

STRUCTURAL SPARSENESS AND COMPLEX NETWORKS

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der
RWTH Aachen University zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

DIPLOM-INFORMATIKER FELIX JON REIDL

aus

FREIBURG IM BREISGAU

Berichter: Univ.-Prof. Dr. Peter Rossmanith
prof. RNDr. Jaroslav Nešetřil, DrSc

Tag der mündlichen Prüfung: 04.12.2015

Diese Dissertation ist auf den Internetseiten der Hochschulbibliothek online verfügbar.

To my loving family.

ABSTRACT

The field of *complex networks* has seen a steady growth in the last decade, fuelled by an ever-growing collection of relational data that our life in the information age generates. While several structural commonalities of complex networks have been observed—*e.g.* low density, heavily skewed degree-distributions, or the small world property—so far no property has been discovered that is algorithmically exploitable on a broad scale.

Concurrently, the theory of *structurally sparse graphs* has been revolutionised by Robertson and Seymour’s graph minors programme. Many tools and techniques, developed as ‘by-products’ in the programme, have had a tremendous impact on the research of parametrised and approximation algorithms. They in particular enabled the development of several *algorithmic meta-theorems*, that is, algorithms that work for a large spectrum of problems on sparse inputs.

In this thesis, we work towards bringing the field of structural sparse graphs and the field of complex networks closer together. We identify two notions of structural sparseness based on the density of shallow minors as keys for this endeavour: classes of *bounded expansion* and *nowhere dense* classes as introduced by Nešetřil and Ossona de Mendez in their seminal work on a robust theory of sparseness. In the following, we demonstrate that these sparse classes admit efficient algorithms for a huge number of problems, some of which have applications in domain-specific areas of network science. We further prove that several fundamental network models exhibit these properties and demonstrate empirically that this also holds true for a selection of real-world networks from various domains.

As a result, we can state that the theory of structurally sparse graphs is applicable to complex networks and, as a corollary, so is the rich algorithmic toolkit it provides. This connection offers researchers from both the field of algorithmic graph theory and network science new approaches, insights, and productive questions.

ZUSAMMENFASSUNG

Durch die stetig wachsende Menge an relationalen Daten, die unser tägliches Leben im Informationszeitalter erzeugt, hat sich das Forschungsgebiet der *komplexen Netzwerke* im letzten Jahrzehnt enorm entwickelt. Obwohl viele strukturelle Gemeinsamkeiten solcher Netzwerke bekannt sind—etwa ihre geringe Dichte, die starke Rechtsschiefe ihrer Gradverteilung oder das *small-world* Phänomen—, kannte man bisher noch keine Eigenschaft, die auf breiter Front algorithmisch nutzbar ist.

Parallel dazu hat das Graph-Minoren-Programm von Robertson und Seymour die Theorie der strukturell dünnen Graphen revolutioniert. Viele Werkzeuge und Techniken, die als ‚Nebenprodukt‘ im Rahmen dieses Programms entwickelt wurden, hatten tiefgreifende Auswirkungen auf die Erforschung von parametrisierten und approximativen Algorithmen. Insbesondere ermöglichten sie die Entwicklung mehrerer *algorithmischer Metatheoreme*, also Algorithmen, die ein großes Spektrum von Problemen auf strukturell dünnen Eingaben lösen.

Ziel dieser Arbeit ist es, das Forschungsgebiet der strukturell dünnen Graphen und das der komplexen Netzwerke näher zusammenzubringen. Zwei Varianten von struktureller Dünnheit, basierend auf der Dichte von seichten Minoren (*shallow minors*), stellen sich dabei als Schlüsselkonzepte heraus: die von Nešetřil und Ossona de Mendez in ihrer wegweisende Arbeit zu einer robusten Definition des Dünnheitsbegriffs eingeführten Graphklassen mit beschränkter Expansion (*bounded expansion*) und Klassen, die nirgends dicht (*nowhere dense*) sind. Im Folgenden wird zum einen dargelegt, dass diese Klassen strukturell dünner Graphen es erlauben, effiziente Algorithmen für eine Vielzahl von Problemen zu entwerfen—darunter insbesondere Probleme, die in Teilgebieten der Netzwerkforschung Anwendung finden. Zum anderen wird bewiesen, dass fundamentale Netzwerkmodelle diese strukturellen Eigenschaften innehaben, und empirisch belegt, dass dieser Sachverhalt erfolgreich auf konkrete Netzwerke aus verschiedensten Domänen übertragbar ist.

Somit ist die Theorie der strukturell dünnen Graphen—und damit insbesondere die Vielzahl der von ihr bereitgestellten algorithmischen Werkzeuge—tatsächlich auf komplexe Netzwerke anwendbar. Sowohl die algorithmische Graphentheorie als auch die Netzwerkforschung erhält aus diesem Brückenschlag neuartige Ansätze, Einsichten und Fragestellungen.

PREFACE

This thesis has several lines of origins that, told in chronological order, would be rather chaotic. The result presented herein are strictly *not* chronological for this exact reason.

My work began when I joined my Peter Rossmanith’s group in 2011. Initially, my colleague, Somnath Sikdar, and I set out to find a kernelisation algorithm for CHORDAL VERTEX DELETION—a notoriously difficult problem, even though it is fixed parameter tractable due to a result by Dániel Marx [179]—by restricting ourselves to sparse graph classes. We quickly found a solution for planar graphs of bounded degree and tried to generalise it to larger graph classes. Inevitably, our line of research crossed paths with the meta-kernelisation result by Bodlaender, Fomin, Lokshtanov, Penninkx, Saurabh, and Thilikos [23]. After a while, we realised that what we set out to do was already solved by this result; but it introduced us to the idea of protrusion replacement. We decided to broaden our scope: our ideas developed for CHORDAL VERTEX DELETION seemed to work well with the general-purpose reduction rule provided by the meta-kernelisation framework. After playing around with the Robertson-Seymour decomposition for a bit, it became apparent that a subsequent paper by Fomin, Lokshtanov, Saurabh, and Thilikos [101] already provided linear kernels for a large number of problems if the inputs are restricted to classes excluding a minor. The natural attempt was to extend the result to the next larger class we were aware of: classes excluding a *topological* minor. We finally succeeded; both by finding the right property that problems need to satisfy in order to be treatable, and by introducing the first variation of what is called the *twin class lemma* in this thesis. Through a convoluted sequence of events, the original version of that paper (co-authored with Alexander Langer and Peter Rossmanith) only exists as a preprint: it was subsumed by our cooperation with Eun Jung Kim, Igansi Sau, and Christophe Paul [156, 157] which streamlined our techniques and introduced the concept of a *protrusion decomposition*, an idea we previously used only implicitly. The focal point of the events that lead up to our joint effort was the Dagstuhl Seminar ‘Data Reduction and Problem Kernels’, where Somnath presented our results. During the subsequent discussions, Dimitrios Thilikos mentioned bounded expansion classes—a concept I until that day had never heard of. I tried to extend our result to these classes during the seminar, but failed: the *twin class lemma* simply did not hold in the generality we needed.

We tried to extend our results again in fall 2012 while visiting Fedor Fomin at the University of Bergen. After some discussions with Daniel Lokshtanov, Saket Saurabh, Michał Pilipczuk, Markus Dregi, and Pål

Drange, we worked out that we cannot possibly obtain a better kernel for FEEDBACK VERTEX SET in bounded expansion classes and decided to focus on DOMINATING SET instead. It would be a problem that haunted us for almost two years until we finally resolved it; and the group of people involved had grown to include Marcin Pilipczuk, Fernando Sánchez Villaamil, and Stefan Kreutzer. But we succeeded and obtain linear kernels for bounded expansion classes and almost-linear kernels for nowhere dense classes. The writing of this paper was, in particular in correspondence with Michał, what made me finally appreciate working with nowhere dense classes. In our many visits to Bergen, it was also Fedor and his group who introduced us to Nešetřil and Ossona de Mendez' book 'Sparsity' that helped shape my further research.

In parallel, also at the end of 2012, Somnath and I travelled to Masaryk University in Brno, following an invitation by Petr Hliněný. Our failure in Bergen had made me think: The property of being *treewidth bounding*, which we crucially used in our work on classes excluding a topological minor, simply did not seem to work well with bounded expansion classes. What if we exchanged treewidth by a different type of width-measures? While preparing a talk on this issue, I recalled that Petr had talked about a measure called *treedepth* in the Dagstuhl seminar. I worked out that a parametrisation by the vertex cover number would—rather easily—extend our meta-kernelisation result to bounded expansion classes and we went to Brno with the question of whether parametrising by a *treedepth modulator* might be the correct way of moving forward. It turned out to work beautifully [114, 115, 116], and along the way we and our co-authors Jakub Gajarský, Petr Hliněný, Jan Obdržálek, Sebastian Ordyniak, Fernando Sánchez Villaamil, and Peter Rossmanith found several interesting ideas of how to improve the meta-kernelisation framework.

I presented this result at the Dagstuhl seminar 'Bidimensional Structures: Algorithms, Combinatorics and Logic' in spring 2013, where I met Blair Sullivan and Erik Demaine. All three of us had, in some form or another, pondered whether structural sparseness could be applied to real-world networks. Blair was primarily working on hyperbolicity, Erik—as far as I know—was focused on classes in which the bidimensionality framework was applicable. We put our heads together and formulated a rough plan of how to tackle the issue; talking primarily about network models and how one would go about proving structural sparseness for them. I introduced the idea of aiming to show bounded expansion, instead; knowing that Erdős-Rényi graphs in the sparse regime have this property (as shown by Nešetřil, Ossona de Mendez, and Wood [193]). Our first ideas—based on the structural sparseness of Erdős-Rényi graphs—turned out to be viable for what we now call *perturbations* of bounded-degree graphs. With a lot more work put in by us and our co-authors Somnath Sikdar, Fernando Sánchez Villaamil, and Peter Rossmanith, we collected a sizeable body

of results concerning the structural sparseness of network models, algorithms solving network-specific problems exploiting bounded expansion, and empirical results on p -centred colourings that suggested that, at least in some cases, the framework of structural sparseness seemed applicable in practice [65]. Some of the work has also been done during the ICERM seminar ‘Towards Efficient Algorithms Exploiting Graph Structure’, which provided an amazing research environment. Due to its interstitial nature, the paper underwent many iterations (for which I have to give Fernando as much credit as I take), and we are still in the process of finding a good venue to publish it.

The above is the (short) history of how this thesis came into existence. Specifically, the development of the *twin class lemma* (Chapter 6) and the characterisation of structural sparseness of the configuration model and Chung–Lu random graphs (Chapter 15) are primarily my work. This is true for many of the smaller results contained in this thesis as well, in particular, I streamlined many parts of the meta-kernelisation results in order to obtain a unified presentation. Our result on the kernelisation of DOMINATING SET is based a lot on my work, but the breakthrough was made by Marcin and Michał Pilipczuk together with Daniel Lokshtanov who extended it to what is here called the *charging lemma*, a technique that paved the way for the final kernel. In my presentation here of our joint work I have introduced several improvements: the running time to compute the kernel is now almost linear and its size depends only on measures related to two-shallow minors (instead of four-shallow minors); a satisfying improvement given the ‘depth-two’ nature of DOMINATING SET.

Finally, some of the work in this thesis is novel and has not been published. Most of it consist of smaller results that fill small gaps left open in the framework of structurally sparse graphs. These results enable us to obtain tighter bounds, the above mentioned improvement for the DOMINATING SET kernel is a consequence of it. A rather large piece of novel work are *dtf-augmentations*, a variation of transitive-fraternal augmentations introduced by Nešetřil and Ossona de Mendez [188], which not only improves the augmentation framework but also provides the basis for an empirical evaluation of a network’s structural density (Chapter 16). In order to apply these augmentations properly, it was necessary to re-prove many known results.

ACKNOWLEDGEMENTS

First and foremost, I sincerely thank my supervisor Peter Rossmanith. The time in Aachen was a great experience to hone my academic skills and challenge myself to delve into new areas. His group was always a place for amicable arguments and discussions and his support enabled me to expand my professional network and provided me with opportunities to present my results.

I thank Professor Nešetřil for reporting on my thesis; I could not have found a better judge for my work.

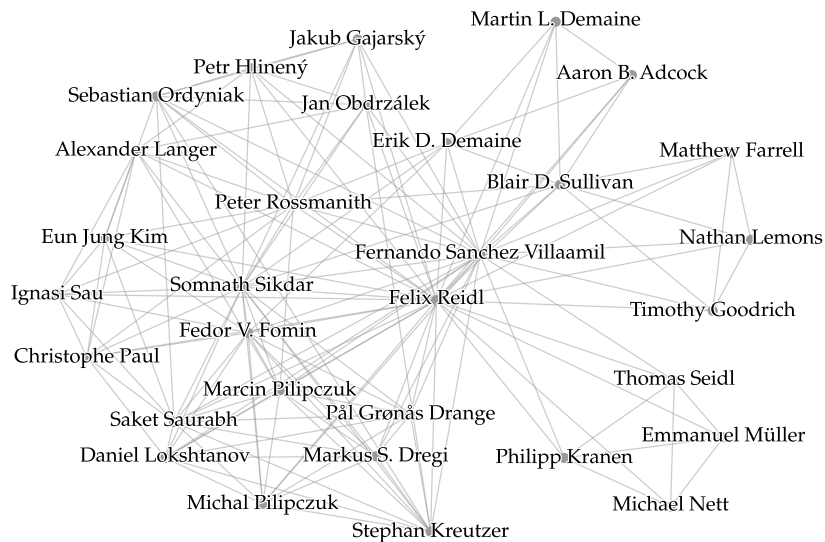
Over the years I have had the pleasure of working with many colleagues who all helped to bring me to this point. I enjoyed all our discussions, debates, disagreements, and diversions. In chronological order, I would like to thank Joachim Kneis, Alexander Langer, Somnath Sikdar, Fernando Sánchez Villaamil, Ling-Ju Hung, and William (Chen Li-Hsuan). Thanks, y'all! I further would like to thank Daniel Meister, Stefan Richter, and Daniel Mölle, who I had the pleasure of meeting while I was working in Peter's group as a student.

All of us owe Birgit Willms, the logistical mastermind and good soul of the group, more than we could possibly thank her. She has moreover been a good friend, with a much better taste in music than I could ever hope for. Thanks for the Jazz!

None of my work would have been possible without the support provided by the Deutsche Forschungsgesellschaft under the grants RO 927/12-1 "Theoretical and Practical Aspects of Kernelization" and RO 927/13-1 "Pragmatic Parameterized Algorithms". The mobility funds by the German Academic Exchange Service have further helped in our fruitful cooperation with the Bergen research group.

I would like to further thank Petr Hliněný for all the invitations to Brno. Even though I can still not share his enthusiasm for Matroids, our joint research has brought me in contact with many interesting topics. The same thanks go out to Fedor Fomin for his ongoing hospitality at Algoritmegruppen Bergen. His student and my good friend Pål Grønås Drange deserves extra mention for being an amazing host during our stays. I further thank all members of the Bergen group for creating such a great work environment. Dieter Kratsch and Henning Fernau deserve my gratitude for hosting the AMT workshop in Metz and Trier, it was always interesting and fun.

I had the fortune to work with many different co-authors in the last five years who all have taught me something new. It seems fitting to provide a small social network here to illustrate our joint work.



Thanks to all of you!

Several people from the parametrised and graph algorithms community have had, whether they know it or not, a very positive influence on me and my work. I would like to thank, in alphabetical order, Hans Bodlaender, Rod Downey, Mike Fellows, Pim van t' Hof, Frances Rosamond, and Dimitrios Thilikos.

Finally, I would like to thank my friends and family. The continuous loving support of my parents, in particular, has given me the peace of mind without which all this work would not have been manageable. Of course, our family would not be the strong collective it is without my siblings Tim and Lisa.

I further owe my best friend, colleague, *and* co-author Fernando a debt of gratitude: we have come a long way together, and I cherish his creativity, knowledge, and humour. ¡Gracias, tio!

In the last months, Catherine has been my constant support and I owe her greatly for that. Our time together is the best antidote for the stress and worry associated with writing a thesis.

CONTENTS

i	STRUCTURALLY SPARSE GRAPHS: A ROBUST NOTION	1
1	INTRODUCTION	3
1.1	Organisation and summary of results	7
2	PRELIMINARIES	9
2.1	Basic definitions and notation	9
2.2	Minors and shallow minors	13
2.3	Algebraic operations on graphs	15
2.4	Well-quasi orderings of graphs	17
2.5	Width measures	18
2.6	Graph measures and limits over classes	21
2.7	Fixed-parameter tractability and kernelisation	22
2.8	Domination problems	23
2.9	Propositional logic over graphs	24
2.10	Graph surgery toolkit	26
3	NOTIONS OF SPARSENESS	29
3.1	Degeneracy	31
3.2	Embeddable graphs	34
3.3	Excluded (topological) minors	36
4	BOUNDED EXPANSION CLASSES	43
5	NOWHERE DENSE CLASSES	47
ii	THE BOUNDED EXPANSION TOOLKIT	51
6	MORE ON SHALLOW MINORS	53
6.1	Stable minors	55
6.2	Weighted topological minors	61
7	AUGMENTATIONS & DECOMPOSITION BY COLOURING	65
8	NEIGHBOURHOOD STRUCTURES	83
8.1	The twin and charging lemma	83
8.2	A characterisation by neighbourhood complexity	86
8.3	Bounding neighbourhood complexity	88
9	FURTHER CHARACTERIZATIONS	95
9.1	Large degree and subgraph-grads	96
9.2	Generalised colouring numbers	96
9.3	Wideness	98
iii	ALGORITHMS AND PREPROCESSING USING STRUCTURAL SPARSENESS	99
10	ALGORITHMIC APPLICATIONS	101
10.1	Faster local search	101
10.2	Using low-treedepth colourings	105
10.3	Approximating r -domination: Dvořák's algorithm	106
10.4	Computing neighbourhood sizes	109

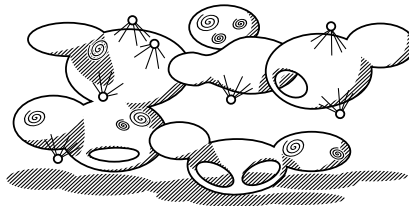
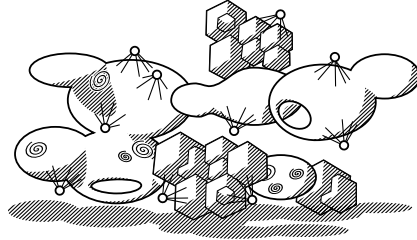
11	THE META-KERNELISATION FRAMEWORK	115
11.1	Finite Integer Index	118
11.2	WQO & representative sets	124
11.3	Protrusions: replacement and decomposition	128
12	META-KERNELISATION BEYOND EXCLUDED MINORS	135
12.1	Classes excluding a topological minor	135
12.2	Interlude: the need for a stronger parameter	139
12.3	Classes of bounded expansion and beyond	140
12.4	Further work and open questions	146
13	A LINEAR KERNEL FOR DOMINATING SET	149
13.1	Finding a small domination core	149
13.2	Computing a kernel	158
iv	COMPLEX NETWORKS AND STRUCTURAL SPARSENESS	163
14	COMPLEX NETWORKS	165
14.1	Basic network statistics	169
14.2	Algorithmic questions	174
15	NETWORK MODELS	179
15.1	Expansion of random graph models	181
15.2	Fixed degree distributions	184
15.3	Degree distributions, tails and expansion	189
15.4	Random graphs and perturbations	201
15.5	Evolutionary and attachment models	203
15.6	Small world models	205
15.7	Kleinberg's model: dense when useful	208
15.8	Other models	210
16	THE GRAD OF REAL-WORLD GRAPHS	213
16.1	Augmentation indegrees and aug-aug plots	213
16.2	Small networks	217
16.3	Medium-sized networks	219
16.4	Large networks	221
16.5	Huge networks	224
16.6	A verdict	227
17	COMPUTING LOW TREEDEPTH COLOURINGS	229
17.1	The small network benchmark	230
17.2	Testing centred colourings efficiently	232
17.3	Colouring the corpus	234
v	CONCLUSION	239
18	STRUCTURAL SPARSENESS AND COMPLEX NETWORKS	241
vi	BIBLIOGRAPHY	245
vii	INDEX	267
viii	APPENDIX	275
A	IMPORTANT PROBLEMS	277
B	LIST OF AUTHOR'S PUBLICATIONS	281

Part I

STRUCTURALLY SPARSE GRAPHS: A ROBUST NOTION

When you ask people, 'What's the opposite of fragile?,' they tend to say robust, resilient, adaptable, solid, strong. That's not it. The opposite of fragile is something that gains from disorder.

— Nassim Nicholas Taleb



INTRODUCTION

The theory of graphs is the fundamental study of relations in their purest, non-trivial form: binary connections between abstract points. And as so often in combinatorics, this simple *assemblage* of trivial objects results in a dazzlingly rich theory of seemingly endless depths. We have far progressed beyond Euler's musings about bridges and Guthrie's colourings of maps: graph theory encompasses cuts and separators, embeddings, decompositions, flows, matchings, random and extremal graphs, well-quasi-orders, infinite graphs, graph limits, and so much more. Squaring these facets with algorithmic questions results in a body of work that cannot be contained in a single area of research any more.

Let us zoom into that part of the above patchwork that provides the context for this thesis. First and foremost, we will consider *sparse* graphs on the one side and *efficient* algorithms on the other. Both terms need further clarification which we will gradually develop by presenting the motivation for this thesis: the applicability of algorithmic frameworks to real-world problems.

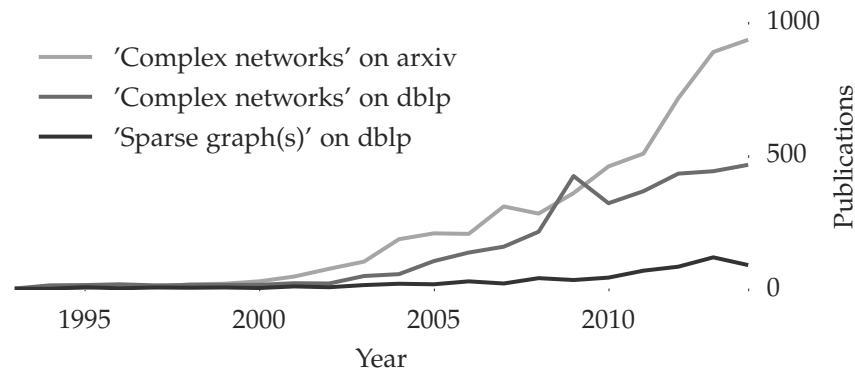
Big data et al.

The past two decades have propelled us head-first into the information age: As Moore's law and its cousins have predicted, in only a fraction of our specie's lifetime we have amassed, transmitted, and processed more data than ever before. Computer science has transformed along this exponential growth curve and brought forth several frameworks and schools-of-thought to deal with the 'data deluge': Parallel computation, database theory, data mining, and big data analytics were all born out of necessity. One important aspect of dealing with data in the petabyte-age (soon to be exa- or zetabyte) are algorithmic matters, the efficient processing of large data sets. Because our focus lies on graphs we need to explicate how they relate to big-data challenges—and for that we have to consider another progeny of the information age: network science.

Network science

Not spear-headed by computer science but by physics and complex system theory, *network science* is the research of large real-world graphs. Data sets have been gathered from biology, social sciences, infrastructures, and various databases pertaining to movies, scientific publications, and political affiliations—under the maxim: The domain does not matter, all networks exhibit the same structure. And indeed, several observations about real-world graphs seem to hold almost universally: these graphs are sparse, have a very small diameter, and while they do contain a good deal of randomness they do *not* resemble random graphs (in the sense of Erdős, Rényi, and Gilbert). However, the silver-bullet approach pursued in the early days of network sci-

ence has not, as hoped, brought forth a grand unifying theory; instead there is a flurry of observations, models, and results that sometimes contradict each other. Nonetheless, the gathered data and observations have inspired a lot of researchers in the last fifteen years to dedicate their time to networks. Consider the following plot showing the number of publications per year (not cumulative!) containing the keyword ‘complex network’ in the arxiv and dblp databases.



The interest in complex networks is, to this day, still on a steep ascend; in computer science, physics, and other fields. But let us also consider the third plot: the number of publications mentioning sparse graphs. Even if we take into consideration the different scales and publishing cultures of applied and theoretical fields, there is hardly a ‘booming interest’ in sparse graphs. And that despite the fact that the last decades have seen immense progress in that field: the work by Robertson and Seymour has infused the algorithmic theory of sparse graphs with a plethora of tools and techniques that have been employed to design a whole spectrum of efficient solutions—be it in the form of approximation, exact, or parametrised algorithms. This difference is what motivated me to write this thesis: the gap between an applied field that works *exclusively* with sparse graphs and a theory that provides efficient algorithms for exactly these structures. Regardless of the tremendous interest in real-world sparse graphs, the gap exists: the algorithmic theory of sparse graphs seems to be left out of the loop. Rephrased as a provocative question: Why are our beautiful algorithms not used to solve real-world problems?

Theory and practice

The first, reflexive, answer is: Because our algorithms are purely theoretical. And indeed, large polynomial dependencies, astronomical hidden constants, or non-constructive tools often prevent algorithms from being practical. That does, however, not mean that such tools are *necessary*. From what I can observe, in that part of the algorithms community I have the pleasure of interacting with, there was and is always a subset of people who *do* care about such matters and who successfully optimise away such blockades—in fact, this ‘race’ of incremental improvements of an algorithm or theorem can be a strong motivational factor. Moreover, mathematics being ‘pure theory’ has

not stopped imaginary numbers from being applied in electrical engineering, group theory in quantum mechanics or elliptic curves in cryptography. And the algorithmic framework we use in this thesis, *parametrised complexity*, has in its very core the idea of being fundamentally *applicable*. As one of the fathers of this field, Rod Downey, put it [72]:

[...] parameterized complexity is a refined complexity analysis, driven by the idea that in real life data is often given to us naturally with an underlying structure which we might profitably exploit.

The multivariate nature of the parametrised approach, in which we do not measure complexity in only the input size n , but also in a secondary *parameter* k which is the ultimate arbiter of tractability and intractability, enables a refined analysis of *hard* problems. In doing so, we can carve out niches of tractability, identifying situations in which there is hope of finding a solution in reasonable time.

Abstract problems

We modify our answer as to why algorithms exploiting sparseness are not found in practice: if at least sometimes the algorithms are indeed applicable, it must be because the problems they solve are simply too abstract. And of course: we usually deal with idealised problems, simplified to a point where we can comfortably reason about them. However, the problems posed in context of networks are often *themselves* idealised. The abstraction from any specific domain automatically demands it. One of the best examples is the so-called motif counting problem (see Section 14.2): for a given small pattern graph H , we want to count how often it appears in a large host-graph G . If we additionally allow the vertices of G to be labelled, we are already dealing with a problem that has seen great popularity in computational biology. Not to mention that many algorithms on sparse graphs are not only applicable to a single problem, but a whole *class* of problems—we will mention several such results later on and prove some ourselves. So not only should some of our algorithms be implementable *and* relevant to network related questions, the broad algorithmic results available tell us that a huge class of problems are tractable in sparse classes. To turn these algorithmic meta-results into practical algorithms is certainly difficult, but as my former colleague and co-author Alexander Langer has proved with his work on Courcelle's theorem, it *is* possible [169, 168, 167].

Sparseness vs. structurally sparse

There is one more factor that could be responsible for the lack of overlap between complex networks and the algorithmic theory of sparse graphs: Maybe our assumptions about sparseness are too strict to be applicable to real-world data. Indeed, we have until now glossed over an important detail, namely, what we mean by *sparse*. Complex networks are sparse in that they have a reasonably low average degree. In contrast, the algorithmic theory of *structurally* sparse graphs deals with stricter notions of sparseness: bounded degree, embeddability on

Applicability

a surface, excluded substructures, *etc.* Complex networks show little sign of such strict organisation, mainly due to their inherent randomness. We have identified the first real obstacle that might be responsible for the divide between network science and graph theory.

With the above motivation we can sharpen the notion of applicability and put it in a concrete context. The qualitative goal of this thesis is the answer to a simple question:

“Can the algorithmic tools derived from the theory of structurally sparse graphs be applied to real-world data?”

Let us deconstruct this question to arrive at a programmatic outline.

1. *What notion of sparseness?* We need to identify a notion of sparseness that applies to real-world networks. Through the dichotomy result by Ossona de Mendez and Nešetřil (Chapter 5) regarding structurally sparse classes we can be certain that either such a notion exists or we have to conclude that complex networks belong to a structurally dense class.
2. *Which algorithmic tools?* Next, we need to demonstrate that the identified notion of sparseness provides enough leverage to build efficient algorithms. In particular, we are interesting in meta-theorems which cover a large range of problems in order to make headway in the programme.
3. *Applied how?* Once we have collated an algorithmic toolkit, it is left to demonstrate that it can be used to solve problems that are interesting in the context of complex networks.

Solving real problems

In particular the first point comes with its own set of problems: since sparseness is a property of graph *classes*, we need a method to relate it to real-world instances. Fortunately, network science has brought forward a range of random *network models* for which the question is well-defined. We further provide empirical evidence that the model we will base our verdict on replicates the structural density of a selection of real-world networks from various domains.

There is a step missing in the above programme that would, without any doubt, demonstrate the applicability of structural sparse theory: an implementation that, on a set of real-world data, solves a relevant problem from a particular domain. I certainly set out to do so; however, the other parts of the programme proved to demand considerable time and effort. Solving real-world problems does not happen by the simple application of a well-designed theoretical algorithm, it needs to be *engineered* to optimise resource usage and to use heuristic short-cuts where possible. Such engineering needs a lot of iterations to succeed and hence demands a significant amount of time. Moreover, identifying and solving a domain-relevant problem needs to happen in cooperation with a domain expert—several publications in the network science literature have fallen short of their promise to have any

meaningful impact, usually due to the proposed methods simply being not applicable in the respective domain. For the above reason, such proof-of-concept implementations have to be postponed. We do, however, demonstrate that certain algorithmic tools that are broad enough to promise real-world utility can be designed *and* engineered to be applicable; thus providing preparation for this ultimate step.

1.1 ORGANISATION AND SUMMARY OF RESULTS

This thesis is divided into five parts. In this part, we lay out its central ideas and programme, paired with an introduction of the relevant notation (Chapter 2). Beside a short introduction to *parametrised complexity* in Section 2.7, we will prove a few basic tools in Section 2.10.

Sparse class hierarchy

We then take a tour through the hierarchy of structurally sparse graph classes and survey important algorithmic results along the way (Chapter 3). The part ends with the topmost class in the hierarchy, so-called *nowhere dense* graph classes and classes that have *bounded expansion* (Chapters 4 and 5).

Algorithmic toolkit

These classes will be the primary focus of the remainder of the thesis and we devote the whole second part to them: in Chapter 6 we introduce a variation of *shallow minors* that proved to be immensely useful. In Chapter 7 we will use these concepts to re-prove the existence of *low treedepth colouring* using a slightly different approach than in the literature. Our tool of choice, so-called *dtf-augmentations*, is an applicable variation of the concept of transitive-fraternal augmentations. In particular, we use dtf-augmentations for empirical results presented in the fourth part. In Chapter 9 we present two entirely novel results: the *twin class lemma*, which is one of the centrepieces in our work on preprocessing of sparse classes (contained in the third part), and a characterisation of bounded expansion classes by a measure called *neighbourhood complexity*. In Chapter 9 we survey several other known characterisations of bounded expansion and nowhere dense classes for completeness.

Algorithms & preprocessing

The third part of the thesis contains algorithmic results obtained by applying results from the second part. Besides some so-far unexplored applications of these tools to local search (Section 10.1), we present a reformulation of Dvořák's approximation algorithm for *r-DOMINATING SET* using dtf-augmentations. This reformulation in particular enables us to show better bounds in our result on preprocessing *DOMINATING SET* in Chapter 13. We furthermore present in Section 10.4 a novel algorithm that, for bounded expansion classes, computes for all vertices the size of their respective *r*-neighbourhood in linear time—an algorithm that has potential applications in complex networks as we explicate in the later Section 14.2.

Chapter 11 contains our novel results for classes excluding a topological minor, bounded expansion and nowhere dense classes: we show that the *meta kernelisation* framework, that demonstrates the existence

of efficient preprocessing for hundreds of problems, can be extended to these graph classes by applying a different interpretation to previous results. We complement these results in Chapter 13 by providing an explicit linear kernel for DOMINATING SET on bounded expansion classes and an almost-linear kernel in nowhere dense classes. This result uses the above-mentioned twin class lemma (alongside an extension of it) and our reformulation of Dvořák’s algorithm; notably, we do not (and cannot) rely on any of the techniques used in smaller graph classes.

Complex networks

In the fourth part we turn our attention to complex networks. After a concise introduction of the topic, we argue that algorithms designed using the toolkit derived from the theory of structurally sparse graphs can be used to answer import network-related questions. To argue that this is indeed possible, we show in Chapter 15 that important random graph models used to simulate complex networks are—in certain parametric settings—structurally sparse with high probability. In fact, we provide a full characterisation of when these random graph models are structurally sparse and when not. This result will play a big role in the subsequent empirical evaluation. Further, we consider the *perturbation* of classes and provide conditions under which such perturbations of structurally sparse classes preserves the sparseness property. The remainder of that chapter is devoted to other, more specialised models used in network science.

Empirical results

The theoretical results on network models strongly suggest that they are structurally sparse. However, this might simply be an aspect where models and real-world networks differ. We therefore devise a test to assess the structural density of concrete instances and apply this test to a corpus of real-world networks. It turns out that we can for most networks confidently state that they *are* structurally sparse and we could not find a single example where that is provably *not* the case (for several networks the result is not statistically conclusive). We further report on our experiment of applying *low treedepth colourings* to the same network corpus and highlight some of the heuristical improvements we successfully applied to the colouring pipeline.

*So perhaps the best thing to do
is to stop writing Introductions and get on with the book.*
— Alan A. Milne

2.1 BASIC DEFINITIONS AND NOTATION

We denote sets by upper case letters and family of sets by calligraphic letters. For a set family \mathcal{F} , we use the shorthand $\bigcup \mathcal{F} := \bigcup_{X \in \mathcal{F}} X$. For two sets A, B we write $A \uplus B$ for their union if A and B are disjoint and we want to emphasize that fact.

 $\bigcup \mathcal{F}, \uplus$

Graphs in the following are simple and finite. For certain definitions, it is useful to assume that vertices of graphs are taken from a countable infinite universe: in particular, this allows us to assign a natural number with every vertex which we call the *natural index* of that vertex. The set of vertices and edges of a graph G are denoted by $V(G)$ and $E(G)$, respectively. The number of vertices and edges will be abbreviated by $|G| := |V(G)|$ and $\|G\| := |E(G)|$. A bijection between the vertex sets of two graphs G, G' is called a *graph isomorphism* if it preserves the edge relation. In the case that such a function exists, we call the graphs G, G' *isomorphic* and write $G \simeq G'$.

 $K(G), \simeq, |\bullet|, \|\bullet\|$

We denote by $K(G)$ the set of connected components; in the case of a digraph \vec{G} we let $K(\vec{G})$ be the set of *weakly* connected components.

Induced subgraph, induced edge set

A graph H is an *induced subgraph* of another graph G if a graph isomorphic to H can be obtained from G by deleting vertices. We write $H \subseteq_i G$ for this relation. For a graph G and a vertex set $W \subseteq V(G)$, we denote by $G[W]$ the subgraph induced by the vertices W , *i.e.* the graph obtained from G by removing all but the vertices contained in W . Similarly, for any edge set $E \subseteq E(G)$ we denote by $E[W] := E \cap \binom{W}{2}$ those edges of E that have both endpoints in W .

A graph H is a *subgraph* of another graph G if a graph isomorphic to H can be obtained from G by deleting vertices and edges. We write $H \subseteq G$ for this relation.

Neighbourhood, closed –

For a vertex $v \in G$, we denote by $N_G(v) = \{u : uv \in E(G)\}$ the *open neighbourhood* of v and by $N_G[v] = N_G(v) \cup \{v\}$ the *closed neighbourhood* of v in G . We extended this notion to sets of vertices $X \subseteq V(G)$ as follows: $N_G[X] = \bigcup_{v \in X} N_G[v]$ and $N_G(X) = N_G[X] \setminus X$. We usually omit the subscript G . The *degree* of a vertex $v \in V(G)$ is the number of its neighbours, *i.e.*, $\deg(v) = |N(v)|$.

Colouring, proper colouring, refinement

A *colouring* of a graph G refers to a function $c : V(G) \rightarrow \mathbb{N}$. It is *proper* if for every $uv \in E(G)$, $c(u) \neq c(v)$. By a slight abuse of notation, we set $c(G) := c(V(G))$. The *cardinality* of a colouring is $|c(G)|$.

Another colouring c' is a *refinement* of c if for every pair $x, y \in G$ it holds that $c(x) \neq c(y) \rightarrow c'(x) \neq c'(y)$.

Complete graphs, tournament,
 $K_n, \omega, \#\omega$

By K_n we denote the *complete graph* on n vertices, also called an *n-clique*. The *clique number* $\omega(G)$ denotes the largest subgraph of G that is complete. By $\#\omega(G)$ we denote the *number* of complete subgraphs in G , not counting single vertices (we do, however, count every edge in G as a K_2). With $K_{i,j}$ we denote the complete bipartite graph with partite sets of size i and j , respectively. A *tournament* is a simple directed graph whose underlying graph is complete.

Paths, P_ℓ

For a path P we denote its *length*, in accordance with the above notation, by $\|P\|$. We use the notation P_ℓ for a path of length $\ell - 1$, i.e. a path with ℓ vertices. In particular, $P_1 = K_1$ and $P_2 = K_2$. For a path P , we denote by $P[i]$ the i^{th} vertex on the path. Hence $P[1]$ is the start and $P[\|P\|]$ the end of the path (we consider the empty graph *not* as a path). For $u, v \in P$ we denote by $P[u, v]$ the sub-path connecting u and v in P including u, v . By $P(u, v)$ we mean $P[u, v] \setminus \{u, v\}$, i.e. the sub-path between u and v excluding u, v . The notation $P[u, v)$ and $P(u, v]$ is used to exclude one of the respective endpoints and including the other. For paths P and P' with $P[\|P\|] = P'[1]$, that is, P ends with a vertex that P' starts with, we write PP' to denote the path obtained from concatenating them. If we particularly care about the endpoints of a path, we will sometimes use the notation uPv to denote a path with endpoints u and v .

Rooted tree, parent, child, sibling

To avoid confusion, we will distinguish between vertices of general graphs and vertices of trees by calling the later *nodes*. A *rooted tree* T is a digraph whose underlying graph is a tree and in which every vertex has in-degree exactly one except one vertex whose in-degree is zero. The source of this digraph is the *root* of T and denoted by $\text{root}(T)$. Nodes with out-degree zero are called *leaves* and the set of leaves is denoted by $\text{leaves}(T)$. The arcs of T represent the *parent relationship* between nodes: for $xy \in T$ we say that x is the *parent* of y whereas y is a *child* of x . Nodes with the same parent are called *siblings*. Hence the children of a node x are exactly the set $N_T^+(x)$. Note that we deviate from the usual orientation of a tree here: letting parent nodes have arcs to their children (instead of from) will make sense in the context of *treedepth decompositions* introduced below.

Spider

A *spider* is a rooted tree in which every internal vertex except the root has exactly degree two. Alternatively, a spider is a subdivision of a star.

Root path, ancestor, descendant,
depth, height

The *root path* of a node $x \in T$ is the unique x - $\text{root}(T)$ -path in the undirected graph underlying T and we denote it by $\text{rpath}_T(x)$. The *depth* of x is $\text{depth}_T(x) = |\text{rpath}_T(x)|$. The *height* of a tree is $\text{height}(T) := \max_{x \in \text{leaves}(T)} \text{depth}_T(x)$. Hence a single vertex constitutes a tree of height one. For nodes x, y with $x \in \text{rpath}_T(y) \setminus \{y\}$ we say that x is an *ancestor* of y , conversely y is a *descendant* of x . In particular, x is neither an ancestor nor a descendant of itself. The set

of ancestors and descendants of x is denoted by $\text{anc}_T(x)$, $\text{desc}_T(x)$, respectively.

Rooted forest

A *rooted forest* F is a disjoint union of rooted trees. We define the height of F as $\text{height}(F) := \max_{T \in \mathcal{K}(F)} \text{height}(T)$. For $x \in F$ extend $\text{rpath}_F(x) := \text{rpath}_T(x)$, $\text{depth}_F(x) := \text{depth}_T(x)$, $\text{anc}_F(x) := \text{anc}_T(x)$, $\text{desc}_F(x) := \text{desc}_T(x)$ where in all cases $T \in \mathcal{K}(F)$ is that tree in F which contains x .

Closure, $\text{clos}(F)$

The *closure* $\text{clos}(F)$ of a forest F is the digraph with node set $V(F)$ and arc set $\{xy \mid y \in \text{desc}(x)\}$. Informally, it is the digraph obtained from F by connecting every node to all its descendants.

Rooted subtree, T_x

For a node x in a rooted tree T , the *subtree of T rooted at x* is defined as $T_x := T[x \cup \text{desc}(x)]$, i.e. the subtree of T induced by the node set $S \cup \{x\}$ with root x .

Least common ancestor, – closure

Given two nodes x, y in a rooted tree T , their *least common ancestor* (lca) is the unique shared ancestor of maximal depth. We denote this vertex by $\text{lca}_T(x, y)$ in the following. For a node set $X \subseteq V(T)$, the *lca closure* $\text{clos}_T^{\text{lca}}(X)$ is defined as the minimal superset $Y \supseteq X$ such that

$$x, y \in Y \implies \text{lca}(x, y) \in Y$$

holds. The following facts about the least common ancestor closure can be considered folklore.

Lemma 1 (Lca closure). *For every rooted tree T and node set $X \subseteq V(T)$ the set $\text{clos}^{\text{lca}}(X)$ is unique and it holds that $|\text{clos}^{\text{lca}}(X)| \leq 2|X|$. Furthermore, every connected component of $T \setminus \text{clos}^{\text{lca}}(X)$ is adjacent to at most two nodes in $\text{clos}^{\text{lca}}(X)$.*

Proof. Given a set X we construct the closure of X iteratively. Set $X_0 = X$ and construct the set X_{i+1} from the set X_i as follows: a node $y \in T \setminus X_i$ is a *closure node* at step i if $V(T_y) \cap X_i$ contains a pair of nodes z, z' such that $y = \text{lca}(z, z')$. We construct X_{i+1} from X_i by adding all closure nodes. We stop the process as soon as $X_{i+1} = X_i$.

The procedure trivially terminates after at most $|V(T)|$ steps. Let Y be the set constructed by the above procedure. That Y is an lca closure follows from the termination of the procedure, that it is unique and minimal follows easily by induction: since for two nodes $x, y \in X$ their lca is uniquely determined in T , it must be contained in any lca closure that contains x and y . The necessity of every single vertex in the set and hence minimality follows inductively.

Consider a connected component of $C \in \mathcal{K}(T \setminus Y)$. Assume it has at least three neighbours $a, b, c \in Y$ where a is closest to $\text{root}(T)$. But then we have that $\text{lca}(b, c) \in C$ and Y would not be an lca closure, contradiction.

Finally, we can bound the size of the closure. Since every connected components of $T \setminus Y$ has at most two neighbours in Y , we can replace it either by an edge (if it has exactly two neighbours) or remove it. The result is a rooted tree T' with node set Y where every internal node

has at least two children (barring trivial cases like the empty tree or a tree of size one). Note that $\text{leaves}(T') \subseteq X$; thus it follows that

$$|Y| = |T'| \leq 2 \text{leaves}(T') \leq 2|X|,$$

as claimed. \square

Graph class, – property, $\mathcal{G}_{\blacklozenge k}$

A *graph class* is a collection of graphs closed under isomorphism. We denote graph classes by calligraphic upper case letters. The term *graph property* is exchangeable with that of a graph class but will be used in a slightly different context. For a graph class \mathcal{G} , an integer k , a function $f: \mathcal{G} \rightarrow \mathbb{N}$ and any relation \blacklozenge on integers we define the subclasses $\mathcal{G}_{f\blacklozenge k} := \{G \in \mathcal{G} \mid f(G) \blacklozenge k\}$. If f is not specified we assume f to be the cardinality function $|\cdot|$. A common use of this notation will be the *truncation* $\mathcal{G}_{\leq k}$ of a class \mathcal{G} which contains all graphs of \mathcal{G} with at most k vertices.

Modulator

For a graph property \mathcal{P} and a graph G , a \mathcal{P} -*modulator* of G is a vertex set $X \subseteq V(G)$ such that $G \setminus X \in \mathcal{P}$. The problem of finding a minimal modulator \mathcal{P} will be called \mathcal{P} -DELETION in the following.

Some common graph properties with the corresponding deletion problem and the corresponding graph parameter are listed in the table below.

Graph property \mathcal{P}	\mathcal{P} -DELETION	$\tau_{\mathcal{P}}$
Edgeless	VERTEX COVER	vc
Cycle-free	FEEDBACK VERTEX SET	–
Bipartite	ODD CYCLE TRANSVERSAL	–
Treedepth $\leq t$	TREEDPTH- t -DELETION	td _{t} $^{\Delta}$
Treewidth $\leq t$	TREEWIDTH- t -DELETION	tw _{t} $^{\Delta}$

Twins, true –, false –

A recurring theme in this thesis will be that of *twin vertices*. Two vertices x, y are *true twins* in a graph G if $N[x] = N[y]$ (which implies that they share an edge) and *false twins* if only $N(x) = N(y)$ holds. We say that a graph is *twin-free* if no pair of vertices in it are false or true twins. We extend this notion to vertex subsets: a set $X \subseteq V(G)$ is *twin-free* if no pair of vertices from X are false or true twins in G .

Twin equivalence, – classes

It is easy to see that both twin relations form equivalence relations over vertices and hence we can partition the vertices of a graph into *twin classes*. Moreover, we can compute these classes in linear time using *partition refinement*.

Given a partition $\mathcal{S} = \{S_1, S_2, \dots, S_p\}$ of some ground set U and a *pivot set* $X \subseteq U$, the refinement of \mathcal{S} by X is the partition

$$\mathcal{S}|X := \{S_1 \cap X, S_1 \setminus X, S_2 \cap X, S_2 \setminus X, \dots\}$$

Proposition 1 (Habib, Paul, Viennot [134]). *There exists a data structure representing a partition \mathcal{S} of a ground set U that, for a pivot set $X \subseteq U$, can be in time $O(|X|)$ modified to represent the partition $\mathcal{S}|X$. The initialization of the data structure with any ordered partition takes time $O(|U|)$.*

Computing true twin classes of a graph G now can be achieved in linear time as follows:

```

Initialise  $\mathcal{X} \leftarrow \{V(G)\}$ 
for  $v \in V(G)$  do
   $\mathcal{X} \leftarrow \mathcal{X} \setminus N[v]$ 

```

After the every vertex has been processed the partition \mathcal{X} contains exactly the true twin classes of G . By Proposition 1 it can be implemented to run in time

$$O(|G| + \sum_{v \in G} |N[v]|) = O(|G| + \|G\|).$$

To compute the false twin classes we modify the algorithm in a very subtle and satisfying manner: we simply refine by the open neighbourhood $N(v)$ instead of the closed neighbourhood $N[v]$.

We conclude this section by packaging the above observations in a lemma which should be folklore but was surprisingly hard to find.

Lemma 2 (Computing twin classes). *We can compute the true or false twin classes of a graph and output the vertices grouped by those classes in linear time.*

Proof. We prove the correctness of the above algorithm here; the running time was already discussed. For an input graph G , consider the partition \mathcal{X} obtained from the trivial partition $\{V(G)\}$ by refining it with every open neighbourhood. We claim that \mathcal{X} exactly contains the false twin classes of G . To verify this, consider two distinct vertices $x, y \in G$.

Case 1: $N(x) \neq N(y)$. Let, wlog, $z \in N(x)$ be a neighbour of x (in the case that $xy \in G$ we might have $z \in \{x, y\}$). Then $x \in N(z)$ while $y \notin N(z)$. Since we refined the partition by $N(z)$, the vertices x and y are necessarily in different classes.

Case 2: $N(x) = N(y)$. Assume towards a contradiction that x, y are contained in different classes. Since in the initial (trivial) partition this was not the case, there must have been a refinement with some pivot neighbourhood $N(z)$ that separated x and y . This is exactly the case if $x \in N(z) \not\leftrightarrow y \in N(z)$. But then z is a neighbour of one of the two and not the other, contradiction.

We conclude that \mathcal{X} indeed contains the true twin classes of G . The proof for false twin classes works analogously. \square

2.2 MINORS AND SHALLOW MINORS

Fix graphs H, G . We define the following *embeddings* of H in G that give rise to natural relationships between graphs.

A *minor embedding* is a function $\phi: V(H) \rightarrow 2^{V(G)}$ such that for every $v \in H$ the subgraph $G[\phi(v)]$ is connected and for every $v \neq u \in H$ we have that

Minor embedding, depth

1. $\phi(v) \cap \phi(u) = \emptyset$ and
2. if $uv \in H$ then there exists $x \in \phi(u), y \in \phi(v)$ such that $xy \in G$.

We define the *depth* of the minor embedding ϕ as follows: for every vertex set $V_x = \phi(x), x \in H$ choose a *centre*, i.e. a vertex v_x such that

$$\max_{v \in \phi(x)} \text{dist}_G(v_x, v) = \text{rad}(G[\phi(v)]).$$

Now the depth of the embedding is defined as

$$\frac{1}{2} \max_{xy \in H} \text{dist}_G(v_x, v_y).$$

This slightly complicated definition enables us to define minor embeddings whose depth is a half-integer and thus a more ‘fine-grained’ distinction.

Minor, shallow $\prec, \preceq_m, \preceq_m^r, \triangleright$

If there exists a minor embedding ϕ for H in G we say that H is a *minor* of G and write $H \preceq_m G$. If there exists a minor embedding at depth r we say that H is an *r -shallow minor* of G and write $H \preceq_m^r G$. We denote the set of all r -shallow minors of G by $G \triangleright r$.

As one would expect, taking a shallow minor of a shallow minor results in a shallow minor of the original graph.

Proposition 2 ([192, Proposition 4.1]). *For every graph G and half-integers a, b it is true that*

$$(G \triangleright a) \triangleright b \subseteq G \triangleright \frac{(2a+1)(2b+1)-1}{2}.$$

Topological minor embedding, depth

A *topological minor embedding* is a pair of functions $\phi_V: V(H) \rightarrow G, \phi_E: E(H) \rightarrow 2^{V(H)}$ where ϕ_V is injective and for every $uv \in H$ we have that

1. $\phi_E(uv)$ is a path in G with endpoints $\phi_V(u), \phi_V(v)$ and
2. for every $u'v' \in H$ with $u'v' \neq uv$ the two paths $\phi_E(uv), \phi_E(u'v')$ are internally vertex-disjoint.

We define the *depth* of the topological minor embedding ϕ_V, ϕ_E as the half-integer $(\max_{uv \in H} |\phi_E(uv)| - 1)/2$.

Topological minor, shallow $\prec, \preceq_t, \preceq_t^r, \tilde{\triangleright}$

If there exists a topological minor embedding ϕ_V, ϕ_E for H in G we say that H is a *topological minor* of G and write $H \preceq_t G$. If there exists a minor embedding at depth r we say that H is an *r -shallow topological minor* of G and write $H \preceq_t^r G$. We denote the set of all r -shallow topological minors of G by $G \tilde{\triangleright} r$.

Again, unsurprisingly, we have the following result for the repeated construction of shallow minors.

Proposition 3 ([192, Proposition 4.2]). *For every graph G and half-integers a, b it is true that*

$$(G \tilde{\triangleright} a) \tilde{\triangleright} b = G \tilde{\triangleright} \frac{(2a+1)(2b+1)-1}{2}.$$

2.3 ALGEBRAIC OPERATIONS ON GRAPHS

We will make use of the following graph products. For graphs G_1, G_2 , the *complete join* $G_1 * G_2$ is the graph obtained by first taking the disjoint union of G_1, G_2 and then connecting every vertex of G_1 to every vertex of G_2 . For example, $G * \bar{K}_2$ is the graph obtained from G by adding two universal vertices. We will also use the notation $G * u$ to signify adding a single universal vertex u to G .

$G_1 \bullet G_2, G_1 * G_2$

The *lexicographic product* $G_1 \bullet G_2$ is the product graph with vertex set $V(G_1) \times V(G_2)$ and edges

$$(u, x)(v, y) \in E(G_1 \bullet G_2) \iff uv \in G_1 \text{ or } (u = v \text{ and } xy \in G_2).$$

In case the graphs have edge-weights ω_1, ω_2 we use the convention that

$$\omega_{1,2}((u, x)(v, y)) = \begin{cases} \omega_1(uv) & \text{if } uv \in G_1 \\ \omega_2(xy) & \text{if } u = v \text{ and } xy \in G_2 \end{cases}$$

provides the edge-weights for the product $G_1 \bullet G_2$.

Labelling, t -labelling

A *labelling* b of a set W is a bijective function $b: L \rightarrow W$ for some $L \subseteq \mathbb{N}$. We say that b is a t -labelling if the domain/codomain of b has cardinality t . For convenience, we will often treat b as a partial function over \mathbb{N} instead of L with the convention that $b(x) = \perp$ for $x \notin L$.

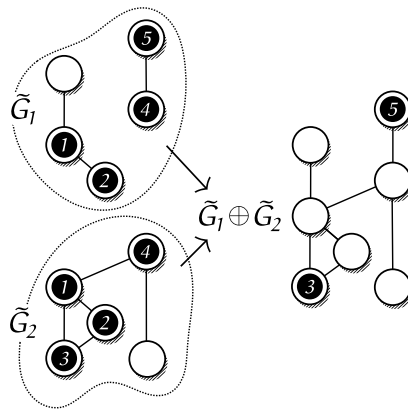
Boundaried graph

Definition 1 (t -boundaried graph). A t -boundaried graph is a triple ${}^\circ G = (V, E, b)$ where (V, E) is a graph and b is a t -labelling of a subset of V . The image of b is the *boundary of* ${}^\circ G$ and denoted by $\partial {}^\circ G = \text{img}(b)$. We denote by $G = (V, E)$ the *underlying graph* of ${}^\circ G$.

Boundaried class, ${}^\circ \mathcal{G}_t$

For a graph class \mathcal{G} we define the t -boundaried class ${}^\circ \mathcal{G}_t$ as those t -boundaried graphs whose underlying graphs are members of \mathcal{G} .

Boundaried graph allow us to build up larger graphs from smaller building blocks in well-defined manner. The *gluing* operator combines the boundaried graphs ${}^\circ G_1, {}^\circ G_2$ according to their labelling. The resulting graph has a boundary consisting of unmatched labels. The basic operation is *gluing* which combines two boundaried graphs into another boundaried graph.



Gluing, \oplus

Definition 2 (Gluing). Let ${}^\circ G_1 = (V_1, E_1, b_1)$ and ${}^\circ G_2 = (V_2, E_2, b_2)$ be t_1 - and t_2 -boundaried graphs, respectively. We assume that $V_1 \cap V_2 = \emptyset$ by an appropriate relabelling of ${}^\circ G_2$. Define $\phi = b_1 \circ b_2^{-1}$.

The *gluing operation* is denoted by $\circ G_3 = \circ G_1 \oplus \circ G_2$ where $\circ G_3$ is a t_3 -boundaried graph (V_3, E_3, b_3) defined via

$$\begin{aligned} V_3 &:= V_1 \cup (V_2 \setminus \partial \circ G_2) \\ E_3 &:= E_1 \cup E_2[V_2 \setminus \partial \circ G_2] \\ &\quad \cup \{\phi(x)\phi(y) \mid xy \in E[\partial \circ G_2] \wedge \phi(x) \neq \perp, \phi(y) \wedge \perp\} \\ b_3(\ell) &\mapsto \begin{cases} b_1(\ell) & \text{for } \ell \in \text{dom}(b_1) \setminus \text{dom}(b_2) \\ b_2(\ell) & \text{for } \ell \in \text{dom}(b_2) \setminus \text{dom}(b_1) \end{cases} \end{aligned}$$

In essence, the graph $\circ G_1 \oplus \circ G_2$ is created by taking the disjoint union of the two boundaried graphs and identifying those vertices $v_1 \in \text{img}(b_1)$, $v_2 \in \text{img}(b_2)$ where $b_1^{-1}(v_1) = b_2^{-1}(v_2)$. Note that the operation does not commute¹ in that $\circ G_1 \subseteq \circ G_1 \oplus \circ G_2$ but $\circ G_2 \not\subseteq \circ G_1 \oplus \circ G_2$. However, we have that $\circ G_1 \oplus \circ G_2$ is isomorphic to $\circ G_2 \oplus \circ G_1$.

Our primary use for boundaried graphs will be *replacement*: We want to remove a part of a graph and put another graph in its place. To this end, we need the following definition of how to decompose a graph into boundaried pieces.

Induced boundaried graph, excision, $\circ G[\bullet], \ominus$

Definition 3 (Induced boundaried graph, excision). Let G be a graph and $W \subseteq V(G)$ a vertex set. Let b be a t -labelling of W where $t = |\partial_G W|$. The *induced boundaried graph* $\circ G[W]$ is defined as

$$\circ G[W]_b := (W, E[W], b).$$

The *excision of W from G under b* creates the t -boundaried graph

$$G \ominus_b W = (\overline{W}, E[\overline{W}], b)$$

where $\overline{W} = V - (W - \partial_G W)$.

Note that in particular $\circ G[W]_b \oplus (G \ominus_b W) = G$. For simplicity, we also allow subgraphs of G to be operands with the understanding that $G \ominus_b G' := G \ominus_b V(G')$. Given these operation, we now can define the replacement operation.

Replacement, $G[\bullet \rightarrow \bullet]$

Definition 4 (Replacement). Let G be a graph, $W \subseteq V(G)$ with $t = |\partial_G W|$. Let $\circ H$ be a t -boundaried graph and b a t -labelling. The operation of *replacing W with $\circ H$ under b* results in the graph

$$G[W \mapsto \circ H]_b = (G \ominus_b W) \oplus \circ H.$$

Subdivision

Note that $G[\overline{W}] \subseteq G[W \mapsto \circ H]_b$ by our definition of gluing.

One special replacement operation is that of a *subdivision*. Denote by $\circ P_\ell$ the path with ℓ vertices and both endpoints as its boundary (since the graph is symmetric we do not care about which endpoint receives which label). For a graph G , an edge $uv \in E(G)$ and an integer $\ell \geq 0$, saying that *uv is subdivided ℓ times* is defined as $G[uv/\ell] := G[\{uv\} \mapsto \circ P_{\ell+2}]$. The opposite operation—replacing a path by an edge—is called *dissolution*. We also write $G[E/\ell]$ to denote the graph obtained from G by subdividing every edge ℓ times.

¹ One can easily define the operation in a way such that it does commute. However, this creates a lot of unpleasant special cases in the following applications.

2.4 WELL-QUASI ORDERINGS OF GRAPHS

The following is loosely based on the very nice work by Fellows, Hermelin, and Rosamond [92]. Let \triangleleft be quasi-order on a class of graphs \mathcal{G} , *i.e.* it is a binary relation that is reflexive and transitive. For an infinite sequence $(G_i)_{i \in \mathbb{N}}$ of graphs in \mathcal{G} , a pair G_i, G_j is *increasing* if $G_i \triangleleft G_j$ and $i < j$. If every infinite sequence contains an increasing pair, then \triangleleft is a well-quasi order of \mathcal{G} . In particular, \mathcal{G} does not contain infinite anti-chains. The following table contains known wqo of graphs.

Graph class	Relation	
All graphs	\preceq_m	Minor relation [217]
All graphs	\preceq_i	Immersion relation [218]
Bounded \mathbf{tw}_1^Δ	\preceq_t	Topological minor relation [92]
Bounded treedepth	\subseteq_i	Induced subgraph relation [192]

Ideal, forbidden set

An *ideal* \mathcal{I} of a well-quasi-ordered class $(\mathcal{G}, \triangleleft)$ is a subclass of \mathcal{G} closed under \triangleleft . The interesting property of ideals is that they are characterised by a finite forbidden set: consider the set

$$\text{forb}(\mathcal{I}) := \{G \in \mathcal{G} \setminus \mathcal{I} \mid G' \not\triangleleft G \text{ for all } G' \in \mathcal{G} \setminus \mathcal{I}\}.$$

Then $\text{forb}(\mathcal{I})$ is finite since its members form an anti-chain in \mathcal{G} . Moreover, for any graph $G \notin \mathcal{I}$ it holds that there exists a member $H \in \text{forb}(\mathcal{I})$ such that $H \triangleleft G$. Hence, recognizing whether a graph is in \mathcal{I} reduces to checking whether any of the finitely many graphs of $\text{forb}(\mathcal{I})$ is *not* ‘contained’ in G . We discuss algorithmic aspects of well-quasi orders in Chapter 3 and will make use of their properties in Chapter 11.

Finite basis

For subsets that do not form ideals, we have a weaker property that nonetheless proves very useful.

Definition 5 (Finite basis). Let $(\mathcal{G}, \triangleleft)$ be a well-quasi ordered graph class and let $\mathcal{H} \subseteq \mathcal{G}$. We define the *basis* of \mathcal{H} as

$$\text{basis}(\mathcal{H}) := \{G \in \mathcal{H} \mid \forall G' \triangleleft G, G' \neq G: G' \notin \mathcal{H}\}.$$

The *basis* of a subclass has the following nice properties.

Lemma 3. *For a wqo graph class $(\mathcal{G}, \triangleleft)$ and for every $\mathcal{H} \subseteq \mathcal{G}$ it holds that*

- $\text{basis}(\mathcal{H})$ is unique and finite and
- for every $G \in \mathcal{H}$ there exists $M \in \text{basis}(\mathcal{H})$ with $M \triangleleft G$.

Proof. Since $\text{min}(\mathcal{H})$ forms an anti-chain, it must be finite by the wqo property. That it is unique follows directly from the definition, we are left to show the second property. Consider a graph $G_0 \in \mathcal{H}$. If $G_0 \in \text{basis}(\mathcal{H})$ we are done, otherwise there exists $G_1 \triangleleft G_0$ with $G_1 \in \mathcal{H}$ and $G_1 \neq G_0$. If $G_1 \in \text{basis}(\mathcal{H})$, the property holds, otherwise we repeat this procedure with G_1 . The constructed sequence

$$G_0 \triangleright G_1 \triangleright G_2 \triangleright G_3 \dots$$

must necessarily be finite, otherwise it would constitute an infinite descending chain and thus contradicting the wqo property. \square

For technical reasons we need to define a way to lift wqo relations to boundaried graphs. In essence, we disregard the boundary and use the ordering of the remainder.

Definition 6 (Boundary-preserving extension). Let \triangleleft be a graph relation and $t \in \mathbb{N}$. We extend \triangleleft to t -boundaried graphs via

$${}^\circ G_1 \triangleleft {}^\circ G_2 \iff G_1 \setminus \partial^\circ G_1 \triangleleft G_2 \setminus \partial^\circ G_2 \text{ and } G_1[\partial^\circ G_1] = G_2[\partial^\circ G_2].$$

Lemma 4. *Let \mathcal{G} be wqo by \triangleleft . Then for every $t \in \mathbb{N}$, the t -boundaried graphs \mathcal{G}_t over \mathcal{G} are wqo by the boundary-preserving extension of \triangleleft .*

Proof. Since there only finitely many graphs on t vertices, and hence only finitely many variants of boundaries, it suffices to prove the statement for a subclass \mathcal{G}'_t whose members have exactly the same boundary. Assume the contrary, *i.e.* there exists an infinite sequence $({}^\circ G_i)_{i \in \mathbb{N}}$ in \mathcal{G}'_t without an increasing pair. Derive the sequence $(H_i)_{i \in \mathbb{N}} = (G_i \setminus \partial^\circ G_i)_{i \in \mathbb{N}}$ from that sequence and note that it contains graphs of \mathcal{G} , hence it must contain an increasing pair $H_i \triangleleft H_j, i < j$. But then, by the definition of the boundary-preserving extension of \triangleleft , we have that

$$\begin{aligned} H_i \triangleleft H_j &\iff G_i \setminus \partial^\circ G_i \triangleleft G_j \setminus \partial^\circ G_j \\ &\iff {}^\circ G_i \triangleleft {}^\circ G_j \end{aligned}$$

contradicting our assumption that the first sequence did not contain an increasing pair. \square

2.5 WIDTH MEASURES

Width measures have a strong connection to both sparse graph theory and efficient algorithms. The notion of treewidth, for example, was popularised by the graph minors programme (see Chapter 3) and famously, by Courcelle's Theorem [51] and the preceding body of work, enables us to design linear-time fpt algorithm for a large collection of problems.

Definition 7 (Tree decomposition). A *tree decomposition* \mathcal{T} of a graph G is a pair (T, χ) , where T is a tree and χ is a function that assigns each tree node t a set $\chi(t) \subseteq V$ of vertices such that the following conditions hold:

1. For every vertex $u \in V$, there is a node $t \in T$ such that $u \in \chi(t)$.
2. For every edge $\{u, v\} \in E(G)$, there is a node $t \in T$ such that $u, v \in \chi(t)$.
3. For every vertex $v \in V(G)$, the set of nodes $\{t \in T \mid v \in \chi(t)\}$ induces a subtree in T .

Bags, treewidth

We call T the *underlying tree* of the decomposition. The sets $\chi(t)$ are called *bags* of the decomposition \mathcal{T} and $\chi(t)$ is the bag associated with the node t . The *width* of a tree decomposition (T, χ) is the size of a largest bag minus one. A tree decomposition of minimum width is called *optimal*. The *treewidth* of a graph G , denoted by $\mathbf{tw}(G)$, is the width of an optimal tree decomposition of G .

It is often convenient to restrict the structure of tree decompositions further. In particular, we will usually work with *rooted tree decompositions* by letting T be a rooted tree. A *nice tree decomposition* is a tree decomposition (T, χ) with the following characteristics:

Nice tree decomposition

- T is a ternary rooted tree.
- For $x \in \text{leaves}(T)$ it holds that $|\chi(x)| = 1$
- For $x \in T$ with $N^+(x) = \{y\}$, then either $|\chi(x) \setminus \chi(y)| = 1$ and we call $\chi(x)$ an *introduce bag*, or $|\chi(y) \setminus \chi(x)| = 1$ and we call it a *forget bag*.
- For $x \in T$ with $N^+(x) = \{y_1, y_2\}$ we have that $\chi(x) = \chi(y_1) = \chi(y_2)$ and we call $\chi(x)$ a *join bag*.

It is folklore that any tree decomposition can be turned into a nice tree decomposition in time linear in the size of the decomposition.

Path decomposition, nice -, pathwidth

A *path decomposition* is a tree decomposition whose underlying tree is a path. The concept of *niceness* translates directly with the additional condition that the root of the path is one of its endpoints. The *pathwidth* of a graph G , denoted by $\mathbf{pw}(G)$, is the width of an optimal path decomposition of G .

Treedepth, - decomposition, nice - decomposition

We will encounter a more restricting width measure, the *treedepth* of a graph, at many points in this thesis: the measure is intricately connected with the theory of structurally sparse graph classes.

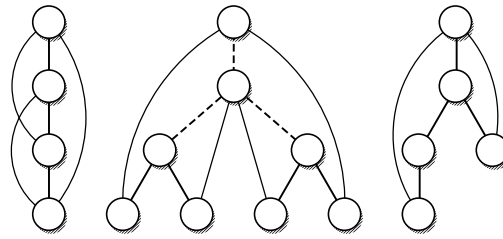
Definition 8 (Treedepth decomposition). A *treedepth decomposition* \mathcal{F} of a graph G is a pair (F, ν) , where F is a rooted forest and ν is an injective function $\nu: V(G) \rightarrow V(F)$ subject to the condition that for each $uv \in G$, either $\nu(u) \in \text{desc}_F(\nu(v))$ or $\nu(v) \in \text{desc}_F(\nu(u))$.

A treedepth decomposition (F, ν) of G is *nice* if $V(F) = V(G)$, ν is the identity function and we have that for every node $x \in F$, the graph $G[F_x]$ is connected.

Underlying tree, Trémaux -

We call F the *underlying forest/tree* of the decomposition. The *width* or *depth* of the decomposition is the height of F . A treedepth decomposition of minimum width is called *optimal*. The *treedepth* of a graph G , denoted by $\mathbf{td}(G)$, is the height of a optimal treedepth decomposition of G . We often will use a forest with the same vertex set as G as a decomposition, then ν is understood to be the identity and will not be mentioned explicitly. Rooted trees with additional edges that obey the ancestor-relationship emerge in depth-first traversal of graphs and are also called *Trémaux trees*. A treedepth-decomposition of a graph

and a Trémaux tree of bounded height differ in that the edges of the Trémaux tree are part of the graph whereas the edges of the treedepth-decomposition might not be.



Optimal treedepth decompositions for K_4 , C_8 and the house graph. Thick edges occur in the graph and the tree, thin edges only in the graph and dashed edges only in the tree.

Two important types of graphs classes that have unbounded treedepth are the class of complete graphs K_n and the class of paths P_n . A complete graph K_n has treedepth n , the underlying tree of the decomposition simply being a path of length n . For a path of length 2^ℓ the canonical (and optimal) decomposition is a binary tree of height ℓ . Since treedepth is a property closed under taking subgraphs, the presence of either a long path or a large clique in a graph will result in a high treedepth for the whole graph. The following well-known proposition will prove quite useful in the following.

Approximating treedepth

Proposition 4. *For every graph G a treedepth decomposition of width $2^{\text{td}(G)}$ can be computed in time $O(|G| + \|G\|)$.*

Proof. We compute a Trémaux tree T (or forest if G is disconnected) by a dfs traversal from an arbitrary vertex of G in the claimed running time. Since any edge of G is either a tree-, forward- or backwards-edge in T and these edges obey the ancestor relationship, the tree T is indeed a treedepth decomposition of G . To see that the height of T is bounded, recall that a path of length 2^t has treedepth t . Therefore, the graph G cannot contain a path of length $> 2^{\text{td}(G)}$ —it follows that the height of T is bounded, as claimed. \square

For the last part of this section, let us compare the three width measures presented. By definition the pathwidth of a graph is at least the treewidth of a graph since every path decomposition is a tree decomposition. The relation between pathwidth and treedepth is similar: any treedepth decomposition (F, ν) of a graph G can easily be converted into a path decomposition of similar width. To that end, consider any dfs-traversal of T and restrict the given ordering of $V(F)$ to the set $\text{leaves}(F)$. Let us denote this order of the leaves by x_1, \dots, x_ℓ . Then (P_ℓ, χ) with $\chi(i) \mapsto \text{rpath}_F(x_i)$ is a path decomposition of G and it follows by construction that

$$\mathbf{width}((P_\ell, \chi)) = \max_{x \in \text{leaves}(F)} |\text{rpath}_F(x)| - 1 = \mathbf{width}((F, \nu)) - 1.$$

In conclusion, for any graph G we have the following relation between the three introduced width measures:

$$\mathbf{tw}(G) \leq \mathbf{pw}(G) \leq \mathbf{td}(G) - 1.$$

For any width-measure \mathbf{wm} that satisfies $\mathbf{tw}(G) \leq \mathbf{wm}(G) + O(1)$ for all graphs we will say that \mathbf{wm} *upper-bounds* \mathbf{tw} .

Further width-measures

We will mention some further width-measures in passing and will not define them here. In their graph-minors programme, Robertson and Seymour defined the *branchwidth* as a measure closely related to treewidth [214]. In fact, they showed that the branchwidth \mathbf{bw} of a graph has the following linear relationship with treewidth:

$$\mathbf{bw}(G) - 1 \leq \mathbf{tw}(G) \leq \frac{3}{2} \mathbf{bw}(G) - 1.$$

Branchwidth has, however, some beneficial algorithmic properties on planar graphs. It can be computed in polynomial time [225, 131] and even approximated within a constant factor in almost-linear time [132].

The *cliquewidth* of graph \mathbf{cw} was introduced as a hypergraph-grammar by Courcelle [52] and provides a ‘dense’ variation of treewidth. It is motivated by the fact that the model checking problem for MSO_1 is decidable in polynomial time on cliques, a fact not captured by Courcelle’s theorem for treewidth. Clique-width provides a measure that in particular reflects the tractability of MSO_1 model-checking [53]. While a clique-decomposition of a graph has very nice algorithmic properties, *finding* such a decomposition seems to be very difficult. It is known that cliquewidth is NP-hard to compute [93], but it is not even known whether it admits polynomial-time algorithms for every *fixed* cliquewidth.

As a consequence, Oum and Seymour provided an approximation for cliquewidth by introducing a width-measure called *rankwidth*. The rankwidth of graph $\mathbf{rw}(G)$ is related to the cliquewidth by

$$\mathbf{rw}(G) \leq \mathbf{cw}(G) \leq 2^{\mathbf{rw}(G)+1} - 1.$$

2.6 GRAPH MEASURES AND LIMITS OVER CLASSES

The notions of *nowhere dense* and *somewhere dense* (defined later in Chapter 5) concern the asymptotic density of graph classes. We need to introduce the following conventions and notations in order to describe the results concisely. The definitions here largely follow those by Nešetřil and Ossona de Mendez [192], but are somewhat simplified.

Graph measure, supremum

A *graph measure* is any function f mapping graphs to \mathbb{R}^+ that is invariant under isomorphism, *i.e.* for $G \simeq G'$ we have that $f(G) = f(G')$. A *parametrised graph measure* is a family of graph measures $(f_r)_{r \in \mathbb{N}}$. We define the *supremum* of a graph measure f on a graph class \mathcal{G} as

$$f(\mathcal{G}) := \sup_{G \in \mathcal{G}} f(G).$$

$\limsup_{G \in \mathcal{G}} \omega(G)$

This in particular allows notation like $\omega(\mathcal{G})$ to denote the clique number of a class. To describe the asymptotic behaviour of graph measures we introduce the following limit:

$$\limsup_{G \in \mathcal{G}} f(G) := \limsup_{n \rightarrow \infty} f(\mathcal{G}_{\geq n}).$$

$\limsup_{H \triangleleft G \in \mathcal{G}}$ For a graph relation \triangleleft we will often use the convenient notation

$$\limsup_{H \triangleleft G \in \mathcal{G}} f(H) := \limsup_{G \in \mathcal{G}} \sup_{H \triangleleft G} f(H).$$

Two parametrised graph measures f, g are *polynomially related* if there exists a function α such that for every graph G and every integer r it holds that

$$f_r(G)^{-\alpha(r)} \leq g_r(G) \leq f_r(G)^{\alpha(r)}.$$

Note that this is exactly the case if $\log f_r(G) = \Theta(\log g_r(G))$.

2.7 FIXED-PARAMETER TRACTABILITY AND KERNELISATION

We roughly follow the formalism of Flum and Grohe [97]. For a finite alphabet Σ , a *language* or *decision problem* is a subset $L \subseteq \Sigma^*$. If L encodes graphs (possibly among other things) we call it a *graph problem*.

Definition 9 (Parametrisation, parametrised problem). For a finite alphabet Σ , a *parametrisation* of Σ^* is a mapping $\kappa: \Sigma^* \rightarrow \mathbb{N}$ that is computable in polynomial time. A *parametrised problem* is a pair (L, κ) where $L \subseteq \Sigma^*$ is a language.

Fpt, - algorithm A selection of important parametrised problems can be found in the appendix and in the following section.

Definition 10 (Fpt-algorithm). Let Σ be a finite alphabet and κ a parametrisation of Σ . An algorithm \mathcal{A} is an *fpt-algorithm* with respect to κ if there exists a computable function f and a polynomial p such that the running time of \mathcal{A} on every input $x \in \Sigma^*$ is at most

$$f(\kappa(x)) \cdot p(|x|).$$

*Single-exponential,
subexponential, uniformly,
non-uniform*

If p is linear, we call \mathcal{A} a *linear fpt-algorithm*. If $f(k) = 2^{O(k)}$, we call it a *single-exponential fpt-algorithm*. If $f(k) = 2^{o(n)}$, we call it a *subexponential fpt-algorithm*. There is further a slight distinction to be made of whether there exists a *single* algorithm that decides a parametrised language or whether we allow one algorithm for every k . In the former case, we call the algorithm *uniform* and in the latter case *non-uniform*.

Definition 11 (Fixed-parameter tractable, FPT). A parametrised problem (L, κ) is *fixed parameter tractable* if there exists an fpt-algorithm which decides L with respect to κ . The class of all fixed-parameter tractable problems is denoted by FPT.

Hardness with respect to fpt algorithms is provided by a hierarchy of complexity classes $W[1] \subseteq W[2] \subseteq \dots$ where $FPT \subseteq W[1]$ with the assumption that $FPT \neq W[1]$; and a suitable notion of fpt-reductions. Examples for $W[1]$ -complete problems are INDEPENDENT SET and MAXIMUM CLIQUE, the canonical example for $W[2]$ -completeness is DOMINATING SET. We therefore expect none of these problems to admit fpt-algorithms. For details, we refer to the books by Downey and Fellows [74] and Flum and Grohe [97].

Kernel

Definition 12 (Kernel). Let (L, κ) be a parametrised problem. A polynomial-time computable function $\mathbf{K}: \Sigma^* \rightarrow \Sigma^*$ is a *kernel* for (L, κ) if there exists a function h such that for all inputs $x \in \Sigma^*$ it holds that

$$x \in L \iff \mathbf{K}(x) \in L \quad \text{and} \quad |\mathbf{K}(x)| \leq h(\kappa(x)).$$

The function h is the *size* of the kernel.

We say that a kernel is *linear* if $h(k) = \Theta(k)$ and *almost linear* if $h(k) = O(k^{1+\varepsilon})$ for every $\varepsilon > 0$ or, equivalently, $h(k) = O(k^{1+o(1)})$. *Quadratic* and *cubic* kernels are defined analogously. It is folklore by now that a parametrised problem is in FPT if and only if it admits a kernel. However, of particular interest are kernels of *polynomial* size. By relating the kernelisation machinery to the concept of distillations in classical complexity theory, Bodlaender, Downey, Fellows, and Hermelin showed that a large set of problems do not admit polynomial kernels unless $\text{NP} \subset \text{coNP}/\text{poly}$ [22]. The canonical problem affected by this machinery is k -PATH. This framework has subsequently been extended by Drucker [77] to include, among other problems, TREEWIDTH and PATHWIDTH. It has been further streamlined by Bodlaender, Jansen, and Kratsch [25]. A totally different technique by Dell and van Melkebeek [242] yields, under similar complexity-theoretic assumption, that problems like VERTEX COVER and FEEDBACK VERTEX SET do not admit linear kernels².

Kernelisation lower bounds

Most kernelisation routines work by exhaustively applying a set of *reduction rules* in polynomial time that successively reduce the size of the input. The meta-kernelisation framework presented and extended in Chapter 11 is no exception; only here the reduction rule is replaced by techniques developed in the context of reduction algorithms for graphs of bounded treewidth.

2.8 DOMINATION PROBLEMS

We will encounter domination problems at several points in this thesis. The classical DOMINATING SET problem is a good litmus test for how well structural sparseness can improve tractability, at least in the parametrised setting—since DOMINATING SET is NP-hard even on planar graphs of maximum degree three ([GT2] in Garey and Johnson [120]), we do not see improvement from the classical perspective. Furthermore, this problem is log-APX-complete, hence it cannot be approximated to within a factor better than $O(\log n)$.

In the parametrised setting, DOMINATING SET is one of the canonical $W[2]$ -complete problems under its natural parametrisation. We therefore assume that it does not admit an fpt-algorithm and in particular no kernelisation. We will outline in Chapter 13 how in the sparse settings even polynomial kernelisation is possible and add new results in even larger classes of sparse graphs.

Dominating set, r - $\text{}$, relative domination, \mathbf{ds}

² This result applies to the *encoding* size. Kernels with a linear number of *vertices* are known for VERTEX COVER.

A *dominating set* of a graph G is a vertex subset $D \subseteq V(G)$ such that $N[D] = V(G)$, i.e. every vertex is either in D or has a neighbour in D . The *domination number* $\mathbf{ds}(G)$ is the size of a minimum domination set of G . Given a vertex set Z , a set D is a Z -*dominator* in G if $N[D] \supset Z$. The *relative domination number* $\mathbf{ds}(G, Z)$ for a set Z is the size of a minimal Z -dominator. For an integer r , an r -*dominating set* is a vertex subset $D \subseteq V(G)$ such that $N^r[D] = V(G)$.

DOMINATING SET parametrised by k

Input: A graph G and an integer k .

Problem: Does G have a dominating set of size at most k ?

A natural extension of domination is to let vertices dominate not only their neighbourhood but all vertices within some distance r . The following problem is also known as (k, r) -CENTRE if we parametrise by both r and k .

r -DOMINATING SET parametrised by k

Input: A graph G and an integer k .

Problem: Does G have an r -dominating set, i.e. a set $D \subseteq V(G)$ such that $V(G) = N^r[D]$, of size at most k ?

Instead of minimising the number of vertices to dominate the whole graph, we can also instead ask how many vertices we are able to dominate given a fixed budget:

PARTIAL DOMINATING SET parametrised by t

Input: A graph G and integers t, k .

Problem: Does G have a set of size at most k that dominates at least t vertices?

2.9 PROPOSITIONAL LOGIC OVER GRAPHS

Several results in this thesis rely on results from model theory, therefore we provide the basic notation and vocabulary here. For a complete introduction into model theory we refer to the book by Ebbinghaus and Flum [82].

Propositional logic

In propositional logic we can join atomic propositions, whose individual truth-value is later given by an *assignment*, recursively by the usual boolean operations (conjunction, disjunction and negation) into *propositional formulas*. An assignment that is evaluated *true* by propositional formula ϕ is a *model* of ϕ .

First-order logic

By adding variables and the quantifiers \exists, \forall to propositional logic we arrive at *first-order logic*. We consider only formulas over graphs here,

seen as logical structures. That is: a graph G is a structure with universe $V(G)$ and a single binary relation $E(G)$ (we also allow a finite number of unary relations in order to have labelled vertices).

The syntax of a first-order formula in the language of graphs consists of the atoms $\text{adj}(x, y)$ and $x = y$. A first-order formula is an atom or recursively constructed from formulas joined through boolean operators or by addition of a quantification over a variable that occurs in it. A variable that is not bound by a quantifier is a *free variable*.

First-order definable, model checking

We write $G \models \phi$ to express that the formula ϕ holds in G . A graph property Π that can be expressed in this fashion, *i.e.* there exists a formula ϕ such that

$$G \in \Pi \iff G \models \phi$$

is called *first-order definable*. The *model checking* problem is, given a graph G and a first-order formula ϕ as input, to decide whether $G \models \phi$. This problem is, in general, PSPACE-complete since it already contains QBF SATISFIABILITY as a degenerate sub-case (and testing a first-order formula with polynomial space is straightforward by brute-force). In Chapter 3 we will see that structural sparseness considerably improves the tractability of the model checking problem, up to even the most general notion of structural sparseness.

Monadic second-order logic

By augmenting first-order formulas with quantification over sets of vertices and sets of edges we arrive at the *monadic second-order logic* (MSO) of graphs. There exist two variations: the logic MSO_1 models graphs as before as structures with universe $V(G)$, hence set quantification is only admissible for vertex-sets. The logic MSO_2 (sometimes *guarded* MSO or GMSO) models graphs as structures whose universe is the union of vertices and edges, but introduces two types of quantifications; either over edge-sets or over vertex-sets. Quantification over ‘mixed’ sets are not possible.

Courcelle’s Theorem classically states that model checking of MSO_2 is possible in linear time on graphs of bounded treewidth, or, in the language of parametrised complexity, there is a linear fpt algorithm for MSO_2 parametrised by the treewidth of the input graph³. Kreutzer and Tazari showed that this result is tight: for graph classes with even very moderately growing treewidth, the MSO_2 model-checking problem is not even possible in polynomial time if the size of the formula is considered a constant [165].

Extended MSO

In order to model decision or optimisation problems, it is often necessary that we minimise or maximise a certain set. Take the problem of deciding whether a vertex cover of size k exists, or the problem of finding a minimum dominating set. These problems are not expressible in MSO since we are lacking any means of talking about set sizes (except constant-sized ones). However, MSO is clearly powerful enough to describe how a feasible solution to these problems looks

³ Technically, this formulation also needs the classical result by Bodlaender that a tree-decomposition can be computed in linear fpt-time.

like. Furthermore, we know that these problems admit hand-crafted dynamic programming routines to solve them—something is still missing in Courcelle’s Theorem. Arnborg, Lagergren, and Seese [13] extended Courcelle’s formalism by introducing *extended* MSO or EMSO: besides the formula, we also provide an *evaluation* which maps free variables of the formula into the domain of integers or reals. In particular, evaluations can be optimised against a target functions. A simple variant of this formalism are formulas of the form $\min X \phi(X)$, where ϕ is an MSO₂ formula with free variable X . If ϕ expresses that X is a vertex cover in the input graph, then by the result of Arnborg *et al.* we can determine the size of a minimal vertex cover in linear time on graphs of bounded treewidth.

2.10 GRAPH SURGERY TOOLKIT

In Chapter 7 we will work on embeddings of shallow minors, trying to ‘salvage’ a large portion of the minors’ edges in order to recover a dense-enough subgraph. I suspect the next lemma is known in some form, but I was unable to find it in the literature.

Lemma 5 (High-density transversal). *Let G be a loopless multi-graph on $2n$ vertices partitioned into n sets V_1, V_2, \dots, V_n each containing two independent vertices. There exists a set X of size n that intersects every V_i in exactly one vertex such that $\|G[X]\| \geq \|G\|/4$.*

Proof. Choose a vertex-colouring into colours black B and white W that split every V_i , $1 \leq i \leq n$. Denote by $w_i := V_i \cap W$ the white vertex and by $b_i := V_i \cap B$ the black vertex of V_i .

If $\|G[B]\| \geq \|G\|/4$ or $\|G[W]\| \geq \|G\|/4$ we are done. Otherwise, we have that $|E(W, B)| > \|G\| - \|G\|/2$.

Claim. *There exists a partite set V_i such that*

$$\|G[W \setminus \{w_i\} \cup \{b_i\}]\| > \|G[W]\|.$$

For a vertex $v \in G$, let $\deg_W(v) := |E(\{v\}, W)|$ denote the number of edges v has to white vertices (the count includes potential double edges). Consider the potential ϕ defined for the partite sets as

$$\phi(V_i) := \deg_W(b_i) - \deg_W(w_i).$$

Note that any partite set with positive potential satisfies the claim: exchanging w_i for b_i will increase the number of monochromatic white edges. Summing up the potential over all partite sets, we see that

$$\begin{aligned} \sum_{i=1}^n \phi(V_i) &= \sum_{i=1}^n \deg_W(b_i) - \sum_{i=1}^n \deg_W(w_i) \\ &= |E(W, B)| - 2\|G[W]\| \\ &> \|G\| - \|G\|/2 - \|G\|/2 = 0 \end{aligned}$$

Thus the total sum of the potential is strictly positive, hence there exists at least one V_i with $\phi(V_i) > 0$. By the above observation, this proves the claim.

The lemma now follows by the following iterative procedure: if either sets B, W already induce a subgraph with at least $\|G\|/c$ edges we return the respective set. Otherwise, by the above claim, we can find a partite set $V_i = (b_i, w_i)$ such that the set $W' := W \setminus \{w_i\} \cup \{b_i\}$ contains more edges than W previously. We iterate the procedure with the colouring $W', B' := V \setminus W'$. Since the number of edges in the white set steadily increases, this procedure must terminate after a finite number of steps by finding a suitable set X . \square

A direct consequence, and one of the applications in the following, is that we can turn a directed graph into a bipartite graph in which every vertex is either a source or a sink.

