induces a (connected) subtree of T.



G T T-decomposition of G

Figure 7.3: A tree-decomposition with width 3

An example is given in Figure 7.3. Each set W_x in $\mathcal{W}$ is called a *bag*. The *width* of $\mathcal{W}$ is $\max\{|W_x|\colon x \in V(T)\} - 1$. A *tree-decomposition* is a T-decomposition for any tree T. The *treewidth* $\mathrm{tw}(G)$ of a graph G is then the minimum width of a tree-decomposition of G.

When working with tree decompositions, it is often convenient to have a slightly reinterpreted version of our original definition in mind. In a T-decomposition of G, every vertex v of G corresponds to a subtree T_v of T with the property that if $uv \in E(G)$ then T_u and T_v intersect. This is often useful in conjunction with the Helly property for subtrees, which says that if $\mathcal{F}$ is a collection of subtrees of a fixed tree T, then either there are two disjoint subtrees or there exists a vertex v that hits all of the subtrees (more generally, if there is no set of $\ell + 1$ disjoint subtrees then there is a hitting set consisting of ℓ vertices).

Treewidth has been a fixture not only of structural graph theory but also graph algorithms. It was first introduced as one of a family of parameters defined by Halin [73] in 1976 in the course of approaching Hadwiger's conjecture. The key property that interested Halin is that treewidth is monotone under taking minors; that is, if H is a minor of G then $\mathrm{tw}(H) \leq \mathrm{tw}(G)$. This means that the class of graphs with treewidth at most some fixed k is a minor-closed class. It is worth noting that this monotonicity property also holds for subgraphs, and if G' is a subdivision of G then $\mathrm{tw}(G) = \mathrm{tw}(G')$. All of these properties are straightforward to prove directly from the definition by operating on bags. Treewidth was then rediscovered by Proskurowski [136] in 1980 in the context of graph algorithms, where high interest came from the realisation that many problems which are NP-complete for general graphs have polynomial-time algorithms when restricted to graphs of bounded treewidth[4]. These include Hamiltonicity, Graph Isomorphism and Vertex Colouring (see [10]).

Informally, the idea of treewidth is that it measures how similar a graph is to a tree. Indeed, a graph has treewidth at most 1 if and only if it is a forest. Cycles

---

[4] A particularly effective strategy for such problems has been to use dynamic programming algorithms on tree decompositions.

have treewidth 2, and graphs with treewidth at most 2 are exactly the series-parallel graphs (K_4-minor-free). On the other hand, the complete graph K_n has treewidth $n-1$, so complete graphs do not have bounded treewidth. This comes from the fact (which follows easily from the Helly property) that in every tree-decomposition of a graph G, each clique of G appears in some bag. Hence, more generally we have that $\mathrm{tw}(G) \geq \omega(G) - 1$, and the treewidth of a k-clique-sum of two graphs G_1 and G_2 is at most $\max\{\mathrm{tw}(G_1), \mathrm{tw}(G_2)\}$, with equality if we take a full clique-sum. There is an equivalent definition of treewidth that brings to light a deeper connection between treewidth and clique-sums. A *k-tree* is any graph obtained by repeated full clique-sums on cliques of size k from cliques of size at most $k+1$. These are precisely the maximal graphs of treewidth k, and $\mathrm{tw}(G)$ is precisely the minimum k such that G is a subgraph of a k-tree. Subgraphs of k-trees are sometimes called *partial k-trees*, which gives another name for the class of treewidth at most k.

The discussion above tells us that having high clique number (or a large clique minor) is one reason to have high treewidth, which raises the questions: how else we can show that a graph has high treewidth, and what graphs have high treewidth? For the former, another parameter that we can use in a similar way to clique number is the chromatic number. Recall that a graph G is *d-degenerate* if every non-empty subgraph of G has a vertex of degree at most d. This provides a measure of sparsity and leads to an easy upper bound on chromatic number. When working with optimal tree-decompositions, observe that we can always assume no bag is a subset of another (if not, then $W_x \subseteq W_{x'}$ for some edge $xx' \in E(T)$, so we may as well contract this edge in T and assign the larger bag $W_{x'}$ to the resulting single vertex). It then follows that if we have a T-decomposition of G with width $\mathrm{tw}(G)$ and ℓ is a leaf in T, there is a vertex $v \in W_\ell$ such that $\Gamma(v) \subseteq W_\ell$ so $|\Gamma(v)| \leq |W_x| - 1 \leq \mathrm{tw}(G)$. Hence, every graph is $\mathrm{tw}(G)$-degenerate, and we have $\chi(G) \leq \mathrm{tw}(G) + 1$.

Another way to view treewidth is that it provides a more global measure of connectivity. By observing that a graph of degeneracy k is at most k-vertex-connected, we see that the (vertex-)connectivity of a graph is always bounded above by its treewidth. This can also be seen directly; given a T-decomposition of G with width k and in which no bag contains another, then for any $xy \in E(T)$, the adhesion set $W_x \cap W_y$ (which has size at most k) is a vertex cutset in G. Vertex cutsets are commonly called *separators* in this context, and having low treewidth is characterised by the existence of certain 'good' separators. Specifically, given a graph G with n vertices, a set of vertices X is a *balanced separator* if every component of $G \setminus X$ has at most $\frac{2}{3}|V(G \setminus X)|$ vertices. For $W \subseteq V(G)$, a set X is a *balanced W-separator* if every component of $G \setminus X$ has at most $\frac{2}{3}|W \setminus X|$ vertices in W. Such separators have important applications to algorithms as well as inductive proofs, and it is highly desirable to have balanced separators of small size. It is easy to see (by a recursive construction) that trees have balanced separators of size 1, and this argument can be extended to tree decompositions. Moreover, having balanced separators implies low treewidth (see [140]). This relationship is summarised below.

Theorem 7.7. *If G has treewidth at most k, then any $W \subseteq V(G)$ of size at least $2k+3$ has a balanced W-separator of size at most $k+1$. Conversely, if every $W \subseteq V(G)$ of size at least $2k+3$ has a balanced W-separator of size at most $k+1$, then* $\mathrm{tw}(G) \leq 4k+3$.

Examining the separators of a graph can give us an idea of its treewidth, but this is not easy in general. Instead, let us introduce a more practical characterisation. A *bramble* of a graph G is a collection $\mathcal{B} \subset \mathcal{P}(V(G))$ of subsets of $V(G)$ such that $G[B]$ is connected for every $B \in \mathcal{B}$, and for all $B_1, B_2 \in \mathcal{B}$ either $B_1 \cap B_2 \neq \varnothing$ or there is an edge with one endvertex in B_1 and the other in B_2 (in which case we say they are incident). A bramble therefore almost provides a minor model for a clique, except that the B_i do not have to be disjoint. A *hitting set* for $\mathcal{B}$ is a set $S \subseteq V(G)$ such that $S \cap B \neq \varnothing$ for all $B \in \mathcal{B}$, and the *order* of a bramble is the size of a smallest hitting set. Seymour and Thomas [156] showed that if a graph has treewidth at least k, then it contains a bramble of order $k+1$. In fact, the opposite implication also holds (and is much easier to show). This means that treewidth is tied to the *bramble number* of a graph G, denoted $\mathrm{bn}(G)$, which is the maximum order among all brambles of G.

Theorem 7.8 (Bramble theorem [156])**.** *For any graph G,* $\mathrm{bn}(G) = \mathrm{tw}(G) + 1$.

The bramble theorem provides a useful strategy for proving lower bounds on treewidth, which notably applies to the planar grids $P_n \square P_n$ [5]. Specifically, a bramble is given by taking $\{B_{i,j}\}_{i,j \in [n-1]}$, where $B_{i,j}$ is the 'cross' formed by taking a union of the paths forming row i and column j in the grid, together with row n and all remaining uncovered vertices as 2 additional sets. This has order $n+1$ ($n-1$ to hit the crosses and then two more for the two extra sets which are disjoint), so $\mathrm{tw}(P_n \square P_n) \geq n$. On the other hand, $P_n \square P_n$ has a P_{n^2-n}-decomposition with $W_s = \{s, s+1, \ldots, s+n\}$ with $s \in [n^2 - n]$, so we really have $\mathrm{tw}(P_n \square P_n) = n$. More generally, $\mathrm{tw}(P_n \square P_m) = \min(n, m)$. Note that grids are 2-degenerate, have clique

---

[5] Here, $G_1 \square G_2$ is the Cartesian product of G_1 and G_2 with vertex set $V(G_1) \times V(G_2)$ and distinct vertices (u_1, v_1) and (u_2, v_2) are adjacent if and only if $u_1 = u_2$ and $v_1 v_2 \in E(G_2)$ or $v_1 = v_2$ and $u_1 u_2 \in E(G_1)$. For example, $P_1 \square P_1 \cong \square$, which accounts for the choice of symbol.

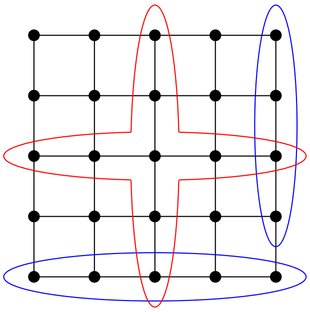


Figure 7.4: Three bramble elements on a grid.

number 2, and no K_5 minor (by planarity), so our earlier tests fail to give a good lower bound.

Planar grids are the archetypal family of graphs with unbounded treewidth in the following sense.

Theorem 7.9 (Grid theorem [147]). *For every n, there exists N such that every graph G with $\mathrm{tw}(G) \geq N$ contains $P_n \,\square\, P_n$ as a minor.*

Any class of graphs that contains the planar graphs also contains all planar grids, and therefore has (as a class) unbounded treewidth. This means that if H is not planar, then $\mathcal{N}_H$ – the class of H-minor-free graphs as in the previous section – has unbounded treewidth. Conversely, if H is planar, then since every planar graph is a minor of a sufficiently large grid there is some planar grid that is not in $\mathcal{N}_H$.

Corollary 7.10. *$\mathcal{N}_H$ has bounded treewidth if and only if H is planar.*

To conclude this section, we mention that there are myriad similar 'width' parameters quantifying closeness to some nice structure. While treewidth was one of the earliest and most prominent, other examples include branchwidth [149], pathwidth [144], rankwidth [134], twin-width [50] and flipwidth [171], all of which are intertwined to some extent.

7.3 The product structure theorems

We already saw one approximate structural characterisation in Theorem 7.5 where graphs excluding a minor are built up by enlarging and combining structures obtained from graphs on surfaces. The goal of graph product structure theory is to simplify

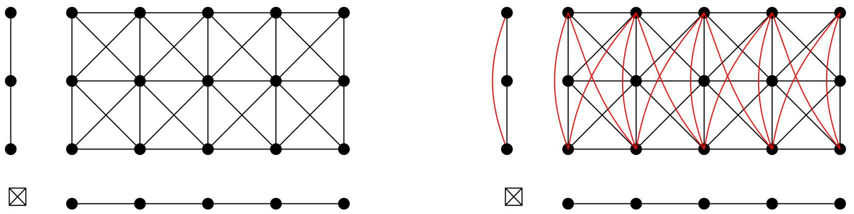


Figure 7.5: Strong products

in the other direction by finding well-structured supergraphs with graphs of bounded treewidth as building blocks.

Before we work our way to these core statements, we should clarify that the product in the name refers to the *strong product* of graphs A and B, denoted[6] $A \boxtimes B$, which has vertex-set $V(A) \times V(B)$, and distinct vertices $(v, x), (w, y) \in V(A) \times V(B)$ are adjacent whenever $v = w$ and $xy \in E(B)$, or $x = y$ and $vw \in E(A)$, or $vw \in E(A)$ and $xy \in E(B)$. The strong products of $P_3 \boxtimes P_5$ and $C_3 \boxtimes C_5$ are given as examples in Figure 7.5.

A generic product structure statement says: every graph G in a given class $\mathcal{C}$ is a subgraph of $F_1 \boxtimes F_2 \ldots \boxtimes F_k$ for some graphs F_i that either have bounded treewidth or some simple explicit structure (such as a path or complete graph). The number of factors is not a major consideration, although for many classes having $k \leq 3$ is enough. These statements give an approximate characterisation of $\mathcal{C}$, and there is no expectation that the subgraphs of such products should all be in $\mathcal{C}$.

The groundwork for product structure theory was laid by Dujmović, Joret, Micek, Morin, Ueckerdt and Wood [40] in 2019 as a side-effect of their work on showing that planar graphs have bounded queue number. Accordingly, they first approached this for planar graphs, and in slightly different language to the statements that have since been most-used. We will take this chance to introduce that earlier language since it will be useful for developing further theory.

If H is any graph, then a *H-partition* of G is a partition $(V_x \colon x \in V(H))$ of $V(G)$ indexed by the nodes of H, such that for every edge vw of G, if $v \in V_x$ and $w \in V_y$, we have $x = y$ or $xy \in E(H)$. The *width* of such a H-partition is $\max\{|V_x| \colon x \in V(H)\}$, and we call H the *quotient* of such a partition. Meanwhile, a *layering* of G is an ordered partition $(V_0, V_1, \ldots, V_\ell)$ of $V(G)$ such that for every edge $vw \in E(G)$ with $v \in V_i$ and $w \in V_j$, we have $|i - j| \leq 1$. That is, there are only edges within the parts and between consecutive parts in the ordering. The typical example of a layering to

---

[6] As with the Cartesian product, this fitting notation captures the fact that $P_1 \boxtimes P_1 = \boxtimes$.

keep in mind comes from running a breadth-first-search from a fixed vertex r where we let $V_i = \{v \in V(G) : d(r, v) = i\}$.

Putting these definitions together, the big innovation in [40] is the concept of a layered partition, which follows the slightly earlier notion of layered treewidth [67, 42]. The *layered width* of any partition of G is the minimum k such that for some layering $(V_0, V_1, \ldots, V_\ell)$ of G, each part of the partition has at most k vertices in each layer V_i. As usual, grids provide a useful tutorial example; the columns of the grid $P_n \square P_n$ are a partition of width 1 where the layering can be taken to be the rows, and the quotient in this case is P_n. Planar graphs also have partitions with low layered width.

Theorem 7.11 (DJMMUW [40]). *Every planar graph G has a partition with layered width 1 such that the quotient graph has treewidth at most 8. Also, every planar G has a partition with layered width 3 such that the quotient has treewidth at most 3.*

The proofs are somewhat involved, and proceed by inductively constructing the partition in question (as well as certifying its width) based on a known partition of planar graphs into geodesics. The following relationship follows from the definitions and illuminates the path towards the elegant formulation using strong products.

Observation 7.12. *Let G and H be graphs. Then G has a H-partition of layered width at most ℓ if and only if $G \subseteq H \boxtimes P \boxtimes K_\ell$ for some path P.*

Combining Observation 7.12 and Theorem 7.11 results in the first product structure theorem.

Theorem 7.13 (DJMMUW [40]). *Let G be a planar graph. Then G is a subgraph of each of the following:*

- *$H \boxtimes P$ where H is some graph with $\mathrm{tw}(H) \leq 8$ and P is a path.*
- *$H \boxtimes P \boxtimes K_3$ where H is some graph with $\mathrm{tw}(H) \leq 3$ and P is a path.*

The trade-off between the treewidth and size of clique in the product here is a general feature that reflects the fact that there is some flexibility in product representations. In some situations it may be more important to minimise one parameter than another, so it is natural to keep multiple statements. Mirroring the work in graph minor theory, the next step is to move to graphs embeddable on surfaces. Let $A + B$ be the *complete join* of graphs A and B, formed from the disjoint union of A and B by adding in all edges of the form ab where $a \in V(A)$, $b \in V(B)$. By using a similar partition and more complicated constructions of layerings, the same group as before prove the following.

Theorem 7.14 (DJMMUW [40])**.** *Let G be a graph of genus g. Then G is a subgraph of each of the following:*

- $H \boxtimes P \boxtimes K_{\max\{2g,1\}}$ *where H is some graph with* $\mathrm{tw}(H) \leq 9$ *and P is some path.*
- $H \boxtimes P \boxtimes K_{\max\{2g,3\}}$ *where H is some graph with* $\mathrm{tw}(H) \leq 4$ *and P is some path.*
- $(H + K_{2g}) \boxtimes P$ *where H is some graph with* $\mathrm{tw}(H) \leq 8$ *and P is some path.*

Taking this even further, it turns out that the graph operations involved in the Graph Minor Structure Theorem are compatible with the partitions and layerings constructed in the proof of Theorem 7.14. As a result, applying the Graph Minor Structure Theorem leads to a product structure theorem that covers every minor-closed class. The treewidth of the building blocks is still bounded as we want, but now depends on the specific k from Theorem 7.5 which we have seen is not explicit. Hence, we state this result in more vague terms. A *k-tree-grid* is a strong product of a k-tree (i.e. a graph with treewidth at most k) and a path. An *a-extended k-tree-grid* is a graph obtained from a k-tree-grid by adding at most a apex vertices.

Theorem 7.15. *For every proper minor-closed class $\mathcal{G}$, there exist integers k, a such that every graph $G \in \mathcal{G}$ can be obtained by taking repeated clique-sums of a-extended k-tree-grids. Equivalently, G is a subgraph of the clique sum of $G_1, G_2 \ldots, G_q$ for some q where each $G_i \subseteq (H_i \boxtimes P_i) + K_a$ with* $\mathrm{tw}(H_i) \leq k$ *and P_i a path.*

We should note that since these theorems have appeared, alternative proofs have followed (for instance in [84] for the planar case). These include strengthenings where the building blocks in the strong product are further simplified or endowed with additional structure (see [37], or [41] for discussion and some limitations).

In addition, various other examples of graphs classes that can be described this way have since been proved. In light of Theorem 7.15, this is particularly interesting for non-minor-closed classes of graphs. Dujmović, Morin and Wood [43] gave such statements for k-planar graphs that admit a drawing in which every edge is involved in at most some fixed number k of crossings, powers of graphs with bounded maximum degree, intersection graphs of curves in the plane (or another surface) called string graphs, and map graphs constructed by taking certain subgraphs of duals of graphs embedded in a surface. We refer to the notes of Dvořák, Huynh, Joret, Liu and Wood [49] for an overview up until mid-2020. Later additions to list by different groups include fan-planar graphs [83] and h-framed graphs [8].

These product structure theorems have proved to be an effective toolbox. As a general plan, using Theorem 7.15 say, if we wish to lift some property of graphs of bounded treewidth to all minor-closed classes, then it suffices to show that each of the operations involved (taking subgraphs, adding apex vertices, clique-sums) preserves this property. The main point is that this avoids having to work with vortices (or, more seriously, embedded graphs), and we are not limited to properties that hold for graphs of bounded genus. We now list a sampling of settings in which product structure theorems have facilitated progress, some long sought-after.

- The motivating question of whether planar graphs have bounded queue number was asked by Heath, Leighton and Rosenberg in 1992. Doing more than asked for, Dujmović, Joret, Micek, Morin, Ueckerdt and Wood [40] showed that graphs of genus g have queue number bounded by a function of g.
- Dujmović, Esperet, Joret, Walczak and Wood [39] proved that planar graphs have nonrepetitive colourings with a bounded number of colours, settling a conjecture of Alon, Grytczuk, Hałuszczak and Riordan from 2002.
- Bonamy, Gavoille and Pilipczuk [14] showed the existence of a graph with $n^{4/3+o(1)}$ vertices that is induced-universal for planar graphs on at most n vertices. This improved on the previous best upper bound of $n^{2+o(1)}$. Soon after, this was improved to $n^{1+o(1)}$ by Dujmović, Esperet, Gavoille, Joret, Micek and Morin [38], which is asymptotically best possible. These results can be almost equivalently phrased in terms of adjacency labelling schemes. By adapting the methods used, Esperet, Joret and Morin also proved an analogous result concerning (non-induced) universal graphs [54].
- Twin-width is another graph width parameter, this time quantifying closeness to a cograph. Bonnet, Kwon and Wood [21] studied extensions of twin-width, and prove that every proper minor-closed class (as well as fixed powers of graph in such a class) are bounded in these stronger parameters. Bekos, Da Lozzo, Hliněný and Kaufmann [8] also improved the bound on the twin-width of planar graphs from 183 to 37, and for 1-planar graphs (where each edge is allowed to participate in 1 crossing) from $O(1)$ to 80.

In fact, although many of the results recorded above are stated for planar graphs, they also extend to some of the other classes that we have mentioned admit product structures.

Chapter 8

Comparable box dimension

In this chapter, we discuss a question in geometric graph theory where we are able to effectively apply product structure theory. Let us begin with a brief background on geometric representations of graphs. Typically, this consists of a collection of geometric objects where the vertices correspond to those objects and the edges are governed by the way in which they interact. The utility of such representations is twofold. On one hand, if we know that a class of graphs can be represented a certain way geometrically, then this opens up avenues to using geometric techniques to learn more about the class. On the other hand, it is surprisingly often the case that the class of graphs admitting a particular type of representation has turned out to have interesting properties, so this is a neat way to generate candidate classes of graphs to play with when exploring problems or parameters. Recently, they have been particularly pertinent to the study of χ-boundedness in graph colouring. In addition, geometric representations of graphs have close ties to work of a more applied nature in computational geometry or computing, such as in graph visualisation and VLSI circuit design (see [72, Chapter 7]).

More formally, for a graph G, a geometric representation of G consists of a system $\mathcal{O}$ of subsets in $\mathbb{R}^d$, with some specified constraints on the geometry of these sets and the dimension, and a bijection $f : V(G) \to \mathcal{O}$ such that there is an edge $uv \in E(G)$ if and only if $f(u)$ and $f(v)$ interact in a specific way. If $f(u) \cap f(v) \neq \emptyset$ if and only if $uv \in E(G)$, then we call f an *intersection representation.* If, in addition, the sets $f(u)$ and $f(v)$ are interior-disjoint for all distinct u and v, then this is a *touching*[1] *representation.* This definition does not cover all representations in literature, but it does cover many of the most-studied ones. One example comes from touching representations by disks (or interior-disjoint circles) in $\mathbb{R}^2$. These are also called coin

---

[1] These are also called contact, kissing, or tangency representations in some circles.

representations, and were made famous by the Circle Packing Theorem first proved by Koebe in 1936.

Theorem 8.1 (Koebe [93]). *A graph is planar if and only if it has a coin representation.*

As one of the first major results on geometric representations of graphs and due to its fascinating interconnections with other areas of mathematics (see [153] for a discussion), the Circle Packing Theorem sparked a good deal of attention in the area leading to work on finding similar geometric characterisations. In 1986, Thomassen proved an analogous result using boxes instead of circles. Throughout this chapter, by 'box' we always mean an *axis-aligned box* in some specified dimension d, which is a Cartesian product $I_1 \times I_2 \times \cdots I_d \subset \mathbb{R}^d$ of closed intervals in $\mathbb{R}$ of non-zero length.

Theorem 8.2 (Thomassen [170]). *Every planar graph has a touching representation by boxes in* $\mathbb{R}^3$.

Note that this is an approximate characterisation; for instance, eight cubes with corners meeting at a point form a representation of K_8, which is non-planar. Felsner and Francis [55] showed that this is still true even when we further impose the restriction that all boxes are the product of three intervals of the same length. That is, every planar graph has a touching representation by cubes in $\mathbb{R}^3$. Allowing overlapping objects, Esperet [53] gave partial extensions of these results toward the usual more general classes from planar graphs, showing that every toroidal graph has an intersection representation by 6-dimensional boxes, and graphs embeddable in a fixed orientable surface with no short non-contractible cycles have an intersection representation by 5-dimensional boxes. Esperet also showed that for any minor-closed class $\mathcal{F}$ of graphs, there is an integer $g = g(\mathcal{F})$ such that any graph in $\mathcal{F}$ of girth at least g has an intersection representation by 3-dimensional boxes.

In these results, we begin to see that there is some trade-off possible between the dimension of the objects and the complexity of graphs that can be represented. This is captured by the notion of *boxicity* of a graph, which was defined by Roberts [143] in 1969 as the minimum dimension d such that the graph can be represented as the intersection graphs of d-dimensional boxes in $\mathbb{R}^d$. As with treewidth, this parameter caught interest as many hard computational problems can be solved much more easily on graphs with bounded boxicity than arbitrary graphs. The results we have mentioned give bounds on the boxicity for several graph classes, but showing that various classes have bounded boxicity and improving these bounds is an active area of research. We refer to [154] for an overview of recent progress.

Given that touching representations are a subclass of intersection representations, it is natural to look at the analogue of boxicity in this setting. Importantly, an attractive feature of touching representations is that it is possible to exclusively represent graph classes that are sparse, by which we mean not containing large complete bipartite graphs as subgraphs. This is in contrast to general intersection representations where the represented class always includes arbitrarily large cliques, and so gives us the chance of obtaining classes of representable graphs that possess properties particular to sparse graphs. Of course, whether the class of touching graphs of objects from a system $\mathcal{O}$ is sparse or not depends on the particular system. For example, all complete bipartite graphs $K_{n,m}$ are touching graphs of boxes in $\mathbb{R}^3$, where the vertices in one part are represented by $m \times 1 \times 1$ boxes and the vertices of the other part are represented by $1 \times n \times 1$ boxes. Dvořák, McCarty and Norin [46] noticed that this issue disappears if we forbid such a combination of long and wide boxes. This condition can be expressed as follows.

For two boxes B_1 and B_2, we write $B_1 \sqsubseteq B_2$ if B_2 contains a translate of B_1. We say that B_1 and B_2 are *comparable* if $B_1 \sqsubseteq B_2$ or $B_2 \sqsubseteq B_1$. A *touching representation by comparable boxes* of a graph G is a touching representation f by boxes such that for every $u, v \in V(G)$, the boxes $f(u)$ and $f(v)$ are comparable. Now let the *comparable box dimension* $\dim_{cb}(G)$ of a graph G be the smallest integer d such that G has a touching representation by comparable boxes in $\mathbb{R}^d$. It is easy to see that the comparable box dimension of every graph G is at most $|V(G)|$, which will follow from Lemma 8.5. For a class $\mathcal{G}$ of graphs, let $\dim_{cb}(\mathcal{G}) := \sup\{\dim_{cb}(G) : G \in \mathcal{G}\}$. The result of Thomassen mentioned earlier implies that planar graphs have comparable box dimension 3.[2] If the comparable box dimension of graphs in $\mathcal{G}$ is not bounded, we write $\dim_{cb}(\mathcal{G}) = \infty$.

Several nice properties of this notion have been proved which indicate that these definitions are good ones, and provide an incentive to bound the comparable box dimension of classes of interest (since this then tells us immediately that they satisfy those properties). Some key facts are listed below. We leave the definitions of technical terms to the cited articles, but the main point is that there is good coverage of properties that we would want in a class of sparse graphs.

---

[2] The graph of an octahedron provides an example of a planar graph that requires 3-dimensional boxes, even if intersections are allowed.

- Dvořák, McCarty and Norin [46] proved that if a class $\mathcal{G}$ has finite comparable box dimension, then it has polynomial strong colouring numbers.
- By an algorithm of Plotkin, Rao and Smith [135], the above point implies that classes with finite comparable box dimension have strongly sublinear separators.
- Dvořák, Pekárek, Ueckerdt and Yuditsky [47] showed that graphs of comparable box dimension 3 have exponential weak colouring numbers. Putting this together with a result of of Dvořák, McCarty and Norin [46] shows that these classes of graphs have polynomial strong colouring numbers and superpolynomial weak colouring numbers, the only known natural class to do so.[3]
- With Dvořák, Gonçalves, Lahiri and Ueckerdt [48], we showed that bounded comparable box dimension implies fractionally treewidth-fragility. This gives arbitrarily precise approximation algorithms for all monotone maximization problems that are expressible in terms of distances between the solution vertices and tractable on graphs of bounded treewidth [45], or expressible in the first-order logic [44]. It also directly implies the existence of sublinear separators.

Our goal for the remainder of this chapter is to explore how comparable box representations interact with various graph operations and parameters. These will be used to prove the following main theorem, which extends the line of work presented above on representability by boxes. It came out of a joint project with Dvořák, Gonçalves, Lahiri and Ueckerdt that was initiated at a Sparse Graphs Coalition meeting in 2021.

Theorem 1.7. *Every proper minor-closed class of graphs has finite comparable box dimension.*

The proof is based on the product structure theorem for minor-closed classes (Theorem 7.15), which immediately provides a plan of attack. In Section 8.1, we relate comparable box dimension to clique number and some variants of chromatic number. These are then used in Section 8.2 to show that the comparable box dimension behaves well under the operations of addition of apex vertices, strong products, taking subgraphs and clique-sums. The results in the two sections we have mentioned are derived purely by manipulating the geometric representations. Finally, we tie these results together when we apply product structure in Section 8.3.

---

[3] The question of whether such graphs exist is attributed to Joret and Wood. Dvořák, McCarty and Norin proved that for every nondecreasing function h that has values at least 3 and tends to ∞, the class of graphs with strong colouring numbers bounded by h has infinite comparable box dimension. The 'unnatural' graphs that provided the first negative answer are obtained by subdividing edges of every graph suitably many times [71].

8.1 Parameters

In this section we bound some basic graph parameters in terms of comparable box dimension. The first result bounds the clique number $\omega(G)$ in terms of $\dim_{cb}(G)$.

Lemma 8.3. *For any graph G, we have $\omega(G) \leq 2^{\dim_{cb}(G)}$.*

Proof. We may assume that G has bounded comparable box dimension witnessed by a box representation f. To represent any clique $A = \{a_1, \ldots, a_w\}$ in G, the corresponding boxes $f(a_1), \ldots, f(a_w)$ have pairwise non-empty intersections. Since axis-aligned boxes have the Helly property (see [32]), there is a point $p \in \mathbb{R}^d$ contained in $f(a_1) \cap \cdots \cap f(a_w)$. As each box is full-dimensional, their interiors each intersect at least one of the 2^d orthants at p. At the same time, it follows from the definition of a touching representation that $f(a_1), \ldots, f(a_d)$ have pairwise disjoint interiors, and hence $w \leq 2^d$. □

Note that a clique with 2^d vertices has a touching representation by comparable boxes in $\mathbb{R}^d$, where each vertex is a hypercube defined as the Cartesian product of intervals of form $[-1, 0]$ or $[0, 1]$. From this together with Lemma 8.3, it follows that $\dim_{cb}(K_{2^d}) = d$.

The remaining bounds pertain to the chromatic number $\chi(G)$ of a graph G, and two of its variants. An *acyclic colouring* (resp. *star colouring*) of a graph G is a proper colouring such that any two colour classes induce a forest (resp. star forest, i.e., a forest in which each component is a star). The *acyclic chromatic number* $\chi_a(G)$ (resp. *star chromatic number* $\chi_s(G)$) of G is the minimum number of colours in an acyclic (resp. star) colouring of G. We will need the fact that all the variants of the chromatic number are at most exponential in the comparable box dimension; this follows from [46], although we include an argument to make the dependence clear.

Lemma 8.4. *For any graph G we have $\chi(G) \leq 3^{\dim_{cb}(G)}$, $\chi_a(G) \leq 5^{\dim_{cb}(G)}$ and $\chi_s(G) \leq 2 \cdot 9^{\dim_{cb}(G)}$.*

Proof. We focus on the star chromatic number and note that the chromatic number and acyclic chromatic number can be bounded by very similar arguments with some tweaks to the numbers. Suppose that G has comparable box dimension d witnessed by a representation f, and let $v_1, \ldots, v_n$ be the vertices of G written so that $\mathrm{vol}(f(v_1)) \geq \ldots \geq \mathrm{vol}(f(v_n))$. Equivalently, we have $f(v_i) \sqsubseteq f(v_j)$ whenever $i > j$. Now define a greedy colouring c so that $c(v_i)$ is the smallest colour such that $c(v_i) \neq c(v_j)$ for any $j < i$ for which either $v_j v_i \in E(G)$ or there exists $m > j$ such that $v_j v_m, v_m v_i \in E(G)$.

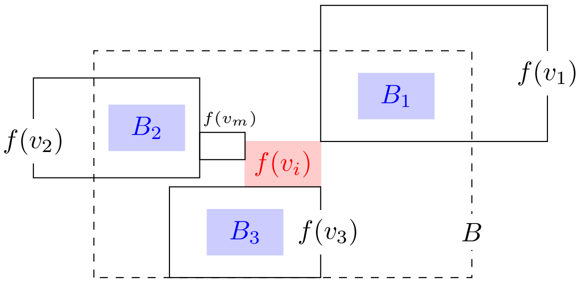


Figure 8.1: Nearby boxes obstructing colours at v_i.

Note that this gives a star colouring, since a path on four vertices always contains a 3-vertex subpath of the form $v_{i_1}v_{i_2}v_{i_3}$ such that $i_1 < i_2, i_3$, and our colouring procedure gives distinct colours to vertices forming such a path.

It remains to bound the number of colours used. Suppose we are colouring v_i. We will bound the number of vertices v_j such that $j < i$ and there exists $m > i$ for which $v_jv_m, v_mv_i \in E(G)$. Let B be the box obtained by scaling up $f(v_i)$ by a factor of 5 while keeping the same center, so that B sees boxes corresponding to vertices at graph distance up to 2 away from v. Since $f(v_m) \sqsubseteq f(v_i) \sqsubseteq f(v_j)$, there exists a translation B_j of $f(v_i)$ contained in $f(v_j) \cap B$ (see Figure 8.1). Two boxes B_j and $B_{j'}$ for $j \neq j'$ have disjoint interiors since their intersection is contained in the intersection of the touching boxes $f(v_j)$ and $f(v_{j'})$, and their interiors are also disjoint from $f(v_i) \subset B$. Thus, the number of such indices j is at most $\operatorname{vol}(B)/\operatorname{vol}(f(v_i)) - 1 = 5^d - 1$.

Near-identical arguments show that the number of indices m such that $m < i$ and $v_mv_i \in E(G)$ is at most $3^d - 1$. Consequently, the number of indices $j < i$ for which there exists m such that $j < m < i$ and $v_jv_m, v_mv_i \in E(G)$ is at most $(3^d - 1)^2$. This means that when choosing the colour of v_i greedily, we only need to avoid colours of at most $(5^d - 1) + (3^d - 1) + (3^d - 1)^2$ vertices, so $2 \cdot 9^d$ colours suffice. $\square$

8.2 Representation of operations

Given a touching representation of a graph G, it is easy to obtain a touching representation by boxes of an induced subgraph H of G by simply deleting the boxes corresponding to the vertices in $V(G) \setminus V(H)$. We show that these representations also behave nicely under several other basic operations on graphs. To describe the boxes, we use Cartesian products of boxes of lower dimension (so that $A \times B$ is the box whose projection on some first number of dimensions gives the box A, while the projection on the remaining dimensions gives the box B), or specify its projections onto every dimension (and in this case write $A[i]$ for the interval obtained from projecting A on its i^{th} dimension).

8.2.1 Adding dimensions

Three of our operations of interest admit touching representations by adding dimensions in a sometimes straightforward, sometimes clever way.

Vertex addition. Let us start with a simple lemma which says that the addition of a vertex increases the comparable box dimension by at most one. In particular, this implies that $\dim_{cb}(G) \leq |V(G)|$.

Lemma 8.5. *For any graph G and $v \in V(G)$, we have* $\dim_{cb}(G) \leq \dim_{cb}(G - v) + 1$.

Proof. Let f be a touching representation of $G - v$ by comparable boxes in $\mathbb{R}^d$, where $d = \dim_{cb}(G - v)$. We define a representation h of G as follows. For each $u \in V(G) \setminus \{v\}$, let $h(u) = [0, 1] \times f(u)$ if $uv \in E(G)$ and $h(u) = [1/2, 3/2] \times f(u)$ if $uv \notin E(G)$. Let $h(v) = [-1, 0] \times [-M, M] \times \cdots \times [-M, M]$, where M is chosen large enough so that $f(u) \subseteq [-M, M] \times \cdots \times [-M, M]$ for every $u \in V(G) \setminus \{v\}$. Then h is a touching representation of G by comparable boxes in $\mathbb{R}^{d+1}$. □

Strong product. To obtain a touching representation of $G \boxtimes H$ it suffices to take a product of representations of G and H, but the resulting representation may contain incomparable boxes. Indeed, in general $\dim_{cb}(G \boxtimes H)$ is not bounded by a function of $\dim_{cb}(G)$ and $\dim_{cb}(H)$; for example, every star has comparable box dimension at most two, but the strong product of the star $K_{1,n}$ with itself contains $K_{n,n}$ as an induced subgraph, and thus its comparable box dimension is at least $\Omega(\log n)$. However, as shown in the following lemma, this issue does not arise if the representation of H consists of translates of a single box; by scaling, we can without loss of generality assume this box is a unit hypercube.

Lemma 8.6. *Consider a graph H having a touching representation h in $\mathbb{R}^{d_H}$ by axis-aligned hypercubes of unit size. Then for any graph G, the strong product $G \boxtimes H$ of these graphs has comparable box dimension at most* $\dim_{cb}(G) + d_H$.

Proof. It suffices to take a product of the two representations. Indeed, consider a touching respresentation g of G by comparable boxes in $\mathbb{R}^{d_G}$, with $d_G = \dim_{cb}(G)$, and the representation h of H. Now define a representation f of $G \boxtimes H$ in $\mathbb{R}^{d_G + d_H}$ by

$$f((u,v))[i] = \begin{cases} g(u)[i] & \text{if } i \leq d_G \\ h(v)[i - d_G] & \text{if } i > d_G. \end{cases}$$

Consider distinct vertices (u, v) and (u', v') of $G \boxtimes H$. The boxes $g(u)$ and $g(u')$ are comparable, say $g(u) \sqsubseteq g(u')$. Since $h(v')$ is a translation of $h(v)$, this implies that $f((u, v)) \sqsubseteq f((u', v'))$. Hence, the boxes of the representation f are pairwise comparable.

The boxes of the representations g and h have pairwise disjoint interiors. Hence, if $u \neq u'$, then there exists $i \leq d_G$ such that the interiors of the intervals $f((u, v))[i] = g(u)[i]$ and $f((u', v'))[i] = g(u')[i]$ are disjoint; if $v \neq v'$, then there exists $i \leq d_H$ such that the interiors of the intervals $f((u, v))[i + d_G] = h(v)[i]$ and $f((u', v'))[i + d_G] = h(v')[i]$ are disjoint. Consequently, the interiors of boxes $f((u, v))$ and $f((u', v'))$ are pairwise disjoint. Moreover, if $u \neq u'$ and $uu' \notin E(G)$, or if $v \neq v'$ and $vv' \notin E(G)$, then the aforementioned intervals (not just their interiors) are disjoint for some i; hence, if (u, v) and (u', v') are not adjacent in $G \boxtimes H$, then $f((u, v)) \cap f((u', v')) = \emptyset$. Therefore, f is a touching representation of a subgraph of $G \boxtimes H$.

Finally, suppose that (u, v) and (u', v') are adjacent in $G \boxtimes H$. Then there exists a point p_G in the intersection of $g(u)$ and $g(u')$, since $u = u'$ or $uu' \in E(G)$ and g is a touching representation of G; and similarly, there exists a point p_H in the intersection of $h(v)$ and $h(v')$. Then $p_G \times p_H$ is a point in the intersection of $f((u, v))$ and $f((u', v'))$. Hence, f is indeed a touching representation of $G \boxtimes H$. $\square$

Taking a subgraph. Unlike when taking induced subgraphs, the comparable box dimension of a (general) subgraph of a graph G may be larger than $\dim_{cb}(G)$ (see the end of this section for an example). However, we show that the comparable box dimension of a subgraph is at most exponential in the comparable box dimension of the whole graph. This is essentially Corollary 25 in [46], but since the setting is somewhat different and the construction of [46] uses rotated boxes, we provide details of the argument.

Lemma 8.7. *If G is a subgraph of a graph G', then* $\dim_{cb}(G) \leq \dim_{cb}(G') + \binom{\chi_s(G')}{2}$.

Proof. By removing boxes that represent vertices of G that are not in G', we may assume that $V(G') = V(G)$. Let f be a touching representation of G' by comparable boxes in $\mathbb{R}^d$, where $d = \dim_{cb}(G')$. Let φ be a star colouring of G' using colours $\{1, \ldots, c\}$, where $c = \chi_s(G')$.

For any distinct colours $i, j \in \{1, \ldots, c\}$, let $A_{i,j} \subseteq V(G)$ be the set of vertices u of colour i such that there exists a vertex v of colour j such that $uv \in E(G') \setminus E(G)$. For each $u \in A_{i,j}$, let $a_j(u)$ denote such a vertex v chosen arbitrarily.

Let us define a representation h by boxes in $\mathbb{R}^{d+\binom{c}{2}}$ by starting from the representation f and, for each pair $i < j$ of colours, adding a dimension $d_{i,j}$ and setting

$$h(v)[d_{i,j}] = \begin{cases} [1/3, 4/3] & \text{if } v \in A_{i,j} \\ [-4/3, -1/3] & \text{if } v \in A_{j,i} \\ [-1/2, 1/2] & \text{otherwise.} \end{cases}$$

Note that the boxes in this extended representation are comparable, as in the added dimensions, all the boxes have size 1.