Corollary 1. *Let \vec{G} be a digraph. Then there exists a subgraph \vec{G}' of \vec{G} in which every node is either a source or a sink and*

$$\|\vec{G}'\| \geq \|\vec{G}\|/4.$$

Proof. Construct an auxiliary-graph H by replacing every vertex $v \in \vec{G}$ with two vertices v^-, v^+ . For every arc $uv \in \vec{G}$ we add the arc u^+v^- to H . Applying Lemma 5 with the partition $\{\{v^-, v^+\}\}_{v \in V(G)}$ to H then tells us which vertices of G can be turned into sources and which into sinks while still retaining at least a quarter of the arcs. \square

And of course we can lift the lemma to larger partitions.

Corollary 2. *Let c be an integer and G be a graph on cn vertices partitioned into n sets V_1, V_2, \dots, V_n each containing $\leq c$ independent vertices. There exists a set X of size n that intersects every V_i in exactly one vertex such that $\|G[X]\| \geq \|G\|/4c^2$.*

Proof. By a simple padding-argument we can assume that all partite sets have exactly size c . Now the case for $c = 2$ is handled by Lemma 5 and the case $c = 1$ is trivial. For $c > 2$, we partition each set V_i into sets B_i, W_i such that $|B_i| = |W_i|$ if $|V_i|$ is even and $|B_i| = |W_i| + 1$ otherwise. Create the auxiliary multi-graph H defined via

$$\begin{aligned} V(H) &:= \{B_1, B_2, \dots, B_n, W_1, W_2, \dots, W_n\} \\ |E_H(X, Y)| &= |E_G(X, Y)| \text{ for } X, Y \in V(H). \end{aligned}$$

Now we apply Lemma 5 and obtain a set X such that $\|H[X']\| \geq \|H\|/4$ and for $1 \leq i \leq n$, the set X exactly one of $\{W_i, B_i\}$. We repeat the argument for $G[\cup X]$ using the partition X (if c is odd we need to add a single padding vertex to every black partite set in X in order to ensure a size of $\lceil c/2 \rceil$). We iterate this procedure until $c \leq 1$; the number of iterations necessary is $\lfloor \log c \rfloor + 1$. Since every application of Lemma 5 yields a quarter of the edges of the input multi-graph, we end up with a total fraction of $4^{\lfloor \log c \rfloor + 1} \leq 4c^2$ edges, as claimed. \square

Spiders and trees

The following innocuous looking (and easy to prove) lemma lies at the heart of the very important equivalence of minors and topological minors in case of shallow embeddings. We will see this application later in Chapter 4.

Lemma 6 (Large spiders in shallow trees). *Let T be a rooted tree of height at least two. Then there exists a spider $S \subseteq T$ of height at least two such that*

$$(\text{leaves}(S) \cap \text{leaves}(T)) \geq \text{leaves}(T)^{1/(h-1)}.$$

Proof. We prove the statement by induction over the height h . It is trivial for $h = 2$, since then T is already a spider. Assume that the statement holds for $h - 1$ and that T has height h . Let $r = \text{root}(T)$ be the root of T . If $\deg(r) \geq \text{leaves}(T)^{1/(h-1)}$ we are done: we turn T into a spider by choosing one leaf per subtree $T_x, x \in N^+(r)$ and removing everything but the paths from the root to these leaves.

Otherwise, we have that $\deg(r) < \text{leaves}(T)^{1/(h-1)}$. Therefore there exists a subtree T_x for some child $x \in N^+(r)$ such that

$$\text{leaves}(T_x) \geq \text{leaves}(T) / \text{leaves}(T)^{1/(h-1)} = \text{leaves}(T)^{\frac{h-2}{h-1}}.$$

Applying the induction hypothesis, we obtain a spider $S \subseteq T_x$ with

$$\min_{S \in \mathcal{K}(F)} (\text{leaves}(S) \cap \text{leaves}(T_x)) \geq \text{leaves}(T_x)^{1/(h-2)} \geq \text{leaves}(T)^{1/(h-1)},$$

as claimed. Since the base case has height two, the resulting spider has height at least two. \square

I like sparseness. There's something about that minimalist feel that can make something have an immediate impact and make it unique. I'll probably always work with that formula; I just don't know how.
— Britt Daniel, singer of Spoon.

Sparseness is a phenomenon seen in all types of applications: from the discretisation of partial differential equations to electrical networks, sparse structures arise wherever we look. A great deal of examples will be mentioned in Chapter 14 where we consider *complex networks* which are simply (labelled) graphs obtained from real-world objects. In this chapter we introduce common notions of *structural sparseness*, beginning with the most basic concept of *degeneracy* and working up towards the full sparse-class dichotomy by Nešetřil and Ossona de Mendez. On the way we encounter graphs embeddable in surfaces and classes defined by forbidden substructures, alongside important algorithmic results, most of which take the form of *algorithmic meta-theorems*, that is, they concern a large body of problems that can be formulated in a certain logic or simply share certain fundamental characteristics. The interplay of structural graph theory and the development of efficient algorithms—mostly in the framework of approximation and parametrised algorithms—is the fundamental message: wherever we obtain strong decomposition theorems for sparse classes, great algorithmic opportunities arise.

The corollary and hypothesis is: structural sparseness is the key to efficient algorithms for real-world data. Even if efficient algorithms are available in general classes, the additional assumption of sparseness can only improve such algorithms—be it in a reduction of running time, space complexity or solution quality. We will see in the fourth part of this thesis (Chapter 15 in particular) that the more well-studied notions of structural sparseness are too strong to be applied to complex networks. Only the most general concepts, *bounded expansion* and *nowhere dense* classes, will be adequate to describe them.

The remainder of this chapter is roughly organised around the hierarchy of sparse classes depicted in Figure 1. We already encountered low-treewidth graphs in Section 2.5 and some of their algorithmic properties in Section 2.9. Classes of bounded degree will primarily be contained in discussions about classes excluding a topological minor. Before ascending the hierarchy, however, we need to discuss the basic notion of *degeneracy* that plays a fundamental role in many results on structural sparseness.

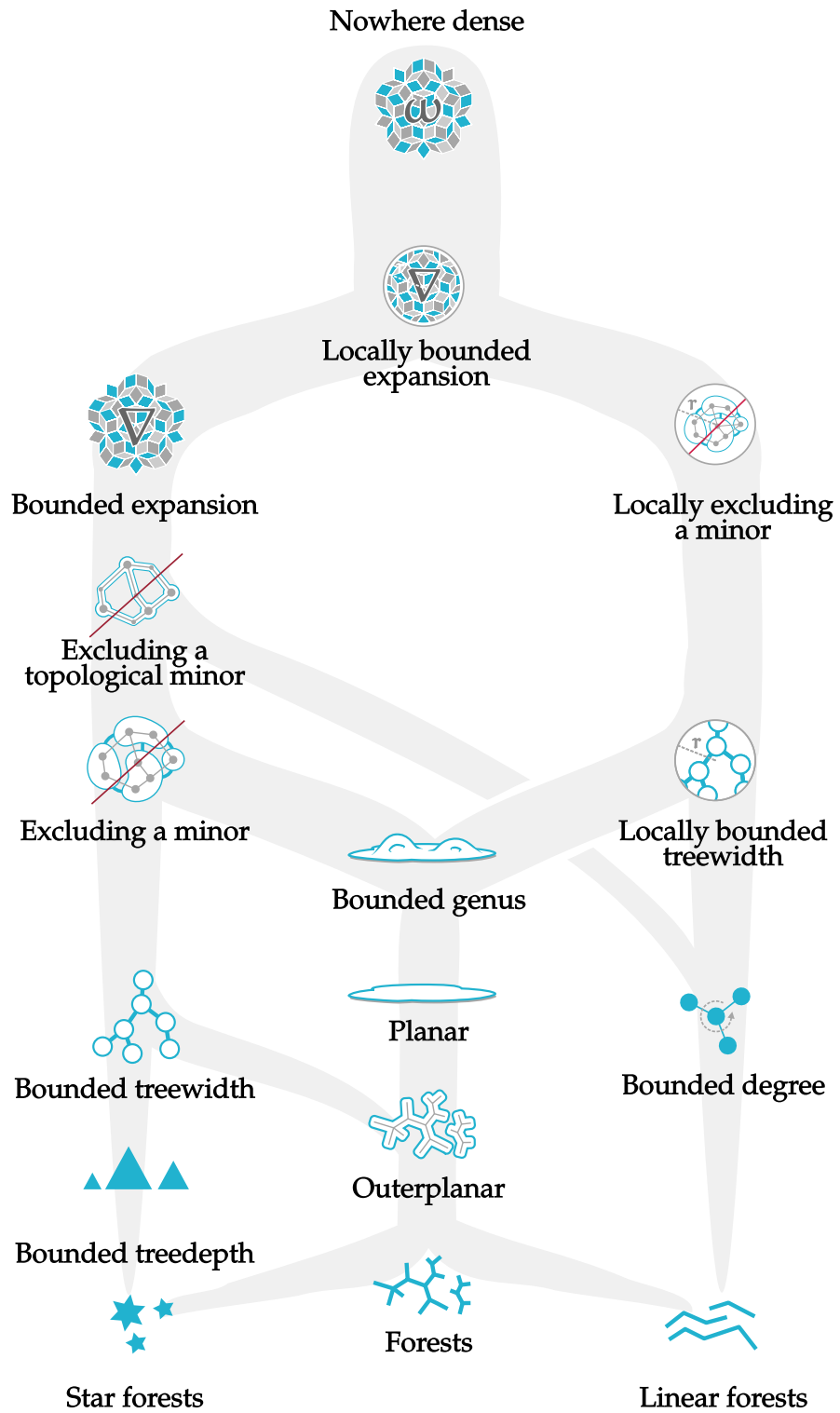


Figure 1: Hierarchy of sparse graph classes. Adapted from a figure in [80].

3.1 DEGENERACY

If a unified definition structural sparseness is the goal of our endeavour, then degeneracy is the low bar with which we have to measure: no sensible notion of sparseness does *not* imply degeneracy¹. And while this notion is still far away from a complete theory of sparse classes, it lays at the heart of several important generalisation which will lead us there. Informally, a graph is degenerate if its density is bounded *and* the densities of all its subgraphs are bounded. In that reading, degeneracy is just hereditary sparseness.

d-degenerate

Proposition 1. *Given a graph G and a natural number d , the following statements are equivalent:*

1. *Every subgraph of G has a vertex of degree at most d ,*
2. *G has colouring number $\text{col}(G) = d + 1$,*
3. *G has an acyclic orientation \vec{G} such that $\Delta^-(\vec{G}) = d$, and*
4. *G has a linear ordering π such that $|N_\pi^<(v)| \leq d$ for every vertex $v \in V(G)$.*

The minimum number d for which any of these statements hold true for a graph is called its *degeneracy*. For such graphs, an acyclic orientation of maximum indegree d or an ordering π with small ‘left-neighbourhoods’ can be computed efficiently: we iteratively remove the vertex of smallest degree to obtain the ordering, the orientation can be obtained by turning all edges into arcs pointing towards the larger vertices in the ordering. With a smart data-structure to track the degrees of the remaining vertices, this is possible in linear time [181].

Graph classes of bounded degeneracy already provide one important feature: every member of such a class only contains degenerate subgraphs. However, they can still host dense substructures other than subgraphs. Take, for example, the class $\{K_\ell[E/2]\}_{\ell \in \mathbb{N}}$ consisting of all one-subdivisions of complete graphs. All members of this class are two-degenerate, but hardly sparse in a structural sense.

However, since most of the graph classes we are about to consider are degenerate, we will spend some time on the basic properties of degenerate graphs. A widely applied fact is that a d -degenerate graph contains only a linear number of complete subgraphs. The following result and proofs can be considered folklore.

Proposition 5. *Let G be d -degenerate. Then $\#\omega(G) \leq 2^d \cdot (|G| - d + 1)$.*

Proof. Let π be a linear ordering such that $|N_\pi^<(v)| \leq d$ for every vertex $v \in V(G)$. Let $n = |G|$ be the number of vertices. We count the number of complete subgraphs by charging them to their respective last vertex in the ordering π .

¹ We will have to revise that statement slightly once we reach nowhere dense-classes and say instead: no sensible notion of sparseness does *not* imply $O(n^{o(1)})$ -degeneracy. This is, however, less catchy.

The first d vertices of the ordering π can induce at most 2^d complete subgraphs. For every vertex v with $\pi(v) > d$, we observe that the number of complete subgraphs whose last vertex in the ordering π is v are bounded by $2^{|N_\pi^<(v)|} \leq 2^d$. In total we count at most

$$2^d + 2^d(n - d) = 2^d \cdot (n - n + 1)$$

complete subgraphs in G . \square

Let us quickly prove a similar bound in terms of the degeneracy and the clique number. While it looks inferior to the above bound, it will be useful for graph classes where the clique number is much smaller than the degeneracy.

Lemma 7. *Let G be d -degenerate. Then $\#\omega(G) \leq \omega(G)(ed)^{\omega(G)}|G|$.*

Proof. Let again π be a linear ordering such that $|N_\pi^<(v)| \leq d$ for every vertex $v \in V(G)$. Let $\omega := \omega(G)$ in the following. We use a similar charging argument as before: a vertex $v \in V(G)$ is the last vertex in the ordering of at most

$$\sum_{k=1}^{\omega-1} \binom{d}{k} \leq \sum_{k=1}^{\omega-1} \frac{d^k e^k}{k^k} \leq (\omega - 1)(ed)^{\omega-1}$$

complete subgraphs. The claimed bound follows. \square

A further useful property of degenerate graphs is that their composition is again degenerate.

Lemma 8. *Let G_1, G_2 be d_1 - and d_2 -degenerate graphs on the same set of vertices V . Then the graph $G = (V, E(G_1) \cup E(G_2))$ is $2(d_1 + d_2)$ -degenerate.*

Proof. Since the number of edges in G is at most $(d_1 + d_2)|V|$ there must exist a vertex v of degree at most $2(d_1 + d_2)$. We can now inductively apply the argument to $G \setminus \{v\}$ which obviously is again a graph obtained by combining a d_1 - and a d_2 -degenerate graph. We conclude that G is $2(d_1 + d_2)$ -degenerate. \square

The above can easily be extended to more than two graphs, in that case the resulting degeneracy is twice the sum of the degeneracies of the input graphs.

We saw above that degenerate graphs can be defined via acyclic orientations of low in-degree and that this characterisation easily proves that graphs with low degeneracy have a low colouring number. The following relaxation to—not necessarily acyclic—orientations of low indegree will prove useful. It is considered folklore and indeed it seems impossible to track down the first time it was proved.

Lemma 9. *The underlying graph of a digraph \vec{G} is $2\Delta^-(\vec{G})$ -degenerate.*

Proof. Since G contains at most $\Delta^-(\vec{G})|G|$ edges, it contains a vertex v of degree at most $2\Delta^-(\vec{G})$. Removing this vertex from \vec{G} cannot increase the maximum indegree, therefore we can apply the same argument to the graph $\vec{G} \setminus \{v\}$ and hence prove inductively that G is in fact $2\Delta^-(\vec{G})$ -degenerate. \square

The following basic observation will crop up in many of the following proofs, hence we put it down here as a convenient shorthand.

Lemma 10. *Let $G = (X, Y, E)$ be a d -degenerate bipartite graph. Then the number of vertices in Y with degree at least d is at most $d|X|$.*

Proof. Let $Y' \subseteq Y$ be the set of vertices with degree at least d . Let further $X' = N(Y')$ and $G' = G[X' \cup Y']$. Since every vertex in Y' has degree larger than d , there exists a vertex of degree at most d in X' . We count the number of vertices in Y' by choosing such a vertex $x \in X'$ with $\deg_{G'} \leq d$, removing it from X' and removing $N_{G'}(x)$ from Y' . Notice that the degree of every vertex in $Y' \setminus N_{G'}(x)$ still has degree larger than d into the set $X' \setminus \{x\}$. Hence repeating this operation until X' is empty gives us that $|Y'| \leq d|X|$, as claimed. \square

k-core

A related notion to degeneracy is that of *cores*. A k -core of a graph G is the maximal subgraph of G with minimal degree at least k . It can be computed in linear time by repeatedly removing the vertex of smallest degree until that vertex has degree at least k or the graph is empty (the linear running time is achieved by keeping track of the vertices' degrees in an array and updating it during deletion). We will need the following folklore result about cores later.

Lemma 11. *Let G be a graph. Then G contains a $\|G\|/2|G|$ -core with at least $\sqrt{\|G\|}$ vertices and at least $\|G\|/2$ edges.*

Proof. We obtain the core by iteratively removing vertices of degree less than $\|G\|/2|G|$. Denote the number of vertices deleted in this fashion by p . The number of edges we have removed is then in total at most $p\|G\|/2|G| < \|G\|/2$. It follows that the core still contains at least $\|G\|/2$ edges, and accordingly at least $\sqrt{\|G\|}$ vertices. \square

We should consider degeneracy as the 'low bar' of structural sparseness: any sparse class *should* be degenerate. And sometimes degeneracy already improves the tractability of problems. Shai and Gutner demonstrated that DOMINATING SET can be solved in linear fpt-time [6] on degenerate graphs. Philip, Raman, and Sikdar showed that DOMINATING SET also admits a kernel of size $O(k^{(d+1)^2})$ in d -degenerate graphs [204]. On the other hand, Cygan, Marcin and Michał Pilipczuk, and Wojtaszczyk, showed that CONNECTED DOMINATING SET (among other problems with connectivity constraint) does not admit a polynomial kernel in degenerate graphs [55] unless NP is contained in coNP/poly. Furthermore, Golovach, and Villanger [126] earlier proved

that PARTIAL DOMINATING SET and (k, r) -CENTRE are $W[1]$ - and $W[2]$ -hard, respectively. In order to obtain broad algorithmic results, we need to go beyond this ‘low bar’ of sparseness.

3.2 EMBEDDABLE GRAPHS



Graphs embedded in the plane have been a topic since the very infancy of graph theory—it might have been different had the bridges of Königsberg crossed each other. There is something fundamentally alluring to plane graphs, maybe fed by our innate propensity towards order. It is not surprising, then, that a rich body of work on the algorithmic properties of planar graphs exist. Let us begin with the famous separator theorem for planar graphs by Lipton and Tarjan.

Theorem 1 (Lipton and Tarjan [174]). *Every planar graph with n vertices has a $\frac{2}{3}$ -separator of size $O(\sqrt{n})$.*

Recall that a $\frac{2}{3}$ -separator divides a graph into two pieces, none of which are larger than $2n/3$. Lipton and Tarjan later demonstrated that the separator theorem can be used to design a PTAS for INDEPENDENT SET [175] and VERTEX COVER [43]. Baker applied a different technique [15] that can also be used to design PTAS for DOMINATING SET, a problem that seems to be impervious to the separator approach. Among other problems, she also showed that VERTEX COVER, INDEPENDENT SET, and SUBGRAPH ISOMORPHISM admit polynomial time approximation schemes in planar graphs.

Baker’s technique

Her technique works as follows: we divide a planar graph into layers by running a bfs starting at some arbitrary vertex. Depending on the approximation quality, we choose a number r and assign each layer a number $(d \bmod r)$, where d is the layer’s depth in the bfs tree. By ‘sacrificing’ a group of layers with the same number i (in the case of INDEPENDENT SET, for example, we would decide that our solution should not contain vertices from these layers), we can decompose the graph into disjoint parts each of treewidth at most $O(r)$ and solve the problem in these independent subinstance optimally. For the ‘correct’ choice of the number i (for which we can simply try all values) this yields a solution whose quality is $1 + O(1/r)$.



Eppstein generalised Baker’s approach beyond planar graphs [84] by showing that the same trick—with some modifications—is possible in

every class of graphs that has the *diameter-treewidth property*, or equivalently *locally bounded treewidth*.

Locally bounded treewidth

A class \mathcal{G} has *locally bounded treewidth* if there exists a function f such that for every $r \in \mathbb{N}$ and every graph $G \in \mathcal{G}$ and every vertex $v \in G$ it holds that

$$\text{tw}(G[N^r[v]]) \leq f(r).$$

In particular, planar graphs and bounded-genus graphs have locally bounded treewidth and therefore admit PTAS for many problems.

Another major algorithmic breakthrough that in particular applies to classes with locally bounded treewidth was achieved by Frick and Grohe. Starting from a result by Seese [222] for classes of bounded degree, they proved the following:

FO model-checking

Theorem 2 (Frick and Grohe [112]). *Let \mathcal{G} be a graph class with locally bounded treewidth and ϕ a first-order definable property. There exists a linear-time algorithm that decides whether a given graph $G \in \mathcal{G}$ has property ϕ .*

We note that the result by Frick and Grohe is actually more general but it in particular applies to graphs with locally bounded treewidth. The property they and before them Seese exploit is that first-order definable properties are inherently *local* by Gaifman's Theorem [113]. Avoiding a detailed excursion, it suffices to say that every first-order definable property ϕ holds in a graph G if and only if we can find a $2r$ -scattered set S whose members all satisfy a *local* first-order formula ϕ^r , which is only evaluated around the r -neighbourhood of each vertex $v \in S$. Note that the quantities r and $|S|$ only depend on ϕ . Roughly speaking, testing the property ϕ then can be done by evaluating ϕ^r for every vertex and then finding a suitable $2r$ -scattered set. In the context of bounded-degree classes, the local formulas can be simply evaluated by brute force. In classes with locally bounded treewidth one can instead make use of Courcelle's theorem.

The implication for parametrised complexity on sparse graphs cannot be overstated: problems like VERTEX COVER, DOMINATING SET, r -DOMINATING SET, SUBGRAPH ISOMORPHISM and in particular k -PATH, and EDGE DOMINATING SET are all first-order definable (by a family of formulas, one for each parameter k) and can therefore be solved in *linear* fpt-time. As is often the case with such meta-theorems, the dependence on the parameter is a tower of exponentials.

We will postpone the discussion of another big result in parametrised complexity to Chapter 11: the meta-kernelisation theorem [23] that showed for large range of problems that they admit linear kernels in planar and bounded genus graphs.

Further results

Beside these algorithmic meta-theorems, there exists a large number of results for parametrised problems in the context of embeddable graphs. Alber, Fernau, and Niedermeier demonstrated that planarity can be utilised to obtain subexponential fpt-algorithm for VERTEX COVER, INDEPENDENT SET and DOMINATING SET [4]. Marcin and

Michał Pilipczuk, Sankowksi and Erik Jan van Leeuwen show that STEINER TREE and STEINER FOREST parametrised by the size of the tree/forest and MULTIWAY CUT admit polynomial kernels when the input is a planar graph [206]. The same authors also demonstrated a subexponential time fpt-algorithm for STEINER TREE [205]. However, many more results are applicable to larger classes of graphs which we will discuss in the next section. We should remember, however, that many insights were first had in the context of graphs that have the pleasing property of being drawable without crossing edges.

3.3 EXCLUDED (TOPOLOGICAL) MINORS

Defining a structure by stating what should *not* occur in it often provides a succinct description of a rich and complicated object. Take graphs excluding the claw (the graph $K_{1,3}$) as an example: this seemingly minor restriction on what subgraphs can occur give rise to a deep and complicated theory with major algorithmic implications (see, e.g., the survey by Faudree, Flandrin, and Ryjáček [90] and the structural characterisation by Chudnovsky and Seymour [45]).

In the context of sparse graphs, Kuratowski's theorem (as formulated by Wagner) is certainly the most iconic instance of such a characterisation: a graph is planar if and only if it does not contain a minor isomorphic to either K_5 or $K_{3,3}$. Such characterisation by forbidden structures are extremely useful, since every graph without the property in question contains a certificate that proves it.

The minor WQO

By the Robertson-Seymour theorem (formerly Wagner's conjecture), finite graphs are well-quasi-ordered under the minor relation. This in particular implies that graph properties that are minor-closed are *always* characterised by a finite obstruction set: if a graph contains one of the obstructions as a minor, it does not have the property. Now if degeneracy is simply hereditary sparseness, why not consider minor-closed sparseness? It certainly nullifies of our above counterexample: a subdivided K_n is not sparse in a minor-closed sense, since we can simply undo the subdivision by taking a minor. Indeed, there is a rich body of work on planar, apex-minor-free², minor-free and topological-minor-free graphs.

Before we go into the algorithmic consequences of the Robertson-Seymour result and its extensions, let us consider some more digestible properties of graph classes that exclude a fixed graphs as a minor. The first important two results shows that graph classes excluding some fixed graphs as a minor are degenerate.

Theorem 3 (Mader [177]). *If G has average degree at least 2^{t-2} , then G contains K_t as a minor.*

The degeneracy of classes excluding a minor then follows from the simple observation that such classes are hereditary. For excluded to-

² An apex graph is a planar graph with a universal vertex added to it.

polynomial minors, the same was later shown by Lomlós and Szemerédi and improved by Bollobás and Thomason.

Theorem 4 (Bollobás & Thomason [32], Komlós & Szemerédi [161]).

There is a constant $\rho \leq 10$ such that, for $r > 2$, every graph with average degree at least ρr^2 contains K_r as a topological minor.

Again, graph classes defined by an excluded topological minor are hereditary, so it follows that they are in particular degenerate and contain at most $2^{\rho r^2} \cdot n$ complete subgraphs.

The grid theorem

The first structural property of graphs excluding a minor is one of the many ‘by-products’ of the work by Robertson and Seymour that had far-reaching consequences in both graph theory and theoretical computer science. It is known as the *grid-minor theorem*:

Theorem 5 (Grid-minor theorem [213]). *There exists a function f such that every graph with treewidth larger than $f(k)$ contains a $k \times k$ -grid as a minor.*

In the original proof the function f was non-elementary, a different proof by Robertson, Seymour and Thomas [219] lowered it to $2^{O(k^5)}$. They conjecture that the bound should actually be polynomial, and after several improvements [152, 170], finally Chekuri and Chuzhoy claimed a polynomial bound [42].

One of the direct implications of this theorem is that graphs excluding a fixed planar graph have bounded treewidth: this follows from the simple fact that every planar graph is contained as a minor in a large enough planar grid. The grid minor theorem can be strengthened if we restrict ourselves to sparse classes. Robertson, Seymour, and Thomas observed that in planar graphs, one can use $f(k) = 6k - 5$. Demaine and Hajiaghayi lifted this result to graphs excluding a fixed graph H as a minor: the corresponding function f is linear and depends only on the excluded graph H [63].

Bidimensionality

The (sparse) grid minor theorem has led to a fruitful line of research called *bidimensionality*. The key insight is that many problems have necessarily large solutions in grids; and that this still holds true if the grid is only present as a minor. Take the FEEDBACK VERTEX SET problem: a $4k \times 4k$ grid contains k disjoint cycles of length 4, therefore we need to remove at least k vertices to turn it into a forest. This still holds true if the grid is actually a minor embedded in a bigger graph, therefore we have the following Win/Win-scenario: either the treewidth of our input graph is bounded, or it contains a large grid and we can conclude that our solution will be too large. This framework only makes sense in the parametrised setting and is particularly powerful in graph classes with a linear grid theorem: then the bound on the treewidth is actually $O(\sqrt{k})$ and one obtains fpt-algorithms whose dependence on k is a subexponential function $2^{O(\sqrt{k})}$. Note that depending on whether we work in planar graphs, graphs of bounded genus, graphs excluding a minor or an apex-minor, the notion of a ‘grid’-graph needs to be adapted (for example, in a plane graphs we can

obtain a triangulated grid that respects the embedding of the graph). For bounded-genus graphs, Demaine, Fomin, Hajiaghayi, and Thilikos show the existence of subexponential parametrised algorithms for VERTEX COVER, FEEDBACK VERTEX SET, DOMINATING SET, EDGE DOMINATING SET, and more [61]. For VERTEX COVER, DOMINATING SET, and SET COVER they even prove such algorithms exist for graphs excluding a fixed minor. Fomin, Lokshtanov, Raman, and Saurabh used bidimensionality to obtain EPTAS for r -DOMINATING SET, CONNECTED DOMINATING SET and CONNECTED VERTEX COVER on classes excluding an apex-minor and MAXIMUM LEAF SPANNING TREE, VERTEX \mathcal{H} -PACKING (among many other problems) on classes excluding a minor [100]. For more results and details on the techniques we refer to the surveys by Demaine and Hajiaghayi [62] and Dorn, Fomin, and Thilikos [70]. We will shortly revisit bidimensionality in the context of kernelisation algorithms later in Chapter 11.

Width measure

Before we move on, we should note that while treewidth was discovered earlier by Halin [136], the graph minor programme popularised it. Not only has this width measure inspired a whole range of algorithmic results (a 1985 survey by Arnborg [12] lists dozens of results and applications) culminating in what is now known as Courcelle’s Theorem [51]—namely that all MSO-expressible problems can be solved in linear fpt-time on graphs of bounded treewidth. It has also inspired a whole range of graph width measures, including clique-width, rankwidth [201], shrubdepth [118], booleanwidth [39] modularwidth [117] matchingwidth [243] treedepth, and treelength [71], all with their own, rich body of work.

Robertson-Seymour decomposition

Finally, we come to the decomposition theorem by Robertson and Seymour. In essence, it states that graphs excluding a fixed minor have a tree decomposition whose bags are either of constant size *or* have a relatively tractable structure related to embeddings in a surface of low genus. We need a few technical definitions to state it, though we will gloss over several details for the sake of brevity. A very approachable write-up can be found in Diestel’s book [67].

Given a tree decomposition $(V_t)_{t \in T}$, the *torso* of the decomposition are the graphs obtained from $G[V_t]$, $t \in T$ by turning all sets $V_t \cap V_s$, $s \in N_T(t)$ into cliques—we think of these sets as the ‘gluing’ points between the bags of the decomposition.

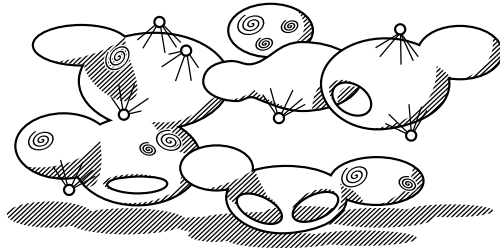
k-nearly embeddable

A graph G is *k-nearly embeddable* in surface S if we can remove up to k vertices from G and decompose it into at most k graphs $H_0 \cup H_1 \cup \dots \cup H_k$ such that

- H_0 can be embedded into S ,
- H_1, \dots, H_k are pairwise disjoint and intersect H_0 in exactly one face of the embedding, and
- the graphs H_1, \dots, H_k each have a path decomposition of width k that contains the vertices of the face the graph intersects with, and these vertices appear in the same order as on the face.

The third point is very vaguely formulated here: for our purposes here, we imagine that G will be embedded in S with the exception of a few vertices and a few parts that can be ‘crammed’ into path decompositions that attach nicely to the rest of the embedding.

Theorem 6 (Robertson and Seymour [216]). *For every $h \geq 5$ there exist a $k \in \mathbb{N}$ such that every graph that excludes K_h as a minor has a tree decomposition whose torsos are k -nearly embeddable in a surface in which K_h is not embeddable.*



We can replace the last part by ‘a surface whose genus only depends on h ’ to make the statement easier to digest. The power of the decomposition lies in the fact that it can be used to lift results on graphs of bounded genus up to graphs excluding a minor. For example, Baker’s approach to design PTAS [15] can be modified using the decomposition theorem to work for graphs excluding a minor [127, 235]. Again, this ‘by-product’ of the programme deeply affected the evolution and understanding of algorithms. Its original purpose was, of course, the big result that took twenty-one years and twenty papers to prove:

Theorem 7 (Graph minor theorem [217]). *Graphs are well-quasi ordered by the minor relation.*

Minor testing

The really far-reaching impact in the theory of computation, however, was created by the constructive minor test based on the decomposition [215]: for a fixed graph H , we can test whether it is contained in a graph G in time $f(H)|G|^3$. In the early days of parametrised complexity, this theorem was the key algorithmic tool for Downey and Fellows to design fpt algorithms (see, e.g., their early overview paper [73]). Take, for example, the VERTEX COVER problem. It is easy to verify that the class of graphs with a vertex cover of size at most k is minor-closed. Since graphs are wqo by the minor relation, therefore this class has a finite obstruction set whose size is some function of k . Hence, we can test in $f'(H)|G|^3$ time whether a given instance of VERTEX COVER belongs to the class. Historically, we can draw a straight line from the graph minor theorem to parametrised complexity, a field that has matured into a solid segment of theoretical computer science and provides efficient algorithms for hard problems! A better proof that the theory of sparse graphs goes hand in hand with the theory of efficient computation could not exist.

Topological wqo?

Returning back to the sparse class hierarchy, the immediate question is: are graphs also wqo by the topological minor order? As it turns out, there is a rather concise counterexample: take the class of all cycles with each edge duplicated. Clearly none of these graphs is a subgraph of another and any vertex dissolution will create a single edge without a parallel counterpart. Therefore they are pairwise incomparable in the topological minor relation, which proves that graphs are not wqo under it (Ding is often cited for the first construction of an anti-chain [68]). Note that if we insist on having simple graphs, taking the above class and subdividing every edge once does the trick.

Topological Bidimensionality?

The other big algorithmic tool, the grid minor theorem, also does not extend to classes excluding a topological minor. In fact, already classes of bounded degree pose an obstacle: a counterexample are d -regular expander graphs, who have treewidth $\Omega(n)$ and in particular do not admit sublinear balanced separators. As a consequence, the relationship between the treewidth and the size of a grid-minor is (at least) quadratic; hence bidimensionality theory cannot be used to obtain subexponential fpt algorithms³.

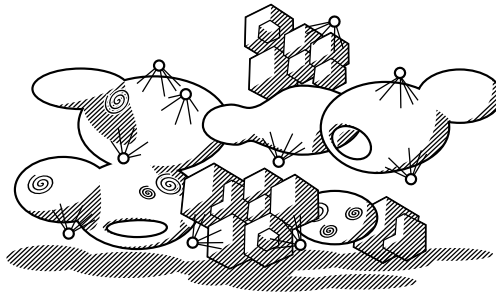
Nonetheless, there exists a decomposition theorem for graphs excluding a topological minor which made it possible to extend several algorithmic results known for classes excluding a minor.

Grohe-Marx decomposition

Theorem 8 (Grohe and Marx [130]). *For every fixed graph H , every graph excluding H as a topological minor has a tree decomposition where every torso*

- *either has bounded degree except for a bounded number of vertices, or*
- *is k -almost embeddable in a surface of bounded genus.*

Such a decomposition can be computed in time $f(H)|H|^{O(1)}$ for some recursive function f .



Grohe and Marx demonstrate the applicability of the decomposition by showing that it can be used to design fpt-algorithms and to test isomorphism in time $n^{f(H)}$. Both the Robertson-Seymour and the Grohe-Marx decomposition have also been applied in the field of kernelisation: Fomin, Lokshtanov, Saurabh, and Thilikos showed that **DOMINATING SET** and **CONNECTED DOMINATING SET** admit linear kernels for

³ Many thanks to Michał Pilipczuk for this explanation, which can also be found in the recently published book on parametrised algorithms [54].

graphs excluding a minor [102] and graphs excluding a topological minor [103].

Low treewidth colouring

A decomposition of a very different flavour for classes excluding a fixed minor—conjectured by Thomas—was proved by Devos *et al.*

Theorem 9 (Low-treewidth 2-colouring [66]). *For every graph H there exist integers k such that every graph excluding H as a minor has a vertex-partition into two parts that each induce a graph of treewidth at most k .*

Their paper contains some stronger but more technical results and also include an edge-partition variant. We further should mention a similar result by Demaine, Hajiaghayi, and Kawarabayashi for so-called odd-minor free graphs [64]. The aforementioned decompositions into two graphs of constant treewidth have, as far as I know, not been exploited algorithmically, save for the possibility of approximating the chromatic number of such graphs within a factor of two. But a variation of this idea will form mayor part of the algorithmic toolkit presented in the next part.

FO model-checking

Given that the first-order properties can be checked in linear fpt-time on graphs of bounded genus, it is natural to ask whether we can obtain a similar results for excluded-minor classes. Flum and Grohe provided a linear-time algorithm for classes excluding a minor [96] which was subsequently generalised by Dawar, Grohe, and Kreutzer to classes *locally* excluding a minor [58] (as one might suspect, these classes exclude a minor H_r in the r -neighbourhood of every vertex, for every $r \in \mathbb{N}$). Moreover, for an important fragment of first-order definable problems; Dawar, Grohe, Kreutzer, and Schweikardt showed that they admit an EPTAS in classes excluding a fixed minor [59]

Let us wrap up this short excursion in graph classes characterised by forbidden substructures. There exist a long list of further notions of substructures like immersions, contractions, and induced minors that encouraged interesting and fruitful research, both in terms of graph theory and algorithms. There has also been interest in ‘dense’ counterparts like vertex minors and pivot minors [200] as well as odd minors [64] whose algorithmic application is just now being explored. Our conclusion for this section is that the theory of sparse graphs has an immense impact on the theory of algorithms, in particular in the parametrised framework. This pattern continues with the graph class we are about to introduce next and that will be the main focus of this thesis.

We finally arrive at the centrepiece and focus of this thesis: the grad of graphs and graph classes of bounded expansion. As we can see in the hierarchy in Figure 1, graph classes of bounded expansion contain both classes with bounded degree and classes excluding a (topological) minor. The key ingredient to define these classes are shallow minors (see Section 2.2) and an appropriate ‘forbidden structure’ definition.

Definition 13 (Topological grad). For a graph G and an integer r , the *topological greatest reduced average density (grad) at depth r* is defined as

$$\tilde{\nabla}_r(G) = \max_{H \in \mathcal{G}_{\tilde{\nabla}_r}} \frac{\|H\|}{|H|}$$

For a graph class \mathcal{G} we define $\tilde{\nabla}_r(\mathcal{G}) = \sup_{G \in \mathcal{G}} \tilde{\nabla}_r(G)$.

In particular we have that $2\tilde{\nabla}_0(G)$ is exactly the degeneracy of G .

Bounded expansion, ∇ .

Now a graph class \mathcal{G} has *bounded expansion* if and only if there exists a function f such that $\tilde{\nabla}_r(\mathcal{G}) < f(r)$. Note that instead of forbidding concrete graphs to occur as substructure, we instead forbid *all* graphs whose density is above $f(r)$. There is an analogous measure, the *grad* ∇_r defined over shallow minors:

$$\nabla_r(G) = \max_{H \in \mathcal{G}_{\nabla_r}} \frac{\|H\|}{|H|}.$$

In particular, the two measures coincide at depth zero, *i.e.* we have that $\nabla_0(G) = \tilde{\nabla}_0(G)$. As a convention we will prefer the symbol ∇_0 in this case. We saw earlier that graph classes defined over excluded minors are quite different from those defined over excluded *topological* minors. A surprising feature of bounded expansion classes is that the flavour of minor does not matter.

Polynomial equivalence of ∇ and $\tilde{\nabla}$.

Proposition 6 (Nešetřil and Ossona de Mendez [192]).

For every graph G and half-integer r it is true that

$$\tilde{\nabla}_r(G) \leq \nabla_r(G) \leq 4(4\tilde{\nabla}_r(G))^{(r+1)^2}$$

Hence the topological grad of a graph class is bounded if and only if its grad is bounded. As we will see later, we can modify the definition of shallow minors even further and again obtain an alternative definition of bounded expansion. This property is one of the extremely robust features that make the theory of shallow minors so exciting. Let us improve the above bound for the case of $\frac{1}{2}$ -shallow minors, it will come in handy later on and gives a good idea how Proposition 6 can be proved.

Lemma 12. *For every graph G we have*

$$\tilde{\nabla}_{1/2}(G) \leq \nabla_{1/2}(G) \leq (2\tilde{\nabla}_{1/2}(G) + 1)^2.$$

Proof. The first inequality is trivial and we only have to show the second one. Consider a $1/2$ -shallow minor H of G with $\nabla_0(H) = \nabla_{1/2}(G)$, an appropriate shallow embedding ϕ_V of H into G . Let $S_1, \dots, S_{|H|}$ be the branch sets of H (that is, the sets $\{\phi_V(u)\}_{u \in H}$) and let M be the subgraph of G induced by these sets. Note that, by minimality, the subgraphs $M[S_i]$ are stars for $1 \leq i \leq |H|$.

Let $W := V(H)$ and $S := V(M) \setminus V(H)$. Every edge in H is either due to an edge with both endpoints in W or due to an edge with exactly one endpoint in W . Partition the edges of H into ‘direct’ edges E_D (those that are due to a W - W -edge in M) and ‘indirect’ edges E_I (those that are due to a W - S -edge in M).

If $|E_D| \geq \nabla_0(H)^{-1/2} \|H\|$ we are done: the subgraph $M[W] \subseteq G$ already has density

$$\frac{\|M[W]\|}{|W|} = \frac{|E_D|}{|W|} \geq \frac{\nabla_0(H)^{-1/2} \|H\|}{|H|} = \nabla_0(H)^{1/2} = \nabla_{1/2}(G)^{1/2}.$$

Otherwise, assume that $|E_I| \geq (1 - \nabla_0(H)^{-1/2}) \|H\|$. We partition the vertices W into two sets W_l, W_h as follows: W_h contains those vertices who have at least $\nabla_0(H)^{1/2}$ neighbours in $M \setminus (E_D \cup E_I)$ and W_l all others. We construct a $1/2$ -shallow topological minor of M by contracting all neighbours of vertices in W_h into (arbitrary) edges. Note that by the minimality of M this cannot create double edges. The density of the resulting minor is at least

$$\frac{1}{2} \cdot \frac{\nabla_0(H)^{1/2} |W_h| + (\nabla_0(H) - \nabla_0(H)^{1-1/2}) |W_l|}{|W_h| + \nabla_0(H)^{1/2} |W_l|}$$

This fraction takes its extreme values in the cases where $|W_h| = |W|$ and $|W_l| = |W|$. In these two cases we obtain the densities

$$\frac{1}{2} \nabla_0(H)^{1/2} \quad \text{and} \quad \frac{1}{2} \nabla_0(H)^{1/2} - \frac{1}{2},$$

respectively. It follows that

$$\frac{1}{2} \nabla_{1/2}(G)^{1/2} - \frac{1}{2} \leq \tilde{\nabla}_{1/2}(G) \iff \nabla_{1/2}(G) \leq (2\tilde{\nabla}_{1/2}(G) + 1)^2,$$

as claimed. \square

Algebraic robustness

The invariance of the bounded expansion property under different notions of shallow minors is no the only feature that makes it ‘robust’. Another feature is that we can modify the members of the class with certain operations such that the resulting class again has bounded expansion. Let us start with a very simple operation: adding a universal vertex, that is, a vertex connected to every other vertex of the graph.

Lemma 13. *For any graph G it holds that*

$$\tilde{\nabla}_r(G) \leq \tilde{\nabla}_r(G * u) \leq \tilde{\nabla}_r(G) + 1.$$

Proof. Observe that a topological minor of G can always be extended by the universal vertex u to obtain a topological minor of $G * u$. In the other direction, any topological minor of $G * u$ that uses u (as either nail or a subdivision vertex) can be turned into a topological minor of G by removing u . In both cases, the embedding of the minor does not change and hence its depths does not increase. Let H, H' be two such minors that only differ in that H' has the additional nail u . Then

$$\frac{\|H\|}{|H|} \leq \frac{\|H'\|}{|H'|} \leq \frac{\|H * u\|}{|H| + 1} \leq \frac{\|H\| + |H|}{|H| + 1} \leq \frac{\|H\|}{|H|} + 1,$$

from which the claim follows. \square

It follows directly that classes of bounded expansion are closed under the addition of a constant number of universal vertices. A second, more surprising, feature is that they are also closed under lexicographic products with constant-sized graphs.

Proposition 7 (Nešetřil and Ossona de Mendez [192]).

For every graph G , integer $p \geq 2$ and half-integer r it is true that

$$\tilde{\nabla}_r(G \bullet K_p) \leq \max\{2r(p-1)+1, p^2\} \cdot \tilde{\nabla}_r(G) + p - 1$$

This stability under lexicographic products is of major help in many of the following proofs and even has direct practical relevance for model of complex networks (see Section 15.2).

In conclusion, the density of shallow minors provides a robust definition for sparse classes. However, to obtain a complete theory of sparse classes, we would like to partition all classes into ‘dense’ and ‘sparse’: a dichotomy of structural sparseness. As indicated by the hierarchy in Figure 1, bounded expansion classes are not the largest possible structurally sparse class.

Before we consider larger classes, let us return to algorithmic applications. Some results will be discussed later in Chapter 10, among them a constant-factor approximation for r -DOMINATING SET by Dvořák and several smaller results following from work by Nešetřil and Ossona de Mendez. A result that needs to be mentioned right here is the continuation of the first-order model checking programme: Dvořák, Král and Thomas showed that again, a linear fpt-time algorithm is possible [81]. They exploit a property of bounded expansion classes that will be revisited in detail in Chapter 7, the so-called low-treewidth colourings. The result even extends to the next larger class, the graphs of *locally* bounded expansion.

Locally bounded expansion

As we saw in the previous sections, an easy way to obtain a larger class from a forbidden-structure characterisation is to ‘localise’ the definition: instead of graph classes that exclude a fixed minor,

we define classes in which every vertex' ℓ -neighbourhood induces a graph that excludes some minor H_ℓ , and obtain graph classes that *locally* exclude a minor. The same can be done for classes of bounded expansion: classes with *locally* bounded expansion have a bivariate function $f(\ell, r)$ depending on the locality parameter ℓ and the minor-depth r . Now, there are graph classes that have globally unbounded but locally bounded expansion—hence we have not yet arrived at a dichotomy of sparseness, which will be the topic of the next chapter.

Let us retrace the development of the structural sparseness dichotomy that makes the framework of shallow minors so alluring. It starts out with the following result, independently proved by Dvořák and Jiang.

Theorem 10 (Dvořák [78], Jian [148]). *Let $\ell \in \mathbb{N}$ and $\varepsilon > 0$. There exists integers $n_{\ell, \varepsilon}$ and c_ε such that every graph G with $n > n_{\ell, \varepsilon}$ vertices and at least $n^{1+\varepsilon}$ edges contains a c_ε -subdivision of K_ℓ .*

In short: dense shallow clique-minors are unavoidable if the density of a class grows as a superlinear polynomial of n . Using this fact, Nešetřil and Ossona de Mendez proved the following dichotomy result:

Proposition 8 (Nešetřil and Ossona de Mendez [191]).

Let \mathcal{G} be an infinite graph class. Then the limit¹

$$\limsup_{H \preceq_r^t G \in \mathcal{G}} \frac{\log \|H\|}{\log |H|}$$

is either ≤ 1 for all $r \in \mathbb{N}$ or it is 2 for some $r_0 \in \mathbb{N}$. In the former case, the class \mathcal{G} is nowhere dense. Otherwise, it is somewhere dense and $\omega(\mathcal{G} \tilde{\vee} r_0) = \infty$.

Proof. The limit obviously lies in the interval $[0, 2]$. If there exists a number m_0 such that $\|G\| \leq m_0$ for all $G \in \mathcal{G}$, then the limit is 0. Otherwise we have already that for any m_0 , there exists a graph $G \in \mathcal{G}$ such that some 0-shallow topological minor—i.e. a subgraph— H of G satisfies $\|H\| \geq m_0$. We can assume that H does not contain isolated vertices, therefore $\|H\| \geq |H|$ and it follows that $\log \|H\| / \log |H| \geq 1$. Accordingly, we see that the limit cannot take any value in $(0, 1)$.

Now assume that for some $r \in \mathbb{N}$ and some $0 < \varepsilon < 1$ the limit takes on a value $\geq 1 + \varepsilon$: hence there is an infinite sequence of graphs H_1, H_2, \dots satisfying $\lim_{i \rightarrow \infty} \log \|H_i\| / \log |H_i| \geq 1 + \varepsilon$.

By Lemma 11, each such graph H_i contains a $\|H_i\|/2|H_i|$ -core H'_i with at least $\|H_i\|/2$ edges and a least $\sqrt{\|H_i\|}$ vertices. This gives rise to the infinite (since $\log_{i \rightarrow \infty} |H'_i|^{1/2} = \infty$) sequence $H'_{f(1)}, H'_{f(2)}, \dots$ of graphs, where f is an appropriate selection function to ensure that $|H'_{f(i)}| < |H'_{f(i+1)}|$, which has the property that

$$\lim_{i \rightarrow \infty} \frac{\log \|H'_{f(i)}\|}{\log |H'_{f(i)}|} \geq \lim_{i \rightarrow \infty} \frac{\log \|H_i\|}{2 \log |H_i|^{1/2}} \geq \lim_{i \rightarrow \infty} \frac{\log \|H_i\|}{\log |H_i|} \geq 1 + \varepsilon$$

¹ We evaluated the fraction $\log \|H\| / \log |H|$ symbolically in that we define it to be 0 in case $\|H\| = 0$.

Since the sequence $\{|H'_{f(i)}|\}_{f(i) \in \mathbb{N}}$ increases monotonically, there exists some $i_0 \in \mathbb{N}$ such that for $f(i) > i_0$ it holds that

$$\frac{\log \|H'_{f(i)}\|}{\log |H'_{f(i)}|} \geq 1 + \frac{\varepsilon}{2} \Leftrightarrow \|H'_{f(i)}\| \geq |H'_{f(i)}|^{1+\varepsilon/2}$$

Now Theorem 10 implies that for every ℓ , there exists $i_{\ell, \varepsilon/2}$ and $c_{\varepsilon/2}$ such that for $f(i) > i_{\ell, \varepsilon/2}$ we have that $K_\ell \preceq_t^{c_{\varepsilon/2}} H'_{f(i)}$, i.e. $H'_{f(i)}$ contains a $\leq c_{\varepsilon/2}$ -subdivision of K_ℓ as a subgraph. Since $H'_{f(i)} \subseteq H_i \preceq_t^r G_{g(i)}$ for some sequence $\{g(i)\}_{i \in \mathbb{N}}$, we have that K_ℓ is a $(c_\varepsilon/2 + 1)(r + 1)$ -shallow minor of $G_{g(i)}$. In other words: the class \mathcal{G} contains arbitrarily large complete subgraphs as r_0 -shallow topological minors, where $r_0 = (c_\varepsilon/2 + 1)(r + 1)$, and hence the above limit is two. \square

ω_* , $\tilde{\omega}_*$

We arrived at the hilltop of our ascend through the sparse hierarchy. From up here we have the complete overview: *every* graph class can be categorised as structurally sparse or dense using the above limit. For a more intuitive definition of how nowhere dense classes look like, we return to shallow minors. To that end, let us introduce a parametrised version of the clique number of a graph. We define

$$\tilde{\omega}_r(G) = \omega(G \tilde{\nabla} r) \text{ and } \omega_r(G) = \omega(G \nabla r).$$

Definition 14 (cf. [192]). A graph class \mathcal{G} is *nowhere dense* if and only if for every integer r it holds that $\tilde{\omega}_r(\mathcal{G}) < \infty$.

It follows immediately that graphs classes of bounded expansion are nowhere dense: we have the simple relation $\tilde{\omega}_r(\mathcal{G}) \leq \tilde{\nabla}_r(\mathcal{G})^2$. The converse is not true, it is in particular possible for nowhere dense classes to have a superlinear density. Nonetheless we can transfer most of the results for classes of bounded expansion to nowhere dense classes. The rule of thumb is that if a quantity is guaranteed to be a constant in a class of bounded expansion, it is asymptotically (for algorithmic purposes) of the order $n^{o(1)}$.

Let us consider a small example to demonstrate the typical line of reasoning. For a nowhere dense class \mathcal{G} , Proposition 8 implies that

$$\limsup_{H \subseteq G \in \mathcal{G}} \log \|H\| / \log |H| \leq 1$$

which means that for every $\varepsilon > 0$, there exists N_ε such that

$$\log \|H\| / \log |H| \leq 1 + \varepsilon \quad \text{for } H \subseteq G \in \mathcal{G}_{\geq N_\varepsilon}.$$

This is of course equivalent to $\|H\| \leq |H|^{1+\varepsilon}$ and means that the degeneracy of a nowhere dense class is asymptotically bounded by $O(n^\varepsilon)$ for every ε , or simply $O(n^{o(1)})$. To see the algorithmic implications of such a bound, consider the algorithm to compute a degeneracy-ordering of a graph: as described in the very beginning of this chapter, this algorithm takes time $O(dn)$ where d is the degeneracy of the input graph. By the previous observation, this algorithm will work in

almost linear time for nowhere dense classes: for any $\varepsilon > 0$, for large enough inputs, it will take time $O(n^{1+\varepsilon})$. Note that by manipulating the ε , we can hide any polynomial dependence on d . For an algorithm with running time $O(d^c n)$, for some constant c , we chose $\varepsilon' = \varepsilon/c$ and choose our graphs large enough such that the degeneracy drops below $n^{\varepsilon'}$. Then the running time of the hypothetical algorithm turns out to be (asymptotically)

$$O(d^c n) = O(n^{\varepsilon' c} n) = O(n^{1+\varepsilon}).$$

FO model-checking

Now, we finally pose the question: is first-order model checking still possible using the most general notion of structural sparseness? The result was claimed by Dawar and Kreutzer [164] and independently by Dvořák, Král, and Thomas. The former set of authors retracted their claim after finding a flaw in their proof, the latter published the weaker statement for graphs of bounded expansion and locally bounded expansion [81]. Finally in 2013, Grohe, Kreutzer, and Siebertz succeeded in proving that first-order properties can be checked in almost linear fpt-time in nowhere dense classes [129]. And it turns out that nowhere dense classes are the limit for efficient first-order model checking:

Theorem 11 (Dawar and Kreutzer [164]). *If \mathcal{G} is a monotone graph class and effectively² somewhere dense, then the first-order model checking problem for \mathcal{G} is not in FPT unless $\text{FPT} = \text{AW}[*]$.*

This even extends to the fragment of Σ_1 -formulas, *i.e.* first-order formulas of the form $\exists x_1 \dots \exists x_p \phi(x_1, \dots, x_p)$ where ϕ is quantifier-free:

Theorem 12 (Dvořák, Král, Thomas [81]). *If \mathcal{G} is a monotone graph class and somewhere dense, then the Σ_1 -model checking problem for \mathcal{G} is not in FPT unless $\text{FPT} = \text{W}[1]$.*

We should also note that Dawar and Kreutzer showed earlier that r -DOMINATING SET is in FPT for nowhere dense classes [60], a result now subsumed by the above first-order meta-theorem.

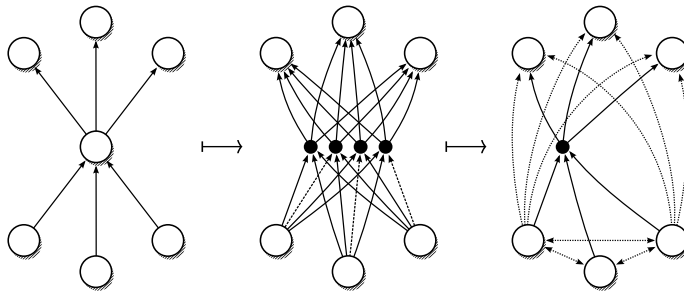
Locally nowhere dense

As a final consideration, note that the ‘localisation-trick’ from earlier does not result in a new class: if a graph class is locally nowhere dense, it is also simply nowhere dense. This is easy to see if one considers an r -shallow clique minor: its embedding has a diameter of $\Theta(r)$, therefore any graph that locally does not contain such a structure also excludes it globally. We have indeed reached the peak of structural sparseness.

² A class is effectively somewhere dense if for every graph H and integer r , one can compute in polynomial time a member G_H of the class that contains an r -subdivision of H .

Part II

THE BOUNDED EXPANSION TOOLKIT



The fundamental proofs presented by Nešetřil and Ossona de Mendez [192] about bounded-expansion classes often alternate between shallow minors and their topological cousins. While every such alternation introduces a worst-case estimate and as such should be used sparingly, the benefit of ‘switching gears’ is often a proof that is much easier to comprehend.

Recognising the strength of having a variety of ‘minor-flavours’, we introduce in this chapter the notions of *weighted* and *stable* minors. Weighted minors simply inherit some edge-weights of the host-graph and will be very useful in Chapter 7, where such weighting will naturally crop up in yet another characterisation of bounded expansion classes. Stable minors simply have the same depth in their branch sets (in the minor variant) or have the same path-lengths (in the topological variant). Fernando Sánchez and I found that this simple restriction can be immensely helpful to reduce the complexity of proofs and, additionally, the loss in ‘precision’ is not too bad (*cf.* our paper with Farrel, Goodrich, Lemons, and Sullivan [89]).

Before we come to these new notions, let us formulate some helpful ideas that permeate all the following proofs. Some vocabulary here is new, but most of the following statements are known.

Definition 15 ($\tilde{\nabla}_r$ -critical). A graph G is $\tilde{\nabla}_r$ -critical if for every proper subgraph $G' \subsetneq G$ it holds that $\tilde{\nabla}_r(G') < \tilde{\nabla}_r(G)$

Critical graphs impose a maximality condition on their density that can be exploited nicely. In particular, we have that ∇_0 -critical graph have a lower bound on their minimal degree.

Lemma 14. *Let G be a ∇_0 -critical graph. Then $\delta(G) \geq \nabla_0(G) = \bar{d}(G)/2$.*

Proof. Consider $v \in G$. Since G is $\tilde{\nabla}_0$ -critical, we have that

$$\frac{\|G\|}{|G|} > \frac{\|G\| - d(v)}{|G| - 1} \Leftrightarrow d(v) > \frac{\|G\|}{|G|}.$$

Further, note that being ∇_0 -critical implies that $\|G\|/|G| = \nabla_0(G)$. \square

Pairing the density-maximality condition with a model-minimality condition, we obtain the following statement.

Lemma 15. *Let G be $\tilde{\nabla}_r$ -critical and let $M \in G \tilde{\nabla} r$ be a minor with the minimal number of vertices that satisfies $\tilde{\nabla}_0(M) = \tilde{\nabla}_r(G)$. Then the following statements hold:*

1. G itself is a model of M .

2. M is $\tilde{\nabla}_0$ -critical.
3. $|M| \leq |G| \leq (1 + 2r\tilde{\nabla}_r(G))|M|$.

Proof. The first statement follows easily: let $H \subseteq G$ be a model of M . Since $\tilde{\nabla}_r(H) = \tilde{\nabla}_0(M) = \tilde{\nabla}_r(G)$ and since G is $\tilde{\nabla}_r$ -critical, it follows that $H = G$. For the second statement, assume there exists a proper subgraph $M' \subset M$ such that $\tilde{\nabla}_0(M') = \tilde{\nabla}_0(M)$. But then $M' \in G \tilde{\nabla} r$ and obviously $|V(M')| < |M|$, contradicting our choice of M . The third statement follows from the first. Since G is a model of M , its size is bounded by the number of nails and subdivision vertices:

$$|G| \leq |M| + 2r||M|| = (1 + 2r\tilde{\nabla}_0(M))|M| = (1 + 2r\tilde{\nabla}_r(G))|M|.$$

The other bound is trivial. \square

Lemma 14 and Lemma 15 imply the following Corollary that bounds the minimal degree of $\tilde{\nabla}_r$ -critical graphs.

Corollary 3. *Let G be a $\tilde{\nabla}_r$ -critical graph. Then the graph G contains at least $|G|/(1 + 2r\tilde{\nabla}_r(G))$ vertices of degree at least $\tilde{\nabla}_r(G)$.*

Sometimes switching from topological to regular shallow minors involves mixing operators like ∇_\bullet and $\tilde{\nabla}$ and we want to ‘normalise’ these operators to the same minor-flavour. The following theorem is one of the tools necessary for such operations.

Theorem 13 (Nešetřil and Ossona de Mendez [192, Theorem 4.2]).
For every graph G , integer r and half-integer s it holds that

$$\tilde{\nabla}_s(G \tilde{\nabla} r) \leq \tilde{\nabla}_s(G \nabla r) \leq 2^{r+2} 3^{(r+1)(r+2)} \tilde{\nabla}_s(G \tilde{\nabla} r)^{(r+1)^2}$$

and for half-integers r

$$\tilde{\nabla}_s(G \tilde{\nabla} r) \leq \tilde{\nabla}_s(G \nabla r) \leq 2^{r+2} 3^{(r+1)(r+2)} \tilde{\nabla}_s(G \tilde{\nabla} r)^{(\lceil r \rceil + 1)^2}.$$

We saw already in Section 2.2 through Propositions 2 and 3 that the repeated operation of taking shallow (topological) minors results in graphs that are themselves shallow (topological) minors. The following lemma is a direct consequence of these observations and helps us to reduce expression that involve both a density-operator like $\tilde{\nabla}_\bullet$ and a minor-operator like $\tilde{\nabla}$.

Lemma 16 (Nešetřil and Ossona de Mendez [192, Prop. 4.1 and 4.2]).
Let a, b be half-integers and let

$$c := \frac{(2a + 1)(2b + 1) - 1}{2}.$$

Then for every graph G it holds that

$$\nabla_b(G \nabla a) \leq \nabla_c(G) \quad \text{and} \quad \tilde{\nabla}_b(G \tilde{\nabla} a) = \tilde{\nabla}_c(G).$$

For similar situations in nowhere-dense classes, we will need the following lemma.

Lemma 17 (Nešetřil and Ossona de Mendez [192, Proposition 5.2]).
Let a be a half-integer. Then for every graph G it holds that

$$\omega(G \tilde{\nabla} a) \leq \omega(G \nabla a) \leq 2\omega(G \tilde{\nabla} (3a + 1))^{\lfloor a \rfloor + 1}.$$

The above results are extremely helpful: we can change mid-proof between shallow minors and their topological variant and will still be able to collect all our terms in the end and express everything in one measure. In what follows we will introduce more shallow minor flavours and prove similar bounds to enrich the minor-related part of the toolkit.

6.1 STABLE MINORS

Let us begin by defining a new minor flavour here which we dubbed *stable*. Again, in the usual settings of excluded-minor, that is, without the bound on the minor's depth, such stable minors would be very different beasts than their non-stable counterpart. By parametrising the minors by their depth and reducing them to their density, however, we again obtain polynomial equivalences with the previously established measures ∇_\bullet and $\tilde{\nabla}_\bullet$.

Stable topological embedding

A *stable topological embedding* is a topological embedding ϕ_E, ϕ_V of a minor H in a graph G if the following two criteria are met: the paths $\phi_E(uv), uv \in H$ are induced paths in G and have all the same length. If such an embedding of H exists with depth $2r + 1$, we say that H is an *r -stable topological minor* of G and write $H \stackrel{r}{\preceq}_t G$. Note that we drop the 'shallow' here since the nomenclature is already verbose enough. In concordance with the established notation, we define $G \stackrel{\check{\nabla}}{\preceq} r$ to be the set of r -stable topological minors of G and define a related parametrised graph measure

$$\check{\nabla}_r(G) := \max_{H \in G \stackrel{\check{\nabla}}{\preceq} r} \frac{\|H\|}{|H|}$$

which we extend to classes via $\check{\nabla}_r(\mathcal{G}) = \sup_{G \in \mathcal{G}} \check{\nabla}_r(G)$, as usual.

Let us first relate this density measure to the topological measure $\tilde{\nabla}_\bullet$. The following proof is quite easy and relies on the observations that a) we can easily enforce the paths of a shallow minor-embedding to be induced and b) we can simply restrict ourselves to those paths whose length appears most frequently.

Lemma 18. *For every graph G and half-integer r it is true that*

$$\check{\nabla}_r(G) \leq \tilde{\nabla}_r(G) \leq (2r + 1)\check{\nabla}_r(G).$$

Proof. The first inequality is trivially true since every r -stable topological minor is in particular r -shallow. To see that the second inequality

holds, fix a half-integer r and let $H \preceq_t^r G$ be such that $\nabla_0(H) = \tilde{\nabla}_r(G)$ with the topological minor embedding ϕ_V, ϕ_E as a witness. Note that we can easily assume that $\phi_E(uv), uv \in H$ is an induced path; otherwise we simply short-circuit it and adapt ϕ_E accordingly.

Now consider the edge-labelling λ of the minor H defined simply via $\lambda(uv) := \|\phi_E(uv)\|$ for $uv \in H$. We partition the edges of H into $E(H) = E_1 \uplus \dots \uplus E_{2r+1}$ where $E_\ell := \{e \in E(H) \mid \lambda(e) = \ell\}$. Then for some $1 \leq \ell \leq 2r+1$ it must hold that $|E_\ell| \geq \|H\|/(2r+1)$ and $H[E_\ell]$ is a stable topological minor of G of the claimed density. \square

A very similar trick can be applied to embedding of shallow minors and we formalise it by introducing r -stable minors here.

Stable minor embedding

A *stable minor embedding* of a graph H in a graph G is a pair of functions $\phi_V: V(H) \rightarrow 2^{V(G)}$ and $\phi_C: V(H) \rightarrow V(G)$ and integers r_1, r_2 where ϕ_V is a minor embedding of H and additionally

- $\phi_C(u) \in \phi_V(u)$, and
- if $uv \in H$ then there exists $x \in \phi_V(u), y \in \phi_V(v)$ such that $xy \in G$ and further $\text{dist}_G(\phi_C(u), x) = r_1$ and $\text{dist}_G(\phi_C(v), y) = r_2$ (or vice-versa).

The *depth* of a stable minor embedding is $r = \max\{r_1, r_2\}$. If an embedding of depth r exists we say that H is an *r -stable minor* of G and we write $H \preceq_m^r G$. Again, we define $G \dot{\nabla} r$ to be the set of r -stable topological minors of G and obtain the corresponding density measure

$$\dot{\nabla}_r(G) := \max_{H \in G \dot{\nabla} r} \frac{\|H\|}{|H|}.$$

The extension to classes works as usual by taking the class-supremum.

Stable minors are structurally a bit more complex and the following annotated version of their model will be useful. Recall that a Trémaux tree is a rooted spanning tree of a graph such that all non-tree edges obey the ancestor-relationship induced by the tree.

Definition 16 (Extended model). Let H be an r -stable minor of G with embedding ϕ_V . The *extended model* of H in G is a collection of trees $\mathcal{T} = \{T_v\}_{v \in H}$ of height at most $r+1$ and a mapping $\phi_E: E(H) \rightarrow E(G)$ such that

- for all $v \in H$ we have that $T_v \subseteq G[\phi_V(v)]$ and T_v is a Trémaux-tree around $\text{root}(T_v)$ in G ,
- for every leaf x of $T_v, v \in H$, there exists an edge $v \bullet \in H$ such that $\phi_E(v \bullet)$ is incident to x , and
- for $uv \in H$ the edge $\phi_E(uv)$ connects $V(T_u)$ and $V(T_v)$ in G .

This definition is very similar to what is called a *retract* [192]: the idea is to have a bare-bone skeletal model that still gives rise to the minor H . Note that in the case of stable minors, an edge $\phi_E(uv)$ connects level r_1 in T_u with level r_2 in T_v (being a bit sloppy, we treat the edges of H

as arcs in this case) and we assume that the centres $\phi_C(H)$ are exactly the roots of the trees \mathcal{T} .

Lemma 19. *For every graph G and half-integer r it is true that*

$$\check{\nabla}_r(G) \leq \nabla_r(G) \leq (r+1)^2 \check{\nabla}_r(G).$$

Proof. Again the first inequality is trivial. For the other direction, assume H is an r -shallow minor of G with $\nabla_0(H) = \nabla_r(G)$. Let \mathcal{T}, ϕ_E be H 's extended model. Let $\lambda: \phi_E(E(H)) \rightarrow [r+1]^2$ be a mapping that labels every edge $xy = \phi_E(uv)$ with the depths of x and y in their respective trees of \mathcal{T} , i.e.

$$\lambda(xy) = (\text{depth}_{T_u}(x), \text{depth}_{T_u}(y)).$$

Partition the edges $\phi_E(E(H))$ into sets $E_{(1,1)}, \dots, E_{(r+1,r+1)}$ by their labels so that $E_{(r_1,r_2)} = \{xy \in \phi_E(E(H)) \mid \lambda(xy) = (r_1, r_2)\}$.

By a simple averaging argument, we have that for some label (r_1, r_2) it holds that $|E_{(r_1,r_2)}| \geq (r+1)^2 \|H\|$. Retaining only these edges in the model gives rise to a minor $H' \subseteq H$ that is an r -stable minor of G (defining ϕ_C to map $v \in H$ to $\text{root}(T_v)$) completes the embedding alongside ϕ_V and r_1, r_2 . \square

We can relate the parametrised graph measures $\check{\nabla}_\bullet$ and ∇_\bullet in the same fashion as Proposition 6.

Lemma 20. *For every graph G and half-integer r it is true that*

$$\check{\nabla}_r(G) \leq \nabla_r(G) \leq (2r+1)2^{(r+1)^2+1} \check{\nabla}_r(G)^{(r+1)^2}.$$

Proof. The first inequality follows easily: the model of any r -stable topological minor H of G uses paths of exactly the same lengths in its embedding, let us denote this length by $\ell \leq 2r+1$. Now we partition the vertices of these paths into two (almost) equal halves and associated them with the closest nails to form branch-sets around the nails. Clearly, every edge between the branch sets connects vertices that are at depth $\lfloor \ell/2 \rfloor$ and $\lceil \ell/2 \rceil$, respectively; it follows that H has a r -stable embedding and is in particular an r -stable minor of G .

For the second inequality, assume that H is an r -stable minor of G with $\nabla_0(H) = \nabla_r(G)$. Let ϕ_V, ϕ_C, r_1, r_2 be the stable embedding of H in G and let \mathcal{T}, ϕ_E be the corresponding extended model of H . We will work on the extended model itself and keep track of the density of its quotient to argue that we can recover a topological r -stable minor with sufficiently high density.

Let \vec{H} be the orientation of H constructed by the following rule: if $uv \in H$ such that $\phi_E(uv)$ points from a depth r_1 -vertex in T_u to a depth r_2 -vertex in T_v , then orient the edges uv to point from u to v ; otherwise orient it from v to u . Applying Corollary 1, we obtain a subgraph \vec{H}' of \vec{H} where every vertex is either a source or a sink and where $\|\vec{H}'\| \geq \|\vec{H}\|/4$. In particular, note that the graph \vec{H}' is bipartite.

We assume that $\delta(H') \geq 2\nabla_0(H')$, otherwise we remove the vertex of minimal degree until this property is satisfied (this operation can only increase the density of H').

Let $\mathcal{T}^+ := \{v \in \vec{H}' \mid \deg^-(v) = 0\}$ be those trees of \mathcal{T} corresponding to sources in \vec{H}' and $\mathcal{T}^- := \mathcal{T} \setminus \mathcal{T}^+$ those corresponding to sinks. Let further $E' = \phi_E(E(H'))$ be those edges of G that connect vertices of \mathcal{T}^+ to vertices of \mathcal{T}^- and participate in the formation of the minor H' .

We prune the trees in both sets by iteratively removing leaves that are not incident to an edge in E' . As a result, all trees in \mathcal{T}^+ have height exactly r_1 and all trees in \mathcal{T}^- height r_2 (assuming that H had no vertex of degree zero, the orientation \vec{H} will not have such a vertex either and we will not have a tree of height zero after pruning).

We will in the remainder assume that \mathcal{T}^- is larger than \mathcal{T}^+ , otherwise the following operations need to be applied in the reverse order. Let us modify every tree $T_v \in \mathcal{T}^+$ to create a new set \mathcal{S}^+ and a new set of edges E'' (which initially is set to E') by proceeding as follows: let \hat{T}_v be the tree derived from T_v by adding to every leaf $x \in \text{leaves}(T_v)$ the vertices

$$\{y \in T_u \in \mathcal{T}^- \mid \phi_E(vu) = xy\}$$

as children. That is, the node x has in \hat{T}_v exactly as many children as arcs going to r_2 -level vertices of trees in \mathcal{T}^- . Using Lemma 6, we construct a spider $\hat{S}_v \subseteq \hat{T}_v$ such that

$$(\text{leaves}(\hat{S}_v) \cap \text{leaves}(\hat{T}_v)) \geq \text{leaves}(\hat{T}_v)^{1/r_1}.$$

Note that Lemma 6 guarantees us that \hat{S}_v height at least two. We therefore can safely let $S_v = \hat{S}_v \setminus (\hat{T}_v \setminus T_v)$ be the spider where we removed the previously added leaves. We add S_v to the new collection \mathcal{S}^+ and continue until every tree of \mathcal{T}^+ has been processed.

We constructed the spiders in a way that encoded not only which leaves of a tree $T_v \in \mathcal{T}^+$ to retain, but also which outgoing edges from that leaf should be kept (hence the extra level of leaves). We keep track of this choice by defining the edge set $E'' \subseteq E'$ as

$$E'' := \{xy \in E' \mid x \in T_v \text{ and } xy \in \hat{S}_v\},$$

i.e. we keep exactly those edges of E' that are still present in the (extended) spiders. Note that the degree of as spider in \mathcal{S}^+ is precisely the number of edges from E'' it is incident to. We therefore have that

$$\text{leaves}(S_v) = (\text{leaves}(S_v) \cap \text{leaves}(T_v)) \geq \deg_{H'}(v)^{1/r_1}.$$

We again prune the trees in \mathcal{T}^- by iteratively removing leaves not incident to E'' and further remove every tree that is incident to less than $\delta(H')^{1/r_1}$ edges of E'' .

Repeating the same construction on the other side for \mathcal{T}^- , we obtain a set of spiders \mathcal{S}^- and edge set $E''' \subseteq E''$ such that

- for every pair of spiders $S_1 \in \mathcal{S}^+, S_2 \in \mathcal{S}^-$ there exists at most one edge $xy \in E'''$ connecting them, and
- every leaf $x \in S \in \mathcal{S}^+ \cup \mathcal{S}^-$ has exactly one edge in E''' incident to it and every non-leaf has no such edge.

We further have for every tree $T_u \in \mathcal{T}^-$ that

$$\text{leaves}(\hat{S}_v) \geq \text{leaves}(\hat{T}_u)^{1/r_2} \geq \delta(H')^{1/(r_1 r_2)},$$

since the number of edges of E'' incident to a tree T_u is exactly the number of leaves of \hat{T}_u . Note that the spiders $\mathcal{S}^+ \cup \mathcal{S}^-$ in conjunction with the edges E''' can be turned into a graph H^* by contracting paths of length at most $r_1 + r_2 - 1$ and with at least

$$\|H^*\| \geq |S^-| \delta(H')^{1/(r_1 r_2)} \geq |S^-| \left(\frac{\nabla_0(H)}{2} \right)^{1/(r_1 r_2)}$$

many edges. By our assumption that $|\mathcal{T}^-| \geq |\mathcal{T}^+|$ and hence $|S^-| \geq |S^+|$ we can rearrange the above inequality and see that

$$\nabla_0(H^*) \geq \frac{\|H^*\|}{|H^*|} \geq \frac{\|H^*\|}{|H|} \geq \frac{\|H^*\|}{2|S^-|} \geq \frac{1}{2} \left(\frac{\nabla_0(H)}{2} \right)^{1/(r_1 r_2)}.$$

Since H^* is a $(\frac{r_1+r_2}{2} - 1)$ -shallow topological minor of G , we conclude (invoking Lemma 18 to obtain a stable topological minor) that

$$\dot{\nabla}_r(G) \leq 2^{r_1 r_2 + 1} \tilde{\nabla}_r(G)^{r_1 r_2} \leq (r_1 + r_2 - 1) 2^{r_1 r_2 + 1} \check{\nabla}_r(G)^{r_1 r_2}.$$

Since $r_1 + r_2 = 2r + 2$ and the term $r_1 r_2$ is maximised when r_1 and r_2 are both equal to $r + 1$ (which can be easily seen by expressing both values as fractions of $r + 1$ and then considering the function $x(1 - x)$), we finally obtain

$$\dot{\nabla}_r(G) \leq (2r + 1) 2^{(r+1)^2 + 1} \check{\nabla}_r(G)^{(r+1)^2}.$$

In the case that r is a half-integer, note that the term $r_1 r_2$ is maximal when it takes the form

$$\lfloor r + 1 \rfloor \cdot \lceil r + 1 \rceil = (r + 1 - 1/2)(r + 1 + 1/2) = (r + 1)^2 - 1/4,$$

and hence the above inequality to half-integers r as well. \square

The notion of stable minors works as well if we consider the ‘denser’ measure ω . We mention the following simple results for completeness, they will not be needed in the following. In this setting, we need to replace the simple averaging-argument by Ramsey theory, where the following formulation of the classic Ramsey Theorem already suffices.

Theorem 14 (Finite Ramsey Theorem). *For integers s, t there exists an integer $r(s, t)$ such that any two-colouring of the edges of $K_{r(s, t)}$ contains either a monochromatic K_s or a monochromatic K_t .*

A well-known upper bound is that

$$r(s, t) \leq \binom{s+t-2}{t-1},$$

which follows from the easily proved recurrence

$$r(s, t) \leq r(s-1, t) + r(s, t-1).$$

The currently best known bound was proved by Thomason [237]: he showed that there exists a constant c such that

$$r(s+1, t+1) \leq e^c \sqrt{\log s - t \log s^{2s}} \binom{s+t}{s}.$$

A common special case occurs when $s = t$, the numbers $r(s) := r(s, s)$ are called the *diagonal* Ramsey numbers. Thomason obtained in the above mentioned paper the bound

$$r(s+1, s+1) \leq s^{c/\sqrt{\log k-1/2}} \binom{2s}{s}$$

for the diagonal case. In the context of this thesis we will be satisfied knowing that $r(s) = O(2^s/\sqrt{s})$. The diagonal two-colour case can be easily extended to multiple colours.

Lemma 21. *For integers s, p there exists an integer $r_p(s)$ such that any p -colouring of the edges of $K_{r_p(s)}$ contain some monochromatic subgraph K_s .*

Proof. We use induction over p , the case for $p = 2$ being described by Theorem 14. Hence, consider some $p > 2$ and assume the statement holds true for $p - 1$.

We claim that $r_p(s) \leq r_{p-1}(r_2(s))$. Consider a complete graph on $r_{p-1}(r_2(s))$ vertices and let c be an edge-colouring of this graph with p colours. Let c' be a colouring derived from c by unifying two arbitrary colours. Since c' colours the edges with only $p - 1$ colours, by the induction hypothesis, there exists a monochromatic $K_{r_{p-1}(s)}$ as a subgraph. If this subgraph is monochromatic under c we are done since $s \leq r_2(s)$. Otherwise, $K_{r_2(s)}$ receives exactly two colours by c (the two colours we merged) and by our induction hypothesis therefore contains a monochromatic K_s as a subgraph. \square

The proof of Lemma 21 also provides us with a recurrence and thus a simple bound on r_p : it can be upper-bounded by a tower of two of height $p - 1$. After this small tangent, let us return to the matter at hand.

Lemma 22. *For every graph G and half-integer r it is true that*

$$\omega(G \heartsuit r) \leq \omega(G \nabla r) \leq r_{(r+1)^2}(\omega(G \heartsuit r)),$$

where $r_p(\bullet)$ denotes the diagonal Ramsey number with p colours.

Proof. Again the first inequality is trivial. For the other direction, assume K_h is an r -shallow minor of G where $h = \omega(G \dot{\vee} r)$. Let \mathcal{T}, ϕ_E be K_h 's extended model in G and let further $\lambda: \phi_E(E(H)) \rightarrow [r+1]^2$ be a mapping that labels every edge $xy = \phi_E(uv)$ with the depths of x and y in their respective trees of \mathcal{T} , i.e.

$$\lambda(xy) = (\text{depth}_{T_u}(x), \text{depth}_{T_u}(y)).$$

Partition the edges $\phi_E(E(K_h))$ into sets $E_{(1,1)}, \dots, E_{(r+1,r+1)}$ by their labels so that $E_{(r_1,r_2)} = \{xy \in \phi_E(E(K_h)) \mid \lambda(xy) = (r_1, r_2)\}$.

Now by Ramsey's theorem (cf. Lemma 21) we have that K_h coloured with the above edge-partition contains a monochromatic complete subgraph of size at least $r_{(r+1)^2}^{-1}(h)$ and the claim follows. \square

The same idea can be applied to the topological operator $\check{\vee}$. We omit the proof here since it follows easily by replacing the averaging in the proof of Lemma 18.

Lemma 23. *For every graph G and half-integer r it is true that*

$$\omega(G \check{\vee} r) \leq \omega(G \tilde{\vee} r) \leq r_{(2r+1)}(\omega(G \tilde{\vee} r)),$$

where $r_p(\bullet)$ denotes the diagonal Ramsey number with p colours.

6.2 WEIGHTED TOPOLOGICAL MINORS

Given a graph with edge weights, we can restrict shallow minors not only by their depth, but by their *weighted* depth. We will need such *weighted topological minors* primarily in the next Chapter in order to obtain better bounds than available by unweighted minors.

*Weighted topological embedding,
- minor*

A topological embedding ϕ_E, ϕ_V of a graph H in a weighted graph (G, ω) is γ -weighted if for every $uv \in H$ it is true that

$$\sum_{e \in E(\phi_E(uv))} \omega(e) \leq \gamma.$$

Accordingly, we define the γ -weighted r -shallow minor relation as $H \preceq_r^\gamma G$ and the corresponding parametrised density measure

$$\tilde{\nabla}_r^\gamma(G) = \max_{H \preceq_r^\gamma G} \frac{\|H\|}{|H|}.$$

The above definition is compatible with the notion of stable topological minors: note that we only insist that the *length* of the paths used in the embedding have the same length, not necessarily their *weight* (such a variation would probably be useful in some cases, but for our application this definition is better suited). Accordingly, we define $\check{\nabla}_r^\gamma$ over the r -stable, γ -weighted minors as usual.

The following is pretty much the proof of Proposition 4.6 in [192]: we demonstrate that it is easily adapted to preserve weighted minors (a fact we need later). The bounds are slightly worse in favour of making the proof more lightweight.

Lemma 24 (Stability under blow-ups). *For every graph G , half-integer r and integer s it is true that*

$$\tilde{\nabla}'_r(G \bullet \bar{K}_s) \leq (4(r+1)(s-1) + 1)s^2 \cdot \tilde{\nabla}'_r(G)$$

Proof. Let H be a topological minor of G with an r -stable and γ -weighted embedding ϕ_E, ϕ_V in $G \bullet \bar{K}_s$ such that $\nabla_0(H) = \tilde{\nabla}'_r(G \bullet \bar{K}_s)$. Assume H is ∇_0 -critical, otherwise we continue with a suitable subgraph of H . We assume without loss of generality that the embedding has depth exactly r .

For a vertex $v \in G$, let v_1, \dots, v_s denote its copies in $G \bullet \bar{K}_s$ and let \equiv_s denote the copy-relation over $V(G \bullet \bar{K}_s)$. Let us define the sets $V_i := \{v_i\}_{v \in G}$ for convenience. Our goal will be to construct an r -shallow minor H' in $G \bullet \bar{K}_s$ whose embedding intersects every equivalence class at most once—such a minor will obviously also be a minor of G . Since H is a stable topological minor, every edge $uv \in H$ is embedded as an induced path $\phi_E(uv)$. It follows that $\phi_E(uv)$ traverses every equivalence class under \equiv_s at most once, otherwise it would contain a chord.

First, let us make sure that no two nails of the minor are mapped to the same copy-class. We invoke Corollary 2 on the graph H and the partition defined by the equivalence relation

$$u \equiv_s^H v \iff \phi_V(u) \equiv_s \phi_V(v).$$

As a result, we obtain an induced subgraph $H' \subseteq_i H$ whose vertices are pairwise not equivalent under \equiv_s^H (therefore their images under ϕ_V are not equivalent under \equiv_s) and whose density is at least

$$\nabla_0(H') \geq \nabla_0(H)/4s^2.$$

Due to the symmetric nature of $G \bullet \bar{K}_s$ we can now assume that ϕ_V maps the vertices of H' into the set V_1 .

Note that for $r = 0$, *i.e.* H is a subgraph of $G \bullet \bar{K}_s$, we are done since then $H' \subseteq (G \bullet \bar{K}_s)[V_1] \simeq G$. We assume in the following that $r \geq 1/2$. Let us define the conflict-graph \vec{G}_{conf} with vertex set $E(H')$ and an arc $e_1 e_2$ if the paths $\phi_E(e_1)$ and $\phi_E(e_2)$ traverse a common copy-class $\{v_1, \dots, v_s\} \subseteq G \bullet \bar{K}_s$ and the vertex used by $\phi_E(e_1)$ has a higher index than the vertex used by $\phi_E(e_2)$. Since the embedding of H' has depth r , every embedding of an edge of H' traverses at most $2r + 2$ classes. Using our assumption that ϕ_V embeds only into V_1 , we conclude that

$$\Delta^-(\vec{G}_{\text{conf}}) \leq (2r + 2)(s - 1)$$

which implies that G_{conf} is $4(r + 1)(s - 1)$ -degenerate (*cf.* Lemma 9). By a simple colouring argument, the conflict graph therefore contains an independent set of size at least

$$\frac{|G_{\text{conf}}|}{(4(r + 1)(s - 1) + 1)} = \frac{\|H'\|}{(4(r + 1)(s - 1) + 1)}.$$

Retaining only the edges of H' corresponding to such an independent set leaves us with a graph $H'' \subseteq H'$ whose edges are pairwise conflict-free. Accordingly, we can modify the embedding ϕ_E to exclusively map into V_1 . Finally, H'' is embedded into $(G \bullet K_s)[V_1] \simeq G$ and hence it is a topological minor of G .

Since all the above operations where deletions of vertices and edges of H and the modification of ϕ_V and ϕ_E only exchanged vertices inside copy-classes, we conclude that H'' is an r -stable and γ -weighted topological minor of G . The density of H'' is at least

$$\frac{\|H''\|}{|H''|} \geq \frac{\nabla_0(H')}{(4(r+1)(s-1)+1)} \geq \frac{\nabla_0(H)}{(4(r+1)(s-1)+1)s^2},$$

from which the claimed bound follows directly. \square

AUGMENTATIONS & DECOMPOSITION BY COLOURING

There is no indication that a Robertson-Seymour style decomposition for graphs of bounded expansion exists and I personally feel pessimistic about the prospect of finding one. The parametrisation of minors by their depths introduces a bias in the forbidden-structure characterisation that put less restrictions on the global scale of the graph and tightens towards the local scale. A powerful global decomposition would need stark restrictions on the global scale (at least, that is my intuition) which seems irreconcilable with bounded expansion and nowhere dense classes.

All is not lost, however: Nešetřil and Ossona de Mendez provided a decomposition theorem that works locally and—unsurprisingly—it is parametrised by something like a depth-parameter. We devote this section to provide an alternative proof for the following theorem, tightening some bounds and developing a machinery which, as demonstrated later, has practical applications. We will in particular be able to apply the stable minors introduced in the previous chapter.

Theorem 15 (Nešetřil, Ossona de Mendez [189, 187]). *Let \mathcal{G} be a graph class of bounded expansion. There exists a function f such that for every integer r , every graph in \mathcal{G} can be coloured with at most $f(r)$ colours such that every set of vertices on at most $k < r - 1$ colours induces a graph of treedepth at most k .*

p-centred colouring, χ_p

As we will see, some result (in particular algorithmic ones) are quite easy to obtain using the above *low-treedepth colourings* but seem absolutely unprovable using minors. These type of colourings generalises the star-colouring number introduced by Fertin, Raspaud, and Reed [94]. In that context, they are usually called *p-centred colouring* with the following equivalent definition: a vertex-colouring c of a graph G is *p-centred* if every connected subgraph $H \subseteq G$ either receives at least p colours by c or there exists (at least) one colour in $c(H)$ that appears exactly once in H . In the latter case, we call the vertex carrying this colour the *centre* of the subgraph. We denote by $\chi_p(G)$ the minimum number of colours necessary in a *p-centred* colouring of G . Note that $\chi_2(G)$ is exactly the chromatic number (every edge receives at least two colours) and $\chi_3(G)$ the star-chromatic number of G (every two colour classes induce a forest of stars).