Suppose $uv \in E(G)$, where $\varphi(u) = i$ and $\varphi(v) = j$ and say $i < j$. We cannot have $u \in A_{i,j}$ and $v \in A_{j,i}$, as then $a_j(u)uva_i(v)$ would be a 4-vertex path in G' in colours i and j. Hence, in any added dimension d', we have $h(u)[d'] = [-1/2, 1/2]$ or $h(v)[d'] = [-1/2, 1/2]$, and thus $h(u)[d'] \cap h(v)[d'] \neq \emptyset$. Since the boxes $f(u)$ and $f(v)$ touch, it follows that the boxes $h(u)$ and $h(v)$ touch as well.

Suppose now that $uv \notin E(G)$. If $uv \notin E(G')$, then $f(u)$ is disjoint from $f(v)$, and thus $h(u)$ is disjoint from $h(v)$. Hence, we can assume $uv \in E(G') \setminus E(G)$, $\varphi(u) = i$, $\varphi(v) = j$ and $i < j$. Then $u \in A_{i,j}$, $v \in A_{j,i}$, $h(u)[d_{i,j}] = [1/3, 4/3]$, $h(v)[d_{j,i}] = [-4/3, -1/3]$, and $h(u) \cap h(v) = \emptyset$.

Consequently, h is a touching representation of G by comparable boxes in dimension $d + \binom{c}{2}$. □

Let us now combine Lemmas 8.4 and 8.7.

Corollary 8.8. *If G is a subgraph of a graph G', then* $\dim_{cb}(G) \leq \dim_{cb}(G') + 2 \cdot 81^{\dim_{cb}(G')} \leq 3 \cdot 81^{\dim_{cb}(G')}$.

We remark that an exponential increase in the dimension is unavoidable: we have $\dim_{cb}(K_{2^d}) = d$, but the graph obtained from K_{2^d} by deleting a perfect matching has comparable box dimension 2^{d-1}. Indeed, for every pair u, v of non-adjacent vertices there is a specific dimension i such that their boxes span intervals $[a, b]$ and $[c, d]$ with $b < c$, while the i^{th} interval of every other box in the representation contains $[b, c]$.

8.2.2 Clique-sum extendability

The last operations we wish to represent are clique-sums, which we recall are defined before Theorem 7.5. This entails taking much greater care with how our initial representations of the original graphs are chosen. The main issue to overcome in obtaining a representation for a (full) clique-sum is that the representations of G_1 and G_2 can be 'degenerate'. Consider, for example, the case where G_1 is represented by unit squares arranged in a grid; here there is no space to attach G_2 at the cliques formed by four squares intersecting in a single corner. This can be avoided by increasing the dimension, but we need to be careful so that the dimension stays bounded even after an arbitrary number of clique-sums. Thus, we introduce the notion of *clique-sum extendable* representations.

Definition 8.9. Consider a graph G with a distinguished clique $C^\star$, called the *root clique* of G. A touching representation h of G by (not necessarily comparable) boxes in $\mathbb{R}^d$ is called *$C^\star$-clique-sum extendable* if the following conditions hold for every sufficiently small $\varepsilon > 0$.

(vertices) For each $u \in V(C^\star)$, there exists a dimension d_u, such that:

(v0) $d_u \neq d_{u'}$ for distinct $u, u' \in V(C^\star)$,

(v1) each vertex $u \in V(C^\star)$ satisfies $h(u)[d_u] = [-1, 0]$ and $h(u)[i] = [0, 1]$ for any dimension $i \neq d_u$, and

(v2) each vertex $v \notin V(C^\star)$ satisfies $h(v) \subset [0, 1)^d$.

(cliques) For every clique C of G, there is a point $p(C) \in [0, 1)^d \cap \big(\bigcap_{v \in V(C)} h(v)\big)$ such that, defining the *clique box* $h^\varepsilon(C)$ by setting $h^\varepsilon(C)[i] = [p(C)[i], p(C)[i]+\varepsilon]$ for every dimension i, the following conditions are satisfied:

(c1) For any two cliques $C_1 \neq C_2$, $h^\varepsilon(C_1) \cap h^\varepsilon(C_2) = \emptyset$ (equivalently, this says that $p(C_1) \neq p(C_2)$).

(c2) A box $h(v)$ intersects $h^\varepsilon(C)$ if and only if $v \in V(C)$, and in that case their intersection is a facet of $h^\varepsilon(C)$ incident to $p(C)$. That is, there exists a dimension $i_{C,v}$ such that for each dimension j,

$$h(v)[j] \cap h^\varepsilon(C)[j] = \begin{cases} \{p(C)[i_{C,v}]\} & \text{if } j = i_{C,v} \\ [p(C)[j], p(C)[j] + \varepsilon] & \text{otherwise.} \end{cases}$$

Note that the root clique can be empty, that is the empty subgraph with no vertices. In that case the clique is denoted $\emptyset$. Let $\dim_{cb}^{ext}(G)$ be the minimum dimension such that G has an $\emptyset$-clique-sum extendable touching representation by comparable boxes.

We remark that a clique-sum extendable representation in dimension d implies the existence of such a representation in higher dimensions as well.

Lemma 8.10. *Let G be a graph with a root clique $C^\star$ and let h be a $C^\star$-clique-sum extendable touching representation of G by comparable boxes in $\mathbb{R}^d$. Then G has such a representation in $\mathbb{R}^{d'}$ for every $d' \geq d$.*

Proof. It clearly suffices to consider the case that $d' = d+1$. Note that the **(vertices)** conditions imply that $h(v') \sqsubseteq h(v)$ for every $v' \in V(G) \setminus V(C^\star)$ and $v \in V(C^\star)$. We extend the representation h by setting $h(v)[d+1] = [0,1]$ for $v \in V(C^\star)$ and $h(v)[d+1] = [0, \frac{1}{2}]$ for $v \in V(G) \setminus V(C^\star)$. The clique point $p(C)$ of each clique C is extended by setting $p(C)[d+1] = \frac{1}{4}$. It is easy to verify that the resulting representation is $C^\star$-clique-sum extendable. □

The following lemma ensures that clique-sum extendable representations behave well with respect to full clique-sums. While it is somewhat technical, the main idea is to translate (allowing also exchanges of dimensions) and scale h_2 to fit in $h_1^\varepsilon(C_1)$.

Lemma 8.11. *Consider two graphs G_1 and G_2, given with a $C_1^\star$- and a $C_2^\star$-clique-sum extendable representations h_1 and h_2 by comparable boxes in $\mathbb{R}^{d_1}$ and $\mathbb{R}^{d_2}$, respectively. Let G be the graph obtained by performing a full clique-sum of these two graphs on any clique C_1 of G_1, and on the root clique $C_2^\star$ of G_2. Then G admits a $C_1^\star$-clique sum extendable representation h by comparable boxes in $\mathbb{R}^{\max(d_1,d_2)}$.*

Proof. By Lemma 8.10, we can assume that $d_1 = d_2$; let $d = d_1$. Consider an $\varepsilon > 0$ sufficiently small so that $h_1^\varepsilon(C_1)$ satisfies all the **(cliques)** conditions, and such that $h_1^\varepsilon(C_1) \sqsubseteq h_1(v)$ for any vertex $v \in V(G_1)$. Let $V(C_1) = \{v_1, \ldots, v_k\}$; without loss of generality, we can assume $i_{C_1,v_i} = i$ for $i \in \{1, \ldots, k\}$, and thus

$$h_1(v_i)[j] \cap h_1^\varepsilon(C_1)[j] = \begin{cases} \{p_1(C_1)[i]\} & \text{if } j = i \\ [p_1(C_1)[j], p_1(C_1)[j] + \varepsilon] & \text{otherwise.} \end{cases}$$

Now let us consider G_2 and its representation h_2. Here the vertices of $C_2^\star$ are also denoted $v_1, \ldots, v_k$, and without loss of generality, the **(vertices)** conditions are satisfied by setting $d_{v_i} = i$ for $i \in \{1, \ldots, k\}$

To define h, for $v \in V(G_1)$, we first set $h(v) = h_1(v)$. Then scale and translate h_2 to fit inside $h_1^\varepsilon(C_1)$. That is, we fix $\varepsilon > 0$ small enough so that the conditions **(cliques)** hold for h_1, we have $h_1^\varepsilon(C_1) \subset [0,1)^d$, and $h_1^\varepsilon(C_1) \sqsubseteq h_1(u)$ for every $u \in V(G_1)$. For each $v \in V(G_2) \setminus V(C_2^\star)$, we set $h(v)[i] = p_1(C_1)[i] + \varepsilon h_2(v)[i]$ for $i \in \{1, \ldots, d\}$. Note that the condition (v2) for h_2 implies $h(v) \subset h_1^\varepsilon(C_1)$. Each clique C of H is a clique of G_1 or G_2. If C is a clique of G_2 then we set $p(C) = p_1(C_1) + \varepsilon p_2(C)$, otherwise we set $p(C) = p_1(C)$. In particular, for subcliques of $C_1 = C_2^\star$, we use the former choice.

Let us now check that h is a $C_1^\star$-clique sum extendable representation by comparable boxes. Firstly, the fact that the boxes are comparable follows from the fact that those of h_1 and h_2 are comparable and from the scaling of h_2: by construction both $h_1(v) \sqsubseteq h_1(u)$ and $h_2(v) \sqsubseteq h_2(u)$ imply $h(v) \sqsubseteq h(u)$, and for any vertex $u \in V(G_1)$ and any vertex $v \in V(G_2) \setminus V(C_2^\star)$, we have $h(v) \subset h_1^\varepsilon(C_1) \sqsubseteq h(u)$.

Next, we check that h is a touching representation of G. For $u, v \in V(G_1)$ (resp. $u, v \in V(G_2) \setminus V(C_2^\star)$) it is clear that $h(u)$ and $h(v)$ have disjoint interiors, and that they intersect if and only if $h_1(u)$ and $h_1(v)$ intersect (resp. if $h_2(u)$ and $h_2(v)$ intersect). Consider now a vertex $u \in V(G_1)$ and a vertex $v \in V(G_2) \setminus V(C_2^\star)$. As $h(v) \subset h^\varepsilon(C_1)$, the condition (v2) for h_1 implies that $h(u)$ and $h(v)$ have disjoint interiors.

Furthermore, if $uv \in E(G)$, then $u \in V(C_1) = V(C_2^\star)$, say $u = v_1$. Since $uv \in E(G_2)$, the intervals $h_2(u)[1]$ and $h_2(v)[1]$ intersect, and by (v1) and (v2) for h_2, we conclude that $h_2(v)[1] = [0, \alpha]$ for some positive $\alpha < 1$. Therefore, $p_1(C_1)[1] \in h(v)[1]$. Since $p_1(C_1) \in \bigcap_{x \in V(C_1)} h_1(x)$, we have $p_1(C_1) \in h(u)$, and thus $p_1(C_1)[1] \in h(u)[1] \cap h(v)[1]$. For $i \in \{2, \ldots, d\}$, note that $i \neq 1 = i_{C_1,u}$, and thus by (c2) for h_1, we have $h_1^\varepsilon(C_1)[i] \subseteq h_1(u)[i] = h(u)[i]$. Since $h(v)[i] \subseteq h_1^\varepsilon(C_1)[i]$, it follows that $h(u)$ intersects $h(v)$.

All that remains is to verify the conditions for $C_1^\star$-clique-sum extendability. The **(vertices)** conditions hold, since (v0) and (v1) are inherited from h_1, and (v2) is inherited from h_1 for $v \in V(G_1) \setminus V(C_1^\star)$ and follows from the fact that $h(v) \subseteq h_1^\varepsilon(C_1) \subset [0, 1)^d$ for $v \in V(G_2) \setminus V(C_2^\star)$. For the **(cliques)** condition (c1), the mapping p inherits injectivity when restricted to cliques of G_2, or to cliques of G_1 not contained in C_1. For any clique C of G_2, the point $p(C)$ is contained in $h_1^\varepsilon(C_1)$, since $p_2(C) \in [0, 1)^d$. On the other hand, if C' is a clique of G_1 not contained in C_1, then there exists $v \in V(C') \setminus V(C_1)$, we have $p(C') = p_1(C') \in h_1(v)$, and $h_1(v) \cap h_1^\varepsilon(C_1) = \emptyset$ by (c2) for h_1. Therefore, the mapping p is injective, and thus for sufficiently small $\varepsilon' > 0$ we have $h^{\varepsilon'}(C) \cap h^{\varepsilon'}(C') = \emptyset$ for any distinct cliques C and C' of G.

The condition (c2) of h is (for sufficiently small $\varepsilon' > 0$) inherited from the property (c2) of h_1 and h_2 when C is a clique of G_2 and $v \in V(G_2) \setminus V(C_2^\star)$, or when C is a clique of G_1 not contained in C_1 and $v \in V(G_1)$. If C is a clique of G_1 not contained in C_1 and $v \in V(G_2) \setminus V(C_2^\star)$, then by (c1) for h_1 we have $h_1^\varepsilon(C) \cap h_1^\varepsilon(C_1) = \emptyset$, and since $h^{\varepsilon'}(C) \subseteq h_1^\varepsilon(C)$ and $h(v) \subseteq h_1^\varepsilon(C_1)$, we conclude that $h(v) \cap h^{\varepsilon'}(C) = \emptyset$. It remains to consider the case that C is a clique of G_2 and $v \in V(G_1)$. Note that

$h^{\varepsilon'}(C) \subseteq h_1^{\varepsilon}(C_1)$. If $v \notin V(C_1)$, then by the property (c2) of h_1, the box $h(v) = h_1(v)$ is disjoint from $h_1^{\varepsilon}(C_1)$, and thus $h(v) \cap h^{\varepsilon'}(C) = \emptyset$.

Otherwise $v \in V(C_1) = V(C_2^\star)$, say $v = v_1$. Note that by (v1), we have $h_2(v) = [-1,0] \times [0,1]^{d-1}$. If $v \notin V(C)$, then by the property (c2) of h_2, the box $h_2(v)$ is disjoint from $h_2^{\varepsilon}(C)$. Since $h_2^{\varepsilon}(C)[i] \subseteq [0,1] = h_2(v)[i]$ for $i \in \{2,\dots,d\}$, it follows that $h_2^{\varepsilon}(C)[1] \subseteq (0,1)$, and thus $h^{\varepsilon'}(C)[1] \subseteq h_1^{\varepsilon}(C_1)[1] \setminus \{p(C_1)[1]\}$. By (c2) for h_1, we have $h(v)[1] \cap h_1^{\varepsilon}(C_1)[1] = h_1(v)[1] \cap h_1^{\varepsilon}(C_1)[1] = p(C_1)[1]$, and thus $h(v) \cap h^{\varepsilon'}(C) = \emptyset$.

If $v \in V(C)$, then by the property (c2) of h_2, the intersection of $h_2(v)[1] = [-1,0]$ and $h_2^{\varepsilon}(C)[1] \subseteq [0,1)$ is the single point $p_2(C)[1] = 0$, and thus $p(C)[1] = p_1(C_1)[1]$ and $h^{\varepsilon'}(C)[1] = [p_1(C_1)[1], p_1(C_1)[1] + \varepsilon']$. Recall that the property (c2) of h_1 implies $h(v)[1] \cap h_1^{\varepsilon}(C_1)[1] = \{p(C_1)[1]\}$, and thus $h(v)[1] \cap h^{\varepsilon'}(C)[1] = \{p(C)[1]\}$. For $i \in \{2,\dots,d\}$, the property (c2) of h_1 implies $h_1^{\varepsilon}(C_1)[i] \subseteq h_1(v)[i] = h(v)[i]$, and since $h^{\varepsilon'}(C)[i] \subseteq h_1^{\varepsilon}(C_1)[i]$, it follows that $h^{\varepsilon'}(C)[i] \subseteq h(v)[i]$. ☐

Moreover, we can pick the root clique at the expense of increasing the dimension by $\omega(G)$. The proof is essentially the same as that of Lemma 8.5, but we include it for completeness.

Lemma 8.12. *For any graph G and any clique $C^\star$, the graph G admits a $C^\star$-clique-sum extendable touching representation by comparable boxes in $\mathbb{R}^d$, for $d = |V(C^\star)| + \dim_{cb}^{ext}(G \setminus V(C^\star))$.*

Proof. Consider a $\emptyset$-clique-sum extendable touching representation h' of $G \setminus V(C^\star)$ by comparable boxes in $\mathbb{R}^{d'}$, with $d' = \dim_{cb}(G \setminus V(C^\star))$, and let $V(C^\star) = \{v_1, \dots, v_k\}$. We construct the desired representation h of G as follows. For each vertex $v_i \in V(C^\star)$, let $h(v_i)$ be the box in $\mathbb{R}^d$ uniquely determined by the condition (v1) with $d_{v_i} = i$. For each vertex $u \in V(G) \setminus V(C^\star)$, if $i \le k$ then let $h(u)[i] = [0, 1/2]$ if $uv_i \in E(G)$, and $h(u)[i] = [1/4, 3/4]$ if $uv_i \notin E(G)$. For $i > k$ we have $h(u)[i] = \alpha h'(u)[i-k]$, for some $\alpha > 0$. The value $\alpha > 0$ is chosen sufficiently small so that $h(u)[i] \subset [0,1)$ whenever $u \notin V(C^\star)$. We proceed similarly for the clique points. For any clique C of G, if $i \le k$ then let $p(C)[i] = 0$ if $v_i \in V(C)$, and $p(C)[i] = 1/4$ if $v_i \notin V(C)$. For $i > k$ we refer to the clique point $p'(C')$ of $C' = C \setminus \{v_1, \dots, v_k\}$, and we set $p(C)[i] = \alpha p'(C')[i-k]$.

By the construction, it is clear that h is a touching representation of G. As $h'(u) \sqsubset h'(v)$ implies that $h(u) \sqsubset h(v)$, and as $h(u) \sqsubset h(v_i)$ for every $u \in V(G) \setminus V(C^\star)$ and every $v_i \in V(C^\star)$, we have that h is a representation by comparable boxes.

For the $C^\star$-clique-sum extendability, the **(vertices)** conditions hold by the construction. For the **(cliques)** condition (c1), let us consider distinct cliques C_1 and

C_2 of G such that $|V(C_1)| \geq |V(C_2)|$, and let $C'_i = C_i \setminus V(C^\star)$. If $C'_1 = C'_2$, there is a vertex $v_i \in V(C_1) \setminus V(C_2)$, and $p(C_1)[i] = 0 \neq 1/4 = p(C_2)[i]$. Otherwise, if $C'_1 \neq C'_2$, then $p'(C'_1) \neq p'(C'_2)$, which implies $p(C_1) \neq p(C_2)$ by construction.

For the **(cliques)** condition (c2), let us first consider a vertex $v \in V(G) \setminus V(C^\star)$ and a clique C of G containing v. In the dimensions $i \in \{1, \dots, k\}$, we always have $h^\varepsilon(C)[i] \subseteq h(v)[i]$. Indeed, if $v_i \in V(C)$, then $h^\varepsilon(C)[i] \subseteq [0, 1/2] = h(v)[i]$, as in this case v and v_i are adjacent. If instead $v_i \notin V(C)$, then $h^\varepsilon(C)[i] \subseteq [1/4, 1/2] \subseteq h(v)[i]$. By the property (c2) of h', we have $h^\varepsilon(C)[i] \subseteq h(v)[i]$ for every $i > k$, except one, for which $h^\varepsilon(C)[i] \cap h(v)[i] = \{p(C)[i]\}$.

Next, let us consider a vertex $v \in V(G) \setminus V(C^\star)$ and a clique C of G not containing v. As $v \notin V(C')$, the condition (c2) for h' implies that $p'(C')$ is disjoint from $h'(v)$, and thus $p(C)$ is disjoint from $h(v)$.

Finally, we consider a vertex $v_i \in V(C^\star)$. Note that for any clique C containing v_i, we have that $h^\varepsilon(C)[i] \cap h(v_i)[i] = [0, \varepsilon] \cap [-1, 0] = \{0\}$, and $h^\varepsilon(C)[j] \subseteq [0, 1] = h(v_i)[j]$ for any $j \neq i$. For a clique C that does not contain v_i we have that $h^\varepsilon(C)[i] \cap h(v_i)[i] \subset (0, 1) \cap [-1, 0] = \emptyset$. Condition (c2) is therefore fulfilled, which completes the proof of the lemma. $\square$

The following lemma provides an upper bound on $\dim_{cb}^{ext}(G)$ in terms of the comparable box dimension and the chromatic number of G.

Lemma 8.13. *For any graph G, $\dim_{cb}^{ext}(G) \leq \dim_{cb}(G) + \chi(G)$.*

Proof. Let h be a touching representation of G by comparable boxes in $\mathbb{R}^d$, with $d = \dim_{cb}(G)$, and let c be a $\chi(G)$-colouring of G. We start with a slightly modified version of h. We first scale h to fit in $(0, 1)^d$, and for a sufficiently small real $\alpha > 0$ we increase each box in h by 2α in every dimension, that is we replace $h(v)[i] = [a, b]$ by $[a - \alpha, b + \alpha]$ for each vertex v and dimension i. Here, we choose α to be sufficiently small so that the boxes representing non-adjacent vertices remain disjoint, and thus the resulting representation h_1 is an intersection representation of the same graph G. Moreover, observe that for every clique C of G, the intersection $I_C = \bigcap_{v \in V(C)} h_1(v)$ is a box with non-zero edge lengths. For any clique C of G, let $p_1(C)$ be a point in the interior of I_C different from the points chosen for all other cliques.

Now we add $\chi(G)$ dimensions to make the representation touching again, and to ensure some space for the clique boxes $h^{\varepsilon}(C)$. Formally we define h_2 as

$$h_2(u)[i] = \begin{cases} h_1(u)[i] & \text{if } i \leq d \\ [1/5, 3/5] & \text{if } i > d \text{ and } c(u) < i - d \\ [0, 2/5] & \text{if } i > d \text{ and } c(u) = i - d \\ [2/5, 4/5] & \text{otherwise (if } c(u) > i - d > 0). \end{cases}$$

For any clique C of G, let $c(C)$ denote the colour set $\{c(u) \mid u \in V(C)\}$. We now set

$$p_2(C)[i] = \begin{cases} p_1(C)[i] & \text{if } i \leq d \\ 2/5 & \text{if } i > d \text{ and } i - d \in c(C) \\ 1/2 & \text{otherwise.} \end{cases}$$

As h_2 is an extension of h_1, and as in each dimension $j > d$, $h_2(v)[j]$ is an interval of length 2/5 containing the point 2/5 for every vertex v, we have that h_2 is an intersection representation of G by comparable boxes. To prove that it is touching consider two adjacent vertices u and v such that $c(u) < c(v)$, and let us note that $h_2(u)[d + c(u)] = [0, 2/5]$ and $h_2(v)[d + c(u)] = [2/5, 4/5]$.

For the $\emptyset$-clique-sum extendability, the **(vertices)** conditions are void. For the **(cliques)** conditions, since p_1 is chosen to be injective, the mapping p_2 is injective as well, implying that (c1) holds.

Consider now a clique C in G and a vertex $v \in V(G)$. If $c(v) \notin c(C)$, then $h_2(v)[c(v) + d] = [0, 2/5]$ and $p_2(C)[c(v) + d] = 1/2$, implying that $h_2^{\varepsilon}(C) \cap h_2(v) = \emptyset$. If $c(v) \in c(C)$ but $v \notin V(C)$, then letting $v' \in V(C)$ be the vertex of colour $c(v)$, we have $vv' \notin E(G)$, and thus $h_1(v)$ is disjoint from $h_1(v')$. Since $p_1(C)$ is contained in the interior of $h_1(v')$, it follows that $h_2^{\varepsilon}(C) \cap h_2(v) = \emptyset$. Finally, suppose that $v \in C$. Since $p_1(C)$ is contained in the interior of $h_1(v)$, we have $h_2^{\varepsilon}(C)[i] \subset h_2(v)[i]$ for every $i \leq d$. For $i > d$ distinct from $d + c(v)$, we have $p_2^{\varepsilon}(C)[i] \in \{2/5, 1/2\}$ and $[2/5, 3/5] \subseteq h_2(v)[i]$, and thus $h_2^{\varepsilon}(C)[i] \subset h_2(v)[i]$. For $i = d + c(v)$, we have $p_2^{\varepsilon}(C)[i] = 2/5$ and $h_2(v)[i] = [0, 2/5]$, and thus $h_2^{\varepsilon}(C)[i] \cap h_2(v)[i] = \{p_2^{\varepsilon}(C)[i]\}$. Therefore, (c2) holds. □

A touching representation of axis-aligned boxes in $\mathbb{R}^d$ is said to be *fully touching* if any two intersecting boxes intersect on a $(d-1)$-dimensional box. Note that the construction above is fully touching. Indeed, two intersecting boxes corresponding to vertices u, v with colours $c(u) < c(v)$ only touch at coordinate 2/5 in the $(d + c(u))^{\text{th}}$ dimension, while they fully intersect in every other dimension. This observation with Lemma 8.4 leads to the following.

Corollary 8.14. *Any graph G has a fully touching representation of comparable axis-aligned boxes in $\mathbb{R}^d$, where $d = \dim_{cb}(G) + 3^{\dim_{cb}(G)}$.*

Together, the lemmas from this section show that comparable box dimension is almost preserved by full clique-sums.

Corollary 8.15. *Let $\mathcal{G}$ be a class of graphs of chromatic number at most k. If $\mathcal{G}'$ is the class of all graphs that can be obtained from $\mathcal{G}$ by repeatedly performing full clique-sums, then $\dim_{cb}(\mathcal{G}') \leq \dim_{cb}(\mathcal{G}) + 2k$.*

Proof. Suppose a graph G is obtained from $G_1, \ldots, G_m \in \mathcal{G}$ by a sequence of full clique-sums. Without loss of generality, the labelling of the graphs is chosen so that we first perform the full clique-sum on G_1 and G_2, then on the resulting graph and G_3, and so on. Let $C_1^\star = \emptyset$ and for $i = 2, \ldots, m$, let $C_i^\star$ be the root clique of G_i on which it is glued in the full clique-sum operation. By Lemmas 8.13 and 8.12, G_i has a $C_i^\star$-clique-sum extendable touching representation by comparable boxes in $\mathbb{R}^d$, where $d = \dim_{cb}(\mathcal{G}) + 2k$. Repeatedly applying Lemma 8.11, we conclude that $\dim_{cb}(G) \leq d$. $\square$

Putting the preceding corollary together with Lemma 8.4 and Lemma 8.7, we now have the following bounds.

Corollary 8.16. *Let $\mathcal{G}$ be a class of graphs of comparable box dimension at most d.*

- *The class $\mathcal{G}'$ of graphs obtained from $\mathcal{G}$ by repeatedly performing full clique-sums has comparable box dimension at most $d + 2 \cdot 3^d$.*
- *The closure of $\mathcal{G}'$ by taking subgraphs has comparable box dimension at most 1250^d.*

Proof. The former bound directly follows from Corollary 8.15 and the bound on the chromatic number from Lemma 8.4. For the latter, we need to bound the star chromatic number of $\mathcal{G}'$. Suppose a graph G is obtained from $G_1, \ldots, G_m \in \mathcal{G}$ by performing full clique-sums. For $i = 1, \ldots, m$, suppose G_i has an acyclic colouring φ_i by at most k colours. Note that the vertices of any clique get pairwise different colours, and thus by permuting the colours, we can ensure that when we perform the full clique-sum, the vertices that are identified have the same colour. Hence, we can define a colouring φ of G such that for each i, the restriction of φ to $V(G_i)$ is equal to φ_i. Let C be the union of any two colour classes of φ. Then for each i, $G_i[C \cap V(G_i)]$ is a forest, and since $G[C]$ is obtained from these graphs by full clique-sums, $G[C]$ is also a forest. Hence, φ is an acyclic colouring of G by at most k colours. By [2], G has a star colouring by at most $2k^2 - k$ colours. Hence, Lemma 8.4 implies that $\mathcal{G}'$ has star chromatic number at most $2 \cdot 25^d - 5^d$. The bound on the comparable box dimension of subgraphs of graphs from $\mathcal{G}'$ then follows from Lemma 8.7. □

8.3 Applying product structure

We are now ready to use product structure to bound comparable box dimension. First, let us extract the specific product structure results that we will need.

Theorem 8.17 (From Theorem 7.14 and Theorem 7.15)**.** *Any graph G of genus at most g is a subgraph of the strong product of a 4-tree-grid and $K_{\max(2g,3)}$. Moreover, for every t, there exists an integer k such that any K_t-minor-free graph G is a subgraph of a graph obtained by repeated clique-sums from k-extended k-tree-grids.*

To apply the first statement, we bound the comparable box dimension of a graph in terms of its genus. As paths and m-cliques admit touching representations with hypercubes of unit size in $\mathbb{R}^1$ and in $\mathbb{R}^{\lceil \log_2 m \rceil}$ respectively, it suffices to bound the comparable box dimension of k-trees and finish by applying Lemma 8.6.

Theorem 8.18. *For any k-tree G, $\dim_{cb}(G) \leq \dim_{cb}^{ext}(G) \leq k+1$.*

Proof. Let H be a complete graph with $k+1$ vertices and let $C^\star$ be a clique of size k in H. By Lemma 8.11, it suffices to show that H has a $C^\star$-clique-sum extendable touching representation by hypercubes in $\mathbb{R}^{k+1}$, since G can be constructed by iterated cliques sums of cliques by definition. Let $V(C^\star) = \{v_1, \ldots, v_k\}$. We construct the representation h so that (v1) holds with $d_{v_i} = i$ for each i; this uniquely determines the hypercubes $h(v_1)$, …, $h(v_k)$. For the vertex $v_{k+1} \in V(H) \setminus V(C^\star)$, we set $h(v_{k+1}) = [0, 1/2]^{k+1}$. This ensures that the **(vertices)** conditions holds.

For the **(cliques)** conditions, let us set the point $p(C)$ for every clique C as follows:

- $p(C)[i] = 0$ for every $i \leq k$ such that $v_i \in C$
- $p(C)[i] = \frac{1}{4}$ for every $i \leq k$ such that $v_i \notin C$
- $p(C)[k+1] = \frac{1}{2}$ if $v_{k+1} \in C$
- $p(C)[k+1] = \frac{3}{4}$ if $v_{k+1} \notin C$

By construction, it is clear that for each vertex $v \in V(H)$, $p(C) \in h(v)$ if and only if $v \in V(C)$.

For any two distinct cliques C_1 and C_2, the points $p(C_1)$ and $p(C_2)$ are distinct. Indeed, by symmetry we can assume that for some i we have $v_i \in V(C_1) \setminus V(C_2)$, and this implies that $p(C_1)[i] < p(C_2)[i]$. Hence, the condition (c1) holds.

Consider now a vertex v_i and a clique C. As we observed before, if $v_i \notin V(C)$, then $p(C) \notin h(v_i)$, and thus $h^\varepsilon(C)$ and $h(v_i)$ are disjoint (for sufficiently small $\varepsilon > 0$). If $v_i \in C$, then the definitions ensure that $p(C)[i]$ is equal to the maximum of $h(v_i)[i]$, and that for $j \neq i$, $p(C)[j]$ is in $h(v_i)[j]$, implying that $h(v_i)[j] \cap h^\varepsilon(C)[j] = [p(C)[j], p(C)[j] + \varepsilon]$ for sufficiently small $\varepsilon > 0$. □

It is worth noting that the bound on the comparable box dimension of Theorem 8.18 actually extends to graphs of treewidth at most k.

Corollary 8.19. *Every graph G satisfies* $\dim_{cb}(G) \leq \operatorname{tw}(G) + 1$.

Proof. Let $k = \operatorname{tw}(G)$. Observe that there exists a k-tree T with the root clique $C^\star$ such that $G \subseteq T - V(C^\star)$. By inspecting the proof of Theorem 8.18 (and Lemma 8.11), it can be seen that we obtain a representation h of $T - V(C^\star)$ in $\mathbb{R}^{k+1}$ such that

- the vertices are represented by hypercubes of pairwise different sizes, and
- if $uv \in E(T - V(C^\star))$ and $h(u) \sqsubseteq h(v)$, then $h(u) \cap h(v)$ is a facet of $h(u)$ incident with its point with minimum coordinates.

If for some $u, v \in V(G)$, we have $uv \in E(T) \setminus E(G)$, where without loss of generality $h(u) \sqsubseteq h(v)$, we now alter the representation by shrinking $h(u)$ slightly away from $h(v)$ in such a way that all other touchings are preserved. Since the hypercubes of h have pairwise different sizes, the resulting touching representation of G is by comparable boxes. □

We have already observed that every planar graph G has $\dim_{cb}(G) \leq 3$, following the results of Thomassen or Felsner and Francis. For graphs with higher genus we can now also derive upper bounds using the product structure. Combining the previous observation on the representations of paths and K_m with Theorem 8.18, Lemma 8.6 and Corollary 8.8, we find that for every graph G of Euler genus g, there exists a supergraph G' of G such that $\dim_{cb}(G') \leq 6 + \lceil \log_2 \max(2g, 3) \rceil$. This has the following direct consequence.

Theorem 1.8. *For every graph G of genus g,* $\dim_{cb}(G) \leq 3 \cdot 81^7 \cdot \max(2g, 3)^{\log_2 81}$.

Similarly, we can deal with proper minor-closed classes.

Proof of Theorem 1.7. Let $\mathcal{G}$ be a proper minor-closed class. Since $\mathcal{G}$ is proper, there exists t such that $K_t \notin \mathcal{G}$. Theorem 8.17, implies that for every t, there exists an integer k such that any K_t-minor-free graph G is a subgraph of a graph obtained by repeated clique-sums from k-extended k-tree-grids. As we have seen, k-tree-grids have comparable box dimension at most $k+2$, and by Lemma 8.5, k-extended k-tree-grids have comparable box dimension at most $2k+2$. By Corollary 8.16, it follows that $\dim_{cb}(\mathcal{G}) \leq 1250^{2k+2}$. □

Note that the graph obtained from K_{2n} by deleting a perfect matching has Euler genus $\Theta(n^2)$ and comparable box dimension n. It follows that the dependence of the comparable box dimension on the genus cannot be subpolynomial (though the degree $\log_2 81$ of the polynomial established in Theorem 1.8 certainly can be improved). The dependence of the comparable box dimension on the size of the forbidden minor that we established is not explicit as Theorem 8.17 is based on the the Graph Minor Structure Theorem. Therefore, it would be interesting to prove Theorem 1.7 by a more constructive approach. Indeed, this is an important open question for most other applications of the graph product structure theorems and the Graph Minor Structure Theorem as well.

Chapter 9

Underlying treewidth

We are now two layers deep into approximate structural characterisation theorems, with Theorem 7.5 allowing us to express minor-closed classes of graphs in terms of graphs embeddable on surfaces, and then Theorem 7.14 allowing us to view graphs embeddable on surfaces in terms of graphs with bounded treewidth. Digging deeper again, we now work to show that graph product structure theory can be used to describe graphs of bounded treewidth in terms of even simpler graphs.

Here, the building blocks are graphs of even smaller treewidth and complete graphs of bounded size. For example, a classical theorem by the referee of [35] can be interpreted as saying that every graph G of treewidth k and maximum degree Δ is a subgraph of $T \boxtimes K_{O(k\Delta)}$ for some tree T. This motivates the following definition. The *underlying treewidth* of a graph class $\mathcal{G}$ is the minimum $c \in \mathbb{N}_0$ such that, for some function f, for every graph $G \in \mathcal{G}$ there is a graph H with $\operatorname{tw}(H) \leq c$ such that G is contained in $H \boxtimes K_{f(\operatorname{tw}(G))}$. If there is no such c, then $\mathcal{G}$ has *unbounded* underlying treewidth. We call f the *treewidth-binding function*, with a secondary goal being to make this small. For example, in this framework, the result mentioned above says that any graph class with bounded degree has underlying treewidth at most 1 with treewidth-binding function $O(k)$.

Our objective in this section is to present an initial study into underlying treewidth. We will first see in Section 9.1 how they relate to existing notions of partition-width. Just as the first product structure theorems arose from certain types of partitions, we introduce a related notion of disjointed partitions in Section 9.2 that characterise the underlying treewidth of any graph class (9.2.1). This provides a practical way to obtain upper bounds on underlying treewidth, which we put to use for certain minor-closed classes of interest (9.2.2). On the other hand, Section 9.3 describes a connection to clustered colourings that allows us to derive lower bounds. We conclude in Section 9.4 with some further known values and consequences.

The work in this section is joint with Campbell, Clinch, Distel, Gollin, Hendrey, Hickingbotham, Huynh, Illingworth, Tamitegama and Wood. It was initiated at the MATRIX Structural Graph Theory Downunder II workshop in 2022, and also completed partly on visit to Monash University.

9.1 In terms of partitions

While it is convenient to use the language of graph products to describe our main results as we have done in the definition of underlying treewidth, the proof is more conveniently phrased with the equivalent notion of graph partitions. We have already seen that these are somewhat interchangeable in Section 7.3. Building on the definitions in that section, we now introduce some more terminology on partitions.

For $c \in \mathbb{N}_0$, a H-partition where $\operatorname{tw}(H) \leq c$ is called a *c-tree-partition.* The *c-tree-partition-width* of a graph G, denoted $\operatorname{tpw}_c(G)$, is the minimum width of a c-tree-partition of G. It follows from the definitions that a graph G has a H-partition of width at most ℓ if and only if G is contained in $H \boxtimes K_\ell$. Note that this is essentially the same observation as that in Observation 7.12. Thus, $\operatorname{tpw}_c(G)$ equals the minimum $\ell \in \mathbb{N}_0$ such that G is contained in $H \boxtimes K_\ell$ for some graph H with $\operatorname{tw}(H) \leq c$. This means that the underlying treewidth of a graph class $\mathcal{G}$ is precisely the minimum $c \in \mathbb{N}_0$ such that, for some function f, every graph $G \in \mathcal{G}$ has c-tree-partition-width at most $f(\operatorname{tw}(G))$. Henceforth, we take this as our working definition of underlying treewidth.

If a graph G has a H-partition for some graph H of treewidth c, then we may assume that H is edge-maximal of treewidth c, making it a c-tree (and justifying the terminology). It is also worth noting that such graphs H are chordal, which means that our results on c-tree-partitions for graphs of bounded treewidth also imply chordal partitions with bounded-size parts. These are well-studied in their own right, and notably have application to Hadwiger's conjecture (see [142, 174]).

Another special case occurs when H is forest ($c = 1$), and then a H-partition is called a *tree-partition.* These were independently introduced by Seese [155] and Halin [74]. The *tree-partition-width*[1] of G, denoted by $\operatorname{tpw}(G) = \operatorname{tpw}_1(G)$, is the minimum width of a tree-partition of G, equal to the minimum $\ell \in \mathbb{N}_0$ for which G is contained in $T \boxtimes K_\ell$ for some forest T. This presents a different way of modelling a graph on a tree to tree-decompositions, with the main difference being that while we previously had bags intersecting so edges between bags correspond to vertices of G

---

[1] Tree-partition-width has also been called *strong treewidth* [155].

in some sense, the bags are now disjoint and edges between bags correspond to edges in G. Although they are not tied for general graphs, bounded tree-partition-width does imply bounded treewidth [155]. This fact generalises for c-tree-partition-width, which has been observed implicitly (see [12]) but is also easy to see explicitly.

Observation 9.1. *For all graphs G and $c \in \mathbb{N}_0$, we have* $\operatorname{tw}(G) \leq (c+1)\operatorname{tpw}_c(G)-1$.

Proof. Let $(V_h \colon h \in V(H))$ be a c-tree-partition of G with width $\operatorname{tpw}_c(G)$, let $(W_x \colon x \in V(T))$ be a tree-decomposition of H with width c. For each $x \in V(T)$, let $U_x \coloneqq \bigcup\{V_h \colon h \in W_x\}$. We show that $(U_x \colon x \in V(T))$ is a tree-decomposition of G. For each edge vw of G, if $v \in V_h$ and $w \in V_j$, then $h = j$ or $hj \in E(H)$; in both cases, there exists $x \in V(T)$ such that $j, h \in W_x$, implying $v, w \in U_x$. For each vertex v of G, if $v \in V_h$, then $\{x \in V(T) \colon v \in U_x\} = \{x \in V(T) \colon h \in W_x\}$; since the latter set induces a subtree of T, so does the former. Thus $(U_x \colon x \in V(T))$ is a tree-decomposition of G and the width is at most $(c+1)\operatorname{tpw}_c(G) - 1$ by construction. □

Of course, $\operatorname{tw}(T) = \operatorname{tpw}(T) = 1$ for every tree T, but in general, $\operatorname{tpw}(G)$ can be much larger than $\operatorname{tw}(G)$. For example, fan graphs on n vertices have treewidth 2 and tree-partition-width $\Omega(\sqrt{n})$. On the other hand, the referee of [35] showed that if the maximum degree and treewidth are both bounded, then so is the tree-partition-width, which is one of the most useful results about tree-partitions.

Lemma 9.2 (Ding and Oporowski [35])**.** *For $k, \Delta \in \mathbb{N}$, every graph of treewidth less than k and maximum degree at most Δ has tree-partition-width at most $24k\Delta$.*

This bound is best possible up to the multiplicative constant [178]. Note that bounded maximum degree is not necessary for bounded tree-partition-width, as illustrated by stars. Ding and Oporowski [36] characterised graph classes with bounded tree-partition-width in terms of excluded topological minors. In the next section, we will give an alternative characterisation which says that graph classes with bounded tree-partition-width are exactly those that have bounded treewidth and satisfy a further 'disjointedness' condition. Furthermore, this result naturally generalises for c-tree-partition-width and thus for underlying treewidth.