In this chapter, we will provide an alternative proof for Theorem 15. The reason for re-proving it are twofold: First, the machinery used here was designed to be applied in practice and differs from the one in the literature. It is therefore necessary to re-prove the above statement

for this machinery in order to apply it with good conscience. Second, I feel that an alternative proof might open up further improvements in the future. Additionally, we show that if the bound on the treedepth does not need to match the number of colour classes, that is, selecting k colour classes results in a subgraph of treedepth at most $h(k)$ for some function h , we actually need less colours. This trade-off make low-treedepth colourings more flexible in practice, since we can balance the number of colours against the treedepth induced by few colours.

tf-augmentation, transitive, fraternal

Let \vec{G} be a directed graph. A *1-transitive fraternal augmentation* is a directed graph \vec{H} with vertex set $V(\vec{G})$ and $\vec{E}(\vec{G}) \subseteq E(\vec{H})$ with the following additional properties:

- If $uv, vw \in E(\vec{G})$, then $uw \in E(\vec{H})$ (transitive arcs)
- If $uw, vw \in E(\vec{G})$, then at least one of uv, vu is in $E(\vec{H})$ (fraternal arcs)

Sequential, fraternal

Such an augmentation \vec{H} is *tight* if every proper subgraph of it violates one of the above two properties. In the following we will silently assume that augmentation are tight and thus drop this specifier for brevity's sake. Vertices u, w that are connect by arcs uv, vw with each other will be called *sequential* in the following¹ and vertices u, w connected by arcs uv, vw *fraternal*.

We are interested in the repeated application of these augmentations. Consider a sequence of directed graphs $\vec{G}_1, \vec{G}_2, \dots$ where \vec{G}_i is a 1-transitive fraternal augmentation of \vec{G}_{i-1} and \vec{G}_1 is any orientation of G (we will later impose more restrictions on \vec{G}_1 in order to ensure additional properties of the augmentations). We call such a sequence simply a *transitive-fraternal augmentation* (tf-augmentation) of G .

dtf-augmentation

Let us introduce the following variant of tf-augmentations: let G be an undirected graph and let \vec{G}_1 be any orientation of G . Then a *distance-constrained transitive-fraternal augmentation* (dtf-augmentation) of G is a sequence $\vec{G}_1, \vec{G}_2, \dots$ of directed graphs which satisfy:

1. $\vec{G}_i \subseteq \vec{G}_{i+1}, i \geq 1$;
2. for arcs $uv \in \vec{G}_i, vw \in \vec{G}_j$ it follows that $vw \in \vec{G}_{i+j}$; and
3. for arcs $uv \in \vec{G}_i, wv \in \vec{G}_j$ it follows that either $vw \in \vec{G}_{i+j}$ or $wv \in \vec{G}_{i+j}$.

To simplify notation we associate a weight function $\omega_i: \vec{E}(\vec{G}_i) \rightarrow [i]$ with the i^{th} dtf-augmentation \vec{G}_i as follows:

$$\omega_i(uv) = \begin{cases} \min\{\omega_{i-1}(uv), i\} & \text{if } uv \in \vec{G}_i \\ \infty & \text{else} \end{cases}$$

In other words: if the arc uv is present in \vec{G}_i but not in \vec{G}_{i-1} , then we have $\omega_{\geq i}(uv) = i$ and $\omega_{< i}(uv) = \infty$.

¹ The term 'transitive' applied to such vertices simply does not sound right.

First, let us prove that not only are only arcs of weight d added in the dtf-step from \vec{G}_{d-1} to \vec{G}_d , but this is exactly the step in which all arcs of weight d appear.

Lemma 25. *For any integers d, f with $d \leq f$ it holds that every arc $uv \in \vec{G}_f$ with $\omega_f(uv) = d$ is already contained in \vec{G}_d with $\omega_d(uv) = d$.*

Proof. The base case $d = 1$ follows immediately: no arc of weight 1 can be added by a dtf-augmentation, hence they all are already present in the orientation \vec{G}_1 of G .

Now choose the smallest $f > 1$ such that there exists an arc $uv \in \vec{G}_f$ with $\omega_f(uv) = d < f$ (the case $f = d$ is trivial) such that $uv \notin \vec{G}_d$. By the choice of f , this implies that $uv \notin \vec{G}_{f-1}$. Since uv is added by a dtf-augmentation from \vec{G}_{f-1} to \vec{G}_f , the vertices u and v are either sequential or fraternal in \vec{G}_{f-1} . In either case, there exist two arcs e_1, e_2 of weights d_1 and d_2 in \vec{G}_{f-1} with $d_1 + d_2 = d$ who are responsible for the addition of uv . But by the choice of f , e_1 and e_2 must be present in \vec{G}_{d_1} and \vec{G}_{d_2} , respectively. Since $d_1, d_2 < d$ we have a contradiction. \square

The following property holds also for transitive-fraternal augmentations [189, 192]. In a way, tf-augmentations behave very similar to graph powers: after d iterations, we have that all vertices at distance at most d are either connected by an arc or they share a common in-neighbour (in any case: they have distance at most two to each other). This property lies at the heart of the low-treedepth colouring, though we need to put in a lot more work to prove it. Let us first see that dtf-augmentations behave the same as their non-constrained counterpart.

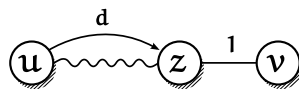
Lemma 26. *For every vertices $u, v \in G$ with $\text{dist}(u, v) = d$ and for every integer $r \geq d$ one of the following holds:*

1. $uv \in \vec{G}_r$ and $\omega_r(uv) = d$,
2. $vu \in \vec{G}_r$ and $\omega_r(vu) = d$,
3. or there exists x such that $xu, xv \in \vec{G}_r$ and $\omega_r(xu) + \omega_r(xv) = d$.

Proof. First note that it suffices to prove the statement for $r = d$. We prove it by induction over d . For $d = 1$ either Case 1 or 2 will hold since \vec{G}_1 is simply an orientation of G_0 with arc weights set to 1. Thus assume that the statement holds for vertices at distance at most d and consider u, v with $\text{dist}(u, v) = d + 1$.

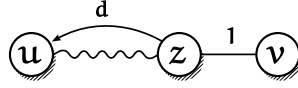
Chose some vertex z in $N(v)$ that has distance d to u , then for u, z the induction hypothesis holds. Note that if u and v are either fraternal or sequential in \vec{G}_d we are done: then \vec{G}_{d+1} must contain either the arc uv or vu with weight $d + 1$. Let us distinguish the following cases:

Case 1: $uz \in \vec{G}_d$ with $\omega_d(uz) = d$.



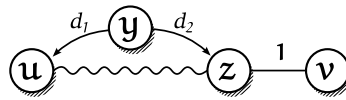
If $zv \in \vec{G}_1$ then u and v are sequential in \vec{G}_d . Otherwise, $vz \in \vec{G}_1$ and u and v are fraternal in \vec{G}_d .

Case 2: $zu \in \vec{G}_d$ with $\omega_r(zu) = d$.



If $zv \in \vec{G}_1$ then the vertex z serves as the common in-neighbour of v and u in \vec{G}_d and we have that $\omega_r(zu) + \omega_r(uv) = d + 1$. If $vz \in \vec{G}_1$ then v and u are sequential in \vec{G}_d .

Case 3: There exists y such that $yu, yz \in \vec{G}_d$ with $d_1 = \omega_d(yu)$ and $d_2 = \omega_d(yz)$ such that $d_1 + d_2 = d$.



If $zv \in G_0$ we are done: then y and v are sequential in \vec{G}_d , hence y is a common in-neighbour of u and v in \vec{G}_{d+1} . Thus assume in the following that $vz \in G_0$. Now y and v are fraternal in \vec{G}_d , if the arc yv is present in \vec{G}_{d+1} we are again done: the remaining case is that $vy \in \vec{G}_d$.

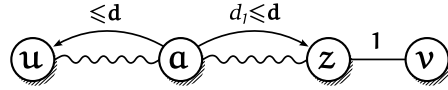
Note that in this remaining case, we have $\omega_d(vy) = d_2 + 1 \leq d$ since $d_1, d_2 \geq 1$ and $d_1 + d_2 = d$. By Lemma 25, the arc vy therefore is already present in \vec{G}_d . Thus v and u are sequential in \vec{G}_d . \square

Out-apex

Before we consider how general subgraphs behave under dtf-augmentations, let us first limit our attention to paths. We will use the term *out-apex* in the following, by which we mean a vertex that contains all other vertices (usually in the context of some subgraph) in its out-neighbourhood.

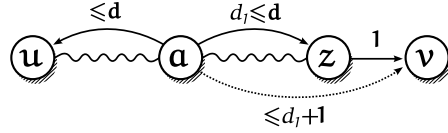
Lemma 27. *Let $\vec{G}_1, \vec{G}_2, \dots$ be dtf-augmentations of G . Let P be a path of length d in G_1 . Then $\vec{G}_d[P]$ contains an out-apex $a \in P$ such that every outgoing arc ax for $x \in P$ has weight $\omega_d(ax) \leq |P[a, x]|$.*

Proof. Consider any such path P with endpoints u and v . The statement is trivially true for $|P| = 2$, hence assume $|P| \geq 3$. Assume the predecessor of v on P is the vertex $z \neq u$. We assume that the statement holds by induction for paths of length d , hence consider the case that $|P| = d + 1$. Since then the subpath $P[u, z]$ has length d , it contains an out-apex in G_d by the induction hypothesis. Let us call this apex a ; note however that a might be identical to u or z . Let $d_1 := \omega_d(a, z)$ with the special case $d_1 = 0$ if $a = z$. By the induction hypothesis, it holds that $d_1 = |P[a, z]|$ and hence $d_1 + 1 = |P[a, v]|$. The following figure depicts the situation in \vec{G}_d .



Let us consider the possible orientations of the edge zv .

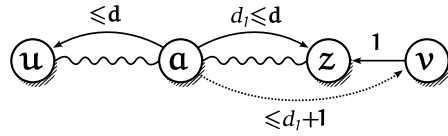
Case 1: $zv \in G_0$.



Then a and v are sequential (via z) and hence the transitive edge av exists in \vec{G}_d with weight $d_i + 1$. It follows that a is an out-apex of P in \vec{G}_{d+1} as claimed.

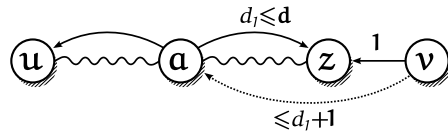
In the remaining cases we assume that $vz \in \vec{G}_0$, hence a and v are fraternal in \vec{G}_d . We distinguish the two possible cases of which arc between a and v is contained in \vec{G}_{d+1} .

Case 2: $vz \in G_0$ and $av \in G_{d+1}$.



Again a is an out-apex of P in \vec{G}_{d+1} with the desired properties.

Case 3: $vz \in G_0$ and $va \in G_{d+1}$.



Note that now in \vec{G}_{d+1} , all vertices in $P \setminus \{v, a\}$ are sequential with v . Hence for vertices in $P[u, a]$ and for the vertex z we already have that v is an out-apex in \vec{G}_{d+1} with the correct arc-weights. For vertices in $P(a, z)$ we have that v is an out-apex in \vec{G}_{d+1} , but it is left to show that the arc-weights are correct. To that end, consider any vertex $x \in P(a, z)$ and let $d_x = |P[d_x, v]|$ be the distance of x to v on P . By Lemma 26 we know that already in \vec{G}_{d_x} , either the arc xv with weight d_x or the arc vx with weight d_x exists; or x and v share a common in-neighbour y such that the sum of the two weights of yx and yv is exactly d_x . If the arc vx exists v we are done. Otherwise, assume the arc xv exists. But then a is sequential with v in $\vec{G}_{d_x} \subset \vec{G}_{d_1}$, hence the arc av would exist in \vec{G}_{d_1+1} , contradicting our assumption. Lastly, assume x and v share a common in-neighbour: since such an

in-neighbour must be contained in $P(a, z)$, again a and v would be sequential in \vec{G}_d and hence the arc av would exist in \vec{G}_{d+1} . We conclude that the only possibility is that the first one: the arc vx exists with weight at most d_x . Therefore we have shown that v is an out-apex for P in \vec{G}_{d+1} . \square

In fact, one can easily prove the following more general statement with the exact same approach:

Lemma 28. *Let $\vec{G}_1, \vec{G}_2, \dots$ be dtf-augmentations of G . Let P be a path of weighted length d in G_ℓ . Then $\vec{G}_d[P]$ contains an out-apex $a \in P$ such that every outgoing arc ax for $x \in P$ has weight $\omega_d(ax) \leq \sum_{e \in \vec{E}(P[a, x])} \omega_\ell(e)$.*

We can take this idea in a slightly different direction by asking in what cases a vertex will be an out-apex of some path: this is essentially determined by whether there exists an arc that spans the whole length of the path. However, this only holds true if the path in question contains the vertex that was ‘responsible’ for the addition of said arc.

Lemma 29. *Let $\vec{G}_1, \vec{G}_2, \dots$ be dtf-augmentations of G and let there be vertices u, v connected by a path of length d in G . Then if $uv \in \vec{G}_d$, then there exists a u - v -path P of length d such that either u or v is an out-apex of $\vec{G}_d[P]$ whose arc-weights correspond to the distances on P .*

Proof. The statement is true for $d = 2$: if the arc uv exists in \vec{G}_2 , then u, v are either transitive or fraternal in \vec{G}_1 via some vertex x . In either case, u is an out-apex of the path uxv .

Fix $d > 2$ and assume by induction that the statement holds for all smaller values for d . First consider a vertex x such that u, v are sequential via x in \vec{G}_{d-1} , i.e. x is the responsible for the addition of the arc uv in \vec{G}_d . Let d_1 be the arc-weight of ux and let d_2 be the arc-weight of xv . By induction, there exist paths P_{ux} from u to x and P_{xv} from x to v of respective length d_1, d_2 where P_{ux} has u and P_{xv} has x as an out-apex. Accordingly, u is sequential with every vertex of $P_{xv} \setminus \{x\}$ and thus is an out-apex of $P_{ux}P_{xv}$ in \vec{G}_d .

Now consider the case where u, v are fraternal via x . Again we have paths P_{ux}, P_{xv} ; only this time u is an out-apex of P_{ux} and v is an out-apex of P_{xv} . By Lemma 27, the path $P_{ux}P_{xv}$ contains an out-apex a in \vec{G}_d . If a is either u or v we are done. Otherwise, a lies either in P_{ux} or P_{xv} (or on both, in case $a = x$). If a lies on P_{ux} , then u is sequential via a with every vertex in P_{xv} and hence is an out-apex for $P_{ux}P_{xv}$. If a lies on P_{xv} , the same holds for v .

Since in both cases we have that the out-apex $a \in \{u, v\}$ is created via transitive arcs, using induction we see that the arc-weights correctly reflect the distances along $P_{ux}P_{xv}$. \square

Lemma 29 has another interesting corollary; namely, that once no arc is added during a step of the augmentation process, then each following step will also not add further arcs.

Corollary 4. *If $\vec{G}_{d+1} = \vec{G}_d$, then $\vec{G}_{\geq d} = \vec{G}_d$.*

Proof. Assume to the contrary that $\vec{G}_{d+1} = \vec{G}_d$ but $\vec{G}_{d+2} \neq \vec{G}_{d+1}$, i.e. \vec{G}_{d+2} contains an arc uv of weight $d + 2$. Applying Lemma 29, we have that uv are connected by a path of length $d + 2$ in G and u is an out-apex of that path. But then there is an arc of weight $d + 1$ emanating from u , contradiction. \square

A simple corollary to Lemma 27 is that every path inside the original graph will—after sufficiently many augmentations—contain a clique whose size is related to the length of said path.

Corollary 5. *Every path of length 2^d in G contains an acyclic orientation of the complete graph K_d in \vec{G}_{2^d} .*

This already suffices to show that a property colouring of \vec{G}_{2^d} is some kind of low-treewidth colouring—it does, however, not have a nice bound on the treewidth as Theorem 15 has.

Corollary 6. *There exists a function f such that every proper colouring of \vec{G}_{2^d} , every set of at most $d - 1$ colours-classes induces a subgraph of tree-depth at most $f(d)$.*

Proof. By Corollary 5, every path of length d in G is augmented to a complete graph in G_{2^d} and therefore receives d colours in any proper colouring of G_{2^d} . Therefore any collection of up to $d - 1$ colour classes must induce a subgraph of G that does not contain any path longer than 2^d . As mentioned in the proof of Proposition 4, any dfs-tree is a treewidth decomposition. Therefore the treewidth of any subgraph induced by up to $d - 1$ colour classes is bounded by 2^{2^d} . \square

To obtain a better bound on the treewidth, we need to see how arbitrary connected subgraphs behave under augmentation. As a sub-case, we will need to bound the number of augmentations necessary to obtain an out-apex inside a complete subgraph.

Lemma 30. *Let $\vec{G}_1, \vec{G}_2, \dots$ be dtf-augmentations of G and let $\vec{H} \subseteq \vec{G}_d$ be an orientation of K_h where $h = |H|$. Then $\vec{G}_{d \lceil \log h \rceil}[H]$ contains an out-apex.*

Proof. The statement is trivially true for $h = 2$ so we apply the argument inductively. Partition the graph H into two subgraphs H_1, H_2 of size $\lceil h/2 \rceil$ and $\lfloor h/2 \rfloor$, respectively. By the induction hypothesis, we assume that $\vec{G}_{f(d, \lceil h/2 \rceil)}[H_1]$ and $\vec{G}_{f(d, \lfloor h/2 \rfloor)}[H_2]$ each contain an out-apex. Since these two apices are connected by an arc in \vec{G}_d , one of them is sequential to all vertices in the respective other half of H . Hence that apex is an apex for *all* of H in $\vec{G}_{f(d, \lceil h/2 \rceil)+d}$. We obtain the claim by resolving the recurrence with the boundary condition that $f(d, 2) = d$. \square

Armed with the above lemma, we can now tackle arbitrary subgraphs of low treewidth. The following proof largely follows the corresponding proof for tf-augmentations by Dvořák and Král [80]. Our version

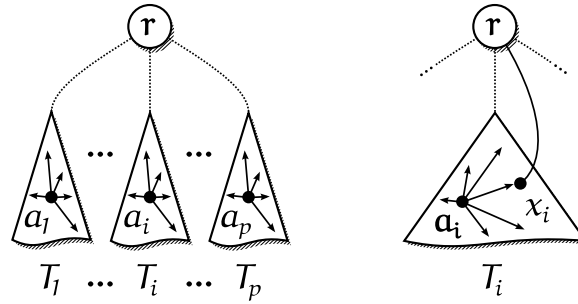
employs Lemma 30 and the arc-weights provided by the dtf-augmentations to improve the final bound².

Lemma 31. *Let $H \subseteq G$ be a subgraph with $\mathbf{td}(H) = t$. Then the induced orientation $\vec{H}_{(2\log t)^t} := \vec{G}_{(2\log t)^t}[V(H)]$ either contains an out-apex $a \in H$ or $\vec{H}_{(2\log t)^t}$ contains some orientation of K_t as a subgraph.*

Proof. It suffices to show this for connected graphs. Note that for $t = 1$ the statement is trivially true.

Let $t + 1 > 1$ in the following and fix any nice treedepth decomposition T of H . Let $r \in H$ be the root of that decomposition with subtrees T_1, \dots, T_p . Since $\mathbf{td}(H[V(T_i)]) \leq t$ for $1 \leq i \leq p$, we can assume by induction that every subgraph $\vec{H}_{f(t)}[V(T_i)]$ contains an out-apex a_i or an orientation of K_{t+1} . Since in the latter case we are done, assume that for every such subgraph we have an out-apex a_i . Note that every arc outgoing from a_i into its respective subgraph has necessarily weight at most $f(t)$.

Fix any $1 \leq i \leq p$ and consider the graph $H^i = H[r \cup V(T_i)]$. Since T is a nice treedepth-decomposition, there exists a vertex $x_i \in T_i$ that is connected to r in H . We consider the following cases.



Case 1: $x_i = a_i$. Then either the arc $a_i r$ or the arc $r a_i$ exists in \vec{H}_1 . In the first case, a_i is already an out-apex of H^i in $\vec{H}_{f(t)}$. Otherwise, the arc $r a_i$ exists in $\vec{H}_{f(t)}$ and hence r is sequential with every vertex in T_i . Thus already in $\vec{H}_{f(t)+1}$, the root r is an out-apex of H^i .

Case 2: $x_i \neq a_i$ and $r x_i \in \vec{H}_1$. Since a_i is an out-apex in $\vec{H}_{f(t)}$, the arc $a_i x_i$ exists. Hence a_i and r are sequential in $\vec{H}_{f(t)}$. It follows that a_i is an out-apex of H^i with the desired properties in $\vec{H}_{f(t)+1}$.

Because the first two cases will give us the desired outcome for *any* vertex in T_i connected to r we assume in the last case that the above two cases do not hold for any choice of x_i .

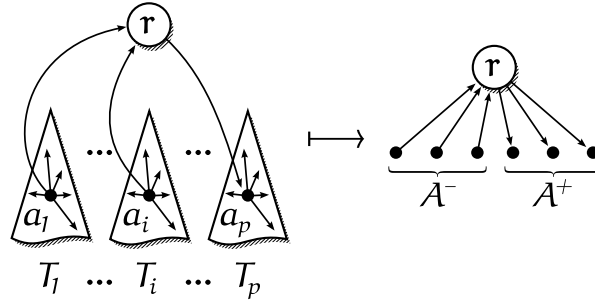
Case 3: For all $x \in N(r) \cap V(T_i)$ it holds that $x \neq a_i$ and $x r \in \vec{H}_1$. Any such vertex x will make a_i and r fraternal in $\vec{H}_{f(t)}$, thus either the arc $a_i r$ or $r a_i$ will be present in $\vec{H}_{f(t)+1}$. The arc $a_i r$ again makes a_i

² Translating their result from tf-augmentations to dtf-augmentations, their bound states that around 2^{t^2} augmentations are sufficient.

the out-apex in $\vec{H}_{f(t)+1}$, therefore assume that $a_i r \notin \vec{H}_{f(t)+1}$ and $ra_i \in \vec{H}_{f(t)+1}$. This already implies that r is an out-apex for H^i in $\vec{H}_{2f(t)+1}$

In the above cases, we have now proved that any H^i , $1 \leq i \leq p$, has either a_i or r as an out-apex in $\vec{H}_{2f(t)+1}$. Note that in the case that a_i is the out-apex, the arc $a_i r$ is already present in $\vec{H}_{f(t)+1}$. Using this fact, we now prove the statement for the whole graph H .

Let $A^- := \{a_j \mid a_j r \in \vec{H}_{2f(t)+1}\}_{1 \leq j \leq p}$ be those apices of subtrees that have arcs towards r and let $A^+ := \{a_j \mid a_j \notin A^-\}_{1 \leq j \leq p}$ be the remaining ones.



If $A^- = \emptyset$ we are done with r as the apex of H . Note that all vertices of A^- are pairwise fraternal already in $\vec{H}_{f(t)+1}$, therefore the digraph $\vec{H}_{2f(t)+2}[A^-]$ is an orientation of $K_{|A^-|}$. Hence if $|A^-| \geq t$ we are done as well.

Assume therefore that $1 \leq |A^-| < t$. Since $\vec{H}_{2f(t)+2}[A^-]$ is an orientation of a complete subgraph, we can apply Lemma 30 and obtain that in augmentation number $(2f(t) + 2)\lfloor \log(t-1) \rfloor$ there exists an out-apex $a \in A^-$. At that point the vertex a is sequential with all vertices in subtrees containing an apex of A^- . Since every apex in A^- has out-arcs with weight at most $f(t)$ into its respective subtree, a is an apex for subtrees with apices in A^- in augmentation number

$$(2f(t) + 2)\lfloor \log(t-1) \rfloor + f(t).$$

For subtrees with apices in A^+ we have that the vertex r is an apex after $2f(t) + 1$ augmentations, and after at most $f(t) + 1$ augmentations the arc ar appears. Hence, the vertex a is an out-apex of H after at most

$$\begin{aligned} f(t+1) &\geq \min\{(2f(t) + 2)\lfloor \log(t-1) \rfloor + f(t), \quad 3f(t) + 2\} \\ &= (2f(t) + 2)\lfloor \log(t-1) \rfloor + f(t) \\ &= f(t)(2\lfloor \log(t-1) \rfloor + 1) + 2\lfloor \log(t-1) \rfloor \end{aligned}$$

augmentations.

The function $f(t+1) = (2\log t)^t$ satisfies the resulting recurrence and hence the claim follow. \square

With the help of Lemma 31 we can now guarantee a better bound on the treedepth of any selection of few colour classes.

Lemma 32. *For any $p \in \mathbb{N}$, any proper vertex-colouring of $G_{(2 \log p)^p}$ is a p -centred colouring of G . Hence any $i < p$ colour classes will induce a graph of treedepth at most i in G .*

Proof. Let $c: V(G) \rightarrow [q]$ be a proper colouring of $G_{(2 \log p)^p}$. Consider any connected subgraph $H \subseteq G$ with treedepth $\geq p$.

By Lemma 31, $\vec{H}_{(2 \log p)^p} = \vec{G}_{(2 \log p)^p}[V(H)]$ either contains an out-apex or some orientation of K_p . But since $\mathbf{td}(H) \geq p$, we actually obtain K_p as a subgraph in both cases: removing a vertex that is an apex in $\vec{H}_{(2 \log p)^p}$ from H results in a graph of treedepth at least $p - 1$. Iterating this procedure will give us an orientation of K_p . We conclude that H receives at least p colours by c .

Now consider any connected subgraph H with treedepth less than p . Again by Lemma 31, the subgraph $\vec{H}_{(2 \log p)^p}$ will contain an out-apex or K_p . In the latter case we are done, since H receives at least p colours by c . Otherwise, the out-apex $a \in \vec{H}_{(2 \log p)^p}$ will receive a unique colour by c , hence a is a centre of H . Applying the same argument to the connected components of $H - a$ yields that c is a centred colouring for H and the claim follows. \square

By Lemma 32 we now have the number of transitive- fraternal augmentations needed to obtain a p -centred colouring. In order to obtain a characterisation of bounded expansion classes, we need to bound the number of colours necessary. To this end, we will show that the indegree of dtf-augmentations is upper-bounded by the density of stable topological minors. We roughly follow the proofs which were used to relate the indegree of tf-augmentations to shallow minors [189, 192]. However, we refrain from switching between shallow minors and shallow topological minors in order to keep the bounds as low as possible.

In the following we now make use of stable and weighted topological minors. The weight function is simply the weight introduced through the dtf-augmentation; by Lemma 26 this weight is simply the distance between its endpoints in the original graph.

Lemma 33. *Let $\vec{G}_1, \vec{G}_2, \dots$ be a dtf-augmentation. Then for every $d \geq 2$ it is true that*

$$\tilde{\nabla}_{r/2}^{\gamma}(G_d) \leq (40r' \lfloor \gamma/d \rfloor \Delta^-(\vec{G}_{d-1}))^6 \cdot \tilde{\nabla}_{r'/2}^{\gamma}(G_{d-1})^2,$$

where $r' = r + 2 + \lfloor \gamma/d \rfloor$.

Proof. Let ϕ_V, ϕ_E be the topological embedding of H with depth $r/2$ that further is γ -weighted, and ℓ -stable for some $\ell \leq r + 1$. We further assume that the minor is ∇_0 -critical and thus by Lemma 14 we have that $\delta(H) \geq \nabla_0(H)$.

We will modify the minor and embedding in several steps to arrive at a minor H' of G_{d-1} with the desired properties; to that end we will create a sequence of minors H^i and embeddings ϕ_E^i (the em-

beddings ϕ_V are clear from the context and we simplify notation by assuming that $V(H^i) \subseteq V(G_{d-1})$. We begin with H^0 equal to H :

$$V(H_0) := V(H)$$

$$E(H_0) := E(H)$$

Define $E_d \subseteq E(G_d)$ as the set of edges with weight d in G_d , i.e. let $E_d := \{e \in E(G_d) \mid \omega_d(e) = d\}$; let further E_d^i be the subset of edges of H_i whose image under ϕ_E^i contains an edge of E_d . Precisely, let $E_d^i := \{e \in H_i \mid E(\phi_E^i(e)) \cap E_d \neq \emptyset\}$ for all future $H_i, i \geq 0$.

Every edge in E_d corresponds to a fraternal or transitive arc in \vec{G}_{d-1} , as such for every edge $uv \in E_d$ there exists a vertex $w \in G$ such that uw are fraternal or sequential via w . Let us call w the *pivot* of the edge uv and let $\mathfrak{P} \subseteq V(G)$ be a set of pivot-vertices of edges in E_d . In the case that an edge has multiple pivots, we only require at least one pivot to be in \mathfrak{P} . Let $\pi: E_d \rightarrow \mathfrak{P}$ be a mapping assigning each edge in E_d to its pivot in \mathfrak{P} . For brevity's sake, we extend π to a function over $E(G_{d-1})$ by setting $\pi(e) = \perp$ for $e \in E(G_{d-1}) \setminus E_d$ and extend it to edge sets via $\pi(E) := \{\pi(e)\}_{e \in E}$.

In general, the vertices \mathfrak{P} might appear as both nails and subdivision vertices of the embedding ϕ_E^0 and make the subsequent operations on H^0 rather tedious. We circumvent this problem by considering the graph $G_{d-1} \bullet \bar{K}_2$, i.e. the graph obtained from G_{d-1} by creating a false twin for every vertex. By Proposition 7, we have for $r' \geq 2$ that

$$\tilde{\nabla}_{r'}(G_{d-1} \bullet \bar{K}_2) \leq (2r' + 1) \cdot \tilde{\nabla}_{r'}(G_{d-1}) + 1.$$

Hence we can construct a minor of $G_{d-1} \bullet \bar{K}_2$ and argue that a minor of similar density must exist in G_{d-1} . Let us label the original vertices of G_{d-1} in the product graph by V and the twin-vertices by V' . Now we can consider the twin-vertices of \mathfrak{P} during the construction which are disjoint from the embedding of the minor inside V . To avoid cumbersome notation, we simply assume in the following that $\phi_V^0(H)$ and $\phi_E^0(E(H))$ are disjoint from \mathfrak{P} and will pay for this convenience with the above loss in density at the very end.

Before we proceed to the next step, let us modify the minor in such a way that each pivot is only responsible to add edges to a single nail in the model of the minor. This will obviously remove some edges of E_d and hence from the minor, but by the following claim we can preserve a high-enough fraction of vertices. For convenience, we define the function $\rho: V(H^0)^2 \rightarrow \mathfrak{P}$ as follows: for an edge $uv \in H^0$, let ux be the first edge in the path $\phi_E^0(uv)$ (i.e. the one incident to u). If $ux \notin E_d$ or the arc $u\mu(ux)$ does not exist in \vec{G}_d we set $\rho(uv) = \perp$. Otherwise, we let $\rho(uv) = \pi(ux)$. Note that ρ is not symmetrical, in general we have that $\rho(uv) \neq \rho(vu)$.

Since some pivots might not add any edges incident to a nail, we partition the pivots as follows:

$$\begin{aligned}\mathfrak{P}^h &:= \{w \in \mathfrak{P} \mid \exists uv \in H^0 : \rho(uv) = w\} \\ \mathfrak{P}^l &:= \mathfrak{P} \setminus \mathfrak{P}^h\end{aligned}$$

Claim. *There exists a mapping $\mu: \mathfrak{P}^h \rightarrow V(H)$ such that $\mu(w) \in N_{d-1}^-(w)$ for $w \in \mathfrak{P}^h$ and the minor H^1 obtained from H^0 by removing all edges $uv \in H^0$ where $u \neq \mu(\rho(ux))$ satisfies $\|H^1\| \geq \|H^0\|/4\Delta^-(\vec{G}_{d-1})^2$.*

Consider the auxiliary graph \mathcal{H} with

$$\begin{aligned}V(\mathcal{H}) &:= \bigcup_{w \in \mathfrak{P}^h} \{(w, v) \mid v \in N_{d-1}^-(w) \cap V(H)\} \\ (w, v)(w', u) \in E(\mathcal{H}) &\iff uv \in H \text{ and } w = \rho(vu), w' = \rho(uv)\end{aligned}$$

The partition $\{V_w\}_{w \in \mathfrak{P}}$ defined via $V_w := \{v \in V(H) \mid (w, v) \in V(\mathcal{H})\}$ has partite sets of size at most $\Delta^-(\vec{G}_{d-1})$ that induce independent sets in \mathcal{H} . Hence we can apply Corollary 2 and obtain a subset $\mathcal{X} \subseteq V(\mathcal{H})$ such that

$$\|\mathcal{H}[\mathcal{X}]\| \geq \|\mathcal{H}\|/4\Delta^-(\vec{G}_{d-1})^2.$$

Note that for every $w \in \mathfrak{P}^h$ there exists per construction exactly one vertex $v \in V(H)$ such that $(w, v) \in \mathcal{X}$. The assignment μ therefore simply assigns this v to w . Therefore any edge in $\mathcal{H}[\mathcal{X}]$ corresponds to an edge *not* deleted in the construction of H^1 and we have that indeed $\|H^1\| \geq \|\mathcal{H}[\mathcal{X}]\|$ (note that some edges of H^0 might not be affected at all, this is reflected by the inequality), as claimed.

Let us assume in the following that H^1 is ∇_0 -critical. This property can be enforced simply by removing nails of degree less than $\nabla_0(H^1)$ and does not interfere with the remaining construction. We therefore assume that H^1 already has this property.

We continue by partitioning the vertices of H^1 into two sets V^h, V^l as follows:

$$\begin{aligned}V^h &:= \{v \in H^1 : |\mu^{-1}(v)| \geq \nabla_0(H)^{1/2}\} \\ V^l &:= V(H^1) \setminus V^h\end{aligned}$$

We construct the minor H^2 from H^1 as follows: for every vertex $v \in V^l$ we add all pivots $\mu^{-1}(v)$ to the nail set. We include the edges from v to $\mu^{-1}(v)$ and reroute the edges $vu \in H^2$ with $\pi(vu) \neq \perp$ to wu where $w = \pi(vu) \in \mu^{-1}(v)$. For vertices in $v \in V^h$ we take every pivot $w \in \mu^{-1}(v)$ and of all the edge incident to v in H^2 with pivot w , we keep an arbitrary single one. We lower-bound the density of the resulting minor by counting the degrees of vertices and obtain:

$$\begin{aligned}2 \frac{\|H^2\|}{|H^2|} &\geq \frac{|V^h| \nabla_0(H)^{1/2} + |V^l| \nabla_0(H^1)}{|H^1| + |V^l| \nabla_0(H)^{1/2}} \\ &\geq \frac{|V^h| \nabla_0(H)^{1/2} + |V^l| \nabla_0(H^0)/4\Delta^-(\vec{G}_{d-1})^2}{|H^0| + |V^l| \nabla_0(H)^{1/2}}\end{aligned}$$

This fraction takes its extreme values in the cases $|V^h| = |V(H^0)|$ and $|V^l| = |V(H^0)|$. With these values we obtain the densities

$$\frac{1}{2}\nabla_0(H)^{1/2} \quad \text{and} \quad \frac{1}{8\Delta^-(\vec{G}_{d-1})^2}\nabla_0(H)^{1/2},$$

respectively. It follows that

$$\nabla_0(H^2) \geq \frac{\nabla_0(H)^{1/2}}{8\Delta^-(\vec{G}_{d-1})^2}.$$

Note that we can adjust the embedding ϕ_E^2 in such a way that no nail of the minor is incident to an edge in E_d : for every nail $v \in V^l$, we have added all pivots $\mu^{-1}(u)$ responsible for edges incident to v as nails—since the edges incident to these pivots have weight at most d , they are not in E_d . For every nail $v \in V^h$ we have reduced the number of edges a single pivot in $\mu^{-1}(v)$ contributes to v to one. Consider such an edge $vu \in H^2$ with $\pi(vu) = w \in \mu^{-1}(v)$ and assume we chose vu to be the single edge w still adds to v . If vx is the first edge of the embedding $\phi_E^1(vu)$, we reroute via the two edges $vw, wx \in G_{d-1}$ instead. We conclude this part of the construction by tracking how the parameters of the embedding have changed so far.

Claim. *The embedding ϕ_E^2 of H^2 in G_{d-1} has depth at most $(r+2)/2$. and weight at most γ .*

The construction of H^1 from H^0 simply removed edges, so we focus on the step from H^1 to H^2 . For every edge $uv \in E(H^1) \cap E(H^2)$ we potentially rerouted the embedding of uv through up to two pivots of \mathfrak{P}^h , hence the embedding of such an edge was lengthened by two edges. For an edge $uv \in H^2$ where one of the endpoints is a pivot, the length of the embedding (on that side) only got shorter. We conclude that the embedding has depth at most $(r+2)/2$. For the weight of the embedding, notice that replacing an edge of E_d by a reroute through its respective pivot adds two edges whose sum is exactly d ; hence the operations so far could only have decreased the weight of the embedding. Consequently, the number of edges of weight d that can occur in any embedding $\phi_E^2(e), e \in H^2$ is bounded by $\lfloor \gamma/d \rfloor$.

We are now left with treating the pivots \mathfrak{P}^l and their respective edges of E_d .

Claim. *Let $xy \in E_d$ be such that $\pi(xy) = w \in \mathfrak{P}^l$ and there exists an edge $uv \in H^2$ with $xy \in \phi_E^2(uv)$. Then x, y are distinct from u, v .*

By the definition of \mathfrak{P}^l we have that x, y cannot be contained in H^0 since then w would be contained in \mathfrak{P}^h . Further, since $xy \in E_d$, neither the vertex x nor the vertex y can be in \mathfrak{P}^h and we conclude that $x, y \notin V(H^2)$. A simple consequence is that in every pivot in \mathfrak{P}^l can only be responsible for up to $\Delta^-(\vec{G}_{d-1})$ edges in the embedding of H^2 .

With the above two claims we can construct the final minor H^3 iteratively as follows: we take an arbitrary remaining edge in $e \in E(H^2)$

and collect all pivots \mathfrak{P}_e that occur as pivots of edges in $\phi_E^2(e)$. We delete all up to $\lfloor \gamma/d \rfloor \Delta^-(\vec{G}_{d-1})$ other edges of H^2 that share any of the pivots \mathfrak{P}_e and reroute the embedding of e to avoid all edges of E_d using \mathfrak{P}_e . This operation cannot increase the weight of the embedding, but it will increase its length by up to $\lfloor \gamma/d \rfloor$ many edges.

Finally, the minor H^3 has an embedding that avoids all edges of E_d and hence is a topological minor of G_{d-1} . Its density is at least

$$\nabla_0(H^3) \geq \frac{\|H^3\|}{|H^3|} \geq \frac{\nabla_0(H^2)}{\lfloor \gamma/d \rfloor \Delta^-(\vec{G}_{d-1})} \geq \frac{\nabla_0(H)^{1/2}}{8 \lfloor \gamma/d \rfloor \Delta^-(\vec{G}_{d-1})^3}.$$

Remember that we constructed H^3 implicitly in $G_{d-1} \bullet \bar{K}_2$ by assuming that the pivots were disjoint from the vertices used by the embedding of H . We conclude that G_{d-1} contains a minor H' of density at least

$$\nabla_0(H') \geq \frac{\nabla_0(H^3) - 1}{2r' + 1} \geq \frac{\nabla_0(H^3)}{5r'} \geq \frac{\nabla_0(H)^{1/2}}{40r' \lfloor \gamma/d \rfloor \Delta^-(\vec{G}_{d-1})^3}$$

and depth at most $r'/2 := (r + 2 + \lfloor \gamma/d \rfloor)/2$ and weight at most γ . Note that we assume that $\nabla_0(H^3) \geq 2$ to make the above simplifications. \square

We directly arrive at the following corollary by observing that an r -shallow topological minor of the d^{th} augmentation can have weight at most $(2r + 1)d$.

Corollary 7. *Let $\vec{G}_1, \vec{G}_2, \dots$ be a dtf-augmentation. Then for every integer d and half-integer r it is true that*

$$\tilde{\nabla}_r(G_d) \leq 2^{12} (2r + 1)^4 \Delta^-(\vec{G}_{d-1})^6 \tilde{\nabla}_{2r+3/2}^d(G_{d-1})^2.$$

Lemma 33 provides us with a recurrence to bound the (stable and weighted) topological grad of dtf-augmentations. However, the bound contains the maximum indegree of the previous augmentation; therefore we need a second recurrence relation to come to a closed form.

Lemma 34. *For every graph G there exists a dtf-augmentation $\vec{G}_1, \vec{G}_2, \dots$ such that*

$$\Delta^-(\vec{G}_d) \leq \Delta^-(\vec{G}_{d-1})^2 + 12 \Delta^-(\vec{G}_{d-1})^3 \cdot \tilde{\nabla}_{1/2}^d(G_{d-1}).$$

Proof. We estimate the above quantities for an orientation \vec{G}_d by counting the transitive and fraternal edges added separately. To that end, let \vec{G}_d^f be the graph \vec{G}_{d-1} with the fraternal arcs of \vec{G}_d added and, analogously, \vec{G}_d^t the graph \vec{G}_{d-1} with the transitive arcs added.

It is easy to see that $\Delta^-(\vec{G}_d^t) \leq \Delta^-(\vec{G}_{d-1})^2$, so let us focus on the fraternal arcs. Our orientation strategy is simple: we will prove that G_d^f has low degeneracy, hence an acyclic orientation of the graph G_d^f will have low maximal indegree.

Claim. Let $s := \Delta^-(\vec{G}_{d-1})^2$. The graph G_d^f is a $1/2$ -stable topological minor of $G_{d-1} \bullet \bar{K}_s$ with weight d .

For every vertex $v \in G_{d-1}$, at most

$$\binom{N_{d-1}^-(v)}{2} \leq \binom{\Delta^-(G_{d-1})}{2} = \frac{\Delta^-(G_{d-1})(\Delta^-(G_{d-1}) - 1)}{2} \leq s - 1$$

fraternal edges will be added in its in-neighbourhood $N_d^-(v)$. Note that the above inequality assumes $\Delta^-(G_{d-1}) \geq 1$, which we can safely assume (the lemma holds trivially for edgeless graphs). Hence, we can construct the graph G_d^f from $G_{d-1} \bullet \bar{K}_s$ as follows: for every vertex $v \in G_{d-1}$, let v^0, v^1, \dots, v^s denote its twin-vertices in the product graph. For each edge that needs to be added in $N_{d-1}^-(G)$ in order to construct G_d^f , we contract a vertex v^i with $i \geq 1$ to obtain it—by our choice of s , we will not run out of vertices. After all fraternal edges have been constructed, we remove all left-over vertices v^i , $i \geq 1$ and arrive at a graph isomorphic to G_d^f . Since we only contracted paths of length 2 with weight d , we conclude that G_d^f is a $1/2$ -stable topological minor of $G_{d-1} \bullet \bar{K}_s$ with the claimed weight.

The claim allows us to apply Lemma 24 to show that

$$\begin{aligned} \nabla_0(G_d^f) &\leq (4(1/2 + 1)(s - 1) + 1)s^2 \cdot \check{\nabla}_{1/2}^d(G_{d-1}) \\ &\leq 6s^3 \cdot \check{\nabla}_{1/2}^d(G_{d-1}). \end{aligned}$$

In order to obtain a bound on the maximum indegree we need to fix and orientation strategy for the transitive edges. Since we saw that the degeneracy of G_d^f is bounded, a natural choice is an acyclic orientation of the graph G_d^f via the degeneracy-ordering. With this strategy we obtain

$$\begin{aligned} \Delta^-(\vec{G}_d) &\leq \Delta^-(G_d^t) + \Delta^-(G_d^f) \\ &\leq \Delta^-(\vec{G}_{d-1})^2 + 12\Delta^-(\vec{G}_{d-1})^3 \cdot \check{\nabla}_{1/2}^d(G_{d-1}), \end{aligned}$$

which is the claimed bound. \square

We now possess all the ingredients to bound the maximum indegree of the d^{th} dtf-augmentation by the (stable topological) grad of the original graph: we only have to resolve the recurrences given by Lemma 33 and by Lemma 34.

Theorem 16. For every graph G there exists a dtf-augmentation $\vec{G}_1, \vec{G}_2, \dots$ such that

$$\Delta^-(\vec{G}_{d+1}) \leq 2^{8 \cdot 5^d} d^{6^d} (d + 1)^{2^d} \cdot (\check{\nabla}_{d+1}^d(G_1) \cdot \Delta^-(\vec{G}_1))^{3 \cdot 5^{d-1}}.$$

Proof. Lemma 33 and Lemma 34 provide us with the recurrences

$$\Delta^-(\vec{G}_d) \leq 13\Delta^-(\vec{G}_{d-1})^3 \cdot \check{\nabla}_{1/2}^d(G_{d-1}) \quad (1)$$

$$\check{\nabla}_{r_i}^{d+1}(G_i) \leq (40r_{i-1} \cdot \Delta^-(\vec{G}_{i-1}))^3 \cdot \check{\nabla}_{r_{i-1}}^{d+1}(G_{i-1})^2 \quad (2)$$

where the depth r_i is recursively defined via

$$\begin{aligned} r_d &:= 1/2 \\ r_i &:= r_{i+1} + 2 + \left\lfloor \frac{d+1}{i+1} \right\rfloor. \end{aligned}$$

This easily expands into a closed form, allowing a simple bound:

$$\begin{aligned} r_i &:= 1/2 + 2(d-i) + \sum_{k=i}^{d-1} \left\lfloor \frac{d+1}{k+1} \right\rfloor \\ &\leq 1/2 + 2(d-i) + (d+1) \sum_{k=i+1}^d \frac{1}{k} \\ &= 1/2 + 2(d-i) + (d+1)(H_d - H_i). \end{aligned}$$

The second recurrence by can be brought into a more manageable form by letting $x_i := \log \tilde{\nabla}_{r_i}^{d+1}(G_i)$, $y_i := \log \Delta^-(\vec{G}_i)$, and taking logarithms on both sides:

$$\begin{aligned} x_i &\leq 2 \log 40r_{i-1} + 6y_{i-1} + 2x_{i-1} \\ \implies x_i &\leq 2^{i-1}x_1 + \sum_{k=1}^{i-1} 2^{i-k}(\log 40 + \log r_k + 3y_k) \\ &= 2^{i-1}x_1 + (2^i - 1) \log 40 + \sum_{k=1}^{i-1} 2^{i-k}(\log r_k + 3y_k). \end{aligned}$$

To bound the sum over $\log r_k$, notice that

$$\begin{aligned} \log r_k &= \log((d+1)(H_d - H_k) + 2(d-k) + 1/2) \\ &= \log((d+1)(H_d - H_k)) + \log\left(1 + \frac{2(d-k) + 1/2}{(d+1)(H_d - H_k)}\right) \\ &\leq \log(d^2 + d) + \frac{2(d-k) + 1/2}{(d+1)\log d} \end{aligned}$$

and therefore

$$\sum_{k=1}^{i-1} 2^{i-k} \log r_k \leq (2^i - 1) \left(\log(d^2 + d) + \frac{5}{2\log d} \right),$$

which simplifies the recurrence to

$$x_i \leq 2^{i-1}x_1 + (2^i - 1) \left(\log 40(d^2 + d) + \frac{5}{2\log d} \right) + \sum_{k=1}^{i-1} 2^{i-k} 3y_k.$$

Re-substitution provides the following reformulation

$$\tilde{\nabla}_{r_i}^{d+1}(G_i) \leq \tilde{\nabla}_{r_1}^{d+1}(G_1)^{2^{d-1}} (40(d^2 + d) 2^{\frac{5}{2\log d}})^{2^{d-1}} \prod_{k=1}^{d-1} \Delta^-(\vec{G}_k)^{3 \cdot 2^{i-k}},$$

of the recurrence (2). Plugging this into (1) yields

$$\Delta^-(\vec{G}_{d+1}) \leq 13 \tilde{\nabla}_{r_1}^{d+1}(G_1)^{2^{d-1}} (40(d^2 + d) 2^{\frac{5}{2\log d}})^{2^{d-1}} \prod_{k=1}^d \Delta^-(\vec{G}_k)^{3 \cdot 2^{d-k}},$$

which we can finally resolve. The solution to the following recurrence relation (related again logarithmically to our recurrence)

$$z_{d+1} = c_d + 3 \sum_{k=1}^d 2^{d-k} z_k$$

has the form

$$z_{d+1} = c_d + 3 \sum_{k=1}^{d-1} 5^{d-k-1} c_k + 3 \cdot 5^{d-1} z_1$$

which can be easily checked via induction and the fact that

$$z_{d+1} = c_d + 3z_d + 2(z_d - c_{d-1}).$$

With

$$c_d := \log \left(13 \check{\nabla}_{r_1}^{d+1}(G_1)^{2^{d-1}} (40(d^2 + d) 2^{\frac{5}{2 \log d}})^{2^{d-1}} \right) \quad \text{and}$$

$$z_d := \log \Delta^-(\vec{G}_d)$$

we see that

$$2^{z_{d+1}} \leq 2^{c_d} \left(\prod_{k=1}^{d-1} 2^{5^{d-k-1} c_k} \right)^3 \cdot 2^{3 \cdot 5^{d-1} z_1},$$

which we can further simplify by considering

$$\begin{aligned} \prod_{k=1}^{d-1} 2^{5^{d-k-1} c_k} &= \prod_{k=1}^{d-1} \left(13 \check{\nabla}_{r_1}^{k+1}(G_1)^{2^{k-1}} (40(k^2 + k) 2^{\frac{5}{2 \log k}})^{2^{k-1}} \right)^{5^{d-k-1}} \\ &\leq 13^{(5^{d-1}-1)/4} (240d^2)^{\frac{1}{12} (5^d - 2^{d+2} + 3)} \check{\nabla}_d(G_1)^{\frac{1}{3} (5^{d-1} - 2^{d-1})} \\ &\leq 2^{6 \cdot 5^{d-1}} d^{5^{d-1}} \check{\nabla}_d(G_1)^{\frac{1}{3} 5^{d-1}} \end{aligned}$$

where we used the fact that $\check{\nabla}_{r_1}^d(G_1) = \check{\nabla}_d(G_1)$ since the edges of the initial graph $G_1 = G$ have all weight one and $d < r_1$. Applying the same argument again and using the just derived bound on the product, we finally arrive at the claimed bound:

$$\begin{aligned} \Delta^-(\vec{G}_{d+1}) &\leq (13 \check{\nabla}_{d+1}^d(G_1)^{2^{d-1}} (240(d^2 + d))^{2^{d-1}}) \\ &\quad \cdot (2^{6 \cdot 5^{d-1}} d^{5^{d-1}} \check{\nabla}_d^{k+1}(G_1)^{\frac{1}{3} 5^{d-1}})^3 \cdot \Delta^-(\vec{G}_1)^{3 \cdot 5^{d-1}} \\ &\leq 2^{8 \cdot 5^d} d^{6^d} (d+1)^{2^d} \cdot (\check{\nabla}_{d+1}^d(G_1) \cdot \Delta^-(\vec{G}_1))^{3 \cdot 5^{d-1}}. \quad \square \end{aligned}$$

Armed with Theorem 16 we can now prove the characterisation of bounded expansion classes by low-treedepth colouring. The technique used above was rather tedious and complicated; but it pays off here since we are able to bound the maximal number of colours necessary for a p -centred colouring by a function of the stable topological grad. This does not quite reach the best known bounds of $Q(\check{\nabla}_{2^{p-2}+1/2}(G))$, where Q is a polynomial of degree 2^{2^p} , but it is better than the known bound based on tf-augmentations which is $P(\nabla_{(8p)^{p-1}}(G))$, where P is a polynomial of a similar degree as Q (cf. [192]).

Theorem 17. *Let \mathcal{G} be a graph class of bounded expansion. Then for every graph $G \in \mathcal{G}$ and every $p \in \mathbb{N}$ we can compute a p -centred colouring with at most*

$$\chi_p(G) \leq 2^{2^{\log^{O(p)} p}} \tilde{\nabla}_{(2 \log p)^p}(G)^{3.5(2 \log p)^p}$$

colours in time

$$2^{2^{\log^{O(p)} p}} \tilde{\nabla}_{(2 \log p)^p}(G)^{6.5(2 \log p)^p} \cdot |G|.$$

Proof. By Lemma 32, a proper colouring of $\vec{G}_{(2 \log p)^p}$ is a p -centred colouring of G . Theorem 16 gives us the bound

$$\chi(G_{(2 \log p)^p}) \leq 2\Delta^-(\vec{G}_{(2 \log p)^p}) \leq 2^{2^{\log^{O(p)} p}} \tilde{\nabla}_{(2 \log p)^p}(G)^{3.5(2 \log p)^p}$$

Since the running time of the d^{th} dtf-augmentation step is bounded by $O(\Delta^-(\vec{G}_d)^2|G|)$, the claimed running time follows. \square

The above bounds seem astronomically high, but we need to keep in mind that they are only worst-case estimates. Our experiments (see Chapter 17) show that real-world graphs behave much tamer than the above bound might suggest.

We can use the same result to show that nowhere dense classes admit low-treedepth colourings. Unsurprisingly, the running time is almost linear. Note that Theorem 20, which is stated later in this thesis, provides us with the fact that $\tilde{\nabla}_r(G) = O(|G|^{o(1)})$ for every fixed r and every graph G from a nowhere dense class

Lemma 35 (cf. [192]). *Let \mathcal{G} be a nowhere dense graph class. For $G \in \mathcal{G}$ and $p \in \mathbb{N}$ it holds that $\chi_p(G) \leq O(|G|^{o(1)})$. Moreover, such a colouring is computable in time $O(|G|^{1+o(1)})$.*

Proof. Again, Theorem 16 gives us the bound

$$\chi(G_{(2 \log p)^p}) \leq 2\Delta^-(\vec{G}_{(2 \log p)^p}) = O(|G|^{o(1)}),$$

where the last bound follows from Theorem 20. Since the running time of the d^{th} dtf-augmentation step is bounded by $O(\Delta^-(\vec{G}_d)^2|G|)$, the claimed running time follows. \square

Open question 1. The bound $(2 \log p)^p$ on the number of dtf-augmentations necessary to obtain a p -centred could be immediately lowered by improving Lemma 31. This might in particular be possible if we fix the orientation strategy already at this point: consider a set of independent but pairwise fraternal vertices to which fraternal arcs are added in the same augmentation step. Then the acyclic orientation strategy for fraternal edges will immediately introduce an out-apex into this vertex set. Can we show a bound of $2^{O(p)}$ to match the best known bound?

An area in the theory of sparse graphs that seem to have garnered little attention so far is the neighbourhood structure. The one example that comes to mind is the result by Grohe, Kreutzer, and Siebertz that nowhere dense classes admit small ‘neighbourhood covers’ [128]. We present here two key ingredients to further unlock the algorithmic potential of bounded expansion classes: the *twin class lemma* and its refinement the *charging lemma*. Both will be extremely useful in our work on preprocessing sparse graphs (Chapters 11 and 13). Inspired by the bounds provided through the twin class lemma, we define the notion of *neighbourhood complexity* and show that it provides an alternative characterisation of bounded expansion classes. Our method of arriving at this result uses some novel insight into how p -centred colourings structure a graph and we find an interesting connection to laminar set families.

8.1 THE TWIN AND CHARGING LEMMA

The following lemma is the basis of much of the work presented in this thesis. I first proved it in the context of graphs excluding a topological minor [156] and later found the right way of extending it to graphs of bounded expansion and nowhere dense classes [114]. It is easiest proved by considering a bipartite graph (X, Y, E) where we want to bound the structure and size of the set Y in terms of $|X|$. It will be heavily featured in the kernelisation results in Chapter 11. The below statement is in its most general form (which some of the applications need). One can mentally replace the values τ, ω by $\tilde{\nabla}_{1/2}(G)$ and $\tilde{\omega}_{1/2}(G)$, respectively, without losing too much of its power.

Lemma 36 (Twin class lemma). *Let $G = (X, Y, E)$ be a bipartite graph, and let*

$$\begin{aligned} \tau &\geq \nabla_0((G \tilde{\nabla}_{1/2})_{\leq |X|}) \quad \text{and} \\ \omega &\geq \omega((G \tilde{\nabla}_{1/2})_{\leq |X|}). \end{aligned}$$

Then it holds that

1. $|\{u \mid \deg(u) > 2\tau\}_{u \in Y}| \leq 2\tau \cdot |X|$, and
2. $|\{N(u)\}_{u \in Y}| \leq (\min\{4^\tau, \omega(e\tau)^\omega\} + 2\tau) \cdot |X|$.

Proof. We construct a sequence G_0, G_1, \dots, G_ℓ of topological $\frac{1}{2}$ -shallow minors of G as follows. Set $G_0 = G$, and for $0 \leq i \leq \ell - 1$ construct G_{i+1} from G_i by choosing a vertex $v \in V(G_i) \setminus X$ such that $N(v) \subseteq X$ contains two non-adjacent vertices u, w in G_i ; then contract v into the

edge vu to obtain G_{i+1} . Note that this necessarily adds the edge uw . It is clear from the construction that $X \subseteq V(G_i) \subseteq X \cup Y$ for $0 \leq i \leq \ell$.

This process clearly terminates, as G_{i+1} has at least one more edge between vertices of X than G_i . Note that $G_i \in G \tilde{\nabla} 1/2$ for $0 \leq i \leq \ell$, as the edges e_1, \dots, e_{i-1} that were contracted to vertices in X in order to construct G_i correspond to paths of length two in G whose internal vertices are pairwise disjoint. We therefore conclude that G_i is a $1/2$ -shallow topological minor of G . Accordingly, $G_\ell[X] \in (G \tilde{\nabla} 1/2)_{\leq |X|}$. This implies that $G_\ell[X]$ is 2τ -degenerate and has at most $2\tau \cdot |X|$ edges. Further, note that $Y \cap V(G_i)$ is, by construction, still an independent set in G_i for each $0 \leq i \leq \ell$.

Let us now prove the first claim. To this end, assume towards a contradiction that there is a vertex $v \in Y \cap V(G_\ell)$ such that $\deg(v) > 2\tau$. We claim that $G_\ell[N(v)]$ is a clique. If not, we could choose a pair of non-adjacent vertices in $G_\ell[N(v)]$ and construct a $(\ell + 1)$ -th graph for the sequence which would contradict the assumption that we exhausted the process. However, a clique of size $|\{v\} \cup N(v)| > 2\tau + 1$ is not 2τ -degenerate. Hence we conclude that no vertex of $Y \cap V(G_\ell)$ has degree larger than 2τ in G_ℓ . Therefore all vertices in Y of degree greater than 2τ in G must have been deleted during the edge contractions that resulted in the graph G_ℓ . As every contraction added one edge between vertices in X and since $G_\ell[X]$ contains at most $2\tau \cdot |X|$ edges, the first claim follows.

For the second claim, consider the set $Y_\ell = Y \cap V(G_\ell)$. The neighbourhood of every vertex $v \in Y_\ell$ induces a clique in $G_\ell[X]$. From the degeneracy of $G_\ell[X]$ and Proposition 5, it follows that $G_\ell[X]$ has at most $2^{2\tau} |G_\ell[X]| = 4^\tau \cdot |X|$ complete subgraphs. Similarly, from Lemma 7 we obtain the bound $\omega(e\tau)^\omega \cdot |X|$.

Therefore the total number of neighbourhoods in X induced by vertices of Y in G is bounded by the number of contractions $\ell \leq 2\tau \cdot |X|$ and the resulting number of complete subgraphs in $G_\ell[X]$. This is exactly the second claim. \square

A related tool developed for our paper on the kernelisation of DOMINATING SET in structurally sparse classes [76] is the following *charging lemma*. As it turns out, not only is the number of twin-classes in the above setting bounded, they are also in a certain sense ‘well-distributed’. We provide an alternative proof here using a more direct approach than the one presented in the paper, which should be attributed to Marcin and Michał Pilipczuk, and Daniel Lokshtanov.

Lemma 37 (Charging lemma, cf. [76]). *Let $G = (X, Y, E)$ be a bipartite graph such that Y is twin-free and every $u \in Y$ has a non-empty neighbourhood. Let again*

$$\begin{aligned} \tau &\geq \nabla_0((G \tilde{\nabla} 1/2)_{\leq |X|}) \quad \text{and} \\ \omega &\geq \omega((G \tilde{\nabla} 1/2)_{\leq |X|}). \end{aligned}$$

Then there exists a mapping $\phi: Y \rightarrow X$ with the following properties:

- $u\phi(u) \in E$ for $u \in Y$, and
- $|\phi^{-1}(v)| \leq 2\nabla_0(G)(\min\{4^\tau, \omega(e\tau)^\omega\} + 2\tau + 1)$ for every $v \in X$.

Proof. We partition the vertices of $Y = Y_\ell \cup Y_h$ according to their degree as follows:

$$Y_\ell := \{x \in Y \mid \deg(x) \leq 2\nabla_0(G)\} \text{ and } Y_h := Y \setminus Y_\ell$$

We construct two mappings $\phi_d: Y_d \rightarrow X$ for $d \in \{h, \ell\}$ and combine them at the end. Accordingly, let G_d be the two bipartite graphs with vertex sets (X, Y_d) , respectively.

Claim. *There exists a mapping $\phi_h: Y_h \rightarrow X$ such that $u\phi_h(u) \in E(X, Y_h)$ for every $u \in Y_h$ and $|\phi^{-1}(v)| \leq 2\nabla_0(G)$ for every $x \in X$.*

The graph G_h is $2\nabla_0(G)$ -degenerate thus contains a vertex of degree at most $2\nabla_0(G)$. By the construction of Y_h then, this vertex must be contained in X . With this observation at hand we proceed as follows to construct ϕ_h : let $v \in X$ be such that $\deg_{G_h} \leq 2\nabla_0(G)$. We set $\phi_h(u) = x$ for every $x \in N_{G_h}(v)$, then we remove the set $N[x]$ from the graph. Note that this operation does not change the remaining vertices of Y_h , hence we can repeat this operation until every vertex of Y_h has been charged to a vertex in X . The bound on ϕ_h^{-1} follows directly from the observed degree bound.

Claim. *There exists a mapping $\phi_\ell: Y_\ell \rightarrow X$ such that $u\phi_\ell(u) \in E(X, Y_\ell)$ for every vertex $u \in Y_\ell$ and $|\phi^{-1}(v)| \leq 2\tau(\min\{4^\tau, \omega(e\tau)^\omega\} + 2\tau)$ for every vertex $x \in X$.*

By Lemma 36 the set Y_ℓ (since it is twin-free) has size at most

$$|Y_\ell| \leq (\min\{4^\tau, \omega(e\tau)^\omega\} + 2\tau) \cdot |X|.$$

Therefore we have that

$$\begin{aligned} \sum_{v \in X} \deg_{G_\ell}(v) &= \sum_{u \in Y_\ell} \deg_{G_\ell}(u) \\ &\leq 2\nabla_0(G) \cdot (\min\{4^\tau, \omega(e\tau)^\omega\} + 2\tau) \cdot |X|, \end{aligned}$$

which implies that the average degree of vertices in X is

$$\frac{1}{|X|} \sum_{v \in X} \deg_{G_\ell}(v) \leq 2\nabla_0(G) \cdot (\min\{4^\tau, \omega(e\tau)^\omega\} + 2\tau).$$

Accordingly, there exists a vertex $v \in X$ that has degree at most $2\nabla_0(G)(\min\{4^\tau, \omega(e\tau)^\omega\} + 2\tau)$ in G_ℓ . With this observation we construct the mapping ϕ_ℓ exactly as in the high-degree case: We take the vertex $v \in X$ of minimal degree, set $\phi(u) = v$ for $u \in N(v)$ and delete $N[v]$ from G_ℓ . Note that in the resulting graph, the remainder of the set Y_ℓ is still twin-free. It follows that we can iterate this procedure until every vertex of Y_ℓ has been charged to a vertex in X . The bound on ϕ_ℓ^{-1} follows directly from the observed degree bound.

To finally obtain the mapping ϕ we simply join the two constructed mappings ϕ_h, ϕ_ℓ . The first desired property follows straight from the construction of the two mappings. For the size bound we have that for every $v \in X$ it holds that

$$|\phi^{-1}(v)| = |\phi_h^{-1}(v)| + |\phi_\ell^{-1}(v)| \leq 2\nabla_0(G)(\min\{4^\tau, \omega(e\tau)^\omega\} + 2\tau + 1),$$

as claimed. \square

The factor $\omega(e\tau)^\omega$ in the above two lemmas might look inferior to the much cleaner 4^τ , however, in the case of nowhere dense classes the quantity τ might depend sublinear on $|X|$ whereas ω will remain a constant. In this case, we prefer a polynomial dependence on τ .

8.2 A CHARACTERISATION BY NEIGHBOURHOOD COMPLEXITY

Neighbourhood complexity

The twin class lemma and the charging lemma have great algorithmic potential as we will see in the next part of the thesis. But one has to wonder: they both work on a strictly local level, formulated as a ‘depth-one’ statement if you will. Most statements and characterisations of bounded expansion classes work beyond that by introducing exactly this depth as a parameter. Can we formulate and prove such a variation? Let us define a parametrised measure which captures the spirit of the twin class lemma.

Definition 17 (Neighbourhood complexity). For a graph G we define the r -neighbourhood complexity as a function ν_r via

$$\nu_r(G) := \max_{H \subseteq G, X \subseteq V(H)} \frac{|\{N^r[v] \cap X\}_{v \in H}|}{|X|}.$$

We extend this definition to graph classes \mathcal{G} via $\nu_r(\mathcal{G}) := \sup_{G \in \mathcal{G}} \nu_r(G)$.

We say that a graph class \mathcal{G} has *bounded neighbourhood complexity* if there exists a function f such that for every r it holds that $\nu_r(\mathcal{G}) < f(r)$.

Note the hereditary nature of the definition: we need to define it over all subgraphs H of G in order to exclude dense graphs—without this detail, a clique would have neighbourhood complexity one (which might be an interesting measure in its own right).

The goal of this section is to show that graphs of bounded expansion are exactly those with neighbourhood complexity.

Theorem 18. *A graph class \mathcal{G} has bounded expansion if and only if it has bounded neighbourhood complexity.*

The main challenge is to prove that graphs from a graph class of bounded expansion have low neighbourhood complexity, the other direction is quite simple. Some definitions will be necessary to prove Lemma 38 which relates the neighbourhood complexity to the centred colouring number.

A *signature* σ over a universe U is a sequence $(u_i)_{1 \leq i \leq \ell}$, $u_i \in U$ where ℓ is the *length* of the signature, also denoted by $|\sigma|$. Accordingly, a ℓ -*signature* is simply a signature of length ℓ . We use the notation $\sigma[i] := u_i$ to signify the i^{th} element of σ . A signature is *proper* if all its elements are pairwise distinct. We impose a total order on all signatures (say, lexicographic). Thus for a set Σ of signatures and a function $f: \Sigma \rightarrow A$ for an arbitrary set A , we employ the notation $(f(\sigma))_{\sigma \in \Sigma}$ to obtain sequences over elements of A derived from that ordering.

Recall that for a path P , we denote by $P[i]$ the i^{th} vertex on the path. Hence, Let G be a graph coloured by $c: V(G) \rightarrow [\zeta]$ for some $\zeta \in \mathbb{N}$. Consider a path $P \in G$, then we denote by σ_P the $|P|$ -signature over $[\zeta]$ with $\sigma_P[i] = c(P[i])$. For a fixed signature σ , we say that $P \in G$ is a σ -*path* if $\sigma_P = \sigma$.

σ -neighbourhood, $N^\sigma(\bullet)$

For a fixed signature σ over $[\zeta]$, we define the σ -*neighbourhood* of a vertex v in G as

$$N^\sigma(v) := \{w \in G \mid \exists vPw \text{ such that } \sigma_{vPw} = \sigma\}$$

Note that $N^\sigma(v) \subseteq N^{|\sigma|}(v)$. Also, if $\sigma[0] \neq c(v)$ then $N^\sigma(v) = \emptyset$. We use the following extension to vertex sets $X \in V(G)$ and sets of signatures Σ over $[\zeta]$:

$$\begin{aligned} N^\Sigma(v) &:= \bigcup_{\sigma \in \Sigma} N^\sigma(v) & N^\sigma(X) &:= \bigcup_{v \in X} N^\sigma(v) \\ N^\Sigma(X) &:= \bigcup_{v \in X} \bigcup_{\sigma \in \Sigma} N^\sigma(v) \end{aligned}$$

Similarly, the σ -*in-neighbourhood* of a vertex v is defined as

$$N^{-\sigma}(v) := \{w \in G \mid \exists wPv \text{ such that } \sigma_{wPv} = \sigma\}$$

and we extend this notation to vertex and signature sets in the same manner as above:

$$\begin{aligned} N^{-\Sigma}(v) &:= \bigcup_{\sigma \in \Sigma} N^{-\sigma}(v) & N^{-\sigma}(X) &:= \bigcup_{v \in X} N^{-\sigma}(v) \\ N^{-\Sigma}(X) &:= \bigcup_{v \in X} \bigcup_{\sigma \in \Sigma} N^{-\sigma}(v) \end{aligned}$$

The following basic fact about σ -neighbourhoods for proper signatures σ is easy to verify.

Observation 1. *Let $u, v \in G$ be distinct vertices and $uP\bullet, vP\bullet$ be two σ -paths for some proper signature σ . Then for any $x \in u\sigma\bullet \cap v\sigma\bullet$ it holds that x has the same index on both $u\sigma\bullet$ and $v\sigma\bullet$ and that x is a centre of $uP\bullet \cup vP\bullet$.*

We can define the neighbourhood complexity via the index of an equivalence relation. This turns out to be a useful perspective in the subsequent proofs.

For $r \in \mathbb{N}$ and $X \subseteq V(G)$, we define the (r, X) -twin equivalence over $V(G)$ as

$$u \simeq_r^{G, X} v \iff N^r[u] \cap X = N^r[v] \cap X$$

which gives rise to the alternative definition

$$v_r(G) = \max_{H \subseteq G, X \subseteq V(G)} \frac{|V(H) / \simeq_r^{H, X}|}{|X|}.$$

We will usually fix a graph in the following and hence omit the superscript G of this relation.

8.3 BOUNDING NEIGHBOURHOOD COMPLEXITY

This section is dedicated to proving the following lemma, which constitutes the challenging direction of Theorem 18.

Lemma 38. *There exists a function f such that for every graph G , every subset $X \subseteq V(G)$ and $r \in \mathbb{N}$ it holds that*

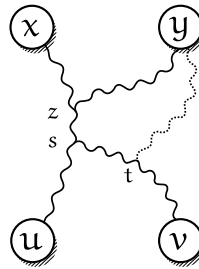
$$|V(G) / \simeq_r^X| \leq f(\chi_{2r+1}(G)) \cdot |X|$$

For the remainder of this section, fix a graph G , a vertex subset $X \subseteq V(G)$, an integer r and a $2r + 2$ -centred colouring $c: V(G) \rightarrow [\xi]$ where $\xi = \chi_{2r+1}(G)$. We will introduce a sequence of equivalence relations over $V(G)$ and prove that they successively refine $\simeq_{X, r}$. To that end, define $\Sigma_{\leq r}$ to be the set of all r -signatures over $[\xi]$.

The following sequence of lemmas will elucidate the connection between centred colourings and proper signatures.

Lemma 39. *For any proper signature $\sigma \in \Sigma_{\leq r}$ and any vertices $u, v \in V(G)$, either $N^\sigma(u) \cap N^\sigma(v) = \emptyset$ or $N^\sigma(u) = N^\sigma(v)$.*

Proof. Assume there exists $x \in N^\sigma(u) \cap N^\sigma(v)$ but $N^\sigma(u) \neq N^\sigma(v)$. Without loss of generality, let $y \in N^\sigma(v) \setminus N^\sigma(u)$.



Fix one σ -path uPx and a σ -path vPx . Let $s \in uPx \cap vPx$ be the first vertex in which both paths intersect (since both paths end in x , such a vertex must exist). Further, fix a σ -path vPy . Now if $vPy \cap uPx$ is non-empty, then y is σ -reachable from u : by Observation 1, there would be a vertex $z \in vPy \cap uPx$ that has the same index on both paths. Since

σ is proper, the subpath of vPy from z to y cannot share a vertex with uPx , thus we can construct a σ -path by first taking the subpath from u to z in uPx and then the subpath from z to y in vPy . Thus, assume vPy and uPx do not intersect. But then the graph $uPx \cup vPx \cup vPy$ is connected and contains every colour of σ at least twice. Since it holds that $|\sigma| \leq 2r + 1$, this contradicts our assumption that the colouring c is $(2r + 2)$ -centred. \square

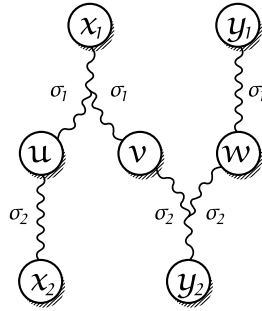
We see that a single proper signature σ imposes a very restricted structure on the respective σ -neighbourhoods in the graph. Even more interesting is the interaction of proper signatures with each other, as described in the following lemma.

Lemma 40. *For every pair of proper signatures $\sigma_1, \sigma_2 \in \Sigma_{\leq r}$ and every pair of vertices $a, b \in N^{-\sigma_1}(X) \cap N^{-\sigma_2}(X)$ we either have $[a]_{\sigma_1} \cap [b]_{\sigma_2} = \emptyset$, $[a]_{\sigma_1} \subseteq [b]_{\sigma_2}$ or $[a]_{\sigma_1} \supseteq [b]_{\sigma_2}$.*

Proof. The statement is trivial if $\sigma_1 = \sigma_2$ or $a = b$. Otherwise, assume that there exist $a \neq b$ such that indeed $[a]_{\sigma_1}$ and $[b]_{\sigma_2}$ are not related in the three above ways—since this is impossible when $|[a]_{\sigma_1}| = 1$ or $|[b]_{\sigma_2}| = 1$, we know that there exists vertices $u, v, w \in N^{-\sigma_1}(X) \cap N^{-\sigma_2}(X)$ with $u \in [a]_{\sigma_1} \setminus [b]_{\sigma_2}$, $v \in [b]_{\sigma_2} \setminus [a]_{\sigma_1}$ and $w \in [a]_{\sigma_1} \cap [b]_{\sigma_2}$.

The respective membership in the classes tell us the following about the vertices u, v, w : $N^{\sigma_1}(u) = N^{\sigma_1}(w) \neq N^{\sigma_1}(v)$ and at the same time $N^{\sigma_2}(u) \neq N^{\sigma_2}(w) = N^{\sigma_2}(v)$. Using Lemma 39 we can strengthen this statement: $N^{\sigma_1}(u) \cap N^{\sigma_1}(v) = \emptyset$ and $N^{\sigma_2}(u) \cap N^{\sigma_2}(v) = \emptyset$ and from the fact that $u, v, w \in N^{-\sigma_1}(X) \cap N^{-\sigma_2}(X)$ we know that all these neighbourhoods are non-empty.

Therefore, we can pick distinct vertices $x_1, y_1, x_2, y_2 \in X$ such that $x_1 \in N^{\sigma_1}(u)$, $y_1 \in N^{\sigma_1}(v)$ and $x_2 \in N^{\sigma_2}(u)$, $y_2 \in N^{\sigma_2}(v)$.



Since $N^{\sigma_1}(w) = N^{\sigma_1}(v)$, we can connect the vertices v, w with two (not necessarily disjoint) σ_1 -paths $P_u^{\sigma_1}, P_w^{\sigma_1}$ that start both in x_1 . Further, there exists a σ_1 -path $P_v^{\sigma_1}$ from y_1 to v . If $P_v^{\sigma_1}$ would intersect either $P_u^{\sigma_1}$ or $P_w^{\sigma_1}$, we could not have that $N^{\sigma_1}(v) \cap N^{\sigma_1}(u) = \emptyset$. We conclude that indeed $P_v^{\sigma_1}$ is disjoint from both $P_u^{\sigma_1}$ and $P_w^{\sigma_1}$.

We repeat the same construction for x_2, y_2 and the signature σ_2 to obtain paths $P_u^{\sigma_2}, P_v^{\sigma_2}, P_w^{\sigma_2}$. This time, $P_u^{\sigma_2}$ is necessarily disjoint from both $P_v^{\sigma_2}$ and $P_w^{\sigma_2}$. We reach a contradiction: observe that the graph

induced by the paths $P_u^{\sigma_1}, P_v^{\sigma_1}, P_w^{\sigma_1}, P_u^{\sigma_2}, P_v^{\sigma_2}, P_w^{\sigma_2}$ is connected, contains every colour of σ_1, σ_2 at least twice and in total at most $2r + 1$ colours. This is impossible if c was indeed $(2r + 2)$ -centred. \square

The above immediately yields a useful bound if we restrict ourselves to vertices that connect to X via the same signatures: we show in the next lemma that their number is linear in the size of X . In order to do so, we need to introduce another equivalence relation, relative to any set of signatures $\hat{\Sigma}$: let in the following

$$u \simeq_{\hat{\Sigma}}^X v \iff (N^\sigma(u) \cap X)_{\sigma \in \hat{\Sigma}} = (N^\sigma(v) \cap X)_{\sigma \in \hat{\Sigma}}.$$

Recall that $N^{-\sigma}(v)$ denotes the σ -in-neighbourhood of v , that is, it contains all vertices from which we can reach v by a σ -path.

Lemma 41. *Let $\hat{\Sigma} \subset \Sigma_{\leq r}$ be a set of proper signatures and let $W_{\hat{\Sigma}} = \bigcap_{\sigma \in \hat{\Sigma}} N^{-\sigma}(X)$ be those vertices in G who have a non-empty σ -neighbourhood in X for every $\sigma \in \hat{\Sigma}$. Then $|W_{\hat{\Sigma}} / \simeq_{\hat{\Sigma}}^X| \leq |\hat{\Sigma}| \cdot |X|$.*

Proof. Define the set family $\mathcal{F} := \bigcup_{\sigma \in \hat{\Sigma}} (W_{\hat{\Sigma}} / \simeq_{\sigma}^X)$ that contains the classes of all equivalence relations defined via the signatures in $\hat{\Sigma}$. By Lemma 40 and our choice of $W_{\hat{\Sigma}}$, the family \mathcal{F} is laminar; i.e. every pair $B_1, B_2 \in \mathcal{F}$ satisfies $B_1 \cap B_2 \in \{\emptyset, B_1, B_2\}$.

Consider a class $B \in W_{\hat{\Sigma}} / \simeq_{\hat{\Sigma}}^X$. Then B is the result of a intersection of at most $|\hat{\Sigma}|$ classes in \mathcal{F} . Since $B \neq \emptyset$ and \mathcal{F} is laminar, it follows that $B \in \mathcal{F}$. We conclude that

$$|W_{\hat{\Sigma}} / \simeq_{\hat{\Sigma}}^X| \leq |\mathcal{F}| \leq |\hat{\Sigma}| \cdot |X| \quad \square$$

In order to apply the above lemma we need to bound the number of possible r -neighbourhoods in X by σ -neighbourhoods of *proper* signatures. We establish this bound by successively refining the (r, X) -twin equivalence. The following figure gives an overview over the proof (using relations yet to be introduced).

$$\begin{array}{ccc} u \simeq_{r-1}^X v & \iff & N^{r-1}[u] \cap X = N^{r-1}[v] \cap X \\ \uparrow \text{Lemma 42} & & \\ u \cong_{r-1}^X v & \iff & (N^i(u) \cap X)_{0 \leq i < r} = (N^i(v) \cap X)_{0 \leq i < r} \\ \uparrow \text{Lemma 43} & & \\ u \simeq_{\Sigma_{\leq r}}^X v & \iff & (N^\sigma(u) \cap X)_{\sigma \in \Sigma_{\leq r}} = (N^\sigma(v) \cap X)_{\sigma \in \Sigma_{\leq r}} \\ \uparrow \text{Lemma 44} & & \\ u \simeq_{\hat{\Sigma}_{\leq r}}^X v & \iff & (N^\sigma(u^1) \cap X^{|\sigma|})_{\sigma \in \hat{\Sigma}_{\leq r}} = (N^\sigma(v^1) \cap X^{|\sigma|})_{\sigma \in \hat{\Sigma}_{\leq r}} \end{array}$$

The last relation is defined with the help of an auxiliary graph \hat{G} and signature set $\hat{\Sigma}_{\leq r}$ whose construction is described later. The bound on the index of this last relation will prove Lemma 38.

Lemma 42. *The equivalence relation \simeq_r^X over $V(G)$ defined via*

$$u \simeq_r^X v \iff (N^i(u) \cap X)_{0 \leq i \leq r} = (N^i(v) \cap X)_{0 \leq i \leq r}$$

is a refinement of \simeq_r^X .

Proof. Assume $u \simeq_r^X v$. Since by definition $N^r[v] = \bigcup_{1 \leq i \leq r} N^r(v)$, we have that $N^r[u] = N^r[v]$ and hence $u \simeq_r^X v$. \square

Lemma 43. *The equivalence relation $\simeq_{\Sigma_{\leq r}}^X$ over $V(G)$ defined via*

$$u \simeq_{\Sigma_{\leq r}}^X v \iff (N^\sigma(u) \cap X)_{\sigma \in \Sigma_{\leq r}} = (N^\sigma(v) \cap X)_{\sigma \in \Sigma_{\leq r}}$$

is a refinement of \simeq_{r-1}^X .

Proof. Assume $u \simeq_{\Sigma_{\leq r}}^X v$. We need to prove that for every $0 \leq i < r$, it holds that $N^i(u) \cap X = N^i(v) \cap X$. To that end, consider the set of signatures $\Sigma_{=i} \subseteq \Sigma_{\leq r}$ of length exactly i . The equivalence of u and v implies that

$$(N^\sigma(u) \cap X)_{\sigma \in \Sigma_{=i}} = (N^\sigma(v) \cap X)_{\sigma \in \Sigma_{=i}}$$

for every $1 \leq i \leq r$. Therefore we have that

$$\begin{aligned} w \in N^i(v) \cap X &\implies \exists \sigma \in \Sigma_{=i}: w \in N^\sigma(v) \cap X \\ &\implies \exists \sigma \in \Sigma_{=i}: w \in N^\sigma(u) \cap X \\ &\implies w \in N^i[u] \cap X \end{aligned}$$

Assume towards a contradiction that actually $w \notin N^i(u)$. Since we can swap u and v in the above, we also have that $w \notin N^i(u) \cap X \implies w \notin N^i[v] \cap X$ which contradicts our above assumption. Hence we necessarily have that $N^i(u) \cap X = N^i(v) \cap X$. \square

We now construct an auxiliary graph and colouring as follows: let $\hat{G} = G \bullet \bar{K}_r$ and let $\hat{c}: V(\hat{G}) \rightarrow [\zeta] \times [r]$. Assuming that $V(K_r) = [r]$ and hence $V(\hat{G}) = V(G) \times [r]$, we will use the shorthand $v^i = (v, i)$ for $v \in V(G), i \in [r]$ and call v^i the i^{th} copy of v . Note that \hat{c} is a $2r + 2$ -centred colouring of \hat{G} : any connected subgraph $\hat{H} \subseteq \hat{G}$ with less than $2r + 2$ colours and no centre would directly imply that the subgraph $H \subseteq G$ with vertex set $V(H) = \bigcup_{1 \leq i \leq r} \{v \in G \mid v^i \in \hat{H}\}$ contains at most $2r + 2$ colours and no centre, contradicting our choice of c .

For a signature $\sigma \in \Sigma_{\leq r}$ we define the proper signature

$$\hat{\sigma} = ((\sigma[i], i))_{1 \leq i \leq |\sigma|}.$$

Accordingly, we define the signature set $\hat{\Sigma}_{\leq r}$ over colours $[\zeta] \times [r]$ as

$$\hat{\Sigma}_{\leq r} = \{\hat{\sigma} \mid \sigma \in \Sigma_{\leq r}\}.$$

The following lemma connects the sigma-equivalence $\simeq_{\Sigma_{\leq r}}^X$ over $V(G)$ with a suitable equivalence defined over the above auxiliary structure.

Lemma 44. *The equivalence relation $\simeq_{\hat{\Sigma}_{\leq r}}^X$ over $V(G)$ defined via*

$$u \simeq_{\hat{\Sigma}_{\leq r}}^X v \iff (N^\sigma(u^1) \cap X^{|\sigma|})_{\sigma \in \hat{\Sigma}_{\leq r}} = (N^\sigma(v^1) \cap X^{|\sigma|})_{\sigma \in \hat{\Sigma}_{\leq r}}$$

is a refinement of $\simeq_{\Sigma_{\leq r}}^X$ where $X^i := \{v^i \mid v \in X\}$.

Proof. Assume $u \simeq_{\hat{\Sigma}_{\leq r}}^X v$. Then for every signature $\sigma \in \Sigma_{\leq r}$ we have that $N^{\hat{\sigma}}(u^1) \cap X^{|\hat{\sigma}|} = N^{\hat{\sigma}}(v^1) \cap X^{|\hat{\sigma}|}$. Since a $\hat{\sigma}$ -path in \hat{G} corresponds to a σ -path in G , this implies that $N^\sigma(u) \cap X^{|\sigma|} = N^\sigma(v)$ and thus we have that $u \simeq_{\Sigma_{\leq r}}^X v$. \square

Lemma 45. $|V(G)/\simeq_{\hat{\Sigma}_{\leq r}}^X| \leq 2^{\xi^{r+1}} \cdot |X|$

Proof. To obtain the bound, we apply Lemma 41 to every subset of signatures $\hat{\Sigma} \subseteq \hat{\Sigma}_{\leq r}$. Note that every signature $\hat{\sigma}$ under consideration ends in a tuple (\bullet, r) . Hence, we can restrict ourselves to the set $\hat{X}^r = X \cap (V(G) \times \{r\})$ and obtain in total:

$$|V(G)/\simeq_{\hat{\Sigma}_{\leq r}}^X| \leq |V(\hat{G})/\simeq_{\hat{\Sigma}_{\leq r}}^{\hat{X}^r}| \leq \sum_{\hat{\Sigma} \subseteq \hat{\Sigma}_{\leq r}} |\hat{\Sigma}| \cdot |\hat{X}^r| = 2^{\xi^{r+1}} \cdot |X|$$

Which proves the claim. \square

Now the proof of the main lemma is only a technicality.

Proof of Lemma 38. By Lemma 42 to 44 we have that

$$|V(G)/\simeq_{r-1}^X| \leq |V(G)/\cong_{r-1}^X| \leq |V(G)/\simeq_{\Sigma_{\leq r}}^X| \leq |V(G)/\simeq_{\hat{\Sigma}_{\leq r}}^X|,$$

which, by Lemma 45, is at most $2^{\chi_{2r+1}(G)^{r+1}} \cdot |X|$. \square

We conclude that bounded expansion implies bounded neighbourhood complexity. Let us quickly prove the other direction to arrive at the full characterisation and thus Theorem 18.

Lemma 46. *For every graph G and every r it holds that $\tilde{\nabla}_r(G) \leq \nu_r(G)$.*

Proof. Fix r and let $H \preceq_t r G$ be an r -shallow topological minor of G with density $\nabla_0(H) = \tilde{\nabla}_r(G)$. Let ϕ_V, ϕ_E be an embedding witnessing this fact.

Since all paths $(\phi_E(e))_{e \in E(H)}$ connect to unique pair of vertices of H in G and the paths have length at most $2r + 1$, we have that

$$\nabla_0(H) \leq |\{N^r[v] \cap V(H)\}_{v \in \phi_E(E(H))}| \leq \nu_r(G),$$

which proves the claim. \square

We can therefore add neighbourhood complexity as yet another tool in our kit. It seem particularly useful in the context of kernelisation, where we usually reason about the structure of a graph G around

some special set X of small size (for example, an approximate solution to the problem obtained beforehand). The linear kernel for DOMINATING SET presented in Chapter 13 makes use of it, however, only at depth one for which the twin class lemma is already sufficient. It seems likely that an extension of this result to ‘deeper’ problems like r -DOMINATING SET will need to employ something like our bound on the neighbourhood complexity.

Open question 2. The r -neighbourhood complexity is related to the size of the chosen set X via an exponential function in the $(2r+1)$ -centred colouring number. Therefore we cannot say anything about nowhere dense classes, where the latter value can only be bounded by $|X|^\epsilon$. Can Lemma 38 be proved with a polynomial dependence on χ_\bullet to also characterise nowhere dense classes?

Open question 3. Can we provide a bound of the neighbourhood complexity that does not rely on p -centred colourings but works directly with shallow minors?

FURTHER CHARACTERIZATIONS

As noted before, the notion of structurally sparse classes is extremely robust in that there are a number of parametrised graph measures over which they can be defined: in this part alone we saw the measures $\check{\nabla}_\bullet$, $\check{\nabla}_\bullet$, $\Delta^-(\vec{G}_\bullet)$ and ν_\bullet . A whole zoo of other measures exist (cf. [192]) and we can define classes of bounded expansion by using any of them:

Theorem 19 (cf. Theorem 13.2 [192], [191]). *For every graph class \mathcal{G} that is nowhere dense and every parametrised graph measure f_\bullet equal to or polynomially related to any of $\{\nabla_\bullet, \check{\nabla}_\bullet, \check{\nabla}_\bullet, \check{\nabla}_\bullet, \chi_\bullet, \nu_\bullet, \text{col}_\bullet, \text{wcol}_\bullet, \text{adm}_\bullet\}$ it holds that for every integer r ,*

$$\limsup_{G \in \mathcal{G}} f_r(G) < \infty.$$

The measures wcol_\bullet , col_\bullet and adm_\bullet will be introduced below in Section 9.2. By observing that these measures are all polynomially related to ∇_\bullet , and every nowhere dense class \mathcal{G} satisfy $\nabla_\bullet(\mathcal{G}) = n^{o(1)}$, we can similarly define nowhere dense classes using these parameters (with the exception of ν_\bullet since we only related it to an exponential function of χ_\bullet).

Theorem 20 (cf. Theorem 13.1 [192], [191]). *For every graph class \mathcal{G} that is nowhere dense and every parametrised graph measure f_\bullet equal to or polynomially related to any of $\{\nabla_\bullet, \check{\nabla}_\bullet, \check{\nabla}_\bullet, \check{\nabla}_\bullet, \chi_\bullet, \text{col}_\bullet, \text{wcol}_\bullet, \text{adm}_\bullet\}$ it holds that for every integer r ,*

$$\limsup_{G \in \mathcal{G}} \log f_r(G) / \log |G| = 0.$$

Nešetřil, Ossona de Mendez, and Wood [193] also introduced a characterisation for bounded expansion classes using *unparametrised* graph measures¹. Let ϕ be a graph measure. We say that ϕ is

- *monotone* if for every subgraph $H \subseteq G$ it holds that $\phi(H) \leq \phi(G)$,
- *degree bound* if for some function f , every graph has degree at most $f(\phi(G))$, and
- *strongly topological* if for some function f , every graph G and every ≤ 1 -subdivision H of G it holds that

$$\phi(G) \leq f(\phi(H)) \text{ and } \phi(H) \leq f(\phi(G)).$$

Strongly topological, monotone, degree bound

Such measures are related to bounded expansion classes as follows:

¹ The use the term ‘graph parameter’ which for reasons of clarity we cannot adopt here.

Lemma 47 (Nešetřil, Ossona de Mendez, Wood [193]). *A class \mathcal{G} has bounded expansion if and only if there exists a graph measure ϕ that is monotone, degree bound, and strongly topological and a constant c such that*

$$\mathcal{G} \subseteq \{G \mid \phi(G) \leq c\}.$$

For example, the width-measure \mathbf{tw} fulfills all three requirements and we can—not very surprisingly—conclude that classes of bounded tree-width have bounded expansion. The authors apply this notion to show that graphs whose crossing number *per edge* is bounded by c form a class \mathcal{G} with $\tilde{\nabla}_r(\mathcal{G}) = O(\sqrt{cr})$. They also applied it to graphs with special embeddings, see their paper for details.

9.1 LARGE DEGREE AND SUBGRAPH-GRADS

The following characterisation will be very useful later to show that certain random graphs have bounded expansion. In essence, it enables us to prove that a class has bounded expansion if its members contain only few vertices of high degree and all dense shallow minors have embedding that span a large fraction of vertices.

Theorem 21 (Nešetřil, Ossona de Mendez, Wood [193]). *A graph class \mathcal{G} has bounded expansion if and only if there exist real-valued functions f_{thresh} , f_{deg} , $f_{\tilde{\nabla}}$, f_H such that for all $G \in \mathcal{G}$ the following two conditions hold:*

1. *For all $\varepsilon > 0$ either $|G| \leq f_{\text{thresh}}(\varepsilon)$ or it holds that*

$$|\{v \in V(G) : \deg(v) \geq f_{\text{deg}}(\varepsilon)\}| \leq \varepsilon \cdot |G|.$$

2. *For all $r \in \mathbb{N}$, all $H \subseteq G$ with $\tilde{\nabla}_r(H) > f_{\tilde{\nabla}}(r)$ satisfy*

$$|H| \geq f_H(r) \cdot |G|.$$

9.2 GENERALISED COLOURING NUMBERS

We have seen earlier how low indegree augmentations of a graph result in a characterisation of bounded expansion classes. In a sense, they generalise the low indegree orientations that exist for degenerate graphs and turn them into a parametrised graph measure. Other characterisations of degeneracy are based on density of subgraphs, which is generalised by the density of shallow minors, and on linear orderings. This latter definition is called the *colouring number*: Let $\Pi(G)$ denote the set of all linear orderings of the vertices of G . The colouring number is then defined as

$$\text{col}(G) := 1 + \min_{\pi \in \Pi(G)} \max_{v \in G} |N_{\pi}^{\leq}(v)|.$$

It gets its name from the simple fact that any graph can be properly coloured (by a simple greedy algorithm) with $\text{col}(G)$ colours—though

its chromatic number might be much smaller. If we seek out to generalised the colouring number by parametrising it with some notion of ‘depth’, we need to replace the left-neighbourhood N_π^\leq by an appropriate parametrised notion.

k-colouring number, weak –

Kierstead and Yang in their study of the game colouring number introduced two such parametrisations, namely, the *k-colouring number* col_k and the *weak k-colouring number* wcol_k [155]. To define them, we need to introduce the notions of *accessible* and *weakly accessible*. Fixing an ordering π , and two vertices x, y with $x <_\pi y$. Then x is weakly k -accessible from y if there exists an x - y -path P of length at most k such that $x \leq_\pi V(P)$, that is, all vertices of P come after x in the ordering π . If even $y \leq_\pi V(P)$ holds then x is k -accessible from y . The set of vertices that are k -accessible from y is denoted by $R_\pi^k(v)$ and the set of weakly k -accessible vertices by $Q_\pi^k(v)$. Note that both sets are exactly N_π^\leq for $k = 1$. Therefore the following notions both coincide with the colouring number for $k = 1$:

$$\begin{aligned} \text{col}_k(G) &= 1 + \min_{\pi \in \Pi(G)} \max_{v \in G} |R_\pi^k(v)|, \\ \text{wcol}_k(G) &= 1 + \min_{\pi \in \Pi(G)} \max_{v \in G} |Q_\pi^k(v)|. \end{aligned}$$

That these two quantities are polynomially related was also shown by Kierstad and Yang: it is true that

$$\text{col}_k(G) \leq \text{wcol}_k(G) \leq \text{col}_k(G)^k.$$

Backconnectivity, k-admissibility

A different variant of $|N_\pi^\leq(v)|$ is the *backconnectivity* $b_\pi^k(v)$. It is defined as the maximum number of paths of length at most k that only intersect in v and whose respective other endpoints come before v in the ordering. Note again that $b_\pi^1(v) = |N_\pi^\leq(v)|$. Accordingly, the *k-admissibility* is defined as

$$\text{adm}_k(G) = \min_{\pi \in \Pi(G)} \max_{v \in G} b_\pi^k(v).$$

This parametrised graph measure was introduced by Kierstead and Trotter in a paper bounding the game chromatic number for planar graphs [154]. From the definition it is immediately clear the k -admissibility is never larger than the k -colouring number: the paths used in the definition of the backconnectivity b_π^k can be assumed to contain exactly one vertex smaller than v , which gives us a lower bound for $|R_\pi^k(v)|$. Dvořák showed a bound in the opposite direction [79] and we therefore have that

$$\text{adm}_k(G) \leq \text{col}_k(G) \leq \text{adm}_k(G)(\text{adm}_k(G) - 1)^{k-1} + 1.$$

It was Xuding Zhu who related the above parametrised graph measures to classes of bounded expansion [253]: he showed that

$$\nabla_{(k-1)/2}(G) \leq \text{col}_k(G)^k \leq F_k(\nabla_{(k-1)/2}(G)),$$

where the function F_k is a polynomial of degree around k^{2k} . As a consequence, we can define bounded expansion classes and nowhere dense classes by the parametrised graph measures col_\bullet , wcol_\bullet or adm_\bullet as formulated in Theorem 19.

9.3 WIDENESS

We saw in Chapter 3 that while often algorithmic applications follow new insights in structural graph theory, sometimes the direction is reversed (consider the definition of graph classes with locally bounded treewidth). A similar advance-by-application can be posited in the *homomorphism preservation* programme. The classical result from mathematical logic that first-order formulas are preserved under structure homomorphisms if and only if they are existential and positive (or equivalent to such a formula) holds in the class of all infinite and in the class of all finite structures. However, it does *not* hold for all *classes* of finite structures. Atserias, Dawar, and Kolaitis showed that the homomorphism preservation holds in classes that are closed under taking substructures and disjoint unions; and whose Gaifman graph excludes a minor [14]. Introducing the notions of *quasi-wideness* (see below), Dawar extended this result to a much larger set of classes—including classes that have bounded expansion.

Wide, almost wide, quasi-wide

A graph class \mathcal{G} is *wide* if for all k, d there exists a threshold $N_{k,d}$ such that every graph of \mathcal{G} with at least $N_{k,d}$ vertices contains a d -scattered set of size k . The simplest example for such a class are bounded-degree graphs.

Similarly, a graph class \mathcal{G} is *almost wide* if there exists s such that we can remove up to s vertices from every graph in \mathcal{G} to turn it into a wide class. The class of stars is a simple example: the removal of the centre vertex allows arbitrarily large scattered sets. Alternatively, adding a few apex vertices to the members of a wide class turns it into an almost wide class. Importantly, the number s depends only on the class and not on k or d .

Finally, a graph class \mathcal{G} is *quasi-wide* if there exists a function s such that we can remove up to $s(d)$ vertices from every graph in \mathcal{G} to turn it into a wide class. It turns out that nowhere dense classes are quasi-wide and the converse holds for hereditary nowhere dense classes:

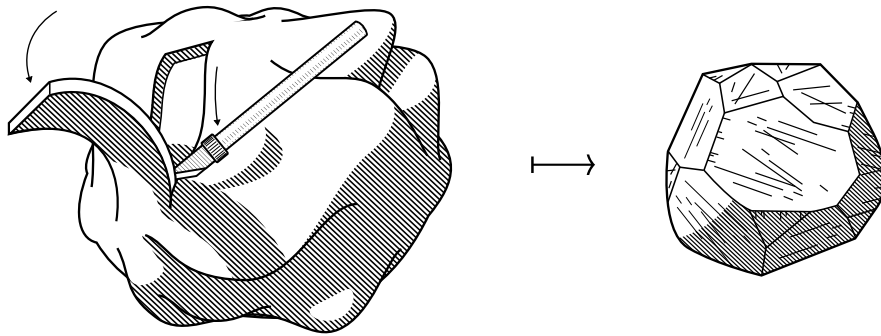
Theorem 22 (Nešetřil, Ossona de Mendez [190]). *Every hereditary graph class is nowhere dense if and only if it is quasi-wide.*

Furthermore, Dawar showed that classes of bounded expansion (as well as classes locally excluding a minor) are quasi-wide and provided a simple bound on the function s .

Theorem 23 (Dawar [57]). *Let \mathcal{G} be a bounded-expansion class. Then \mathcal{G} is quasi-wide with $s(r) = 2\tilde{\nabla}_{r+1}(\mathcal{G})$.*

Part III

ALGORITHMS AND PREPROCESSING USING
STRUCTURAL SPARSENESS



The bounded expansion toolkit presented in the preceding chapters—in particular the low-treedepth colourings introduced by Nešetřil and Ossona de Mendez and the first-order model checking by Dvořák, Král, and Thomas—has interesting applications to well-established algorithms designed for structurally sparse classes. Both of these general-purpose algorithms run themselves in linear time (or almost linear time in the case of nowhere dense classes) and we can apply these tools to algorithms designed for classes that are much smaller, say, planar graphs, to improve their polynomial dependence on the input size. The reason for this is twofold: first, in many fields (like parametrised algorithms), only recently have people taken interest in reducing the polynomial dependence of algorithms. Second, some tools developed in the context of bounded expansion classes simply have not had a comparable counterpart in smaller classes—in fact, most smaller classes provide much more powerful tools with high polynomial dependence (think Robertson-Seymour decomposition) that are so convenient to use that it might not have occurred to researchers to look for less powerful but faster replacements. Working in classes of bounded expansion, however, does not leave us any other choice: we have to re-think many ideas and work around the limits imposed by our tools.

Preprocessing

Two good examples of how fruitful this forced change of perspective can be will be described later in this part: the extension of the meta-kernels framework, where the restrictions already imposed by classes excluding a topological minor led to the shift to structural parametrisation and ultimately to an extension to bounded expansion and nowhere dense classes, and a ‘handcrafted’ linear kernel for DOMINATING SET that works in an entirely different manner than its predecessors in smaller classes. As a consequence, we can confidently state that structurally sparse classes—although far from being fully explored in terms of preprocessing—admit efficient (almost) linear-time preprocessing routines for a number of pivotal problems. We preface these results by smaller observations and algorithmic applications that we deem interesting and further support our claim that these classes are rich with algorithmic potential.

10.1 FASTER LOCAL SEARCH

Fellows *et al.* demonstrated a remarkably practical application of parametrised complexity in a paper on local search in graphs with locally bounded treewidth [91]. They observed that the usual k -exchange-

rule for local search, *i.e.* exchanging k elements of the current solution for $k - 1$ elements to obtain a smaller solution (in the case of minimisation), runs in time $n^{O(k)}$. Such running time is clearly prohibitive already for very small values of k . Since the quality of local search heuristics scales with the size of the exchanges, it is desirable to design algorithms that reduce the dependence on k . This is a perfect application for parametrised complexity: we decouple the dependence of the input size from the parameter and obtain an algorithm that will scale better in terms of k . Concretely, the authors provide k -exchange algorithms for the weighted version of r -CENTRE, VERTEX COVER, ODD CYCLE TRANSVERSAL, MAX-CUT and MIN-BISECTION.

We will demonstrate that for a whole class of problems—which includes VERTEX COVER, DOMINATING SET, r -DOMINATING SET and r -CENTRE—the unweighted k -exchange problem can be solved in linear time and the weighted k -exchange in almost linear time. For simplicity we will work only with vertex-exchanges here, but all techniques presented here generalise to edge-exchanges and mixed-exchanges as well.

min-II-VERTEX EXCHANGE parametrised by k

Input: A graph G and a feasible solution $X \subseteq G$ to Π .

Problem: Is there a tuple (X^-, X^+) with $X^- \subseteq X$, $|X^-| = k$ and $X^+ \cap X = \emptyset$, $|X^+| = k - 1$ such that $(X \setminus X^-) \cup X^+$ is a feasible solution of G to Π ?

FO-Certifiable Our approach is to exploit the first-order model checking machinery. In order to do so, we consider the following class of problems.

Definition 18 (FO-Certifiable). Let $\phi(X)$ be a first-order formula with one free set variable S . We say that a vertex-set optimisation problem is *first-order certifiable* if for every instance G every subset $X \subseteq V(G)$ is a feasible solution of G iff S satisfies ϕ .

Feasible k -exchange

This notion is very similar to first-order definable optimisation problems [59], without the monotonicity constraint. Given an instance G of a problem that is first-order certifiable by ϕ and given a feasible solution X , we will call a k -exchange (X^+, X^-) *feasible* if the set $(X \setminus X^-) \cup X^+$ satisfies ϕ .

Lemma 48. *Let Π be an optimisation problem that is first-order verifiable. Then for every $k \in \mathbb{N}$, the problem Π - k -VERTEX EXCHANGE is first-order definable.*

Proof. We assume that Π is a minimisation problem, the proof for maximisation problems works analogously. Let $\phi(X)$ be the first-order formula that certifies the feasible solutions of Π . We construct a first-order formula $\hat{\phi}(\bar{x}^+, \bar{x}^-)$ that additionally has a relation R_X in its vocabulary with $\bar{x}^+ = (x_i^+)_{i \in [k-1]}$ and $\bar{x}^- = (x_i^-)_{i \in [k]}$ from $\phi(X)$ as follows.

Let $\hat{\phi}$ be the formula $\phi(X)$ where every atom of the form Xv is replaced by the formula

$$(R_X v \wedge \bigwedge_{i=1}^k x_i^+ \neq v) \vee (\neg R_X v \wedge \bigwedge_{i=1}^{k-1} x_i^- \neq v).$$

We also add clauses to $\hat{\phi}$ to ensure that $\neg R_x x_i^+$ and $R_x x_i^-$ for all variables in \bar{x}^+, \bar{x}^- ; and a clause to ensure that $x_i^+ \neq x_j^+, x_i^- \neq x_j^-$ for $i \neq j$ and $x_i^+ \neq x_j^-$ for all i, j . It is easy to verify that, by construction, the formula

$$\exists \bar{x}^+ \exists \bar{x}^- \hat{\phi}(\bar{x}^+, \bar{x}^-)$$

is true for (G, X) (that is: the structure obtain from G, X by augmenting the canonical structure of G with the relation R_X that contains exactly the member of X) if and only if there exists a feasible k -exchange (X^+, X^-) for X in G . \square

Using the result by Dvořák, Král, and Thomas (see Chapter 4) and the result by Grohe, Kreutzer, and Siebertz (see Chapter 5) we immediately obtain the following:

Corollary 8. *Let Π be an optimisation problem that is first-order verifiable and \mathcal{G} be a graph class. Then Π - k -VERTEX EXCHANGE can be decided in non-uniform linear fpt-time on instances from \mathcal{G} if \mathcal{G} has bounded expansion and in non-uniform almost-linear fpt-time if \mathcal{G} is nowhere dense.*

As a consequence, we can solve VERTEX COVER k -VE, INDEPENDENT SET k -VE, DOMINATING SET k -VE, r -DOMINATING SET k -VE in linear fpt-time on bounded-expansion classes and almost linear fpt-time on nowhere-dense classes. Similar results can be obtained for H -PACKING, i.e. the problem of finding k vertex-disjoint (or edge-disjoint) subgraphs isomorphic to H ; we only need to define a suitable notion of exchange in first-order logic—for example, exchanging one subgraph of the packing for two others is a possibility.

Handling weights

Finally, let us consider weighted variants of the exchange problem: given G , a weight function $\omega : V(G) \rightarrow \mathbb{N}$ and a feasible solution $X \subseteq V(G)$ we want to find a k -exchange (X^+, X^-) that minimises the resulting weight $\omega((X \setminus X^-) \cup X^+)$. We prove only the minimisation variant, the proof for maximisation works exactly the same.

Lemma 49. *Let Π be a weighted minimisation problem that is first-order verifiable by $\phi(S)$ and let G, ω be an instance of Π with a feasible solution X . For every $k \in \mathbb{N}$ we can in time $f(k) \cdot |G| \log w$ find a feasible k -exchange (X^+, X^-) such that $\omega((X \setminus X^-) \cup X^+)$ is minimal among all feasible k -exchanges, where w is the total number of different weights in G .*

Proof. Define the weight function $\omega'(v) = W - \omega(v)$ where $W = \max_{v \in G} \omega(v)$ is the largest weight assigned by ω . Note that minimising $\omega((X \setminus X^-) \cup X^+)$ is equivalent to minimising $\omega(X^+) + \omega'(X^-)$.

Given $\phi(S)$, we construct the formula $\hat{\phi}(\bar{x}^+, \bar{x}^-)$ as in the proof of Lemma 48. Let $R_{\leq \tau}$ be a relation that contains exactly the vertices

$$R_{\leq \tau} := \{v \in G \setminus S \mid \omega(\tau) \leq \tau\} \cup \{v \in S \mid \omega'(\tau) \leq \tau\},$$

and let $\phi_{\leq \tau}(\bar{x}^+, \bar{x}^-)$ be a formula with this relation in its vocabulary derived from $\hat{\phi}$ by adding the additional clause

$$\bigwedge_{x \in \bar{x}^+, \bar{x}^-} R_{\leq \tau} x.$$

Now $\hat{\phi}_{\leq \tau}(\bar{x}^+, \bar{x}^-)$ has a model in G, S exactly if there exists a k -exchange whose vertices all have weight at most τ . Using binary search on τ and the first-order model checking algorithms, we can find this maximal weight \hat{w} in time $g(k)|G| \log w$ for some function g . We then guess which of the $2k - 1$ variables \bar{x}^+, \bar{x}^- are assigned vertices with this maximal weight \hat{w} in a minimal k -exchange and add our guess to the formula $\hat{\phi}$ in the form of clauses $R_{=\hat{w}}x$ for every guessed variable $x \in \bar{x}^+, \bar{x}^-$, obtaining the formula $\hat{\phi}'$. We then repeat the above construction to obtain formulas $\phi'_{\leq \tau'}$ with the constraint that the guessed variables are excluded from the relation $R_{\leq \tau}$.

Since every iteration will fix one more weight of the variables, we have to iterate at most $2k - 1$ times. Taking into consideration our guesses, the whole process has to be repeated at most $2^{(2k-1)^2}$ times and the claimed running time follows. \square

The same process can be easily adapted to nowhere-dense graphs. Since $\log w \leq \log n$, the additional factor even hurts us less since it can be hidden in the running time of the model-checking procedure.

Corollary 9. *Let Π be a weighted optimisation problem that is first-order verifiable and \mathcal{G} be a nowhere dense graph class. Then WEIGHTED Π - k -VERTEX EXCHANGE can be decided in non-uniform almost linear fpt-time on instances from \mathcal{G} .*

In conclusion, structurally sparse classes possess properties that lets us design local-search routines that scale much better than brute-force enumeration. The local-search setting seems to be the perfect playing field for parametrised algorithms and, as we will argue and experimentally verify in the last third of this thesis, a lot of real-world graphs are best modelled by bounded-expansion classes. These algorithms seem to be a great starting point to show that results from the theory of structurally sparse graphs can be applied in practice.

Open question 4. Can we obtain a uniform fpt-algorithm for the k -exchange variant of some practically relevant problem?

Open question 5. Can we improve the running time of the weighted variant for bounded-expansion classes? The first-order enumeration algorithm by Kazana and Segoufin [153] in bounded-expansion classes might provide a good starting point.

10.2 USING LOW-TREEDEPTH COLOURINGS

We will now demonstrate how low-treewidth colourings can be employed to approximate or solve many problems relating to finding small structures in graphs. Given that graphs of bounded expansion admit low-treewidth colourings with only a constant number of colours, these algorithms constitute important building blocks for more involved results and even practical applications (see Section 14.2).

Theorem 24 (Nešetřil, Ossona de Mendez [192, 189]). *Let G, H be graphs and c a $(|H|+1)$ -centred colouring of G . Then one can count*

1. *how often H appears as a (induced) subgraph of G ,*
2. *the number of homomorphisms from H to subgraphs of G , or*
3. *the number of isomorphisms from H to subgraphs of G*

in time $|c(G)|^{O(1)}|G|$.

Proof sketch. Regardless of what type of substructure related to H we want to count, we know that it is contained in subgraphs of G with at most $|H|$ vertices. These subgraphs will receive at most $|H|$ colours by c , hence iterating through all $\binom{|c(G)|}{k}$, $1 \leq k \leq |H|$ colour sets reduces the problem to counting in a graph of bounded treewidth F (and some inclusion-exclusion post processing to recover the correct number). This can be done by basic branching on what vertices of H to map/assign to those of F (cf. ‘Sparsity’ [192]). \square

In our work on applying the theory of structurally sparse graphs to real-world instances we improved this result by developing a faster dynamic programming routine for graphs of bounded treewidth.

Theorem 25 (Demaine *et al.* [65]). *Given graphs G, H and a treewidth decomposition of G of height t , one can count*

1. *how often H appears as a (induced) subgraph of G ,*
2. *the number of homomorphisms from H to subgraphs of G , or*
3. *the number of isomorphisms from H to subgraphs of G*

in time $O(6^h t^h h^2 \cdot |G|)$ and space $O(3^h t^h h t \cdot \log |G|)$, where $h = |H|$.

This immediately implies that the running time in Theorem 24 is at most $O(|c(G)|^{2h} 6^h h^2 \cdot |G|)$.

Centred colourings are very suited to locate small structures in a larger host-graph and we can, for example, use it to approximate problems related to removing specific subgraphs from a host graph.

Corollary 10. *Let \mathcal{F} be a finite set of graphs with $p := \max_{H \in \mathcal{F}} |H|$ and let G a graph with a $(p+1)$ -centred colouring c . Then the problems \mathcal{F} -DELETION and INDUCED- \mathcal{F} -DELETION can be approximated within a factor of $|c(G)|^p$ in time $|c(G)|^p \cdot 2^{O(p)} \cdot |G|$.*

Input: A graph G and an integer r .
Output: A r -dominating set D and $2r + 1$ -scattered set A .

Initialise $D \leftarrow \emptyset$ and $A \leftarrow \emptyset$
 Compute a dtf-augmentations \vec{G}_{2r} of G
 ① Calculate dominating set
for $v \in V(G) \setminus N^r[D]$ **do**
 $D \leftarrow D \cup N_{2r}^-(v) \cup \{v\}$
 $A \leftarrow A \cup \{v\}$
 ② Construct auxiliary graph
 $H \leftarrow (A, \emptyset)$ **for** $u, v \in A$ **do**
 if $\text{dist}_G(u, v) \leq 2r + 1$ **then**
 $E(H) \leftarrow E(H) \cup \{uv\}$
 ③ Construct scattered set
 Colour H with $(\nabla_0(H) + 1)$ colours
 Let $S \subseteq A$ be the largest colour class
return D and S

Algorithm 1: Approximating r -DOMINATING SET in linear time in graphs of bounded expansion.

Proof. Since every forbidden (induced) subgraph receives at most p colours, we can solve the problem—using dynamic programming—for every collection of subgraph induced by at most p colours. Since c is $(p+1)$ -centred, these subgraphs have treedepth at most p . Taking the union of all local solutions gives us an approximate solution that is at most $|c(G)|^p$ times larger than an optimal solution. The time taken on each subgraph is single-exponential, hence the above claimed running time follows. \square

For nowhere dense classes this results implies $|G|^\varepsilon$ -approximations in time $2^{O(p)} \cdot |G|^{1+\varepsilon}$.

10.3 APPROXIMATING r -DOMINATION: DVOŘÁK'S ALGORITHM

The following Lemma was proved by Dvořák in [79] using weak colouring numbers. We prove it here using dtf-augmentations instead to obtain a more direct bound in terms of the topological grad. This bound will help use to obtain tighter bounds in Chapter 13 where we prove that DOMINATING SET has a linear kernel in graphs of bounded expansion. As part of the kernelisation routine, we will need to invoke Dvořák's algorithm to find large 2-scattered sets, *i.e.* obstructions towards having a small domination number.

Lemma 50 (cf. [79]). *Let \mathcal{G} be a graph class of bounded expansion. For every integer r there exists a constant c_r such that for every graph $G \in \mathcal{G}$, we can*

in linear time construct an r -dominating set $D \subseteq G$ and a $(2r+1)$ -scattered set $S \subseteq D$ of size at least $|S| \geq |D|/c_r$.

Proof. Consider Algorithm 1. Note that we can calculate the $2r^{\text{th}}$ dtf-augmentation \vec{G}_{2r} with weights ω_{2r} in linear time (cf. Theorem 16). Then by Lemma 26, for every pair of vertices uv with $\text{dist}_G(u, v) \leq 2r$, we either have that uv or $vu \in \vec{G}_{2r}$ with weight $\omega_{2r}(uv) = \text{dist}_G(u, v)$; or there exists a third vertex $w \in N_{2r}^-(u) \cap N_{2r}^-(v)$ such that $\omega_{2r}(wu) + \omega_{2r}(vw) = \text{dist}(u, v)$.

Consider the sequence of sets D_0, D_1, \dots and A_0, A_1, \dots as they are constructed by the loop in Step 1, i.e. $A_0, D_0 = \emptyset$ and A_i, D_i are the sets A, D after the i^{th} iteration. Let further v_1, v_2, \dots be the sequence of vertices added to A ; hence $A_i = A_{i-1} \cup \{v_i\}$. Note that Step 1 clearly terminates after at most $|G|$ iterations since in every loop at least one vertex is added to D .

Claim. For every distinct pair of vertices v_i, v_j with $i < j$ and for every $w \in N_{2r}^-(v_i) \cap N_{2r}^-(v_j)$ we have that $\omega_{2r}(wv_j) \geq r + 1$.

Assume the contrary, i.e. there exists such a vertex w with $\omega_{2r}(wv_j) \leq r$. Now $D_i \subseteq D_{j-1}$ and by construction $N_{2r}^-(v_i) \subseteq D_i$. Since the vertex $w \in N_{2r}^-(v_i)$ and $\text{dist}_G(w, v_j) \leq \omega_{2r}(wv_j) \leq r$, it follows that w dominates v_j . This contradicts that v_j was not dominated by D_{j-1} at step j and proves the claim.

Claim. For every vertex v_j and $w \in N_{2r}^-(v_j)$ with $\omega_{2r}(wv_j) \geq r + 1$ there exists at most one vertex $v_i, i < j$ such that $w \in N_{2r}^-(v_i)$ and $\omega_{2r}(wv_i) \leq r$.

Assume otherwise: let $v_h, v_i, h < i$, be vertices both connected to w with an in-arc of weight at most r . Then $\text{dist}_G(w, v_h), \text{dist}_G(w, v_i) \leq r$ and hence w already r -dominates v_i before step i . Contradiction.

Claim. The auxiliary graph H is $2(\Delta^-(\vec{G}_{2r}) + \nabla_0(\vec{G}_{2r}))$ -degenerate.

We focus on those edges $uv \in H$ that are not present in G_{2r} . By the above properties of \vec{G}_{2r} , these edges occur exactly between those vertices $v_i v_j \in A$ that are connected via a third vertex $w \in N_{2r}^-(v_i) \cap N_{2r}^-(v_j)$. As proved in the previous claim, v_j can have at most $N_{2r}^-(v_j)$ vertices preceding it that are connected via such a third vertex. We conclude that the number of edges present in H but not present in G_{2r} is bounded by $|A| \cdot \Delta^-(\vec{G}_{2r})$. By Lemma 8, the graph H is therefore $2(\Delta^-(\vec{G}_{2r}) + \nabla_0(\vec{G}_{2r}))$ -degenerate.

Finally, since S is by construction an independent set in H , it follows that S is $(2r + 1)$ -scattered in G . The size of S is at least

$$\begin{aligned} 2|S| &\geq |H| / (\Delta^-(\vec{G}_{2r}) + \nabla_0(\vec{G}_{2r})) \\ &\geq |A| / (\Delta^-(\vec{G}_{2r}) + \nabla_0(\vec{G}_{2r})) \\ &\geq |D| / ((\Delta^-(\vec{G}_{2r}) + \nabla_0(\vec{G}_{2r})) \Delta^-(\vec{G}_{2r})). \end{aligned}$$

Choosing $c_r := 2(\Delta^-(\vec{G}_{2r}) + \nabla_0(\vec{G}_{2r})) \Delta^-(\vec{G}_{2r})$ and noting that by Theorem 16 the function $\Delta^-(\vec{G}_{2r})$ is a function of $2r$ and $\vec{V}_{2r}(G)$ then proves the claim. \square

The following corollary rephrases this result in a more algorithmic way and foreshadows how we will apply it later on.

Corollary 11. *For every graph G and integers r, k we can either compute a dominating set of size $c_r k$ in time $O(c_r \|G\|)$ or conclude that $\mathbf{ds}(G) > k$ by obtaining a $(2r + 1)$ -scattered set of size $k + 1$ as a witness. The ratio c_r depends only on $\nabla_0(G_{2r})$, where \vec{G}_{2r} is the $2r^{\text{th}}$ dtf-augmentation of G .*

In Chapter 13 we will need to solve an annotated domination problem, in which we are given a subset of vertices that need to be dominated. We can use Dvořák's approximation algorithm as a black box to achieve that. The variable c_r is the same as in the proof of Lemma 50.

Lemma 51. *For every graph G , vertex set $Z \subseteq V(G)$ and integers r, k we can either compute in time $O(c_r \|G\|)$ a Z -dominator of size $c_r k$ or conclude that $\mathbf{ds}(G, Z) > k$ by obtaining a $(2r + 1)$ -scattered set $S \subseteq Z$ of size $k + 1$ as a witness.*

Proof. We construct a graph G^* from G as follows: add two labelled vertices $a, x \notin G$ to the graph and connect every vertex $v \in G \setminus Z \cup \{x\}$ to a by a path P_v of length r (with endpoints v and a). In the case of $r = 1$, this construction adds a universal vertex a to $G \setminus Z$ and adds a pendant vertex x to a .

We apply Corollary 11 with parameter $k + 1$ to G^* . First assume that the algorithm returns a r -dominating set D of G^* with $|D| \leq c_r(k + 1)$.

Claim. *There exists an r -dominating set D' of G^* that satisfies*

- $|D'| \leq |D|$,
- $a \in D'$, and
- $P_v \cap D' \subseteq \{a, v\}$ for every $v \in V(G) \setminus Z \cup \{x\}$.

We start out with $D' = D$ and modify it until it has the desired properties. Assume $a \notin D'$: then there must exist a vertex $v \in D' \cap P_x$ since otherwise x would not be dominated. We set $D' \leftarrow D' \setminus \{v\} \cup \{a\}$, this neither changes the size nor the fact that D' is an r -dominating set of G^* . Now assume there exists a vertex $v \in V(G) \setminus Z$ such that there is $u \in P_v \cap D'$ with $u \notin \{v, a\}$. Since $a \in D'$, all vertices of P_v are already dominated and we can set $D' \leftarrow D' \setminus P_v \cup \{v, a\}$. This at most decreases the size of D' and keeps its domination property. Applying this procedure for every offending vertex $v \in V(G) \setminus Z$ results in a set D' with the claimed properties.

Since D' is in particular a Z -dominator of the claimed size we have proved the first part of the lemma. Let us now assume that the algorithm returns a $(2r + 1)$ -scattered set S of size $k + 2$ (recall that we called it with parameter $k + 1$).

Claim. *There exists a $(2r + 1)$ -scattered set $S' \subseteq Z$ with $|S'| = k + 1$.*

Simply observe that the subgraph $G^* \setminus Z$ has diameter $2r$, therefore at most one vertex of it can be contained in the scattered set. Removing this offending vertex from S yields the set S' with the desired properties. This proves the claim and concludes the proof of the lemma. \square

It is worthwhile to calculate what the approximation factor c_1 for DOMINATING SET turns out to be. By Lemma 34, we have for every graph G there exist a dtf-augmentation $\vec{G}_1, \vec{G}_2, \dots$ such that

$$\Delta^-(\vec{G}_d) \leq \Delta^-(\vec{G}_{d-1})^2 + 12\Delta^-(\vec{G}_{d-1})^3 \cdot \tilde{\nabla}_{1/2}(G_{d-1})$$

which for the case of $d = 2$ implies that

$$\begin{aligned} \Delta^-(\vec{G}_2) &\leq \Delta^-(\vec{G}_1)^2 + 12\Delta^-(\vec{G}_1)^3 \cdot \tilde{\nabla}_{1/2}(G) \\ &\leq \nabla_0(G)^2 + 12\nabla_0(G)^3 \cdot \tilde{\nabla}_{1/2}(G) \end{aligned}$$

where the last inequality follows from our assumption that \vec{G}_1 is obtained by orienting the edges of G according to its degeneracy-order. Hence we have that c_1 is at most

$$\begin{aligned} c_1 &\leq 2(\Delta^-(\vec{G}_2) + \nabla_0(G_2))\Delta^-(\vec{G}_2) \leq 6\Delta^-(\vec{G}_2)^2 \\ &\leq 2^{10}\nabla_0(G)^6\tilde{\nabla}_{1/2}(G)^2. \end{aligned}$$

where we used the bound $\tilde{\nabla}_0(G_2) \leq 2\Delta^-(\vec{G}_2)$ from Lemma 9 and our usual assumption that $\nabla_0(G) \geq 1$. While this constant is not absurdly large, it clearly lies more in the theoretical than the practical realm. However, it is quite probable that our worse-case assumptions are far off the mark when it comes to practical instances.

10.4 COMPUTING NEIGHBOURHOOD SIZES

Nešetřil and Ossona de Mendez noted that tf-augmentations can carry information about local distances in graphs, an observation that inspired the dtf-augmentations in Chapter 7.

Lemma 52 (cf. [192]). *For every graph G and every integer d , one can compute a data structure in time $\tilde{\nabla}_d(G)^{2^{O(d)}}|G|$ that can return for every pair $u, v \in G$ either their distance $\text{dist}_G(u, v)$ or ascertain that it is larger than d in time $\tilde{\nabla}_d(G)^{2^{O(d)}}$.*

Proof. The complexity of calculating the d^{th} dtf-augmentation crucially depends on the maximum in-degree, which by Theorem 16 can be bounded by

$$\Delta^-(\vec{G}_d) = \tilde{\nabla}_d(G)^{2^{O(d)}}.$$

By Lemma 26, for every pair of vertices uv with $\text{dist}_G(u, v) \leq d$, we either have that uv or $vu \in \vec{G}_d$ with weight $\omega_d(uv) = \text{dist}_G(u, v)$; or there exists a third vertex $w \in N_d^-(u) \cap N_d^-(v)$ such that $\omega_d(wu) +$

$\omega_d(vw) = \text{dist}(u, v)$. In the former case, we can simply return the arc-weight as the correct distance. In the latter case, we need to consider all vertices in the intersection $N_d^-(u) \cap N_d^-(v)$. By the above bound on the maximum in-degree, this takes time $\tilde{\nabla}_d(G)^{2^{O(d)}}$. \square

In this section, we ask whether this idea can be taken further: can we compute a data structure that for a query-vertex v returns *how many* vertices have distance r from v ? In other words, for a fixed r we want to compute a data structure in linear time that tells us the neighbourhood-sizes $|N^i(\bullet)|$ for any query-vertex in constant time. This is indeed the case, and we can even generalise it by allowing weights on the vertices of our input graph. The proof of the following theorem will be the scope of the remaining section.

Theorem 26. *Let \mathcal{G} be a class of bounded expansion. Then for $G \in \mathcal{G}$ with vertex-weights α we can, in linear time, compute a table C such that*

$$C[v][d] = \sum_{u \in N^d(v)} \alpha(u).$$

In the following we fix a graph G weighted by the function α and a dtf-augmentation $\vec{G}_1, \vec{G}_2, \dots$. For simplicity, we write $N_r^-(v)$ instead of the unwieldy $N_{\vec{G}_r}^-(v)$. We will assume that the weights assigned by α are polynomial in $|G|$ and hence addition and subtraction of all values are assumed to take constant time.

Definition 19. Let $\vec{G}_1, \vec{G}_2, \dots$ be a dtf-augmentation of G and let α be vertex-weights. Fix a vertex $v \in G$ and let $\emptyset \neq X \subseteq N_r^-(v)$ be subset of v 's in-neighbourhood and let $\vec{d} \in [r]^{|X|}$ be a distance vector of matching length. We define

$$N(v, X, \vec{d}) := \{u \in V(G) \mid N_r^-(v) \cap N_r^-(u) = X \text{ and } \text{dist}(u, X) = \vec{d}\}$$

as those vertices whose in-neighbourhood in \vec{G}_r overlap with the in-neighbourhood of v in exactly X and whose distance-vector to X is exactly \vec{d} . We further define the *query-function* c_α is defined as

$$c_\alpha(v, X, \vec{d}) := \sum_{u \in N(v, X, \vec{d})} \alpha(u).$$

Lemma 53. *Let $\Delta = \Delta^-(\vec{G})$. One can compute a data structure R in time $O(2^\Delta n)$ such that queries $c_\alpha(v, X, \vec{d})$ can be answered in time $O(\Delta 2^\Delta)$ by R .*

Proof. We define an auxiliary dictionary R indexed by vertex sets X for which a vertex v exists with $X \subseteq N_r^-(v)$. At each entry $v \in \vec{G}_r$, we will store another dictionary indexed by distance vectors which in turn stores a simple counter. We initialise R as follows: for every $v \in \vec{G}_r$, $X \subseteq N_r^-(v)$ and every distance vector $\vec{d} \in [r]^{|X|}$, set $R[X][\vec{d}] = 0$. Note that in total, R contains at most $2^{\Delta^-(\vec{G}_r)} n$ entries.

We can implement R as a hash-map to achieve the desired (expected) constant-time for insertion and look-up, though this would yield

a randomised algorithm. A possible way to implement R on a RAM deterministically is the following: We store the key $X = \{x_1, x_2, \dots, x_p\}$ at address $x_1 + n \cdot x_2 + \dots + n^p \cdot x_p$. This uses addresses up to size n^c , for some constant c , but since we only insert $O(n)$ keys the set-up takes only linear time. Our later queries to R will be restricted to keys that are guaranteed to be contained in the dictionary, thus we will never visit a register that has not been initialised.

Now, for every $v \in G$, every $X \subseteq N_r^-(v)$, we increment the counter $R[X][\text{dist}(v, X)]$ by $\alpha(v)$:

```

for  $v \in G$  do
    for  $X \subseteq N_r^-(v)$  do
         $R[X][\text{dist}(v, X)] \leftarrow R[X][\text{dist}(v, X)] + \alpha(v)$ 
    
```

We now claim that queries of the form $c_\alpha(v, X, \bar{d})$ can be computed using inclusion-exclusion as follows:

$$c_\alpha(v, X, \bar{d}) = \sum_{X \subseteq Y \subseteq N_r^-(v)} (-1)^{|Y \setminus X|} \sum_{\bar{d}' : \bar{d}'|_X = \bar{d}} R[Y][\bar{d}'].$$

Here $\bar{d}'|_X$ denotes the restriction of the distance-vector \bar{d}' to those entries that correspond to the vertices in X . Computing this value is possible in time $O(|N_r^-(v)| \cdot 2^{|N_r^-(v)|})$ using fast Möbius inversion. Let us now prove that it indeed computes the quantity $c_\alpha(v, X, \bar{d})$.

First, consider a vertex $u \in \vec{G}_r$, such that $N_r^-(u) \cap N_r^-(v) = X$ and $\text{dist}(u, X) = \bar{d}$. Now since $\alpha(u)$ is not counted by $R[Y][\cdot]$ with $Y \supsetneq X$, it appears exactly once in the above sum when $Y = X$.

Second, consider a vertex $u \in \vec{G}_r$ such that $\text{dist}(u, X) \neq \bar{d}$. The weight of such a vertex is not counted by the above sum, since $\alpha(u)$ is only counted in entries of R that do not occur as summands.

Finally, consider a vertex $u \in \vec{G}_r$ with $N_r^-(u) \cap N_r^-(v) = Z$ where $X \subsetneq Z \subseteq N_r^-(v)$ such that $\text{dist}(u, X) = \bar{d}$. The weight of this vertex is counted in each term of

$$\sum_{X \subseteq Y \subseteq Z} (-1)^{|Y \setminus X|} R[Y][\text{dist}(u, Z)|_Y],$$

and since

$$\sum_{X \subseteq Y \subseteq Z} (-1)^{|Y \setminus X|} = \sum_{0 \leq k \leq n} (-1)^k \binom{n}{k} = 0,$$

we know that the signs cancel out and thus $\alpha(u)$ does not contribute to $c_\alpha(v, X, \bar{d})$. We conclude that the inclusion-exclusion formula above computes the query $c_\alpha(v, X, \bar{d})$, as claimed. \square

Given the auxiliary data structure R we can now compute the weights of all r -neighbourhoods.

Lemma 54. *Given R , one can compute a table C with*

$$C[v][d] = \sum_{w \in N^d(v)} \alpha(w) \text{ for } v \in G, d \leq r$$

in time $O(\Delta(4r)^\Delta |G|)$, where $\Delta = \Delta^-(\vec{G}_r)$.

Proof. Let $\alpha_d(v) := \sum_{w \in N^d(v)} \alpha(w)$, where $v \in G$ and $1 \leq d \leq r$, for convenience. To compute the quantity $\alpha_d(v)$ for all vertices and distances smaller than r , we initialise a table C with $C[v][d] = 0$ for every $v \in G, 1 < d \leq r$.

Now we update the table by the following algorithm:

```

for  $v \in G$  do
  for  $X \subseteq N_r^-(v)$  do
    for  $\bar{d} \in [r]^{|X|}$  do
       $d_{min} \leftarrow \min(\bar{d} + \text{dist}(v, X))$ 
      if  $d_{min} \leq r$  then
         $C[v][d_{min}] \leftarrow C[v][d_{min}] + c_\alpha(v, X, \bar{d})$ 
       $d_{self} = 2 \min(\text{dist}(v, X))$ 
      if  $d_{self} \leq r$  then
         $C[v][d_{self}] \leftarrow C[v][d_{self}] - 1$ 

```

The number of iterations of this update is $O((2r)^\Delta |G|)$. Each iteration can query R to calculate c_α , hence the total time taken is $O(\Delta(4r)^\Delta |G|)$.

At this point, $C[v][d]$ contains (by the definition of c_α) the sum of weights of vertices u for which $\min(\text{dist}(v, X) + \text{dist}(u, X)) = d$ for non-empty $X = N_r^-(v) \cap N_r^-(u)$: We have counted the ‘indirect’ neighbours in \vec{G}_r .

By Lemma 26, every pair of vertices of distance at most r in G either is connected by an arc or they share a common in-neighbour in \vec{G}_r . To count these ‘direct’ neighbours, we update C as follows (recall that ω_r is the arc-weight function of \vec{G}_r):

```

for  $uv \in \vec{G}_r$  do
  if  $N_r^-(u) \cap N_r^-(v) = \emptyset$  then
    ① Not counted yet
     $C[v][\omega(uv)] \leftarrow C[v][\omega(uv)] + \alpha(u)$ 
     $C[u][\omega(uv)] \leftarrow C[u][\omega(uv)] + \alpha(v)$ 
  else if  $vu \notin \vec{G}_r$  or  $u < v$  then
    ② Wrong distance
     $d_{wrong} = \min(\text{dist}(u, X) + \text{dist}(v, X))$ 
     $C[v][d_{wrong}] \leftarrow C[v][d_{wrong}] - \alpha(u)$ 
     $C[u][d_{wrong}] \leftarrow C[u][d_{wrong}] - \alpha(v)$ 
     $C[v][\omega(uv)] \leftarrow C[v][\omega(uv)] + \alpha(u)$ 
     $C[u][\omega(uv)] \leftarrow C[u][\omega(uv)] + \alpha(v)$ 

```

Vertices for which block ① is executed do not share common in-neighbours, hence before the execution of the above algorithm the weight of u was not counted in $C[v][\bullet]$ and vice-versa.

Vertices for which block ② is executed were already counted in C , however, their distance according to the common in-neighbours might be larger than the actual distance given by ω_r . We therefore correct the

counts in C ; note that if the distance was correct the above adjustments cancel out. To avoid applying the correction twice in case \vec{G}_r contains both arcs uv and vu , we introduce the condition $u < v$ to enter the block. At this point, $C[v][d]$ contains the sum of weights of vertices $u \neq v$ for which either

- $uv \notin \vec{G}_r$, the intersection $X = N_r^-(v) \cap N_r^-(u)$ is not empty and $\min(\text{dist}(v, X) + \text{dist}(u, X)) = d$, or
- $uv \in \vec{G}_r$ and $\omega_r(uv) = d$.

It follows from Lemma 26 that $C[v][d] = \alpha_d(v)$ for $d < r$ and $v \in G$. The claimed running time follows by seeing that the second algorithm takes time $O(\|\vec{G}_r\|) = O(\Delta|G|)$. \square

Now Theorem 26 follows directly from the just proved Lemma 53 and the bound $\Delta^-(\vec{G}_d) = \tilde{V}_d(G)^{2^{O(d)}}$ given by Theorem 16. In practice, the dependence on $\Delta^-(\vec{G}_d)$ is much easier to measure and certainly better than our worse-case upper bound.

Some practical test (executed by our student Xinyu Ge in his master thesis) showed that while the above algorithm on its own is too slow in practice—the quantity $\Delta^-(\vec{G}_d)$ is rather high, but only very few vertices attain the maximum—a heuristic based on it has good chances of being faster than truncated bfs from every vertex of the graph for very large instances.

Open question 6. Can we remove the exponential dependence on $\Delta^-(\vec{G}_d)$ in order to extend Theorem 26 to nowhere-dense classes?

*It seems that perfection is attained not when there is nothing more to add,
but when there is nothing more to remove.*

— Antoine de Saint Exupéry, *Terre des Hommes*

The first steps towards a kernelisation meta-theorem appeared in a paper by Guo and Niedermeier who provided a framework to design linear kernels on planar graphs for graph problems which satisfy a certain distance property [133]. Their work built on the seminal paper by Alber, Fellows, and Niedermeier who showed that DOMINATING SET has a linear kernel on planar graphs [3]. This was followed by the first true meta-theorem in this area by Bodlaender, Fomin, Lokshantov, Penninkx, Saurabh and Thilikos [23] who showed that graph problems that have finite integer index (f.i.i., defined below) and satisfy a property called *quasi-coverable*¹, admit linear kernels on bounded genus graphs.

Shortly after the meta-kernelisation paper was published, Fomin et al. [101] proved a meta-theorem for linear kernels on classes excluding a minor, a graph class that strictly contains graphs of bounded genus. A rough statement of their main result states that any graph problem that has finite integer index, is *contraction bidimensional*, and satisfies a *separation property* has a linear kernel on graphs that exclude a fixed graph as minor.

Keeping the structurally sparse hierarchy in mind, we ask the natural question: can the same be achieved for graph classes excluding a topological minor? Can we go even higher? We answer these questions in the positive, with some caveats. After a revision of the important properties like finite integer index (Section 11.1), well-quasi orders (Section 11.2) and protrusions (Section 11.3) we prove the following two extensions of the meta-kernel framework. First, we obtain the following for classes defined by an excluded topological minor.

Theorem 27. *Let Π be a graph problem that has f.i.i. on graphs of bounded treewidth and let \mathcal{G} be a graph class that excludes a fixed graph H as a topological minor. For every $d \in \mathbb{N}$ there exists an algorithm that takes as input $(G, \xi) \in \mathcal{G} \times \mathbb{N}$ and outputs in time $O(|G|)$ an instance (G', ξ') such that*

1. $(G, \xi) \in \Pi \iff (G', \xi') \in \Pi$ with $\xi' \leq \xi$,
2. $H \not\leq_t G'$, and
3. $|G'| = O(\mathbf{tw}_d^\Delta(G))$.

¹ This property was called *quasi-compactness* in the earlier version of their paper.

For example, the problems FEEDBACK VERTEX SET, CHORDAL VERTEX DELETION INTERVAL VERTEX DELETION, PROPER INTERVAL VERTEX DELETION, COGRAPH VERTEX DELETION, and EDGE DOMINATING SET all have the property that they bound the treewidth of the graph surrounding the solution². There are two interpretations of the above result: one is that it holds only for problems that are *treewidth-bounding*, that is, a small solution guarantees the existence of small treewidth-modulator. The other, and in hindsight more useful interpretation, is that a large number of problems—those that have finite integer index—admit linear kernels in classes excluding a topological minor if we *parametrise* them by the size of a small treewidth-modulator. Further discussion on this shift to a *structural* parametrisation can be found in Section 12.2.

Larger classes

This structural interpretation of our result enables us to continue the programme: using a modulator to constant *treedepth*.

Theorem 28. *Let Π be a graph problem that has f.i.i. on graphs of bounded treedepth. Let \mathcal{G} be a graph class of bounded expansion and let $t \in \mathbb{N}$ be a constant. Then there is an algorithm that takes as input $(G, \xi) \in \mathcal{G} \times \mathbb{N}$ and, in time $O(|G| + \log \xi)$, outputs (G', ξ') such that*

1. $(G, \xi) \in \Pi$ if and only if $(G', \xi') \in \Pi$;
2. G' is an induced subgraph of G ; and
3. $|G'| = O(\mathbf{td}_t^\Delta(G))$.

Theorem 29. *Let Π be a graph problem that has f.i.i. on graphs of bounded treedepth. Let \mathcal{G} be a nowhere dense graph class and let $t \in \mathbb{N}$ be a constant. Then there is an algorithm that takes as input $(G, \xi) \in \mathcal{G} \times \mathbb{N}$ and, in time $O(|G|^{1+o(1)})$, outputs (G', ξ') such that*

1. $(G, \xi) \in \Pi$ if and only if $(G', \xi') \in \Pi$;
2. G' is an induced subgraph of G ; and
3. $|G'| = O(\mathbf{td}_t^\Delta(G)^{1+o(1)})$.

We immediately have that DOMINATING SET, r -DOMINATING SET, CONNECTED DOMINATING SET, CONNECTED VERTEX COVER, FEEDBACK VERTEX SET, CONNECTED FEEDBACK VERTEX SET, INDEPENDENT SET, among others, admit linear kernel in bounded expansion classes and almost-linear kernels in nowhere-dense classes when parametrised by a tree-depth-modulator. For CONNECTED VERTEX COVER, this also implies a linear kernel (almost linear kernel) using its natural parametrisation—since this problem does not admit a polynomial kernel in general graphs unless $\text{NP} \subseteq \text{coNP}/\text{poly}$ [69], we have another example of structural sparseness improving tractability.

The introduction of a structural parameter also makes it possible to include problems like HAMILTONIAN PATH and HAMILTONIAN CYCLE who simply lack a suitable natural parameter.

Further improvements

² For CHORDAL VERTEX DELETION this holds because we work in a class with bounded clique number.

We also extend the machinery in two further directions. By proving these results in such a way that it is enough for a problem to have finite integer index only when restricted to graphs of bounded treewidth/treedepth, we expand on the number of problems for which our results hold. For example, the above two theorems are also applicable to k -PATH, k -CYCLE, EXACT s, t -PATH, and EXACT CYCLE, problems that do not have finite integer index in general. We provide a proof of this fact in Section 11.1.2. Further, we proved that TREEWIDTH has finite integer index if restricted to graphs of bounded treewidth, PATHWIDTH for graphs of bounded pathwidth, and BRANCHWIDTH for graphs of bounded branchwidth [114, 116, 115]. As a consequence, these problems admit kernels if parametrised by a treedepth-modulator in structurally sparse classes; in contrast, note that the smallest known kernel for a structural parametrisation of TREEWIDTH in general graphs is $O(\mathbf{vc}(G)^3) = O(\mathbf{td}_1^\Delta(G)^3)$ [24].

As a second improvement to the meta-kernelisation framework, we address the problem of protrusion replacement for graph classes that themselves do not have finite index by using well-quasi ordering relations. In the context of nowhere dense classes, our algorithm returns an induced subgraph of the input graph. Consequently, the only additional condition we need to impose on the target graph class is that it is hereditary. The same idea can be applied to the kernelisation of planar, bounded-genus, and classes excluding a minor: the resulting graph is then a minor of the input graph (note that while planarity is a property with finite index, there are planar graph classes that do not have finite index).

Let us state the consequences of the above theorems and smaller results in the rest of the chapter succinctly.

Corollary 12. *The following graph problems have finite integer index in general graphs, and hence admit*

- *linear kernels in classes excluding a topological minor when parametrised by a modulator to constant treewidth,*
- *linear kernels in hereditary bounded-expansion classes when parametrised by a modulator to constant treedepth, and*
- *almost linear kernels in hereditary nowhere-dense classes when parametrised by a modulator to constant treedepth.*

DOMINATING SET, r -DOMINATING SET, EDGE DOMINATING SET, CONNECTED DOMINATING SET, CONNECTED VERTEX COVER, INDEPENDENT SET, FEEDBACK VERTEX SET, EFFICIENT DOMINATING SET, HAMILTONIAN PATH, HAMILTONIAN CYCLE, INDUCED MATCHING, CHORDAL VERTEX DELETION, ODD CYCLE TRANSVERSAL, INDUCED d -DEGREE SUBGRAPH, MIN LEAF SPANNING TREE, MAX FULL DEGREE SPANNING TREE.

wm- w -DELETION

Consider the following generic parametrised problem: fix a constant w and a with-measure **wm**. For a given graph G and budget k , can we find a set $X \subseteq V(G)$ of size at most k such that $\mathbf{wm}(G \setminus X) \leq w$?

In graph classes excluding a topological minor, Fomin, Oum, and Thilikos demonstrated that the width-measures \mathbf{cw}, \mathbf{rw} , and \mathbf{tw} are all linearly related to each other [104]. The same is true for \mathbf{bw} and \mathbf{tw} in general graphs, as shown by Robertson and Seymour [214]. Together with Theorem 27 we therefore obtain the following.

Corollary 13. \mathbf{wm} - w -DELETION admits a linear kernel in graphs excluding a topological minor, where \mathbf{wm} is either $\mathbf{tw}, \mathbf{rw}, \mathbf{cw}$ or \mathbf{bw} .

11.1 FINITE INTEGER INDEX

Recall that for a graph class \mathcal{G} we write ${}^\circ\mathcal{G}_t$ to denote all t -boundaried graphs whose underlying graphs are contained in \mathcal{G} .

Finite index, f.i.

Definition 20 (Finite index, f.i.). Let \mathcal{P} be a graph property and let ${}^\circ\mathcal{G}_1, {}^\circ\mathcal{G}_2$ be two t -boundaried graphs. We write ${}^\circ\mathcal{G}_1 \equiv_{\mathcal{P}} {}^\circ\mathcal{G}_2$ if for all t -boundaried graphs ${}^\circ H$ it holds that

$${}^\circ\mathcal{G}_1 \oplus {}^\circ H \in \mathcal{P} \iff {}^\circ\mathcal{G}_2 \oplus {}^\circ H \in \mathcal{P}.$$

We say that \mathcal{P} has *finite index in the class \mathcal{G}* if, for every $t \in \mathbb{N}$, the quantity $|{}^\circ\mathcal{G}_t / \equiv_{\mathcal{P}}|$ is finite.

Crucially, all properties that are CMSO-definable have finite index [74]. However, the notion of finite index is yet too weak to apply it to decision problems. Bodlaender and van Antwerpen-de Fluiter introduced the property *finite integer index* in the context of reducing graphs of low treewidth [26] which forms a cornerstone of the meta-kernelisation framework. We adapt this notion slightly by allowing the finiteness-condition to hold only in a certain graph class instead of all graphs.

Finite integer index, f.i.i.

Definition 21 (Finite integer index; f.i.i.). Let Π be a graph problem and let ${}^\circ\mathcal{G}_1, {}^\circ\mathcal{G}_2$ be two t -boundaried graphs. We write ${}^\circ\mathcal{G}_1 \equiv_{\Pi} {}^\circ\mathcal{G}_2$ if there exists an integer $\Delta_{\Pi}({}^\circ\mathcal{G}_1, {}^\circ\mathcal{G}_2)$ such that for all t -boundaried graphs ${}^\circ H$ and for all $\zeta \in \mathbb{N}$ it holds that

$$({}^\circ\mathcal{G}_1 \oplus {}^\circ H, \zeta) \in \Pi \iff ({}^\circ\mathcal{G}_2 \oplus {}^\circ H, \zeta + \Delta_{\Pi}({}^\circ\mathcal{G}_1, {}^\circ\mathcal{G}_2)) \in \Pi.$$

We say that Π has *finite integer index in the class \mathcal{G}* if, for every $t \in \mathbb{N}$, the quantity $|{}^\circ\mathcal{G}_t / \equiv_{\Pi}|$ is finite.

Note that the constant $\Delta_{\Pi}({}^\circ\mathcal{G}_1, {}^\circ\mathcal{G}_2)$ depends on Π, t , and the *ordered pair* $({}^\circ\mathcal{G}_1, {}^\circ\mathcal{G}_2)$ so that $\Delta_{\Pi}({}^\circ\mathcal{G}_1, {}^\circ\mathcal{G}_2) = -\Delta_{\Pi}({}^\circ\mathcal{G}_2, {}^\circ\mathcal{G}_1)$. We point out that the *restriction* of \equiv_{Π} to a class \mathcal{G} does *not* mean that the graph ${}^\circ H$ in the above definition is contained in \mathcal{G} .

Ⓜ

For a graph problem Π and a graph property \mathcal{P} , we denote by $\Pi \pitchfork \mathcal{P}$ the YES-instances $(G, \bullet) \in \Pi$ where additionally $G \in \mathcal{P}$. Since graph properties and classes are exchangeable, the same notation can be used to restrict a problem to instances of a certain class.

The following basic observations demonstrate how properties with finite index interact with each other, with graph problems that have

finite integer index, and with graph classes. The proofs are simple arguments about refinements and we omit them here.

Observation 2. Let \mathcal{P}, \mathcal{Q} be graph properties that have finite index on a graph class \mathcal{G} . Then the property $\mathcal{P} \cap \mathcal{Q}$ has finite index on \mathcal{G} .

Observation 3. Let Π be graph problem that has finite integer index on a graph class \mathcal{G} and let \mathcal{P} be a graph property that has finite index on \mathcal{G} . Then the equivalence relation $\equiv_{\Pi \cap \mathcal{P}}$ defined as

$${}^\circ G_1 \equiv_{\Pi \cap \mathcal{P}} {}^\circ G_2 \iff {}^\circ G_1 \equiv_{\Pi} {}^\circ G_2 \text{ and } {}^\circ G_1 \equiv_{\mathcal{P}} {}^\circ G_2$$

has finite integer index on \mathcal{G} .

Observation 4. If \mathcal{P} has finite index on \mathcal{G} , then \mathcal{P} has finite index on every subclass $\mathcal{G}' \subseteq \mathcal{G}$. If Π has finite integer index on \mathcal{G} , then Π has finite integer index on every subclass $\mathcal{G}' \subseteq \mathcal{G}$.

A great tool provided by the original meta-kernelisation paper to show that a specific problem has f.i.i. is the notion of *strong monotonicity*: While expressibility in a certain logic does not seem to delineate problems that do have f.i.i. from those that do not, we can at least prove f.i.i. for this subset of problems expressible in min/max-CMSO. We recapitulate the definition for minimization problems here, the maximization variant works analogously.

Strong monotonicity

Definition 22 (Strong monotonicity). A problem that is expressible as a min-CMSO formula ψ is *strongly monotone* if there exists a function f with the following property: for every t -boundaried graph ${}^\circ G$ there must exist a subset $W \subseteq V({}^\circ G)$ such that for every other t -boundaried graph ${}^\circ G'$ and subset $S' \subseteq V({}^\circ G')$ we have that either

$$\forall S : {}^\circ G \oplus {}^\circ G' \not\models \psi(S \cup S'),$$

or we have that

$$\begin{aligned} {}^\circ G \oplus {}^\circ G' &\models \psi(W \cup S') \\ \text{and } |W| &\leq |\arg \min_{S \subseteq V({}^\circ G)} \{ {}^\circ G \oplus {}^\circ G' \models \psi(S \cup S') \}| + f(t). \end{aligned}$$

The intuition for strong monotonicity is that an optimal solution of a problem computed in a boundaried graph can always be extended to an almost optimal solution in a graph obtained from gluing something to it. The ‘loss’ in optimality should only be a function of the boundary size. For example, VERTEX COVER is strongly monotone: by taking all t boundary vertices into a solution, we can always ensure that a local optimal solution can be extended to an *almost* optimal solution since we are now essentially dealing with two independent subinstances.

Another helpful tool is the following result: for finite families of connected graph \mathcal{H} , the problems \mathcal{H} -MINOR DELETION and \mathcal{H} -MINOR PACKING have f.i.i. [23]³.

³ Note that in the original proof it is assumed that \mathcal{H} contains at least one planar graph. This additional property is needed to show quasi-coverability, not f.i.i.

Similarly, let \mathcal{F} be a family of finite connected graphs. Then the problems \mathcal{F} -DELETION and \mathcal{F} -PACKING have f.i.i. if the input instance does not contain a so-called *redundant vertex* [23].

Redundant vertex

In this context, a *redundant vertex* is a vertex which does not belong to any subgraph isomorphic to a graph in \mathcal{F} . A simple preprocessing rule, which removes such vertices, then leaves a modified instance on which the above mentioned problems do indeed have f.i.i. and thus are amenable to the kernelisation framework. We can improve this result slightly using the bounded expansion toolkit: the preprocessing rule can be implemented to run in linear time (or almost linear time in nowhere dense classes).

Lemma 55. *Let \mathcal{F} be a finite set of graph with at most p vertices. Let \mathcal{G} be a graph class. Then we can remove all vertices from a graph $G \in \mathcal{G}$ that are redundant with respect to \mathcal{F} in time $O(|G|)$ if \mathcal{G} has bounded expansion and in time $O(|G|^{1+\epsilon})$ if \mathcal{G} is nowhere dense.*

Proof. By Theorem 17, we can compute a $(p+1)$ -centred colouring of G in linear time or in almost-linear time in the nowhere-dense case. Since every occurrence of a subgraph isomorphic to a graph in \mathcal{F} receives at most p colours, we can then find the irrelevant vertices using Courcelle's Theorem in the subgraphs induced by at most p colours. Since p is a constant, the claimed running time follows. \square

On the other hand, to show that a problem Π does not have f.i.i., we need to find an infinite set of graphs that are pairwise not equivalent. To show in turn that two distinct boundaried graphs ${}^\circ G_1, {}^\circ G_2$ are not equivalent, we need to provide for every offset $\Delta \in \mathbb{N}$ a *distinguishing pair* ${}^\circ H, \zeta$ such that

$$({}^\circ G_1 \oplus {}^\circ H, \zeta) \in \Pi \iff ({}^\circ G_2 \oplus {}^\circ H, \zeta + \Delta) \notin \Pi.$$

We can streamline this procedure by looking at a minimal or maximal solution. Define

$\zeta_{\min}^\Pi(\bullet)$

$$\zeta_{\min}^\Pi(G) = \arg \min_{\zeta \in \mathbb{N}} \{(G, \zeta) \in \Pi\}$$

as the minimal parameter such that G is a yes-instance (we will drop the superscript Π in the following if the problem is clear from the context). Now if indeed ${}^\circ G_1 \not\equiv_{\Pi} {}^\circ G_2$ then there exist two graph ${}^\circ H, {}^\circ H'$ such that

$$\zeta_{\min}({}^\circ G_1 \oplus {}^\circ H) - \zeta_{\min}({}^\circ G_2 \oplus {}^\circ H) \neq \zeta_{\min}({}^\circ G_1 \oplus {}^\circ H') - \zeta_{\min}({}^\circ G_2 \oplus {}^\circ H').$$

In order to show that a problem is not f.i.i. in a graph classes \mathcal{G} , we simply need to take care to choose an infinite set of pairwise not equivalent boundaried graphs whose underlying graphs are contained in \mathcal{G} . Since the graph classes that proved to be useful consist of graphs of bounded treewidth or treedepth, we will focus on classes defined by width-measures in the following.

11.1.1 Selected negative results

Lemma 56. INDEPENDENT DOMINATING SET does not have f.i.i. on graphs of treedepth two.

Proof. Let $\{\circ S_\ell\}_{\ell \in \mathbb{N}}$ be the class of stars with their centre as the single boundary vertex and let $\{\circ S'_\ell\}_{\ell \in \mathbb{N}}$ be the class of stars with a single leaf as their boundary. In both cases, the integer ℓ denotes the number of leaves of the star. The following tableaux shows that the graphs $\circ S_\ell$ and $\circ S_{\ell'}$ for $2 \leq \ell < \ell'$ are not equivalent.

$$\begin{array}{cc} \xi_{\min}(\circ S_\ell \oplus \circ S'_{\ell+1}) = \ell & \xi_{\min}(\circ S_\ell \oplus \circ S'_{\ell'+1}) = \ell + 1 \\ \hline \xi_{\min}(\circ S_{\ell'} \oplus \circ S'_{\ell+1}) = \ell & \xi_{\min}(\circ S_{\ell'} \oplus \circ S'_{\ell'+1}) = \ell' + 1 \\ 0 & \ell - \ell' \end{array}$$

We conclude that INDEPENDENT DOMINATING SET does not have finite integer index on graphs of treedepth two. \square

Lemma 57. k -PATH and TREEDEPTH do not have f.i.i. on graphs of pathwidth one.

Proof. Consider the graph class $\{\circ PP_\ell\}_{\ell \in \mathbb{N}}$ consisting of two paths of length ℓ with each one endpoint in the boundary. Let $\circ P_2$ be a single edge with both endpoints in the boundary. The following tableaux shows that the graphs $\circ PP_\ell$ and $\circ PP_{\ell'}$ for $1 \leq \ell < \ell'$ are not equivalent under k -PATH.

$$\begin{array}{cc} \xi_{\min}(\circ PP_\ell \oplus \circ P_2) = 2\ell + 1 & \xi_{\min}(\circ PP_\ell \oplus \circ PP_\ell) = 2\ell \\ \hline \xi_{\min}(\circ PP_{\ell'} \oplus \circ P_2) = 2\ell' + 1 & \xi_{\min}(\circ PP_{\ell'} \oplus \circ PP_\ell) = \ell + \ell' \\ 2\ell - 2\ell' & \ell - \ell' \end{array}$$

The proof for TREEDEPTH works analogous since a path of length 2^ℓ has treedepth ℓ . \square

Lemma 58. CHROMATIC NUMBER does not have f.i.i. on general graphs.

Proof. The counterexample is the class of complete graphs with an empty boundary. Consider integers $1 \leq \ell < \ell'$ and note that

$$\begin{array}{cc} \chi(K_\ell \cup K_\ell) = \ell & \chi(K_\ell \cup K_{\ell'}) = \ell' \\ \hline \chi(K_{\ell'} \cup K_\ell) = \ell' & \chi(K_{\ell'} \cup K_{\ell'}) = \ell' \\ \ell - \ell' & 0 \end{array}$$

We conclude that CHROMATIC NUMBER does not have finite integer index on general graphs. \square

11.1.2 Selected positive results

Let us show for some problems that do not have finite integer index in general graphs that they do have this property if restricted to a suitable small class.

Lemma 59. CHROMATIC NUMBER has f.i.i. on degenerate graphs.

Proof. Fix an integer t and let \mathcal{G}_t be the class of all t -boundaried graphs. For simplicity, we identify the vertices of a graph $\mathring{G} \in \mathcal{G}_t$ with the numbers $1, \dots, t$. We abbreviate the problem CHROMATIC NUMBER by χ in the following.

We define a *configuration* with respect to \mathcal{G}_t as a $(t+1)$ -tuple $C = (p, c_1, \dots, c_t)$, where $p, c_i \in \mathbb{N}$. A t -boundaried graph \mathring{G} satisfies this configuration if the partial colouring c_1, \dots, c_t of its boundary-vertices can be extended to a proper colouring of G_t using in total p different colours.

The *signature* $\sigma[\mathring{G}]$ of a graph $\mathring{G} \in \mathcal{G}_t$ is a function from the configurations into $\{0, 1\}$ where $\sigma[\mathring{G}](C) = 1$ if and only if \mathring{G} satisfies C . We define:

$$\mathring{G}_1 \simeq_\sigma \mathring{G}_2 \iff \sigma[\mathring{G}_1] \equiv \sigma[\mathring{G}_2] \text{ for } \mathring{G}_1, \mathring{G}_2 \in \mathcal{G}_t.$$

We claim that the equivalence relation \simeq_σ is a refinement of the canonical equivalence \equiv_χ .

Assume the contrary, that $\sigma[\mathring{G}_1] \equiv \sigma[\mathring{G}_2]$ while $\mathring{G}_1 \not\equiv_\chi \mathring{G}_2$. The latter implies that for all constants $q \in \mathbb{N}$, there exists a graph $\mathring{G}_3 \in \mathcal{G}_t$ such that $\mathring{G}_1 \oplus \mathring{G}_3$ is q -colourable while $\mathring{G}_2 \oplus \mathring{G}_3$ is not. We choose $c = 0$ to arrive at a contradiction. Let c be a proper colouring of $\mathring{G}_1 \oplus \mathring{G}_3$ with q colours. By projecting c on $\partial\mathring{G}_1$ and counting the number of colours in $c(V(\mathring{G}_1))$, we identify a configuration C such that $\sigma[\mathring{G}_1](C) = 1$. By our above assumption, $\sigma[\mathring{G}_2] \equiv \sigma[\mathring{G}_1]$ and therefore $\sigma[\mathring{G}_2](C) = 1$ as well. But then the colouring $c(V(\mathring{G}_3))$ can be extended to a proper colouring of $\mathring{G}_2 \oplus \mathring{G}_3$ with exactly q colours, a contradiction. Since in the above argument we can exchange \mathring{G}_1 and \mathring{G}_2 we conclude that \simeq_σ is a refinement of \equiv_χ .

Finally, observe that the index of σ is finite in d -degenerate graphs since the number of signatures is bounded by $2d^{t+1}$: we need at most d additional colours to extend a colouring of the boundary to a proper colouring. We conclude that the index of σ is at most $2^{2d^{t+1}}$ and the claim follows. \square

For CHROMATIC NUMBER it is beneficial to bound the maximal number of colours ‘behind’ the boundary. Note that the above lemma easily extends to graph classes whose colouring number is bounded by other means than degeneracy. For connectivity problems like k -PATH we show a similar result, this time we need that one side of the boundary does not contain arbitrarily long paths. We prove it here for bounded-treedepth graphs since this is the setting that is applicable in the context of Theorems 28 and 32.

Lemma 60. *k-PATH, k-CYCLE, EXACT s, t-PATH, EXACT CYCLE have f.i.i. on graphs of bounded treedepth.*

Proof. Let \mathcal{D} be a class whose treedepth is bounded by a constant d . Let Π be any one of the mentioned problems. Consider the class \mathcal{G}_t of all t -boundaried graphs, and let $T = \{0, 1, \dots, t\}$.

We define a *configuration* of Π with respect to ${}^\circ\mathcal{G}_t$ as a multiset

$$C = \{(s_1, d_1, t_1), \dots, (s_p, d_p, t_p)\}$$

of triples from $(T \times \mathbb{N} \times T)$. We say a t -boundaried graph ${}^\circ G \in {}^\circ\mathcal{G}_t$ *satisfies* the configuration C if there exists a set of (distinct) paths P_1, \dots, P_p in G such that

- s_i, t_i can only be endvertices of P_i , $V(P_i) \cap \partial({}^\circ G) \subseteq \{s_i, t_i\}$, and $|P_i| = d_i$, for $1 \leq i \leq p$,
- $V(P_i) \cap V(P_j) \subseteq \partial({}^\circ G)$ for $1 \leq i < j \leq p$,
- $V(P_i) \cap V(P_j) \cap V(P_k) = \emptyset$ for $1 \leq i < j < k \leq p$.

Note that, for simplicity, we identify the boundary vertices in $\partial({}^\circ G)$ with their labels $1, \dots, t$ from T . Moreover, s_i, t_i can take the value 0 which is not contained in $\partial({}^\circ G)$: semantically these tuples describe paths that intersect the boundary of ${}^\circ G$ at only one or no vertex. Another special case are tuples with $s_i = t_i$ and $d = 0$: those describe single vertices of the boundary. In short, a graph satisfies a configuration if it contains internally non-intersecting paths of length and endvertices prescribed by the tuples of the configuration, and no three of the paths are prescribed to have the same endvertex (hence some configurations are not satisfiable at all, but this is a small technicality).

The *signature* $\sigma[{}^\circ G]$ of a graph ${}^\circ G \in {}^\circ\mathcal{G}_t$ is a function from the configurations into $\{0, 1\}$ where $\sigma[{}^\circ G](C) = 1$ if and only if ${}^\circ G$ satisfies C . We define:

$${}^\circ G_1 \simeq_\sigma {}^\circ G_2 \iff \sigma[{}^\circ G_1] \equiv \sigma[{}^\circ G_2] \text{ for } {}^\circ G_1, {}^\circ G_2 \in {}^\circ\mathcal{G}_t.$$

We claim that the equivalence relation \simeq_σ is a refinement of the canonical equivalence \equiv_Π . We provide only a sketch for $\Pi = k\text{-PATH}$, the proofs for the other problems work analogously.

Assume the contrary, that $\sigma[{}^\circ G_1] \equiv \sigma[{}^\circ G_2]$ while ${}^\circ G_1 \not\equiv_\Pi {}^\circ G_2$. Up to symmetry, this means that for all integers c there exists a graph ${}^\circ G_3 \in {}^\circ\mathcal{G}_t$ such that $({}^\circ G_1 \oplus {}^\circ G_3, \ell) \in \Pi$ but $({}^\circ G_2 \oplus {}^\circ G_3, \ell + c) \notin \Pi$. We choose $c = 0$ and show the contradiction. Thus the graph ${}^\circ G_1 \oplus {}^\circ G_3$ contains a path P of length ℓ but ${}^\circ G_2 \oplus {}^\circ G_3$ does not.

Using the implicit order given through the vertex order of P , we sort the subpaths of P contained in $P \cap G_1$ and obtain a sequence of paths $P_1, \dots, P_q \subseteq G_1$, each with at most two vertices—the ends—in $\partial({}^\circ G_1)$. By identifying each subpath P_i with the tuple (s_i, d_i, t_i) , where $d_i = |P_i|$ and s_i is the label of the start of P_i in $\partial({}^\circ G_1)$ (or 0 if $s_i \notin \partial({}^\circ G_1)$), and t_i the label of the end of P_i in $\partial({}^\circ G_1)$ (ditto), we obtain a

configuration $C_P = \{(s_1, d_1, t_1), \dots, (s_q, d_q, t_q)\}$. Now, ${}^\circ G_1$ satisfies C_P by the definition. Since $\sigma[{}^\circ G_1](C_P) = \sigma[{}^\circ G_2](C_P)$, there exists a set of paths $Q_1, \dots, Q_q \subseteq G_2$ witnessing that ${}^\circ G_2$ satisfies C_P . But then Q_1, \dots, Q_q together with $P \cap G_3$ form a path Q of length ℓ in ${}^\circ G_2 \oplus {}^\circ G_3$, a contradiction.

Second, although \simeq_σ is generally of infinite index, we claim that for every t , only a finite number of equivalence classes of \simeq_σ carry a representative of $\text{treedepth} \leq d$, and hence \simeq_σ is of finite index when restricted to graphs from \mathcal{D} . This is rather easy since graphs of $\text{treedepth} \leq d$ do not contain paths of length $2^d - 1$ or longer, and so a graph ${}^\circ G \in \mathcal{D}_t$ can satisfy a configuration $C = \{(s_1, d_1, t_1), \dots, (s_p, d_p, t_p)\}$ only if $d_i \in \{0, 1, \dots, 2^d - 2\}$ for $1 \leq i \leq p$. Recall, each boundary vertex label occurs at most twice among $s_1, t_1, \dots, s_p, t_p$ in a satisfiable configuration. Hence only finitely many such configurations C can be satisfied by a graph from \mathcal{D}_t , and consequently, finitely many function values of $\sigma[{}^\circ G]$ are non-zero for any ${}^\circ G \in \mathcal{D}_t$ and the number of the non-empty classes of \simeq_σ restricted to \mathcal{D}_t is finite. \square

Further, we proved the following results relating to width-measures.

Lemma 61 ([116]). *The problem BRANCHWIDTH has finite integer index on graphs of bounded branchwidth.*

Lemma 62 ([116, 115]). *The problem PATHWIDTH has f.i.i. on graphs of bounded pathwidth and TREEWIDTH has f.i.i. on graphs of bounded treewidth.*

11.2 WQO & REPRESENTATIVE SETS

Representative set

Definition 23 (Representative set). Let $(\mathcal{G}, \triangleleft)$ be a wqo graph class, let Π be a graph problem, and let further t be an integer.

We call a set of t -boundaried graphs \mathcal{R} a *representative set* of \mathcal{G}_t if for all ${}^\circ G \in \mathcal{G}_t$ there exists ${}^\circ H \in \mathcal{R}$ with ${}^\circ H \equiv_\Pi {}^\circ G$. We call ${}^\circ H$ the *representative* of ${}^\circ G$ and will denote it by $\mathcal{R}({}^\circ G)$ in the following.

Obviously we want the set of representatives to be as small as possible and we will see in the following that the concept of finite integer index helps to achieve this. Before that, we impose certain restrictions on the set of representatives which will simplify the following proofs.

*Monotone, order-preserving,
canonical*

Definition 24. Let $(\mathcal{G}, \triangleleft)$ be a wqo graph class, let Π be a graph problem, and let further t be an integer.

Let \mathcal{R} be a set of representative for \mathcal{G}_t . We call $\mathcal{R} \dots$

\dots *monotone* if $\Delta_\Pi(\mathcal{R}({}^\circ G), {}^\circ G) \geq 0$

\dots *canonical* if $\partial \mathcal{R}({}^\circ G) = \partial {}^\circ G$

\dots *order-preserving* if $\mathcal{R}({}^\circ G) \triangleleft {}^\circ G$

holds for all ${}^\circ G \in \mathcal{G}_t$. For a set \mathcal{R} with any of these properties, we denote by $\mathcal{R}({}^\circ G)$ the representative of ${}^\circ G$ that satisfies exactly these properties.

Lemma 63. *Let $(\mathcal{G}, \triangleleft)$ be a wqo graph class, Π be a graph problem that has f.i.i. in \mathcal{G} . Then for every $t \in \mathbb{N}$ there exists a finite set of monotone, canonical, and order-preserving representatives \mathcal{R} of ${}^\circ\mathcal{G}_t$.*

Proof. Consider the equivalence classes $\mathcal{C}_1, \dots, \mathcal{C}_p$ of ${}^\circ\mathcal{G}_t / \equiv_\Pi$; by the definition of f.i.i. we know that p is finite—this already proves that there is a finite representation, it is left so show that we can enforce the additional properties.

We start by showing that we can find a canonical set. This can be simply enforced by refining the classes $\mathcal{C}_1, \dots, \mathcal{C}_p$ further: we partition each individual class \mathcal{C}_i into at most $2^{O(t^2)}$ subclasses according to the boundaries of its members. The following construction will choose a set of graphs from each such equivalence class, therefore the resulting set will be canonical.

Fix a class $\mathcal{C} = \mathcal{C}_i$ and first assume that for every graph ${}^\circ H$ and every ${}^\circ G \in \mathcal{C}$ we have that $({}^\circ G \oplus {}^\circ H, \bullet) \notin \Pi$ —such a class essentially contains only no-instances of Π . Then we can assume that $\Delta_\Pi({}^\circ G_1, {}^\circ G_2)$ is zero for all members ${}^\circ G_1, {}^\circ G_2 \in \Pi$ and therefore any choice of a representative will be monotone for members of \mathcal{C} .

Now assume that \mathcal{C} does not contain such a graph. Consider the relation defined via

$${}^\circ G_1 \leq {}^\circ G_2 \iff \Delta_\Pi({}^\circ G_1, {}^\circ G_2) \geq 0.$$

Claim. *The relation \leq is a linear order.*

Since $\Delta_\Pi({}^\circ G, {}^\circ G) = 0$, the relation is reflexive and, since we defined it on the equivalence class \mathcal{C} , it is total. Unsurprisingly, it is also transitive which we can quickly verify: consider a triple of graphs from \mathcal{C} with ${}^\circ G_1 \leq {}^\circ G_2$ and ${}^\circ G_2 \leq {}^\circ G_3$. Let $c_{12} = \Delta_\Pi({}^\circ G_1, {}^\circ G_2)$ and $c_{23} = \Delta_\Pi({}^\circ G_2, {}^\circ G_3)$. For every ${}^\circ H$ and $\xi \in \mathbb{N}$ we have that

$$\begin{aligned} ({}^\circ G_1 \oplus {}^\circ H, \xi) \in \Pi &\iff ({}^\circ G_2 \oplus {}^\circ H, \xi + c_{12}) \in \Pi \\ &\iff ({}^\circ G_3 \oplus {}^\circ H, \xi + c_{12} + c_{23}) \in \Pi, \end{aligned}$$

which immediately implies that ${}^\circ G_1 \leq {}^\circ G_2$ since $c_{12}, c_{23} \geq 0$.

Fix a graph ${}^\circ H$ such that for all ${}^\circ G \in \mathcal{C}$ it holds that $({}^\circ G \oplus {}^\circ H, \bullet) \in \Pi$. Such a graph must exist by our choice of \mathcal{C} . Define

$$\Phi({}^\circ G) = \min\{\xi \in \mathbb{N} \mid ({}^\circ G \oplus {}^\circ H, \xi) \in \Pi\}.$$

It is easy to check that for ${}^\circ G_1, {}^\circ G_2 \in \mathcal{C}_i$ it holds that

$$\Phi({}^\circ G_1) = \Phi({}^\circ G_2) + \Delta_\Pi({}^\circ G_1, {}^\circ G_2).$$

Therefore, we have that ${}^\circ G_1 \leq {}^\circ G_2$ implies $\Phi({}^\circ G_1) \leq \Phi({}^\circ G_2)$. Any graph in \mathcal{C} that minimizes Φ will be a minimum according to the order defined by \leq ; choosing a representative from this set would therefore ensure the monotonicity—however, we need to work a bit more to make sure that our choice is also order-preserving.

Recall that Lemma 4 implies that $\circ\mathcal{G}_t$ is wqo by the boundary-preserving extension of \triangleleft . Our class \mathcal{C} does not need to be closed under \triangleleft ; we will therefore rely on the finite set $\text{basis}(\mathcal{C})$ (cf. Lemma 3 for the necessary properties). Now the members of $\text{basis}(\mathcal{C})$ need not be minimal according to Φ , our construction of the representative set is not yet done.

Let us partition \mathcal{C} into slices \mathcal{C}^k , $k \geq 0$, where $\mathcal{C}^k := \mathcal{C}_{\Phi=k}$. Since $\text{basis}(\mathcal{C})$ has finite cardinality, there exists a number λ such that for all $k \geq \lambda$ we have $\mathcal{C}^k \cap \text{basis}(\mathcal{C}) = \emptyset$. For all these slices the two properties already hold if $\text{basis}(\mathcal{C})$ is contained in the representative set; it is left to show that we can ensure it for the lower slices as well. To that end, consider the following iterative construction:

$$\mathcal{R}_k := \bigcup_{0 \leq i \leq k} \text{basis}(\mathcal{C}^i).$$

We claim that $\bigcup_{0 \leq k \leq \lambda} \mathcal{R}_k$ has the above two properties for all graphs in the lower slices.

Claim. *Let $\circ\mathcal{G} \in \mathcal{C}^k$ for $k < \lambda$. Then there exists $\circ\mathcal{H} \in \mathcal{R}_k$ such that $\circ\mathcal{H} \triangleleft \circ\mathcal{G}$ and $\Phi(\circ\mathcal{H}) \geq \Phi(\circ\mathcal{G})$.*

We prove the claim by induction over k . For $k = 0$, the statement clearly holds. Assume the statement holds for $k - 1 > 0$. If there exists $\circ\mathcal{H} \triangleleft \circ\mathcal{G}$, $\circ\mathcal{H} \neq \circ\mathcal{G}$ such that $\circ\mathcal{H} \in \mathcal{C}^{<k}$ we are done by induction and the transitivity of \triangleleft . Otherwise we have that every $\circ\mathcal{H} \triangleleft \circ\mathcal{G}$ is either not in \mathcal{C} or it is in \mathcal{C}^k . But then, for some $\circ\mathcal{H} \triangleleft \circ\mathcal{G}$, it must hold that $\circ\mathcal{H} \in \text{basis}(\mathcal{C}^k) \subseteq \mathcal{R}^k$.

We are finally done: the set $\mathcal{R}_{\mathcal{C}} := \mathcal{R}^k \cup \text{basis}(\mathcal{C})$ contains all representatives for the class \mathcal{C} . Taking the union over all such sets for all classes $\mathcal{C}_1, \dots, \mathcal{C}_p$ the set \mathcal{R} is monotone, canonical and order-preserving. This concludes the proof. \square

The above lemma tells us that problems with finite integer index behave well with respect to constant-sized separators: in a sense, there is only small amount of information which can be passed through such a bottleneck. Even more, the necessary information can be reduced to a constant-sized (boundaried) instance of the *same problem*. While we will in the following assume that such a set of representatives is simply part of algorithm, we will need to take care of identifying the correct representative.