9.2 Disjointed coverings

We now introduce disjointed coverings. We will then show that they can be used to characterise bounded c-tree-partition-width and underlying treewidth, and then apply these to minor-closed classes of graphs.

An ℓ-*covering* of a graph G is a set $\beta \subseteq \mathcal{P}(V(G))$ such that $|B| \leq \ell$ for every $B \in \beta$, and $\bigcup\{B \colon B \in \beta\} = V(G)$.[2] If $B_1 \cap B_2 = \emptyset$ for all distinct $B_1, B_2 \in \beta$, then β is an ℓ-*partition*. An ℓ-covering β of a graph G is (c,d)-*disjointed* if for every c-tuple $(B_1, \dots, B_c) \in \beta^c$ and every component X of $G - (B_1 \cup \dots \cup B_c)$ there exists $Q \subseteq V(X)$ with $|Q| \leq d$ such that for each component Y of $X - Q$, for some $i \in \{1, \dots, c\}$ we have $V(Y) \cap N_G(B'_i) = \emptyset$, where $B'_i := B_i \setminus (B_1 \cup \dots \cup B_{i-1})$. This is illustrated in Figure 9.1. If some $B'_i = \emptyset$, then we can take $Q = \emptyset$ since $N_G(\emptyset) = \emptyset$.

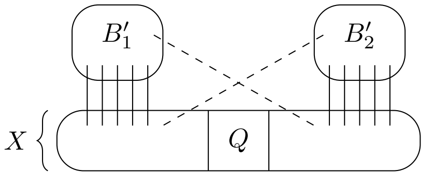


Figure 9.1: A disjointed partition with $c = 2$, where non-edges are dashed.

Here is a similar notion. Let β be an ℓ-covering of a graph G. For $t \in \mathbb{N}$, we define $\beta[t] := \{\bigcup \mathcal{B} \colon \mathcal{B} \subseteq \beta, |\mathcal{B}| \leq t\}$, which is a $t\ell$-covering of G. For a function $f \colon \mathbb{N} \to \mathbb{R}^+$, we say that β is (c,f)-*disjointed* if $\beta[t]$ is $(c, f(t))$-disjointed for every $t \in \mathbb{N}$.

While (c,d)-disjointed coverings are conceptually simpler than (c,f)-disjointed coverings, they are roughly equivalent (Theorem 9.4). Moreover, (c,f)-disjointed coverings are essential for the main proof (Lemma 9.8) and give better bounds on the c-tree-partition-width, leading to smaller treewidth-binding functions. Often, we just consider the singleton partition $\beta := \{\{v\} \colon v \in V(G)\}$ of a graph G, which is (c,f)-disjointed if and only if, for every $t \in \mathbb{N}$, every t-partition of G is $(c, f(t))$-disjointed.

9.2.1 Characterisation by disjointedness

This section characterises c-tree-partition-width in terms of (c,d)-disjointed coverings (or partitions) and (c,f)-disjointed coverings (or partitions). We start with an observation that deals with the $c = 0$ case.

Observation 9.3. *The following are equivalent for any graph G and $d \in \mathbb{N}$:*

- *G has a $(0,d)$-disjointed covering*;
- *every covering of G is $(0,d)$-disjointed*;
- *each component of G has at most d vertices*;
- *G has 0-tree-partition-width at most d.*

---

[2] Our definition of ℓ-covering differs from the standard usage where it refers to a covering in which each element of the ground set is covered ℓ times.

This implies that a graph class $\mathcal{G}$ has underlying treewidth 0 if and only if there is a function f such that every component of every graph $G \in \mathcal{G}$ has at most $f(\text{tw}(G))$ vertices. Our setup make it possible to prove the following characterisation of bounded c-tree-partition-width for general c. Previously, no such result existed even in the $c = 1$ case.

Theorem 9.4. *Let $\mathcal{G}$ be a class of graphs with bounded treewidth. For a fixed $c \in \mathbb{N}_0$, the following are equivalent:*

(a) *$\mathcal{G}$ has bounded c-tree-partition-width*;

(b) *for some $d, \ell \in \mathbb{N}$, every graph in $\mathcal{G}$ has a (c, d)-disjointed ℓ-partition*;

(c) *for some $d, \ell \in \mathbb{N}$, every graph in $\mathcal{G}$ has a (c, d)-disjointed ℓ-covering*;

(d) *for some $\ell \in \mathbb{N}$ and function f, every graph in $\mathcal{G}$ has a (c, f)-disjointed ℓ-partition*;

(e) *for some $\ell \in \mathbb{N}$ and function f, every graph in $\mathcal{G}$ has a (c, f)-disjointed ℓ-covering.*

The following serves as a roadmap for the lemmas that we will subsequently prove.

Proof of Theorem 9.4. Observation 9.3 handles the $c = 0$ case. Now assume that $c \geq 1$. Lemma 9.6 below says that (a) implies (b). Since every ℓ-partition is an ℓ-covering, (b) implies (c), and (d) implies (e). Then, Lemma 9.5 below says that (c) implies (d). Finally, Lemma 9.8 below says that (e) implies (a). □

By definition, every (c, f)-disjointed ℓ-covering is $(c, f(1))$-disjointed. The next lemma gives a qualitative converse to this.

Lemma 9.5. *Let $\ell, c, d \in \mathbb{N}$, and let β be a (c, d)-disjointed ℓ-covering of a graph G. Then β is (c, f)-disjointed, where $f(t) := dt^c$ for each $t \in \mathbb{N}$.*

Proof. Fix $t \in \mathbb{N}$. Let $B_1, \dots, B_c \in \beta[t]$. Let X be a component of $G - (B_1 \cup \dots \cup B_c)$. For each $i \in \{1, \dots, c\}$, let $\mathcal{B}_i$ be a set of at most t elements of β whose union is B_i. Let $\mathcal{F} := \mathcal{B}_1 \times \dots \times \mathcal{B}_c$, and for each $y = (A_1, \dots, A_c) \in \mathcal{F}$, define Q_y as follows. Let X_y the component of $G - (A_1 \cup \dots \cup A_c)$ containing X. Since β is (c, d)-disjointed, there exists $Q_y \subseteq V(X_y)$ of size at most d such that for every component Y of $X_y - Q_y$ there is some $i \in \{1, \dots, c\}$ such that $V(Y) \cap N_G(A_i \setminus (A_1 \cup \dots \cup A_{i-1})) = \emptyset$. Now let $Q := \bigcup_{y \in \mathcal{F}} Q_y$, and note that $|Q| \leq d|\mathcal{F}| \leq dt^c$.

Suppose for a contradiction that for some component Y of $X - Q$ and each $i \in \{1, \dots, c\}$, there is a vertex $b_i \in N_G(Y) \cap B'_i$, where $B'_i := B_i \setminus (B_1 \cup \dots \cup B_{i-1})$.

Let $y = (A_1, \dots, A_c) \in \mathcal{F}$ be such that $(b_1, \dots, b_c) \in A_1 \times \dots \times A_c$, and consider that component Y' of $X_y - Q_y$ containing Y. By the definition of Q_y, there is some $i \in \{1, \dots, c\}$ such that Y' contains no neighbour of a vertex in $A_i \setminus (A_1 \cup \dots \cup A_{i-1})$. In particular, all neighbours of vertices of Y are either vertices of Y' or neighbours of vertices of Y', so b_i is not a neighbour of any vertex of Y which is a contradiction. $\square$

Now we prove that having a (c, d)-disjointed partition is necessary for bounded c-tree-partition-width.

Lemma 9.6. *For all $c, \ell \in \mathbb{N}_0$, every graph G with c-tree-partition-width ℓ has a $(c, c\ell)$-disjointed ℓ-partition.*

Proof. By assumption, G has a H-partition $\beta = (V_h \colon h \in V(H))$ where H is a graph of treewidth at most c and $|V_h| \leq \ell$ for all h. We first show that the singleton partition of H is (c, c)-disjointed. Let $v_1, \dots, v_c \in V(H)$ and let X be a component of $H - \{v_1, \dots, v_c\}$. Let $(W_x \colon x \in V(T))$ be a tree-decomposition of H where $|W_x| \leq c + 1$ for all $x \in V(T)$. We may assume that $W_x \neq W_y$ whenever $x \neq y$. For each $i \in \{1, \dots, c\}$, let T_i be the subtree of T induced by $\{x \in V(T) \colon v_i \in W_x\}$.

First suppose that $V(T_i) \cap V(T_j) = \emptyset$ for some $i, j \in \{1, \dots, c\}$. Let $z \in V(T_i)$ be the closest node (in T) to T_j. Let $Q \coloneqq W_z \cap X$. Note that $Q \subseteq W_z \setminus \{v_i\}$ so $|Q| \leq c$. Any path from v_i to v_j in H passes through W_z, so each component of $X - Q$ is disjoint from $N_H(v_i)$ or $N_H(v_j)$.

Now assume that $V(T_i) \cap V(T_j) \neq \emptyset$ for all $i, j \in \{1, \dots, c\}$. Let T_X be the subgraph of T induced by $\{x \in V(T) \colon V(X) \cap W_x \neq \emptyset\}$. Since X is connected, T_X is a subtree of T. Suppose that $V(T_i) \cap V(T_X) = \emptyset$ for some i. Since $N_H(v_i) \subseteq \bigcup(W_x \colon x \in V(T_i))$, it follows that $N_H(v_i) \cap V(X) = \emptyset$ so we can take $Q \coloneqq \emptyset$ in this case. Now assume that $V(T_i) \cap V(T_X) \neq \emptyset$ for all $i \in \{1, \dots, c\}$. By the Helly property, $\tilde{T} \coloneqq T_1 \cap \dots \cap T_c \cap T_X$ is a non-empty subtree of T. For $x \in V(\tilde{T})$, we have $|W_x| \leq c + 1$ and so $W_x = \{v_1, \dots, v_c, u\}$ for some $u \in V(X)$. First suppose that $|V(\tilde{T})| \geq 2$. Then there are adjacent $x, y \in V(\tilde{T})$ with $W_x = \{v_1, \dots, v_c, u\}$ and $W_y = \{v_1, \dots, v_c, v\}$ for $u, v \in V(X)$. Since $W_x \neq W_y$, we have $u \neq v$ and thus there is no (u, v)-path in $H - \{v_1, \dots, v_c\}$, contradicting the connectedness of X. Hence $\tilde{T}$ consists of a single vertex z; thus $W_z = \{v_1, \dots, v_c, u\}$ for some $u \in V(X)$. Let $Q \coloneqq \{u\}$ and consider a component Y of $X - Q$. Let T_Y be the subtree of T induced by $\{y \in V(T) \colon V(Y) \cap W_y \neq \emptyset\}$. Since T_Y is connected and does not contain z, it is disjoint from some T_i. As above, $N_H(v_i) \cap V(Y) = \emptyset$, as required.

Having shown that the singleton partition of H is (c, c)-disjointed, we now turn back to G. By assumption, β is an ℓ-partition of G. Let $V_{v_1}, \dots, V_{v_c}$ be parts in

β, and let X be a component of $G-(V_{v_1}\cup\cdots\cup V_{v_c})$. Then $X\subseteq\bigcup\{V_h\colon h\in X'\}$ where X' is a component of $H-\{v_1,\dots,v_c\}$. Since H is (c,c)-disjointed, there exists $Q'\subseteq V(X')$ of size at most c such that each component of $X'-Q'$ is disjoint from some $N_H(v_i)$. Let $Q:=\bigcup\{V_h\colon h\in Q'\}$, which has size at most $c\ell$. Each component of $X-Q$ is disjoint from some $N_G(V_{v_i})$. □

Note that (c,f)-disjointedness is preserved when restricting to a subgraph.

Lemma 9.7. *If β is a (c,f)-disjointed ℓ-covering of a graph G, then for every subgraph $\tilde{G}$ of G, the restriction $\tilde{\beta}:=\{B\cap V(\tilde{G})\colon B\in\beta\}$ is a (c,f)-disjointed ℓ-covering of $\tilde{G}$.*

Proof. Fix $t\in\mathbb{N}$. Let $\tilde{B}_1,\dots,\tilde{B}_c\in\tilde{\beta}[t]$ and let $\tilde{X}$ be a component of $\tilde{G}-(\tilde{B}_1\cup\cdots\cup\tilde{B}_c)$. For each $i\in\{1,\dots,c\}$, there is a subset $S_i\subseteq\beta$ of size at most t such that $\tilde{B}_i=\bigcup_{B\in S_i}(B\cap V(\tilde{G}))$. Let $(B_1,\dots,B_c):=(\bigcup S_1,\dots,\bigcup S_c)$, and let β'' be the $t\ell$-covering of G given by $\beta\cup\{B_1,\dots,B_c\}$. Let X be the component of $G-(B_1\cup\cdots\cup B_c)$ that contains $\tilde{X}$, and for each $i\in\{1,\dots,c\}$ let $B_i':=B_i\setminus(B_1\cup\cdots\cup B_{i-1})$. Since β is (c,f)-disjointed, there is a subset Q of $V(X)$ of size at most $f(t)$ such that each component of $X-Q$ disjoint from $N_G(B_i')$ for some $i\in\{1,\dots,c\}$. Let $\tilde{Q}:=Q\cap V(\tilde{X})$, and note that $|\tilde{Q}|\leq|Q|\leq f(t)$. Each component of $\tilde{X}-\tilde{Q}$ is contained in a component of $X-Q$, and hence is disjoint from $N_{\tilde{G}}(\tilde{B}_i\setminus(\tilde{B}_1\cup\cdots\cup\tilde{B}_{i-1}))\subseteq N_G(B_i')$ for some $i\in\{1,\dots,c\}$. Hence $\tilde{\beta}$ is (c,f)-disjointed. □

The following lemma lies at the heart of the work in this chapter.

Lemma 9.8. *Let $k,c,\ell\in\mathbb{N}$ and $f\colon\mathbb{N}\to\mathbb{R}^+$. For any graph G, if $\operatorname{tw}(G)<k$ and G has a (c,f)-disjointed ℓ-covering, then G has c-tree-partition-width $\operatorname{tpw}_c(G)\leq\max\{12\ell k,2c\ell f(12k)\}$.*

In order to prove Lemma 9.8, we will proceed by induction using the following hypothesis.

Lemma 9.9. *Let $k,c,\ell\in\mathbb{N}$ and let $f\colon\mathbb{N}\to\mathbb{R}^+$. Let G be a graph of treewidth less than k and let $\beta\subseteq\mathcal{P}(V(G))$ be a (c,f)-disjointed ℓ-covering of G. Let $S_1,\dots,S_{c-1},R$ be subsets of $V(G)$, where $S_i\in\beta[12k]$ for each $i\in\{1,\dots,c-1\}$ and $4k\leq|R|\leq f(12k)$. Then there exists a c-tree-partition $(V_x\colon x\in V(H))$ of G of width at most $W:=\max\{12\ell k,2c\ell f(12k)\}$, and there exists a c-clique $\{x_1,\dots,x_{c-1},y\}$ of H such that $V_{x_i}=S_i\setminus(S_1\cup\cdots\cup S_{i-1})$ for each $i\in\{1,\dots,c-1\}$, and $R\setminus(S_1\cup\cdots\cup S_{c-1})\subseteq V_y$ with $|V_y|\leq2\ell(|R|-2k)$.*

Proof. We proceed by induction on $|V(G)|$. Let $S := S_1 \cup \dots \cup S_{c-1}$.

(Case A) $V(G) = R \cup S$: Let H be the complete graph on vertices $x_1, \dots, x_{c-1}, y$. Let $V_{x_i} := S_i \setminus (S_1 \cup \dots \cup S_{i-1})$ for each i and let $V_y := R$. Then $(V_x : x \in V(H))$ is a c-tree-partition of G of width at most W and $|V_y| = |R| \leq 2(|R| - 2k) \leq 2\ell(|R| - 2k)$. From now on, we assume that $G - (R \cup S)$ is non-empty.

(Case B) $4k \leq |R| \leq 12k$: Since β is an ℓ-covering, and $|R| \leq 12k$, we can pick $S_c \in \beta[12k]$ such that $R \subseteq S_c$ and $|S_c| \leq \ell|R| \leq 2\ell(|R| - 2k)$.

Let $G_1, \dots, G_a$ be the connected components of $G - (S \cup S_c)$. For each $i \in [c]$, let $S'_i := S_i \setminus (S_1 \cup \dots \cup S_{i-1})$. To complete this case, we first prove the following.

Claim 1. *For each $j \in [a]$, the induced subgraph $G[V(G_j) \cup S \cup S_c]$ has a c-tree-partition $(V^j_h : h \in V(H_j))$ of width at most W such that there is a c-clique K on the vertices $\{x_1, \dots, x_c\}$ in H_j, where $S'_i = V^j_{x_i}$ for each $i \in [c]$.*

Proof. If $|V(G_j)| < 4k$, then take H_j to be the complete graph on vertices $x_1, \dots, x_c, z$ with the partition $V^j_{x_i} := S'_i$ for $i \in \{1, \dots, c\}$ and $V^j_z := V(G_j)$. Then this gives us the desired c-tree-partition of $G[V(G_j) \cup S \cup S_c]$. So assume $|V(G_j)| \geq 4k$. Note that $\beta[12k]$ is a $(c, f(12k))$-disjointed $12k\ell$-covering containing $S_1, \dots, S_c$, so there is a subset $Q'_j \subseteq V(G_j)$ of size at most $f(12k)$ and there is a partition $\{A_1, \dots, A_c\}$ of $V(G_j - Q'_j)$ such that each A_i is a union of vertex sets of components of $G_j - Q'_j$ that do not intersect $N_G(S'_i)$. Let Q_j be a set such that $Q'_j \subseteq Q_j \subseteq V(G_j)$ and $4k \leq |Q_j| \leq f(12k)$. As illustrated in Figure 9.2, consider the subgraph

$$\begin{aligned} F_i &:= G[A_i \cup Q_j \cup S_1 \cup \dots \cup S_{i-1} \cup (S_{i+1} \setminus S'_i) \cup \dots \cup (S_c \setminus S'_i)] \\ &= G[A_i \cup Q_j \cup S'_1 \cup \dots \cup S'_{i-1} \cup S'_{i+1} \cup \dots \cup S'_c]. \end{aligned}$$

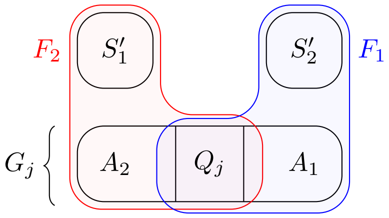


Figure 9.2: The graphs F_1 and F_2 in the case $c = 2$.

By Lemma 9.7, the restriction of β to $V(F_i)$ is a (c, f)-disjointed ℓ-covering of F_i. Apply induction to F_i with the sets $S_1, \dots, S_{i-1}, S_{i+1} \setminus S'_i, \dots, S_c \setminus S'_i$ in place of the sets $S_1, \dots, S_{c-1}$ and the set Q_j in the place of R. For each $i \in \{1, \dots, c\}$, this gives a graph L_i of treewidth at most c containing a c-clique $\{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_c, y\}$

such that F_i has an L_i-partition $(V^{j,i}_x \colon x \in V(L_i))$ with $V^{j,i}_{x_m} = S'_m$ for all $m \in \{1, \ldots, i-1, i+1, \ldots, c\}$ and $Q_j \setminus ((S \cup S_c) \setminus S'_i)) \subseteq V^{j,i}_y$ where

$$|V^{j,i}_y| \leq 2\ell(|Q_j| - 2k) \leq 2\ell f(12k).$$

Let L^+_i be obtained by taking L_i together with a vertex x_i adjacent to the clique $\{x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_c, y\}$, so we have $\text{tw}(L^+_i) \leq c$. Set $V^{j,i}_{x_i} := S'_i$. Then L^+_i contains the $(c+1)$-clique $K^+ := \{x_1, \ldots, x_c, y\}$ and $(V^{j,i}_h \colon h \in V(L^+_i))$ is an L^+_i-partition of $G[A_i \cup Q_j \cup S \cup S_c]$. Now we may assume that $V(L^+_1), \ldots, V(L^+_c)$ pairwise intersect in exactly the clique K^+. Let $H_j := L^+_1 \cup \cdots \cup L^+_c$. Since each L^+_i has treewidth at most c, so does H_j. For $x \notin K^+$, set $V^j_x := V^{j,i}_x$ for the unique i for which $x \in V(L^+_i)$, for $i \in \{1, \ldots, c\}$ set $V^j_{x_i} := S'_i$, and set $V^j_y := \bigcup_{i \in \{1,\ldots,c\}} V^{j,i}_y$. Since $|V^j_y| \leq 2c\ell f(12k)$, the partition $(V^j_x \colon x \in V(H_j))$ has width at most W. Setting $K := \{x_1 \ldots, x_c\}$, the claim follows. □

We may assume that $V(H_1), \ldots, V(H_a)$ pairwise intersect in exactly the clique K. Let $H := H_1 \cup \ldots \cup H_a$. For $x \in V(H)$, setting $V_x := V^j_x$ if $x \in V(H_j)$ is well defined, and yields a H-partition $(V_h \colon h \in V(H))$ of G. Since each H_i has treewidth at most c, so does H. Let $y := x_c$. Then $R \setminus S \subseteq S'_c = V_y$, and, as noted at the start of this case, $|V_y| = |S'_c| \leq |S_c| \leq 2\ell(|R| - 2k)$. Hence, the width of this partition is at most W, as required.