Lemma 64 (Finding representatives). *Fix a graph class \mathcal{G} , a problem Π that has f.i.i. on \mathcal{G} and an integer t . Let \mathcal{R} be a representative set of $\circ\mathcal{G}_t$.*

Then for every t -boundaried graph $\circ\mathcal{G}$, the representative $\mathcal{R}(\circ\mathcal{G})$ for $\circ\mathcal{G}$ is computable.

Proof. Consider first the case that \mathcal{R} contains only two graphs $\circ\mathcal{H}_1, \circ\mathcal{H}_2$ and that $\circ\mathcal{G}_t / \equiv_{\Pi}$ has only two equivalence classes $\mathcal{C}_1, \mathcal{C}_2$ represented by them. In this setting, given a t -boundaried graph from $\circ\mathcal{G}_t$, we need to determine to which of the two classes it belongs.

Since ${}^\circ H_1 \not\equiv_{\Pi} {}^\circ H_2$, for every $d \in \mathbb{N}$ there exists a (minimal) witness ${}^\circ \Theta_d, \zeta_d$ such that

$$({}^\circ H_1 \oplus {}^\circ \Theta_d, \zeta_d) \in \Pi \iff ({}^\circ H_2 \oplus {}^\circ \Theta_d, \zeta_d + d) \notin \Pi.$$

For every witness graph ${}^\circ \Theta_d$ we define, as we did in the proof of Lemma 63, the function

$$\Phi_d({}^\circ G) := \min\{\zeta \mid ({}^\circ G \oplus {}^\circ \Theta_d, \zeta) \in \Pi\}.$$

Now if ${}^\circ G \equiv_{\Pi} {}^\circ H_i$, then for every $d \geq 0$ it holds that $\Delta_{\Pi}({}^\circ G, {}^\circ H_i) = \Phi_d({}^\circ G) - \Phi_d({}^\circ H_i)$. Define $\delta = \Phi_0({}^\circ H_2) - \Phi_0({}^\circ H_1)$. Assume without loss of generality that $({}^\circ H_1 \oplus {}^\circ \Theta_{\delta}, \zeta_{\delta}) \in \Pi$. Then the two following implications hold:

$$\begin{aligned} {}^\circ G \equiv_{\Pi} {}^\circ H_1 &\implies ({}^\circ G \oplus {}^\circ \Theta_{\delta}, \zeta_{\delta} + \Phi_0({}^\circ G) - \Phi_0({}^\circ H_1)) \in \Pi \\ {}^\circ G \equiv_{\Pi} {}^\circ H_2 &\implies ({}^\circ G \oplus {}^\circ \Theta_{\delta}, \zeta_{\delta} + \delta + \Phi_0({}^\circ G) - \Phi_0({}^\circ H_2)) \notin \Pi \\ &\iff ({}^\circ G \oplus {}^\circ \Theta_{\delta}, \zeta_{\delta} + \Phi_0({}^\circ G) - \Phi_0({}^\circ H_1)) \notin \Pi \end{aligned}$$

Therefore, we have the following test for membership in classes $\mathcal{C}_1, \mathcal{C}_2$: the graph ${}^\circ G$ is equivalent to ${}^\circ H_1$ if and only if

$$({}^\circ G \oplus {}^\circ \Theta_{\delta}, \zeta_{\delta} + \Phi_0({}^\circ G) - \Phi_0({}^\circ H_1)) \in \Pi \leftrightarrow ({}^\circ H_1 \oplus {}^\circ \Theta_{\delta}, \zeta_{\delta}) \in \Pi.$$

Let us now generalise to the case for arbitrarily many equivalence classes $\mathcal{C}_1, \dots, \mathcal{C}_p$. We use the above test between each pair of distinct classes and note the outcome in the form of a tournament; that is, for a pair C_i, C_j we add the arc $C_i C_j$ if the binary tests come out in favour of ${}^\circ G \in C_j$ and $C_j C_i$ otherwise. Since we know that there exists a representative $\mathcal{R}({}^\circ G)$, this tournament contains a sink—the class ‘wins’ all binary tests against all other classes. Since a tournament can contain at most one sink, we correctly identify the class and representative $\mathcal{R}({}^\circ G)$. Note that all involved quantities, in particular δ , depend on ${}^\circ G, \mathcal{R}$. Hence the total time necessary to determine a representative is a function of these two inputs only. \square

Corollary 14. *Assume the conditions of Lemma 64 and additionally, let \mathcal{R} be any combination of monotone, canonical, and order-preserving. For the last case, assume that $(\mathcal{G}, \triangleleft)$ is wqo for a decidable relation \triangleleft .*

Then we can compute for every graph ${}^\circ G \in {}^\circ \mathcal{G}_t$ a representative $\mathcal{R}({}^\circ G)$ that obeys these constraints.

Proof. For all three properties, we use a simple preprocessing routine to identify a subset of representatives on which we proceed as before.

- If we need to find a canonical representative, we only retain those representatives that have a boundary identical to the query graph.
- If the representative needs to be monotone, we disregard all candidate representatives where the prospective Δ_{Π} -value is negative (using, for example, Φ_0 as in the proof of Lemma 64).

- If the representative is supposed to be order-preserving, we test for every candidate representative ${}^\circ H$ and query graph ${}^\circ G$ whether ${}^\circ H \triangleleft {}^\circ G$ and keep those representatives that pass this test. We rely on the assumption that \triangleleft is decidable.

We conclude that the three constraints, in any combination, do not change that a representative is computable from the representative set alone. \square

11.3 PROTRUSIONS: REPLACEMENT AND DECOMPOSITION

The notion of a *protrusion* was introduced in the original meta-kernels paper [23]. We adapt and extend the definition as follows.

Definition 25 ((α, β) -protrusion^{wm}). Given a graph G and a width-measure \mathbf{wm} , a set $W \subseteq V(G)$ is a (α, β) -protrusion^{wm} of G if

- $|\partial_G W| \leq \alpha$ and
- $\mathbf{wm}(G[W]) \leq \beta$.

We call $\partial_G W$ the *boundary*, α the *adhesion*, β the *width* and $|W|$ the *size* of the protrusion W .

Note that for a (α, \bullet) -protrusion W in G we have that ${}^\circ G[W]_\bullet$ is a α -boundaried graph. If a suitable labelling b of the boundary is fixed we will therefore use the shorthand ${}^\circ W := {}^\circ G[W]_b$.

With the tools proved in Section 11.1 at hand we can finally prove the cornerstone of the meta-kernelisation framework: protrusion replacement. We already saw in the previous section that problems with f.i.i. allow us to replace small boundaried graphs with even smaller graphs taken from a finite selection of representatives. As a thought experiment, consider a ‘trivial’ protrusion: a graph of bounded tree-width. Iteratively replacing small boundaried graphs will, in essence, persistently eat away at the graph until only a constant-sized, equivalent instance of our original problem remains.

The same procedure is possible even if a protrusion attaches to some larger set—as shown in the next lemma.

Protrusion replacement

Lemma 65 (Protrusion replacement). *Let \mathbf{wm} be a with-measure that upper-bounds \mathbf{tw} . Let \mathcal{G} be a graph class, \triangleleft a graph relation, α, β integers and Π be a graph problem such that*

- \mathbf{wm} is monotone under \triangleleft ,
- $\mathcal{G}_{\mathbf{wm} \leq \alpha + \beta}$ is wqo by \triangleleft and
- Π has f.i.i. on $\mathcal{G}_{\mathbf{wm} \leq \alpha + \beta}$.

Let further \mathcal{R} be a set of monotone, canonical, and order-preserving representatives for Π on $\mathcal{G}_{\mathbf{wm} \leq \alpha + \beta}$.

Let (G, k) be an input instance of Π with $G \in \mathcal{G}$ and let $W \subseteq G$ be a (α, β) -protrusion^{wm} of G , and b a suitable α -labelling of ∂W . There exists an algorithm that computes in time $O(|W|)$ a representative $\mathcal{R}(\circ G[W]_b)$ for $\circ G[W]_b$.

Proof. We first compute a tree-decomposition $\mathcal{T} = (T, \chi)$ of $G[W]$ of width at most $\alpha + \beta - 1$ that satisfies the following conditions:

1. The tree-decomposition is nice and the leaf bags contain one vertex, and
2. the bag of $r = \text{root}(T)$ is exactly the boundary of W , i.e. we have $\chi(r) = \partial_G W$.

The first condition can be achieved by simply modifying the graph $G[W]$ so that $\partial_G W$ induces a clique; if a bag containing $\partial_G W$ we make it the root of T , otherwise we introduce such a node (since $\partial_G W$ will be contained in some bag there must be a way of attaching it). This construction also justifies the width $\alpha + \beta - 1$, using the fact that

$$\mathbf{tw}(G[W]) \leq \mathbf{wm}(G[W]) \leq \beta.$$

We abbreviate $\gamma := \alpha + \beta$ in the following.

For a node $x \in T$, we let $\circ G_x$ denote the γ -boundaried graph induced by the vertices in the bags of the rooted subtree T_x . That is,

$$\circ G_x := \circ G \left[\bigcup_{y \in T_x} W_y \right]_{b_x} \text{ and } \partial \circ G_x = \chi(x),$$

where b_x is some fixed labelling of $\chi(x)$. Note that $\circ G_r = \circ G[W]$. Since Π has f.i.i. on $\mathcal{G}_{\mathbf{wm} \leq w}$ there exists a finite set of representatives \mathcal{R} for Π with properties as defined in Lemma 63. In particular, we have that for every node $x \in T$, there exists a representative $\mathcal{R}(\circ G_x) \in \mathcal{R}$ for $\circ G_x$ such that

- $\mathcal{R}(\circ G_x) \equiv_{\Pi} \circ G_x$ with $\Delta_{\Pi}(\mathcal{R}(\circ G_x), \circ G_x) \geq 0$,
- $\partial \mathcal{R}(x) = \partial \circ G_x$, and
- $\mathcal{R}(\circ G_x) \triangleleft \circ G_x$.

Note that the size of $\mathcal{R}(\circ G_x)$ is only depends on our ‘constants’: γ, Π, \mathcal{G} and \triangleleft . For brevity’s sake, we introduce the function

$$\mu(x) := \Delta_{\Pi}(\mathcal{R}(\circ G_x), \circ G_x).$$

$\mu(x)$

Our first goal is to prove the following claim, under the assumption that the finite set \mathcal{R} is part of the algorithm.

Claim. *The values $\mathcal{R}(\circ G_r), \mu(r)$ can be computed in time $O(|\circ G_r|)$.*

We use a bottom-up computation over the tree decomposition. For a leaf-node $x \in \text{leaves}(T)$, Lemma 64 and Corollary 14 tell us that we can compute a suitable representative $\mathcal{R}(\circ G_x)$ of $\circ G_x$. In the following, we consider nodes $x \in T$ and assume that the values $\mathcal{R}(\bullet), \mu(\bullet)$ are known for all its children.

Case 1: $x \in T$ is an introduce- or forget-node. Let $y \in N_T^+(x)$ be the unique child of x . Consider the γ -boundaried graph

$${}^\circ G'_x := ({}^\circ G_x \ominus_{b_y} {}^\circ G_y) \oplus_{b_y} \mathcal{R}({}^\circ G_y),$$

where b_y is the labelling fixed for $\chi(y)$. We claim that ${}^\circ G'_x \equiv_{\Pi} {}^\circ G_x$; to prove this we need to show that for all γ -boundaried graphs ${}^\circ H$ and for all $\xi \in \mathbb{N}$,

$$({}^\circ G_x \oplus_{b_x} {}^\circ H, \xi) \in \Pi \iff ({}^\circ G'_x \oplus_{b_x} {}^\circ H, \xi + \Delta_{\Pi}({}^\circ G_x, {}^\circ G'_x))$$

for some yet to be specified constant $\Delta_{\Pi}({}^\circ G_x, {}^\circ G'_x)$ and some labelling b_x of the boundary $\partial {}^\circ G_x$. Observe that, by some basic ‘arithmetic’,

$$\begin{aligned} {}^\circ G'_x \oplus_{b_x} {}^\circ H &= ({}^\circ G_x \ominus_{b_y} {}^\circ G_y) \oplus_{b_y} \mathcal{R}({}^\circ G_y) \oplus_{b_x} {}^\circ H \\ &= ({}^\circ G_x \oplus_{b_x} {}^\circ H) \ominus_{b_y} {}^\circ G_y \oplus_{b_y} \mathcal{R}({}^\circ G_y). \end{aligned}$$

Therefore, we have that

$$\begin{aligned} &({}^\circ G'_x \oplus_{b_x} {}^\circ H, \xi) \in \Pi \\ \iff &(({}^\circ G_x \oplus_{b_x} {}^\circ H) \ominus_{b_y} {}^\circ G_y \oplus_{b_y} \mathcal{R}({}^\circ G_y), \xi) \in \Pi \\ \stackrel{{}^\circ G_y \equiv_{\Pi} \mathcal{R}({}^\circ G_y)}{\iff} &(({}^\circ G_x \oplus_{b_x} {}^\circ H) \ominus_{b_y} {}^\circ G_y \oplus_{b_y} {}^\circ G_y, \xi - \mu(y)) \in \Pi \\ \iff &({}^\circ G_x \oplus_{b_x} {}^\circ H, \xi - \mu(y)) \in \Pi, \end{aligned}$$

proving the claim with $\Delta_{\Pi}({}^\circ G_x, {}^\circ G'_x) = -\mu(y)$. Since ${}^\circ G'_x$ has constant size, we can safely invoke Lemma 64/Corollary 14 to determine $\mathcal{R}({}^\circ G_x) = \mathcal{R}({}^\circ G'_x)$ and $\mu(x) = \mu(y) + \Delta_{\Pi}({}^\circ G'_x, \mathcal{R}({}^\circ G'_x))$ in constant time. Further, since $\mathcal{R}({}^\circ G'_x) \triangleleft {}^\circ G_x$, we have that $\mathbf{wm}({}^\circ G'_x) \leq \mathbf{wm}({}^\circ G_x)$.

Case 2: $x \in T$ is a join-node. Let $y, z \in N_T^+(x)$ be the two children of x . Since the tree-decomposition is nice, we have that $\chi(x) = \chi(y) = \chi(z)$ and therefore also $\partial {}^\circ G_x = \partial {}^\circ G_y = \partial {}^\circ G_z$. Let $b := b_x$ be the fixed labelling of the shared boundary $\partial {}^\circ G_x$. Consider the γ -boundaried graph ${}^\circ G_{yz}$ defined as

$$G_{yz} := \mathcal{R}(y) \oplus_b \mathcal{R}(z) \text{ and } \partial {}^\circ G_{yz} := \chi(x).$$

Analogous to the previous case, we can show that for all ${}^\circ H$ and $\xi \in \mathbb{N}$,

$$({}^\circ G_{yz} \oplus_b {}^\circ H, \xi) \in \Pi \iff ({}^\circ G_x \oplus_b {}^\circ H, \xi - \mu(y) - \mu(z)) \in \Pi.$$

Again, the graph ${}^\circ G_{yz}$ has, by induction, constant size and it follows that we can compute $\mathcal{R}({}^\circ G_x)$ and $\mu(x)$ in constant time.

This concludes the proof of the claim: we are able to compute $\mathcal{R}({}^\circ G_r)$ and $\mu(r)$ in time $O(|{}^\circ G_r|)$. Actually, we have now proved the Lemma—the pair ${}^\circ H := \mathcal{R}({}^\circ G_r)$ and $k' := \mu(r)$ with boundary $b := b_r$ satisfies the above conditions. \square

We now have the template for a reduction rule: using a suitable width-measure and graph relation, we are able to safely replace protrusions while preserving important properties—like membership in a graph class (more on this later). The second ingredient to prove a kernels is, as always, the proof that the exhaustive application of the reduction rule will output an instance of bounded size. This second step will crucially depend on the graph class we are working in; in order to unify the following results as much as possible we will rely on the following graph decomposition.

Protrusion decomposition

Definition 26 (Protrusion decomposition). For a graph G , a vertex partition $Y_0 \uplus Y_1 \uplus \dots \uplus Y_\ell$ is a *protrusion^{wm} decomposition* of width β and adhesion α if $K(G - Y_0)$ refines the partition $\{Y_1, \dots, Y_\ell\}$ and for all $1 \leq i < j \leq \ell$ we have that

- $N(Y_i) \neq N(Y_j)$ and
- Y_i is a protrusion^{wm} of adhesion α and width β .

We call the set Y_0 the *core* of the decomposition.

Note that decompositions is entirely defined by its core: given $Y_0 \subseteq G$, we group the components $K(G \setminus Y_0)$ according to their boundaries in Y_0 to obtain $Y_0 \uplus Y_1 \uplus \dots \uplus Y_\ell$. Moreover, we can compute the decomposition for a given core in linear time: we construct an bipartite auxiliary graph with vertex set $K(G \setminus Y_0) \cup Y_0$ and add the edges from $C \in K(G \setminus Y_0)$ to $\partial_G C \subseteq Y_0$. Computing the twin classes of this bipartite graph then produces the correct grouping of the connected components; by Lemma 2 this can be done in linear time.

Low-adhesion –

The following lemma shows how the structure of protrusion decompositions can be improved further by reducing the adhesion at the cost of slightly increasing the size of the core. The exact cost of this increase will depend on the graph class we are working in, therefore we simply return a suitable witness-structure whose size is related to the increase *and* which can be bounded in sparse classes.

Lemma 66 (Low-adhesion protrusion decomposition). *Let G be a graph with protrusion^{tw} decomposition $Y_0 \uplus Y_1 \uplus \dots \uplus Y_\ell$ of width β and arbitrary adhesion. There is an algorithm that for every $\alpha > 0$ computes in linear time*

- a witness structure of pairwise disjoint vertex subsets W_1, \dots, W_p where $W_j \cap Y_0 = \emptyset$, $G[W_j]$ connected and $N(W_j) \cap Y_0 \geq \alpha$; and
- a protrusion^{tw} decomposition $Y'_0 \uplus Y'_1 \uplus \dots \uplus Y'_\ell$ for G of width β and adhesion at most $\alpha + 2\beta + 2$ where $Y'_0 \supseteq Y_0$, $Y'_i \subseteq Y_i$ for $1 \leq i \leq \ell$ and further $|Y'_0| \leq 2p(\beta + 1) + |Y_0|$.

Proof. The procedure is listed as Algorithm 2. We prove the different claims in the statement one by one for the output of the algorithm. Accordingly, we denote by $Y'_0 \uplus Y'_1 \uplus \dots \uplus Y'_\ell$ and W_1, \dots, W_p the output

Input: A graph G with a protrusion^{tw} decomposition $Y_0 \uplus Y_1 \uplus \dots \uplus Y_\ell$ and an integer $\alpha > 0$.
Output: A protrusion^{tw} decomposition $Y'_0 \uplus Y'_1 \uplus \dots \uplus Y'_{\ell'}$ and witness structure W_1, \dots, W_p as described in Lemma 66.

For every $C \in \mathcal{K}(G - Y_0)$ with $|N(C) \cap Y_0| \geq \alpha$, compute an optimal rooted tree-decomposition $\mathcal{T}_C = (T_C, \chi_C)$

Set $M \leftarrow \emptyset$ as the set of marked nodes
Set $\mathcal{W} \leftarrow \emptyset$ as the list of witnesses
Repeat the following loop for every rooted tree-decomposition \mathcal{T}_C :
while T_C contains an unprocessed node **do**

Let $x \in T$ be an unprocessed node at the farthest distance from $\text{root}(T)$

① LCA marking
if x is the LCA of two nodes in M **then**
 $M \leftarrow M \cup \{x\}$

② Large-subgraph marking
else if $G[\bigcup_{y \in T_x} \chi(y) \setminus \bigcup_{z \in M} \chi(z)]$ contains a connected component W such that $|N(W) \cap Y_0| \geq \alpha$ **then**
 $M \leftarrow M \cup \{x\}$
 $\mathcal{W} \leftarrow \mathcal{W} \cup \{W\}$

Node x is now processed

③ Decomposition construction
 $Y'_0 = Y_0 \cup V(\mathcal{M})$
for $1 \leq i \leq \ell$ **do**
 $Y'_i = Y_i \setminus V(\mathcal{M})$

return $Y'_0 \uplus Y'_1 \uplus \dots \uplus Y'_{\ell'}$ and $\mathcal{W} = W_1, \dots, W_p$

Algorithm 2: Computing a low adhesion protrusion decomposition.

of the algorithm for given inputs G and $Y_0 \uplus Y_1 \uplus \dots \uplus Y_\ell$. We immediately see that, by construction, Y'_0 is a superset of Y_0 and that each protrusion Y'_i is a subset of the protrusion Y_i . This implies that the width of $Y'_0 \uplus Y'_1 \uplus \dots \uplus Y'_\ell$ is at most β , since \mathbf{tw} is monotone under taking subgraphs. Let us examine the remaining claims.

Claim. *The witness sets W_1, \dots, W_p are pairwise disjoint, induce connected subgraphs and have at least α neighbours in Y_0 .*

Consider that iteration i in the algorithm where W_j was added to the list of witnesses in Step 2: from the if-clause it directly follows that the witness sets induce connected subgraphs and they contain at least α neighbours in Y_0 . Furthermore, since the tree node x is added to the set of marked nodes, we have—by the properties of tree-decompositions—that none of the vertices in W will be added to any other witness set. Hence they are pairwise disjoint.

Claim. $|Y'_0| \leq 2p(\beta + 1) + |Y_0|$.

Each marked node in Algorithm 2 adds the vertices of a bag to Y_0 in order to construct Y'_0 ; since the protrusion decomposition has width β , each such bag contains at most $\beta + 1$ vertices. It remains to be shown that at most $2p$ nodes are marked in total. Every node marked in Step 2 will create a witness and hence increase the final number of witnesses p . Since Step 1 simply computes the lca closure of the so-far marked nodes, we have by Lemma 1 that the total number of marked nodes is at most $2p$.

Claim. $Y'_0 \uplus Y'_1 \uplus \dots \uplus Y'_\ell$ has adhesion at most $\alpha + 2\beta + 2$.

For every protrusion Y'_i we have that each connected component of $G[Y'_i]$ has at most α neighbours in Y_0 (as ensured by Step 2). However, these components are potentially adjacent to $Y'_0 \setminus Y_0$. We finally see the reason behind computing the lca closure: by Lemma 1, every connected component of unmarked nodes in the tree decomposition has at most two marked neighbours. Since a connected component of $G[Y'_i]$ lives inside bags corresponding to unmarked nodes in the tree decomposition, we have that $G[Y'_i]$ can be adjacent to at most two bags worth of vertices in $Y'_0 \setminus Y_0$. Hence, in total the adhesion of Y'_i and subsequently $Y'_0 \uplus Y'_1 \uplus \dots \uplus Y'_\ell$ is bounded by $\alpha + 2\beta + 2$. \square

Because the only modification Algorithm 2 applies to a protrusion decomposition is that it pushes some vertices from the protrusions into the core, the above lemma holds for width measures like pathwidth or treedepth, as well. The following corollary formalises the conditions under which the results can be carried over for general width measures.

Corollary 15. *Let $\mathbf{wm} \geq \mathbf{tw}$ be a width measure that is monotone under taking induced subgraphs. Then Lemma 66 can be applied to a protrusion ^{\mathbf{wm}} -decomposition as well.*

META-KERNELISATION BEYOND EXCLUDED MINORS

Having introduced the necessary ingredients—finite integer index, protrusions, and protrusion decompositions—we will now see how they interact with structurally sparse classes. In essence, the kernelisation routine simply reduces all protrusions; the decomposition is then used in combination with the graph classes’ properties to show that the resulting graph is indeed a kernel.

12.1 CLASSES EXCLUDING A TOPOLOGICAL MINOR

Recall that \mathbf{tw}_d^Δ denotes the size of a minimal treewidth- d -modulator. This section is dedicated to proving the first of the two main theorems. Let us restate it here:

Theorem 27. *Let Π be a graph problem that has f.i.i. on graphs of bounded treewidth and let \mathcal{G} be a graph class that excludes a fixed graph H as a topological minor. For every $d \in \mathbb{N}$ there exists an algorithm that takes as input $(G, \xi) \in \mathcal{G} \times \mathbb{N}$ and outputs in time $O(|G|)$ an instance (G', ξ') such that*

1. $(G, \xi) \in \Pi \iff (G', \xi') \in \Pi$ with $\xi' \leq \xi$,
2. $H \not\prec_t G'$, and
3. $|G'| = O(\mathbf{tw}_d^\Delta(G))$.

For the remainder of this section, let \mathcal{G} be a graph class that excludes the fixed graph H as a topological minor. Recall that, by Theorem 4, there exists a constant $\rho \leq 10$ such that \mathcal{G} is $\rho|H|^2$ -degenerate. We fix a problem Π that has f.i.i. on graphs of bounded treewidth, i.e. for every $d \in \mathbb{N}$, it has f.i.i. if restricted to $\mathcal{G}_{\mathbf{tw} \leq d}$.

Let us begin with a technical but highly useful lemma on graphs excluding a topological minor—this was, historically, the predecessor to Lemma 36 and uses a very similar proof.

Lemma 67. *Let G be a graph that excludes K_h as a topological minor and let $X \subseteq V(G)$ be a vertex set. Then it holds that*

1. $|\{C : |\partial C| \geq h\}_{C \in K(G \setminus X)}| \leq \rho h^2 |X|$, and
2. $|\{\partial C\}_{C \in K(G \setminus X)}| \leq (2^{\rho h^2} + \rho h^2) |X|$.

Proof. We construct a sequence of graphs G_0, G_1, \dots, G_ℓ while ensuring that $G_i \prec_t G$ for all $0 \leq i \leq \ell$ as follows. Set $G_0 = G$, and for $0 \leq i \leq \ell - 1$ construct G_{i+1} from G_i by choosing a component $C \in K(G \setminus X)$ such that $\partial C \subseteq X$ contains two non-adjacent vertices

u, w in G_i . We construct G_{i+1} by adding the edge uw and removing C ; to see that $G_{i+1} \preceq_t G_i$, note that there exists a u - w -path whose internal vertices lie completely in C . Hence we can contract this path into the edge uw and remove the remainder of C to obtain G_{i+1} via legal topological minor operations.

This process clearly terminates, as G_{i+1} has at least one more edge between vertices of X than G_i . Since $G_{i+1} \preceq_t G_i$ we in particular have that $G_i \preceq_t G_0$ for all i . Therefore, $G_\ell[X]$ excludes K_h as a topological minor and is ρh^2 -degenerate.

Let us now prove the first claim. To this end, assume towards a contradiction that there is a connected component $C \in K(G_\ell \setminus X)$ such that $|\partial C| > h$. If $G_\ell[\partial C]$ contains a pair of non-adjacent vertices, we can construct the $(\ell + 1)^{\text{th}}$ graph in the sequence contradicting the assumption that G_ℓ was the last graph of the sequence. But then $G_\ell[\partial C]$ induces a clique of size at least h , contradicting that G does not contain K_h as a topological minor.

We conclude that no connected component of $G_\ell \setminus X$ has a boundary of size larger than h . Therefore components of $G \setminus X$ with boundaries of size $\geq h$ must have been deleted in the sequence of edge contractions. As every contraction adds exactly one edge between vertices in X and since we established that $G_\ell[X]$ contains at most $\rho h^2 |X|$ edges, the first claim follows.

For the second claim, consider the set $\mathcal{C}_\ell = K(G_\ell - X)$. By the above observation, the boundary of every component $C \in \mathcal{C}_\ell$ induces a clique in $G_\ell[X]$. From the degeneracy of $G_\ell[X]$ and Theorem 4, it follows that $G_\ell[X]$ contains at most $2^{\rho h^2} \cdot |G_\ell[X]| = 2^{\rho h^2} \cdot |X|$ complete subgraphs. Therefore the total number of boundaries in X of components in $G \setminus X$ is bounded by the number of contractions $\ell \leq \rho h^2 |X|$ and the resulting number of complete subgraphs in $G_\ell[X]$. This is exactly the second claim. \square

Consider the first statement of Lemma 67 and recall the ominous witness structure constructed by Algorithm 2. The claim was designed to bound the size of that witness structure, which in turn will bound the increase of the core set necessary to ensure that a protrusion decomposition has low adhesion.

We will assume, for now, that we are given an instance (G, ξ) of Π with $G \in \mathcal{G}$ and a treewidth- d modulator $X \subseteq G$ and see in the following how the protrusion reduction rule enables us to compress the instance to size $O(|X|)$. Afterwards, we tie in known results to derive Theorem 27.

Lemma 68. *Let $Y_0 \uplus Y_1 \uplus \dots \uplus Y_\ell$ be the protrusion decomposition of G with $Y_0 = X$ of width d . Then the witness structure obtained by applying Algorithm 2 with parameter $\alpha = h$ contains at most $\rho h^2 |X|$ sets.*

Proof. Recall that the witness structure consist of pairwise disjoint vertex subsets W_1, \dots, W_p contained entirely in $G \setminus Y_0$ that each in-

duce a connected subgraph and have at least h neighbours in X (cf. Lemma 66). To prove the claim, we show that $p \leq \rho h^2 |X|$.

To that end, construct the auxiliary subgraph $H \subseteq G$ with vertex set $X \cup \bigcup_{1 \leq i \leq p} W_i$ and edge set $E[X] \cup \bigcup_{1 \leq i \leq p} E[W_i]$. Note that every subgraph $G[W_i]$ is a connected component of $H \setminus X$; therefore we can apply the first statement of Lemma 67 to see that

$$p = |K(H \setminus X)| \leq \rho h^2 |X|,$$

as claimed. \square

We can now already prove a major part of Theorem 27: if we are given a modulator to constant treewidth as part of the input then we can compute a compressed instance that excludes H as a topological minor.

Lemma 69. *Given the instance (G, ξ) and the treewidth- d modulator X , we can, in linear time, compute an instance (G', ξ') such that*

1. $(G, \xi) \in \Pi \iff (G', \xi') \in \Pi$ with $\xi' \leq \xi$,
2. $H \not\prec_t G'$, and
3. $|G'| = O(|X|)$.

Proof. We first construct a protrusion decomposition $Y_0 \uplus Y_1 \uplus \dots \uplus Y_\ell$ around the core $Y_0 = X$ in linear time. Applying Algorithm 2 with parameter $\alpha = h$ gives us, by Lemma 66, a protrusion decomposition $Y'_0 \uplus Y'_1 \uplus \dots \uplus Y'_\ell$ with adhesion at most $\alpha' := h + 2d + 2$ and a witness structure W_1, \dots, W_p . As proved above in Lemma 68, the number of witness sets is bounded by $p \leq \rho h^2 |X|$.

It follows (cf. again Lemma 66) that the size of the new core Y'_0 is

$$|Y'_0| \leq 2p(d+1) + |Y_0| = 2\rho h^2(d+1)|X| + |X| = O(|X|).$$

Let us fix for every protrusion Y'_i a $\leq \alpha'$ -labelling of its boundary b_i , which enables us to refer to the $\leq \alpha'$ -boundaried graphs $\circlearrowleft Y'_i := G[Y'_i]_{b_i}$ in the following and avoid clumsy notation.

Now, applying the protrusion reduction machinery, we replace every protrusion Y'_i , $1 \leq i \leq \ell$, by a constant sized representative. More precisely, we apply Lemma 65 choosing the following parameters:

- The problem is $\Pi \upharpoonright \mathcal{G}$;
- the width-measure is treewidth;
- the graph relation is the minor relation \preceq_m ; and
- the parameters are $\alpha = \alpha' = h + 2d + 2$ and $\beta = d$.

Since Π has f.i.i. on graphs of bounded treedepth, we assume that the finite set of canonical, monotone and minor-preserving representatives \mathcal{R} for Π on $\mathcal{G}_{\text{tw} \leq \alpha + \beta}$ is contained in the algorithm.

Note that since \mathcal{G} is the class of all graphs excluding a fixed topological minor, and this fact is expressible in CMSO, we have that \mathcal{G} has

finite index. By Observation 3, the problem $\Pi \cap \mathcal{G}$ therefore has finite integer index.

Applying Lemma 65 with the above parameters to (G, ξ) and a protrusion Y'_i with boundary b_i provides us in time $O(|Y'_i|)$ with the representative $\circ Y''_i := \mathcal{R}(\circ Y'_i)$. Given all such representatives $\circ Y''_i$ for indices $1 \leq i \leq \ell$, we compute a sequence of instances (G_i, ξ_i) as follows: start with $(G_0, \xi_0) := (G, \xi)$. Then (G_i, ξ_i) is obtained from (G_{i-1}, ξ_{i-1}) by replacing a protrusion and adjusting the parameter accordingly:

$$\begin{aligned} G_i &:= G_{i-1}[Y'_i \mapsto \circ Y''_i]_{b_i} \\ \xi_i &:= \xi_{i-1} - \Delta_{\Pi \cap \mathcal{G}}(\circ Y''_i, \circ Y'_i). \end{aligned}$$

Since Lemma 65 computes representatives that are canonical, we know that $G[Y'_0]$ is not affected by these operations and therefore the sequence can be constructed as described above. Because the representatives are monotone, we have that $\Delta_{\Pi \cap \mathcal{G}}(\circ Y''_i, \circ Y'_i) \geq 0$ and therefore it holds that $\xi_i \leq \xi_{i-1}$. Since we used the problem $\Pi \cap \mathcal{G}$, we have that $G_i \in \mathcal{G}$ and thus $H \not\leq_t G_i$.

Obviously, but most importantly, we have that

$$(G_i, \xi_i) \in \Pi \iff (G_{i-1}, \xi_{i-1}) \in \Pi.$$

We set $(G', \xi') = (G_{\ell'}, \xi_{\ell'})$; the first two claims then follow via induction over the just constructed sequence. It is left to show the size bound. Note that the total size of the graph is

$$|G'| \leq |Y'_0| + \sum_{i=0}^{\ell'} |Y''_i| \leq O(|X|) + \ell' \cdot \max_{R \in \mathcal{R}} |R|.$$

Since \mathcal{R} is finite, the maximum in the above bound is a constant—it therefore remains to bound the number of (non-empty) protrusions ℓ .

We finally invoke the second statement of Lemma 67: applying it to Y'_0 it states that

$$|\{\partial C\}_{C \in K(G' \setminus Y'_0)}| \leq (2^{\rho h^2} + \rho h^2) |Y'_0| = O(|X|).$$

Now since every protrusion in $Y'_0 \uplus Y'_1 \uplus \dots \uplus Y'_{\ell'}$ has—by definition—its unique boundary in Y'_0 and the replacement does not change the respective boundaries, we can bound the number of protrusions by the number of possible boundaries. Specifically:

$$\ell' = |\{\partial Y'_i\}_{1 \leq i \leq \ell'}| \leq |\{\partial C\}_{C \in K(G' \setminus Y'_0)}| = O(|X|).$$

We conclude that $|G'| = O(|X|)$ as claimed. \square

The one crucial missing piece from Lemma 69 to Theorem 27 is that we are not given a treewidth modulator as input. However, note that the proof of Lemma 69 only needs the modulator to construct the protrusion decomposition, and the decomposition to reduce the protrusions—if we can replace protrusions *without* using the decomposition, we are done.

To achieve this feat in the claimed running time we employ the following nice result by Fomin, Lokshtanov, Misra, Ramanujan, and Saurabh (improving on a previous randomised algorithm [99]).

Theorem 30 (Fomin *et al.* [98]). *Let Π be a problem that has a protrusion replacer which replaces r -protrusions of size at least r' and let s and β be constants such that $r \geq 3(\beta + 1)$ and $s \geq 2^r r'$.*

Given an instance (G, k) as input, there exists a linear-time algorithm that produces an equivalent instance (G', k') with $|G'| \leq |G|$ and $k' \leq k$. If additionally G has a protrusion-decomposition $Y_0 \uplus Y_1 \uplus \dots \uplus Y_\ell$ of width at most β and satisfies $\max\{|Y_0|, \ell\} \leq \frac{n}{244s}$, then we have $|G'| \leq (1 - \delta)|G|$ for some constant δ .

The bottom line is: we can replace ‘enough’ protrusions using a linear-time algorithm such that the size of the resulting kernel is only larger by a constant factor compared to the kernel obtained by replacing every protrusion. Combining Lemma 69 with Lemma 30 proves Theorem 27.

12.2 INTERLUDE: THE NEED FOR A STRONGER PARAMETER

As it turns out, we cannot obtain linear kernels using treewidth-modulators as parameter for the next larger sparse class; graphs of bounded expansion. To see this, we combine the following result with a simple observation.

Theorem 31 (Dell and van Melkebeek [242]). *Let \mathcal{P} be a monotone graph property that is satisfied by infinitely many but not all graphs. Let ε be a positive real. If $\text{coNP} \not\subseteq \text{NP}/\text{poly}$, there is no protocol of cost $O(k^{2-\varepsilon})$ for deciding whether a graph satisfying \mathcal{P} can be obtained from a given graph by removing at most k vertices, even when the first player is co-nondeterministic*

Luckily, the authors spell out the consequences of their results for people not versed in communication complexity: the k -vertex deletion problem for such properties \mathcal{P} does not admit subquadratic kernels. Famous problems that can be modelled with such properties are VERTEX COVER (where \mathcal{P} is being an edgeless graph), FEEDBACK VERTEX SET (where \mathcal{P} is being free of cycles) or CHORDAL VERTEX DELETION (where \mathcal{P} is, well, being chordal). To avoid confusion: both VERTEX COVER and FEEDBACK VERTEX SET allow kernels that have only $2k$ vertices, however, the above states that we cannot give any guarantee about the number of edges in general graphs.

We make the following simple observation:

Lemma 70. *Fix a function $f \geq 2$ and let \mathcal{G}_f be the class of all graphs G with $\nabla_r(G) \leq f(r)$ for all $r \geq 0$. Let Π a parametrised problem that is closed under edge subdivision. Then if Π admits kernels of size $g(k)$ in \mathcal{G}_f then Π admits kernels of size $g(k)$ in general graphs.*

Proof. Since Π is closed under edge subdivision, we have that for every graph G and all integers k, ℓ it holds that

$$(G, k) \in \Pi \iff (G[E/\ell], k) \in \Pi.$$

Since $\nabla_r(G[E/|G|]) \leq 2 \leq f(r)$ we have that $G[E/|G|] \in \mathcal{G}_f$. As such, a kernel of size $g(k)$ for $G[E/|G|]$ is also a kernel for G . \square

As a corollary to Theorem 31 and Lemma 70 we see that the property of ‘just’ having bounded expansion is not enough to facilitate linear kernels for certain problems—we seem to need a parametrisation which restricts the structure sufficiently.

Corollary 16. *Fix a function $f \geq 2$ and let \mathcal{G}_f be the class of all graphs G with $\nabla_r(G) \leq f(r)$ for all $r \geq 0$. Unless $\text{coNP} \not\subseteq \text{NP/poly}$, FEEDBACK VERTEX SET restricted to \mathcal{G}_f does not admit a kernel of size $O(k^{2-\epsilon})$.*

In general, we can state that problems that are invariant under edge subdivision cannot profit from structural sparseness. As an aside, this also holds true for the graph isomorphism problem; which is solvable in time $n^{f(H)}$ for classes excluding a topological minor [130].

Structural parameters

In the parametrised framework, we have a way of addressing this issue rather elegantly: we simply change the parametrisation. The first important realisation is that all previous meta-kernelisation results, be it in planar, bounded-genus or minor-excluding classes, can be rephrased using a *structural* parameter. For example, combined properties of bidimensionality and separability (used to prove the result on minor-free graphs) imply that the problem is treewidth-bounding (cf. Lemma 3.2 and 3.3 in [101]). That is: the natural parameter *implies* the existence of a treewidth-modulator, where the treewidth of the remaining graph depends only on the problem. Quasi-coverability on bounded genus graphs implies the same (cf. Lemma 6.4 in [23]). These two meta-theorems alongside the one proved in the previous section can therefore be rephrased using the size of treewidth-modulator as a *structural* parameter.

In the context of bounded-expansion and nowhere dense classes, parametrising by a treewidth-modulator is out of the question: Again, this parameter is closed under edge subdivisions. But keeping with the theme of modulating to a width-measure, another candidate parameter emerges naturally: a modulator to bounded treedepth.

12.3 CLASSES OF BOUNDED EXPANSION AND BEYOND

We will first prove that problems that have finite integer index (either in general graphs or in graphs of bounded treedepth) admit linear kernels if a) inputs are restricted to a hereditary bounded-expansion class and b) they are parametrised by a modulator to constant treedepth.

Theorem 28. *Let Π be a graph problem that has f.i.i. on graphs of bounded treedepth. Let \mathcal{G} be a graph class of bounded expansion and let $t \in \mathbb{N}$ be a*

constant. Then there is an algorithm that takes as input $(G, \xi) \in \mathcal{G} \times \mathbb{N}$ and, in time $O(|G| + \log \xi)$, outputs (G', ξ') such that

1. $(G, \xi) \in \Pi$ if and only if $(G', \xi') \in \Pi$;
2. G' is an induced subgraph of G ; and
3. $|G'| = O(\mathbf{td}_t^\Delta(G))$.

The proof of the corresponding theorem for nowhere dense classes, Theorem 32, works very similarly and we elaborate on the necessary changes after the proof of Theorem 28. As a first ingredient, we need to be able to compute an (approximate) treedepth modulator which then forms the core around which the kernelisation routine works.

Lemma 71. *Fix a constant $t \in \mathbb{N}$. Given a graph G from a class \mathcal{G} , a treedepth- d modulator of size at most $2^d \mathbf{td}_t^\Delta(G)$ can be computed in time $O(\mathbf{td}_t^\Delta(G) \|G\|)$. If \mathcal{G} has bounded expansion, we can compute such a modulator in time $O(|G|)$.*

Proof. Note that any path of length $2^d + 1$ must intersect any treedepth- d modulator. We will call such a path a *long* path in the following.

We start with an empty vertex set S_0 . By Proposition 4 we can compute a decomposition F of width at most $2^{\mathbf{td}(G)}$ in time $O(|G| + \|G\|)$ by computing a dfs forest. If the height of F is larger than 2^d , we find a long path P in F and include it into S_0 . We remove $V(P)$ from the graph and iterate this procedure, until the resulting decomposition F' of $G \setminus S_0$ has height at most 2^d . This gives us a tree (path) decomposition of the graph of width at most $2^d - 1$. Now use standard tools (e.g., Courcelle’s Theorem) to obtain an optimal treedepth- d modulator S_1 in $G \setminus S_0$, and set $S = S_0 \cup S_1$. Since the treewidth of $G \setminus S_0$ is a constant, the latter algorithm runs in time linear in the size of the graph. The overall size of the modulator S is at most 2^d times the optimal solution.

For graph classes of bounded expansion we change the computation of S_0 by using Corollary 10 with a $(2^d + 2)$ -centred colouring computed in linear time (cf. Theorem 17): With the forbidden family $\mathcal{F} = \{P_{2^d+1}\}$ the approximate solution to \mathcal{F} -DELETION yields S_0 . Note that the approximation factor only depends on d and the graph class. \square

For the remainder of this section, let \mathcal{G} be a graph class whose expansion is bounded by f .

Recall the first statement of Lemma 36 which says that in a bipartite graph with partite sets X, Y the number of vertices in Y of large degree is bounded linearly in $|X|$ with a factor that depends on $\nabla_{1/2}$. The following is a simple corollary to that fact and we will use it, as before, to bound the size of the witness structure computed by Algorithm 2.

Corollary 17. *Suppose that for $G \in \mathcal{G}$ and $X \subseteq V(G)$, the connected components $\{C_1, \dots, C_p\} = \mathbf{K}(G \setminus X)$ of $G \setminus X$ satisfy the following two conditions:*

1. $\text{diam}(G[V(C_i)]) \leq r$ and
2. $|N(C_i) \cap X| > 2f(r+1)$ for $1 \leq i \leq p$.

Then $p \leq 2f(r+1) \cdot |X|$.

Proof. We construct an auxiliary bipartite graph G' with partite sets X and $Y = \{C_1, \dots, C_p\}$. There is an edge between C_i and $x \in X$ in G' iff $x \in N(C_i) \cap X$. Note that G' is a depth- r shallow minor of G with branch sets C_i .

First, let us prove that for every $H \preceq_t^1 G'$ it holds that $H \preceq_t^{r+1} G$. To this end, let ϕ' be the minor embedding witnessing that $H \preceq_t^1 G'$. Note that the branch sets $\{\phi'(x)\}_{x \in H}$ induce stars with centres either in X or Y . Therefore, if we construct a minor embedding ϕ of H in G by simply mapping the vertex C_i in G' to the vertex set C_i in G , the depth of ϕ is at most $1 + \max_i \text{diam}(C_i) \leq r+1$. It follows that $H \preceq_t^{r+1} G$.

Consequently, we can apply Lemma 36 to G via G' and conclude that for $\tau = f(r+1) \geq \nabla_{r+1}(G) \geq \nabla_1(G')$ it holds that

$$p \leq 2\tau \cdot |X| = 2f(r+1) \cdot |X|,$$

as claimed. \square

We can use the second statement of Lemma 36 in a similar manner. Recall that it stated that in a bipartite graph with partite sets X, Y the number of possible neighbourhoods of vertices contained in Y is bounded linearly in $|X|$ with a factor that, again, depends on $\nabla_{1/2}$. We will use it to bound the number of protrusions in a protrusion decomposition, using the fact that there is a one-to-one correspondence between the protrusions and the boundaries attaching them to the core.

Corollary 18. *Suppose that for $G \in \mathcal{G}$ and $X \subseteq V(G)$, the connected components $\{C_1, \dots, C_p\} = \mathcal{K}(G \setminus X)$ of $G \setminus X$ are partitioned into equivalence classes $\mathcal{C}_1, \dots, \mathcal{C}_\ell$ where*

$$C, C' \in \mathcal{C} \iff N(C) \cap X = N(C') \cap X.$$

Let further $r \geq \max_{1 \leq j \leq p} \text{diam}(C_j)$ be a bound on the diameter of the components of $G \setminus X$. Then the number of classes ℓ is bounded by $\ell \leq (4^{f(r+1)} + 2f(r+1)) \cdot |X|$.

Proof. As in the proof of Corollary 17, we construct a bipartite graph G' with partite sets X and $Y = \{C_1, \dots, C_r\}$. By the previous observation in the proof of Corollary 17, we have that every graph $H \preceq_t^1 G'$ satisfies $H \preceq_t^{r+1} G$. Choosing again $\tau := f(r+1) \geq \nabla_{r+1}(G) \geq \nabla_1(G')$ we apply the second statement of Lemma 36 and obtain the bound

$$\ell \leq (4^\tau + 2\tau) \cdot |X| = (4^{f(r+1)} + 2f(r+1)) \cdot |X|$$

as claimed. \square

With the above lemmas at hand we can now prove the main theorem of this section. The proof is similar to that of Theorem 27.

Proof of Theorem 28. Given an instance (G, ξ) of Π with $G \in \mathcal{G}$, we use Lemma 71 to calculate a treedepth- d modulator X of size at most $2^d \mathbf{td}_d^\Delta(G)$ in linear time.

We construct a protrusion^{td} decomposition $Y_0 \uplus Y_1 \uplus \dots \uplus Y_\ell$ around the core $Y_0 = X$ in linear time. Since the protrusions have treedepth at most d they have diameter at most 2^d . We therefore apply Algorithm 2 with parameter $\alpha = 2f(2^d + 1)$ and obtain, by Lemma 66 a protrusion decomposition $Y'_0 \uplus Y'_1 \uplus \dots \uplus Y'_\ell$ with adhesion at most¹ $\alpha' := \alpha + 2d$ and a witness structure W_1, \dots, W_p . By applying Corollary 17 to the subgraph induced by the witness structure together with X , we obtain that the number of witness sets is bounded by $p \leq 2f(2^d + 1) \cdot |X|$

It follows (cf. again Lemma 66) that the size of the new core Y'_0 is

$$|Y'_0| \leq 2p(d + 1) + |Y_0| = 2f(2^d + 1) \cdot |X| + |X| = O(|X|).$$

Let us fix for every protrusion Y'_i a $\leq \alpha'$ -labelling of its boundary b_i , which enables us to refer to the $\leq \alpha'$ -boundaried graphs ${}^\circ Y'_i := G[Y'_i]_{b_i}$ in the following and avoid clumsy notation.

Applying the protrusion reduction machinery, we replace every protrusion Y'_i , $1 \leq i \leq \ell$, by a representative of constant size. To this end, we apply Lemma 65 choosing the following parameters:

- The problem is Π ;
- the width-measure is treedepth;
- the graph relation is the induced subgraph relation;
- the parameters are $\alpha = \alpha' = 2f(2^d + 1) + 2d$ and $\beta = d$.

Since Π has f.i.i. on graphs of bounded treedepth, we assume that the finite set of canonical, monotone, and induced-subgraph-preserving representatives \mathcal{R} for Π on $\mathcal{G}_{\mathbf{td} \leq \alpha + \beta}$ is part of the algorithm.

Applying Lemma 65 with the above parameters to (G, ξ) and a protrusion Y'_i with boundary b_i provides us in time $O(|Y'_i|)$ with the representative ${}^\circ Y''_i := \mathcal{R}({}^\circ Y'_i)$. Given all such representatives ${}^\circ Y''_i$ for every $1 \leq i \leq \ell$, we compute a sequence of instances (G_i, ξ_i) as follows: let the first instance be $(G_0, \xi_0) := (G, \xi)$. Then (G_i, ξ_i) is obtained from (G_{i-1}, ξ_{i-1}) by replacing a protrusion and adjusting the parameter accordingly:

$$\begin{aligned} G_i &:= G_{i-1}[Y'_i \mapsto {}^\circ Y''_i]_{b_i} \\ \xi_i &:= \xi_{i-1} - \Delta_{\Pi \cap \mathcal{G}}({}^\circ Y''_i, {}^\circ Y'_i). \end{aligned}$$

Since Lemma 65 computes representatives that are canonical, we know that $G[Y'_0]$ is not affected by these operations and therefore the sequence can be constructed as described above. Because the representatives are monotone, we have that $\Delta_{\Pi \cap \mathcal{G}}({}^\circ Y''_i, {}^\circ Y'_i) \geq 0$ and therefore

¹ the missing offset of two traces back to the bound $\mathbf{tw} \leq \mathbf{td} - 1$

it holds that $\xi_i \leq \xi_{i-1}$. Since the representatives are simply induced subgraphs, we have that $G_i \subseteq_i G_{i-1}$.

Most importantly, we have that

$$(G_i, \xi_i) \in \Pi \iff (G_{i-1}, \xi_{i-1}) \in \Pi.$$

We set $(G', \xi') = (G_{\ell'}, \xi_{\ell'})$; the first two claims then follow via induction over the above sequence. It is left to show the size bound. Note that the total size of the graph is

$$|G'| \leq |Y'_0| + \sum_{i=0}^{\ell'} |Y'_i| \leq O(|X|) + \ell' \cdot \max_{R \in \mathcal{R}} |R|.$$

Since \mathcal{R} is finite, the maximum in the above bound is a constant—it therefore remains to bound the number of (non-empty) protrusions ℓ .

We now invoke Corollary 18: applying it to Y'_0 it states that

$$|\{\partial C\}_{C \in \mathcal{K}(G' \setminus Y'_0)}| \leq (4^{f(2^d+1)+2f(2^d+1)}) \cdot |Y'_0| = O(|X|)$$

using again the fact that the diameter of the reduced protrusions is at most 2^d .

Since every protrusion in $Y'_0 \uplus Y'_1 \uplus \dots \uplus Y'_{\ell'}$ has—by definition—its unique boundary in Y'_0 and the replacement does not change the respective boundaries, we can bound the number of protrusions by the number of possible boundaries. Specifically:

$$\ell' = |\{\partial Y'_i\}_{1 \leq i \leq \ell'}| \leq |\{\partial C\}_{C \in \mathcal{K}(G' \setminus Y'_0)}| = O(|X|).$$

We conclude that $|G'| = O(|X|) = O(\mathbf{td}_d^\Delta(G))$ as claimed. \square

The proof of Theorem 32, restated below, works exactly as the one above; we only have to replace Corollary 17 and 18 by suitable alternatives for the nowhere dense case.

Theorem 32. *Let Π be a graph problem that has f.i.i. on graphs of bounded treedepth. Let \mathcal{G} be a nowhere dense graph class and let $t \in \mathbb{N}$ be a constant. Then there is an algorithm that takes as input $(G, \xi) \in \mathcal{G} \times \mathbb{N}$ and, in time $O(|G|^{1+o(1)})$, outputs (G', ξ') such that*

1. $(G, \xi) \in \Pi$ if and only if $(G', \xi') \in \Pi$;
2. G' is an induced subgraph of G ; and
3. $|G'| = O(\mathbf{td}_t^\Delta(G)^{1+o(1)})$.

Let us quickly prove the two necessary corollaries. Again, we essentially apply Lemma 36 to a suitable shallow minor in order to obtain a bound in the witness-structure outputted by Algorithm 2 and a bound on the number of protrusions.

Corollary 19. *Suppose that for $G \in \mathcal{G}$, where \mathcal{G} is nowhere dense, and $X \subseteq V(G)$, the connected components $\{C_1, \dots, C_p\} = \mathcal{K}(G \setminus X)$ of $G \setminus X$ satisfy the following two conditions:*

1. $\text{diam}(G[V(C_i)]) \leq r$ and
2. $|N(C_i) \cap X| > \tilde{\nabla}_0((G \tilde{\nabla}(r+1))_{\leq |X|})$ for $1 \leq i \leq p$.

Then $p = O(|X|^{1+o(1)})$.

Proof. We construct an auxiliary bipartite graph G' with partite sets X and $Y = \{C_1, \dots, C_p\}$. There is an edge between C_i and $x \in X$ in G' iff $x \in N(C_i) \cap X$. Note that G' is a depth- r shallow minor of G with branch sets C_i . As shown in the proof of Corollary 17, for every $H \preceq_t^1 G'$ it holds that $H \preceq_t^{r+1} G$.

Consequently, we can apply Lemma 36 to G via G' and conclude that for $\tau = \tilde{\nabla}_0((G \tilde{\nabla}(r+1))_{\leq |X|}) \geq \tilde{\nabla}_0((G' \tilde{\nabla} 1/2)_{\leq |X|})$, it holds that

$$p \leq 2\tau \cdot |X| = 2|X|^{o(1)} \cdot |X| = 2|X|^{1+o(1)},$$

where we used that $\tilde{\nabla}_0((G \tilde{\nabla}(r+1))_{\leq |X|}) = O(|X|^{o(1)})$ in nowhere dense classes. \square

Corollary 20. *Suppose that for $G \in \mathcal{G}$, where \mathcal{G} is nowhere dense, and $X \subseteq V(G)$, the connected components $\{C_1, \dots, C_p\} = K(G \setminus X)$ of $G \setminus X$ are partitioned into equivalence classes $\mathcal{C}_1, \dots, \mathcal{C}_\ell$ where*

$$C, C' \in \mathcal{C} \iff N(C) \cap X = N(C') \cap X.$$

Let further $r \geq \max_{1 \leq j \leq p} \text{diam}(C_j)$ be a bound on the diameter of the components of $G \setminus X$. Then the number of classes ℓ is bounded by $O(|X|^{1+o(1)})$.

Proof. As in the proof of Corollary 17, we construct a bipartite graph G' with partite sets X and $Y = \{C_1, \dots, C_r\}$. By the previous observation in the proof of Corollary 17, we have that every graph $H \preceq_t^1 G'$ satisfies $H \preceq_t^{r+1} G$. Choosing again $\tau := \tilde{\nabla}_0((G \tilde{\nabla}(r+1))_{\leq |X|})$ and $\omega = \omega((G \tilde{\nabla}(r+1))_{\leq |X|}) \geq \omega((G' \tilde{\nabla} 1/2)_{\leq |X|})$, we apply the second statement of Lemma 36 and obtain the bound

$$\ell \leq (\omega(e\tau)^\omega + 2\tau) \cdot |X| = \omega(e\omega)^\omega |X|^{1+o(1)} + 2|X|^{1+o(1)},$$

and this last expression is $O(|X|^{1+o(1)})$, since ω is a constant. \square

Proof of Theorem 32. We proceed exactly as in the proof of Theorem 28 and construct a reduced protrusion-decomposition $Y'_0 \uplus Y'_1 \uplus \dots \uplus Y'_{\ell'}$. Using Corollary 17, we obtain that the number of witness sets returned by Algorithm 2 is at most $O(|X|^{1+o(1)})$ and therefore that $|Y'_0| = O(|X|^{1+o(1)})$.

Applying Corollary 20 to Y'_0 now yields

$$|\{\partial C\}_{C \in K(G \setminus Y'_0)}| \leq O(|Y'_0|^{1+o(1)}) = O(|X|^{1+o(1)}),$$

using again the fact that the diameter of the reduced protrusions is at most 2^d .

We again bound the number of protrusions by the number of possible boundaries:

$$\ell' = |\{\partial Y'_i\}_{1 \leq i \leq \ell'}| \leq |\{\partial C\}_{C \in K(G \setminus Y'_0)}| = O(|X|^{1+o(1)}).$$

We conclude that $|G'| = O(|X|^{1+o(1)}) = O(\mathbf{td}_d^\Delta(G)^{1+o(1)})$ as claimed. \square

12.4 FURTHER WORK AND OPEN QUESTIONS

As argued in Section 12.2, it is unlikely that we can extend the meta-kernelisation framework to bounded expansion classes using a weaker parameter. We will see in the next chapter that there is hope to obtain linear kernels for some problems using their natural parametrisation. Aside from this line of work, several interesting developments can be noted since we introduced the ‘structural’ perspective into meta-kernelisation. At this point we should note that it was Jansen’s programme on kernelisation by structural parameters (see, e.g., his dissertation [144]) who introduced the idea of replacing the natural parameter by a more promising, structure-based one.

First, we should ask where the limitation of the protrusion-based approach lie. The number of problems it can be applied to is impressive, but there exist natural problem for which even the parametrisation to a treedepth-modulator does not seem to help. As shown in Section 11.1.1, INDEPENDENT DOMINATING SET does not have finite integer index even in the class of stars, *i.e.* graphs of treedepth two. Some form of monotonicity therefore seems to be an indispensable property in order to apply the meta-kernelisation framework, even if we restrict ourselves to very simple forms of protrusions.

Garnero, Paul, Sau, and Thilikos have recently demonstrated that for a large subset of problems, the protrusion replacement used for Theorems 27, 28 and 32 can be made uniform by using a dynamic programming [121].

Ganian, Slivovsky, and Szeider independently extended the meta-kernelisation framework by using a structural parametrisation that is orthogonal to modulators [119]: specifically, they parametrise by the number of modules into which the graph can be partitioned such that each module induces a graph of bounded rankwidth. Using this parameter, all problems expressible in MSO_1 admit polynomial kernels.

Recently, Eiben, Ganian, and Szeider refined the notion of structural parameters combining the ideas of a modulator to some bounded with-measure with the ‘internal’ complexity of such a set (the exact definition is rather technical) [83]. As a result, they can apply Theorem 28 and 32 with a potentially smaller parameter.

Modulators to bounded treewidth have been used to obtain *uniform* polynomial kernels for \mathcal{H} -MINOR DELETION where \mathcal{H} contains at least one planar graph [122] (uniform meaning that the polynomial does not depend on \mathcal{H}). In the same paper, the authors also demonstrate a $O(k^6)$ kernel for TREEDPTH- t -DELETION. Modulators can also be used to define *backdoors* in SAT-formulas. In particular, one can in *fpt*-time decide whether a formula contains such a *backdoor to small*

treewidth [98]. Modulators to bounded treewidth have also been employed to preprocess ILP instances [146].

Open question 7. In the spirit of the work by Garnero *et al.*, can the non-uniformity of the protrusion replacement be circumvented for *all* problems that admit dynamic programming routines on graphs with bounded width-measures?

Open question 8. Can the non-uniformity of the protrusion replacement be avoided using Turing-kernelisation? Jansen showed that in the case of k -PATH, this is possible for planar and other classes [145].

Open question 9. Can the meta-kernelisation framework be applied to *dense* classes? In particular, can we use modulators to a dense width-measure like cliquewidth or rankwidth and obtain polynomial kernels in a suitable graph class?

Open question 10. One implication of the above result is that a large set of problems is fpt in structurally sparse classes if parametrised by a treedepth-modulator. Is there a direct way—maybe using dynamic programming on the protrusions—of designing such algorithms?

Reading the original meta-kernelisation paper, one cannot help but feel a certain polarity between the problems affected by the result. On the one side, we have problems like `FEEDBACK VERTEX SET` and `\mathcal{H} -MINOR DELETION` (with at least one planar graph in \mathcal{H}) whose solution structures the remainder of the graph by providing a bound on the treewidth. On the other side, problems like `DOMINATING SET` and `INDEPENDENT SET` do not structure the graph outside a solution in any direct way, but we know that all vertices have to lie within a constant radius around the solution. This mix of problems is still approachable in apex-minor free graphs [101], but for classes excluding an arbitrary minor only the first flavour of problems can be tamed. This pattern continued the previous chapters, where we have seen how the meta-kernelisation framework can be lifted for the second flavour of problems only by reintroducing additional structure through a suitable parametrisation.

A novel approach

However, in a different line of work Fomin, Lokshantov, Saurabh, and Thilikos showed that `DOMINATING SET` and `CONNECTED DOMINATING SET` admit linear kernels in classes excluding a minor [102] and classes excluding a topological minor [103]. Both times the respective decomposition theorems [216, 130] were the key. Motivated by these results, we show that indeed `DOMINATING SET` still admits a linear kernel in bounded expansion classes and an almost-linear kernel in nowhere dense classes under its natural parametrisation. Moreover, the approach presented here is radically different from previous approaches: without a suitable decomposition theorem, we had to resort to more fundamental arguments. The twin and the charging lemma introduced in Section 8.1 make another appearance here, as well as our reformulation of Dvořák’s algorithm from Section 10.3.

Domination core

The kernelisation routine is split into two phases: in the first phase we identify a suitably small subset of vertices called the *domination core*. This set has the property of acting as a ‘proxy’: if every vertex in the core is dominated, then so is the rest of the graph. Once we have computed a small core, the kernelisation proceeds by reducing the remainder of the graph until its size is comparable to that of the core.

13.1 FINDING A SMALL DOMINATION CORE

The central concept that facilitates the kernel will be that of a *domination core*: a subset of vertices in a graph that can ‘witness’ that another set is dominating *the whole graph*. Similar ideas are abundant in

combinatorial arguments; the famous sunflower reduction rule being probably the most prominent example. However, the idea of a core specifically for DOMINATING SET needs to be credited to Dawar and Kreutzer who used it to obtain an fpt-algorithm for r -DOMINATING SET in nowhere dense classes [60].

Definition 27 (Domination core). Let G be a graph and $Z \subseteq V(G)$ a vertex set. We say that Z is a *domination core* in G if every minimum-size Z -dominator of G is also a dominating set in G .

Clearly, the vertex set $V(G)$ is a domination core; the goal will be to find a sufficiently small one and later build a kernelisation routine around it. Note that if Z is a domination core, then $\mathbf{ds}(G) = \mathbf{ds}(G, Z)$. As the first important step we will prove the following theorem.

Theorem 33. *Given an instance (G, k) of DOMINATING SET, we can in time $\|G\| \cdot \tilde{\nabla}_2(G)^{\omega_2(G)^{O(1)}}$ either*

1. *prove that $\mathbf{ds}(G) > k$ or*
2. *find a domination core $Z \subseteq V(G)$ with $|Z| \leq c_Z \cdot k$ where c_Z (defined below) depends only on $\tilde{\nabla}_2(G)$ and $\omega_{1/2}(G)$.*

Let us fix an instance (G, k) for the remainder of this section. We prove Theorem 33 by starting with the trivial domination core $Z = V(G)$ and gradually reduce $|Z|$.

For convenience and readability, we define the following quantities related to the density of G and will use them lavishly for the remainder of this section:

$$\begin{aligned} \tau &:= \tilde{\nabla}_{1/2}(G) & \omega &:= \tilde{\omega}_{1/2}(G) & \xi &:= \min\{2\tilde{\nabla}_0(G), \omega\} \\ \rho &:= 2\tilde{\nabla}_0(G)(2\omega(e\tau)^\omega + 1), \end{aligned}$$

where c_1 is the approximation ratio of Lemma 50 for $r = 1$. We also will need the rather horrible quantities

$$\begin{aligned} c_Z &:= (\omega^2 + 1)(2\omega(e\tau)^\omega) \\ &\quad \cdot (1 + c_1 \cdot 2^{12\omega_2(G)+55} \rho^{18} \cdot \omega_2(G)^2 \cdot \tilde{\nabla}_2(G)^{18\omega_2(G)+11})^\xi \cdot c_1 \\ c_X &:= (1 + c_1 \cdot 2^{12\omega_2(G)+55} \rho^{18} \cdot \omega_2(G)^2 \cdot \tilde{\nabla}_2(G)^{18\omega_2(G)+11})^{\xi-1} \cdot c_1 \\ c_{SX} &:= 2^{12\omega_2(G)+55} \rho^{18} \cdot \omega_2(G)^2 \cdot \tilde{\nabla}_2(G)^{18\omega_2(G)+11}. \end{aligned}$$

The following lemma is the crucial tool in reducing the size of Z .

Lemma 72. *Given a domination core Z of size $> c_Z \cdot k$, we are able to either conclude that $\mathbf{ds}(G) > k$, or find a non-empty subset $Z^* \subseteq Z$ such that $Z \setminus Z^*$ is still a domination core in time $O(\|G\| \cdot c_{SX} \cdot c_X \cdot k)$.*

Theorem 33 follows directly from Lemma 72, we dedicate the rest of this section to its proof.

The first step of the core reduction is an algorithm that computes sets (X, S) where X is a Z -dominator of size $O(k)$ and $S \subseteq Z$ is a

sufficiently large 3-scattered set in $G \setminus X$. With this pair of sets it will be possible to find irrelevant vertices in Z .

To construct (X, S) , we first apply Lemma 51 with $r = 1$ and Z, k as given. We either obtain a two-scattered set of size $k + 1$, in which case we conclude that $\mathbf{ds}(G) > k$, or a dominating set X_0 of G with $|X_0| \leq c_1 k$. In the latter case we construct a sequence $(X_i)_{i \in \mathbb{N}}$ beginning with the dominating set X_0 as follows. Define the functions

$$\begin{aligned} g(x) &:= 2^{12\omega_2(G)+55} \rho^{18} \cdot \omega_2(G)^2 \cdot x^{18\omega_2(G)+11} \\ f_i(x) &:= (1 + c_1 \cdot g(x))^i \cdot c_1 \quad \text{for } i \geq 0. \end{aligned}$$

Throughout the construction, we will maintain the invariant $|X_i| \leq f_i(\tilde{\nabla}_2(G)) \cdot k$. Note that this invariant is indeed satisfied for $i = 0$. We iteratively proceed as follows:

1. Apply the algorithm of Lemma 51 to the graph $G \setminus X_i$, using the set $Z \setminus X_i$, the integer $g(\tilde{\nabla}_2(G)) \cdot |X_i|$ and $r = 1$ as parameters.
2. If the algorithm has found a 3-scattered set $S \subseteq Z \setminus X_i$ of cardinality larger than $g(\tilde{\nabla}_2(G)) \cdot |X_i|$, we let $X = X_i$ and continue with the second phase using the tuple (X, S) .
3. Otherwise, the algorithm found a $(Z \setminus X_i)$ -dominator D_{i+1} in $G \setminus X_i$ of size at most

$$|D_{i+1}| \leq c_1 g(\tilde{\nabla}_2(G)) \cdot |X_i|.$$

We set $X_{i+1} = X_i \cup D_{i+1}$ and repeat with the next i . Note that

$$\begin{aligned} |X_{i+1}| &\leq (1 + c_1 g(\tilde{\nabla}_2(G))) \cdot |X_i| \\ &\leq (1 + c_1 g(\tilde{\nabla}_2(G))) \cdot f_i(\tilde{\nabla}_2(G)) \cdot k \\ &= f_{i+1}(\tilde{\nabla}_2(G)) \cdot k, \end{aligned}$$

so the invariant on the size of X_i is maintained.

In a nutshell, the above algorithm iteratively extracts Z -dominators D_1, D_2, \dots and constructs sets X_1, X_2, \dots until the second case occurs. Before we show that the algorithm necessarily terminates after few steps, let us make use of the helping functions f_i , to rewrite c_Z, c_X into the more digestible form

$$\begin{aligned} c_Z &:= (\omega^2 + 1)(2\omega(e\tau)^\omega) \cdot f_\xi(\tilde{\nabla}_2(G)) \\ c_X &:= f_{\xi-1}(\tilde{\nabla}_2(G)) \\ c_{SX} &:= g(\tilde{\nabla}_2(G)). \end{aligned}$$

This should give some intuition how these values were chosen in the first place. Let us return our attention to the above algorithm.

Lemma 73. *The above algorithm terminates with a tuple (X, S) after less than $\xi = \min\{2\tilde{\nabla}_0(G), \omega\}$ iterations, provided that $|Z| \geq c_Z \cdot k$.*

Proof. Assume towards a contradiction that the algorithm iterates ξ or more steps, in particular it constructs the set X_ξ . By the invariant, we have that $|X_\xi| \leq f_\xi(\tilde{\nabla}_2(G)) \cdot k$. Consider any vertex in $Z \setminus X_\xi$: it is dominated by every set $X_0, D_1, D_2, \dots, D_\xi$ and hence has at least one neighbour in each of these sets.

Recall that Lemma 10 states that a d -degenerate bipartite graph (A, B, E) cannot contain more than $d|A|$ vertices of degree at least d in B . Applying this fact to X_ξ and Z means that either $|Z| \leq 2\nabla_0(G) \cdot |X_\xi|$ or $\xi < 2\tilde{\nabla}_0(G)$. Since the former does not hold due to our choice of c_Z , the latter has to and we have shown the first part of the claim.

Applying Lemma 36 to X_ξ and $Z \setminus X_\xi$ we obtain that Z has at most

$$\begin{aligned} & (2\omega(e\tau)^\omega) \cdot |X_\xi| \\ & \leq (2\omega(e\tau)^\omega) \cdot f_\xi(\tilde{\nabla}_2(G)) \cdot k \end{aligned}$$

different twin-classes with respect to X_ξ . This term should look familiar, indeed if we divide $c_Z \cdot k \leq |Z|$ by it to estimate the *size* of the twin-classes we see that

$$\frac{(\omega^2 + 1)(2\omega(e\tau)^\omega) \cdot f_\xi(\tilde{\nabla}_2(G)) \cdot k}{(2\omega(e\tau)^\omega) \cdot f_\xi(\tilde{\nabla}_2(G)) \cdot k} = \omega^2 + 1,$$

from which we deduce that there exists at least one twin-class of size at least $\omega^2 + 1$. Then necessarily $\xi < \omega$: otherwise this twin-class alongside its neighbourhood in X induces a supergraph of $K_{\omega+1, (\omega+1)^2}$; which in turn contains $K_{\omega+1}$ has a $\frac{1}{2}$ -shallow topological minor—a contradiction. \square

It will be convenient in the following to bound the number of neighbours a vertex in S has inside X . The following lemma shows that we can enforce this property that by sacrificing a small fraction of S .

Lemma 74. *There exists a subset $S' \subseteq S$ such that*

$$|S'| \geq (g(\tilde{\nabla}_2(G))/2\tau) \cdot |X|$$

and every vertex $x \in S'$ satisfies $|N(x) \cap X| \leq 2\tau$.

Proof. Note that $2\tau \geq 2\nabla_0(G)$. By Lemma 10 the bipartite subgraph $(X, S, E(X, S))$ contains at most

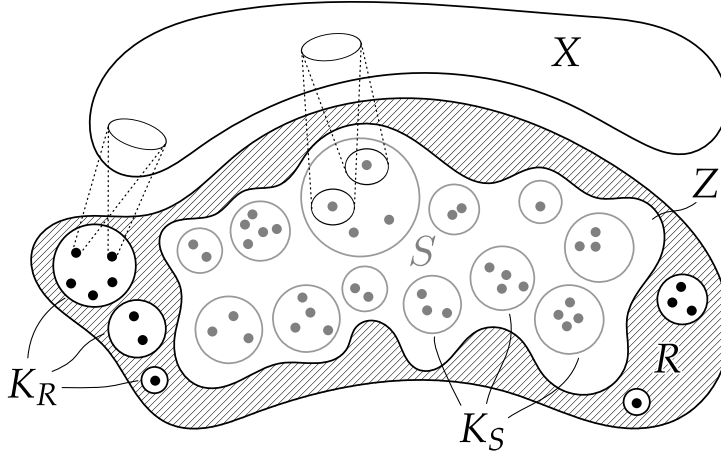
$$2\nabla_0(G) \cdot |X| \leq 2\tau \cdot |X|$$

vertices in S of degree larger than 2τ . Since by construction $|S| > g(\tilde{\nabla}_2(G))$ removing the high-degree vertices from S results in the claimed set S' . \square

To avoid even further variables we will from now on simply set $S := S'$. Since the above is hard to keep in mind, let us summarise the situation: we have our instance (G, k) alongside vertex sets X, S , and Z , where

- the domination core Z has size at least $c_Z k$,
- X is a Z -dominator of size at most $c_X k$,
- $S \subseteq Z \setminus X$ is 3-scattered in $G \setminus X$ with $|S| > c_{SX}|X|$, and
- $|N(v) \cap X| \leq 2\tau$ for every $v \in S$.

The goal is to show that we can remove vertices from Z in a way that leaves us with new domination core. The picture below summarises the situation at hand and contains some further sets defined below.



To proceed, let us define the following sets and auxiliary structures: let $R := N(S) \setminus X$ be those vertices in $G \setminus X$ that are not adjacent to S . Let B be the bipartite graph obtained from G by the following operation: we remove R from the graph and contract for every $x \in S$ the set $N(x) \setminus X$ onto x (naming the resulting vertex x for convenience) and then remove all edges between vertices in S and X . Since we do not keep edges between the vertices of S , the resulting minor is a half-shallow minor of G . We define $N'(s) := N_{G'}(s)$ in the following.

We will now group the vertices of S and R into equivalence classes according to how they can be dominated by the set X : let \equiv_X be defined via

$$\begin{aligned} u \equiv_X v \text{ for } u, v \in S &\text{ iff } N(u) \cap X = N(v) \cap X \text{ and } N'(u) = N'(v), \\ x \equiv_X y \text{ for } x, y \in R &\text{ iff } N(x) \cap X = N(y) \cap X, \\ u \not\equiv_X y &\text{ otherwise.} \end{aligned}$$

Note that S and R are, by construction, disjoint and therefore the above constraints can be met.

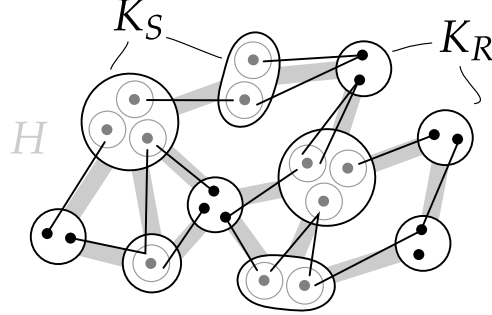
Next we construct an auxiliary graph H called the *class graph* to capture the interaction between the equivalence classes. Let $K_S := S / \equiv_X$ and $K_R := R / \equiv_X$ be the equivalence classes of \equiv_X . Let $\Gamma: K_S \cup K_R \rightarrow 2^{V(G)}$ be a function defined as

$$\Gamma(\kappa) := \begin{cases} \bigcup_{u \in \kappa} N'(u) \cup \kappa & \text{for } \kappa \in K_S \\ \kappa & \text{for } \kappa \in K_R. \end{cases}$$

We define the class graph H to contain the edge $\kappa_1\kappa_2$ iff the vertex sets $\Gamma(\kappa_1)$ and $\Gamma(\kappa_2)$ are connected by an edge in G , i.e.

$$\kappa_1\kappa_2 \in H \iff E_G(\Gamma(\kappa_1), \Gamma(\kappa_2)) \neq \emptyset.$$

The following pictorial representation should give some intuition about how the edges of the class graph reflect interactions between the equivalence classes.



Lemma 75 (Class graph size). *The number of classes is bounded by*

- $|K_S| \leq 8 \cdot 2^{12\omega_2(G)} \cdot \omega_2(G)^2 \cdot \tilde{\nabla}_{3/2}(G)^{18\omega_2(G)+1} \cdot |X|$ and
- $|K_R| \leq 2\omega(e\tau)^\omega \cdot |X|$,

respectively, where again $\tau = \tilde{\nabla}_{1/2}(G)$ and $\omega = \tilde{\omega}_{1/2}(G)$.

Proof. We invoke in both cases the twin class lemma (Lemma 36). For the classes K_R the bound follows directly: every class corresponds to exactly one unique neighbourhood set in X from vertices in R . The number of twin classes in R is hence at most

$$|K_R| \leq 2\omega(e\tau)^\omega \cdot |X|.$$

For the classes K_S we use the same argument, however, we need to work in the graph G' . Since G' is a $1/2$ -shallow minor of G , the number of twin-classes of S in G' is by the twin class lemma at most

$$\begin{aligned} & 2\tilde{\omega}_{1/2}(G')(e\tilde{\nabla}_{1/2}(G'))^{\tilde{\omega}_{1/2}(G')} \cdot |X| \\ & \leq 2\omega_2(G)(952 \cdot \tilde{\nabla}_{3/2}(G)^9)^{\omega_2(G)} \cdot |X| \end{aligned}$$

where we used that (since $G' \in G \nabla 1/2$)

$$\tilde{\omega}_{1/2}(G \nabla 1/2) = \omega((G \nabla 1/2) \tilde{\nabla} 1/2) \leq \omega(G \nabla 2) = \omega_2(G)$$

and, using Theorem 13, that

$$\tilde{\nabla}_{1/2}(G \nabla 1/2) \leq 350 \cdot \tilde{\nabla}_{1/2}(G \tilde{\nabla} 1/2)^9 = 350 \cdot \tilde{\nabla}_{3/2}(G)^9.$$

Because every vertex in S has at most 2τ neighbours in X , every set of vertices in S that share the same direct neighbourhood in X will be refined into at most $4\omega(e\tau)^\omega \tau$ equivalence classes under \equiv_X . Hence, we obtain the total bound

$$\begin{aligned} |K_S| & \leq 8\omega(e\tau)^\omega \tau \cdot \omega_2(G)(952 \cdot \tilde{\nabla}_{3/2}(G)^9)^{\omega_2(G)} \cdot |X| \\ & \leq 8 \cdot 2^{12\omega_2(G)} \cdot \omega_2(G)^2 \tilde{\nabla}_{3/2}(G)^{18\omega_2(G)+1} \cdot |X| \end{aligned}$$

and thus have shown the lemma. □

Not only is the size of the graph class only dependent on X and parameters related to two-shallow minors of G ; it turns out that its grad is closely related to that of G as well.

Lemma 76 (Class graph grad). *There exists a function f_H such that for every half-integer r it holds that $\tilde{\nabla}_r(H) \leq f_H(\tilde{\nabla}_{(10r+4)/2}(G))$. In particular, we have that*

$$\tilde{\nabla}_0(H) \leq (4(2\rho + 1)^2 \tilde{\nabla}_2(G) + 2\rho)^9$$

where ρ is defined as before.

Proof. We show that H is a 2-shallow minor of $G \bullet \bar{K}_{2\rho+1}$. Let $\hat{R} \subseteq R$ be a minimal set that contains exactly one representative per class K_R and $\hat{S} \subseteq S$ an analogous set for S .

Define the bipartite graphs $G_{\hat{R}} := (X \cup \hat{R}, E_G(X, \hat{R}))$ and $G_{\hat{S}} := (X \cup \hat{S}, E_G(X, \hat{S}))$. Applying Lemma 37, we obtain mappings $\phi_{\hat{R}}: \hat{R} \rightarrow X$ and $\phi_{\hat{S}}: \hat{S} \rightarrow X$ such that

- $x\phi(x) \in E(G_{\hat{R}})$ for $x \in R$
- $u\phi(u) \in E(G_{\hat{S}})$ for $u \in S$
- $|\phi_{\hat{R}}^{-1}(v)|, |\phi_{\hat{S}}^{-1}(v)| \leq \rho$ for every $v \in X$.

Note that for a class $\kappa \in K_R$ with representative $x = \kappa \cap \hat{R}$ it holds that every vertex $y \in \kappa$ is connected to $\phi_{\hat{R}}(x)$ in G (the same holds true for a class in K_S with respect to $\phi_{\hat{S}}$).

We combine these two mappings and obtain $\phi: \hat{R} \cup \hat{S} \rightarrow X$ in the obvious way. Now we can construct a supergraph H' of H as follows: starting with G , we replace every vertex $x \in X$ by $2\rho + 1$ copies $x_1, \dots, x_{2\rho+1}$. Note that this graph is a subgraph of $G \bullet \bar{K}_{2\rho+1}$. Now for every class $\kappa \in K_R \cup K_S$ we choose a distinct copy x_i of the vertex $x = \phi(\kappa \cap (\hat{R} \cup \hat{S}))$ —since the mapping ϕ charges x at most 2ρ times, each class can choose its own distinct copy—and we contract the vertex set $\Gamma(\kappa) \cup \{x_i\}$ into a single vertex labelled κ . A careful look at this construction reveals that the graph H' is a 2-shallow minor of $G \bullet \bar{K}_{2\rho+1}$, using the copies x_i , $x \in X$ as centres. Therefore we have for every $r \in \mathbb{N}$ that

$$\begin{aligned} \tilde{\nabla}_r(H) &\leq \tilde{\nabla}_r((G \bullet \bar{K}_{2\rho+1}) \nabla 2) \\ &\leq 2^3 3^{12} \tilde{\nabla}_r((G \bullet \bar{K}_{2\rho+1}) \tilde{\nabla} 2)^9 && \text{by Theorem 13} \\ &\leq 2^3 3^{12} \tilde{\nabla}_{(10r+4)/2}(G \bullet \bar{K}_{2\rho+1})^9 && \text{by Lemma 16} \\ &\leq 2^3 3^{12} (\sigma \tilde{\nabla}_{(10r+4)/2}(G) + 2\rho)^9 && \text{by Proposition 7,} \end{aligned}$$

where $\sigma := \max\{20r\rho + 8\rho + 1, (2\rho + 1)^2\}$. By the same argument we obtain the bound

$$\begin{aligned} \tilde{\nabla}_0(H) &\leq \nabla_2(G \bullet \bar{K}_{2\rho+1}) \\ &\leq 4(4\tilde{\nabla}_2(G \bullet \bar{K}_{2\rho+1}))^9 && \text{by Proposition 6} \\ &\leq 4(4(2\rho + 1)^2 \tilde{\nabla}_2(G) + 2\rho)^9 && \text{by Proposition 7,} \end{aligned}$$

as claimed¹. □

Having established that H is small and sparse we can now show that there exists a class in K_S whose members do not all need to be in Z . First, let us show the existence of a class with suitable properties.

Lemma 77 (Large class κ^*). *There exist a class $\kappa^* \in K_S$ of size at least*

$$|\kappa^*| \geq 2\tau(\deg_H(\kappa^*) + 1) + 1.$$

Proof. Let τ, ω be defined as before. We define the potential function $\phi: K_S \rightarrow \mathbb{R}$ as

$$\phi(\kappa) = |\kappa| - (2\tau(\deg_H(\kappa) + 1) + 1).$$

Consider the sum

$$\begin{aligned} \sum_{\kappa \in K_S} \phi(\kappa) &= \sum_{\kappa \in K_S} |\kappa| - 2\tau \sum_{\kappa \in K_S} \deg_H(\kappa) - (2\tau + 1) \cdot |K_S| \\ &\geq |S| - 4\tau \|H\| - 3\tau |K_S|. \end{aligned}$$

By Lemma 76, we have that

$$\begin{aligned} \|H\| &\leq \tilde{\nabla}_0(H) \cdot |H| \leq (4(2\rho + 1)^2 \tilde{\nabla}_2(G) + 2\rho)^9 |H| \\ &\leq (38\rho^2 \tilde{\nabla}_2(G))^9 |H|, \end{aligned}$$

where ρ is defined as in the proof of Lemma 76. Using the bounds from Lemma 75 on K_S, K_R we see that $|H| \leq 2|K_S|$. Adding the fact that $|S| > c_{SX}|X|$ we obtain for the right hand side

$$\begin{aligned} &|S| - 4\tau(38\rho^2 \tilde{\nabla}_2(G))^9 |H| - 3\tau |K_S| \\ &\geq |S| - 14\tau(38\rho^2 \tilde{\nabla}_2(G))^9 |K_S| \\ &> (c_{SX} - 2^{12\omega_2(G)+55} \rho^{18} \cdot \omega_2(G)^2 \cdot \tilde{\nabla}_2(G)^{18\omega_2(G)+11}) \cdot |X| = 0. \end{aligned}$$

We infer that $\sum_{\kappa \in K_S} \phi(\kappa) > 0$, accordingly there exists at least one class $\kappa^* \in K_S$ with

$$\phi(\kappa^*) > 0 \implies |\kappa^*| > (2\tau(\deg_H(\kappa^*) + 1) + 1). \quad \square$$

We finally arrive at a situation in which we can safely decrease the size of Z by removing some vertices of $\kappa^* \subseteq S$ from it. The important intuition is that the vertices of κ^* are necessarily dominated from vertices in X . Similar to the famous sunflower-reduction, we can prune $\kappa^* \cap Z$ without losing this property.

Lemma 78 (Shrinking κ^*). *Let $\hat{\kappa}^* \subset \kappa^*$ be an arbitrary subset with*

$$|\hat{\kappa}^*| = 2\tilde{\nabla}_{1/2}(G)(\deg_H(\kappa^*) + 1) + 1.$$

Then the set $Z' := (Z \setminus \kappa^) \cup \hat{\kappa}^*$ is a domination core of G .*

¹ Assuming $\rho \geq 1$, otherwise the term $(2\rho + 1)^2$ needs to be replaced by $8\rho + 1$.

Proof. Assume D is a minimal Z' -dominator in G , we need to show that it also is a Z -dominator. Recall that across all vertices $v \in \kappa^*$, we have that $N'(v)$ and $N(v) \cap X$ are the same two subsets of X . We denote these sets by $N'(\kappa^*)$ and $N(\kappa^*) \cap X$ in the following.

First, assume that $D \cap N'(\kappa^*)$ is non-empty. Then D dominates all of κ^* and is therefore a Z -dominator. We thus consider the case that $D \cap N'(\kappa^*) = \emptyset$ and will show a contradiction with respect to the minimality of D . To that end, let us construct a different set D' from D as follows:

1. Remove all vertices contained in $\Gamma(\kappa^*)$, then
2. for $\kappa_S \in N_H[\kappa^*] \cap K_S$, add the vertices $N(\kappa_S) \cap X$ and
3. for $\kappa_R \in N_H[\kappa^*] \cap K_R$, add an arbitrary vertex from $N(\kappa_R) \cap X$ (if such a vertex exists).

Claim. $|D'| < |D|$.

Since $D \cap N'(\kappa^*)$ is empty every vertex of $\hat{\kappa}^*$ must be dominated by a neighbour in $G \setminus X$. Since the vertices $\hat{\kappa}^*$ are 3-scattered in $G \setminus X$, we have that $|D \cap \Gamma(\hat{\kappa}^*)| \geq |\hat{\kappa}^*|$. Hence the first step in the construction of D' removes at least $|\kappa^*|$ vertices. The second and third step now add at most

$$2\tilde{\nabla}_{1/2}(G)(\deg_H(\kappa^*) + 1) < |\hat{\kappa}^*|$$

vertices, hence D' contains less vertices than D .

Claim. D' is a Z -dominator.

We show that every vertex $x \in Z$ is dominated by distinguishing the following cases. Note that since X is a Z -dominator, every vertex of Z is either contained in X or has at least one neighbour in it.

Case 1: $x \in \Gamma(\kappa_S)$ for $\kappa_S \in K_S$. If $\kappa_S \notin N_H[\kappa^*]$ then x is dominated by the same vertices that dominated it from D . Otherwise, the second step in the construction of D' in particular added the vertices $N(\kappa_S) \cap X$ to D' . Since every vertex of $\Gamma(\kappa_S) \cap Z$ has at least one neighbour in X and hence in $N(\kappa_S) \cap X$, we conclude that x is dominated by D' .