(Case C) $12k < |R| \leq f(12k)$: Since $\text{tw}(G) < k$ the separator lemma of Robertson and Seymour [146, (2.6)] tells us that there is a partition (A, B, C) of $V(G)$ with no edges between A and B, where $|C| \leq k$ and $|A \cap R|, |B \cap R| \leq \frac{2}{3}|R \setminus C|$. Let $G_1 := G[A \cup C]$ and $G_2 := G[B \cup C]$. Let $R_1 := (R \cap A) \cup C$ and $R_2 := (R \cap B) \cup C$. Since $|R| \geq 12k$,

$$\begin{aligned} &|R_1| = |A \cap R| + |C| \leq \tfrac{2}{3}|R| + k < |R| \quad \text{and} \\ &|R_1| \geq |R| - |B \cap R| \geq |R| - \tfrac{2}{3}|R \setminus C| \geq \tfrac{1}{3}|R| \geq 4k. \end{aligned}$$

Hence, $4k \leq |R_1| \leq f(12k)$ and similarly $4k \leq |R_2| \leq f(12k)$. Also $|V(G) - V(G_1)| = |B| \geq |R_2| - |C| \geq 4k - k > 0$, so $|V(G_1)| < |V(G)|$ and likewise for G_2. Fix $j \in \{1, 2\}$. Let β_j be the restriction of β to $V(G_j)$. By Lemma 9.7, β_j is a (c, f)-disjointed ℓ-covering of G_j. Let $S^j_i := S_i \cap V(G_j)$ for each $i \in \{1, \ldots, c-1\}$; note that each set S^j_i is a union of at most $12k$ elements of β_j.

Apply induction to G_j with S^j_i in place of S_i and R_j in place of R. Thus, there is a graph H_j of treewidth at most c, a c-clique $\{x_1, \ldots, x_{c-1}, y\}$ of H_j, and a H_j-partition $(V^j_x \colon x \in V(H_j))$ of G_j of width at most W such that $V^j_{x_i} = S^j_i \setminus (S^j_1 \cup \cdots \cup S^j_{i-1})$ for each i and $R_j \setminus (S_1^j \cup \cdots \cup S_{c-1}^j) \subseteq V_y^j$ with $|V_y| \leq 2\ell(|R_j| - 2k)$. We may assume that the intersection of $V(H_1)$ and $V(H_2)$ is equal to the clique $K \coloneqq \{x_1, \dots, x_{c-1}, y\}$.

Let H be the union of H_1 and H_2 and consider the H-partition $(V_x \colon x \in V(H))$ of G given by $V_x \coloneqq V_x^j$ for $x \in V(H_j) \setminus V(H_{3-j})$ and $V_x \coloneqq V_x^1 \cup V_x^2$ for $x \in K$. As H_1 and H_2 both have treewidth at most c, so does H.

Now we have that $\{x_1, \dots, x_{c-1}, y\}$ is a c-clique, $V_{x_i} = S_i \setminus (S_1 \cup \cdots \cup S_{i-1})$, and $R \setminus S \subseteq V_y$ where the size of V_y is at most

$$|V_y^1| + |V_y^2| \;\leq\; 2\ell(|R_1| + |R_2| - 4k) \;\leq\; 2\ell(|R| + 2|C| - 4k) \;\leq\; 2\ell(|R| - 2k).$$

The other parts do not change, and so this is the desired H-partition of G. □

With the preceding key lemma in hand, the proof of Lemma 9.8 is straightforward.

Proof of Lemma 9.8. Let β be a (c, f)-disjointed ℓ-covering of G. If $|V(G)| < 4k$, then the trivial partition $\{V(G)\}$ satisfies the claim. Otherwise $|V(G)| \geq 4k$. Let $R \subseteq V(G)$ with $|R| = 4k$. Let $S_1, \dots, S_{c-1}$ be arbitrary elements of β. Since β is an ℓ-covering, $|S_i| \leq \ell \leq 12\ell k$ for each i. Now Lemma 9.8 follows from Lemma 9.9. □

Putting Lemma 9.8 and Lemma 9.5 together also allows us to deduce the following.

Corollary 9.10. *Let $k, c, d, \ell \in \mathbb{N}$. For any graph G, if $\operatorname{tw}(G) < k$ and G has a (c, d)-disjointed ℓ-covering, then G has c-tree-partition-width $\operatorname{tpw}_c(G) \leq 2cd\ell(12k)^c$.*

Note that the singleton partition of any graph with maximum degree Δ is $(1, \Delta)$-disjointed. So Corollary 9.10 with $c = \ell = 1$ and $d = \Delta$ implies Lemma 9.2 (even with the same constant 24). Indeed, the proof of Lemma 9.8 in the case of graphs with bounded degree is equivalent to the proof of Lemma 9.2.

To conclude this section, Lemma 9.6 and Lemma 9.8 imply the following characterisation of underlying treewidth, fulfilling our initial goal.

Theorem 9.11. *The underlying treewidth of a graph class $\mathcal{G}$ is the minimum $c \in \mathbb{N}_0$ such that, for some function $g \colon \mathbb{N} \to \mathbb{N}$, every graph $G \in \mathcal{G}$ has a $(c, g(\operatorname{tw}(G)))$-disjointed $g(\operatorname{tw}(G))$-partition.*

9.2.2 Application to minor-closed classes

The results of the last section mean that if we can find a disjointed partition for a particular class, then this immediately gives an upper bound on its underlying treewidth which in particular would be bounded. We now put this to use for several minor-closed classes with the eventual target of planar graphs with bounded treewidth and graphs embeddable on other surfaces. However, on our way to these results we actually address the more general classes of K_t-minor-free graphs and $K_{s,t}$-minor-free graphs. These can be handled simultaneously by the next definition. For $s, t \in \mathbb{N}$, let $\mathcal{K}_{s,t}$ be the class of graphs G for which there is a partition $\{A, B\}$ of $V(G)$ such that

- $|A| = s$ and $|B| = t$,
- $vw \in E(G)$ for all $v \in A$ and $w \in B$, and
- $G[B]$ is connected.

It is clear that every graph in $\mathcal{K}_{s,t}$ contains $K_{s,t}$. On the other hand, we can obtain K_t as a minor of any $G \in \mathcal{K}_{t-2,t}$ by contracting a matching between A and B of size $t-2$ whose end-vertices are distinct from the end-vertices of some edge of $G[B]$.

There are two main tools that we'll use to show disjointedness. The first is specialised to the $\mathcal{K}_{s,t}$-minor-free setting, and gives a bound on the size of parts in certain partitions. The proof of this uses a well-established technique (see [133]).

Lemma 9.12. *Let G be a graph with no minor in $\mathcal{K}_{s,t}$. Assume $\{A, B\}$ is a partition of $V(G)$ such that $G[B]$ is connected and every vertex in B has at least s neighbours in A. Then $|B| \leq \delta|A|$ for some $\delta = \delta(s,t)$.*

Proof. We assign vertices in B to pairs of vertices in A as follows. Initially, no vertices in B are assigned. If there is an unassigned vertex $v \in B$ adjacent to distinct vertices $x, y \in A$ and no vertex in B is already assigned to $\{x, y\}$, then assign v to $\{x, y\}$. Repeat this operation until no more vertices in B can be assigned. Let B_1 and B_2 be the sets of assigned and unassigned vertices in B, respectively. Let G' be the graph obtained from $G[A \cup B_1]$ by contracting vx into x, for each vertex $v \in B_1$ assigned to $\{x, y\}$. Thus G' is a minor of G with $V(G') = A$ and $|E(G')| \geq |B_1|$. For each vertex $v \in B_2$, the set $N_G(v) \cap A$ is a clique in G' (otherwise v would have been assigned to a pair of non-adjacent neighbours) of size at least s. For each s-clique C in G', there are at most $t-1$ vertices $v \in B_2$ with $C \subseteq N_G(v)$, otherwise G has a minor in $\mathcal{K}_{s,t}$ (since $G[B]$ is connected). Since G' has no minor in $\mathcal{K}_{s,t}$, we have $|E(G')| \leq \delta|A|$ for some

$\delta = \delta(s,t)$.[3] Thus $|B_1| \le \delta|A|$. Moreover, G' is 2δ-degenerate. It is known that every d-degenerate graph on n vertices has at most $\binom{d}{s-1}n$ cliques of order s [179, Lemma 18]. Thus $|B_2| \le (t-1)\binom{2\delta}{s-1}|A|$. Hence $|B| = |B_1| + |B_2| \le (\delta + (t-1)\binom{2\delta}{s-1})|A|$. □

The second tool is the following Erdős–Pósa type result.

Lemma 9.13. *Let $\mathcal{H}$ be a set of connected subgraphs of a graph G. Then, for every integer $0 \le \ell \le |\mathcal{H}|$, either there are $\ell+1$ vertex-disjoint graphs in $\mathcal{H}$ or there is a set $Q \subseteq V(G)$ of size at most $\ell(\operatorname{tw}(G)+1)$ such that $G - Q$ contains no graph of $\mathcal{H}$.*

Lemma 9.13 follows directly from a strengthened version of the Helly property for trees that we mentioned in Section 7.2. This is given in the next lemma.

Lemma 9.14. *Let $\mathcal{F}$ be a set of subtrees of a tree T. For every non-negative integer $0 \le \ell \le |\mathcal{F}|$, either there are $\ell+1$ vertex-disjoint trees in $\mathcal{F}$ or there is a set $S \subseteq V(T)$ of size at most ℓ such that $T' \cap S \neq \emptyset$ for all $T' \in \mathcal{F}$.*

Proof. Let G be the intersection graph of $\mathcal{F}$. If $\alpha(G) \ge \ell+1$ then there are $\ell+1$ vertex-disjoint trees in $\mathcal{F}$. So we may assume that $\alpha(G) \le \ell$. Since G is chordal and thus perfect, there is a partition $X_1, \dots, X_\ell$ of $V(G)$ into cliques in G. For each i, the subtrees in X_i pairwise intersect. By the Helly property, there is a node x_i in every subtree in X_i. Thus $S \coloneqq \{x_1, \dots, x_\ell\}$ is the desired set. □

With the above tools in hand, we can now show that $\mathcal{K}_{s,t}$-minor-free graphs have the desired disjointed partition.

Lemma 9.15. *For fixed $s,t \in \mathbb{N}$, every graph G with treewidth k and no minor in $\mathcal{K}_{s,t}$ has s-tree-partition-width $O(k^2)$.*

Proof. By Lemma 9.8 it suffices to show that the singleton partition of G is (s,f)-disjointed, where $f(n) \coloneqq \delta s n(k+1)$ and $\delta \coloneqq \delta(s,t)$ from 9.12. Let $S_1, \dots, S_s$ be subsets of $V(G)$ of size at most n, let $S \coloneqq S_1 \cup \dots \cup S_s$, and for each $i \in \{1, \dots, s\}$ let $S'_i \coloneqq S_i \setminus (S_1 \cup \dots \cup S_{i-1})$. Let X be a connected component of $G - S$. Let $\mathcal{H}$ be the set of connected subgraphs H of X such that $H \cap N(S'_i) \neq \emptyset$ for each $i \in \{1, \dots, s\}$. Say $\mathcal{R}$ is a maximum-sized set of pairwise disjoint subgraphs in $\mathcal{H}$. We may assume that $\bigcup\{V(R) \colon R \in \mathcal{R}\} = V(X)$. Let X' be the graph obtained from $G[S \cup V(X)]$ by contracting each subgraph $R \in \mathcal{R}$ into a vertex v_R. So $V(X') = S \cup \{v_R \colon R \in \mathcal{R}\}$. Since X is connected, $\{v_R \colon R \in \mathcal{R}\}$ induces a connected subgraph of X'. By construction, in X', each vertex v_R has at least s neighbours in S.

---

[3] A lot is known on explicit bounds on δ. We refer to [139] for a fairly recent account.

By Lemma 9.12, $|\mathcal{R}| \leq \delta|S|$. By Lemma 9.13, there is a set $Q \subseteq V(X)$ of size at most $\delta|S|(k+1) \leq f(n)$ such that $X - Q$ contains no graph in $\mathcal{H}$, so each component Y of $X - Q$ satisfies $V(Y) \cap N_G(S'_i) = \emptyset$ for some $i \in \{1, \ldots, s\}$. Hence, the singleton partition of G is (s, f)-disjointed. $\square$

Let us now specialise to our classes of interest. Since K_t is a minor of every graph in $\mathcal{K}_{t-2,t}$, Theorem 9.15 implies that every K_t-minor-free graph of treewidth k has $(t-2)$-tree-partition-width $O(k^2)$. Thus, the underlying treewidth of the class of K_t-minor-free graphs is at most $t-2$. Similarly, since $K_{s,t}$ is a subgraph of every graph in $\mathcal{K}_{s,t}$, we have that every $K_{s,t}$-minor-free graph G of treewidth k has s-tree-partition-width $O(k^2)$ and the underlying treewidth of the class is at most s. For $s \leq t \leq 2$, however, this can be improved to $s-1$ with a little case analysis.

- Every $K_{1,1}$-minor-free graph G has no edges, and so $\mathrm{tpw}_0(G) \leq 1$.
- Every $K_{1,2}$-minor-free graph G has at most one edge in each component, and so $\mathrm{tpw}_0(G) \leq 2$.
- Each block of any $K_{2,2}$-minor-free graph G is a triangle, an edge or an isolated vertex; so $\mathrm{tpw}_1(G) \leq 2$.

Since planar graphs are K_5- and $K_{3,3}$-minor-free, the preceding deductions imply that the underlying treewidth of the class of planar graphs is at most 3. Moreover, Euler's formula implies that every graph with genus at most g is $K_{3,2g+3}$-minor-free. Thus, the underlying treewidth of the class of graphs embeddable on any fixed surface Σ is just 3. In particular, every graph embeddable in Σ and of treewidth k has 3-tree-partition-width $O(k^2)$.

We will see from constructions in the next section that all of the best bounds on underlying treewidth deduced here are actually tight.

9.3 In terms of clustered colouring

In an (improperly) vertex-coloured graph, a *monochromatic component* is a connected component of the subgraph induced by all the vertices of one colour. A graph G is *c-colourable with clustering ℓ* if each vertex can be assigned one of c colours such that each monochromatic component has at most ℓ vertices. The *clustered chromatic number* of a graph class $\mathcal{G}$ is the minimum $c \in \mathbb{N}$ such that, for some $\ell \in \mathbb{N}$, every graph in $\mathcal{G}$ has a c-colouring with clustering ℓ. We refer to the survey of Wood [180] for a more comprehensive account of clustered graph colouring. For us, the important thing to note is that a graph G is c-colourable with clustering ℓ if and only if G is contained in $H \boxtimes K_\ell$ for some graph H with $\chi(H) \leq c$. This hints at the connection to product structure and underlying treewidth, which we now establish.

Consider a graph class $\mathcal{G}$ with underlying treewidth c where every graph in $\mathcal{G}$ has treewidth at most k. This means that every graph $G \in \mathcal{G}$ is contained in $H \boxtimes K_\ell$ for some graph H with $\text{tw}(H) \leq c$, where $\ell := \max\{f(0), \dots, f(k)\}$ and f is from the definition of underlying treewidth. Since $\chi(H) \leq \text{tw}(H) + 1$ for every graph H, it follows that $\mathcal{G}$ has clustered chromatic number at most $c + 1$. It follows that lower bounds on the clustered chromatic number of a graph class with bounded treewidth provide lower bounds on the underlying treewidth of the class. For very small values, the aforementioned bound is tight.

Proposition 9.16. *A minor-closed class $\mathcal{G}$ has underlying treewidth 0 if and only if $\mathcal{G}$ has clustered chromatic number 1.*

Proof. Suppose that $\mathcal{G}$ has clustered chromatic number 1. Then there exists $\ell \in \mathbb{N}$ such that every connected graph in $\mathcal{G}$ has at most ℓ vertices. Hence $\mathcal{G}$ has underlying treewidth 0 with treewidth-binding function $f(k) = \ell$. Conversely, suppose that $\mathcal{G}$ has underlying treewidth 0 with treewidth-binding function f, and consider a spanning tree T of a non-trivial component of a graph in $\mathcal{G}$. Since $T \in \mathcal{G}$ and $\text{tw}(T) = 1$, we have $|V(T)| \leq f(1)$. As such, every non-trivial component of any graph in $\mathcal{G}$ has at most $f(1)$ vertices and thus $\mathcal{G}$ has clustered chromatic number at most 1 with clustering $\max\{1, f(1)\}$. □

Proposition 9.17. *A minor-closed class $\mathcal{G}$ has underlying treewidth at most 1 if and only if $\mathcal{G}$ has clustered chromatic number at most 2.*

Proof. Say $\mathcal{G}$ has clustered chromatic number at most 2. By [129, Theorem 14], for some $k \in \mathbb{N}$, the k-fan, the k-fat path and the k-fat star are not in $\mathcal{G}$. Then by [36, Theorem 1.2], $\mathcal{G}$ has underlying treewidth at most 1. Conversely, suppose that $\mathcal{G}$ has underlying treewidth 1 with treewidth-binding function f. The class of fan graphs has underlying treewidth at least 2 (consider $G_{2,\ell}$ in Lemma 9.19 below), so some fan graph is not in $\mathcal{G}$. Since fan graphs are planar, $\mathcal{G}$ has bounded treewidth, say at most k. Hence every graph in $\mathcal{G}$ is contained in $T \times K_\ell$ for some tree T and $\ell := \max\{f(1), \dots, f(k)\}$, and is therefore 2-colourable with clustering ℓ. □

Of course, it is another non-trivial problem to determine the clustered chromatic number of a graph class. For minor-closed classes of graphs, Norin, Scott, Seymour and Wood [129] provide some exact values and a conjecture. There is also a relationship between clustered colouring and underlying treewidth of minor-closed classes in the other direction.

Proposition 9.18. *Every minor-closed graph class $\mathcal{G}$ with underlying treewidth c has clustered chromatic number at most $2(c+1)$.*

Proof. Let f be the treewidth-binding function for $\mathcal{G}$. It was shown in [33] that, for some $w = w(\mathcal{G})$, for every graph $G \in \mathcal{G}$ there is a partition V_1, V_2 of $V(G)$ such that $\operatorname{tw}(G[V_i]) \leq w$ for $i \in \{1, 2\}$. An alternative proof based on graph products can be found in [40]. Now since $\mathcal{G}$ is minor-closed, we have $G[V_1], G[V_2] \in \mathcal{G}$. Since $\mathcal{G}$ has underlying treewidth at most c, for some graphs H_1 and H_2 of treewidth at most c, we have $G[V_1]$ is contained in $H_1 \boxtimes K_\ell$ and $G[V_2]$ is contained in $H_2 \boxtimes K_\ell$ where $\ell := \max\{f(0), \dots, f(w)\}$. Thus $G[V_1]$ and $G[V_2]$ can both be $c+1$ coloured with clustering ℓ. Using distinct colours for $G[V_1]$ and $G[V_2]$, the graph G can be $2(c+1)$ coloured with clustering ℓ, as required. □

At this point, we introduce two constructions which commonly crop up in relation to clustered colouring (see, for instance, [129, 174, 87]), that will provide some more concrete lower bounds. For a graph G and $\ell \in \mathbb{N}$, let $\ell\, G$ be the union of ℓ vertex-disjoint copies of G. Let $\widehat{G}$ be the graph obtained from G by adding one dominant vertex. Note that $\operatorname{tw}(\widehat{G}) = \operatorname{tw}(G) + 1$ and $\operatorname{tw}(\ell\, G) = \operatorname{tw}(G)$ for any $\ell \in \mathbb{N}$, which implies $\operatorname{tw}(\widehat{\ell\, G}) = \operatorname{tw}(G) + 1$. For $c, \ell \in \mathbb{N}$, define graphs $G_{c,\ell}$ and $C_{c,\ell}$ recursively as follows. First, $G_{1,\ell} := P_{\ell+1}$ is the path on $\ell + 1$ vertices, and $C_{1,\ell} := K_{1,\ell}$ is the star with ℓ leaves. Further, for $c \geq 2$, let $G_{c,\ell} := \widehat{\ell\, G_{c-1,\ell}}$ and $C_{c,\ell} := \widehat{\ell\, C_{c-1,\ell}}$. If we define the *closure* of a rooted tree T to be the graph G with $V(G) = V(T)$ and $vw \in E(G)$ whenever v is an ancestor of w or w is an ancestor of v, we may note that $C_{c,\ell}$ is just the closure of the rooted complete ℓ-ary tree of height c.

The next lemma collects together some useful properties of $G_{c,\ell}$ and $C_{c,\ell}$, the last two of which are specially included to provide lower bounds for the classes we looked at in the last section. These results are not new, but are somewhat scattered.

Lemma 9.19. *For all $c, \ell \in \mathbb{N}$,*

(i) $\operatorname{tw}(G_{c,\ell}) = \operatorname{tw}(C_{c,\ell}) = c$;

(ii) *for any ℓ-partition of $G \in \{G_{c,\ell}, C_{c,\ell}\}$, there is a $(c+1)$-clique in G whose vertices are in distinct parts*;

(iii) *$G_{c,\ell}$ and $C_{c,\ell}$ both have $(c-1)$-tree-partition-width greater than ℓ*;

(iv) $G_{2,\ell}$ *is outerplanar and* $G_{3,\ell}$ *is planar*;

(v) $G_{c,\ell}$ *is* $K_{c,\max\{c,3\}}$*-minor-free*;

Proof. Since $\operatorname{tw}(\widehat{\ell G}) = \operatorname{tw}(G) + 1$ for any graph G and $\ell \in \mathbb{N}$, part (i) follows by induction.

We establish (ii) by induction on c. In the case $c = 1$, every ℓ-partition of $P_{\ell+1}$ or $K_{1,\ell}$ contains an edge whose endpoints are in different parts, and we are done. Now assume the claim for $c-1$ ($c \geq 2$) and let $G \in \{G_{c-1,\ell}, C_{c-1,\ell}\}$. Consider an ℓ-partition of $\widehat{\ell G}$. At most $\ell - 1$ copies of G contain a vertex in the same part as the dominant vertex v. Thus, some copy G_0 of G contains no vertices in the same part as v. By induction, G_0 contains a c-clique K whose vertices are in distinct parts. Since v is dominant, $K \cup \{v\}$ satisfies the induction hypothesis.

Let $G \in \{G_{c,\ell}, C_{c,\ell}\}$. Consider a H-partition of G of width at most ℓ. By (ii), G contains a $(c+1)$-clique whose vertices are in distinct parts. So $\omega(H) \geq c+1$, implying $\operatorname{tw}(H) \geq c$. This establishes (iii).

Observe that $G_{2,\ell}$ is a fan graph, and hence outerplanar. The disjoint union of outerplanar graphs is outerplanar and the graph obtained from any outerplanar graph by adding a dominant vertex is planar; thus $G_{3,\ell}$ is planar, establishing (iv).

Lastly, we show that $G_{c,\ell}$ is $K_{c,\max\{c,3\}}$-minor-free. $G_{1,\ell}$ is a path and so has no $K_{1,3}$-minor. $G_{2,\ell}$ is outerplanar and so has no $K_{2,3}$-minor. Let $c \geq 3$ and assume the result holds for smaller c. Suppose that $G_{c,\ell}$ contains a $K_{c,c}$-minor. Since $K_{c,c}$ is 2-connected, some copy of $G_{c-1,\ell}$ in $G_{c,\ell}$ contains a $K_{c-1,c}$-minor. This contradiction establishes (v). □

The underlying treewidth of the class of graphs of partial k-trees is certainly at most k. In light of Lemma 9.19 (i) and (iii), we can conclude that this is exact.

Corollary 9.20. *The underlying treewidth of the class of graphs of treewidth at most* k *is precisely* k.

We next show there are classes of graphs with unbounded underlying treewidth.

Corollary 9.21. *The classes* $\{G_{c,\ell} \colon c, \ell \in \mathbb{N}\}$ *and* $\{C_{c,\ell} \colon c, \ell \in \mathbb{N}\}$ *both have unbounded underlying treewidth.*

Proof. Suppose that $\{G_{c,\ell} \colon c, \ell \in \mathbb{N}\}$ has underlying treewidth b. Thus, for some function f, for all $c, \ell \in \mathbb{N}$, we have $\operatorname{tpw}_b(G_{c,\ell}) \leq f(\operatorname{tw}(G_{c,\ell})) = f(c)$. In particular, with $c \coloneqq b+1$ and $\ell \coloneqq f(c)$, we have $\operatorname{tpw}_{c-1}(G_{c,\ell}) \leq \ell$, which contradicts 9.19 (iii). The proof for $\{C_{c,\ell} \colon c, \ell \in \mathbb{N}\}$ is analogous. □

Finally, we note that Lemma 9.19 gives us lower bounds for the minor-closed classes in Section 9.2.2. In particular, for any function f, we have a graph $G_{t-2,f(t-2)}$ that is K_t-minor-free (since it has treewidth $t-2$) and satisfies $\mathrm{tpw}_{t-3}(G_{t-2,f(t-2)}) > f(t-2) = f(\mathrm{tw}(G_{t-2,\ell}))$. This means the class of K_t-minor-free graphs must have underlying treewidth at least $t-2$. Similarly, for fixed s and t with $t \geq \max\{s,3\}$, we can consider $G_{s,f(s)}$ (which is $K_{s,t}$-minor-free) for any function f to show that $K_{s,t}$-minor-free graphs have underlying treewidth at least s. For the class of graphs embeddable on a fixed surface (including the class of planar graphs), it suffices to consider $G_{3,\ell}$, which we know is planar.

9.4 Further consequences

Putting together the relevant discussions from Section 9.2.2 and Section 9.3, the following theorem summarises what we have determined for minor-closed classes.

Theorem 9.22. *Let $s,t \in \mathbb{N}$ be fixed.*

(i) *If $t \geq 2$, the class of K_t-minor-free graphs has underlying treewidth $t-2$.*

(ii) *If $t \geq \max\{s,3\}$, the class of $K_{s,t}$-minor-free graphs has underlying treewidth s.*

(iii) *If $s \leq t \leq 2$, the class of $K_{s,t}$-minor-free graphs has underlying treewidth $s-1$.*

(iv) *The class of graphs embeddable on any fixed surface (in particular, planar graphs) has underlying treewidth* 3.

In terms of the original product definition, recall that this implies, for instance, that every planar graph of treewidth k is contained in $H \boxtimes K_{O(k^2)}$ where H is a graph of treewidth 3. Moreover, this bound on the treewidth of H is best possible.

Beyond these, our results from Section 9.2.1 have been put to quite extensive use for numerous classes defined by excluding other forbidden structures. We limit ourselves to an overview of the results here, the proofs of which all appear in [29]. Let us begin by replacing minors with topological minors. We will say that a graph is H-topo-minor-free if it does not contain H as a topological minor.

Theorem 9.23. *Let $s,t \in \mathbb{N}$ be fixed.*

(i) *If $t \in \{2,3,4\}$, the class of K_t-topo-minor-free graphs has underlying treewidth $t-2$.*

(ii) *If $t \geq 5$, the class of K_t-topo-minor-free graphs has underlying treewidth t.*

(iii) *If $s \leq 3$ and $t \geq \max\{s, 3\}$, then the class of $K_{s,t}$-topo minor-free graphs has underlying treewidth s.*

In addition, it can be shown that for every fixed graph X with p vertices, every X topological-minor-free graph G of treewidth k has p-tree-partition-width $O(k^2)$. By also showing that tree-partition-width is well-behaved under subdivisions, we get a characterisation of monotone classes with bounded underlying treewidth.

Theorem 9.24. *A monotone graph class $\mathcal{G}$ has bounded underlying treewidth if and only if $\mathcal{G}$ excludes some fixed topological minor.*

This has immediate implications. For example, it is well-known that every graph G has a 1-planar subdivision (that is, a subdivision with a drawing in which every edge crosses at most one other edge). To see this, take an arbitrary drawing of G and for each edge e add a subdivision vertex between consecutive crossings on e. Since the class of 1-planar graphs is monotone, Theorem 9.24 implies that the class of 1-planar graphs has unbounded underlying treewidth.

Next, we examine what happens for classes defined by excluding a fixed subgraph. Let us start by allowing non-induced subgraphs. For a graph H, let $\mathcal{G}_H$ be the class of graphs that do not contain a subgraph isomorphic to H. For a finite set of graphs $\mathcal{H}$, let $\mathcal{G}_{\mathcal{H}}$ be the class of graphs that are not isomorphic to any member of $\mathcal{H}$. It turns out that the underlying treewidth of these classes is controlled by a very specific family of graphs, which we now define.

A *star-forest* is a forest in which every component is a star, and a *spider* is a subdivision of a star. For $s, t \in \mathbb{N}$ with $s \geq 2$, the *(s,t)-spider* (denoted $S_{s,t}$) is the $(t-1)$-subdivision of $K_{1,s}$. A *spider-forest* is a subdivision of a star-forest. Containing a spider forest actually determines whether a class $\mathcal{G}_{\mathcal{H}}$ has bounded underlying treewidth.

Theorem 9.25. *For all $\ell, n, s, t \in \mathbb{N}$ where $n, s \geq 3$ and $\ell \geq 2$, and for every finite set $\mathcal{H}$ of graphs,*

(i) *$\mathcal{G}_{\mathcal{H}}$ has bounded underlying treewidth if and only if $\mathcal{H}$ contains a spider-forest*;

(ii) *the underlying treewidth of $\mathcal{G}_{P_n}$ is $\lfloor \log n \rfloor - 1$;*

(iii) *the underlying treewidth of $\mathcal{G}_{\ell P_n}$ is $\lfloor \log n \rfloor$;*

(iv) *the underlying treewidth of $\mathcal{G}_{S_{s,t}}$ is $\lfloor \log t \rfloor + 1$;*

(v) *the underlying treewidth of $\mathcal{G}_{\ell S_{s,t}}$ is $\lfloor \log t \rfloor + 2$.*

We get a similar answer for hereditary classes, where we now restrict to forbidden induced subgraphs. For a graph H, let $\mathcal{I}_H$ be the class of graphs with no induced subgraph isomorphic to H.

Theorem 9.26. *For any graph H,*

(i) *$\mathcal{I}_H$ has bounded underlying treewidth if and only if H is a star-forest*;

(ii) *If H is a star-forest, then $\mathcal{I}_H$ has underlying treewidth at most* 2;

(iii) *$\mathcal{I}_H$ has underlying treewidth at most* 1 *if and only if H is a star or each component of H is a path on at most three vertices*;

(iv) *$\mathcal{I}_H$ has underlying treewidth* 0 *if and only if H is a path on at most three vertices, or $E(H) = \emptyset$.*

The last couple of families of classes mentioned above are quite well-understood. It would be very interesting if we could do this much for minor-closed classes. In particular, it is an open problem to determine the underlying treewidth of a given minor-closed class. There is currently no known minor-closed class for which the clustered chromatic number is at least 2 more than its underlying treewidth. Could it be possible that Proposition 9.16 and Proposition 9.17 reflect the general situation, so that the simple bound we derived might be tight? More generally, there is the bigger problem of studying the underlying treewidth of classes that arise in ways other than those we have considered.

Finally, we mention that there is also scope to consider other underlying parameters. For a graph parameter θ and a graph class $\mathcal{G}$, the *underlying* θ of $\mathcal{G}$ is the minimum $c \in \mathbb{N}$ such that, for some function f, every graph $G \in \mathcal{G}$ is contained in $H \boxtimes K_{f(\theta(G))}$ for some graph H where $\theta(H) \leq c$. Equivalently, G has a H-partition of width at most $f(\theta(G))$ where $\theta(H) \leq c$. This immediately gives a family of new candidate structural parameters. A shallow exploration of underlying chromatic number, for example, suggests that it may be promising with respect to clustered colouring.

Bibliography

[1] V. Abatangelo and S. Dragomir. Discrete Fourier Calculus and Graph Reconstruction. *Interdiscip. Inform. Sci.*, **13**:163–180, 2007.

[2] M. Albertson, G. Chappell, H. Kierstead, A. Kündgen and R. Ramamurthi. Coloring with no 2-colored P_4's. *Electron. J. Combin.*, **R26**, 2004.

[3] T. Andrae. Note on the reconstruction of infinite graphs with a fixed finite number of components. *Discrete Math.*, **39**:1–7, 1982.

[4] D. Archdeacon. A Kuratowski Theorem for the Projective Plane. *J. Graph Theory*, **5**:243–246, 1981.

[5] D. Archdeacon and P. Huneke. A Kuratowski theorem for nonorientable surfaces. *J. Combin. Theory Ser. B*, **46**(2):173–231, 1989.

[6] E. Arjomandi and D. Corneil. Unicyclic graphs satisfy Harary's Conjecture. *Canad. Math. Bull.*, **17**(4):593–595, 1974.

[7] K. J. Asciak and J. Lauri. On disconnected graph with large reconstruction number. *Ars Combin.*, **62**:173–182, 2002.

[8] M. A. Bekos, G. Da Lozzo, P. Hliněný and M. Kaufmann. Graph Product Structure for h-Framed Graphs. In (S. W. Bae and H. Park, editors) *33rd International Symposium on Algorithms and Computation (ISAAC 2022)*, volume 248 of *LIPIcs*, pages 23:1–23:15, Dagstuhl, Germany, 2022.

[9] M. Bilinski, Y. S. Kwon and X. Yu. On the reconstruction of planar graphs. *J. Combin. Theory Ser. B*, **97**(5):745–756, 2007.

[10] H. L. Bodlaender. A tourist guide through treewidth. *Acta Cybernet.*, **11**(1-2):1–21, 1993.

[11] H. L. Bodlaender. A partial k-arboretum of graphs with bounded treewidth. *Theoret. Comput. Sci.*, **209**(1-2):1–45, 1998.

[12] H. L. Bodlaender, J. R. Gilbert, H. Hafsteinsson and T. Kloks. Approximating treewidth, pathwidth, frontsize, and shortest elimination tree. *J. Algorithms*, **18**(2):238–255, 1995.

[13] B. Bollobás. Almost every graph has reconstruction number 3. *J. Graph Theory*, **14**:1–4, 1990.

[14] M. Bonamy, C. Gavoille and M. Pilipczuk. Shorter labeling schemes for planar graphs. In (S. Chawla, editor) *Proc. ACM-SIAM Symposium on Discrete Algorithms (SODA '20)*, pages 446–462, 2020.

[15] J. A. Bondy. On Kelly's congruence theorem for trees. *Math. Proc. Cambridge Philos. Soc.*, **65**(2):387–397, 1969.

[16] J. A. Bondy. On Ulam's conjecture for separable graphs. *Pacific J. Math.*, **31**:281–288, 1969.

[17] J. A. Bondy. A graph reconstructor's manual. In (A. D. Keedwell, editor) *Surveys in Combinatorics*, volume 166 of *London Mathematical Society Lecture Note Series*, pages 221–252. Cambridge University Press, 1991.

[18] J. A. Bondy. Beautiful conjectures in graph theory. *European J. Combin.*, **37**:4–23, 2014.

[19] J. A. Bondy and R. L. Hemminger. Graph reconstruction – a survey. *J. Graph Theory*, **1**(3):227–268, 1977.

[20] J. A. Bondy and F. Mercier. Switching reconstruction of digraphs. *J. Graph Theory*, **67**(4):332–348, 2011.

[21] E. Bonnet, O. Kwon and D. R. Wood. Reduced bandwidth: a qualitative strengthening of twin-width in minor-closed classes (and beyond). *arXiv:2202.11858* preprint, 2022.

[22] P. Borwein, T. Erdélyi and G. Kós. Littlewood-type problems on [0, 1]. *Proc. Lond. Math. Soc.*, **79**(1):22–46, 1999.

[23] P. Borwein and C. Ingalls. The Prouhet-Tarry-Escott problem revisited. *Enseign. Math.*, **40**(2):3–27, 1994.

[24] A. Bowler, P. Brown and T. Fenner. Families of pairs of graphs with a large number of common cards. *J. Graph Theory*, **63**(2):146–163, 2010.

[25] A. Bowler, P. Brown, T. Fenner and W. Myrvold. Recognizing connectedness from vertex-deleted subgraphs. *J. Graph Theory*, **67**(4):285–299, 2011.

[26] N. Bowler, J. Erde, P. Heinig, F. Lehner and M. Pitz. A counterexample to the reconstruction conjecture for locally finite trees. *Bull. Lond. Math. Soc.*, **49**(4):630–648, 2017.

[27] N. Bowler, J. Erde, P. Heinig, F. Lehner and M. Pitz. Non-reconstructible locally finite graphs. *J. Combin. Theory Ser. B*, **133**:122–152, 2018.

[28] P. Brown and T. Fenner. The size of a graph is reconstructible from any $n-2$ cards. *Discrete Math.*, **341**(1):165–174, 2018.

[29] R. Campbell, K. Clinch, M. Distel, J. P. Gollin, K. Hendrey, R. Hickingbotham, T. Huynh, F. Illingworth, J. Tan, Y. Tamitegama and D. Wood. Product structure of graph classes with bounded treewidth. *arXiv:2206.02395* preprint, 2022.

[30] Z. A. Chernyak. Some additions to an article by B. Manvel: “Some basic observations on Kelly's conjecture for graphs” (Russian). *Vestsı Akad. Navuk BSSR Ser. Fz.-Mat. Navuk*, **126**:44–49, 1982.

[31] S. Conilh. Note on Kocay's 3-hypergraphs and Stockmeyer's tournaments. *J. Combin Theory Ser. B*, **70**:312–316, 1997.

[32] L. Danzer, B. Grünbaum and V. Klee. Helly's theorem and its relatives. In (V. Klee, editor) *Convexity*, Proc. of Symposia in Pure Math., vol. 7, A.M.S., pages 100–181, 1963.

[33] M. DeVos, G. Ding, B. Oporowski, D. P. Sanders, B. Reed, P. Seymour and D. Vertigan. Excluding any graph as a minor allows a low tree-width 2-coloring. *J. Combin. Theory Ser. B*, **91**(1):25–41, 2004.

[34] R. Diestel. *Graph Theory*, volume 173 of *Graduate Texts in Mathematics*. Springer-Verlag, Heidelberg, 3rd edition, 2005.

[35] G. Ding and B. Oporowski. Some results on tree decomposition of graphs. *J. Graph Theory*, **20**(4):481–499, 1995.

[36] G. Ding and B. Oporowski. On tree-partitions of graphs. *Discrete Math.*, **149**(1–3):45–58, 1996.

[37] M. Distel, R. Hickingbotham, T. Huynh and D. R. Wood. Improved product structure for graphs on surfaces. *Discrete Math. Theor. Comput. Sci.*, **24**(6), 2022.

[38] V. Dujmović, L. Esperet, C. Gavoille, G. Joret, P. Micek and P. Morin. Adjacency labelling for planar graphs (and beyond). *J. ACM*, **68**(6):42:1–42:33, 2021.

[39] V. Dujmović, L. Esperet, G. Joret, B. Walczak and D. R. Wood. Planar graphs have bounded nonrepetitive chromatic number. *Adv. Comb.*, **5**, 2020.

[40] V. Dujmović, G. Joret, P. Micek, P. Morin, T. Ueckerdt and D. R. Wood. Planar graphs have bounded queue-number. *J. ACM*, **67**(4):#22, 2020.

[41] V. Dujmović, G. Joret, P. Micek, P. Morin and D. Wood. Bounded-degree planar graphs do not have bounded-degree product structure. *arXiv:2212.02388* preprint, 2022.

[42] V. Dujmović, P. Morin and D. R. Wood. Layered separators in minor-closed families with applications. *J. Combin. Theory Ser. B*, **127**:111–147, 2017.

[43] V. Dujmović, P. Morin and D. R. Wood. Graph product structure for non-minor-closed classes. *J. Combin. Theory Ser. B*, **162**:34 – 67, 2023.

[44] Z. Dvorák. Approximation metatheorem for fractionally treewidth-fragile graphs. *arXiv:2103.08698* preprint, 2021.

[45] Z. Dvorák and A. Lahiri. Approximation schemes for bounded distance problems on fractionally treewidth-fragile graphs. In *29th Annual European Symposium on Algorithms, ESA 2021, September 6-8, 2021, Lisbon, Portugal (Virtual Conference)*, volume 204 of *LIPIcs*, pages 40:1–40:10. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021.

[46] Z. Dvořák, R. McCarty and S. Norin. Sublinear separators in intersection graphs of convex shapes. *arXiv:2001.01552* preprint, 2020.

[47] Z. Dvořák, J. Pekárek, T. Ueckerdt and Y. Yuditsky. Weak coloring numbers of intersection graphs. *arXiv:2103.17094* preprint, 2021.

[48] Z. Dvořák, D. Gonçalves, A. Lahiri, J. Tan and T. Ueckerdt. On comparable box dimension. *Proc. 38th International Symposium on Computational Geometry (SoCG)*, pages 38:1–38:14, 2022.