Case 2: $x \in \Gamma(\kappa_R)$ for $\kappa_R \in K_R$. If $\kappa_R \notin N_H[\kappa^*]$ then x is dominated by the same vertices that dominated it from D . Otherwise, the third step in the construction of D' added a vertex from $N(\kappa_R) \cap X$. Since per definition all vertices of κ_R have the exact same neighbourhood in X , we conclude that x is dominated by D' .

Case 3: $x \in X$. The only vertices in X affected by the construction of D' are those in $N(\kappa^*) \cap X$. Since this set is added to D' in the second step, they are clearly all dominated by D' .

This concludes the proof of the claim and leaves us with a contradiction: D' is smaller than D and a Z -dominator, contradicting the minimality of D . Therefore Z' is still a domination core of G , as claimed. \square

We can finally prove the main technical contribution of this section.

Proof of Lemma 72. As long as $|Z|$ has size at least $c_Z \cdot k$, we are guaranteed by Lemma 73 that the extraction-algorithm from the first phase results in a decomposition (X, S) with $|S| \geq c_{SX}|X|$ (glossing over the minor step applied through Lemma 74). The time needed for this phase is $O(c_1 \xi \cdot \|G\|)$, where $\xi = \min\{2\nabla_0(G), \omega\}$ and c_1 is the approximation ratio of the algorithm presented in Lemma 50.

Using the class graph H derived from (X, S) we are guaranteed the existence of a large equivalence class κ^* in S via Lemma 77 as long as $|S| > c_{SX}|X|$, from which as per Lemma 78 we can safely remove vertices. If the resulting set S' now is smaller or equal to $c_{SX}|X|$, we are done. Otherwise we want to identify another large class: to that end, we simply adapt the class graph H by removing those edges that are now lost due to the removal of vertices from κ^* . Note that the resulting class graph H' is a subgraph of H , the upper bounds on $|H'|$ and $\|H'\|$ therefore still apply. We conclude that we can iterate the procedure until $|S| \leq c_{SX}|X|$. The time taken here is at most

$$O(\|G\| \cdot |H|) = O(\|G\| \cdot c_{SX}|X|) = O(\|G\| \cdot c_{SX}c_Xk)$$

since we might need to shrink every single class until the resulting set S is small enough.

After the above procedure is done, we are left with a set \hat{S} such that $\hat{Z} := \hat{S} \cup (X \cap Z)$ is a domination core of G . The size of this set is bounded by

$$\begin{aligned} |\hat{Z}| &\leq |X| + |\hat{S}| \leq (c_{SX} + 1)|X| \\ &\leq (c_{SX} + 1) \cdot c_Xk. \end{aligned}$$

as claimed. This concludes the proof. \square

Having shown the existence of a small domination core, we now proceed to construct the kernel by reducing the remaining vertices around the core.

13.2 COMPUTING A KERNEL

We did all the heavy lifting in the previous section: given a graph G and a parameter k , we can compute a domination core Z of G of size at most $c_Z \cdot k$ where c_Z essentially depends on $\tilde{\nabla}_2(G)$ and $\omega_2(G)$. We will use the quantities τ, ω, ρ defined in Section 13.1.

Lemma 79 (Reducing dominators). *For a graph G and integer k with domination core Z we can, in linear time, compute an induced subgraph G' of G such that $\mathbf{ds}(G') = \mathbf{ds}(G)$ and further*

$$|G'| \leq (2\omega(e\tau)^\omega + 1)|Z|.$$

Proof. First, observe that any vertex that has distance two or more to Z can be safely removed: any optimal Z -dominator will not make use of such a vertex². Therefore assume that $N[Z] = V(G)$. Let $Y = V(G) \setminus Z$.

We construct the auxiliary bipartite graph $H = (Z, Y, E_G(Z, Y))$ and partition Y into twin-classes $\gamma_1, \dots, \gamma_p$. Using Proposition 1, this is possible in time $O(|H| + \|H\|)$. Note that by Lemma 36, we have that

$$p \leq (2\omega(e\tau)^\omega) \cdot |Z|,$$

where we used the fact that $\tau \geq \tilde{\nabla}_{1/2}(H)$ and $\omega \geq \tilde{\omega}_{1/2}(H)$.

We construct G' from G by retaining one representative vertex v_i per class γ_i , i.e. we let $G' = G[Z \cup \{v_1, \dots, v_p\}]$. By the above bound on p we immediately see that $|G'|$ is indeed bounded from above as claimed. It is left to show that $\mathbf{ds}(G') = \mathbf{ds}(G)$.

Claim. *It is true that $\mathbf{ds}(G') \leq \mathbf{ds}(G)$.*

Assume $D \subseteq V(G)$ is a minimal dominating set in G , i.e. $|D| = \mathbf{ds}(G)$. Then D is also a Z -dominator in G and we can easily construct a set D' as follows: we start out with $D' = D \cap Z$ and then add for every twin-class γ_i with $|\gamma_i \cap D| \geq 1$ its representative v_i . Clearly $|D'| \leq |D|$. Further, note that D' is a Z -dominator in G : the vertices of Z that were formerly dominated by some class γ_i with non-empty intersection with D are now dominated by the vertex $v_i \in D'$. As such, D' is actually a dominating set in G . Now by construction $D' \subseteq V(G')$, and therefore D' is also a dominating set in the induced subgraph G' . We conclude that

$$\mathbf{ds}(G') \leq |D'| \leq |D| = \mathbf{ds}(G).$$

Claim. *It is true that $\mathbf{ds}(G) \leq \mathbf{ds}(G')$.*

Assume $D' \subseteq V(G')$ is a minimal dominating set in G' . Therefore, it dominates Z in G' and accordingly—by construction—in G . We conclude that

$$\mathbf{ds}(G) = \mathbf{ds}(G, Z) \leq |D'| = \mathbf{ds}(G').$$

This concludes the proof of the lemma. \square

We can finally state and prove the main theorems of this chapter.

Theorem 34. *Let \mathcal{G} be a hereditary graph class of bounded expansion. Then DOMINATING SET parametrised by the solution size admits a linear kernel in \mathcal{G} and the kernel can be computed in time $O(\|G\|k)$.*

Proof. We assume that \mathcal{G} is infinite since otherwise the result is trivial. Given an instance (G, k) of DOMINATING SET with $G \in \mathcal{G}$, we invoke Theorem 33. As a result, we either conclude that $\mathbf{ds}(G) > k$ or we obtain a domination core $Z \subseteq V(G)$ of size at most $c_Z \cdot k$. In the previous

² In our construction of Z such vertices will not exist, but for the sake of generality we address this possibility here.

case we output a trivial NO-instance (since \mathcal{G} is hereditary and infinite we have $\bar{K}_2 \in \mathcal{G}$ and hence can output $(\bar{K}_2, 1)$). In the latter case, we invoke Lemma 79 and obtain a subgraph $G' \in \mathcal{G}$ with $\mathbf{ds}(G') = \mathbf{ds}(G)$ of size at most

$$\begin{aligned} |G'| &\leq (2\omega(e\tau)^\omega) \cdot |Z| \\ &\leq (2\omega(e\tau)^\omega) \cdot c_Z k. \end{aligned}$$

Since \mathcal{G} has bounded expansion, we can treat τ, ω, ρ and by extension c_Z as constants and therefore conclude that

$$O(|G'| + \|G'\|) = O(|G'|) = O(k).$$

Hence, (G', k) is a linear kernel of (G, k) . \square

Theorem 35. *Let \mathcal{G} be a nowhere-dense and hereditary graph class. Then DOMINATING SET parametrised by the solution size admits a kernel of size $O(k^{1+\varepsilon})$ for every $\varepsilon > 0$ on \mathcal{G} . Moreover, such a kernel can be computed in time $O(\|G\| \log \log |G| \cdot k)$.*

Proof. We proceed exactly as in the proof of Theorem 34, up to the point where we construct the subgraph G' . Again, the size of G' is bounded by

$$|G'| \leq (2\omega(e\tau)^\omega) \cdot c_Z \cdot k,$$

however, we cannot simply regard the quantity τ and thus c_Z (which also depends on $\tilde{\nabla}_2(G)$) as a constant here. By (cf. Theorem 20), we can, however, choose any $\varepsilon > 0$ and assume that

$$\tau = \tilde{\nabla}_{1/2}(G) \leq \tilde{\nabla}_2(G) \leq |G|^\varepsilon$$

provided that $|G|$ is larger than some threshold N_ε depending on \mathcal{G} and ε . The quantities ω and $\omega_2(G)$, however, are indeed constants only dependent on \mathcal{G} . Recall that c_Z was defined as

$$\begin{aligned} c_Z &= (\omega^2 + 1)(2\omega(e\tau)^\omega) \\ &\quad \cdot (1 + c_1 \cdot 2^{12\omega_2(G)+55} \rho^{18} \cdot \omega_2(G)^2 \cdot \tilde{\nabla}_2(G)^{18\omega_2(G)+11})^{\tilde{\zeta}} \cdot c_1. \end{aligned}$$

where $\tilde{\zeta} \leq \omega$. Simplifying the bounds, we see that

$$|G'| = \tilde{\nabla}_2(G)^{\omega_2(G)^{O(1)}} \cdot k \leq |G|^{\varepsilon'} \cdot k$$

where $\varepsilon' = \omega_2(G)^{O(1)}\varepsilon$ can be chosen as small as desired. The final trick—which I have to credit Michał Pilipczuk for who developed it for our paper [76]—is to apply the data reduction exhaustively: assume we construct a sequence of graphs $G = G_0, G_1, \dots$ where G_i is obtained from G_{i-1} by applying the data reduction. At some point we reach a stage where $G_i = G_{i-1}$; now observe that

$$|G_i| = O(|G_{i-1}|^{\varepsilon'} k) = O(|G_i|^{\varepsilon'} k) \iff |G_i|^{1-\varepsilon'} = O(k).$$

Recalling the series expansion for $1/(1-x)$, we see that

$$k^{\frac{1}{1-\varepsilon'}} \leq k^{1+2\varepsilon'}$$

for $0 < \varepsilon' < 1$. Thus for any ε'' we choose $\varepsilon = \varepsilon''/2(19\omega^2 + 36\omega + 3)$ and obtain a kernel of size $O(k^{1+\varepsilon''})$. The last question is how often we have to apply the data reduction. Observe that the sizes of the G_i obey

$$|G_i| = O(|G|^{\varepsilon'^i} \cdot k^{1+\varepsilon'+\varepsilon'^2+\dots+\varepsilon'^i}) = O(|G|^{\varepsilon'^i} \cdot k^{1+2\varepsilon'}).$$

Thus after $(\log \log |G| / -\log \varepsilon')$ iterations, the exponent ε'^i becomes

$$\varepsilon'^{\frac{1}{-\log \varepsilon'} \log \log |G|} = \left(\frac{1}{e}\right)^{\log \log |G|} = \frac{1}{\log |G|}$$

which in turns means that the term $|G|^{\varepsilon'^i}$ turns into a constant. We conclude that the total time needed to compute a kernel is at most $O(|G| \log \log |G| \cdot k)$. \square

We conclude that DOMINATING SET can be efficiently preprocessed in structurally sparse graphs. Recently, Kreutzer, Michał Pilipczuk, and Siebertz have extended this result and proved a linear kernel for r -DOMINATING SET in bounded expansion classes³. I further suspect that all ‘local’ problems should admit polynomial kernels. And this locality might even be necessary:

Theorem 36 (Drange *et al.* [76]). *There exists a hereditary class \mathcal{G} of bounded expansion such that CONNECTED DOMINATING SET does not admit a polynomial kernel when restricted to \mathcal{G} , unless $\text{NP} \subseteq \text{coNP}/\text{poly}$.*

However, locality as captured by first-order properties is still too strong. Note that k -PATH is (non-uniformly) first-order expressible, but unless the polynomial hierarchy collapses to the third level, it does not admit a polynomial kernel even in sparse classes [22].

Open question 11. Do other domination problems like DOUBLE DOMINATION, or EFFICIENT DOMINATING SET admit polynomial kernels in structurally sparse classes?

Open question 12. Is there a characterisation of problems that admit (almost) linear kernels in structurally sparse classes?

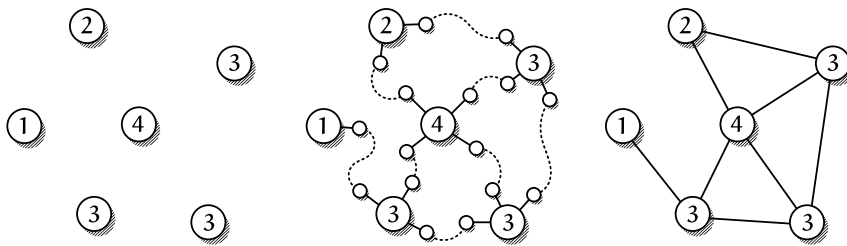
³ Personal communication.

Part IV

COMPLEX NETWORKS AND
STRUCTURAL SPARSENESS

Interstitiality is a theme that is simultaneously genuinely interesting and potentially quite useful, and also a terrible cliché, so if you're going to use it, it helps to be at least respectfully skeptical about the wilder claims of its theoretical partisans, I think.

— China Miéville, *The City & the City*



*What they don't tell you is that it's impossible to move, to live,
to operate at any level without leaving traces, bits,
seemingly meaningless fragments of personal information.*
— William Gibson, *Johnny Mnemonic*

Graph theory is traced back to Leonard Euler's musing about the Königsberger bridges in 1736: in modern parlance, he proved that the multigraph obtained from the topology of the then-German city does not contain an Eulerian cycle¹. His deep contribution was, obviously, not the solution to the puzzle but instead his mathematical abstraction. And while graphs did stay a primarily theoretical topic until just recently, they did crop up in practical application.

Operations research

The most notable examples of 'real-world' graphs before the information age can be found in two fields: sociology and operations research. Operations research originated as a field in WWII in an effort to optimize British war logistics. Graphs primarily appear as road or telecommunication networks, as exemplified by the second version of the Soviet rail network in Figure 2. Despite (or because?) its high-stakes and fundamentally applied nature, operations research gave rise to indispensable tools like the min-cut max-flow theorem and the simplex algorithm. Interestingly, already back then the restriction to certain structurally sparse cases was seen as a viable way of improving algorithmic tractability: Ford and Fulkerson in their seminal work on flows in networks [105] note that the operations-research inspired application of network flows is best seen as a problem in a special type of planar network, for which they describe a method of finding augmenting paths in the residual network. The idea that structural sparseness improves tractability is certainly all but new.

Sociology

In a very different (and more peaceful) line of research, anthropologists and sociologists began as early as the 1930s to systemically map out the relation between individuals in what we today call *social networks*. The following examples are taken from the comprehensive and extremely interesting book by Freeman about the historical development of social network analysis [111]. One of the earliest available example of a rigorously mapped out social network is the *Southern Women Data Set* collated by Davis, Gardner, Gardner *et al.* in the thirties. Their study was conducted in the context of Warner's "Deep South" project at Harvard in 1933. The goal of the study was to assess

¹ Besides its nationality and name, the city also changed its topology due to its destruction in WWII and subsequent rebuilding efforts. It now admits a Eulerian cycle (pointed out by J. Kåhre on his homepage <http://www.mattheory.info/>).

how class, social status, and race² shaped social interaction. They observed the social activity of 18 women from Natchez, Mississippi, over a period of nine months by recording their attendance to 14 social events. The resulting bipartite graph as well as the original tabulated data are displayed in Figure 3. While the scale of the data seems meagre compared to what is available today, it has spurred very creative approaches to extract information about social strata from it (see, e.g., Freeman’s survey on the southern women network [110]).

While the Southern Women Study was one of the earliest in which sociologist embraced a *structural approach*, that is, they focused on the interaction instead of the individual, the concept of social networks is much older. Freeman attests the structural perspective to Auguste Comte himself, the father of sociology, but further acknowledges that it seems to be human nature to track relationships between individuals [111]. It seems that to a certain degree, we were always curious about the ties between us.

Complex networks

To summarise: the roots of network science (or what one might call *applied graph theory*) clearly trace back to the early twentieth century. What distinguishes this early work from what we now understand as network science is its narrow scope: the crucial insight that spawned the explosive growth of the field was that most graphs lifted from the real world are *inherently similar*. Whether we look at research collaboration in high-energy physics, user-supplied movie recommendations on Amazon, or the neural connectome of *Drosophila melanogaster*; there are indisputable structural similarities to be found. These observations crucially relied on better ways to collect, store and organise large sets of data which explains why these insights occurred only recently: network science could not have existed without the digital revolution.

While there is certainly some hype attached to this still young field, the urgent need summarised in the adage “drowning in data, yet starving for information” must ultimately be addressed. Today networks are growing around us with every action undertaken. Our communication, friendship, business relations, sexual encounters, physical locations, and financial transactions are mapped out by various government agencies, banks, and private industries—all in the form of inter-relating data. The mixture of disciplines found in network science certainly has the potential to provide the essential tools to make sense of the ever-growing heap of data our lives generate.

Structural sparseness

In this part of the thesis, we want to demonstrate that the theory of structurally sparse graphs has a great algorithmic potential for real-world networks. This claim needs to be dissected in two parts: first, we need to show that complex networks *are* structurally sparse (and define what we mean by that!). Second, we need to demonstrate that this fact can be exploited to design algorithms for relevant problems. Demonstrating the second part of the claim is rather difficult: a concrete implementation that solves a domain-specific problem us-

² The city of Natchez, Mississippi, in which the study was conducted was segregated.

NAMES OF PARTICIPANTS OF GROUP I	CODE NUMBERS AND DATES OF SOCIAL EVENTS REPORTED IN <i>Old City Herald</i>													
	(1) 6/27	(2) 3/2	(3) 4/12	(4) 9/26	(5) 2/25	(6) 5/19	(7) 3/15	(8) 9/16	(9) 4/8	(10) 6/10	(11) 3/23	(12) 4/7	(13) 11/21	(14) 8/3
1. Mrs. Evelyn Jefferson.....	X	X	X	X	X	X		X	X					
2. Miss Laura Mandeville.....	X	X	X		X	X	X	X						
3. Miss Theresa Anderson.....		X	X	X	X	X	X	X	X					
4. Miss Brenda Rogers.....	X		X	X	X	X	X	X						
5. Miss Charlotte McDowd.....			X	X	X		X							
6. Miss Frances Anderson.....			X		X	X	X	X						
7. Miss Eleanor Nye.....					X	X	X	X						
8. Miss Pearl Oglethorpe.....						X	X	X	X					
9. Miss Ruth DeSand.....					X		X	X	X					
10. Miss Verne Sanderson.....							X	X	X			X		
11. Miss Myra Liddell.....								X	X	X			X	
12. Miss Katherine Rogers.....								X	X	X		X	X	X
13. Mrs. Sylvia Avondale.....							X	X	X	X		X	X	X
14. Mrs. Nora Fayette.....						X		X	X	X	X	X	X	X
15. Mrs. Helen Lloyd.....							X	X	X	X	X	X		
16. Mrs. Dorothy Murchison.....								X	X					
17. Mrs. Olivia Carleton.....									X		X			
18. Mrs. Flora Price.....									X		X			

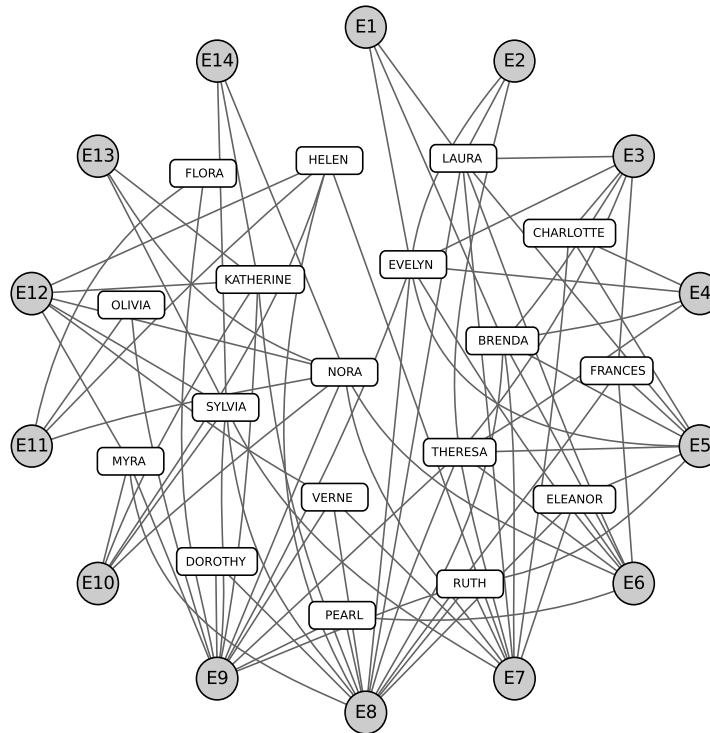


Figure 3: The Southern Women data set, both in its original form as published by Davis *et al.* [56] and as a modern representation as a labelled bipartite graph. There exist two conflicting data sets, the one reproduced here seems to be favoured as the probably correct one (*cf.* Freeman’s survey [110]).

ing techniques presented here would be the ultimate confirmation, but such an endeavour constitutes its own research programme. Consequently, we have to settle for less here: we will demonstrate that several problems that have been deemed interesting in the network science literature can be solved efficiently using tools presented earlier. Since we will demonstrate this for problems of very different flavours they should serve as a sufficient proxy to support our claim.

Efficiency

Finally, we should address what should be considered an ‘efficient algorithm’ here. Since we are yet at a stage too far removed from actual applications, the concept of efficiency is necessarily fuzzy: we can only define its rough outlines by enumerating constraints and try to design algorithms as defensive as possible. First and foremost, since network science inherently deals with huge amounts of data, any algorithm with running time worse than $n \log^{O(1)} n$ will be useless for a large portion of interesting data sets. If possible, we prefer algorithms that are parallelisable, run out-of-core³ or process the data in a stream. If an algorithm should be applicable in practice, it should not contain elements that make these approaches impossible. Luckily, the toolkit provided by structural sparseness are all linear-time algorithms and seem to lend themselves to parallelisation. Consider, for example, algorithms based on low-treedepth colourings: the colouring-approach naturally divides the instance into much easier substances whose individual solutions can be combined easily (usually by some inclusion-exclusion type calculation).

In the following we will first give an basic overview over network statistics in Section 14.1 and algorithmic questions related to networks in Section 14.2. Afterwards we will investigate several network models in Chapter 15, where in Section 15.1 we first define the necessary tools to talk about structural sparseness of random graph models.

14.1 BASIC NETWORK STATISTICS

The study of networks ultimately relies on observable quantities or *statistics* which give us insight into the network’s function. Historically these statistics are very much derived from the physical sciences, like percolation theory and the theory of complex systems. As a consequence, the value of these statistics usually has little impact on algorithmic properties of networks. However, statistics are also a valuable tool to guide our intuition of networks; given that there is little chance of visualising networks above a certain size⁴. In the following we present some of the most commonly cited statistics. For a much more complete overview we refer to the survey by Brinkmeier and Schank [35] and the survey by Newman [195].

Density

³ Meaning the data to be processed is not stored in main memory.

⁴ Such attempts usually end up in a representations called a ‘hair ball’ of little scientific value.

DENSITY The probably first statistic one naturally considers is the density. As mentioned several times before, we find that networks are empirically sparse. The question whether they are *structurally* sparse is the main topic of Chapter 15.1 and we will see some good reasons to believe that they are which will be further substantiated by experiments conducted in Chapter 16.

Densification?

Before we begin to make our case, we need to address a claim regarding network density made by Leskovec, Kleinberg and Faloutsos [171]. They propose—contrary to wide-spread opinion—that complex networks densify and exhibit a shrinking diameter during their growth. Specifically, they claim that empirically the number of edges grows as $\Theta(n^{1+\alpha})$ for some fixed $\alpha > 0$ depending on the network.

Their claim is only problematic in so far as they insist that the densification happens infinitely long. This immediately clashes with almost every imaginable domain: consider the case of a citation network (one of the three types of networks they analyse). If we direct edges from a citing paper towards the cited one, it is clear that our network necessarily has a bounded out-degree—the number of papers cited in a publication might vary a lot, but there is a technical upper limit to it. The same holds true for collaboration between scientists or movie actors: as every collaboration holds a finite number of authors⁵ the total number of edges in the graph is bounded linearly in the number of collaborations. Indeed, the heuristic the authors used to establish the claim has come under criticism [49]:

A common way to probe for power-law behaviour, therefore, is to measure the quantity of interest x , construct a histogram representing its frequency distribution, and plot that histogram on doubly logarithmic axes. If in so doing one discovers a distribution that approximately falls on a straight line, then one can, if one is feeling particularly bold, assert that the distribution follows a power law, with a scaling parameter α given by the absolute slope of the straight line. Typically this slope is extracted by performing a least-squares linear regression on the logarithm of the histogram. This procedure dates back to Pareto's work on the distribution of wealth at the close of the 19th century [...]. Unfortunately, this method and other variations on the same theme generate significant systematic errors under relatively common conditions [...] and as a consequence the results they give cannot not be trusted.

One further criticism can be voiced by recalling the theorem by Dvořák and Jian (Theorem 10): if these networks indeed have $\Theta(n^{1+\alpha})$ edges, then we have to conclude that they contain arbitrarily large shallow clique-minors. This seems to contradict any notion of efficiency in such networks: using a large number of disjoint paths to pairwise connect

⁵ Disregarding papers published by CERN.

the same set of endpoints, instead of optimising such connections by introducing a few intermediary nodes.

A simple reconciliation of the densification claim is that networks have a *growth phase* in which the number of edges grows superlinear after which they enter a *saturated phase* which anneals the number of edges back to a linear function. This hypothesis is in accordance with the observation presented in [171]: the presented networks are—with two exceptions—clearly young and presumably in their growth phase. For the two exceptions, a movie collaboration networks spanning the last century and a network derived from patents, one can only speculated at this point. Since the authors fitted linear functions to log-log-plots of the network densities over times, there is a clear possibility that a late recede of the growth is hidden by the data transformation.

Without going into this matter further (which warrants and deserves further research!), we postulate that the claimed densification during network evolution seems to be an early-stage phenomenon and as such does not contradict our claim of structural sparseness.

DEGREES The second-most accessible statistic is the distribution of degrees in a network. On a coarse level, we see certain similarity across networks: first, the maximal possible degree is usually much lower than $n - 1$. Second, if one creates a histogram for the degrees occurring in the network, one sees that this *degree distribution* tends to be heavily right-skewed. This has given rise to the claim that real-world networks are ‘heavy-tailed’ and ‘scale-free’. More specific, a lot of distributions where said to follow a power law: the fraction of vertices of degree d in a network would be proportional to $d^{-\alpha}$, for α usually between two and three.

Degree distribution

These early claims (following the work by Barabási and Albert [18]) need some revision. The heuristics used to detect power laws by log-log-plots have again generated a lot of false positives (see Chapter 16 for an analysis of the degree distribution of several networks). A more careful analysis indicates that degree distributions are more diverse than claimed and very few of them follow a pure power law [49]. And with good reason: in most domains, network edges attach a real-world cost to their endpoints, a cost that is deducted from finite resources. Friendships or business relationships in social networks cost time and effort; connections in autonomous system networks increase the routing table size; and synaptic connections between neurons increase energy consumption [137]. In such situations there needs to be a sharp limit on the maximum possible degree which is not reflected in a pure power law distribution. See also the work by Tanaka, Yi, and Doyle that casts doubt on whether protein interaction data exhibits power laws [234], and the work by Pržulj, Corneil, and Jurisica that establishes *geometric random graphs* as a better model for protein interaction networks [209].

In contrast, directed networks in which the cost of a connection is primarily paid by one of the two endpoints—like followers in the twitter networks or send emails in a company—have potentially vertices with large indegrees and hence their degree distribution might very well follow a power law.

r-Neighbourhood

A ‘deeper’ variation of the degree of a vertex is the size of $N^r(\bullet)$, *i.e.* the number of vertices that have distance at most r from it. The *hop plot* P is defined via the r -neighbourhoods as

$$P(r) = \sum_{v \in G} |N^r(v)|$$

and is applied to reason about information transmission in a network.

Small world property

DIAMETER So far, we have seen only statistics related to rather local properties of a network which aggregate information of single vertices or edges. But complex networks also exhibit highly interesting behaviour when it comes to questions of connectivity. In particular, most vertices in a complex network lie rather close together: this is what we call the *small world property*. Hungarian writer Frigyes Karinthy was maybe the first to formulate it on paper in his 1929 story ‘Chain-Links’ [150]:

One of us suggested performing the following experiment to prove that the population of the Earth is closer together now than they have ever been before. We should select any person from the 1.5 billion inhabitants of the Earth - anyone, anywhere at all. He bet us that, using no more than five individuals, one of whom is a personal acquaintance, he could contact the selected individual using nothing except the network of personal acquaintances.

The first rigorous study to quantify this phenomenon—subsequently popularised as ‘six degrees of separation’—was famously undertaken by Milgram 1967 [184] and later analysed by Travers and Milgram [239]. The experiment itself used chain-letters to estimate the average path-length between random individuals and a fixed target in the US.

There are different variations of what constitutes a small world. We stick to the most basic definition here and say that a network has the small world property if the diameter of its connected components is of the order $\log n$ or smaller.

Clustering

CLUSTERING Watts and Strogatz introduced the *clustering coefficient* as a measure of how likely connections between nodes are given the knowledge that these nodes share a common neighbour [247]. Let $\tau(v)$ denote the number of neighbours of v that are not connected, *i.e.*

$$\tau(v) = \binom{\text{deg}(v)}{2},$$

and let us extend this statistic to graphs via

$$\tau(G) = \sum_{v \in G} \tau(v).$$

Similarly, let $\lambda(v)$ denote the number of triangles incident to v and $\lambda(G)$ the total number of triangles in G .

The clustering coefficient C is now defined as

$$C(G) = \frac{1}{|G|} \sum_{v \in G} \frac{\lambda(v)}{\tau(v)}.$$

A similar measure is the *transitivity* defined by Newman, Strogatz, and Watts [198] as

$$T(G) = \frac{3\lambda(G)}{\tau(G)}.$$

There are several subtleties in the above definitions, for example, how to treat vertices of degree less than two. For the context of this thesis the high-level idea of clustering is sufficient and we do not have to go into the differences between the different variations measuring it.

Notably, most real-world networks exhibit a much higher clustering than random network models. In particular the *local clustering coefficient* of a vertex $\lambda(v)/\tau(v)$ has been studied extensively in social sciences (see Newman's survey for further references).

Centrality indices

CENTRALITY One of the, let us say, *central* questions in the analysis of networks is the relative importance of its members. While in a specific domain this notion of importance is usually well-defined and measurable, the question arises whether it can be derived using only the network structure. This leads to the notion of *centrality measures* c which assign either a numerical value to the vertices (or edges) of a network, with the usual understanding that high values signify higher importance. The concept itself is often traced back to a paper by Jordan from 1869 [149].

Degree centrality

The most basic index c_D is the degree itself: a node is important if it has a large degree. While readily available, it uses only very local information and we cannot expect to conduct a fine-grained analysis by relying on it.

*Closeness centrality, harmonic –,
Lin's –*

One type of centrality index measures how far a node lies in the 'middle' of a network. Sabidussi introduced the *closeness centrality* [220]

$$c_C(v) = \left(\sum_{u \in G} d(v, u) \right)^{-1}.$$

Related indices are the *harmonic centrality* introduced by Opsahl, Agneessens, and Skvoretz [199]

$$c_H(v) = \sum_{u \in G} d(v, u)^{-1},$$

and *Lin's centrality* [173]

$$c_L(v) = \frac{|\{v \mid d(v, v) < \infty\}|^2}{\sum_{u \in V(G): d(v, u) < \infty} d(v, u)}.$$

Stress centrality, betweenness – We will revisit these three distance-based indices in the next section.

A different kind of importance arises if one views a network in terms of communication, as Shimbel did by introducing the *stress centrality* [227]. Let $\sigma_{st}(v)$ denotes the number of shortest s - t -paths that contain the vertex v and let σ_{st} denote the total number of shortest s - t -paths. Then the centrality is defined by

$$c_S(v) = \sum_{v \neq s, t \in G} \sigma_{st}(v).$$

Assuming that communication happens along shortest paths in a network and between uniformly random endpoints, the stress centrality reveals how much communication load a node has to bear. A related concept is the *betweenness centrality* introduced by Freeman [109] and Anthonisse [10]:

$$c_B(v) = \sum_{v \neq s, t \in G} \frac{\sigma_{st}(v)}{\sigma_{st}}.$$

Eccentricity

The *eccentricity* of a vertex is defined as $e(v) = \max\{d(u, v)\}_{u \in G}$, i.e. the length of the longest shortest path starting at v . It gives rise to the centrality index $c_E(v) = e(v)^{-1}$ introduced by Hage and Harary [135] who applied it to a facility-location-type problem and argued that existing centrality measures (in particular closeness and betweenness) were not suitable.

There are a host of further measures capturing other notions of importance. We refer to the survey by Koschützki *et al.* which also contains interesting historical notes [162] and the book by Wasserman and Faust [246].

In conclusion, there are already a lot of algorithmic questions attached to the basic properties of complex networks. Further, the efficient computation or approximation of important statistics directly impacts how well we can research them.

14.2 ALGORITHMIC QUESTIONS

Since this thesis sets out to bring the field of structural graph theory and network science closer together, we should demonstrate this endeavour's potential pay-off. To this end, we survey several algorithmic questions related to real-world networks that seem to be solvable—at least in theory—with the algorithmic toolkit presented earlier.

The first tier of algorithmic question lies simply with network statistics, some of which we presented in the last section. Obviously, we want to be able to compute these basic measures efficiently and reliably in

order to reason about large-scale datasets. However, some of these properties take quadratic or even cubic time to compute—at least in a naive implementation. Take, for example, the clustering coefficient and the simple cubic-time algorithm to compute it. If we postulate that our network is degenerate, suddenly the computation is possible in linear time! This observation can be extended to arbitrary complete subgraphs; as Eppstein, Löffler, and Strash demonstrated [85]. Their algorithm is superior to the Bron-Kerbosh algorithm in sparse graphs and consequently should replace it in virtually all practical applications—that this is *not* the case is a testimony to the fact that structural graph theory needs to be promoted more vigorously.

Localised closeness

Let us consider another statistic in which our algorithmic toolkit could improve existing algorithms—assuming that the provided inputs are structurally sparse. Closeness centrality and its variants require knowledge of the distances between all vertex-pairs, a problem for which the fastest known algorithm runs in time $O(n(n + m))$ [34]. Even under the sparseness assumption, the known approach therefore still takes quadratic time. However, as we saw in Chapter 10.4, it is possible to compute the sizes of all distance r -neighbourhoods around *all* vertices in linear time! Hence the following ‘localised’ variants of distance-based centrality measures can, on structurally sparse inputs, be computed quickly as stated in Theorem 26.

Measure	Definition	Localised
Closeness	$c_C(v) = \left(\sum_{u \in G} d(v, u) \right)^{-1}$	$c_C^r(v) = \left(\sum_{u \in N^r(v)} d(v, u) \right)^{-1}$
Harmonic	$c_H(v) = \sum_{u \in G} d(v, u)^{-1}$	$c_H^r(v) = \sum_{u \in N^r(v)} d(v, u)^{-1}$
Lin’s	$c_L(v) = \frac{ \{v \mid d(v, v) < \infty\} ^2}{\sum_{u \in G: d(v, u) < \infty} d(v, u)}$	$c_L^r(v) = \frac{ N^r[v] ^2}{\sum_{u \in N^r[v]} d(v, u)}$

In these localised variants, we compute the index of a vertex with respect to its r^{th} neighbourhood rather than with respect to the whole graph. The first natural question is the utility of such localised variants (and their accuracy in reflecting the global measure). We remark that Marsden demonstrated that for some networks, calculating the measure for a vertex v inside its closed neighbourhood $G[N[v]]$ can be used as a viable substitute for the full measure [178]. In the context of computer networks, Pantazopoulos, Karaliopoulos, and Stavrakakis [202] consider local variants which lend themselves to distributed computing and found a close correlation to the full measures on a sample of networks. We see that in certain applications, local knowledge is either all that is necessary, or all that is available.

But we can also show experimentally that our variants reliably capture the top ten percent in (arbitrarily) selected networks of our real-world corpus (see Chapter 16 for a description of the used networks). To that end, we compare in Figure 4 the 10% of highest ranking ver-

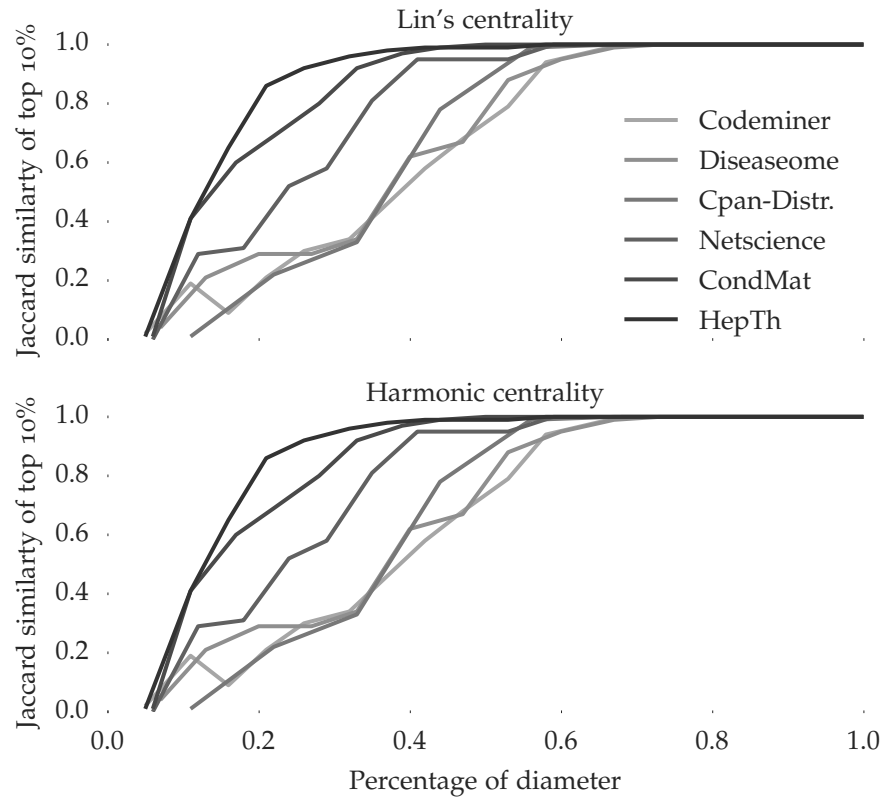


Figure 4: Top-10% similarity of localised indices compared with their respective global variant. The measurements are scaled according to the networks' diameter in order to be comparable. The results for closeness centrality follow the same trend and is omitted here.

tices, as identified by our localised variants, to those 10 percent identified by the respective full centrality measure. Specifically, we use the Jaccard index [143]—defined as $|A \cap B|/|A \cup B|$ for two sets A, B —to measure similarity⁶. We can conclude that for most of the tested networks, already a fifth of the diameter is sufficient to reliably identify most high-ranked vertices. It seems that the larger the network, the better this relation becomes: HepTh and CondMat have around 7000 and 16000 vertices, respectively, Netscience around 1500, Cpan-distributions around 2700 and the other two below 1500. This is probably owed to the fact that the top 10% set grows with the networks size and therefore is simply more stable with respect to the approximate measure.

We conclude that the localised centralities seem—according to this preliminary study—to be a viable alternative to their non-local counterparts. And using techniques based on structural sparseness, their computation is, in theory, much more efficient. Note another application of Theorem 26: we can also use it to compute hop-plots (intro-

⁶ Since we compare sets of equal size, the measures precision and recall—and accordingly the F_1 -score—are all the same. The Jaccard index is better suited for this situation.

duced above) more efficiently, as long as we truncate them above a certain threshold r .

Let us consider a second application of the structural sparseness toolkit. Recall that by using low-treedepth colourings, we can count the occurrence of small patterns—be it as subgraphs, induced subgraphs, homo- or isomorphisms—in linear time (cf. Theorem 24 and 25). As it turns out, such counting algorithms could be immensely useful in the area of computational biology. We highlight three domain-specific applications of calculating the frequency of small fixed pattern graphs inside a network to make our case.

Network motif

A *network motif* is a subgraph (not necessarily induced and possibly labelled) that appears with a significantly higher frequency in a real-world network than one would expect by pure chance. Introduced by Milo *et al.* [185] under the hypothesis that such frequently occurring structures have a functional significance, motifs have been identified in a plethora of different domains—including protein-protein-interaction networks [5], brain networks [230] and electronic circuits [141]. We refer to the surveys of Kaiser, Ribeiro, and Silva [210] and Masoudi-Nejad, Schreiber, and Kashani [180] for an extensive overview.

Graphlets, – degree distribution

Graphlets are a related concept, though their application is in an entirely different scope. While motifs are used to identify and explain local structure in networks, graphlets are used to ‘fingerprint’ them. Pržulj introduced the *graphlet degree distribution* as a way of measuring network similarity [208]. To compute it, one enumerates all connected graphs up to a fixed size (five in the original paper) and computes for each vertex of the target graph how often it appears in a subgraph isomorphic to one of those patterns. Since some graphlets exhibit higher symmetry than others, the computation takes into account all possible automorphisms. The degree distribution then describes for each graphlet G_i , how many vertices of the target graph are contained in $0, 1, 2, \dots$ subgraphs isomorphic to G_i —more precisely, in how many orbits of the respective automorphism groups it appears in. Note that if the set of graphlets only contains the single-edge graph this computation yields exactly the classical degree distribution.

The application of this distribution is two-fold: On the one hand, it can be used to measure similarity of multiple networks [226], in particular for networks derived from biological data [139]. On the other hand, the local structure around a vertex can reveal domain-specific functions. This is the case for protein-protein interaction networks where local structure correlates with biological activity [183], which has been applied to identify cancer genes [182] and construct phylogenetic trees [166]. Graphlets have further been used in analysis of workplace dynamics [236], photo cropping [38] and Denial-of-Service attack detection [207].

A third application of subgraph counting was given by Ugander, Backstrom, and Kleinberg *et al.* [240]: their empirical analysis and subsequent modelling of social networks revealed that there is an inherent

Domination problems

bias towards the occurrence of certain subgraphs. Thus the frequencies of small subgraphs seems an important indicator for the social domain, similar to the role of graphlet frequencies in biological networks.

A final important area of algorithmic questions related to complex networks are *domination problems*. We saw previously that structurally sparse classes admit constant-factor approximation of r -DOMINATING SET and linear or almost linear kernels for DOMINATING SET; further, by the first-order meta-result, we know that r -DOMINATING SET is fixed-parameter tractable in nowhere dense classes. Hence bridging the gap between complex networks and the theory of structurally sparse graphs provides us with more efficient means to solve domination-style problems. Dominating sets themselves can be used to compute efficient sensor placement for disease detection (see, *e.g.*, the work by Eubank *et al.* [88]). Partial dominating sets, *i.e.* those that maximize the number of dominated vertices, have applications as a centrality index in social network that identifies important actors who govern the information-flow [33]. Feng Wang *et al.* [244] defined *positive dominating sets* to model stable influences in networks. A set D is a positive dominating set if every node of a network has at least half of its neighbour in D . Finally, connected dominating sets have applications in wireless routing schemes (see, *e.g.*, the paper by Jie Wu and Hailan Li [142]).

We conclude that if we are able to demonstrate that real-world networks are structurally sparse, then the above mentioned applications of algorithms designed using the toolkit based on bounded expansion and nowhere dense classes make a strong case for bringing these techniques into scientific application. In combination with well-tested heuristical improvements (*e.g.* [248], [211]) these techniques might be able to push the boundary of computational feasibility far beyond where it lies today.

*Remember that all models are wrong;
the practical question is how wrong do they have to be to not be useful.*
—George E.P. Box

Science has a long history of tackling complicated natural phenomena by constructing *models*. The philosophical and scientific theory behind models is rich and difficult, but their every-day usage in the scientific endeavour is almost instinctive. Given the byzantine structure of real-world networks and the recent interest by researchers from Physics and Complex System Theory, it is not surprising that a plethora of (often competing) models exist.

Generative models

We can categorise modelling of networks into two broad approaches. The first type of models mimic a simplified evolution of a network, building it up piece-by-piece. The most common scenario is an ‘attachment model’ in which nodes are successively added to the network and randomly connected to the existing network by a specific probability distribution. This ‘bottom-up’ approach has great appeal: not only can we generate artificial networks and match their statistics against real-world data, at the same time we also obtain an explanation for these statistics—‘the network has a power law degree distribution because members of it attach by a preferential attachment rule’. Their huge appeal is, however, also a drawback: it is very easy to jump to conclusion and accept the proposed mechanism as *the* explanation, rendering us blind to possible alternatives.

Let us call such models *generative* (a term borrowed from machine learning theory and philosophy of science): by introducing a hidden process—hidden in the sense that we do not know that it really was *this* process that is responsible for generating our data in the real world—we seek to explain the final form. We necessarily expect these models to replicate or approximate a whole range of statistics, otherwise we have to call into question the model’s ability to *explain* network formation.

Descriptive models

Polar opposites to generative models are *descriptive models*. They are designed to replicate a fixed set of statistics without any claim that other statistics or even the network’s evolution is reproduced. Such ‘top-down’ approaches have the benefit of being far removed from intuition and hence bias, but on the flip-side they lack explanatory power: if they work, they work, but the only explanation we can proffer is that the model replicates the measured statistic well.

Random graphs

In essence, network models are simply random graphs and we will use the tools developed in random graph theory to reason about them. Since the term ‘random graph’ has become somewhat synonymous

Erdős-Rényi graphs

with the Erdős-Rényi model, we adapt the term *random graph model* in the following. This term also reminds us that the random graphs we are working with have a semantic attached to them and are not just detached mathematical objects.

The classical study of random graphs was initiated by Erdős and Rényi [86], and Gilbert [123] in 1959. By the available metrics¹, however, their 1960 paper ‘On the evolution of random graphs’ [87] seems to be the one that sparked the study of random graphs. In hindsight, their stated motivation seems almost prophetic:

It seems plausible that by considering the random growth of more complicated structures (*e.g.* structures consisting of different sorts of “points” and connections of different types) one could obtain fairly reasonable models of more complex real growth processes (*e.g.* the growth of a complex communication net consisting of different types of connections, and even of organic structures of living matter, *etc.*).

The Erdős-Rényi model was, indeed, for a long while the only real contender in modelling networks—probably due to the fact that the statistics it does *not* replicate well only become obvious if the networks grow large enough. Even today the model serves as a mathematically tractable baseline (in contrast to the flurry of other models that so far have only been analysed empirically).

Phase transition

We will use Gilbert’s model in the following (which is asymptotically equivalent to the model proposed by Erdős and Rényi): by $G(n, p)$ we denote the random graph on n vertices obtained by adding every possible edge independently with probability p . As usual, we allow p to be function of n .

Threshold

The probably most famous property of Erdős-Rényi graphs is the behaviour of their connectivity [87]: the graph $G(n, p)$ with $np = 1$ contains almost surely a connected component of size $\Omega(n^{2/3})$, in the case of $np > 1$ even of size $\Omega(n)$, but for $np < 1$ it will almost surely contain not component of size larger than $O(\log n)$. Because the transition from not exhibiting a property (no large components) to exhibiting it (having a large connected component) is extremely sharp with respect to the one model parameter p , the term *phase transition* was loaned from the physical sciences and has stuck².

Most interesting properties seem to follow such a phase transition. In fact, a property being monotone already suffices to show that such it must exist. Note that any monotone property Π has the very intuitive trait that

$$\Pr[G(n, p) \in \Pi] < \Pr[G(n, p + \varepsilon) \in \Pi]$$

¹ Google scholar attributes a citation index of 3 for [86] and 6391 for [87]. CiteseerX does not even list [86].

² With respect to random graphs, the term was already used in 1970 by Stepanov [232].

for $\varepsilon > 0$ (this can be easily seen by a simple coupling argument). The *threshold* of a property Π is a function $p_0(n)$ that satisfies

$$\Pr[G(n, p) \in \Pi] \rightarrow \begin{cases} 0 & \text{if } p \in o(p_0), \\ 1 & \text{if } p_0 \in o(p). \end{cases}$$

Bollobás and Thomas showed that every monotone graph property has a threshold [31].

In the context of this thesis we are interested in whether networks exhibit structural sparseness and we will use (as so often) Erdős-Rényi graphs as a first-order approximation to that question. Let us first convince ourselves that stronger notions of sparseness do *not* apply.

Theorem 37 (Fountoulakis, Kühn, Osthus,[106]). *For every $\mu > 1$ there exists a constant δ such that a.a.s. $K_t \preceq_m G(n, \mu/n)$ with $\delta\sqrt{n} \leq t \leq 2\sqrt{cn}$.*

Therefore $G(n, \mu/n)$ does not exclude any constant-sized graph as a minor. The question whether they contain arbitrarily dense *topological* minors seems to be open. On the positive side, Nešetřil, Ossona de Mendez, and Wood proved that $G(n, \mu/n)$ has bounded expansion a.a.s. [193]. We extrapolate that bounded expansion seems to be the most sensible option to show that complex networks exhibit structural sparseness.

15.1 EXPANSION OF RANDOM GRAPH MODELS

Before we proceed, we need to clarify what we mean by a random graph model and what we mean by such a model being structurally sparse—after all, these terms are defined for deterministic objects. We will use a conservative approach here that stands on solid historical ground: a random graph model has a property \mathcal{P} if and only if the asymptotic probability that it generates graphs exhibiting \mathcal{P} tends to one in the limit. Let us first fix the relevant notation.

$P[\bullet], E[\bullet], \text{Var}[\bullet], M[\bullet]$

We usually denote random variables by upper-case letters. Probabilities are denoted by $P[\bullet]$, expectation, variance, and median by $E[\bullet]$, $\text{Var}[\bullet]$, $M[\bullet]$, respectively. If we need to clarify which probability measures we employ, we use subscripts like $P[\bullet]_M$. For a sequence of random variables $(X_n)_{n \in \mathbb{N}}$ and a random variable Y , recall that (X_n) *converges in distribution* to X if it holds that

Convergence in distribution

$$\forall k \lim_{n \rightarrow \infty} P[X_n \leq k] = P[X \leq k].$$

We denote the convergence in distribution with $(X_n) \xrightarrow{d} X$.

Random graph model

A *random graph model* is a sequence of random variables $(G_n)_{n \in \mathbb{N}}$ over n -vertex graphs. For simplicity, we fix $V(G_n) = [n]$. The *parametrisation* of the model is a function $\rho: \mathbb{N} \rightarrow \mathbb{R}^t$ that creates a tuple of t parameters depending on n which in turn determine the probability distribution of each variable G_n . By $G(n, \rho(n))$ we denote the random

Minors of random graphs

variable G_n with the probability distribution prescribed by the model with parameters $\rho(n)$. In order to distinguish random models we will introduce superscripts like G^{CL} or G^{KL} .

For a random graph model $G(n, \rho(n))$ and an integer r the notation $G(n, \rho(n)) \tilde{\nabla} r$ denotes a random variable over sets of graphs with at most n vertices whose probability distribution is given by

$$\Pr[G(n, \rho(n)) \tilde{\nabla} r = A] = \sum_{G:A=G \tilde{\nabla} r} \Pr[G(n, \rho(n)) = G],$$

where A is a set of graphs. With this definition the quantities $\tilde{\nabla}_\bullet$ and $\tilde{\omega}_\bullet$ are well-defined as rational-valued random variables. As noted above, we study the properties of random graphs in the limit and hence define the property of having bounded expansion as follows.

Structural density of random graphs

Definition 28. A graph model $G(n, \rho(n))$ has *bounded expansion asymptotically almost surely (a.a.s.)* if there exists a function f such that for all $r \geq 0$

$$\lim_{n \rightarrow \infty} \Pr[\tilde{\nabla}_r(G(n, \rho(n))) < f(r)] = 1.$$

It has *bounded expansion with high probability (w.h.p.)* if for every $c \geq 1$ there exists a function f such that, again for all $r \geq 0$,

$$\Pr[\tilde{\nabla}_r(G(n, \rho(n))) < f(r)] \geq 1 - O(n^{-c}).$$

The very same definition is possible to define when a random graph model is nowhere dense.

Definition 29. A graph model $G(n, \rho(n))$ is *a.a.s. nowhere dense* if there exists a function f such that for all $r \geq 0$

$$\lim_{n \rightarrow \infty} \Pr[\tilde{\omega}_r(G(n, \rho(n))) < f(r)] = 1.$$

It is *nowhere dense w.h.p.* if for every $c \geq 1$ there exists a function f such that, again for all $r \geq 0$,

$$\Pr[\tilde{\omega}_r(G(n, \rho(n))) < f(r)] \geq 1 - O(n^{-c}).$$

The following notions are needed to prove negative results about random models.

Definition 30. A graph model $G(n, \rho(n))$ is *a.a.s. somewhere dense* if there exists $r \in \mathbb{N}$ such that for all functions f it holds that

$$\lim_{n \rightarrow \infty} \Pr[\tilde{\omega}_r(G(n, \rho(n))) > f(r)] = 1.$$

It is *not a.a.s. nowhere dense* if there exists $r \in \mathbb{N}$ such that for all functions f it holds that

$$\lim_{n \rightarrow \infty} \Pr[\tilde{\omega}_r(G(n, \rho(n))) > f(r)] > 0.$$

Note that the above definition for random graphs is, in contrast to the definition of structural sparseness for graph classes, not a dichotomy: we can easily build an artificial random graph model where the probability that a dense shallow minor exists converges to some value bounded away from zero and one. This might now just seem like a technicality, but we will see two examples of random graph models developed to replicated certain aspects of complex networks that fall exactly in this category.

We will base our positive proofs on the characterisation of bounded expansion classes stated in Theorem 21. Let us recall the statement:

Theorem 21 (Nešetřil, Ossona de Mendez, Wood [193]). *A graph class \mathcal{G} has bounded expansion if and only if there exist real-valued functions f_{thresh} , f_{deg} , $f_{\tilde{\nabla}}$, f_H such that for all $G \in \mathcal{G}$ the following two conditions hold:*

1. For all $\varepsilon > 0$ either $|G| \leq f_{\text{thresh}}(\varepsilon)$ or it holds that

$$|\{v \in V(G) : \deg(v) \geq f_{\text{deg}}(\varepsilon)\}| \leq \varepsilon \cdot |G|.$$

2. For all $r \in \mathbb{N}$, all $H \subseteq G$ with $\tilde{\nabla}_r(H) > f_{\tilde{\nabla}}(r)$ satisfy

$$|H| \geq f_H(r) \cdot |G|.$$

Finding the first pair of functions $f_{\text{thresh}}, f_{\text{deg}}$ is usually straightforward, the real challenge lies in proving that functions $f_{\tilde{\nabla}}, f_H$ exist. Consider the contra-positive of the second condition: in terms of random graph models, we want to show that subgraphs that span at most a $f_H(r)$ -fraction of the vertices of G have their (topological) grad bounded by $f_{\tilde{\nabla}}$ with high probability.

The following basic lemma will help to simplify the following proofs by enabling us to work with randomly chosen subgraphs.

Lemma 80. *Let X_1, \dots, X_n be binary random variables and let $S = \sum_{i=1}^n X_i$. Let further $I \in [n]$ be uniformly distributed. Then*

$$\Pr[S \geq 1] \leq n \cdot \Pr[X_I = 1].$$

Proof. By Markov's inequality we have that

$$\Pr[S \geq 1] \leq E[S] = \sum_{i=1}^n \Pr[X_i = 1].$$

Observe that

$$\Pr[X_I = 1] = \sum_{i=1}^n \Pr[I = i] \cdot \Pr[X_i = 1] = \frac{1}{n} \sum_{i=1}^n \Pr[X_i = 1],$$

and hence

$$\Pr[S \geq 1] \leq n \cdot \Pr[X_I = 1]. \quad \square$$

We apply this statement to subgraphs in a random graph and obtain the following corollary that lies closer to our application.

Corollary 21. *Let $G(n, \rho(n))$ be a random graph model with parametrisation ρ and Π be a graph property. Then*

$$\Pr[\exists X \subseteq V(G) : G[X] \in \Pi] \leq \sum_{k=1}^n \binom{n}{k} \Pr[G[Y_k] \in \Pi],$$

where Y_k is a k -vertex subset of $V(G)$ chosen uniformly at random.

15.2 FIXED DEGREE DISTRIBUTIONS

As noted by Newman [195], the observations that the degree distribution is a qualitative important statistic date back to Rapoport's treatment of random networks³ in the 1950s (see, *e.g.*, the work of Solomonoff and Rapoport on the connectivity of random graphs [229]). A first improvement of the random graph model, whose degree distribution does not match those observed in real world networks, is therefore to *prescribe* the degree distribution and sample randomly from all graphs that exhibit that distribution. While this does not solve some of the other problems associated with the Erdős-Rényi model—other statistics like the clustering coefficient do also not match empirical values—it enables us to work with a model which is still mathematical tractable and at least *closer* to real-world phenomena.

Before getting into the gory technical details, we should briefly discuss which degree distributions are commonly observed. Early claims that all networks follow a power law degree distribution have been subsequently refuted; as it turns out, distinguishing such functions from other candidate distributions is an intricate statistical problem [49]. Modern approaches seem to favour a multitude of distributions (see Table 1 for an overview).

Name	Definition $f(d)$	Parameters
Power law	$d^{-\gamma}$	$\gamma > 2$
Power law w/ cutoff	$d^{-\gamma} e^{-\lambda d}$	$\gamma > 2, \lambda > 0$
Exponential	$e^{-\lambda d}$	$\lambda > 0$
Stretched exponential	$d^{\beta-1} e^{-\lambda d^\beta}$	$\lambda, \beta > 0$
Gaussian	$\exp\left(-\frac{(d-\mu)^2}{2\sigma^2}\right)$	μ, σ
Log-normal	$d^{-1} \exp\left(-\frac{(\log d - \mu)^2}{2\sigma^2}\right)$	μ, σ

Table 1: A selection of established functions used to model degree-distributions of complex networks [49]. The functions here are listed without the necessary normalization factor.

³ He used the term 'axone density'.

Let us now formalise the degree distributions and sequences. Given a graph G , the *degree sequence* $D(G)$ of G is the sequence $(\deg(v))_{v \in G}$. We say that two graphs G_1 and G_2 have the same degree sequence, if $D(G_1) = \pi(D(G_2))$ for some permutation π . A sequence of n integers $(d_i)_{1 \leq i \leq n}$ is a degree sequence if it can be realised by a graph, i.e. there exists a graph G with $D(G) = (d_i)_{1 \leq i \leq n}$. Such sequences are also called *graphical*.

To define degree distributions, consider a random variable D_n describing the degree of a vertex chosen uniformly at random from an n -vertex graph G . The pmf f_n for D_n is then given by

$$f_n(d) = \frac{b_n(d)}{n} = \frac{1}{n} \sum_{v \in V(G)} \mathbb{1}[\deg(v) = d],$$

where $b_n(d)$ denotes the absolute frequency of degree- d vertices in G .

Definition 31 (Degree distribution). A n -vertex degree distribution is a random variable D with probability mass function f such that

1. $f(d) = 0$ for $d \leq 0$ and $d \geq n - 1$, and
2. $nf(d) \in \mathbb{N}_0$ for all $d \in \mathbb{N}$.

We say that a degree sequence $(d_i)_{1 \leq i \leq n}$ matches an n -vertex degree distribution D_n with pmf f_n if for every $1 \leq k \leq n - 1$ it holds that

$$\sum_{i=1}^n \mathbb{1}[d_i = k] = nf_n(k).$$

Consequently, a graph G matches a degree distribution D_n if its degree sequence does.

Degree distribution sequence

Since we will consider sequences of random graphs we need to introduce a related notation of sequences of degree distributions.

Definition 32 (Degree distribution sequence, limit, sparse). A *degree distribution sequence* is an infinite sequence (D_n) of n -vertex degree distributions. A random variable D is the *limit* of (D_n) if $(D_n) \xrightarrow{d} D$. We say that \mathcal{D} is *sparse* if $E[D] < \infty$ and $(E[D_n])_{n \in \mathbb{N}} \rightarrow E[D]$.

To motivate the definition of *sparse* sequences, note that for a degree sequence D_n we have that

$$E[D_n] = \sum_d df_n(d) = \frac{1}{n} \sum_{d=1}^{n-1} db_n(d),$$

thus for a graph G with degree distribution D_n it holds that $E[D_n]$ is exactly its average degree $\bar{d}(G)$. The condition of a degree sequence being sparse will not quite suffice to prove structural sparseness: the situation parallels how graph class with bounded average degree can still harbour dense structures. Consider, for example, the degree distribution sequence of the class $K_n[E/1]$ (the class consisting of all one-subdivided cliques). While the sequence has constant expectation, the graphs matching it are certainly not structurally sparse. This observation motivates the following stronger condition.

Tail-bound

Definition 33 (Tail-bound). A degree distribution sequence (D_n) with limit D has the function h as its *tail-bound* if there exists a constant $\tau \geq 0$ such that for all $d \geq \tau$ it holds that

$$\Pr[D \geq d] = O\left(\frac{1}{h(d)}\right).$$

Referring back to Table 1, all listed function have a tail-bound that is at least quadratic. This is self-evident for power-laws with $\gamma \geq 2$; for the other functions we simply note that their second moment exists and hence by Chebyshev's inequality satisfy

$$\Pr[D \geq d] \leq \frac{\text{Var}[D^2]}{d^2 + \mathbb{E}[D^2]} = O\left(\frac{1}{d^2}\right).$$

for $d \geq \mathbb{E}[D]$.

Observation 5. Let (D_n) be a sparse degree distribution sequence with limit D . Then $\mathbb{M}[D_n] \rightarrow \mathbb{M}[D]$.

Given a degree distribution sequence $\mathcal{D} = (D_n)_{n \in \mathbb{N}_0}$ with limit D and an integer n , we are faced with the task to sample graphs uniformly at random from the set

$$\{G \mid G \text{ has degree distribution } D_n\}.$$

Two methods to accomplish this task—with certain caveats—have been put forward: the *configuration model* as described by Bender and Canfield [19] and the model proposed by Chung and Lu [47, 46]. The latter is a special case of what has been discussed in the mathematical literature as inhomogeneous random graphs (see, e.g., the work by Bollobás, Janson, and Riordan on the phase transition of such graph [27]).

Configuration model

To sample a graph according to the configuration model, we proceed as follows:

1. Build a degree sequence $(d_i)_{1 \leq i \leq n}$ that matches D_n .
2. Construct a vertex set $V_C = \{v_i^1, \dots, v_i^{d_i}\}_{1 \leq i \leq n}$, i.e. create d_i copies (called *stubs*) for what will be the vertex v^i in the final graph.
3. Generate an auxiliary graph H with vertex set V_C and a random matching as its edge set.
4. Assemble the multi-graph G' with vertex set $\{v_i\}_{1 \leq i \leq n}$ and

$$|E(v_i, v_j)| = |E_H(\{v_i^1, \dots, v_i^{d_i}\}, \{v_j^1, \dots, v_j^{d_j}\})|,$$

that is, we connect the vertices v_i and v_j with as many edges as we find between their respective copy-classes in H .

5. Return the graph G derived from G' by removing all parallel edges and loops.

Graphs generated this way will henceforth be denoted by $G^{\text{CF}}(D_n)$. The name *stubs* derives from the following picture of the process: we affix to every vertex a number of half-edges, the stubs, that matches its degree according to the generated degree sequence, and then obtain the multi-graph by randomly wiring the stubs to each other. Instead of generating a matching between the stubs, we can instead choose a random permutation of them and match them up pair-by-pair according to that permutation. This view of the process will be beneficial later on.

It is a priori not clear that this procedure generates graphs with the correct degree sequence. The intermediate multi-graph G' trivially exhibits the degree sequence $(d_i)_{1 \leq i \leq n}$ and thus has the degree distribution D_n . The last step, however, might skew the result by removing parallel edges and loops—we therefore need the probability that G' contains such offending edges to be reasonably low. The conditions under which this is the case have been proved by Molloy and Reed [186] and subsequently improved by Janson, whose result we present here using our own notation.

Theorem 38 (Janson [147]). *Let $(D_n)_{n \in \mathbb{N}_0}$ be a degree distribution sequence. Then we have that*

$$\liminf_{n \rightarrow \infty} \mathbb{P}[G^{\text{CF}}(D_n) \text{ is simple}] > 0 \iff \mathbb{E}[D_n^2] = O(\mathbb{E}[D_n]).$$

The original formulation of the theorem's condition is that

$$\begin{aligned} \sum_{d \geq 0} d^2 b_n(d) = O\left(\sum_{d \geq 0} d b_n(d)\right) &\iff \limsup_{n \rightarrow \infty} \frac{\sum_{d \geq 0} d^2 b_n(d)}{\sum_{d \geq 0} d b_n(d)} < \infty \\ &\iff \limsup_{n \rightarrow \infty} \frac{\sum_{d \geq 0} d^2 f_n(d)}{\sum_{d \geq 0} d f_n(d)} < \infty \\ &\iff \limsup_{n \rightarrow \infty} \frac{\mathbb{E}[D_n^2]}{\mathbb{E}[D_n]} < \infty, \end{aligned}$$

which justifies our reformulation. Note that in the case of sparse degree distribution sequences this condition is equivalent to saying that $\text{Var}[D]$ is finite—this already limits the possible degree distributions to those that have a finite variance. Luckily, this is the case for all degree distributions listed in Table 1 with the exceptions of the power law where only for $\gamma > 3$ the variance is finite.

Chung–Lu model

The second method to sample graphs with a prescribed degree distribution, the Chung–Lu model, forgoes the above problems by generating graphs whose *expected* degree distribution matches the given one. Given D_n , it constructs a random graph as follows:

1. Build a degree sequence $(d_i)_{1 \leq i \leq n}$ that matches D_n . We will call d_i the *weight* of the vertex i .
2. Create a graph on n vertices v_1, \dots, v_n and connect each pair of vertices v_i, v_j with probability $d_i d_j / m$ where $m = \sum_{k=0}^n d_k$.

Random graphs generated according to this procedure will henceforth be denoted by $G^{\text{CL}}(D_n)$. The model ensures that the expected degree of a vertex matches the degree we assigned it in the degree sequence. While it is unclear how close the degree distribution is matched, the models' simplicity makes it both easy to analyse and implement. It is noteworthy that similar models have been proposed to generate scale-free networks in cases where preferential attachment (presented in the next section) is an implausible mechanism [40].

Generalised random graphs

A variation of the above model was introduced by van der Hofstad [241] under the name *generalised random graphs*. He cites it as an important special case of the more general *inhomogeneous random graphs* introduced by Bollobás, Janson, and Riordan [27]. We denote the model by $G^{\text{GRG}}(D_n)$ here. The only difference to the Chung–Lu model lies in how the edge-probability is computed from the weights: for vertices i, j with weights d_i, d_j the edge ij is present in the graph with probability $d_i d_j / (m + d_i d_j)$. Again $m = \sum_{k=0}^n d_k$ is the total sum of weights.

Van der Hofstad proved that the Chung–Lu model is essentially equivalent to generalised random graphs if the degree distribution has a certain shape:

Theorem 39 (van der Hofstad [241, Theorem 6.19]). *The random graphs $G^{\text{CL}}(D_n)$ and $G^{\text{GRG}}(D_n)$ are asymptotically equivalent if $E[D_n^3] = o(\sqrt{n})$.*

Since the Chung–Lu model is slightly easier to handle we will not work directly with generalised random graphs but instead use the above theorem to carry results over.

Household structure

As network models, both the configuration and the Chung–Lu model suffer from some shortcomings. While they, by design, generate graphs with the correct degree-distribution and small diameter, other statistics found in complex networks are not replicated. In particular, both models have a vanishing clustering-coefficient (see, e.g., Newman's survey [195]).

Since this statistic is critical in e.g. disease propagation research, methods to 'fix' these models have been put forward. A notable example is the *configuration model with household structure* as defined by Ball, Sirl, and Trapman [16]. For this variant, one samples a graph with a prescribed degree sequence and then replaces every vertex by a constant-sized 'household'-graph (for example a clique), distributing the edges incident to a household uniformly to the vertices that comprise it. The resulting graph has provably a constant clustering coefficient.

Hybrid models

A different approach proposed by Chung and Lu is to take two graphs, one with the small-world property, the other with 'high local connectivity' and combine them into a hybrid graph [48] (another variation is proposed in [8]). They claim that the resulting graph has both the small-world property and high clustering, however, their definition of clustering remains rather vague. It is in particular unclear what the clustering coefficient of the resulting graph is.

15.3 DEGREE DISTRIBUTIONS, TAILS AND EXPANSION

The main goal of this section will be the proof of the following theorems. To this end, we will first show the behaviour of the Chung–Lu random graphs with respect to $\tilde{\nabla}_0$ and ω . By an appropriate coupling arguments, we can then relate the statistics $\tilde{\nabla}_r$ and $\tilde{\omega}_r$ of different graph models to these base cases. As it turns out, the structural sparseness of the Chung–Lu and the configuration model is entirely determined by the *tail* of the prescribed degree distribution:

Theorem 40. *Let (D_n) be a sparse degree distribution sequence with tail $h(d)$. Both the configuration model $G^{\text{CF}}(D_n)$ and the Chung–Lu model $G^{\text{CL}}(D_n)$, with high probability,*

- *have bounded expansion for $h(d) = \Omega(d^{3+\epsilon})$,*
- *are nowhere dense (with unbounded expansion) for $h(d) = \Theta(d^{3+o(1)})$,*
- *and are somewhere dense for $h(d) = O(d^{3-\epsilon})$.*

Since we can emulate household structures (cf. Section 15.2) by simply taking the lexicographic product with some constant-size clique and then taking a subgraph, Theorem 40 immediately implies the same for those variants.

Corollary 22. *Let (D_n) be a sparse degree distribution sequence. Then the configuration model $G^{\text{CF}}(D_n)$ as well as the Chung–Lu model $G^{\text{CL}}(D_n)$ with households have bounded expansion w.h.p. if the sequence has a supercubic tail-bound and nowhere dense w.h.p. if it has a cubic tail-bound.*

By Theorem 39, the random graphs $G^{\text{CL}}(D_n)$ and $G^{\text{GRG}}(D_n)$ are asymptotically equivalent when $E[D_n^3] = o(\sqrt{n})$. Taken together with Theorem 40, we obtain:

Corollary 23. *Let (D_n) be a sparse degree distribution sequence with tail $h(d)$. Then the generalised random graphs $G^{\text{GRG}}(D_n)$, asymptotically almost surely,*

- *have bounded expansion for $h(d) = \Omega(d^{3+\epsilon})$, and*
- *are nowhere dense (with unbounded expansion) for $h(d) = \Theta(d^{3+o(1)})$.*

In both the Chung–Lu and the configuration model, the first phase consist of assigning weights to vertices according to a degree distribution D_n . This random experiment chooses degrees without replacement, therefore we need to ensure that the important properties of D_n carry over even if a fraction of the vertices have been uncovered already, that is, we know their weight and hence cannot assume they are randomly distributed.

Since in the following proofs low weights are always preferable, we consider a ‘worst-case’ variable describing the degree of a vertex after at most n/c other vertices have been assigned the lowest available degrees. Luckily, sparse degree distributions are robust under truncation up to the median. Using this idea, the following lemma shows that we can, up to a point, assume that the vertex degrees are drawn independently according to a modified distribution.

Lemma 81. Let (D_n) be a sparse degree distribution sequence with limit D and a tail-bound h . Let $\mu_{1/2} = \mathbf{M}[D]$ be the median of D . Then the sequence (\hat{D}_n) defined via

$$\Pr[\hat{D}_n = d] = \Pr[D_n = d \mid D_n \geq \mu_{1/2}]$$

is sparse and has tail-bound h .

Proof. Let \hat{D} be defined as the conditioned random variable $D \mid D > \tau$. Then we have that

$$\begin{aligned} \mathbf{E}[\hat{D}] &= \sum_{d>0} d \Pr[D = d \mid D \geq \mu_{1/2}] = \sum_{d \geq \mu_{1/2}} d \cdot \frac{\Pr[D = d]}{\Pr[D \geq \mu_{1/2}]} \\ &= \frac{\mathbf{E}[D]}{\Pr[D \geq \mu_{1/2}]} \leq \frac{\mathbf{E}[D]}{2}. \end{aligned}$$

Hence $\mathbf{E}[\hat{D}]$ is finite and (\hat{D}_n) is sparse.

Let τ be the constant such that for $d \geq \tau$ it holds that $\Pr[D \geq d] = O(1/h(d))$. To see that the same holds for (\hat{D}_n) , note that

$$\Pr[\hat{D} \geq d] = \Pr[D \geq d \mid D \geq \mu_{1/2}] = \frac{\Pr[D \geq \max\{d, \mu_{1/2}\}]}{\Pr[D \geq \mu_{1/2}]}.$$

Hence for $\hat{\tau} \geq \max\{\tau, \mu_{1/2}\}$ for all $d \geq \hat{\tau}$ the bound

$$\Pr[\hat{D} \geq d] = O\left(\frac{1}{h(d)}\right)$$

holds, as claimed. \square

In the remainder of this section, we work towards a proof of Theorem 40. Ultimately, our approach couples the density of shallow minors for graphs with degree distribution D to the density of subgraphs of a random graph with distribution ηD , where η is an appropriate scaling factor. This factor ultimately traces back to the value of the sum $\sum_{d=1}^{\Delta} \frac{1}{d^{\gamma-2}}$ which is a constant for $\gamma > 3$, roughly $\log \Delta$ for $\gamma = 3$ and around $\Delta^{3-\gamma}$ for $\gamma < 3$.

15.3.1 The supercubic regime

We begin our investigation by considering degree distributions whose tail is bounded by a supercubic function. As discussed earlier, most real-world degree distributions seem to fall into this pattern since they exhibit a sharp *exponential* cut-off.

Lemma 82. Consider vertices $\{v_i\}_{i \in [k]}$ with associated weights $\{d_i\}_{d_i \in [\Delta]}$, for some $\Delta \in \mathbb{N}$. Let G be a random graph on these vertices where each edge $v_i v_j$ is independently present with probability $\leq \beta d_i d_j / n$. Then

$$\Pr[\|G\| \geq \xi k] \leq \left(\frac{e\beta D^2}{2n\xi k e^{D^2/2n}} \right)^{\xi k}$$

if $D := \sum_i d_i$ satisfies $\beta D^2 \leq 2n\xi k$.

Proof. We associate a random variable X_{ij} with every edge $v_i v_j$. The expected number of edges is then

$$E\left[\sum_{ij} X_{ij}\right] = \sum_{ij} E[X_{ij}] \leq \sum_{ij} \frac{\beta d_i d_j}{n} = \frac{\beta}{2n} \left(\sum_i d_i\right)^2 = \frac{\beta D^2}{2n}.$$

We apply the Chernoff-bound

$$\Pr\left[\sum_{ij} X_{ij} \geq (1 + \delta) \frac{\beta D^2}{2n}\right] \leq \left(\frac{e^\delta}{(1 + \delta)^{1 + \delta}}\right)^{\beta D^2 / 2n}$$

choosing $\delta = \frac{2n\zeta k}{\beta D^2} - 1$ and obtain

$$\Pr\left[\sum_{ij} X_{ij} \geq \zeta k\right] \leq \frac{e^{\zeta k - \beta D^2 / 2n \zeta k}}{\left(\frac{2n\zeta k}{\beta D^2}\right)^{\zeta k}} = \left(\frac{e\beta D^2}{2n\zeta k e^{D^2 / 2n}}\right)^{\zeta k},$$

as claimed. \square

The first important puzzle piece is that the assumed tail-bound ensures the presence of only few high-weight vertices:

Lemma 83. *Let (D_n) be a sparse degree distribution sequence with limit D and with a polynomial tail-bound $h(d) = d^{\alpha+1}, \alpha > 1$. Then for every graph G that realises D_n and any set $X \subseteq V(G)$ of k vertices, the sum-of-degrees of X is at most*

$$\sum_{v \in X} \deg_G(v) \leq \sqrt{2} n^{1/\alpha} k^{\alpha-1/\alpha}.$$

Proof. Let $\Delta = h^{-1}(n) = n^{1/(\alpha+1)}$ be the maximal realizable degree. To maximize the sum-of-degrees we obviously need to pick the k vertices of highest degree. Hence we want the largest threshold δ such that

$$\sum_{d=\delta}^{\Delta} \frac{n}{d^{\alpha+1}} \geq k.$$

Bounding the sum by integrals, we see that

$$\sum_{d=\delta}^{\Delta} \frac{n}{d^{\alpha+1}} \sim \frac{n}{\alpha \delta^\alpha} \implies \delta \sim \left(\frac{n}{\alpha k}\right)^{1/\alpha}.$$

Given δ , the maximal possible sum-of-degrees is at most

$$\begin{aligned} \sum_{d=\delta}^{\Delta} \frac{n}{d^{\alpha+1}} \cdot d &= \sum_{d=\delta}^{\Delta} \frac{n}{d^\alpha} \sim \frac{n}{(\alpha-1)\delta^{\alpha-1}} \\ &= \frac{\alpha^{(\alpha-1)/\alpha}}{\alpha-1} \cdot \frac{nk^{(\alpha-1)/\alpha}}{n^{(\alpha-1)/\alpha}} \leq \sqrt{2} n^{1/\alpha} k^{\alpha-1/\alpha}, \end{aligned}$$

as claimed. \square

Using this property, we can now show that Chung–Lu graphs generated with degree distributions that have a supercubic tail will not contain dense subgraphs. Afterwards we will related the general case back to this basic case by coupling.

Lemma 84. *Let (D_n) be a sparse degree distribution sequence with limit D with a polynomial tail-bound $h(d) = d^{\alpha+1}, \alpha > 2$. Then there exists a function f such that for every $\zeta \geq 4ae^2$, every $c \geq 4e$ and every $n \geq f(\zeta, \alpha)$ it holds that*

$$\Pr[\exists H \subseteq G^{CL}(D_n) : |H| \leq n/c \text{ and } \nabla_0(H) \geq \zeta] \leq \frac{1}{n^\zeta}.$$

Proof. Using Lemma 80, we can bound the probability that a dense subgraph on k vertices exists by considering the probability that k randomly chosen vertices form a dense subgraph. Taking the union bound of all possible k (note that we need at least $2\zeta + 1$ vertices for a subgraph of density ζ , we simplify this lower bound to 2ζ), the probability of the aforementioned event is at most

$$\sum_{k=2\zeta}^{n/c} \binom{n}{k} \Pr[\|G^{CL}(D_n)[X_k]\| \geq k\zeta],$$

where X_k is a set of k vertices chosen uniformly at random.

By applying Lemma 82 to the random subgraph $G^{CL}(D_n)[X_k]$, we can bound the above probability by

$$\begin{aligned} & \sum_{k=2\zeta}^{n/c} \binom{n}{k} \sum_{d=k}^{k\Delta} \left(\frac{ecd^2}{2n\zeta k e^{d^2/2n}} \right)^{\zeta k} \Pr[D_k = d] \\ & \leq \sum_{k=2\zeta}^{n/c} \frac{n^k e^k}{k^k} \left(\frac{ec}{2n\zeta k} \right)^{\zeta k} \sum_{d=k}^{k\Delta} \frac{d^{2\zeta k}}{e^{\zeta k d^2/2n}} \Pr[D_k = d], \end{aligned}$$

where D_k denotes the degree-sum of k vertices chosen from D_n .

By Lemma 83 we know that for a degree-distribution with tail-bound $h(d) = d^{\alpha+1}$ the degree-sum of k vertices cannot exceed $\hat{D}_k := \sqrt{2}n^{1/\alpha}k^{\alpha-1/\alpha}$. Since

$$\sum_{d=k}^{\hat{D}_k} d^{2\zeta k} \leq \hat{D}_k^{2\zeta k+1} = (\sqrt{2}n^{1/\alpha}k^{\alpha-1/\alpha})^{2\zeta k+1},$$

we can upper-bound every term of the above sum by

$$\begin{aligned} & \frac{n^k e^k}{k^k} \left(\frac{ec}{2n\zeta k} \right)^{\zeta k} (\sqrt{2}n^{1/\alpha}k^{\alpha-1/\alpha})^{2\zeta k+1} \\ & \leq \frac{(2ec)^{(\zeta+1)k}}{(2\zeta)^{\zeta k}} \cdot \frac{n^{\frac{1}{\alpha}2\zeta k+1} k^{\frac{\alpha-1}{\alpha}(2\zeta k+1)}}{n^{(\zeta-1)k} k^{(\zeta+1)k}}. \end{aligned}$$

Choosing $\zeta >$ large enough, the first factor can be decreased below one (in particular, this is true for $\zeta \geq 4e^2$). Hence, we focus on the second factor:

$$\frac{n^{\frac{1}{\alpha}(2\zeta k+1)} k^{\frac{\alpha-1}{\alpha}(2\zeta k+1)}}{n^{(\zeta-1)k}} = \frac{k^{(\frac{\alpha-1}{\alpha}2\zeta - \zeta - 1)k + \frac{\alpha-1}{\alpha}}}{n^{(\zeta - \frac{1}{\alpha}2\zeta - 1)k - \frac{1}{\alpha}}} = \frac{k^{(\frac{\alpha-2}{\alpha}\zeta - 1)k + \frac{\alpha-1}{\alpha}}}{n^{(\frac{\alpha-2}{\alpha}\zeta - 1)k - \frac{1}{\alpha}}}.$$

Therefore the probability that a subgraph of density at least ζ exists is at most

$$\sum_{k=2\zeta}^{n/c} \frac{k^{\left(\frac{\alpha-2}{\alpha}\zeta-1\right)k+\frac{\alpha-1}{\alpha}}}{n^{\left(\frac{\alpha-2}{\alpha}\zeta-1\right)k-\frac{1}{\alpha}}}.$$

Consider the ratio of two consecutive terms: we see that

$$\begin{aligned} & \frac{(k-1)^{\left(\frac{\alpha-2}{\alpha}\zeta-1\right)(k-1)+\frac{\alpha-1}{\alpha}}}{n^{\left(\frac{\alpha-2}{\alpha}\zeta-1\right)(k-1)-\frac{1}{\alpha}}} \cdot \frac{n^{\left(\frac{\alpha-2}{\alpha}\zeta-1\right)k-\frac{1}{\alpha}}}{k^{\left(\frac{\alpha-2}{\alpha}\zeta-1\right)k+\frac{\alpha-1}{\alpha}}} \\ &= \left(\frac{n}{k}\right)^{\frac{\alpha-2}{\alpha}\zeta-1} \cdot \left(1-\frac{1}{k}\right)^{\left(\frac{\alpha-2}{\alpha}\zeta-1\right)(k-1)+\frac{\alpha-1}{\alpha}} \\ &\geq \left(\frac{n}{k}\right)^{\frac{\alpha-2}{\alpha}\zeta-1} \left(\frac{1}{e}\right)^{\frac{\alpha-2}{\alpha}\zeta-1} \left(\frac{1}{2}\right)^{\frac{\alpha-1}{\alpha}}. \end{aligned}$$

Since k is at most n/c , this ratio is at least

$$\frac{c^{\frac{\alpha-2}{\alpha}\zeta-1}}{e^{\frac{\alpha-2}{\alpha}\zeta-1} 2^{\frac{\alpha-1}{\alpha}}} \geq \left(\frac{c}{2e}\right)^{\frac{\alpha-2}{\alpha}\zeta-1},$$

where the last simplification holds for $\zeta \geq 4e^2 > 2^{\frac{\alpha-1}{\alpha-2}}$. We conclude that for $c \geq 4e$ the whole ratio is at least 2; and accordingly the bound

$$\sum_{k=2\zeta}^{n/c} \frac{k^{\left(\frac{\alpha-2}{\alpha}\zeta-1\right)k+\frac{\alpha-1}{\alpha}}}{n^{\left(\frac{\alpha-2}{\alpha}\zeta-1\right)k-\frac{1}{\alpha}}} \leq 2 \frac{(2\zeta)^{\left(\frac{\alpha-2}{\alpha}\zeta-1\right)2\zeta+\frac{\alpha-1}{\alpha}}}{n^{\left(\frac{\alpha-2}{\alpha}\zeta-1\right)2\zeta-\frac{1}{\alpha}}} \leq \frac{1}{n^{\zeta}}.$$

holds, where the last inequality follows for large enough n (depending only on α, ζ). \square

We pair the above lemma and Theorem 21 with a coupling argument to arrive at the following result. For simplicity, we define $\mathcal{E}_{\zeta,r}^X$ as the event that a random graph contains an r -shallow topological minor of density at least ζ whose embedding ϕ_V maps into X .

Lemma 85. *Let $G^R(n)$ be a random graph model with the following property: for every r there exists a sparse degree distribution (D_n) with tail-bound $h(d) = d^{\alpha+1}$, $\alpha > 2$ such that for every $\zeta \geq 4ae^2$ it holds that*

$$\Pr[\mathcal{E}_{\zeta,r}^X] \leq \Pr[\nabla_0(G^{\text{CL}}(D_n)[X]) \geq \zeta],$$

where X is a random set of at most $n/4e(r\zeta + 1)$ vertices. Then $G^R(n)$ has bounded expansion with high probability.

Proof. Note that if the event $\mathcal{E}_{\zeta,r}^X$ occurs, it already occurs in a subgraph of size $|X| + r\zeta|X|$. Therefore the maximal size of X that needs to be considered in order to apply Theorem 21 is

$$|X| + r\zeta|X| \leq \frac{n}{4e} \iff |X| \leq \frac{n}{4e(r\zeta + 1)}.$$

Exchanging the probability $\Pr[\mathcal{E}_{\zeta,r}^X]$ by $\Pr[\tilde{\nabla}_r(G^R(n)[X]) \geq \zeta]$ in the proof of Lemma 84 immediately shows that

$$\sum_{k=2\zeta}^{n/4e(r\zeta+1)} \binom{n}{k} \Pr[\mathcal{E}_{\zeta,r}^X] \leq \frac{1}{n^{\zeta}}$$

 $\mathcal{E}_{\zeta,r}^X$

for suitably large n . By Lemma 80, therefore the probability that any set of at most $n/4e$ vertices form the nails of a dense r -shallow minor is at most $n^{-\xi}$. Accordingly, setting $f_{\bar{v}} = \xi$ and $f_H = 1/4e$, the second condition of Theorem 21 holds with probability at least $1 - n^{-\xi}$. It is left to show that functions $f_{\text{thresh}}, f_{\text{deg}}$ exist.

Claim. *Let (f_n) be the pmfs and D the limit of (D_n) . Then every graph matching (D_n) satisfies Condition 1 of Theorem 21.*

Recall that Condition 1 states that there exist functions $f_{\text{thresh}}, f_{\text{deg}}$ such that for all ε we either have $|G| \leq f_{\text{thresh}}(\varepsilon)$ or it holds that

$$|\{v \in V(G) : \deg(v) \geq f_{\text{deg}}(\varepsilon)\}| \leq \varepsilon \cdot |G|.$$

This translates to D_n as follows: for every $\varepsilon > 0$ there exists an integer $0 \leq d \leq n - 1$ such that

$$n \sum_{k=d}^{n-1} f_n(k) \leq \varepsilon n \iff \sum_{k=d}^{n-1} f_n(k) \leq \varepsilon.$$

We apply Markov's inequality to find that

$$\sum_{k=d}^{n-1} f_n(k) = \Pr[D_n \geq d] \leq \frac{E[D_n]}{d}.$$

Since $E[D_n] \rightarrow E[D]$ and $E[D]$ is finite, the right hand side can be made small enough by choosing $f_{\text{deg}}(\varepsilon) = d = E[D_n]/\varepsilon$ and n large enough. This proves the existence of appropriate functions f_{thresh} and f_{deg} and the claim.

Hence, we conclude that Theorem 21 is applicable to $G^R(n)$ with probability at least $(1 - n^{-\xi})$ and the claim follows. \square

Having shown the supercubic coupling lemma, we proceed to the next range of distributions.