[49] Z. Dvořák, T. Huynh, G. Joret, C.-H. Liu and D. R. Wood. Notes on Graph Product Structure Theory. In (J. de Gier, C. Praeger and T. Tao, editors) *2019-20 MATRIX Annals*, volume 4 of *MATRIX Book Series*, pages 513–533. Springer, 2021.

[50] Édouard Bonnet, E. J. Kim, S. Thomassé and R. Watrigant. Twin-width I: Tractable FO Model Checking. *J. ACM*, **69**(1):1 – 46, 2021.

[51] M. N. Ellingham. Vertex-switching reconstruction and folded cubes. *J Combin Theory Ser B*, **66**:361–364, 1996.

[52] M. N. Ellingham and G. F. Royle. Vertex-switching reconstruction of subgraph numbers and triangle-free graphs. *J Combin Theory Ser B*, **54**:167–177, 1992.

[53] L. Esperet. Box representations of embedded graphs. *Discrete Comput. Geom.*, **57**:590–606, 2017.

[54] L. Esperet, G. Joret and P. Morin. Sparse universal graphs for planarity. *arXiv:2010.05779* preprint, 2020.

[55] S. Felsner and M. C. Francis. Contact representations of planar graphs with cubes. In *Proceedings of the twenty-seventh annual symposium on computational geometry*, pages 315–320, 2011.

[56] S. Fiorini and J. Lauri. The reconstruction of maximal planar graphs, I: Recognition. *J. Combin. Theory Ser. B*, **30**:188–195, 1981.

[57] S. Fiorini and J. Lauri. The reconstruction of maximal planar graphs, II: Reconstruction. *J. Combin. Theory Ser. B*, **30**:196–214, 1981.

[58] S. Fiorini and B. Manvel. A theorem on planar graphs with an application to the reconstruction problem II. *J. Comb. Inf. Syst. Sci.*, **3**:200–216, 1978.

[59] J. Fisher, R. L. Graham and F. Harary. A simpler counterexample to the reconstruction conjecture for denumerable graphs. *J. Combin. Theory Ser. B*, **12B**:203–204, 1972.

[60] D. Geller and B. Manvel. Reconstruction of cacti. *Canad. J. Math.*, **21**:1354–1360, 1969.

[61] W. B. Giles. On reconstructing maximal outerplanar graphs. *Discrete Math.*, **8**(2):169–172, 1974.

[62] W. B. Giles. The reconstruction of outerplanar graphs. *J. Combin. Theory Ser. B*, **16**(3):215–226, 1974.

[63] W. B. Giles. Point deletions of outerplanar blocks. *J. Combin. Theory Ser. B*, **20**(2):103–116, 1976.

[64] W. B. Giles. Reconstructing trees from two point deleted subtrees. *Discrete Math.*, **15**(4):325–332, 1976.

[65] D. L. Greenwell. Reconstructing graphs. *Proc. Amer. Math. Soc.*, **30**(3):431–433, 1971.

[66] D. L. Greenwell and R. L. Hemminger. Reconstructing the n-connected components of a graph. *Aequationes Math.*, **9**:19–22, 1973.

[67] D. L. Greenwell and R. L. Hemminger. New representation results for planar graphs. In *29th European Workshop on Computational Geometry (EuroCG 2013)*, pages 177–180, 2013.

[68] C. Groenland, H. Guggiari and A. Scott. Size reconstructibility of graphs. *J. Graph Theory*, **96**(2):326–337, 2021.

[69] C. Groenland, T. Johnston, A. Kupavskii, K. Meeks, A. Scott and J. Tan. Reconstructing the degree sequence of a sparse graph from a partial deck. *J. Combin. Theory Ser. B*, pages 283–293, 2022.

[70] C. Groenland, T. Johnston, A. Scott and J. Tan. Reconstructing trees from small cards. *arXiv:2103.13359* preprint, 2021.

[71] M. Grohe, S. Kreutzer, R. Rabinovich, S. Siebertz and K. Stavropoulos. Coloring and covering nowhere dense graphs. *SIAM J. Discrete Math.*, **32**:2467–2481, 2018.

[72] J. Gross and J. Yellen. *Handbook of graph theory.* CRC Press, 2nd edition, 2014.

[73] R. Halin. S-functions for graphs. *J. Geom.*, **8**:171–186, 1976.

[74] R. Halin. Tree-partitions of infinite graphs. *Discrete Math.*, **97**:203–217, 1991.

[75] F. Harary. On the reconstruction of a graph from a collection of subgraphs. In *Theory of Graphs and its Applications (Proc. Sympos. Smolenice, 1963)*, pages 47–52. Publishing House Czechoslovak Academy of Sciences Prague, 1964.

[76] F. Harary. *Graph theory.* Addison-Wesley series in mathematics. Addison-Wesley, Reading, Mass.; London, 1969.

[77] F. Harary. A survey of the reconstruction conjecture. *Graphs Combin.*, **406**:18–28, 1974.

[78] F. Harary and E. M. Palmer. The reconstruction of a tree from its maximal subtrees. *Canad. J. Math.*, **18**:803–810, 1966.

[79] F. Harary and E. M. Palmer. On the problem of reconstructing a tournament from subtournaments. *Monatsh. Math.*, **71**:14–23, 1967.

[80] F. Harary and M. Plantholt. The graph reconstruction number. *J. Graph Theory*, **9**:451–454, 1985.

[81] F. Harary, A. J. Schwenk and R. L. Scott. On the reconstruction of countable forests. *Publ. Inst. Math.*, **166**:39—-42, 1972.

[82] D. J. Harvey and D. R. Wood. Parameters tied to treewidth. *J. Graph Theory*, **84**(4):364–385, 2017.

[83] R. Hickingbotham and D. R. Wood. Shallow minors, graph products and beyond planar graphs. *arXiv:2111.12412* preprint, 2021.

[84] F. Illingworth, A. Scott and D. R. Wood. Product structure of graphs with an excluded minor. *arXiv:2104.06627* preprint, 2021.

[85] P. J. Kelly. *On isometric transformations.* PhD thesis, University of Wisconsin, 1942.

[86] P. J. Kelly. A congruence theorem for trees. *Pacific J. Math.*, **7**(1):961–968, 1957.

[87] J. M. Kleinberg, R. Motwani, P. Raghavan and S. Venkatasubramanian. Storage management for evolving databases. In *38th Annual Symp. on Foundations of Computer Science* (FOCS '97), pages 353–362. IEEE, 1997.

[88] W. L. Kocay. On reconstructing spanning subgraphs. *Ars Combin.*, pages 301–313, 1981.

[89] W. L. Kocay. On Stockmeyer's nonreconstructible tournaments. *J. Graph Theory*, **9**(4):473–476, 1985.

[90] W. L. Kocay. A family of nonreconstructible hypergraphs. *J. Combin. Theory Ser. B*, **42**:46–63, 1987.

[91] W. L. Kocay. Hypomorphisms, orbits, and reconstruction. *J. Combin. Theory Ser. B*, **44**:187–200, 1988.

[92] W. L. Kocay. A note on non-reconstructible 3-hypergraphs. *Graphs Combin.*, **32**:1945–1963, 2016.

[93] P. Koebe. Kontaktprobleme der konformen Abbildung. *Math.-Phys. Kl.*, **88**:141–164, 1936.

[94] A. V. Kostochka, M. Nahvi, D. B. West and D. Zirlin. 3-regular graphs are 2-reconstructible. *arXiv:1908.01258* preprint, 2019.

[95] A. V. Kostochka, M. Nahvi, D. B. West and D. Zirlin. Degree lists and connectedness are 3-reconstructible for graphs with at least seven vertices. *Graphs Combin.*, **36**:491–501, 2020.

[96] A. V. Kostochka, M. Nahvi, D. B. West and D. Zirlin. Acyclic graphs with at least $2\ell+1$ vertices are ℓ-recognizable. *arXiv:2103.12153* preprint, 2021.

[97] A. V. Kostochka and D. B. West. On reconstruction of graphs from the multiset of subgraphs obtained by deleting ℓ vertices. *IEEE Trans. Inform. Theory*, **67**(1):3278 – 3286, 2021.

[98] I. Krasikov. A note on the vertex-switching reconstruction. *Internat. J. Math. & Math. Sci.*, **11**:825–827, 1988.

[99] I. Krasikov. Degree conditions for vertex switching reconstruction. *Discrete Math.*, **160**:273–278, 1996.

[100] I. Krasikov, A. Lev and B. D. Thatte. Upper bounds on the automorphism group of a graph. *Discrete Math.*, **256**(1–2):489–493, 2002.

[101] I. Krasikov and S. Litsyn. On integral zeros of krawtchouk polynomials. *J. Combin. Theory Ser. A*, **74**(1):71–99, 1996.

[102] I. Krasikov and Y. Roditty. Balance equations for reconstruction problems. *Arch. Math.*, **48**:458–464, 1987.

[103] I. Krasikov and Y. Roditty. Switching Reconstruction and Diophantine Equations. *J Combin Theory Ser B*, **54**:189–195, 1992.

[104] I. Krasikov and Y. Roditty. More on vertex-switching reconstruction. *J Combin Theory Ser B*, **60**:40–55, 1994.

[105] M. Krawtchouk. Sur une généralisation des polynômes d'Hermite. *Comptes Rendus Mathématique*, **189**:620–622, 1929.

[106] J. B. Kruskal. Well-quasi-ordering, the tree theorem, and vazsonyi's conjecture. *Trans. Amer. Math. Soc.*, **95**(2):210–225, 1960.

[107] J. B. Kruskal. The theory of well-quasi-ordering: A frequently discovered concept. *J. Combin. Theory Ser. A*, **13**(3):297–305, 1972.

[108] C. Kuratowski. Sur le probleme des courbes gauches en topologie. *Fund. Math.*, **15**(1):271–283, 1930.

[109] J. Lauri. Proof of Harary's conjecture on the reconstruction of trees. *Discrete Math.*, **43**(1):79–90, 1983.

[110] J. Lauri and R. Scapellato. *Topics in graph automorphisms and reconstruction.* London Mathematical Society Lecture Note Series **432**. Cambridge University Press, 2nd edition, 2016.

[111] L. Lovász. A note on the line reconstruction problem. In *Classic Papers in Combinatorics*, pages 451–452. Springer, 1972.

[112] L. Lovász. A note on the line reconstruction problem. *J. Combin. Theory Ser. B*, **13**:309–310, 1972.

[113] L. Lovász. Graph minor theory. *Bull. Amer. Math. Soc.*, **43**(1):75–86, 2006.

[114] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes.* North-Holland, Amsterdam; New York; Oxford, 1977.

[115] B. Manvel. Reconstruction of unicyclic graphs. In (F. Harary, editor) *Proof techniques in graph theory*, pages 103–107. Academic Press, New York, 1969.

[116] B. Manvel. Reconstruction of trees. *Canad. J. Math.*, **22**(1):55—60, 1970.

[117] B. Manvel. Reconstruction of maximal outerplanar graphs. *Discrete Math.*, **2**(3):269–278, 1972.

[118] B. Manvel. Some basic observations on Kelly's conjecture for graphs. *Discrete Math.*, **8**(2):181–185, 1974.

[119] B. Manvel. On reconstructing graphs from their sets of subgraphs. *J. Combin. Theory Ser. B*, **21**:156–165, 1976.

[120] B. D. McKay. Small graphs are reconstructible. *Australas. J. Combin.*, **15**:123–126, 1997.

[121] B. D. McKay. Reconstruction of small graphs and digraphs. *arXiv:2102.01942* preprint, 2021.

[122] B. D. McKay and P. Schweitzer. Switching reconstruction of digraphs. *J. Graph Theory*, **76**(4):279–296, 2014.

[123] B. Mohar and C. Thomassen. *Graphs on Surfaces.* The Johns Hopkins University Press, Baltimore; London, 2001.

[124] W. Myrvold. *Ally and adversary reconstruction problems.* PhD thesis, University of Waterloo, 1988.

[125] W. Myrvold. The ally-reconstruction number of a tree with five or more vertices is three. *J. Graph Theory*, **14**(2):149–166, 1990.

[126] W. Myrvold. The degree sequence is reconstructible from $n - 1$ cards. *Discrete Math.*, **102**(2):187–196, 1992.

[127] W. Myrvold and J. Woodcock. A large set of torus obstructions and how they were discovered. *Electron. J. Combin.*, **P1.16**, 2018.

[128] P. Niesink. The vertex-switching reconstruction problem. Master's thesis, University of Ottawa, 2010.

[129] S. Norin, A. Scott, P. Seymour and D. R. Wood. Clustered colouring in minor-closed classes. *Combinatorica*, **39**(6):1387–1412, 2019.

[130] V. Nýdl. A note on reconstructing of finite trees from small subtrees. *Acta Universitatis Carolinae. Mathematica et Physica*, **31**(2):71–74, 1990.

[131] V. Nýdl. Finite undirected graphs which are not reconstructible from their large cardinality subgraphs. *Discrete Math.*, **108**(1-3):373–377, 1992.

[132] I. C. Oliveira and B. D. Thatte. An algebraic formulation of the graph reconstruction conjecture. *Discrete Math.*, **81**:351–363, 2016.

[133] P. Ossona de Mendez, S. Oum and D. R. Wood. Defective colouring of graphs excluding a subgraph or minor. *Combinatorica*, **39**(2):377–410, 2019.

[134] S. Oum and P. D. Seymour. Approximating clique-width and branch-width. *J. Combin. Theory Ser. B*, **96**(4):514 – 528, 2006.

[135] S. Plotkin, S. Rao and W. D. Smith. Shallow excluded minors and improved graph decompositions. In *Proceedings of the fifth annual ACM-SIAM symposium on Discrete algorithms*, pages 462–470. Society for Industrial and Applied Mathematics, 1994.

[136] A. Proskurowski. k-trees: representations and distances. *Discrete Math.*, **29**:785 – 794, 1980.

[137] L. Pyber. The edge-reconstruction of hamiltonian graphs. *J. Graph Theory*, **14**:173–179, 1990.

[138] S. Ramachandran. On a new digraph reconstruction conjecture. *J. Combin Theory Ser. B*, **31**(2):143–149, 1981.

[139] B. Reed and D. R. Wood. Forcing a sparse minor. *Combin. Probab. Comput.*, **25**(2):300–322, 2016.

[140] B. A. Reed. Finding approximate separators and computing tree width quickly. In *Proceedings of the twenty-fourth annual ACM symposium on Theory of computing*, pages 221–228, 1992.

[141] B. A. Reed. Algorithmic aspects of tree width. In *Recent advances in algorithms and combinatorics*, volume 11, pages 85–107. Springer, 2003.

[142] B. A. Reed and P. Seymour. Fractional colouring and Hadwiger's conjecture. *J. Combin. Theory Ser. B*, **74**(2):147–152, 1998.

[143] F. Roberts. On the boxicity and cubicity of a graph. In *Recent Progresses in Combinatorics*, pages 301–310. Academic Press, New York, 1969.

[144] N. Robertson and P. D. Seymour. Graph Minors. I. Excluding a forest. *J. Combin. Theory Ser. B*, **35**(1):39–61, 1983.

[145] N. Robertson and P. D. Seymour. Graph Minors. III. Planar tree-width. *J. Combin. Theory Ser. B*, **36**:49–64, 1984.

[146] N. Robertson and P. D. Seymour. Graph Minors. II. Algorithmic aspects of tree-width. *J. Algorithms*, **7**(3):309–322, 1986.

[147] N. Robertson and P. D. Seymour. Graph Minors. V. Excluding a planar graph. *J. Combin. Theory Ser. B*, **41**:92–114, 1986.

[148] N. Robertson and P. D. Seymour. Graph Minors. VIII. A Kuratowski theorem for general surfaces. *J. Combin. Theory Ser. B*, **48**(2):255–288, 1990.

[149] N. Robertson and P. D. Seymour. Graph Minors. X. Obstructions to tree-decomposition. *J. Combin. Theory Ser. B*, **52**(2):153–190, 1991.

[150] N. Robertson and P. D. Seymour. Graph Minors. XIII. The disjoint paths problem. *J. Combin. Theory Ser. B*, **63**:65–110, 1995.

[151] N. Robertson and P. D. Seymour. Graph Minors. XVII. Taming a vortex. *J. Combin. Theory Ser. B*, **77**(1):162–210, 1999.

[152] N. Robertson and P. D. Seymour. Graph Minors. XX. Wagner's conjecture. *J. Combin. Theory Ser. B*, **92**(2):325–357, 2004.

[153] H. Sachs. Coin graphs, polyhedra, and conformal mapping. *Discrete Math.*, **134**:133–138, 1994.

[154] A. Scott and D. R. Wood. Better bounds for poset dimension and boxicity. *Trans. Amer. Math. Soc.*, **373**(3):2157–2172, 2020.

[155] D. Seese. Tree-partite graphs and the complexity of algorithms. In (L. Budach, editor) *Proc. Int'l Conf. on Fundamentals of Computation Theory*, volume 199 of *LNCS*, pages 412–421. Springer, 1985.

[156] P. D. Seymour and R. Thomas. Graph searching and a min-max theorem for tree-width. *J. Combin. Theory Ser. B*, **58**(1):22–33, 1993.

[157] H. Spinoza and D. B. West. Reconstruction from the deck of k-vertex induced subgraphs. *J. Graph Theory*, **90**(4):497–522, 2019.

[158] C. St. J. A. Nash-Williams. The reconstruction problem. *Selected Topics in Graph Theory*, **1**:205–236, 1978.

[159] C. St. J. A. Nash-Williams. Reconstruction of locally finite connected graphs with at least three infinite wings. *J. Graph Theory*, **11**(4):497–505, 1987.

[160] C. St. J. A. Nash-Williams. Reconstruction of infinite graphs. *Discrete Math.*, **95**:221–229, 1991.

[161] C. St. J. A. Nash-Williams. Reconstruction of locally finite connected graphs with two infinite wings. *Discrete Math.*, **92**:227–249, 1991.

[162] R. P. Stanley. Reconstruction from vertex-switching. *J. Combin. Theory Ser. B*, **38**(2):132–138, 1985.

[163] P. K. Stockmeyer. The falsity of the reconstruction conjecture for tournaments. *J. Graph Theory*, **1**:19–25, 1977.

[164] P. K. Stockmeyer. A census of non-reconstructible digraphs I. Six related families. *J. Combin Theory Ser. B*, **31**:232–239, 1981.

[165] P. K. Stockmeyer. Erratum to: "The falsity of the reconstruction conjecture for tournaments". *J. Graph Theory*, **62**(2):199–200, 2009.

[166] P. K. Stockmeyer. A census of non-reconstructable digraphs, II: A family of tournaments. *arXiv:2110.06158* manuscript, 2021.

[167] T. Stoll. Reconstruction problems for graphs, Krawtchouk polynomials and Diophantine equations. In (M. Dehmer, editor) *Structural Analysis of Complex Networks*, pages 293–317. Birkhäuser, Boston, 2011.

[168] S. Tarkowski. On the comparability of dendrites. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.*, **8**:39–41, 1960.

[169] R. Taylor. Reconstructing degree sequences from k-vertex-deleted subgraphs. *Discrete Math.*, **79**(2):207–213, 1989/90.

[170] C. Thomassen. Interval representations of planar graphs. *J. Combin. Theory Ser. B*, **40**(1):9–20, 1986.

[171] S. Toruńczyk. Flip-width: cops and robber on dense graphs. *arXiv:2302.00352* preprint, 2023.

[172] W. T. Tutte. All the King's Horses (A Guide to Reconstruction). In (J. A. Bondy and U. S. R. Murty, editors) *Graph Theory and Related Topics*. Academic Press, New York, 1979.

[173] S. M. Ulam. *A Collection of Mathematical Problems*, volume 8 of *Interscience tracts in pure and applied mathematics*. Interscience Publishers, 1960.

[174] J. van den Heuvel and D. R. Wood. Improper colourings inspired by Hadwiger's conjecture. *J. Lond. Math. Soc.*, **98**:129–148, 2018.

[175] J. H. van Lint. *Introduction to Coding Theory*, volume 86. Springer Science & Business Media, 1998.

[176] M. Von Rimscha. Reconstructibility and perfect graphs. *Discrete Math.*, **47**:283–291, 1983.

[177] K. Wagner. Über eine Eigenschaft der ebenen Komplexe. *Math. Ann.*, **114**:570–590, 1937.

[178] D. R. Wood. On tree-partition-width. *European J. Combin.*, **30**(5):1245–1253, 2009.

[179] D. R. Wood. Cliques in graphs excluding a complete graph minor. *Electron. J. Combin.*, **23**(3):P3.18, 2016.

[180] D. R. Wood. Defective and clustered graph colouring. *Electron. J. Combin.*, **DS23**, 2018.

[181] Y. Yang. The reconstruction conjecture is true if all 2-connected graphs are reconstructible. *J. Graph Theory*, **12**:237–243, 1988.