15.3.2 The cubic regime

As before, we will work towards a coupling argument which translates probabilities for events on shallow topological minors to events on subgraphs. Starting at the bottom, we begin our investigation by proving an analogue of Lemma 84 for degree distributions that are scaled by a factor of $\log^{\Theta(1)} n$.

Lemma 86. *Let (D_n) be a sparse degree distribution sequence with limit D and tail-bound $h(d) = \Theta(d^{3+o(1)})$. Then for $\xi > 9$ it holds that*

$$\Pr[\omega(G^{\text{CL}}(\log^{\Theta(1)} n D_n)) \geq \xi] = O(n^{-\xi}).$$

Proof. Let $G = G^{\text{CL}}(\log^{\Theta(1)} n D_n)$. Using Lemma 80, we can bound the probability that a clique on k vertices exists by considering the probability that k randomly chosen vertices form a clique. Taking the

union bound of all possible k , the probability of the aforementioned event is at most

$$\sum_{k=\xi}^n \binom{n}{k} \Pr[\omega(G[X_k]) \geq k],$$

where X_k is a set of k vertices chosen uniformly at random. In the following, we will talk about the vertex weights assigned in the creation of the Chung–Lu random graph G by factoring out the common $\log^{\Theta(1)} n$ -factor: a vertex v has weight $d_v \log^{\Theta(1)} n$.

An edge uv exists in G with probability $d_u d_v \log^{\Theta(1)} n / \mu n$, where again $\mu = E[D_n]$. For k vertices of weights d_1, \dots, d_k , the probability that they form a complete graph is therefore at most

$$\prod_{i < j} \frac{d_i d_j \log^{\Theta(1)} n}{\mu n} = \left(\frac{\log^{\Theta(1)} n}{\mu n} \right)^{\frac{k(k-1)}{2}} \left(\prod_i d_i \right)^{k-1}.$$

Note that the tail-bound H of $\log^{\Theta(1)} n D_n$ is

$$H(d) = h\left(\frac{d}{\log^{\Theta(1)} n}\right) = \left(\frac{d}{\log^{\Theta(1)} n}\right)^{3+o(1)},$$

hence the maximal degree is $\Delta = H^{-1}(n) = n^{1/(3+o(1))} \log^{\Theta(1)} n$.

Taking the union bound of all possible products D of weight k , we obtain the bound

$$\begin{aligned} \sum_{k=\xi}^n \binom{n}{k} \Pr[\omega(G[X_k]) \geq k] &\leq \sum_{k=\xi}^n \binom{n}{k} \left(\frac{\log^{\Theta(1)} n}{\mu n}\right)^{\frac{k(k-1)}{2}} \sum_{D=k}^{\Delta^k} D^{k-1} \\ &\leq \sum_{k=\xi}^n \frac{n^k e^k}{k^k} \left(\frac{\log^{\Theta(1)} n}{\mu n}\right)^{\frac{k(k-1)}{2}} \Delta^{k^2}. \end{aligned}$$

It is easy to check that this sum is supergeometric, hence we can estimate its value by twice its largest term. This results in

$$\begin{aligned} \sum_{k=\xi}^n \binom{n}{k} \Pr[\omega(G[X_k]) \geq k] &\leq 2 \frac{n^\xi e^\xi}{\xi^\xi} \left(\frac{\log^{\Theta(1)} n}{\mu n}\right)^{\frac{\xi(\xi-1)}{2}} \Delta^{\xi^2} \\ &= \frac{2e^\xi}{(\xi\mu)^\xi} \frac{(\log^{\Theta(1)} n)^{\frac{3}{2}\xi^2 - \frac{1}{2}\xi}}{n^{(1/2-1/(3+o(1)))\xi^2 - \frac{3}{2}\xi}}. \end{aligned}$$

Since this last term is, for $\xi > 9$ and large enough n , lies in $O(n^{-\xi})$, the claim follows. \square

\mathcal{K}_r^X

Let in the following \mathcal{K}_r^X denote the event that the vertices of X form the nails of an r -subdivision of a complete graph. Then we can rephrase the above lemma as follows:

Corollary 24. *Let $G^R(n)$ be a random graph model with the following property: for every r there exists a sparse degree distribution (D_n) with tail-bound $h(d) = \Theta(d^{3+o(1)})$ and $\xi > 9$ such that*

$$\Pr[\mathcal{K}_r^X] \leq \Pr[G^{\text{CL}}(\log^{\Theta(1)} n D_n)[X] \simeq K_{|X|}],$$

where X is a random set of at most ξ vertices. Then $G^R(n)$ is nowhere dense with high probability.

The above lemma will later provide the positive statement, namely, that large shallow clique-minors are improbable due to a coupling to the occurrence to large cliques in the above setting. However, Theorem 40 also states that the cubic degree distribution do *not* result in graphs with bounded expansion. The following lemma provides us with the necessary negative statement.

Lemma 87. *Let (D_n) be a sparse degree distribution sequence with tail-bound $h(d) = \Theta(d^{3+o(1)})$. Then*

$$\tilde{V}_{1/2}(G^{CL}(D_n)) = \Omega(\log n)$$

with high probability.

Proof. Consider the $\log n$ vertices V_h of highest weight in $G^{CL}(D_n)$. Then the minimum weight δ of these vertices can be estimated by seeing that (using again the bound from Lemma 83)

$$\begin{aligned} N &= \sum_{d=\delta}^{\Delta} \frac{n}{d^{3+o(1)}} \sim \frac{n}{(2+o(1))\delta^{2+o(1)}} \\ \implies \delta &\sim \left(\frac{n}{(2+o(1)) \log n} \right)^{1/(2+o(1))}. \end{aligned}$$

We will show that there exists a dense $1/2$ -shallow topological minor with V_h as its nails. Note that at most $\log^2 n$ vertices can participate as subdivision-vertices in this minor. By assuming that only the $n - (\log n + \log^2 n)$ vertices of lowest weight can be used as branch vertices, we can simulate the density of the minor in a simpler random graph: we connect vertices $s, t \in V_h$ independently at random with probability p such that

$$p \leq \Pr[\exists w \in V(G) \setminus V_h : sw, st \in G^{CL}(D_n)].$$

Clearly, the surrogate graph's density is (stochastically) smaller than the minor's density.

First, let us estimate the maximal available vertex weight according to our worst-case assumptions: we need a value η that satisfies

$$\begin{aligned} \frac{n}{(2+o(1))\eta^{2+o(1)}} &\geq \log n + \log^2 n \\ \implies \eta &\leq \left(\frac{n}{(2+o(1))(\log n + \log^2 n)} \right)^{1/(2+o(1))}. \end{aligned}$$

For simplicity, we choose $\eta = (n/4\log^2 n)^{1/(2+o(1))}$ which satisfies the above constraint for large enough n . Now the probability that two nails are both connected to a branch vertex is at least

$$\begin{aligned} p &:= \sum_{d=1}^{\eta} \frac{n}{d^{3+o(1)}} \cdot \frac{\delta^2 d^2}{\mu^2 n^2} = \frac{\delta^2}{n} \sum_{d=1}^{\eta} \frac{1}{d^{1+o(1)}} \\ &\sim \frac{\delta^2 \log \eta}{\mu^2 n} \geq \frac{\log n - \log(4 \log^2 n)}{10 \mu^2 n^{o(1)} \log^{1+o(1)} n}. \end{aligned}$$

For large enough n , we therefore have that p is bounded from below by a constant. Hence the expected number of edges in the minor is $\Omega(\log^2 n)$ and its density $\Omega(\log n)$. Since these surrogate edges are independent, their number and hence the lower-bound for the minor's density are sharply concentrated around the mean: the probability of falling below a factor of $(1-t)\log n$ density is only $O(n^{-t^2/2})$. We conclude that the minor has density $\Omega(\log n)$ with high probability. \square

Having fully characterised the near-cubic regime, we proceed to the final range; degree distributions with a subcubic tail.

15.3.3 The subcubic regime

As it turns out, the 'phase transition' from the nowhere-dense regime at d^3 is very sharp: a degree distribution with a tail lower-bounded by $d^{3-\varepsilon}$ for any $\varepsilon > 0$ will result in the presence of shallow dense clique minors of arbitrary size. Since we can leverage the powerful Theorem 10, the proof is quite straightforward.

Lemma 88. *Let (D_n) be a sparse degree distribution sequence with a tail lower-bounded by $h(d) = O(d^{3-\varepsilon})$ for some $\varepsilon > 0$. Then $G^{\text{CL}}(D_n)$ is somewhere dense with high probability.*

Proof. Consider the N vertices of highest degree, where N will be fixed later. Let $\gamma = 3 - \varepsilon$. Then the minimum weight δ of these vertices can be estimated by seeing that (using the same bound as in Lemma 83)

$$N = \sum_{d=\delta}^{\Delta} \frac{n}{d^\gamma} \sim \frac{n}{(\gamma-1)\delta^{\gamma-1}} \implies \delta \sim \left(\frac{n}{(\gamma-1)N} \right)^{1/(\gamma-1)}.$$

Hence the expected number of edges within these N vertices is at least

$$N^2 \frac{1}{\mu n} \left(\frac{n}{(\gamma-1)N} \right)^{2/(\gamma-1)} = N^{\frac{2(\gamma-2)}{\gamma-1}} n^{\frac{3-\gamma}{\gamma-1}} \mu^{\frac{2}{\gamma-1}} (\gamma-1)^{-\frac{2}{\gamma-1}}.$$

Our goal is to show that for a suitable choice of N depending on n , the number of expected edges is $N^{1+\varepsilon'}$ for some $\varepsilon' > 0$. This condition yields:

$$\begin{aligned} N^{\frac{2(\gamma-2)}{\gamma-1}} n^{\frac{3-\gamma}{\gamma-1}} \mu^{\frac{2}{\gamma-1}} (\gamma-1)^{-\frac{2}{\gamma-1}} &\geq N^{1+\varepsilon'} \\ \iff n^{\frac{3-\gamma}{\gamma-1}} \mu^{\frac{2}{\gamma-1}} (\gamma-1)^{-\frac{2}{\gamma-1}} &\geq N^{\frac{3-\gamma}{\gamma-1} + \varepsilon'}. \end{aligned}$$

Choosing, for example, $\varepsilon' = (3 - \gamma)/(\gamma - 1)$ we can let $N = O(\sqrt{n})$. Now by Theorem 10, this implies that for every p the graph $G^{\text{CL}}(D_n)$ contains a $c_{\varepsilon'}$ -subdivision of K_p if we find at least the expected number of edges between the $O(\sqrt{n})$ vertices of highest degree. By the Chernoff bound, this does *not* happen only with probability

$$\Pr[E_N \leq (1-t)n^{\varepsilon'/2}] \leq \exp\left(-\frac{n^{\varepsilon'/2}t^2}{2}\right) = \exp\left(-\frac{n^{\varepsilon/(4-2\varepsilon)}t^2}{2}\right),$$

hence the claim follows. \square

15.3.4 The proof of Theorem 40

We begin with the proof for the Chung–Lu model since the coupling arguments are straightforward. Afterwards, we will show how they can be adapted to extend the proof to the configuration model. The important puzzle piece for the coupling is expressed in the following lemma: the probability that a short path exists in $G^{\text{CL}}(D_n)$ crucially depends on the weight of its endpoints and the shape of the degree distribution's tail.

Lemma 89. *Let (D_n) be a sparse degree-distribution whose tail is bounded by h for degrees above τ . Let s, t be vertices in $G^{\text{CL}}(D_n)$. Then for every integer r it holds that*

$$\Pr[\exists P_{st} \subseteq G^{\text{CL}}(D_n), |P_{st}| = r \mid d_s, d_t, F] = \frac{d_s d_t}{n} O(E[D_n^2]^{r-1})$$

and this bound still holds if up to $n/2$ weights have been uncovered.

Proof. Consider the probability that a fixed path $P_{st} := s, v_1, \dots, v_{r-1}, t$ is realised in $G := G^{\text{CL}}(D_n)$ if the assigned vertex weights are, respectively, $d_s, d_1, \dots, d_{r-1}, d_t$: this probability is given by

$$\Pr[P_{st} \subseteq G \mid d_s, d_{v_1}, \dots, d_{v_{r-1}}, d_t] = \frac{d_s d_t \prod_i d_i^2}{\mu^r n^r},$$

where $\mu = E[D_n]$ is independent of n .

Since only a fraction of the vertex weights have been uncovered, instead of choosing weights without replacement, we can use \hat{D}_n from Lemma 81 to sample weights independently. Accordingly, we let the variables $\hat{D}_{1,n}, \dots, \hat{D}_{r-1,n}$ be independent copies of D_n .

Recall that (\hat{D}_n) has the same tail-bound h as (D_n) . Note now that

$$E[D_n^2] = \sum_{d=1}^{\tau-1} \Pr[D_n = d] \cdot d^2 + \sum_{d=\tau}^{\Delta} \frac{d^2}{h(d)} = \Theta(E[\hat{D}_n^2]).$$

Since the $D_{i,n}$ are independent, we can estimate the union-bound over all weights

$$\begin{aligned} \sum_{d_1, \dots, d_{r-1}} \prod_i d_i^2 \cdot \Pr[\hat{D}_{i,n} = d_i] &= \prod_{d_1, \dots, d_{r-1}} \sum_i d_i^2 \cdot \Pr[\hat{D}_{i,n} = d_i] \\ &= \prod_{d_1, \dots, d_{r-1}} E[\hat{D}_{i,n}^2] \\ &= \Theta(E[D_n^2]^{r-1}). \end{aligned}$$

We arrive at the upper bound

$$\Pr[P_{st} \subseteq G \mid d_s, d_t, F] = \frac{d_s d_t}{\mu^r n^r} \Theta(E[D_n^2]^{r-1}).$$

Since a path of length r has $r - 1$ internal vertices, the probability that *some* s - t -paths of length r exists is (implicitly using Lemma 80) then bounded by

$$\Pr[\exists P_{st} \subseteq G \mid d_s, d_t] = \frac{d_s d_t}{n} O(E[D_n^2]^{r-1}),$$

as claimed. \square

We finally have all the ingredients for the main proof.

Proof for $G^{\text{CL}}(D_n)$

Proof of Theorem 40 for $G^{\text{CL}}(D_n)$. First consider a sparse degree distribution sequence (D_n) with tail-bound $h(d) = \Omega(d^{3+\varepsilon})$ for some $\varepsilon > 0$. By Lemma 89, the probability of an s - t -path of length r existing in $G := G^{\text{CL}}(D_n)$ is

$$\begin{aligned} \Pr[\exists P_{st} \subseteq G] &\leq \Pr[st \in G^{\text{CL}}(\Theta(\sqrt{E[D_n^2]}^{r-1})D_n)] \\ &\leq \Pr[st \in G^{\text{CL}}(\Theta(D_n))] \end{aligned}$$

and this relation is still true if conditioned by the knowledge of up to $n/4e$ vertex-weights. Hence for any random set X of at most $n/4e$ vertices we have that

$$\Pr[\mathcal{E}_{\xi,r}^X]_{G^{\text{CL}}(D_n)} \leq \Pr[\nabla_0(G^{\text{CL}}(\sqrt{c_r}D_n)[X]) \geq \xi].$$

Thus by Lemma 85 and the fact that $\Theta(D_n)$ is sparse and has the same tail-bound as D_n , the model $G^{\text{CL}}(D_n)$ has bounded expansion with high probability.

Next, assume (D_n) has a tail $h(d) = \Theta(d^{3+o(1)})$. By Lemma 89, the probability of an s - t -path of length r existing in $G := G^{\text{CL}}(D_n)$ is

$$\begin{aligned} \Pr[\exists P_{st} \subseteq G] &\leq \Pr[st \in G^{\text{CL}}(\Theta(\sqrt{E[D_n^2]}^{r-1})D_n)] \\ &\leq \Pr[st \in G^{\text{CL}}(\Theta(\log^{\Theta(1)} n D_n))] \end{aligned}$$

and this relation is still true if conditioned by the knowledge of up to $n/4e$ vertex-weights. Let again \mathcal{K}_r^X denote the event that an r -subdivision of a complete subgraph with nails X exists in G . Since the graph G is sparse with high probability, we focus on the case $r \geq 1$. Now for any random set X of at most $\sqrt{n/4er}$ vertices we have that

$$\Pr[\mathcal{K}_r^X]_{G^{\text{CL}}(D_n)} \leq \Pr[G^{\text{CL}}(\log^{\Theta(1)} D_n)[X] \simeq K_{|X|}].$$

and thus by Corollary 24 it follows that $G^{\text{CL}}(D_n)$ is nowhere dense with high probability. By Lemma 87, we further have that already the measure $\tilde{\nabla}_{1/2}(G)$ grows at a rate of at least $\Omega(\log n)$, hence $G^{\text{CL}}(D_n)$ has unbounded expansion.

Finally, assume (D_n) has a tail-bound $h(d) = O(d^{3-\varepsilon})$ for some $\varepsilon > 0$. By Lemma 88 we already have that $G^{\text{CL}}(D_n)$ is somewhere dense with high probability. \square

This proof can be extended to the configuration model, the main difficulty here is that edges are not sampled independently of each other. We first prove a variant of Lemma 89. The bound proved here crucially depends on the number of *unmatched* stubs: recall that we can, instead of matching up stubs by choosing a random matching, match them up pair-by-pair. From this perspective, we can stop the process at any point and express the probabilities at this stage in terms of the remaining number of stubs.

Lemma 90. Let (D_n) be a sparse degree-distribution whose tail is bounded by h for degrees above τ . Let s, t be vertices in $G^{\text{CF}}(D_n)$. Then for every $r \in \mathbb{N}$ it holds that

$$\Pr[\exists P_{st} \subseteq G^{\text{CF}}(D_n), |P_{st}| = r \mid d_s, d_t] = \frac{d_s d_t}{m} O(E[D_n^2]^{r-1}).$$

where m is the number of unmatched stubs.

Proof. Let $G = G^{\text{CF}}(D_n)$. By $M(n) := (n-1)!!$ we denote the number of matchings on n vertices, where $!!$ denotes the double factorial:

$$n!! := \begin{cases} n \cdot (n-2) \cdot \dots \cdot 5 \cdot 3 \cdot 1 & \text{for } n > 0 \text{ odd,} \\ n \cdot (n-2) \cdot \dots \cdot 6 \cdot 4 \cdot 2 & \text{for } n > 0 \text{ even, and} \\ 1 & n \in \{0, -1\}. \end{cases}$$

We will need the following bound for $k < n$:

$$\frac{M(n-k)}{M(n)} \leq \left(\frac{(2e)^k (n-k)^{n-k}}{n^n} \right)^{1/2} \leq \left(\frac{2e}{n} \right)^{k/2}.$$

The number of available stubs decreases with each edge added to the graph and hence the probability of an edge crucially depends on the number m of *remaining* stubs.

Fix a path P_{st} of length r and let d_1, \dots, d_{r-1} denote the weights of its internal vertices. The probability of this path existing in G , conditioned on the weight of its endpoints, is bounded by

$$\begin{aligned} \Pr[P_{st} \subseteq G \mid d_s, d_t] &\leq d_s d_t \frac{M(m-2r)}{M(m)} \sum_{d_1, \dots, d_{r-1}} \prod_{i=1}^{r-1} d_i^2 \Pr[\hat{D}_n = d_i] \\ &\leq \frac{d_s d_t}{m^r} O(E[D_n^2]^{r-1}). \end{aligned}$$

Therefore the probability that *some* s - t -path of length r exists is

$$\Pr[\exists P_{st} \subseteq G \mid d_s, d_t] \leq \frac{d_s d_t}{m} O(E[D_n^2]^{r-1}),$$

as claimed. □

Proof for $G^{\text{CF}}(D_n)$

Proof of Theorem 40 for $G^{\text{CF}}(D_n)$. By Lemma 90, the probability of an s - t -path of length r existing in $G := G^{\text{CF}}(D_n)$ is

$$\Pr[\exists P_{st} \subseteq G^{\text{CF}}(D_n), |P_{st}| = r \mid d_s, d_t] = \frac{d_s d_t}{m} O(E[D_n^2]^{r-1}).$$

Note that this probability looks almost identical to the one given by Lemma 89, provided that the number of remaining stubs m is $\Theta(n)$. Since we want to estimate the probability of the event $\mathcal{E}_{r, \zeta}^X$, only up to $r\zeta \lfloor X \rfloor$ edges need to be considered at once; meaning that at least

$$\hat{m} - 2r\zeta \lfloor X \rfloor \geq 2\mu n - 2r\zeta n / 4e(r\zeta + 1) \geq (\mu - 1)2n$$

stubs remain (again we let $\mu = E[D_n]$). By coupling we see that

$$\Pr[\mathcal{E}_{r,\xi}^X]_{G^{\text{CF}}(D_n)} \leq \Pr[\mathcal{E}_{r,\xi}^X]_{G^{\text{CL}}(\Theta(D_n))}$$

and we conclude that $G^{\text{CF}}(D_n)$ has bounded expansion for distributions with tail-bound $\Omega(d^{3+\epsilon})$. Similarly, the event \mathcal{K}_r^X for any set of vertices $|X| \leq \sqrt{n/4er}$ concerns at most $n/4e$ edges and hence the number of stubs left is $m = \Theta(n)$. Thus

$$\Pr[\mathcal{K}_r^X]_{G^{\text{CF}}(D_n)} \leq \Pr[\mathcal{K}_r^X]_{G^{\text{CL}}(\Theta(D_n))}$$

and therefore $G^{\text{CF}}(D_n)$ is nowhere dense for distributions with a tail that is in $\Theta(d^{3+o(1)})$. The lower bounds provided by Lemma 87 and Lemma 88 can be easily adapted in a similar way to apply to the configuration model. \square

15.4 RANDOM GRAPHS AND PERTURBATIONS

An interesting application of the well-understood Erdős-Rényi random graphs is as a *perturbation* of some n -vertex base graph G^* . We will use the notation $G = G^* \cup G(n, \mu/n)$ to denote the graph obtained from G^* by adding in every possible edge not already contained in G^* independently with probability μ/n . Obviously, the graph G^* itself might be a random graph itself.

Uniform perturbation can be seen as the baseline for more complicated models, like the small-world model by Kleinberg (described below), models used in percolation theory (e.g. [20]) and the hybrid model by Chung and Lu [48] (described above). The central question is: what graph classes are still structurally sparse after the addition of few random edges?

Theorem 41. *Let \mathcal{G} be a class of bounded-degree graphs and μ a constant. Then the composite model $G^{\cup}(\mathcal{G}) + G(n, \mu/n)$ has bounded expansion with high probability.*

Note that this theorem in particular applies to $G(n, \mu/n)$ itself. The result carries over to the *stochastic block model*, if the parameters involved are small enough. This model was first studied in mathematical sociology by Holland, Laskey, and Leinhardt in 1983 [140] and extended by Wang and Wong to directed graphs [245]. We will supplement the above result by demonstrating that there exist very sparse graph classes of unbounded degree for which such a perturbation results in dense clique minors.

The following technical lemma subsumes Theorem 41.

Lemma 91. *Let \mathcal{G} be a class of graphs with the following properties:*

- \mathcal{G} has bounded expansion, and
- for $G \in \mathcal{G}$ and every $r \in \mathbb{N}$ the distribution of $|N^r|$ has a tail-bound h with $h(d) = \Omega(d^{3+\epsilon})$ for some $\epsilon > 0$.

Then $\mathcal{G} + G(n, \mu/n)$ has bounded expansion with high probability.

Proof. Let $G_\nabla \in \mathcal{G}$, $\tilde{G} = G(n, \mu/n)$ and let $G = G_\nabla + \tilde{G}$. Assume H is an r -shallow topological minor of G and consider an embedding ϕ_V, ϕ_E of H witnessing this fact. Since $\tilde{\nabla}_r(G_\nabla)$ is a constant, most of H 's density must depend on random edges, *i.e.*

$$|\{e \in H \mid \phi_E(e) \cap E(\tilde{G}) = \emptyset\}| \leq \tilde{\nabla}_r(G_\nabla)|H|.$$

Therefore it suffices to bound the density of topological minors whose embedding use at least one edge of \tilde{G} for each edge of the minor. Consider a path P of length r in G that uses at least one edge of \tilde{G} : each component in $K(P \setminus E(\tilde{G}))$ is contained in a subgraph $G_\nabla[N^r(v)]$ for some vertex v . Let N_1, N_2, \dots, N_p be these subgraphs of the path P : then we can bound the probability that P exists by considering the probability that there exist at least one edge between N_i and N_{i+1} in \tilde{G} , for $1 \leq i \leq p-1$.

Define the random variable $D_{r,G}$ with distribution given by

$$\Pr[D_{r,G} = d] = \frac{|\{v \in G : |N^r(v)| = d\}|}{|G|}.$$

Since the probability that two r -neighbourhoods $N^r(u), N^r(v)$ in G_∇ are connected by an edge in \tilde{G} is at most

$$\frac{\mu|N^r(u)||N^r(v)|}{n}$$

we can couple the occurrence of r -paths in G to the occurrence of edges in $G^{\text{CL}}(D_{r,G})$. Hence we have that

$$\tilde{\nabla}_r(G) - \tilde{\nabla}_r G_\nabla \leq \tilde{\nabla}_r(G^{\text{CL}}(D_{r,G}))$$

in the stochastic sense. Since the latter has bounded expansion with high probability by Theorem 40, we conclude that $\mathcal{G} + G(n, \mu/n)$ does as well. \square

The same proof can be applied to nowhere dense classes by replacing $\tilde{\nabla}_r$ with $\tilde{\omega}_r$ and a slightly different condition on the 'neighbourhood-statistic' $|N^r|$.

Lemma 92. *Let \mathcal{G} be a class of graphs with the following properties:*

- \mathcal{G} is nowhere dense, and
- for $G \in \mathcal{G}$ and every $r \in \mathbb{N}$ the distribution of $|N^r|$ has a tail $h(d) = \Omega(d^{3+o(1)})$.

Then $\mathcal{G} + G(n, \mu/n)$ is nowhere dense with high probability.

An obstruction

The above poses the question: are there structurally sparse classes which do *not* stay sparse under perturbation? The answer is yes: consider the class of graphs consisting of $\Theta(\sqrt{n})$ copies of $S_{\sqrt{n}}$. The probability that two such stars will be connected by a randomly added

edge is lower-bounded by some constant, hence the minor obtain by contracting the former stars has density $\Theta(n)$ (while only having \sqrt{n} vertices). Hence, the perturbed class is actually somewhere dense with high probability.

This example can be easily generalised: the presence of n^α vertices X to which we can assign at least n^β of their respective r -neighbours (not assigning any neighbour to more than one vertex of X), for some constant r , will yield an r -shallow minor whose density is concentrated around $n^{3\alpha+\beta-1}$. Hence for all α, β that satisfy $3\alpha + \beta > 2$, the perturbed class is somewhere dense with high probability.

Open question 13. The ‘Bernoulli-noise’ we used above is only the most simple case of a graph perturbation. The noise can easily be defined to depend on the graph to which it is added—in fact, the Kleinberg model discussed below could be seen as using a distance-dependant perturbation to a grid. Which structurally sparse classes retain their sparseness under such different notions of perturbations?

15.5 EVOLUTIONARY AND ATTACHMENT MODELS

The side of network model research treated in the previous chapter has focused on how to extend random graphs—which come with a rich mathematical foundation to stand on—in order to turn them into descriptive models for complex networks. In essence, the random graph models are simply tweaked until their output looks more ‘network-like’. An orthogonal approach popular with researchers coming from a statistical mechanics background are models that try to emulate the process of network formation, the *generative* models. These models are usually much harder to analyse theoretically and often our knowledge of them rests on empirical evidence.

Barabási-Albert

One of the most popular papers by Barabási and Albert [18] that undoubtedly helped catalyse the steady growth of the network science field put forward *preferential attachment* as a network formation process that results in a power law degree-distribution. The claims of universality made in the paper are controversial today. In particular the widely-cited power law distribution of ‘the internet’ and its alleged consequences for attack-resilience and susceptibility to epidemics seem to be founded on faulty data—see the rather damning article by Willinger, Anderson, and Doyle [250]. Nonetheless, the appealingly simple model (which essentially can already be found in Yule’s model of evolutionary processes from 1925 [251]) helped to popularise network science as a field. And at least in some domains, the proposed mechanism seems a plausible approximation for network formation.

To generate a graph using the Barabási-Albert model, one fixes a constant initial graph G_0 and an integer k . The graph G_i is constructed from G_{i-1} by attaching a new vertex v_i and connected it to k vertices

of the graph G_{i-1} , where the probability that a vertex is chosen is proportional to its degree in G_{i-1} .

Barabási and Albert suggested that such a network asymptotically displays a power law degree distribution, *i.e.* the fraction of nodes of degree d is proportional to $d^{-\gamma}$. They observed experimentally that $\gamma = 2.9 \pm 0.1$ and suggested that γ is actually 3. This model was rigorously analysed by Bollobás, Riordan, Spencer, and Tusnády [30] who proved that $\gamma = 3$. Later Bollobás and Riordan showed [28] that the diameter of the graphs generated by this model is asymptotically $\log n / \log \log n$. The same authors later showed that the clustering coefficient is asymptotically $\frac{k-1}{8} \frac{\log^2 n}{n}$ [29]. This value differs greatly from the often-reported experimental coefficient of $n^{-3/4}$. In any case, the clustering coefficient vanishes for large n .

The graphs generated by the model are frequently called *scale-free* graphs in the literature. This terminology is problematic since graphs exhibiting a power law degree distribution can be generated any number of ways, of which the preferential-attachment kind might only be a small subset.

Forest fire

A variant of preferential attachment that achieves the ‘rich-get-richer’ effect without explicitly modelling it is the so-called *forest fire* model by Leskovec, Kleinberg, and Faloutsos [171]. The attachment process is governed by two parameters p, r and works as follows. Given the directed graph G_i , the directed graph G_{i+1} is constructed by adding a new vertex v and attaching it to G_i as follows:

1. We choose an *ambassador* vertex w uniformly at random from G_i and add the arc vw .
2. We select k_o out-neighbours and k_i in-neighbours of w , where k_o is randomly selected with geometric distribution around the mean $p/(1-p)$ and k_i around the mean $rp/(1-rp)$. If k_i, k_o are larger than the available number of neighbours we select as many as possible.
3. For all vertices selected in the previous step, we apply the above process recursively using that node as the ambassador (in particular, we add an arc from v to the node). Nodes that were visited by the process earlier are not re-visited.

The model generates, for suitable choices of p, r , graphs that densify according to a power-law $m = \Theta(n^{1+\alpha})$, with $\alpha > 0$. The densification does crucially depend on the choice of p, r and in other regimes we obtain sparse graphs.

Despite the issue of whether densification can really be observed in networks, the model constitutes an interesting approach: the ‘burning’ process provides a natural explanation of the rich-get-richer effect (since high-degree nodes are likely to be discovered via their large neighbourhood). Furthermore, other statistics of the model—like the occurrence of a community-like structure and small diameter—are claimed to reflect those of real-world networks.

15.5.1 *The aberration of attachment*

We would like to see where attachment models fall in the sparseness-dichotomy. However, the following lemma shows that there cannot be a satisfying answer in the current framework:

Lemma 93. *Let (G_i) be a sequence of graphs generated by an attachment process that*

- *attaches a new vertex by at least two edges, and*
- *has a non-zero attachment probability for all vertices.*

Then for every $t \geq 1$, G_n contains a one-subdivision of K_t with probability at least $f(t) > 0$, for some function f that depends on the model.

Proof. Fix some t and consider the graph G_{t+1} . By our assumption on the model, the vertex v_{t+1} added in this step has a non-zero probability to select the vertices v_1, v_2 as its neighbours. Similarly, the vertex v_{t+2} in G_{t+2} has a non-zero probability to select v_1 and v_3 . Hence the probability that the vertices v_{t+1} to $v_{\binom{t}{2}}$ all connect to a unique pair of the first t vertices and thus generate a graph containing K_t is some function of t and as a result bounded away from zero. \square

An immediate consequence of the above lemma is that every attachment model discussed is somewhere dense with non-vanishing probability. This limbo-state is rather unsatisfactory and we cannot reasonably conclude that the graphs generated by such models are structurally dense or sparse. We can only resolve this situation by either changing the models (while preserving their intended properties) or our definition of how we determine the sparseness of random graphs.

Open question 14. Is there a useful relaxation of what it means for a random graph model to be structurally sparse that can be applied to attachment models?

15.6 SMALL WORLD MODELS

Watts-Strogatz model

Two notable models have arisen in researching the small world phenomenon. Watts and Strogatz proposed the following model [247] with parameters $n, k > \log n$ and a probability p : We start with a cycle C_n and connect every vertex on the ring to all vertices in its k^{th} neighbourhood. The resulting graph is equivalent to $(C_n)^{\log k/2}$. Then for every vertex v , we *rewire* each of its edges with probability p , that is, we chose a different endpoint uniformly at random from all $n - 1$ other vertices.

Their original description leaves out some details of this process, for example it is not clear whether an edge can be rewired more than once—however, it seems plausible that these details do not influence

the outcome too much. A more rigorous analysis for a slightly different model was later supplied by Newman and Watts [197] and it confirmed the original findings. Depending on the rewiring probability p , this model generate graphs with average clustering coefficient $C(p)$ and average diameter $D(p)$ with $C(0) \sim 3/4$ and $D(0) \sim n/2k$ on the one end of the spectrum, and $C(1) \sim k/n$ and $D(1) \sim \log(n - k)$ on the other end⁴. Hence one can trade off, via the parameter p , a small diameter against a small clustering and vice-versa.

Kleinberg model

A refinement of the above small-world model was proposed by Kleinberg [158] in his work on small-world *routing*: while the small-world property of networks explains why an experiment like Milgram's can succeed, it falls short of explaining it entirely. We are missing the *algorithmic* nature of the experiment! Every participant of the experiment has only local knowledge, and still, the routing strategy of selecting the neighbour geographically (or socially) closest to the target succeeded with very few steps.

Greedy routing

Kleinberg's model is designed to approximate the setting of the experiment and explain not only why the network has a small diameter, but also why the *greedy routing* strategy is successful in finding short paths. The model itself has parameter $n, p, q \in \mathbb{N}$ and $\gamma \in \mathbb{R}$ from which a graph is generated as follows: starting out with $n \times n$ vertices arranged in the plane, where the vertex $v_{i,j}$ has coordinates (i, j) , we connect every vertex to all neighbouring vertices within distance p , using the Manhattan-distance

$$\text{dist}_M(v_{x_1, y_1}, v_{x_2, y_2}) = \|(x_1, y_1) - (x_2, y_2)\|_1.$$

Next, we add to every vertex u a total of q outgoing *long-range arcs*, where a vertex v is chosen as the endpoint of the arc with probability proportional to $\text{dist}_M(u, v)^{-\gamma}$.

The parameters p, q have relatively little effect on the model. The parameter r , however, is crucial for the small-world property and the success of greedy routing. When $\gamma = 0$, the long-range contacts are uniformly distributed throughout the grid, and one can show that there exist paths between every pair of nodes of length bounded by $\log^{O(1)} n$. Kleinberg showed that in the range $0 \leq \gamma < 2$, the expected delivery time of *every* decentralised algorithm (one that uses only local information) is $\Omega(n^{2/3})$. When $\gamma = 2$, then short paths exist *and* there is a decentralised algorithm to transmit a message that takes $O(\log^2 n)$ time in expectation between any two randomly chosen points. For a value $\gamma > 2$, the delivery time is $\Omega(n^{(\gamma-2)/(\gamma-1)})$.

The intuitive explanation given by Kleinberg for why the number two has this special property is that since the number of nodes within distance d grows quadratic with d , the distribution of the *length* of long-range edges is uniform. Therefore at every step of the routing

⁴ To be precise, Newman and Watts determined the expected value of these statistics and claim that it approximates the average value well.

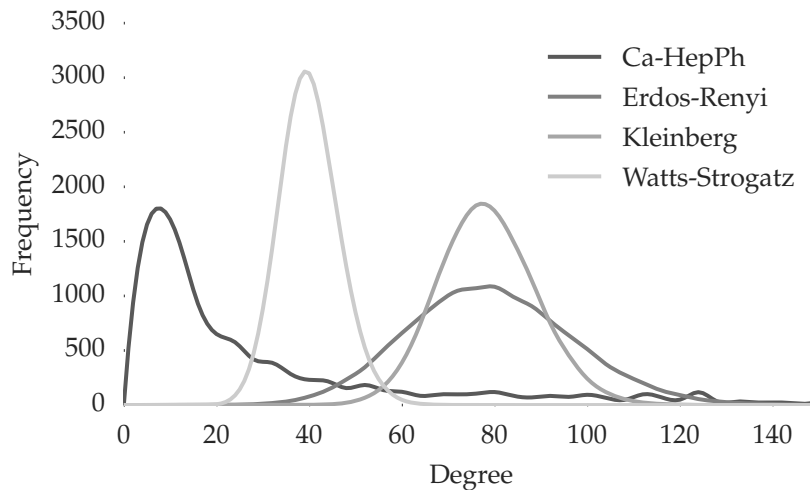


Figure 5: Degree distribution of several small-world graph models compared to that of a collaboration network of High Energy Physics [171]. The model's parameter were chosen to reflect the size and density of the network and averaged over ten iterations.

process, the probability that one can use a long-range arc to make progress is high.

Both the Watts-Strogatz and the Kleinberg model were created to explain specific statistics of networks. For this reason, they are not 'general-purpose' models that can be used domain-independent. For example, the degree-distribution of both is atypical (*cf.* Figure 5). In case of Watts-Strogatz, the useful range of parameters furthermore generates rather dense graphs with $\Theta(n \log n)$ edges.

Extensions

The question of efficient routing has subsequently been researched in several variations of Kleinberg's model. Franceschetti and Meester extended the result to continuous objects, again confirming that the parameter $\gamma = 2$ is the precise point at which greedy routing is possible [108]. Draief and Ganesh proved very similar results for graphs generated through a Poisson process in the plane [75]. Coppersmith, Gamarnik, and Sviridenko [50] considered a d -dimensional lattice and showed that the diameter of the resulting graph is $\Theta(\log n / \log \log n)$ for $\gamma = d$ and at least polynomial for $\gamma > 2d$. In the range $d < \gamma < 2d$ the diameter e is $\log^{\Theta(1)n}$. The routing properties of this extension seem to be unresolved. Notable is also the work by Fraigniaud who showed that greedy routing schemes exist in graphs of small treewidth and treelength⁵ [107].

⁵ Although he formulated it via the *chordality* of the graph which in particular bounds its treelength.

15.7 KLEINBERG'S MODEL: DENSE WHEN USEFUL

Let us continue by considering Kleinberg's model and the relation of γ to the structural sparseness. Recall that γ governs the probability of long-range edges in the second phase of the model and that $\gamma = 2$ is the 'sweet-spot' for greedy routing. From the construction it is clear that the model is sparse: the total number of edges is $(p + q)n^2$ (keep in mind that the number of vertices is n^2 , not n). And the superficial similarity of the model to a perturbation of a grid graphs, which by Theorem 41 has bounded expansion, might tempt us to conclude that Kleinberg's model should be structurally sparse. The answer is, as so often, a little bit more complicated.

The parameters p and q of the model do not influence the asymptotic density, hence we will usually assume them to be one. We denote by Γ_n^p the $n \times n$ grid embedded on a torus such that a vertex is connect to all vertices within distance p . For the second phase of the model, let $\Lambda_n^{\gamma,q}$ be the $n \times n$ random graph (with the same embedding as Γ_n^p) with the long-range arcs added. The final graph then is $G^{\text{KL}}(n, p, q, \gamma) = \Gamma_n^p + \Lambda_n^{\gamma,q}$.

Normalisation

We begin by estimating the normalisation constant c_γ : recall that the probability that a vertex u chooses a vertex v as its long-range neighbour is

$$\Pr[uv \in \Lambda_n^{\gamma,q}] = \frac{1}{c_\gamma \cdot \text{dist}_M(u, v)^\gamma},$$

where $\text{dist}(u, v)_M$ denote the Manhattan distance between u and v in the embedding. For the vertex u , the normalisation will be

$$c_\gamma(u) = \sum_{u \neq v \in G} \frac{1}{\text{dist}_M(u, v)^\gamma}.$$

Since we used a torus for our embedding, the normalisation constant is the same for every vertex (which is not true for embeddings in the plane) and we can compute it alternatively as

$$c_\gamma = \sum_{d=1}^{n-1} \frac{f(d)}{d^\gamma},$$

where $f(d)$ counts the number of vertices at distance exactly d from any vertex of Γ_n^p . It is easy to verify that

$$f(d) = \begin{cases} \Theta(d) & \text{for } d \leq n/2, \\ \Theta(n - d) & \text{for } d \geq n/2. \end{cases}$$

We obtain that

$$c_\gamma = \sum_{d=1}^{n/2-1} \frac{\Theta(d)}{d^\gamma} + \sum_{d=n/2}^{n-1} \frac{\Theta(n - d)}{d^\gamma}.$$

Accordingly, the normalisation is

$$c_\gamma = \begin{cases} \Theta(n^{2-\gamma}) & \text{for } 1 \leq \gamma < 2, \\ \Theta(\log n) & \text{for } \gamma = 2, \text{ and} \\ \Theta(1) & \text{for } \gamma > 2. \end{cases}$$

Let us first investigate the 'useful' regime of $\gamma = 2$. As it turns out, the model contains clique-minors of arbitrary size with almost certainty!

Lemma 94. *For any $p, q \geq 1$ the model $G^{\text{KL}}(n, p, q, 2)$ is somewhere dense with high probability.*

Proof. We prove the statement for $q = 1$, clearly the cases for $q > 1$ follow from that. Consider a $2 \log \log n \times 2 \log \log n$ subgrid of Γ_n^p . Fix any set V_K of $\log \log n$ vertices and $\log \log^2 n$ edges E_K whose endpoints are mutually disjoint inside this subgrid, along with a mapping $\phi: E_K \rightarrow V_K \times V_K$. We want to find a lower-bound for the probability that the vertices in an edge $st \in E_K$ select endpoints $(u, v) = \phi(st)$ for their long-range arc—in which case we have a two-subdivision of $K_{\log \log n}$ in the graph.

For a single pair of u, v vertices inside the subgrid, note that u chooses v as its long-range neighbour with probability

$$\Theta\left(\frac{1}{\log n \text{dist}_M(u, v)^2}\right) \geq \Theta\left(\frac{1}{\log n}\right).$$

and therefore an edge $st \in E_K$ chooses the 'correct' long-range neighbours $\phi(st)$ with probability at least $\Omega(\log^{-2} n)$. The probability that *all* edges E_K choose the neighbours a prescribed by ϕ is therefore at least $\Omega(\log^{-2 \log \log^2 n} n)$.

We can partition the $n \times n$ grid into $n^2/4 \log \log^2 n$ subgrids of size $2 \log \log n \times 2 \log \log n$. Hence the probability that in *none* of these subgrids the above event occurs is at most

$$\left(1 - \Omega(\log^{-2 \log \log^2 n} n)\right)^{\frac{n^2}{4 \log \log^2 n}} = \exp\left(-\Omega\left(\frac{n^2}{\log^{2 \log \log^2 n} n}\right)\right).$$

We conclude that $G^{\text{KL}}(n, p, 1, 2)$ and thus $G^{\text{KL}}(n, p, q, 2)$ for any $q \geq 2$ contains a two-subdivision of $K_{\log \log n}$ with high probability. \square

The above proof can be easily extended to higher dimensional lattices.

What about the other ranges of γ ? For the regime $\gamma > 2$ we run into the same problem as for the attachment models: the probability that $\Theta(t) \times \Theta(t)$ subgrid contains K_t as an r -shallow minor is only a function of t and r :

Observation 6. *For any $p, q \geq 1$ and $\gamma > 2$ the model $G^{\text{KL}}(n, p, q, \gamma)$ contains a one-subdivision of K_t with probability at least $f(t) > 0$, for some function f .*

As with the attachment models, the research of the regime $\gamma > 2$ therefore needs to take place under a different framework than asymptotic convergence to zero or one. For $\gamma = 1$ the model degenerates to a perturbation of the grid Γ_n^p , which by Theorem 41 has bounded expansion with high probability.

Open question 15. Is the Kleinberg-model structurally sparse for the parametric range $1 < \gamma < 2$?

15.8 OTHER MODELS

The network literature is full of models and keeping up with them is often hopeless. The above selection of models constitutes a solid baseline on which we can build further results. We mention some more well-known models here and try to relate them to the above results.

Random intersection graphs

Karónski and Singer-Cohen [228, 151] introduced *random intersection graphs*. They are formed by drawing a random bipartite graph between n objects and m attributes and projecting the attribute-connections into the objects (*i.e.* two objects are connected in the final graph if they share a common attribute). Among other uses, the model has applications in network epidemiology [36, 17] and cybersecurity [21].

The model has a tunable parameter α which determines the number of attributes relative to the number of objects as well as the edge probability p in the bipartite graph: we have that $m = \Theta(n^\alpha)$ and $p = \Theta(n^{-(\alpha+1)/2})$ (the hidden constants in the Landau notation are further parameters not important for this discussion).

The graph model is generally denoted by $G(n, m, p)$ and we obtained the following result with respect to its structural sparseness:

Theorem 42 ([89]). Fix a constant $\alpha > 0$. Let $m = \Theta(n^\alpha)$ and $p = \Theta(n^{-(\alpha+1)/2})$. Then, with high probability, it holds that the model $G(n, m, p)$

- is somewhere dense and has degeneracy $\Omega(n^{(1-\alpha)/2})$ for $\alpha < 1$,
- is somewhere dense and has degeneracy $\Omega(\log n / \log \log n)$ for $\alpha = 1$,
- and it has bounded expansion for $\alpha > 1$.

That random intersection graphs are ultimately sparse is not too surprising: a result by Fill, Scheinerman, Singer-Cohen shows that for $\alpha > 6$, the model is asymptotically equivalent to sparse Erdős-Rényi graphs [95]. They conjecture that this could already be true for $\alpha \geq 3$, based on similarities of several phase-transitions in both models.

Geometric random graphs

At least in some domains, like protein interaction networks [209], *geometric random graphs* seem to be a viable baseline model. Such a graph is generated using a *Poisson point process*: in a fixed metric space, usually $[0, 1]^2$, vertices are randomly placed and connected to all neighbours within some fixed distance r . Such graphs have a degree distribution that converges towards a Poisson distribution. Appel and Russo showed [11] that in such a graph, the probability that any two

vertices are connected is at most $(2r - r^2)^2$. Therefore, for small radii (around $r = n^{-4}$) we know that the resulting graphs are structurally sparse, since their density can be stochastically bounded by the density of a sparse Erdős-Rényi graph. For large r on the other hand, these graphs are essentially cliques. It is an interesting question at what point the phase transition of its structural density lies and how this point compares to the parameters used in network modelling.

RMAT, Kronecker networks

There is another line of network models based on matrix products that has received significant attention. Based on the earlier recursive matrix model (RMAT [41]), Leskovec and Faloutsos devised *Stochastic Kronecker graphs* as fast, parallelisable descriptive network models⁶. In a nutshell, the model generates a graph by iteratively adding edges to it, where the selection of the edges' endpoints follows a stochastic process that implicitly uses Kronecker-product matrix derived from a small (usually two-by-two) seed matrix. Stochastic Kronecker networks are not only widely used as a baseline network model that can be fitted to the statistics of a given network, they are also part of the Graph 500 supercomputer benchmark⁷. The model is, however, not without critics. Seshadhri, Pinar, and Kolda empirically and theoretically established [224] that the degree distribution of graphs generated using this process contains some very large oscillations not found in real-world networks. The same set of authors also demonstrated earlier that graphs generated via the Chung-Lu model behave almost identical to Stochastic Kronecker graphs [223]—without the expensive fitting procedure needed to estimate the latter model's parameters. Moreover, they prove that if the parameters obey certain constraints, the two models are simply equivalent.

Since these matrix-based models are quite hard to analyse, we cannot consider them in the scope of this thesis. If the recently noted similarity to the Chung-Lu model is any indication, we should expect these graphs to follow a very similar dynamic as the characterisation provided by Theorem 40: if the tail-bounds of the resulting degree distributions are sufficiently strong, we should expect these graphs to be structurally sparse.

Open question 16. For which range of the parameter r are geometric random graphs structurally sparse and for which structurally dense?

Open question 17. Assume we are given two random graph models G_1^R, G_2^R with respective parametrisations ρ_1, ρ_2 , and assume that both models have bounded expansion a.a.s. Under what circumstances does the joint model

$$G_1^R(n, \rho_1(n)) \cup G_2^R(n, \rho_2(n))$$

again have bounded expansion?

⁶ The classification of *descriptive* and *generative* is a little blurry here since one could potentially argue that *some* kind of generative process is modelled.

⁷ See <http://www.graph500.org/>.

Determining the grad of a given graph exactly turns out to be rather difficult: enumerating all possible minor models is certainly infeasible, even for $1/2$ -shallow topological minors, and so far no better algorithm exists. While finding fixed graphs as minors or topological minors is possible in polynomial time, in the context of structural sparseness we are not interested in a *fixed* shallow minors; we are interested in their *density*. In particular, we do not know whether the densest shallow minor will be tiny or span the whole host graph! Therefore, we need a good proxy to estimate the structural sparseness (or lack thereof); that is, we need one of the parametrised graph measures introduced in Chapter 3 that can be used to define graphs of bounded expansion.

Classes vs. instances

The second problem is that we are working with single instances here; and all ‘flexible’ notions of sparseness (excluding a minor, having bounded degree, bounded expansion, *etc.*) are difficult to apply. We will begin by presenting a possible solution to this problem: a comparison of a graph’s structural properties to random baseline model from which we know that it (asymptotically) generates sparse or dense instances. We introduce a special type of plot that lets us reason, and sometimes conclude, whether a network should be considered structurally sparse or dense. On the way, we will introduce the different data sets used in our experiments.

16.1 AUGMENTATION INDEGREES AND AUG-AUG PLOTS

As demonstrated in Chapter 7, the maximum indegree of dtf-augmentations constitutes an alternative definition for classes with bounded expansion. Moreover, the computation of these augmentations is relatively fast and feasible for even large networks. To normalise the resulting data, we compare it to graphs generated with the configuration model (alternatively the Chung–Lu model) replicating the degree distribution of the target network: With Theorem 40 as a solid theoretical backdrop, we know whether or not to expect structurally sparse graphs depending on the tail of the degree distribution. This shifts the problem into a well-studied area of statistical analysis and we can build on previous work (and implementations!) to come to a conclusion. The detailed procedure looks as follows. Choose a precision parameter $N \in \mathbb{N}$ and execute these four steps:

Aug-aug plots

1. Generate dtf-augmentations $\vec{G}_1, \vec{G}_2, \dots$ of the network until no further arc can be added.

2. Generate dtf-augmentations $\vec{H}_1^i, \vec{H}_2^i, \dots$ of a random graph $H^i = G^{\text{CF}}$ replicating the degree distribution of G for $1 \leq i \leq N$. Again, we augment until no further arc can be added.
3. Compute the sequence $s_t := \Delta^-(G_t) / M[(\Delta^-(H_t^i))_{1 \leq i \leq N}]$, i.e. divide the indegrees of the network's augmentation by the median of the random graphs' indegrees.
4. Normalise the sequence $(s_t)_{t \in \mathbb{N}}$ such that $\max_t s_t = 1$.

Note that by Corollary 4, computing dtf-augmentations until no further arc is added is enough: if $\vec{G}_{d+1} = \vec{G}_d$, then necessarily $\vec{G}_{\geq d} = \vec{G}_d$.

With the above measurement, we then create the *aug-aug plot* by plotting (s_t) . To compare the densities of the network and the random baseline, we compare the plot to the 45°-line. If (s_t) lies above it, we conclude that the network has higher structural density than a random graph with the same degree sequence. If (s_t) lies below, it has less. The decision of whether G is structurally sparse or not then depends on whether the random graph is structurally sparse. This decision, in turn, can be made using Theorem 40 and known techniques: if the tail of the degree distribution of G is best described by cubic or supercubic function, the random graphs generated from it are structurally sparse. If the tail is subcubic, they are structurally dense.

Estimating tails

Let us begin at the last step: the shape of the degree distributions. Luckily, the theoretical work by Clauset, Shalizi, and Newman [49] has subsequently been turned into a user-friendly python package by Alstott, Bullmore, and Plenz [7]. For a given data set it provides fitting for all distributions in Table 1 (except Gaussian) and enables us to pairwise compare likelihood-ratios for these fittings. The fact that we are interesting in the shape of the tail gives us some leeway; we do not have to decide which distribution shape matches best but only in which regime (subcubic, cubic, supercubic) the tail of the distribution resides. Note that for all distributions except power laws, all higher moments exists and therefore, by Chebyshev's inequality for higher moments, we have arbitrarily strong polynomial tail-bounds.

We devise the following statistical test: given the likelihood ratios $R(D_1, D_2)$ for all distribution-pairs D_1, D_2 , we sum up the positive values for all distributions. For example, if $R(D_1, D_2) = 0.5$, we add 0.5 to the total weight of the distribution D_1 . In the case $R(D_1, D_2) = -0.3$, we would add 0.3 to the total weight of D_2 . If the total weights then favour exponential or stretched exponential distributions we can immediately conclude that the degree distribution is best described to have a superpolynomial—and therefore supercubic—tail. We then expect by Theorem 40 that graphs generated from this distribution are structurally sparse. If the total weights favour truncated power laws or lognormal distributions we need to consider the fitted parameters: if the truncated power law distribution has a reasonable cutoff-parameter we can still conclude that the distribution has cubic or supercubic tail-bounds. We can conclude the same if the lognormal fit is

not degenerate (*i.e.* it has reasonable parameters μ, σ). In all other cases we consider the fitted power-law exponent and decided on whether it lies below three, is exactly three or lies above three.

In the following sections, we review our findings for our corpus of real-world networks using the above methodology. Additionally, we give references and short descriptions of the networks. For better readability, we have partitioned the networks into four sets depending on their respective number of vertices. The small networks contain up to 115 vertices, the medium-sized ones between 235 and 839, the large networks between 1419 and 4941, and the huge networks between 6474 and 58228. The corpus contains a mixture of domains, ranging from neural networks (Drosophila, C.Elegans) over protein-protein interaction networks (Yeast) to social (Twittercrawl, Bergen) and co-authorship networks (HepTh, CondMat).

Calibrating the tail estimation

Before we begin to assess real-world data, let us test our methodology in order to get a feeling for the margin of error we have to expect. We know that the Barabási-Albert preferential attachment model generates graphs whose degree distribution follow a power law with exponent $\gamma = 3$. Hence, our statistical test should be able to clearly identify the power law in these graphs. In the following table, we summarised the results for graphs generated according to the Barabási-Albert model, using a parameter of $d = 5$ (each newly added vertex attaches to five vertices in the already existing graph).

n	Power law	Truncated power law	Lognormal	Stretched exponential	Exponential	γ	λ
100	6.22	6.44	4.92	3.63	0.00	3.10	$3 \cdot 10^{-4}$
500	3.76	5.38	3.70	3.29	0.00	3.05	$4 \cdot 10^{-3}$
1000	7.66	7.87	11.72	5.57	0.00	2.98	$3 \cdot 10^{-3}$
5000	9.11	9.11	11.39	6.72	0.00	3.05	$6 \cdot 10^{-7}$
10000	10.58	10.51	10.49	7.26	0.00	3.09	$3 \cdot 10^{-7}$

The column γ denotes the proposed pure power law coefficient, the column λ the proposed exponent for the truncated power law. Note that the power law coefficient for the truncated power law usually differs from γ ; but we will not need it in the following analysis.

We see that the identified power law coefficient is in all cases close to the theoretical value of 3 and that the test identifies the distribution with highest confidence in power law, truncated power law and lognormal. The difficulty to distinguish between a lognormal distribution and a power law is well-known. Accordingly, we should favour the hypothesis of a subcubic tail over that of a supercubic tail in the following cases:

1. The power law distribution clearly attains the highest score and the power law-coefficient is γ is below 3.2;

2. or the highest scores are attained by the distributions power law, truncated power law and lognormal, the proposed power law coefficient is below 3.2 *and* the exponential cutoff λ for the truncated power law is on the order of 10^{-3} or smaller.

We chose the threshold 3.2 to err on the side of caution, informed by the above calibration with the Barabási-Albert model. In the following tables we will, for reasons of space, not display the truncation-parameter λ for every network, but mention it in the subsequent text if the need arises.

Calibrating aug-aug-plots

Now in order to verify that aug-aug plots are a useful tool to assess structural sparseness, we need to verify that they can reliably detect dense substructures—that this is the case in theory is clear, but in practice the signal-to-noise ratio might be too low. To this end, we generate the complete graph K_{50} and subdivide it r times, for $r \in \{0, 1, 5, 10\}$. We expect to see a sudden jump in the plot at the augmentation \vec{G}_{r+1} ; the point where the augmentations ‘discover’ the dense substructure hidden in the graph. Note that the corresponding random graphs are relatively sparse since the average vertex degree is quite low (except for $r = 0$). We expect to see the significant difference in structural density to be reflected in the plot.

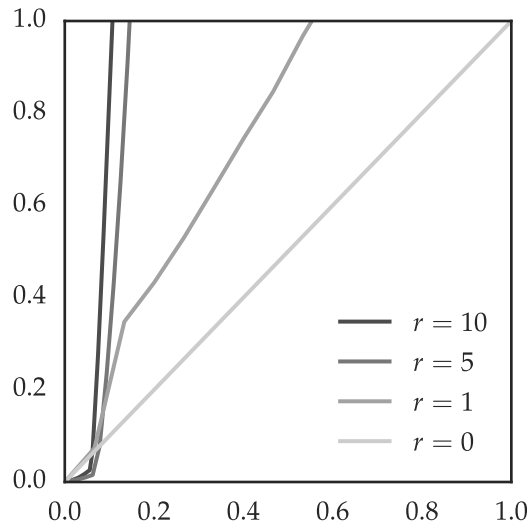


Figure 6: Aug-aug plot for r -subdivisions of K_{50} .

And indeed, we see that the plots reflect our expectations. For $r = 0$, the generated graph is simply K_{50} and the corresponding random graph is very similar (since we are using the configuration model, it might be missing a few edges). For $r > 0$ we see a jump in the aug-aug plot after precisely r augmentations, that is, in the graph \vec{G}_{r+1} .

With these two calibrations we can be confident that our assessment of the structural density of real-world network data is well-founded both theoretically and empirically. We begin our programme with the

smallest networks in the corpus and work our way up to the very large ones.

16.2 SMALL NETWORKS

KARATE The famous network collated by Zachary in 1977 from a university karate club [252]. The 34 vertices represent members of the club and the edges personal ties between them.

BERGEN A network of the University of Bergen’s algorithm group and their first neighbourhood according to the dblp database. Edges represent co-authorship of a paper listed in dblp. The data was collected by Pål Grønås Drange.

DOLPHINS Lusseau *et al.* [176] observed 62 bottle-nose dolphins near Doubtful Sounds, New Zealand, during the years 1994–2001. By observing which dolphins appeared in a school together, they mapped out the relationship graph. According to the authors the group composition of the dolphins was extremely stable due to geographical factors, and this stability is reflected in the statistics of the network.

LESMISERABLES This network, created by Knuth [160], contains 77 characters from Victor Hugo’s *Les Misérables*. The (weighted) edges represent co-occurrence in the same chapter.

POLBOOKS Valdis Krebs regularly surveyed the political publishing landscape of the US. He started with political books listed on best-seller lists and then expanded the network using Amazon’s recommendation system (which is based on user preferences). The data consists of 105 books that share an edge if Amazon customers who bought both books. This network represents his findings in 2004 [163]. In particular, he saw a clear distinction into a conservative and a liberal cluster, only bridged by two books¹.

WORD-ADJACENCIES In his work on community detection using eigenvectors, Newman tested his approach using a word-adjacency network derived from Charles Dickens’ *David Copperfield* [196]. The network contains 112 vertices, derived from the 60 most common nouns and the 60 most common adjectives. A connection between them indicates that they occur in direct succession in the text (eight words with degree zero were excluded). Since adjectives are most commonly succeeded by nouns in the English language, the graph should fall into two ‘anti-communities’, that is, an approximately bipartite graph.

¹ See also the article by the New York Times about this network:
<http://www.nytimes.com/2004/03/13/arts/13BOOK.html>

FOOTBALL Girvan and Newman proposed a community-detection algorithm [124] based around the ideas of centrality indices. One of the data sets they collated in order to test their approach is based on American football teams competing in the 2000 season of Division I. The 115 vertices represent university teams and an edge corresponds to a played match. The graph exhibits some natural community structure due to the set-up of the league.

The following table contains the data resulting from our statistical analysis of the small network's degree-distribution. Some of the data sets are quite small and the tail-estimation needs to be taken with a grain of salt, in particular in case the result is close to a tie.

	Power law	Truncated power law	Lognormal	Stretched exponential	Exponential	γ
Karate	1.57	3.02	1.52	1.50	0.00	2.87
Bergen	0.00	3.68	0.63	1.72	0.56	3.08
Dolphins	0.00	2.84	0.44	1.03	1.12	5.13
Lesmiserables	2.21	2.22	4.55	1.42	0.00	4.68
Polbooks	1.23	7.83	1.79	2.15	0.00	2.92
Word-adjacencies	3.17	3.17	3.54	2.06	0.00	4.00
Football	11.50					75.70

According to this analysis, we should consider the tails of all except two small networks as supercubic: the exceptions are Karate, whose power-law exponent lies far below three (and, not pictured, the truncated power law damping factor is negligible small), and Polbooks. The latter network's degree distribution seems to be best described by the truncated power law $d^{-2.34}e^{-0.04d}$ (excluding the normalisation factor). Since the exponential damping is quite small and the network contains only 105 vertices, it is probably better to state that the distribution is subcubic, but we cannot conclude this with certainty. Finally, Football caused the power law package to fail with divide-by-zeros and the above value cannot be taken seriously. A manual inspection of the degree distribution shows that almost half of the vertices have degree eleven and a quarter of them ten; with a maximal degree of twelve. This hardly looks like a power law distribution and we have to dismiss the above data on this network.

Moving over to the aug-aug plot in Figure 7, we can conclude that the networks Bergen, Lesmiserables and Word-adjacencies are structurally sparse: their degree distribution generate structurally sparse graphs and the aug-aug plot shows that the indegree of their dtf-augmentations grows slower than that of those random graphs. The degree distribution of Karate has a subcubic tail, but the aug-aug plot shows it way below the 45°-degree line. We cannot judge with certainty the

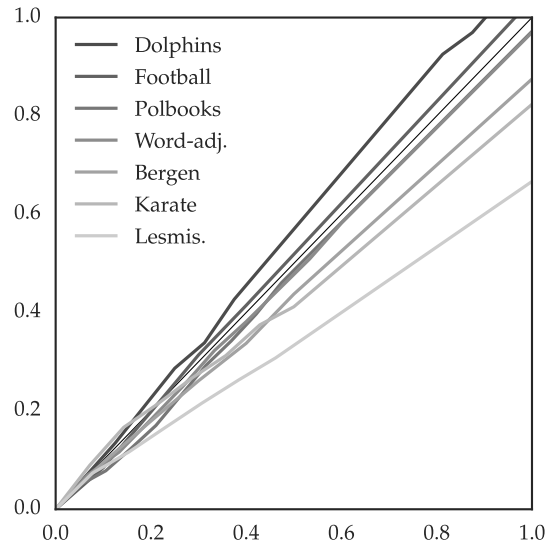


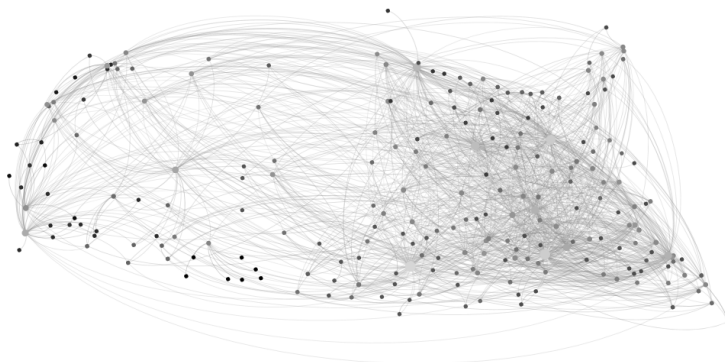
Figure 7: Aug-aug plot of small networks (less than 115 nodes). The networks are sorted from bottom to top according to largest value of their respective plot.

density of this network. The same holds for Polbooks, since here even the degree distribution did not clearly fall into any category.

Finally, Dolphins and Football fall above the 45° -degree line: they are structurally denser than the graphs generate from their respective degree distribution. For both networks we cannot make a judgement on their structural density.

16.3 MEDIUM-SIZED NETWORKS

AIRLINES The origin of this network is illusive: its was part of the example datasets provided with the Gephi graph visualisation software, but the corresponding web page claims an ‘unknown origin’.



It contains 235 US airports (with geographical meta-data) and it is reasonable to assume that the edges represent direct flights.

SP-DATA-SCHOOL Stehlé *et al.* collected data on face-to-face interactions between 228 students and 10 teachers of a French primary school using RFID badges. The network used here is the resulting data set for the second day [231].

C.ELEGANS This network represents the nervous system (the ‘connectome’) of the worm *Caenorhabditis Elegans*, a small nematode worm living in soil. The original data was collected by White *et al.* [249] and compiled by Watts and Strogatz for their work on small-world networks [247]. The network has 297 vertices and 2148 edges. It is not quite clear what synaptic connections the edges represent, the primary source (Watts’ homepage) is unavailable and the total number of edges is only 75% of what other sources would indicate².

CODEMINER Another network from the Gephi package [1], this network represents the call-graph of Java program. It consists of 724 vertices representing packages (one vertex), classes (79 vertices), fields (175 vertices), and methods (469 vertices). Arcs between the vertices represent calls (from method to method) and containments (packages to classes, between classes, and classes to methods).

CPAN-AUTHORS This network is a snapshot of the relationship between Perl developers in the Comprehensive Perl Archiver Network (CPAN). Two developers are connected by an edge if they use the same module in their code. The data is part of the Gephi package [1].

The following table contains the data resulting from our statistical analysis of the medium-sized network’s degree-distribution.

	Power law	Truncated power law	Lognormal	Stretched exponential	Exponential	γ
Airlines	5.17	5.55	9.46	4.06	0.00	2.72
Sp-data-school	0.00	1.11	1.40	2.22	1.40	31.60
C.Elegans	3.50	3.50	6.58	2.43	0.00	3.54
Codeminer	0.97	3.08	1.20	1.15	0.00	3.24
Cpan-authors	7.87	8.12	6.79	4.78	0.00	2.27

We see that the networks Sp-data-school, C.Elegans and Codeminer are best described as having a degree distribution with a supercubic tail. Airlines seem to follow a lognormal distribution and Cpan-authors a truncated power law. However, the proposed parameters for the truncation ($\lambda \approx 10^{-5}$) is so small that we should err on the side of caution and conclude the degree distribution of Cpan-authors might follow a power law with exponent around 2.27—clearly in the subcubic regime.

² See, for example, the worm atlas <http://www.wormatlas.org>.

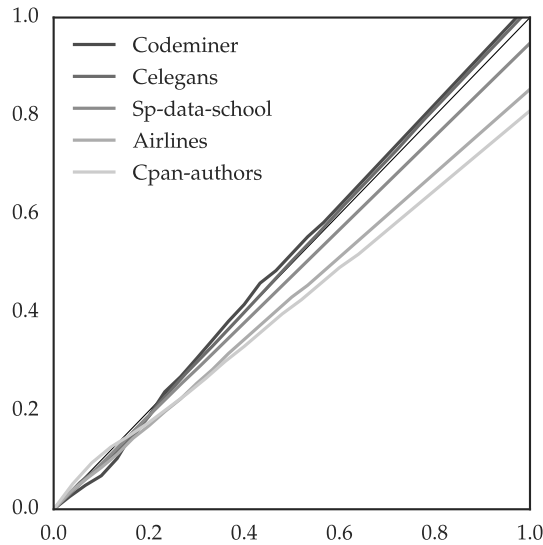


Figure 8: Aug-aug plot of medium-sized networks (between 235 and 839 nodes). The networks are sorted from bottom to top according to largest value of their respective plot.

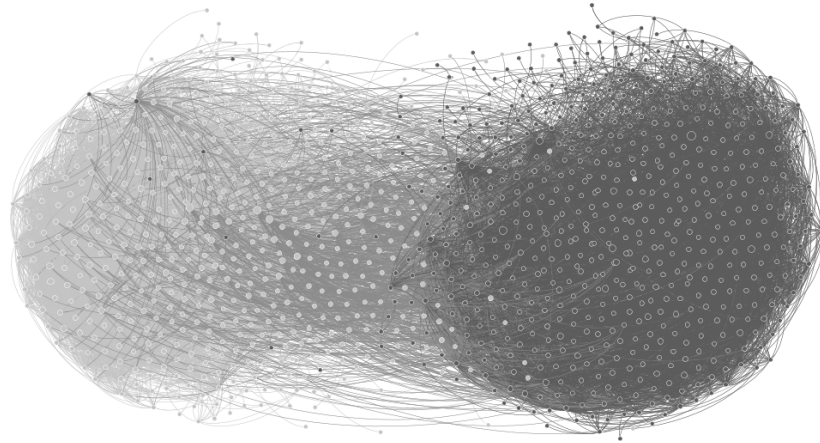
We can see from the plot in Figure 8 that the networks Sp-data-school, Airlines and Cpan-authors are structurally sparser than random networks with the same degree distribution. We therefore conclude that the former two can be called structurally sparse; for Cpan-authors we cannot make a solid statement. The networks C.Elegans and Codeminer appear to be slightly denser than their random counterparts, hence we strictly speaking cannot conclude anything. However, since their fitted power-law exponent lies way above 3.3, and their plots lie only slightly above the 45°-line we can carefully suggest that it is likely that they are structurally sparse.

16.4 LARGE NETWORKS

DISEASOME Goh *et al.* [125] created a bipartite graph containing 1777 genes and 1284 genetic disorders listed by OMIM³. An edge between a gene and a disease indicates a known correlation between the expression of the gene and the disease. The network data itself was obtained from the Gephi package and contains 1419 vertices and 2738 edges. It is unclear why the network is smaller than described in the original paper, potentially it is simply the largest connected component.

POLBLOGS Adamic and Glance surveyed the link-patterns between American bloggers just before the US presidential election 2004 [2].

³ Online Mendelian Inheritance in Man. <http://www.omim.org/>



The 1224 blogs clearly fall into a left-vs-right pattern. The above embedding was obtained with a simple force-layout, yet it reflects the political leanings contained in the original data (light grey indicates leaning towards Democrats, dark towards Republicans) incredibly well.

NETSCIENCE Another data set collated by Newman [196] which reflects the co-authorship of researchers in the networks science field. The whole network contains 1461 researchers, but the largest connected component contains only 379 of them.

DROSOPHILA In their work on the motion-detection of *Drosophila melanogaster*, Takemura *et al.* map out the connectome of a certain subsystem contained within the *medulla* neuropolis of the optic lobe [233]. The resulting network contains 1781 of these subsystems with a total of 8911 connections between them.

YEAST Bu *et al.* collected protein-protein interaction data of *Saccharomyces cerevisiae* (Brewer's yeast) and examined the resulting network for dense patterns [37]. Such dense patterns, the authors conclude, seem to correlate well with biologically similar functions and they use this graph-theoretic approach to predict the function of uncharacterised proteins. The network contains 2361 vertices with 7182 edges, of which 536 are loops.

CPAN-DISTRIBUTIONS Similar to the earlier described Cpan-authors, this network is based on the CPAN network [1]. Instead of connecting authors through packages, this network contains the dependencies of the 2719 software packages themselves.

TWITTERCRAWL Another network from the Gephi package [1], this network maps out retweets and mentions between twitter users. Unless other social networks, this network is particularly dense. With 3656 users and 154824, the average degree is almost 85. This can probably be attributed to the fast-paced nature of the twitter platform.

POWER Watts and Strogatz obtained a network representation of the western US power grid from Phadke and Thorp, who used the data in their book on computer-controlled relays in power substations [203]. The network contains 4941 power stations connected by a total of 6594 in, as expected, a single connected component. Already the network’s density marks it as an outlier in our corpus; moreover, the network is—for obvious reasons—close to being planar.

The following table contains the data resulting from our statistical analysis of the large network’s degree-distribution.

	Power law	Truncated power law	Lognormal	Stretched exponential	Exponential	γ
Diseasome	5.24	5.23	7.78	3.59	0.00	3.39
Polblogs	0.78	4.81	2.01	3.03	0.00	3.69
Netscience	6.88	6.88	15.55	4.70	0.00	4.17
Drosophila	2.08	2.08	4.21	1.35	0.00	3.97
Yeast	1.85	11.85	3.22	3.96	0.00	3.46
Cpan-distr.	25.46	25.39	19.01	8.86	0.00	3.67
Twittercrawl	3.55	3.55	8.22	2.38	0.00	5.75
Power	4.70	4.67	6.66	2.58	0.00	10.44

According to this data, we postulate that all listed network have degree distributions whose tail is best described by a supercubic function. This might seem surprising for Twitter, but a subjectively high average does not imply the presence of dense structures.

The aug-aug plot for large networks in Figure 9 clearly shows that Power, Diseasome, Netscience, Drosophila and Polblogs are sparse. The interesting behaviour of the Power-plot is probably due to it being close to a planar graph and having a comparatively high diameter (around 46). For such networks the comparison to G^{CF} is probably not a good choice since its statistics, other than the degree-distribution, are not well-replicated. However, we can use our domain-knowledge to make a judgement: since the power network will certainly only have a small number of crossings per edge in its real-world embedding, we can conclude that it has bounded expansion [193].

The plots of Yeast and Cpan-distributions lie slightly above the 45° -line. We therefore cannot assess the structural properties of these networks with finality, but since the plots do not diverge drastically and the networks’ degree distributions are supercubic by a comfortable margin, we suggest that it is more likely that they are structurally sparse than dense.

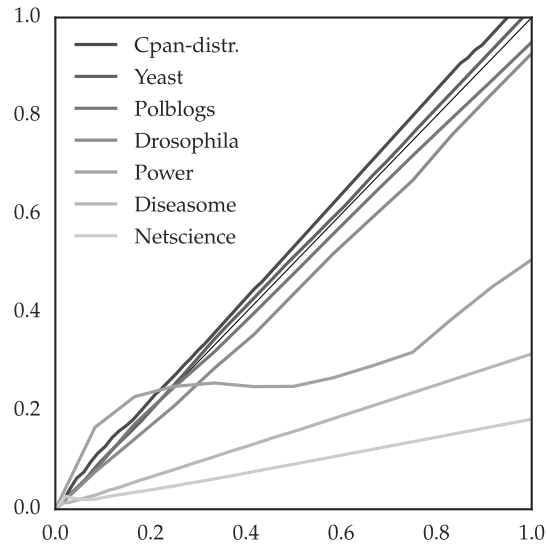


Figure 9: Aug-aug plot of large networks (between 1419 and 4941 nodes). For the networks Power, Polblogs and Drosophila only the first twelve augmentations were computed.

16.5 HUGE NETWORKS

AS JAN 2000 Autonomous systems are logical subunits of the internet, networks controlled by a single administrative entity. The Route Views project⁴ by the University of Oregon regularly provides routing information. As part of the Border Gateway Protocol (BGP), each router of an AS facing the outside world maintains a list of *peers* with which it exchanges routing information. This peer-relationship between 6474 autonomous systems forms, aggregated for over one month, the network. The network is part of the Stanford Network Analysis Project (SNAP) collection [172].

HEPTh In his work on scientific collaboration networks, Newman collected several data sets from four databases [194]. The HepTh network consists of 7610 scientists from the field of theoretical high-energy physics. An edge between scientists indicates a co-authored paper within a five year window (1995–1999). The average number of collaborations is around four and the largest connected component contains roughly 76% of all nodes.

GNUTELLA Gnutella is a peer-to-peer file sharing network that, as many other such networks, was established in the early 2000s. Ripeanu, Iamnitchi, and Foster used a crawler in August 2002 to measure the growth, traffic, and topology of the Gnutella network [212]. In total

⁴ <http://www.routeviews.org/>

edges represent communication between them (and edges exists if at least one email was sent from one address to the other).

BRIGHTKITE Brightkite was a location-based social network, meaning that users could ‘check-in’ at certain locations (usually restaurants, cafes *etc.*). Cho, Myers, and Leskovec collected the social network data alongside these check-ins; providing not only social but also geographical information [44]. With 36692 users connected by 183831 friendships (an average of 10 friends per user) this network is the largest one in our test corpus.

The following table contains the data resulting from our statistical analysis of the huge network’s degree-distribution.

	Power law	Truncated power law	Lognormal	Stretched exponential	Exponential	γ
AS Jan 2000	6.05	6.82	8.18	4.42	0.00	2.17
HepTh	0.30	15.86	3.26	5.94	0.00	3.78
Gnutella	18.92	18.92	19.88	10.43	0.00	5.40
ca-HepPh	8.57	44.20	18.78	27.51	0.00	2.11
CondMat	3.85	10.61	5.33	4.66	0.00	3.80
ca-CondMat	2.81	8.42	4.42	4.58	0.00	3.58
Enron	53.43	62.62	47.66	32.81	0.00	2.11
Brightkite	7.65	16.90	12.00	14.45	0.00	2.56

We can immediately conclude that the degree distribution of HepTh, Gnutella, CondMat, and ca-CondMat is best described by a function with a supercubic tail. The network AS Jan 2000 seems to best follow a lognormal distribution for which we could conclude the same, however, the statistical evidence of this distribution over a low-degree power law is not extremely high. A similarly fuzzy picture is presented by ca-HepPh: the evidence for a truncated power law seems extremely high, however, the truncation parameter is quite small ($\lambda \approx 0.003$), while proposed power law coefficient is below two (for the truncated case it is $\alpha = 1.64$). Similarly, we find an very small cutoff parameter ($\gamma \approx 0.004$) for the Enron network and we suggest that it is best described by a function with a subcubic tail, although we stress that the evidence is not conclusive. Brightkite poses a similar problem and we cannot conclude with absolute certainty how its degree distribution should be described.

The aug-aug plot in Figure 10 suggests that all networks except Gnutella appear to be sparser than the corresponding random graphs.

Considering the collaboration networks HepTh, Condmat, ca-HepPh, and ca-CondMat, it seems reasonable to conclude that they are structurally sparse—we see that this is clearly true for HepTh and there is

no reason to believe that the further augmentations of the other three networks behave differently. While varying densities due to different publishing cultures (in particular the typical number of authors on a publication) is plausible, a varying *structural* density is not: the underlying processes generating these networks are too similar.

Note that HepTh shows the same trend in the first three augmentations as Condmat, ca-HepPh, and ca-CondMat (it is hard to make out due to scaling for HepTh). We can see a similar trend in smaller networks, but it is undeniably pronounced in the very large ones—and from what is known about complex networks, we should expect it! Since real-world data often shows high clustering, which is not replicated in random graphs generated with the configuration or the Chung–Lu model, the *local* density of the real data should be higher than its artificial replicate. We will discuss this issue further in this chapter’s conclusion.

We are left with the two largest networks of our corpus, Enron and Brightkite. If the other large networks are any indication, we should expect these two networks to be sparser than their random counterparts. But since we could only compute the first five augmentations, this can only be a preliminary assessment.

16.6 A VERDICT

In this chapter we tried to answer the question whether real-world networks are structurally sparse empirically, using the results from

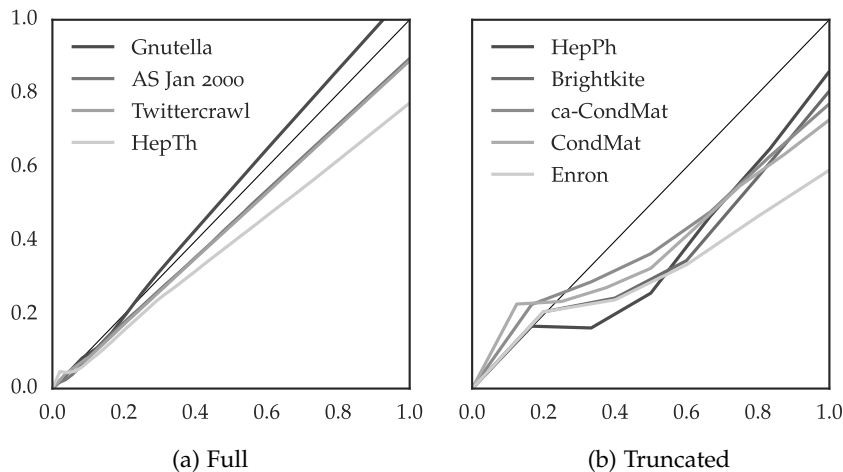


Figure 10: Aug-aug plot of huge networks (between 6474 and 58228 nodes). The plots on the left were computed for the full set of augmentation, the plots on the right had to be truncated due to memory and time consumption. For Brightkite and Enron we computed the first five augmentations, for ca-HepPh and ca-CondMat the first six and for CondMat the first eight.

Chapter 15 as the theoretical foundation. The convenient statistical assessment available through the power law package made it, in most cases, easy to decide how the degree distribution of a network should be classified; and the juxtaposition of real-world structural density with that of corresponding random data in the aug-aug plots gave us the means to extend this classification to our network corpus. In summary, we decided that the networks Bergen, Lesmiserables, Wordadjacencies, Sp-data-school, Airlines, Diseasome, Polblogs, Netscience, Drosophila, Power, HepTh, ca-HepPh, CondMat, and ca-CondMat are structurally sparse. We suggest that the networks Celegans, Codeminer, Cpan-authors, Yeast, Cpan-distributions, Enron, and Brightkite are structurally sparse but cannot conclude it with certainty due to either missing data or an unclear assessment of their degree distribution. Finally, for the networks Karate, Dolphins, Poolbooks, and Football our framework fails and we cannot provide any answer.

While it would be more satisfying to have a full judgement on every single network presented here, note that what we did *not* find: not a single network in the corpus presented clear evidence of being structurally dense. The emerging picture is clear: complex networks are structurally sparse.

One of the most versatile tools developed in the structural sparseness programme by Nešetřil and Ossona de Mendez are the low-treedepth colourings: These colourings are the closest we have to a decomposition of sparse graphs. However, the algorithms presented so far [189, 192] to compute such colourings turn out to be impractical. Recall that the basic idea is to stepwise augment a directed version of the target graph with fraternal and transitive edges, for a total $f(p)$ number of steps. Any proper colouring of the resulting graph $\vec{G}_{f(p)}$ then is a p -centred colouring of the original graph.

In practice, however, the number of augmentation needed is far less than the theoretical bound $f(p)$ (which grows faster than exponential in p). Hence it makes sense to interleave the augmentation process by a routine which colours the graph and checks whether the colouring is already p -centred. To implement such a check efficiently is a challenge in itself. Moreover, what colouring heuristic should one use? Since the augmentations have provably bounded degeneracy if the target graph has bounded expansion, we could simply use the colouring number of the augmentations. Again, in practice colouring heuristics improve that theoretical bound by a wide margin. The orientation of fraternal edges (which for theoretical purposes one simply does by choosing the acyclic orientation available due to bounded degeneracy) is another subtle component that might benefit from heuristical tweaking.

As a result of all these uncertainties, we created a framework which allowed us to easily exchange the different modules of the algorithm and to test different configurations. Kevin Jasik deserves credit for implementing most of the framework.

Our experiments showed that the choice of colouring routine and orientation strategy has an influence on the resulting number of colours, the effect is however not consistent enough across the tested instance to determine a universally superior combination. Three other improvements of the algorithm did, however, show remarkable improvements:

1. Using dtf-augmentations

The dtf-augmentations introduced in Chapter 7 do not only allow slightly better theoretical bounds on the number of augmentations necessary, they also solve an inherent problem of tf-augmentations. Namely, that transitive arcs added at step t tend to span much larger distances than fraternal arcs. The introduction of distances in the dtf-augmentations forces both types of arcs to ‘grow’ at a similar rate.

2. Removing high-degree vertices

By simply assigning the top k vertices of highest degree a unique colour, we can remove them from the graph and use the algorithm to obtain a p -centred colouring for the remainder. Since networks are very heterogeneous, this heuristic both improves the running time and the number of colours significantly. By trying some values for k , one can hone in on a good value from below, until either the colouring number does not decrease any more or the running time grows too large.

3. Postprocessing

Given a p -centred colouring c and a routine that verifies that property, it is straightforward to test whether any two colour classes can be merged to obtain a colouring c' with less colours.

17.1 THE SMALL NETWORK BENCHMARK

In order to assess what components of the tf-framework work best, we tested different configuration on a set of small networks. Since the number of total configuration is prohibitively large, we variate only a single component at a time. There is the remote possibility that a weird combination of otherwise inferior components work incredibly well together. We deem, however, this possibility unlikely enough to forgo the brute-force testing of all variations.

Choice of augmentation

TF-AUGMENTATIONS VS. DTF-AUGMENTATIONS We settle the fundamental question first: do dtf-augmentations help our case?

Network	p					p				
	2	3	4	5	6	2	3	4	5	6
Bergen	22	27	27	27	27	22	26	26	26	26
Dolphins	13	34	34	41	41	13	26	32	32	32
Karate	10	15	15	15	15	10	15	15	15	15
Lesmis.	23	23	29	29	29	23	23	28	28	28
Polbooks	21	37	60	60	60	21	30	45	45	59
	(a) Tf-augmentations					(b) Dtf-augmentations				

Table 2: Comparison of colouring numbers on small networks using tf- and dtf-augmentations. The colouring heuristic used was 'Greedy' and no postprocessing for colour reduction was applied.

The above table clearly shows that it does in our benchmark setting. All other components kept equal, the number of colours decreases in almost all cases, sometimes significantly. It is to be expected that ultimately, the two should roughly arrive at the same number. However,

the slower growth makes dtf-augmentations clearly the better practical choice.

Colouring

COLOURING HEURISTICS Within the tf-framework, we tested four well-known colouring heuristics. First, the *Greedy* heuristic iterates through the vertices in an arbitrary order and assigns a vertex v the lowest¹ colour that does not appear in its neighbourhood. The *Maximum Degree* heuristic iteratively chooses the next uncoloured vertex of highest degree and colours it by the lowest available colour. The heuristics *Minimum Degree*, *Maximum Indegree* and *Minimum Indegree* work accordingly, where the latter two take into account the directed nature of the augmentations. Finally, the *Degree Saturation* heuristic chooses that vertex that sees the largest number of colours in its neighbourhood and again colours it by the next available colour.

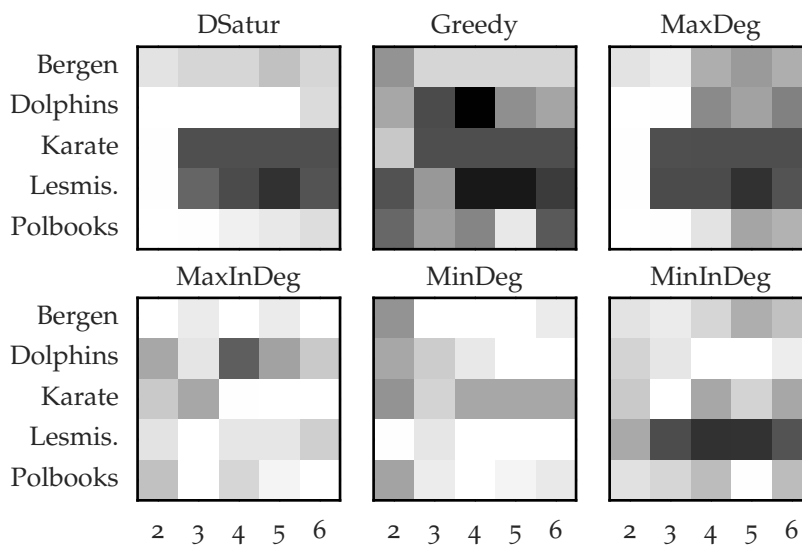


Figure 11: Relative performance of six colouring heuristics on five small network instances. The horizontal axis denotes the colouring number χ_p for $2 \leq p \leq 6$. A white square indicates that the heuristic provided the minimum number of colours among all heuristics. The darker the square, the higher the number of colours.

The benchmark results are visualised in Figure 11. The *Degree Saturation* seems to be favourable in three of five instances and might be well complemented by *Maximum Indegree* and *Minimum Degree*. Since the variation is not incredibly large for these latter three and *Degree Saturation* performs well across all instances, we chose it as the colouring heuristic in the sequel.

Orientation

ORIENTATION HEURISTIC When we proved the bounds on the growth of dtf-augmentations in Chapter 7, orienting the fraternal edges in an acyclic fashion was an important ingredient of the proof. Since

¹ We consider colours as ordered.

our goal is to bound the maximum indegree of the resulting orientation, we can improve this orientation strategy by taking into consideration the already present indegree of vertices. To be more precise: in the d^{th} dtf-augmentation, *i.e.* building the graph \vec{G}_d from \vec{G}_{d-1} , let F be the set of fraternal edges. In the *unweighted* strategy (the one used in the proof of Lemma 34), we compute an acyclic orientation of $G[V(F)]$, using its degeneracy-order, to choose the orientation for F . In the *weighted* strategy we adjust the computation of the degeneracy-order as follows: let $S \subseteq G[V(F)]$ be the vertices already removed from the graph. Instead of choosing the vertex $v \in G[V(F) \setminus S]$ of minimal degree next, we choose that vertex v that minimises the quantity

$$d_{\vec{G}_{d-1}}^-(v) + \text{deg}_{G[V(F) \setminus S]}(v).$$

Again, we compared the two strategies using our benchmark networks.

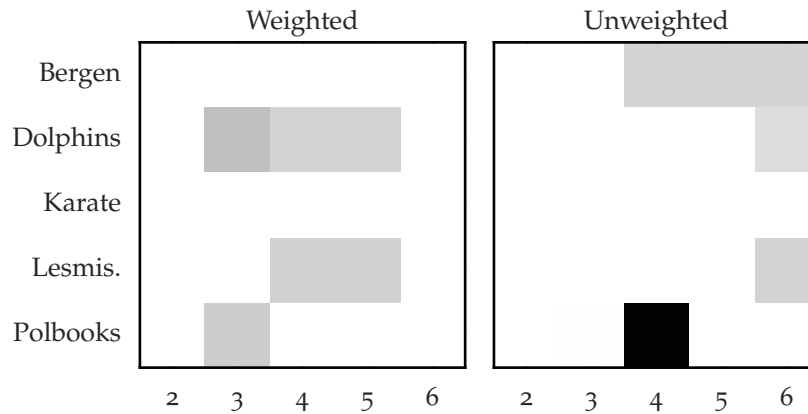


Figure 12: Relative performance of the two orientation strategies on five small network instances. The horizontal axis denotes the colouring number χ_p for $2 \leq p \leq 6$. A white square indicates that the strategy provided the minimum number of colours among all heuristics. The darker the square, the higher the number of colours.

The result are visualised in Figure 12. The absolute difference in performance, *i.e.* the number of colours, is not very high. The weighted strategy seems to perform better for higher colouring numbers, but in conclusion the orientation strategy seems to have far less of an impact than the colouring heuristic employed.

17.2 TESTING CENTRED COLOURINGS EFFICIENTLY

As alluded to before, the one expensive and recurrent task which makes heuristical improvements even possible is an efficient algorithm that tests whether a given colouring is p -centred. The benchmark here is not simply the asymptotic complexity of such an algorithm (we of course demand that it is reasonable), but rather whether it performs

well in the given task. Let us start our with a simple, efficient algorithm and engineer it towards an *applicable* one.

The basic algorithm

In order to test whether a colouring c of a graph G is p -centred, we need to proceed as follows: for every subset $C \subseteq c(G)$ with $|C| < p$, we test whether c is a centred colouring for the subgraph $G[C]$, *i.e.* every connected subgraph of $G[C]$ contains a centre.

Testing inductively

The first obvious improvement is using induction over the colour sets. Assume we have verified that all colour sets with a most i colours induce centred subgraphs. Testing a colour set C of cardinality $i + 1$ then reduces to testing whether $G[c^{-1}(C)]$ contains a centre. To see that, observe that if $a \in C$ is the colour of a centre of $G[c^{-1}(C)]$, then the components of

$$G[c^{-1}(C \setminus \{a\})] = G[c^{-1}(C) \setminus c^{-1}(a)]$$

each receive at most i colours. By induction, the colouring c is centred in these components, and therefore the same holds for $G[C]$.

Testing by connectivity

If we stick with the inductive approach, another improvement reveals itself: assume we have verified that a set $C \subseteq c(G)$ induces a centred subgraph in G . Let $a \in c(G) \setminus C$ be a colour not contained in C . Then verifying that $G[C \cup \{a\}]$ is centred boils down to the following basic question: if a vertex v coloured a connects several connected components of $G[C]$, is the colouring still centred? If no other vertex coloured a connects to any of these components, the answer is—by induction—yes. If on the other hand more vertices coloured a are connected to them, then clearly a is not the centre colour. But to determine whether the resulting connected component of $G[C \cup \{a\}]$ has a centre, it is enough to know the frequency of each colour in members of $K(G[C])$. In fact, it is enough to know whether the frequency is zero, one or above one.

Let us define a data structure D with the following semantic: for a colour set $C \subseteq c(G)$ and a vertex $v \in c^{-1}(C)$, $K_C[v]$ contains the connected component of v in $G[c^{-1}(C)]$, *i.e.*

$$K_C[v] = K_C[w] \iff v, w \text{ are connected in } G[c^{-1}(C)].$$

Additionally, for each component in K_C , we store the corresponding colour frequencies F . Let $b \in C$ be a colour, then

$$F_C[K_C[v]][b] = |c^{-1}(b) \cap K_C[v]|.$$

With the help of these data structures we can implement the following algorithm (listed as Algorithm 3) to test p -centred colourings: by using a traversal of the p -sized subsets of $c(G)$ that before listing a set C will always list some subset with $C' = C \cup \{a\}$, we can compute K_C from $K_{C'}$ using very efficient union-find operations. If at any point we find that $F_C[K_C[v]]$, for some vertex v , contains no colour of frequency one, then c is not p -centred.

Instead of storing the actual frequencies, we can simply truncate all values in F_C above 2 and use only two bits per counter. Furthermore,

because we use the above described traversal of the colours, only p data structures K, F need to be kept in memory at a time. Kevin Jasik implemented Algorithm 3 which made it possible to compute reasonable colouring in the first place.

Input: A graph G with vertex-colouring c , and integer p .

Output: True iff c is a p -centred colouring of G .

Let $c(G) = [\ell]$

$C \leftarrow [1]$

while $C \neq \emptyset$ **do**

for $v \in c^{-1}(C)$ **do**

if $\nexists b: F_C[K_C[v]][b] = 1$ **then**

return false

if $|C| < p$ **and** $C[0] < \ell$ **then**

$a \leftarrow C[0] + 1$

$F_{C+a}, K_{C+a} \leftarrow \text{merge}(F_C, K_C, a)$

$C \leftarrow [a] + C$

else

while $|C| > 0$ **and** $C[0] = \ell$ **do**

$C \leftarrow C \setminus C[0]$

if $|C| > 0$ **then**

$C[0] \leftarrow C[0] + 1$

return true

function $\text{merge}(F_C, K_C, a)$

$K_{C+a} \leftarrow K_C$

$K_{C+a}[v] \leftarrow v$ **and** $F_{C+a}[K_{C+a}[v]][a] \leftarrow 1$ **for** $v \in c^{-1}(a)$

$F_{C+a}[K_{C+a}[v]][b] \leftarrow 0$ **for** $v \in c^{-1}(a), b \in C$

for $v \in c^{-1}(a)$ **do**

for $u \in N(v) \cap c^{-1}(C)$ **do**

for $b \in C \cup \{a\}$ **do**

$F_{C+a}[K_{C+a}[v]][b] \leftarrow F_{C+a}[K_C[u]][b]$

$K_{C+a}[u] \leftarrow K_{C+a}[v]$

return F_{C+a}, K_{C+a}

Algorithm 3: Efficient verification of a p -centred colouring.

17.3 COLOURING THE CORPUS

With the above mentioned improvements and heuristics, we applied the resulting algorithm to the previously described networks. A lot of the work was done by hand—although it could easily be automated. This includes the choice for how many high-degree vertices have been removed from the network before the colouring algorithm was applied

and how long the subsequent colouring-improvement routing was allowed to run.

In Table 3 we have summarised our findings. We computed, if possible, the first six centred colouring numbers, where the column $p = 2$ simply contains the size of a proper colouring and $p = 6$ a six-centred colouring, meaning that up to five colours induce a graph of low treedepth. We further, using an implementation Tobias Räderscheidt based on an algorithm by Fernando Sánchez Villaamil, computed an upper-bound for the treedepth of the network.

The results of our colouring look mixed. Some networks have a decently low colouring numbers (sometimes even very low treedepth as in the case of Netscience). We can expect that algorithms based on these colourings have little trouble to run on instances like Codeminer, Cpan-authors, Diseasome, Netscience, Cpan-distributions, and Power.

For those networks with high colouring number, there are several issues to keep in mind. Some networks (*e.g.* Twittercrawl, Sp-data-school) already have a large colouring number (the column $p = 2$): we established that they are structurally sparse, but the constants involved might doom any attempt at applying our toolkit in practice to fail. For others, like Gnutella, CondMat, ca-CondMat, Enron, and Brightkite, there is still hope that our algorithm is simply not specialised enough. Note also that the post-processing step is rather efficient if the colouring already contains few colours (and in that case, often reduces the number of colours by a sizeable fraction), but absolutely fails in case of many colours. This bias increases the gap between low and high values and might skew the results significantly.

Future improvements

Finally, note that our algorithm sometimes (*cf.* the result on Polbooks, Netscience, Yeast, CondMat) uses more colours than the treedepth of the respective network. This clearly shows that there is potential for improvement since it holds that $\chi_p(G) \leq \mathbf{td}(G)$ for all $p \in \mathbb{N}$. It remains an intriguing open question whether this can be achieved by a different colouring routine, a different orientation heuristic, or whether an entirely different type of algorithm is needed.

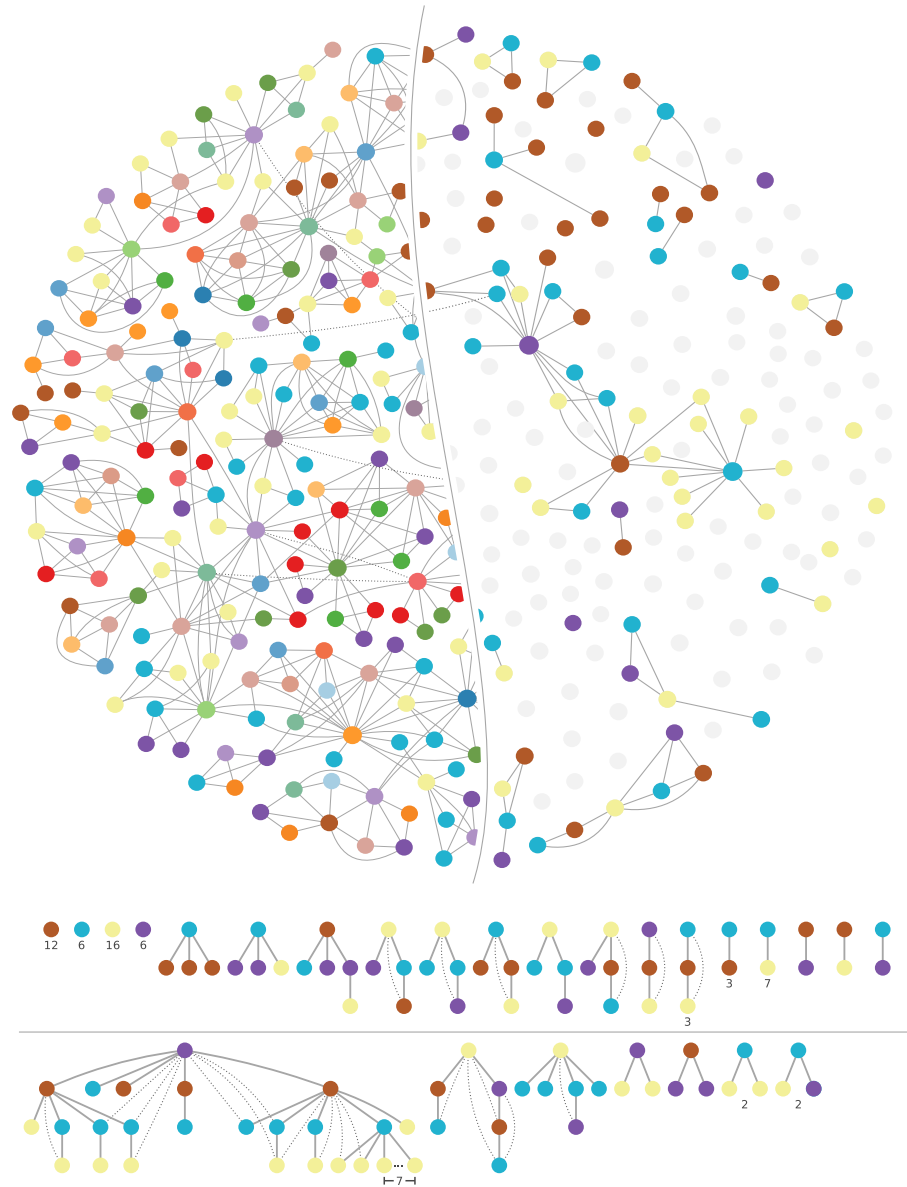


Figure 13: A five-centred colouring (using 21 colours) of Netscience (largest connected component shown). The right half is restricted to a subgraph formed by four colour classes. Below, the corresponding representation by graphs of treedepth at most four (with multiplicities noted).

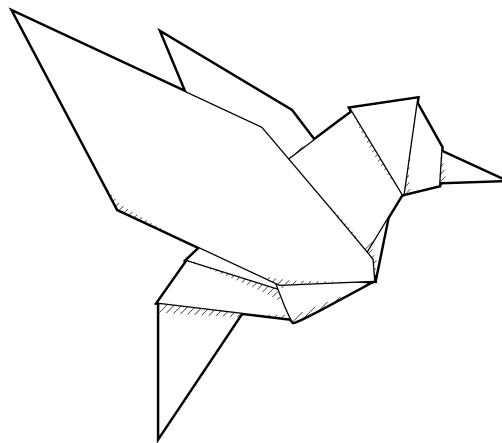
Network	Nodes	Edges	P					
			2	3	4	5	6	∞
Karate	35	78	6	7	9	9	10	8
Bergen	53	272	17	19	20	20	20	22
Dolphins	62	159	7	11	17	18	19	24
Lesmiserables	77	254	10	15	16	16	16	16
Polbooks	105	441	8	16	22	29	31	30
Word-adj.	112	425	8	18	27	35	41	48
Football	115	613	9	22	33	49	62	69
Airlines	235	1297	11	28	39	47	55	64
Sp-data-school	238	5539	23	100	138	157	168	171
C.Elegans	306	2148	8	36	74	83	118	153
Codeminer	724	1017	5	10	15	17	23	51
Cpan-authors	839	2212	9	24	34	43	47	224
Polblogs	1224	16715	30	118	286	354	392	603
Diseasome	1419	2738	12	17	22	25	30	30
Netscience	1589	2742	20	20	20	21	24	20
Drosophila	1781	8911	12	65	137	188	263	395
Yeast	2284	6646	12	38	178	254	431	408
Cpan-distr.	2719	5016	5	14	32	42	56	224
Twittercrawl	3656	154824	89	561	1206	1285	1341	-
Power	4941	6594	6	12	20	21	34	95
AS Jan 2000	6474	13895	12	29	70	102	151	357
HepTh	7610	15751	24	25	104	328	360	558
Gnutella	10876	39994	8	43	626	-	-	-
ca-HepPh	12008	118489	239	296	1002	-	-	-
CondMat	16264	47594	18	47	255	1839	-	1310
ca-CondMat	23133	93497	26	89	665	-	-	-
Enron	36692	183831	27	214	1428	-	-	-
Brightkite	58228	214078	39	193	1421	-	-	-

Table 3: Number of colours in p -centred colourings computed on real world networks and upper bounds of their respective treedepth.

Part V

CONCLUSION

*But just because something does not have an ending
doesn't mean it doesn't have a conclusion.*
— Iain M. Banks, *Use of Weapons*



We motivated the programme of this thesis by observing that the rising popularity of *network science*, a field fundamentally involved with sparse graphs, currently does not translate into a rising popularity of the theory of structurally sparse graphs. Our explanatory attempts left us with the conclusion that a primary factor in this issue is lacking evidence of this theory being *applicable* to network science. Accordingly, we framed this issue in the programmatic question

“Can the algorithmic tools derived from the theory of structurally sparse graphs be applied to real-world data?”.

And now, having thoroughly treated several different aspects of this question, we can formulate an answer: Yes, the overall indications towards the applicability of such algorithmic tools are positive.

The ‘right’ notion of sparseness

Nowhere dense and bounded expansion classes turned out to be the points at which the magisteria of structural graph theory and network science overlap: we have seen both theoretically and empirically that network data exhibits these notions of sparseness. We can be reasonably certain that no stricter notion of sparseness holds in these half-organised, half-disorganised data sets: the innate robustness provided through the Nešetřil-Ossona de Mendez dichotomy, which is in particular exhibited through invariance under minor-flavours and algebraic stability, is simply missing from other frameworks.

Algorithmic aspects

The existence of broad algorithmic meta-theorems—in form of first-order model checking and our improvements of the meta-kernelisation framework—clearly show that structural sparseness can even in its most general form be successfully exploited to improve algorithmic tractability. While these results are purely theoretical, they provide the umbrella under which further research into more applicable algorithms can flourish—not unlike the graph minor theorem and Courcelle’s theorem did for parametrised complexity. And hand-crafted algorithms, like Dvořák’s approximation for r -DOMINATING SET and our routine to compute r -neighbourhood sizes, demonstrate that such applicable algorithms are within reach.

We further saw how the restrictions imposed by working with these most general notions resulted in a kernelisation for DOMINATING SET that has a decent chance of being actually applicable, since it does not rely on practically infeasible decomposition theorems¹—provided that

¹ As Dániel Marx put it once: “If the constants in the Roberston-Seymour theorem are your friends, you don’t need enemies.”

our worst-case estimates are not actually met in practice, a reasonable assumption.

The algorithmic toolkit that exploits structural sparseness has great potential in the area of big-data graph analysis: not only run the provided tools in linear time, they also seem to favour parallel implementations due to their inherently local nature. And this is *before* we engineer the algorithms further to factor in network-specific properties like the small-world property or the skewed degree distribution. As our colouring experiments showed, real-world networks are much more tractable than our pessimistic worst-case estimates derived from the bounded expansion property would let us believe.

Questions related to complex networks

Several of the presented results from structural graph theory can, through the connection to complex networks presented here, be turned around to ask intriguing questions. The aforementioned claim of *densification* [171] is irreconcilable with our claim of structural sparseness, unless we posit that densification is an early-stage phenomenon that subsides during a networks' evolution. Since domain-specific considerations stand in stark contradiction with densification and the employed methodology of detecting power laws has noted flaws, further research into when a network's early-stage development undergoes a transition into a steady-state evolution is warranted.

The very unsatisfactory state of attachment models with respect to our framework highlights a flaw in how we treat properties of random graphs; we need to certainly go beyond the all-or-nothing asymptotics in order to evaluate structural density of models in practical settings. On the other hand, this also highlights a flaw in the models: with some non-vanishing probability, they generate extremely degenerate cases. Our case—the appearance of an arbitrarily large one-subdivided complete subgraph—is only one possible manifestation of this aberration; one could think of any possible scenario involving shallow minors and find that the probability of it occurring will be bounded away from zero, resulting in the generation of atypical instances.

Finally, our application of dtf-augmentations to our network corpus has uncovered an interesting fact: the first augmentation of a network has usually a higher maximum indegree than the randomly generated baseline networks, while the subsequent augmentations have a lower maximum indegree. This ties in beautifully with the known drawback of the configuration model and the Chung–Lu model: they do not replicate clustering, which is exactly the distance-two anomaly seen in the aug-aug-plots. These plots might therefore provide a viable tool to assess a model's ability, when fitted to real-world data, to replicate structural sparseness *and* clustering. More refined analysis could take into consideration the average degree of dtf-augmentation; or even how the augmentations' degree distribution evolves in order to characterise a network's structural properties at different depths. Such variations might provide valuable new statistics that capture the meso-scale properties of complex networks.

The question of applicability

The programme is not at its completion. We made great headway, settling the fundamental questions to assess whether attempts at solving real-world problems using the presented tools and techniques are feasible. Knowing that this is indeed the case, that the framework of structural sparseness *is* excellently suited to improve algorithmic tractability, we can promote our tools as viable solutions and seek applications that benefit from them. Our initial results on the p -centred colouring numbers of real-world networks stresses that such applications are within reach, but also highlight that much more engineering has to be applied in order to tailor the framework as closely as possible to complex networks. Certainly, the intersection between the algorithmic theory of structurally sparse graphs and network science promises to be a fruitful area for future research.

Part VI

BIBLIOGRAPHY

BIBLIOGRAPHY

- [1] Gephi datasets. <https://github.com/gephi/gephi/wiki/Datasets>.
- [2] L. Adamic and N. Glance. The political blogosphere and the 2004 US election: divided they blog. In *Proceedings of the 3rd International Workshop on Link Discovery*, pages 36–43. ACM, 2005.
- [3] J. Alber, M. R. Fellows, and R. Niedermeier. Polynomial-time data reduction for Dominating Set. *JACM*, 51:363–384, 2004.
- [4] J. Alber, H. Fernau, and R. Niedermeier. Parameterized complexity: exponential speed-up for planar graph problems. *Journal of Algorithms*, 52(1):26–56, 2004.
- [5] I. Albert and R. Albert. Conserved network motifs allow protein-protein interaction prediction. *Bioinformatics*, 20(18):3346–3352, 2004.
- [6] N. Alon and S. Gutner. Linear time algorithms for finding a dominating set of fixed size in degenerated graphs. *Algorithmica*, 54(4):544–556, 2009.
- [7] J. Alstott, E. Bullmore, and D. Plenz. powerlaw: A python package for analysis of heavy-tailed distributions. *PLoS ONE*, 9(1):e85777, 01 2014.
- [8] R. Andersen, F. Chung, and L. Lu. Modeling the small-world phenomenon with local network flow. *Internet Mathematics*, 2(3):359–385, 2005.
- [9] G. Anderson, U. Baur, M. Berger, F. Borcharding, A. Brandt, D. Denisov, S. Eno, T. Han, S. Keller, D. Khazins, et al. Summary of the very large hadron collider physics and detector workshop. *arXiv preprint hep-ph/9710254*, 1997.
- [10] J. Anthonisse. The rush in a directed graph. Technical Report BN 9/71, Stichting Mathematisch Centrum, 1971.
- [11] M. Appel and R. Russo. The maximum vertex degree of a graph on uniform points in $[0, 1]^d$. *Advances in Applied Probability*, pages 567–581, 1997.
- [12] S. Arnborg. Efficient algorithms for combinatorial problems on graphs with bounded decomposability—a survey. *BIT Numerical Mathematics*, 25(1):1–23, 1985.
- [13] S. Arnborg, J. Lagergren, and D. Seese. Easy problems for tree-decomposable graphs. *Journal of Algorithms*, 12(2):308–340, 1991.

- [14] A. Atserias, A. Dawar, and P. G. Kolaitis. On preservation under homomorphisms and unions of conjunctive queries. *Journal of the ACM*, 53(2):208–237, 2006.
- [15] B. Baker. Approximation algorithms for NP-complete problems on planar graphs. *Journal of the ACM (JACM)*, 41(1):153–180, 1994.
- [16] F. Ball, D. Sirl, and P. Trapman. Threshold behaviour and final outcome of an epidemic on a random network with household structure. *Advances in Applied Probability*, 41(3):765–796, 09 2009.
- [17] F. Ball, D. Sirl, and P. Trapman. Epidemics on random intersection graphs. *Annals of Applied Probability*, 24(3):1081–1128, 06 2014.
- [18] A. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.
- [19] E. A. Bender and E. R. Canfield. The asymptotic number of labeled graphs with given degree sequences. *Journal of Combinatorial Theory, Series A*, 24(3):296–307, 1978.
- [20] M. Biskup. On the scaling of the chemical distance in long-range percolation models. *Annals of Probability*, pages 2938–2977, 2004.
- [21] S. Blackburn, D. Stinson, and J. Upadhyay. On the complexity of the herding attack and some related attacks on hash functions. *Designs, Codes and Cryptography*, 64(1-2):171–193, 2012.
- [22] H. L. Bodlaender, R. G. Downey, M. R. Fellows, and D. Hermelin. On problems without polynomial kernels. *Journal of Computer and System Sciences*, 75(8):423–434, 2009.
- [23] H. L. Bodlaender, F. V. Fomin, D. Lokshtanov, E. Penninkx, S. Saurabh, and D. M. Thilikos. (meta) kernelization. In *Proceedings of the 50th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 629–638. IEEE Computer Society, 2009.
- [24] H. L. Bodlaender, B. M. P. Jansen, and S. Kratsch. Preprocessing for treewidth: A combinatorial analysis through kernelization. *SIAM Journal on Discrete Mathematics*, 27(4):2108–2142, 2013.
- [25] H. L. Bodlaender, B. M. P. Jansen, and S. Kratsch. Kernelization lower bounds by cross-composition. *SIAM Journal on Discrete Mathematics*, 28(1):277–305, 2014.
- [26] H. L. Bodlaender and B. van Antwerpen-de Fluiter. Reduction algorithms for graphs of small treewidth. *Information and Computation*, 167(2):86–119, 2001.

- [27] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures & Algorithms*, 31(1):3–122, 2007.
- [28] B. Bollobás and O. Riordan. The diameter of a scale-free random graph. *Combinatorica*, 18(3):279–290, 2001.
- [29] B. Bollobás and O. Riordan. Mathematical results on scale-free random graphs. *Handbook of graphs and networks: from the genome to the internet*, pages 1–34, 2003.
- [30] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády. The degree sequence of a scale-free random graph process. *Random Structures and Algorithms*, 24(1):5–34, 2004.
- [31] B. Bollobás and A. Thomason. Threshold functions. *Combinatorica*, 7(1):35–38, 1987.
- [32] B. Bollobás and A. Thomason. Proof of a conjecture of Mader, Erdős and Hajnal on topological complete subgraphs. *European Journal of Combinatorics*, 19(8):883–887, 1998.
- [33] S. Borgatti. Identifying sets of key players in a social network. *Computational & Mathematical Organization Theory*, 12(1):21–34, 2006.
- [34] U. Brandes. A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, 25(2):163–177, 2001.
- [35] M. Brinkmeier and T. Schank. Network statistics. In *Network Analysis*, volume 3418 of *Lecture Notes in Computer Science*, pages 293–317. Springer, 2005.
- [36] T. Britton, M. Deijfen, A. Lagerås, and M. Lindholm. Epidemics on random graphs with tunable clustering. *Journal of Applied Probability*, 45(3):743–756, 2008.
- [37] D. Bu, Y. Zhao, L. Cai, H. Xue, X. Zhu, H. Lu, J. Zhang, S. Sun, L. Ling, N. Zhang, et al. Topological structure analysis of the protein–protein interaction network in budding yeast. *Nucleic acids research*, 31(9):2443–2450, 2003.
- [38] J. Bu, C. Chen, X. Liu, M. Song, L. Zhang, and Q. Zhao. Probabilistic graphlet transfer for photo cropping. *Transactions on Image Processing*, 22(2):802–815, 2013.
- [39] B. Bui-Xuan, J. Telle, and M. Vatshelle. Boolean-width of graphs. *Theoretical Computer Science*, 412(39):5187–5204, 2011.
- [40] G. Caldarelli, A. Capocci, P. D. L. Rios, and M. Muñoz. Scale-free networks from varying vertex intrinsic fitness. *Physical Review Letters*, 89(25):258702, 2002.

- [41] D. Chakrabarti, Y. Zhan, and C. Faloutsos. R-MAT: A recursive model for graph mining. In *Proceedings of the Fourth SIAM International Conference on Data Mining (ICDM)*, pages 442–446. SIAM, 2004.
- [42] C. Chekuri and J. Chuzhoy. Polynomial bounds for the grid-minor theorem. In *Proceedings of the 46th ACM Symposium on Theory of Computing (STOC)*, pages 60–69, 2014.
- [43] N. Chiba, T. Nishizeki, and N. Saito. Applications of the lipton and tarjan’s planar separator theorem. *Journal of Information Processing*, 4(4):203–207, 1981.
- [44] E. Cho, S. A. Myers, and J. Leskovec. Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pages 1082–1090. ACM, 2011.
- [45] M. Chudnovsky and P. Seymour. The structure of claw-free graphs. *Surveys in combinatorics*, 327:153–171, 2005.
- [46] F. Chung and L. Lu. The average distances in random graphs with given expected degrees. *Proceedings of the National Academy of Sciences*, 99(25):15879–15882, 2002.
- [47] F. Chung and L. Lu. Connected components in random graphs with given expected degree sequences. *Annals of combinatorics*, 6(2):125–145, 2002.
- [48] F. Chung and L. Lu. The small world phenomenon in hybrid power law graphs. In *Complex networks*, pages 89–104. Springer, 2004.
- [49] A. Clauset, C. Shalizi, and M. Newman. Power-law distributions in empirical data. *SIAM review*, 51(4):661–703, 2009.
- [50] D. Coppersmith, D. Gamarnik, and M. Sviridenko. The diameter of a long range percolation graph. In *Proceedings of the 13th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 329–337. SIAM, 2002.
- [51] B. Courcelle. The monadic second-order theory of graphs. I. Recognizable sets of finite graphs. *Information and Computation*, 85:12–75, 1990.
- [52] B. Courcelle, J. Engelfriet, and G. Rozenberg. Handle-rewriting hypergraph grammars. *Journal of Computer and System Sciences*, 46(2):218–270, 1993.
- [53] B. Courcelle, J. A. Makowsky, and U. Rotics. Linear time solvable optimization problems on graphs of bounded clique-width. *Theory of Computing Systems*, 33(2):125–150, 2000.

- [54] M. Cygan, F. V. Fomin, L. Kowalik, D. Lokshtanov, D. Marx, M. Pilipczuk, M. Pilipczuk, and S. Saurabh. *Parameterized Algorithms*. Springer, 2015.
- [55] M. Cygan, M. Pilipczuk, M. Pilipczuk, and J. Wojtaszczyk. Kernelization hardness of connectivity problems in d -degenerate graphs. *Discrete Applied Mathematics*, 160(15):2131–2141, 2012.
- [56] A. Davis, B. G. Bradford, and M. Gardner. *Deep south*. University of Chicago Press, 1969.
- [57] A. Dawar. Homomorphism preservation on quasi-wide classes. *Journal of Computer and System Sciences*, 76(5):324–332, 2010.
- [58] A. Dawar, M. Grohe, and S. Kreutzer. Locally excluding a minor. In *Proceedings of the 22nd Annual IEEE Symposium on Logic in Computer Science (LICS)*, pages 270–279. IEEE Computer Society, 2007.
- [59] A. Dawar, M. Grohe, S. Kreutzer, and N. Schweikardt. Approximation schemes for first-order definable optimisation problems. In *Proceedings of the 21th Annual IEEE Symposium on Logic in Computer Science (LICS)*, pages 411–420. IEEE Computer Society, 2006.
- [60] A. Dawar and S. Kreutzer. Domination problems in nowhere-dense classes. In *Proceedings of the 29th Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, volume 4 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 157–168. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2009.
- [61] E. D. Demaine, F. V. Fomin, M. Hajiaghayi, and D. M. Thilikos. Subexponential parameterized algorithms on bounded-genus graphs and H -minor-free graphs. *Journal of the ACM*, 52(6):866–893, 2005.
- [62] E. D. Demaine and M. Hajiaghayi. The bidimensionality theory and its algorithmic applications. *The Computer Journal*, 51(3):292–302, 2008.
- [63] E. D. Demaine and M. Hajiaghayi. Linearity of grid minors in treewidth with applications through bidimensionality. *Combinatorica*, 28(1):19–36, 2008.
- [64] E. D. Demaine, M. Hajiaghayi, and K. Kawarabayashi. Decomposition, approximation, and coloring of odd-minor-free graphs. In *Proceedings of the 21st ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 329–344. SIAM, 2010.

- [65] E. D. Demaine, F. Reidl, P. Rossmanith, F. Sánchez Villaamil, S. Sikdar, and B. D. Sullivan. Structural sparsity of complex networks: Bounded expansion in random models and real-world graphs. *CoRR*, 2014.
- [66] M. DeVos, G. Ding, B. Oporowski, D. Sanders, B. Reed, P. Seymour, and D. Vertigan. Excluding any graph as a minor allows a low tree-width 2-coloring. *Journal of Combinatorial Theory, Series B*, 91(1):25–41, 2004.
- [67] R. Diestel. *Graph Theory, 4th Edition*, volume 173 of *Graduate texts in mathematics*. Springer, 2012.
- [68] G. Ding. Excluding a long double path minor. *Journal of Combinatorial Theory, Series B*, 66(1):11–23, 1996.
- [69] M. Dom, D. Lokshтанov, and S. Saurabh. Incompressibility through colors and ids. In *Proceedings of the 36th International Colloquium on Automata, Languages, and Programming (ICALP)*, number 5555 in *Lecture Notes in Computer Science*, pages 378–389. Springer, 2009.
- [70] F. Dorn, F. V. Fomin, and D. M. Thilikos. Subexponential parameterized algorithms. *Computer Science Review*, 2(1):29–39, 2008.
- [71] Y. Dourisboure and C. Gavaille. Tree-decompositions with bags of small diameter. *Discrete Mathematics*, 307(16):2008–2029, 2007.
- [72] R. Downey. Parameterized complexity for the skeptic. *IEEE 27th Conference on Computational Complexity*, 0:147, 2003.
- [73] R. G. Downey and M. R. Fellows. Fixed parameter tractability and completeness. In *Complexity Theory: Current Research, Dagstuhl Workshop.*, pages 191–225. Cambridge University Press, 1992.
- [74] R. G. Downey and M. R. Fellows. *Parameterized Complexity*. Monographs in Computer Science. Springer, 1999.
- [75] M. Draief and A. Ganesh. Efficient routing in poisson small-world networks. *Journal of Applied Probability*, pages 678–686, 2006.
- [76] P. G. Drange, M. S. Dregi, F. V. Fomin, S. Kreutzer, D. Lokshтанov, M. Pilipczuk, M. Pilipczuk, F. Reidl, S. Saurabh, F. Sánchez Villaamil, and S. Sikdar. Kernelization and sparseness: the case of dominating set. *CoRR*, abs/1411.4575, 2014.
- [77] A. Drucker. New limits to classical and quantum instance compression. In *Proceedings of the 53rd IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 609–618. IEEE Computer Society, 2012.

- [78] Z. Dvořák. *Asymptotical Structure of Combinatorial Objects*. PhD thesis, Charles University, Faculty of Mathematics and Physics, 2007.
- [79] Z. Dvořák. Constant-factor approximation of the domination number in sparse graphs. *European Journal of Combinatorics*, 34(5):833–840, 2013.
- [80] Z. Dvořák and D. Král. Algorithms for classes of graphs with bounded expansion. In *Proceedings of the 35th International Workshop on Graph-Theoretic Concepts in Computer Science (WG)*, number 5911 in *Lecture Notes in Computer Science*, pages 17–32. Springer, 2009.
- [81] Z. Dvořák, D. Král, and R. Thomas. Deciding first-order properties for sparse graphs. In *Proceedings of the 51st IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 133–142. IEEE Computer Society, 2010.
- [82] H. Ebbinghaus and J. Flum. *Finite model theory*. Perspectives in Mathematical Logic. Springer, 1995.
- [83] E. Eiben, R. Ganian, and S. Szeider. Meta-kernelization using well-structured modulators. *CoRR*, abs/1507.05544, 2015.
- [84] D. Eppstein. Diameter and treewidth in minor-closed graph families. *Algorithmica*, 27(3):275–291, 2000.
- [85] D. Eppstein, M. Löffler, and D. Strash. Listing all maximal cliques in sparse graphs in near-optimal time. In *Proceedings of the 21st International Symposium on Algorithms and Computation (ISAAC)*, number 6506 in *Lecture Notes in Computer Science*, pages 403–414. Springer, 2010.
- [86] P. Erdős and A. Rényi. On random graphs I. *Publicationes Mathematicae Debrecen*, 6:290–297, 1959.
- [87] P. Erdős and A. Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5:17–61, 1960.
- [88] S. Eubank, H. Guclu, V. A. Kumar, M. Marathe, A. Srinivasan, Z. Toroczkai, and N. Wang. Modelling disease outbreaks in realistic urban social networks. *Nature*, 429(6988):180–184, 2004.
- [89] M. Farrel, T. D. Goodrich, N. Lemons, F. Reidl, F. Sánchez Villamil, and B. D. Sullivan. Hyperbolicity, degeneracy, and expansion of random intersection graphs. Accepted at the 12th Workshop on Algorithms and Models for the Web-graph (WAW).
- [90] R. Faudree, E. Flandrin, and Z. Ryjáček. Claw-free graphs—a survey. *Discrete Mathematics*, 164(1):87–147, 1997.

- [91] M. R. Fellows, F. V. Fomin, D. Lokshtanov, F. A. Rosamond, S. Saurabh, and Y. Villanger. Local search: Is brute-force avoidable? *Journal of Computer and System Sciences*, 78(3):707–719, 2012.
- [92] M. R. Fellows, D. Hermelin, and F. A. Rosamond. Well quasi orders in subclasses of bounded treewidth graphs and their algorithmic applications. *Algorithmica*, 64(1):3–18, 2012.
- [93] M. R. Fellows, F. A. Rosamond, U. Rotics, and S. Szeider. Clique-width is NP-complete. *SIAM Journal on Discrete Mathematics*, 23(2):909–939, 2009.
- [94] G. Fertin, A. Raspaud, and B. A. Reed. On coloring star of graphs. In *Proceedings of the 27th International Workshop on Graph-Theoretic Concepts in Computer Science (WG)*, volume 2204 of *Lecture Notes in Computer Science*, pages 140–153. Springer, 2001.
- [95] J. A. Fill, E. R. Scheinerman, and K. Singer-Cohen. Random intersection graphs when $m = \omega(n)$: An equivalence theorem relating the evolution of the $G(n, m, p)$ and $G(n, p)$ models. *Random Structures & Algorithms*, 16(2):156–176, 2000.
- [96] J. Flum and M. Grohe. Fixed-parameter tractability, definability, and model-checking. *SIAM J. Comput.*, 31(1):113–145, 2001.
- [97] J. Flum and M. Grohe. *Parameterized Complexity Theory (Texts in Theoretical Computer Science. An EATCS Series)*. Springer, 2006.
- [98] F. V. Fomin, D. Lokshtanov, N. Misra, M. S. Ramanujan, and S. Saurabh. Solving d -SAT via backdoors to small treewidth. In *Proceedings of the 26th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 630–641. SIAM, 2015.
- [99] F. V. Fomin, D. Lokshtanov, N. Misra, and S. Saurabh. Planar \mathcal{F} -deletion: Approximation, kernelization and optimal FPT algorithms. In *Proceedings of the 53rd IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 470–479. IEEE Computer Society, 2012.
- [100] F. V. Fomin, D. Lokshtanov, V. Raman, and S. Saurabh. Bidimensionality and EPTAS. In *Proceedings of the 22nd ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 748–759. SIAM, 2011.
- [101] F. V. Fomin, D. Lokshtanov, S. Saurabh, and D. M. Thilikos. Bidimensionality and kernels. In *Proceedings of the 21st ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 503–510. SIAM, 2010.
- [102] F. V. Fomin, D. Lokshtanov, S. Saurabh, and D. M. Thilikos. Linear kernels for (connected) dominating set on H -minor-free

- graphs. In *Proceedings of the 23rd ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 82–93. SIAM, 2012.
- [103] F. V. Fomin, D. Lokshtanov, S. Saurabh, and D. M. Thilikos. Linear kernels for (connected) dominating set on graphs with excluded topological subgraphs. In *Proceedings of the 30th Symposium on Theoretical Aspects of Computer Science (STACS)*, volume 20 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 92–103. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2013.
- [104] F. V. Fomin, S. Oum, and D. M. Thilikos. Rank-width and tree-width of h-minor-free graphs. *European Journal of Combinatorics*, 31(7):1617–1628, 2010.
- [105] L. R. Ford and D. R. Fulkerson. Maximal flow through a network. *Canadian Journal of Mathematics*, 8(3):399–404, 1956.
- [106] N. Fountoulakis, D. Kühn, and D. Osthus. The order of the largest complete minor in a random graph. 33:127–141, 2008.
- [107] P. Fraigniaud. Greedy routing in tree-decomposed graphs: a new perspective on the small-world phenomenon. In *Proceedings of the 13th European Symposium on Algorithms (ESA)*, number 3669 in *Lecture Notes in Computer Science*, pages 791–802. Springer, 2005.
- [108] M. Franceschetti and R. Meester. Navigation in small-world networks: A scale-free continuum model. *Journal of Applied Probability*, pages 1173–1180, 2006.
- [109] L. C. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, pages 35–41, 1977.
- [110] L. C. Freeman. Finding social groups: A meta-analysis of the southern women data. *Dynamic social network modeling and analysis*, pages 39–97, 2003.
- [111] L. C. Freeman. The development of social network analysis. *A Study in the Sociology of Science*, 2004.
- [112] M. Frick and M. Grohe. Deciding first-order properties of locally tree-decomposable structures. *Journal of the ACM*, 48(6):1184–1206, 2001.
- [113] H. Gaifman. On local and non-local properties. In *Proceedings of the Herbrand Symposium, Logic Colloquium*, volume 81, pages 105–135, 1982.
- [114] J. Gajarský, P. Hliněný, J. Obdržálek, S. Ordyniak, F. Reidl, P. Rossmanith, F. Sánchez Villaamil, and S. Sikdar. Kernelization using structural parameters on sparse graph classes. In

- Proceedings of the 21st European Symposium on Algorithms (ESA)*, number 8125 in *Lecture Notes in Computer Science*, pages 529–540. Springer, 2013.
- [115] J. Gajarský, P. Hliněný, J. Obdržálek, S. Ordyniak, F. Reidl, P. Rossmanith, F. Sánchez Villaamil, and S. Sikdar. Finite integer index of pathwidth and treewidth. In *Proceedings of the 9th International Workshop on Parameterized and Exact Computation (IWPEC)*, number 8894 in *Lecture Notes in Computer Science*, pages 258–269. Springer, 2014.
- [116] J. Gajarský, P. Hliněný, J. Obdržálek, S. Ordyniak, F. Reidl, P. Rossmanith, F. Sánchez Villaamil, and S. Sikdar. Kernelization using structural parameters on sparse graph classes. *To appear in Journal of Computer and System Sciences*, 2015.
- [117] J. Gajarský, M. Lampis, and S. Ordyniak. Parameterized algorithms for modular-width. In *Parameterized and Exact Computation*, pages 163–176. Springer, 2013.
- [118] R. Ganian, P. Hliněný, J. Nešetřil, J. Obdržálek, P. Ossona de Mendez, and R. Ramadurai. When trees grow low: Shrubs and fast MSO₁. In *Mathematical Foundations of Computer Science 2012*, number 7464 in *Lecture Notes in Computer Science*, pages 419–430. Springer, 2012.
- [119] R. Ganian, F. Slivovsky, and S. Szeider. Meta-kernelization with structural parameters. In *Proceedings of the 38th Conference on Mathematical Foundations of Computer Science (MFCS)*, number 8087 in *Lecture Notes in Computer Science*, pages 457–468. Springer, 2013.
- [120] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.
- [121] V. Garnero, C. Paul, I. Sau, and D. M. Thilikos. Explicit linear kernels via dynamic programming. In *Proceedings of the 31st Symposium on Theoretical Aspects of Computer Science (STACS)*, volume 25 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 312–324. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2014.
- [122] A. C. Giannopoulou, B. M. P. Jansen, D. Lokshantov, and S. Saurabh. Uniform kernelization complexity of hitting forbidden minors. volume 9134 of *Lecture Notes in Computer Science*, pages 629–641. Springer, 2015.
- [123] E. N. Gilbert. Random graphs. *Annals of Mathematical Statistics*, 30(4):1141–1144, 12 1959.

- [124] M. Girvan and M. Newman. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821–7826, 2002.
- [125] K. Goh, M. Cusick, D. Valle, B. Childs, M. Vidal, and A. Barabási. The human disease network. *Proceedings of the National Academy of Sciences*, 104(21):8685–8690, 2007.
- [126] P. A. Golovach and Y. Villanger. Parameterized complexity for domination problems on degenerate graphs. In *Proceedings of the 34th International Workshop on Graph-Theoretic Concepts in Computer Science (WG)*, number 5344 in Lecture Notes in Computer Science, pages 195–205. Springer, 2008.
- [127] M. Grohe. Local tree-width, excluded minors, and approximation algorithms. *Combinatorica*, 23(4):613–632, 2003.
- [128] M. Grohe, S. Kreutzer, and S. Siebertz. Characterisations of nowhere dense graphs (invited talk). In *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2013*, volume 24 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 21–40. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2013.
- [129] M. Grohe, S. Kreutzer, and S. Siebertz. Deciding first-order properties of nowhere dense graphs. In *Proceedings of the 46th ACM Symposium on Theory of Computing (STOC)*, pages 89–98, 2014.
- [130] M. Grohe and D. Marx. Structure theorem and isomorphism test for graphs with excluded topological subgraphs. In *Proceedings of the 44th ACM Symposium on Theory of Computing (STOC)*, pages 173–192, 2012.
- [131] Q. Gu and H. Tamaki. Optimal branch-decomposition of planar graphs in $O(n^3)$ time. *ACM Transactions on Algorithms*, 4(3), 2008.
- [132] Q. Gu and H. Tamaki. Constant-factor approximations of branch-decomposition and largest grid minor of planar graphs in $O(n^{1+\epsilon})$ time. In *Proceedings of the 20th International Symposium on Algorithms and Computation (ISAAC)*, number 5878 in Lecture Notes in Computer Science, pages 984–993. Springer, 2009.
- [133] J. Guo and R. Niedermeier. Linear problem kernels for NP-hard problems on planar graphs. In *Proceedings of the 34th International Colloquium on Automata, Languages, and Programming (ICALP)*, number 4596 in Lecture Notes in Computer Science, pages 375–386. Springer, 2007.
- [134] M. Habib, C. Paul, and L. Viennot. A synthesis on partition refinement: A useful routine for strings, graphs, boolean matrices and automata. In *Proceedings of the 25th Symposium on Theoretical*

- Aspects of Computer Science (STACS)*, volume 08001 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 25–38. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2008.
- [135] P. Hage and F. Harary. Eccentricity and centrality in networks. *Social networks*, 17(1):57–63, 1995.
- [136] R. Halin. S-functions for graphs. *Journal of Geometry*, 8(1-2):171–186, 1976.
- [137] J. J. Harris, R. Jolivet, and D. Attwell. Synaptic energy use and supply. *Neuron*, 75(5):762–777, 2012.
- [138] T. Harris and F. Ross. Fundamentals of a method for evaluating rail net capacities. Technical report, DTIC Document, 1955.
- [139] W. Hayes, K. Sun, and N. Pržulj. Graphlet-based measures are suitable for biological network comparison. *Bioinformatics*, 29(4):483–491, 2013.
- [140] P. Holland, K. Laskey, and S. Leinhardt. Stochastic blockmodels: First steps. *Social networks*, 5(2):109–137, 1983.
- [141] S. Itzkovitz, R. Levitt, N. Kashtan, R. Milo, M. Itzkovitz, and U. Alon. Coarse-graining and self-dissimilarity of complex networks. *Physical Review E*, 71(1):016127, 2005.
- [142] J. J. Wu and H. Li. On calculating connected dominating set for efficient routing in ad hoc wireless networks. In *Proceedings of the 3rd international workshop on Discrete algorithms and methods for mobile computing and communications*, pages 7–14. ACM, 1999.
- [143] P. Jaccard. Étude comparative de la distribution florale dans une portion des alpes et des jura. *Bulletin del la Société Vaudoise des Sciences Naturelles*, 37:547–579, 1901.
- [144] B. M. P. Jansen. *The Power of Data Reduction: Kernels for Fundamental Graph Problems*. PhD thesis, Utrecht University, 2013.
- [145] B. M. P. Jansen. Turing kernelization for finding long paths and cycles in restricted graph classes. In *Proceedings of the 22nd European Symposium on Algorithms (ESA)*, number 8737 in Lecture Notes in Computer Science, pages 579–591. Springer, 2014.
- [146] B. M. P. Jansen and S. Kratsch. A structural approach to kernels for ILPs: Treewidth and total unimodularity. *CoRR*, abs/1506.07729, 2015.
- [147] S. Janson. The probability that a random multigraph is simple. II. *Journal of Applied Probability*, 51:123–137, 2014.

- [148] T. Jiang. Compact topological minors in graphs. *Journal of Graph Theory*, 67(2):139–152, 2011.
- [149] C. Jordan. Sur les assemblages de lignes. *J. Reine Angew. Math*, 70(185):81, 1869.
- [150] F. Karinthy. Chain-links. *Everything is the Other Way*, 1929. English translation by Adam Makkai.
- [151] M. S. E. Karoński and K. Singer-Cohen. On random intersection graphs: the subgraph problem. *Combinatorics, Probability and Computing*, 8:131–159, 1999.
- [152] K. Kawarabayashi and Y. Kobayashi. Linear min-max relation between the treewidth of h-minor-free graphs and its largest grid. In *Proceedings of the 29th Symposium on Theoretical Aspects of Computer Science (STACS)*, volume 14 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 278–289. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2012.
- [153] W. Kazana and L. Segoufin. Enumeration of first-order queries on classes of structures with bounded expansion. In *Proceedings of the 32nd ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS)*, pages 297–308. ACM, 2013.
- [154] H. A. Kierstead and W. T. Trotter. Planar graph coloring with an uncooperative partner. *Journal of Graph Theory*, 18(6):569–584, 1994.
- [155] H. A. Kierstead and D. Yang. Orderings on graphs and game coloring number. *Order*, 20(3):255–264, 2003.
- [156] E. J. Kim, A. Langer, C. Paul, F. Reidl, P. Rossmanith, I. Sau, and S. Sikdar. Linear kernels and single-exponential algorithms via protrusion-decompositions. In *Proceedings of the 40th International Colloquium on Automata, Languages, and Programming (ICALP) Part I*, number 7965 in *Lecture Notes in Computer Science*, pages 613–624. Springer, 2013.
- [157] E. J. Kim, A. Langer, C. Paul, F. Reidl, P. Rossmanith, I. Sau, and S. Sikdar. Linear kernels and single-exponential algorithms via protrusion decompositions. *ACM Trans. Algorithms*, 12(2):21, 2016.
- [158] J. Kleinberg. The small-world phenomenon: An algorithmic perspective. In *Proceedings of the 32rd ACM Symposium on Theory of Computing (STOC)*, pages 163–170, 2000.
- [159] B. Klimt and Y. Yang. Introducing the Enron corpus. In *CEAS*, 2004.
- [160] D. Knuth. *The Stanford GraphBase: a platform for combinatorial computing*, volume 37. Addison-Wesley Reading, 1993.

- [161] J. Komlós and E. Szemerédi. Topological cliques in graphs II. *Combinatorics, Probability and Computing*, 5(01):79–90, 1996.
- [162] D. Koschützki, K. A. Lehmann, L. Peeters, S. Richter, D. Tenfelde-Podehl, and O. Zlotowski. Centrality indices. In *Network analysis*, volume 3418 of *Lecture Notes in Computer Science*, pages 16–61. Springer, 2005.
- [163] V. Krebs. <http://orgnet.com/divided.html>.
- [164] S. Kreutzer and A. Dawar. Parameterized complexity of first-order logic. *Electronic Colloquium on Computational Complexity (ECCC)*, 16:131, 2009.
- [165] S. Kreutzer and S. Tazari. Lower bounds for the complexity of monadic second-order logic. In *Proceedings of the 25th Annual IEEE Symposium on Logic in Computer Science (LICS)*, pages 189–198. IEEE Computer Society, 2010.
- [166] O. Kuchaiev, T. Milenković, V. Memišević, W. Hayes, and N. Pržulj. Topological network alignment uncovers biological function and phylogeny. *Journal of the Royal Society Interface*, 7(50):1341–1354, 2010.
- [167] A. Langer. Schnelle algorithmen für zerlegbare graphen. In *Ausgezeichnete Informatikdissertationen 2013*, volume D-14 of *LNI*, pages 101–110. GI, 2013.
- [168] A. Langer, F. Reidl, P. Rossmanith, and S. Sikdar. Evaluation of an MSO-solver. In *Proceedings of the 14th Meeting on Algorithm Engineering & Experiments (ALENEX)*, pages 55–63. SIAM / Omnipress, 2012.
- [169] A. Langer, F. Reidl, P. Rossmanith, and S. Sikdar. Practical algorithms for MSO model-checking on tree-decomposable graphs. *Computer Science Review*, 13-14:39–74, 2014.
- [170] A. Leaf and P. Seymour. Tree-width and planar minors. *Journal of Combinatorial Theory, Series B*, 111:38–53, 2015.
- [171] J. Leskovec, J. Kleinberg, and C. Faloutsos. Graphs over time: densification laws, shrinking diameters and possible explanations. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pages 177–187. ACM, 2011.
- [172] J. Leskovec and A. Krevl. SNAP Datasets: Stanford large network dataset collection. <http://snap.stanford.edu/data>, June 2014.
- [173] N. Lin. *Foundations of Social Research*. McGraw-Hill New York, 1976.

- [174] R. Lipton and R. Tarjan. A separator theorem for planar graphs. *SIAM Journal on Applied Mathematics*, 36(2):177–189, 1979.
- [175] R. Lipton and R. Tarjan. Applications of a planar separator theorem. *SIAM Journal on Computing*, 9(3):615–627, 1980.
- [176] D. Lusseau, K. Schneider, O. Boisseau, P. Haase, E. Slooten, and S. Dawson. The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations. *Behavioral Ecology and Sociobiology*, 54(4):396–405, 2003.
- [177] W. Mader. Homomorphieeigenschaften und mittlere kantendichte von graphen. *Mathematische Annalen*, 174(4):265–268, 1967.
- [178] P. V. Marsden. Egocentric and sociocentric measures of network centrality. *Social Networks*, 24(4):407–422, 2002.
- [179] D. Marx. Chordal deletion is fixed-parameter tractable. *Algorithmica*, 57(4):747–768, 2010.
- [180] A. Masoudi-Nejad, F. Schreiber, and Z. R. M. Kashani. Building blocks of biological networks: a review on major network motif discovery algorithms. *IET Systems Biology*, 6:164–174, 2012.
- [181] D. W. Matula and L. L. Beck. Smallest-last ordering and clustering and graph coloring algorithms. *Journal of the ACM*, 30(3):417–427, 1983.
- [182] T. Milenković, V. Memišević, A. K. Ganesan, and N. Pržulj. Systems-level cancer gene identification from protein interaction network topology applied to melanogenesis-related functional genomics data. *Journal of the Royal Society Interface*, 7(44):423–437, 2010.
- [183] T. Milenković and N. Pržulj. Uncovering biological network function via graphlet degree signatures. *Cancer Informatics*, 6(6):257, 2008.
- [184] S. Milgram. The small-world problem. *Psychology Today*, 61(1), 1967.
- [185] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824–827, 2002.
- [186] M. Molloy and B. A. Reed. A critical point for random graphs with a given degree sequence. *Random Structures & Algorithms*, 6(2/3):161–180, 1995.
- [187] J. Nešetřil and P. Ossona de Mendez. Tree-depth, subgraph coloring and homomorphism bounds. *European Journal of Combinatorics*, 27(6):1022–1041, 2006.

- [188] J. Nešetřil and P. Ossona de Mendez. Grad and classes with bounded expansion I. decompositions. *European Journal of Combinatorics*, 29(3):760–776, 2008.
- [189] J. Nešetřil and P. Ossona de Mendez. Grad and classes with bounded expansion II. algorithmic aspects. *European Journal of Combinatorics*, 29(3):777–791, 2008.
- [190] J. Nešetřil and P. Ossona de Mendez. First order properties on nowhere dense structures. *J. Symb. Log.*, 75(3):868–887, 2010.
- [191] J. Nešetřil and P. Ossona de Mendez. On nowhere dense graphs. *European Journal of Combinatorics*, 32(4):600–617, 2011.
- [192] J. Nešetřil and P. Ossona de Mendez. *Sparsity: Graphs, Structures, and Algorithms*, volume 28 of *Algorithms and Combinatorics*. Springer, 2012.
- [193] J. Nešetřil, P. Ossona de Mendez, and D. R. Wood. Characterisations and examples of graph classes with bounded expansion. *European Journal of Combinatorics*, 33(3):350–373, 2012.
- [194] M. Newman. The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences*, 98(2):404–409, 2001.
- [195] M. Newman. The structure and function of complex networks. *SIAM review*, 45(2):167–256, 2003.
- [196] M. Newman. Finding community structure in networks using the eigenvectors of matrices. *Physical Review E*, 74(3):036104, 2006.
- [197] M. Newman and D. Watts. Renormalization group analysis of the small-world network model. *Physics Letters A*, 263(4):341–346, 1999.
- [198] M. Newman, D. Watts, and S. Strogatz. Random graph models of social networks. *Proceedings of the National Academy of Sciences*, 99(suppl 1):2566–2572, 2002.
- [199] T. Opsahl, F. Agneessens, and J. Skvoretz. Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, 32(3):245–251, 2010.
- [200] S. Oum. Rank-width and vertex-minors. *Journal of Combinatorial Theory, Series B*, 95(1):79–100, 2005.
- [201] S. Oum and P. D. Seymour. Approximating clique-width and branch-width. *Journal of Combinatorial Theory, Series B*, 96(4):514–528, 2006.

- [202] P. Pantazopoulos, M. Karaliopoulos, and I. Stavrakakis. On the local approximations of node centrality in internet router-level topologies. In *Proceedings of the 7th IFIP International Workshop on Self-Organizing Systems IWSOS*, pages 1–12, 2013.
- [203] A. Phadke and J. Thorp. *Computer relaying for power systems*. John Wiley & Sons, 2009.
- [204] G. Philip, V. Raman, and S. Sikdar. Polynomial kernels for dominating set in graphs of bounded degeneracy and beyond. *ACM Transactions on Algorithms*, 9(1):11, 2012.
- [205] M. Pilipczuk, M. Pilipczuk, P. Sankowski, and E. J. van Leeuwen. Subexponential-time parameterized algorithm for steiner tree on planar graphs. In *Proceedings of the 30th Symposium on Theoretical Aspects of Computer Science (STACS)*, volume 20 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 353–364. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2013.
- [206] M. Pilipczuk, M. Pilipczuk, P. Sankowski, and E. J. van Leeuwen. Network sparsification for steiner problems on planar and bounded-genus graphs. In *Proceedings of the 55th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 276–285. IEEE Computer Society, 2014.
- [207] P. Pongpailool, S. Pukkawanna, and V. Visoottiviseth. Lightweight detection of DoS attacks. In *Proceedings of the 15th IEEE Conference on Networks (ICON)*, pages 77–82. IEEE, 2007.
- [208] N. Pržulj. Biological network comparison using graphlet degree distribution. *Bioinformatics*, 23(2):177–183, 2007.
- [209] N. Pržulj, D. Corneil, and I. Jurisica. Modeling interactome: scale-free or geometric? *Bioinformatics*, 20(18):3508–3515, 2004.
- [210] P. Ribeiro, F. Silva, and M. Kaiser. Strategies for network motifs discovery. In *Proceedings of the 5th International Conference on E-Science*, pages 80–87. IEEE, 2009.
- [211] P. M. P. Ribeiro and F. M. A. da Silva. G-tries: a data structure for storing and finding subgraphs. *Data Mining and Knowledge Discovery*, 28(2):337–377, 2014.
- [212] M. Ripeanu, A. Iamnitchi, and I. T. Foster. Mapping the gnutella network. *IEEE Internet Computing*, 6(1):50–57, 2002.
- [213] N. Robertson and P. Seymour. Graph minors. V. excluding a planar graph. *Journal of Combinatorial Theory, Series B*, 41(1):92–114, 1986.
- [214] N. Robertson and P. Seymour. Graph minors. X. obstructions to tree-decomposition. *Journal of Combinatorial Theory, Series B*, 52(2):153–190, 1991.

- [215] N. Robertson and P. Seymour. Graph minors. xiii. the disjoint paths problem. *Journal of combinatorial theory, Series B*, 63(1):65–110, 1995.
- [216] N. Robertson and P. Seymour. Graph minors. XVI. excluding a non-planar graph. *Journal of Combinatorial Theory, Series B*, 89(1):43–76, 2003.
- [217] N. Robertson and P. Seymour. Graph minors. XX. wagner’s conjecture. *Journal of Combinatorial Theory, Series B*, 92(2):325–357, 2004.
- [218] N. Robertson and P. Seymour. Graph minors XXIII. nash-williams’ immersion conjecture. *Journal of Combinatorial Theory, Series B*, 100(2):181–205, 2010.
- [219] N. Robertson, P. Seymour, and R. Thomas. Quickly excluding a planar graph. *Journal of Combinatorial Theory, Series B*, 62(2):323–348, 1994.
- [220] G. Sabidussi. The centrality index of a graph. *Psychometrika*, 31(4):581–603, 1966.
- [221] A. Schrijver. On the history of the transportation and maximum flow problems. *Mathematical Programming*, 91(3):437–445, 2002.
- [222] D. Seese. Linear time computable problems and first-order descriptions. *Mathematical Structures in Computer Science*, 6(6):505–526, 1996.
- [223] C. Seshadhri, A. Pinar, and T. G. Kolda. The similarity between stochastic kronecker and chung-lu graph models. In *Proceedings of the 12th SIAM International Conference on Data Mining (ICDM)*, pages 1071–1082. SIAM, 2012.
- [224] C. Seshadhri, A. Pinar, and T. G. Kolda. An in-depth analysis of stochastic kronecker graphs. *Journal of the ACM*, 60(2):13, 2013.
- [225] P. D. Seymour and R. Thomas. Call routing and the ratcatcher. *Combinatorica*, 14(2):217–241, 1994.
- [226] N. Shervashidze, S. V. N. Vishwanathan, T. Petri, K. Mehlhorn, and K. M. Borgwardt. Efficient graphlet kernels for large graph comparison. In *Proceedings of the 12th International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 5 of *JMLR Proceedings*, pages 488–495. JMLR.org, 2009.
- [227] A. Shimbel. Structural parameters of communication networks. *The bulletin of mathematical biophysics*, 15(4):501–507, 1953.
- [228] K. Singer-Cohen. *Random Intersection Graphs*. PhD thesis, Johns Hopkins University, 1995.

- [229] R. Solomonoff and A. Rapoport. Connectivity of random nets. *The bulletin of mathematical biophysics*, 13(2):107–117, 1951.
- [230] O. Sporns and R. Kötter. Motifs in brain networks. *PLoS Biology*, 2(11):369, 2004.
- [231] J. Stehlé, N. Voirin, A. Barrat, C. Cattuto, L. Isella, J. Pinton, M. Quaggiotto, W. Van den Broeck, C. Régis, B. Lina, and P. Vanhems. High-resolution measurements of face-to-face contact patterns in a primary school. *PLOS ONE*, 6(8):e23176, 08 2011.
- [232] V. Stepanov. Phase transitions in random graphs. *Theory of Probability & Its Applications*, 15(2):187–203, 1970.
- [233] S. Takemura., me A. Bharioke, Z. Lu, A. Nern, S. Vitaladevuni, P. Rivlin, W. Katz, D. Olbris, S. Plaza, P. Winston, et al. A visual motion detection circuit suggested by drosophila connectomics. *Nature*, 500(7461):175–181, 2013.
- [234] R. Tanaka, T.-M. Yi, and J. Doyle. Some protein interaction data do not exhibit power law statistics. *FEBS letters*, 579(23):5140–5144, 2005.
- [235] S. Tazari. Faster approximation schemes and parameterized algorithms on (odd-)H-minor-free graphs. *Theoretical Computer Science*, 417:95–107, 2012.
- [236] S. Teso, J. Staiano, B. Lepri, A. Passerini, and F. Pianesi. Ego-centric graphlets for personality and affective states recognition. In *Proceedings of the International Conference on Social Computing (SocialCom) 2013*, pages 874–877. IEEE, 2013.
- [237] A. Thomason. An upper bound for some ramsey numbers. *Journal of Graph Theory*, 12(4):509–517, 1988.
- [238] A. Tolstoï. Metody nakhozheniya naimen'shego summovogo kilometrazha pri planirovanii perevozok v prostranstve [russian; methods of finding the minimal total kilometrage in cargo-transportation planning in space]. *Planirovanie Perevozok, Sbornik pervyĭ [Russian; Transportation Planning, Volume I], Transpechat'NKPS [TransPress of the National Commissariat of Transportation]*, Moscow, pages 23–55, 1930.
- [239] J. Travers and S. Milgram. An experimental study of the small world problem. *Sociometry*, pages 425–443, 1969.
- [240] J. Ugander, L. Backstrom, and J. Kleinberg. Subgraph frequencies: Mapping the empirical and extremal geography of large graph collections. In *Proceedings of the 22nd International Conference on World Wide Web (WWW)*, pages 1307–1318. IW3C2, 2013.
- [241] R. van der Hofstad. Random graphs and complex networks. <http://www.win.tue.nl/rhofstad/NotesRGCN.pdf>, 2009.

- [242] D. van Melkebeek and H. Dell. Satisfiability allows no non-trivial sparsification unless the polynomial-time hierarchy collapses. *Electronic Colloquium on Computational Complexity (ECCC)*, 17:38, 2010.
- [243] M. Vatshelle. *New width parameters of graphs*. PhD thesis, 2012.
- [244] F. Wang, H. Du, E. Camacho, K. Xu, W. Lee, Y. Shi, and S. Shan. On positive influence dominating sets in social networks. *Theoretical Computer Science*, 412(3):265–269, 2011.
- [245] Y. Wang and G. Wong. Stochastic blockmodels for directed graphs. *Journal of the American Statistical Association*, 82(397):8–19, 1987.
- [246] S. Wasserman and K. Faust. *Social network analysis: Methods and applications*, volume 8. Cambridge university press, 1994.
- [247] D. Watts and S. Strogatz. Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684):440–442, 1998.
- [248] S. Wernicke. Efficient detection of network motifs. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 3(4):347–359, 2006.
- [249] J. White, E. Southgate, J. Thomson, and S. Brenner. The structure of the nervous system of the nematode *Caenorhabditis elegans*: the mind of a worm. *Phil. Trans. R. Soc. Lond.*, 314:1–340, 1986.
- [250] W. Willinger, D. Alderson, and J. Doyle. *Mathematics and the internet: A source of enormous confusion and great potential*. Defense Technical Information Center, 2009.
- [251] G. Yule. A mathematical theory of evolution, based on the conclusions of Dr. J.C. Willis, F.R.S. *Philosophical Transactions of the Royal Society of London. Series B*, pages 21–87, 1925.
- [252] W. Zachary. An information flow model for conflict and fission in small groups. *Journal of anthropological research*, pages 452–473, 1977.
- [253] X. Zhu. Colouring graphs with bounded generalized colouring number. *Discrete Mathematics*, 309(18):5562–5568, 2009.

Part VII

INDEX

SYMBOL INDEX

- $E[\bullet]$, 9
 $G(n, \bullet) \tilde{\nabla} r$, 181
 $G[E/\ell]$, 16
 $G[\bullet]$, 9
 $G_1 * G_2$, 15
 $G_1 \bullet G_2$, 15
 K_n , 10
 $K_{i,j}$, 10
 $N(\bullet)$, 9
 $N[\bullet]$, 9
 $N^\sigma(\bullet)$, 87
 P_ℓ , 10
 \mathfrak{m} , 118
 $\Delta_\Pi(\bullet, \bullet)$, 118
 $E[\bullet]$, 181
 ${}^\circ G_1 \oplus {}^\circ G_2$, 15
 ${}^\circ G$, 15
 $M[\bullet]$, 181
 $P[\bullet]$, 181
 $\text{Var}[\bullet]$, 181
 $\bigcup \mathcal{F}$, 9
 \perp , 15
 $\chi_p(\bullet)$, 65
 $\text{clos}(T)$, 11
 $K(G)$, 9
 $\mathbf{ds}(G)$, 24
 $\mathbf{ds}(G, Z)$, 24
 $\mathcal{R}({}^\circ G)$, 124
 $\text{forb}(\mathcal{I})$, 17
 $\nabla_r(\bullet)$, 43
 \subseteq_i , 9
 $\text{clos}^{\text{lca}}(\bullet)$, 11
 $\text{leaves}(T)$, 10
 $\limsup_{G \in \mathcal{G}}$, 21
 $\limsup_{H \triangleleft G \in \mathcal{G}}$, 22
 $\mathcal{E}_{\xi, r}^X$, 193
 \mathcal{K}_r^X , 195
 $\mathcal{G}_{\leq k}$, 12
 $\mathcal{G}_{f \blacklozenge k}$, 12
 \preceq_{mv} , 14
 $\text{basis}(\mathcal{H})$, 17
 ∇ , 14
 $\#\omega(G)$, 10
 $\omega(G)$, 10
 $\omega(\mathcal{G})$, 21
 $\omega_r(\bullet)$, 48
 $\mathbf{pw}(G)$, 19
 $\text{root}(T)$, 10
 $\sigma[i]$, 87
 \preceq_{m}^r , 14
 \preceq_{t}^r , 55
 \preceq_{t}^r , 14
 $\tilde{\nabla}_r(\bullet)$, 55
 $\tilde{\nabla}_r^{\text{!}}(\bullet)$, 61
 \subseteq , 9
 ${}^\circ \mathcal{G}_t$, 15
 $\mathbf{td}(G)$, 19
 $\mathbf{td}_t^\Delta(G)$, 12
 \preceq_{b}^r , 14
 $\tilde{\nabla}_r(\bullet)$, 43
 $\tilde{\nabla}_r^{\text{!}}(\bullet)$, 61
 $\tilde{\nabla}$, 14
 $\tilde{\omega}_r(\bullet)$, 48
 $\mathbf{tw}(G)$, 19
 $\mathbf{tw}_t^\Delta(G)$, 12
 \uplus , 9
 $\mathbf{vc}(G)$, 12
 $\xi_{\min}^\Pi(\bullet)$, 120
 uPv , 10

PROBLEM INDEX

- H*-PACKING, 103
 Π -*k*-VERTEX EXCHANGE, 103
 \mathcal{F} -DELETION, 105, 120, 141, 272
 \mathcal{F} -PACKING, 120, 272
 \mathcal{H} -MINOR DELETION, 119, 146, 149, 272
 \mathcal{H} -MINOR PACKING, 119, 273
 \mathcal{P} -DELETION, 12, 273
 min- Π -VERTEX EXCHANGE, 102
 wm - w -DELETION, 117, 118, 273
k-CYCLE, 117, 123, 272
k-PATH, 23, 35, 117, 121–123, 147, 161, 272
r-CENTRE, 102
r-DOMINATING SET, 7, 24, 35, 38, 45, 49, 93, 102, 106, 116, 117, 150, 161, 178, 241
 (*k, r*)-CENTRE, 24, 34
 BRANCHWIDTH, 117, 124
 CHORDAL VERTEX DELETION, ix, 116, 117, 139, 271
 CHROMATIC NUMBER, 121, 122
 COGRAPH VERTEX DELETION, 116
 CONNECTED DOMINATING SET, 33, 38, 40, 116, 117, 149, 161
 CONNECTED FEEDBACK VERTEX SET, 116, 271
 CONNECTED VERTEX COVER, 38, 116, 117
 DOMINATING SET, x, xi, 7, 8, 22, 23, 24, 33–35, 38, 40, 84, 93, 101, 102, 106, 109, 115–117, 149, 150, 159–161, 178, 241
 DOUBLE DOMINATION, 161
 EDGE DOMINATING SET, 35, 38, 116, 117
 EFFICIENT DOMINATING SET, 117, 161
 EXACT *s, t*-PATH, 117, 123, 272
 EXACT CYCLE, 117, 123, 272
 EXACT PATH, 272
 FEEDBACK VERTEX SET, x, 12, 23, 37, 38, 116, 117, 139, 140, 149, 271
 HAMILTONIAN CYCLE, 116, 117
 HAMILTONIAN PATH, 116, 117
 INDEPENDENT DOMINATING SET, 121, 146, 271
 INDEPENDENT SET, 22, 34, 35, 116, 117, 149, 271
 INDUCED *d*-DEGREE SUBGRAPH, 117
 INDUCED MATCHING, 117
 INDUCED- \mathcal{F} -DELETION, 105, 272
 INTERVAL VERTEX DELETION, 116
 MAX FULL DEGREE SPANNING TREE, 117
 MAX-CUT, 102
 MAXIMUM CLIQUE, 22
 MAXIMUM LEAF SPANNING TREE, 38
 MIN LEAF SPANNING TREE, 117
 MIN-BISECTION, 102
 MULTIWAY CUT, 36
 ODD CYCLE TRANSVERSAL, 12, 102, 117, 271
 PARTIAL DOMINATING SET, 24, 34
 PATHWIDTH, 23, 117, 124
 PROPER INTERVAL VERTEX DELETION, 116
 QBF SATISFIABILITY, 25
 SET COVER, 38
 STEINER FOREST, 36
 STEINER TREE, 36
 SUBGRAPH ISOMORPHISM, 34, 35, 273
 TREEDEPTH, 121

- TREEDPTH- t -DELETION, 12, 146
see **wm- w -DELETION**
- TREewidth, 23, 117, 124
- TREewidth- t -DELETION, 12 see **wm- w -DELETION**
- VERTEX COVER, 12, 23, 34, 35, 38,
39, 102, 119, 139 **271**
- VERTEX \mathcal{H} -PACKING, 38
- WEIGHTED Π - k -VERTEX EXCHANGE,
104

INDEX

- p*-centred colouring, 73, 86–93
 - linear, 71
- accessible, 97
 - weakly, 97
- k*-admissibility, 97
- almost wide, 98
- ancestor, 10
- aug-aug plot, 214
- backconnectivity, 97
- bags, 19
- basis, 17
- boundaried graph, 15, 118–128
 - induced, 16
 - replacement, 16
- boundary, 15
- bounded expansion, 43
 - a.a.s./ w.h.p., 182
 - locally, 45
- branchwidth, 21, 124
- centrality measures, 173
- centre, 65
- p*-centred colouring, 65
- charging lemma, 84, 155
- child, 10
- Chung–Lu model, 189–199, 213
- class graph, 153
- clique, 10
- clique number, 10
- cliquewidth, 21
- closeness centrality, 173
 - localised, 175
- clustering coefficient, 172
 - local, 173
- CMSO, 119
 - definable, 118
- colouring, 9
 - cardinality, 9
 - proper, 9
 - refinement, 10
- k*-colouring number, 97
- complete graph, 10
- complete join, 15
- configuration model, 186, 189, 199, 213
- convergence in distribution, 181
- k*-core, 33
- $\tilde{\nabla}_r$ -critical, 53
- decision problem, 22
- degenerate, 31
- degree bound, 95
- degree centrality, 173
- degree distribution, 171
- degree sequence, 185
 - matching, 185
- depth, 10
 - stable minor, 56
- descendant, 10
- descriptive model, 179
- dissolution, 16
- dominating set, 24
 - connected, 178
 - partial, 178
 - positive, 178
- r*-dominating set, 24
- domination core, 150
- domination number, 24
 - relative, 23
- Z-dominator, 24
- dtf-augmentation, 66
- eccentricity, 174
- embedding, 13
- EMSO, 25
- excision, 16
- extended model, 56
- f.i., *see* finite index
- f.i.i., *see* finite integer index
- finite index, 118
- finite integer index, 118, 118–146
- first-order certifiable, 102
- first-order definable, 25

- first-order logic, 24
- fixed parameter tractable, 22
- forest, *see* tree
- forest fire, 204
- FPT, 22
- fpt, *see* fixed parameter tractable
- fpt-algorithm, 22
 - non-uniform, 22
 - uniform, 22
- fraternal, 66
- generalised random graphs, 188
- generative model, 179
- geometric random graphs, 210
- gluing, 15
- grad
 - topological, 43
- graph class, 12
 - truncation, 12
- graph measure, 21
 - parametrised, 21
 - polynomially related, 22
 - supremum, 21
- graph parameter
 - degree bound, 95
 - monotone, 95
 - strongly topological, 95
- graph problem, 22
- graph property, 12
- graphlet, 177
 - degree distribution, 177
- greedy routing, 206
- harmonic centrality, 173
 - localised, 175
- height, 10
- ideal, 17
- increasing pair, 17
- induced edge set, 9
- inhomogeneous random graphs, 188
- kernel, 23
- labelling, 15
- laminar, 90
- language, *see* decision problem
- least common ancestor, 11
 - closure, 11
- lexicographic product, 15
- Lin's centrality, 174
 - localised, 175
- low-treedepth colourings, 65
- minor
 - depth, 14
 - embedding, 13
 - shallow, 14
 - stable, 56
- model, 24
- model checking, 25
- modulator, 12
- monadic second-order logic, *see* MSO
- monotone, 95
- motif, 177
- MSO, 25, 119
- natural index, 9
- neighbourhood, 9
- σ -neighbourhood, 87
- neighbourhood complexity, 83, 86–93
- nowhere dense, 47, 48
 - a.a.s./ w.h.p., 182
 - locally, 49
- out-apex, 68
- parametrisation, 22
- parametrised problem, 22
- parent, 10
 - relationship, 10
- partition refinement, 12
- path decomposition, 19
- pathwidth, 19, 124
- protrusion, 128
 - adhesion, 128
 - boundary, 128
 - replacement, 128
 - size, 128
 - width, 128
- protrusion decomposition, 131
 - core, 131
 - low-adhesion, 131
- quasi-wide, 98

- random graph model, 180, 181
 - parametrisation, 181
- random intersection graph, 210
- rankwidth, 21
- redundant vertex, 120
- representative set, 124, 126
 - canonical, 124
 - monotone, 124
 - order-preserving, 124
- RMAT, 211
- root path, 10

- sequential, 66
- siblings, 10
- signature, 86
 - proper, 86
- small world property, 172
- somewhere dense, 47
- Stochastic Kronecker graphs, 211
- stress centrality, 174
- strongly monotone, 119
- strongly topological, 95
- stubs, 186
- subdivision, 16

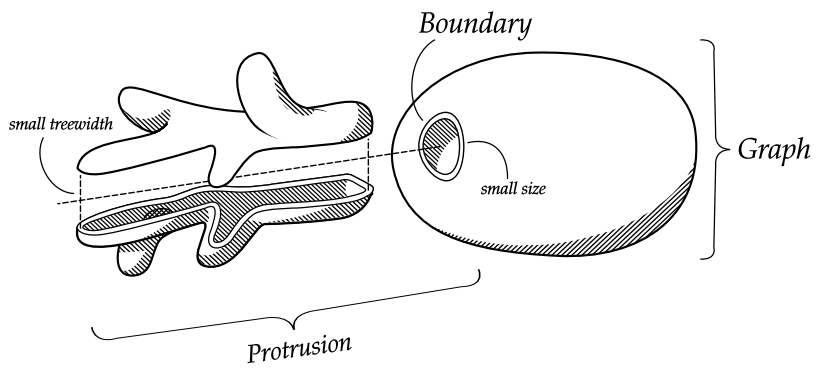
- tail-bound, 186
- tf-augmentation, 66
- topological minor, 14
 - depth, 14
 - embedding, 14
 - shallow, 14
 - stable, 55
 - weighted, 61
- tournament, 10, 127
- transitivity, 173
- tree
 - closure, 11
 - rooted, 10
 - rooted subtree, 11
- tree decomposition, 18
 - nice, 19
 - rooted, 19
- treedepth, 19
 - decomposition, 19
 - nice decomposition, 19
- treewidth, 19, 124
 - locally bounded, 35
- Trémaux tree, 19, 56

- twin class lemma, 83, 141, 142, 144, 151, 154, 158
- twin classes, 12
- twin vertices, 12
- twin-free, 12

- underlying graph, 15

- weak k -colouring number, 97
- well quasi order (wqo), 17, 124
- wide, 98
- width, 19

Part VIII
APPENDIX



IMPORTANT PROBLEMS

VERTEX COVER parametrised by k

Input: A graph G and an integer k .

Problem: Is there a subset $X \subseteq V(G)$ with at least k vertices such that $G \setminus X$ is edgeless?

INDEPENDENT SET parametrised by k

Input: A graph G and an integer k .

Problem: Is there a subset $X \subseteq V(G)$ with at least k vertices such that $G[X]$ is edgeless?

If we further ask that X is a dominating set then the above becomes INDEPENDENT DOMINATING SET.

FEEDBACK VERTEX SET parametrised by k

Input: A graph G and an integer k .

Problem: Is there a subset $X \subseteq V(G)$ with at most k vertices such that $G \setminus X$ is a forest?

If we further ask that $G[X]$ is connected then the above becomes CONNECTED FEEDBACK VERTEX SET.

ODD CYCLE TRANSVERSAL parametrised by k

Input: A graph G and an integer k .

Problem: Is there a subset $X \subseteq V(G)$ with at most k vertices such that $G \setminus X$ is bipartite?

CHORDAL VERTEX DELETION parametrised by k

Input: A graph G and an integer k .

Problem: Is there a subset $X \subseteq V(G)$ with at most k vertices such that $G \setminus X$ is chordal?

k -PATH parametrised by k

Input: A graph G and an integer k .

Problem: Does G contain a path of length at least k ?

k -CYCLE parametrised by k

Input: A graph G and an integer k .

Problem: Does G contain a path of length at least k ?

If we ask for the above problems that the cycle or path has *exactly* length k , we obtain the problems EXACT PATH and EXACT CYCLE. If we further specify two vertices s, t and require them to be the endpoints of the path of length k , we obtain EXACT s, t -PATH.

\mathcal{F} -DELETION parametrised by k

Input: A graph G and integer k .

Problem: Does G have a set X of size at most k such that $G \setminus X$ does not contain an graph from \mathcal{F} as subgraph?

INDUCED- \mathcal{F} -DELETION parametrised by k

Input: A graph G and integer k .

Problem: Does G have a set X of size at most k such that $G \setminus X$ does not contain an graph from \mathcal{F} as an induced subgraph?

\mathcal{F} -PACKING parametrised by k

Input: A graph G and an integer k .

Problem: Does G contain k vertex-disjoint subgraphs that each contain a subgraph isomorphic to a graph in \mathcal{F} ?

\mathcal{H} -MINOR DELETION parametrised by k

Input: A graph G and an integer k .

Problem: Is there a set $X \subseteq V(G)$ of at most k vertices such that $G \setminus X$ does not contain any graph of \mathcal{H} as a minor?

\mathcal{H} -MINOR PACKING parametrised by k

Input: A graph G and an integer k .

Problem: Does G contain k vertex-disjoint subgraphs that each contain some graph of \mathcal{H} as a minor?

\mathbf{wm} - w -DELETION parametrised by k

Input: A graph G and an integer k .

Problem: Does G contain a set $X \subseteq V(G)$ of size at most k such that $\mathbf{wm}(G \setminus X) \leq w$?

\mathcal{P} -DELETION parametrised by k

Input: A graph G and an integer k .

Problem: Does G contain a set $X \subseteq V(G)$ of size at most k such that $G \setminus X \in \mathcal{P}$?

SUBGRAPH ISOMORPHISM parametrised by $|H|$

Input: Graphs G and H .

Problem: Does G contain a subgraph that is isomorphic to H ?

LIST OF AUTHOR'S PUBLICATIONS

 JOURNAL PUBLICATIONS

2015 A. B. Adcock, E. D. Demaine, M. L. Demaine, M. P. O'Brien, F. Reidl, F. Sánchez Villaamil, and B. D. Sullivan. Zig-zag numberlink is NP-complete. *Journal of Information Processing (JIP)*, 23(3):239–245, 2015.

E. J. Kim, A. Langer, C. Paul, F. Reidl, P. Rossmanith, I. Sau, and S. Sikdar. Linear kernels and single-exponential algorithms via protrusion-decompositions. *To appear in Transaction on Algorithms*, 2015.

J. Gajarský, P. Hliněný, J. Obdržálek, S. Ordyniak, F. Reidl, P. Rossmanith, F. Sánchez Villaamil, and S. Sikdar. Kernelization using structural parameters on sparse graph classes. *To appear in Journal of Computer and System Sciences*, 2015.

CONFERENCE AND WORKSHOP PROCEEDINGS

2015 P. G. Drange, F. Reidl, F. Sánchez Villaamil, and S. Sikdar. Fast biclustering by dual parameterization. Accepted at the *10th International Symposium on Parameterized and Exact Computation (IPEC)*, 2015.

2014 A. Langer, F. Reidl, P. Rossmanith, and S. Sikdar. Practical algorithms for MSO model-checking on tree-decomposable graphs. *Computer Science Review*, 13-14:39–74, 2014.

J. Gajarský, P. Hliněný, J. Obdržálek, S. Ordyniak, F. Reidl, P. Rossmanith, F. Sánchez Villaamil, and S. Sikdar. Finite integer index of pathwidth and treewidth. In *Proceedings of the 9th International Workshop on Parameterized and Exact Computation (IWPEC)*, number 8894 in Lecture Notes in Computer Science, pages 258–269. Springer, 2014.

F. Reidl, P. Rossmanith, F. Sánchez Villaamil, and S. Sikdar. A faster parameterized algorithm for treedepth. In *Proceedings of the 41st International Colloquium on Automata, Languages, and Programming (ICALP) Part I*, number 8572 in Lecture Notes in Computer Science, pages 931–942. Springer, 2014.

- 2013 J. Gajarský, P. Hliněný, J. Obdržálek, S. Ordyniak, F. Reidl, P. Rossmanith, F. Sánchez Villaamil, and S. Sikdar. Kernelization using structural parameters on sparse graph classes. In *Proceedings of the 21st European Symposium on Algorithms (ESA)*, number 8125 in Lecture Notes in Computer Science, pages 529–540. Springer, 2013.
- E. J. Kim, A. Langer, C. Paul, F. Reidl, P. Rossmanith, I. Sau, and S. Sikdar. Linear kernels and single-exponential algorithms via protrusion-decompositions. In *Proceedings of the 40th International Colloquium on Automata, Languages, and Programming (ICALP) Part I*, number 7965 in Lecture Notes in Computer Science, pages 613–624. Springer, 2013.
- 2012 A. Langer, F. Reidl, P. Rossmanith, and S. Sikdar. Evaluation of an MSO-solver. In *Proceedings of the 14th Meeting on Algorithm Engineering & Experiments, ALENEX 2012*, pages 55–63. SIAM / Omnipress, 2012.
- 2011 P. Kranen, F. Reidl, F. Sánchez Villaamil, and T. Seidl. Hierarchical clustering for real-time stream data with noise. In *Proceedings of the 23rd International Conference on Scientific and Statistical Database Management (SSDBM)*, volume 6809 of *Lecture Notes in Computer Science*. Springer, 2011.
- 2010 E. Müller, P. Kranen, M. Nett, F. Reidl, and T. Seidl. Air-indexing on error prone communication channels. In *Proceedings of the 15th International Conference on Database Systems for Advanced Applications (DASFAA), Part I*, volume 5981 of *Lecture Notes in Computer Science*, pages 505–519. Springer, 2010.

PREPRINTS

- 2015 M. Farrel, T. D. Goodrich, N. Lemons, F. Reidl, F. Sánchez Villaamil, and B. D. Sullivan. Hyperbolicity, degeneracy, and expansion of random intersection graphs. Submitted to the *12th Workshop on Algorithms and Models for the Web-graph (WAW)*.
- 2014 Erik D. Demaine, F. Reidl, P. Rossmanith, F. Sánchez Villaamil, S. Sikdar, and B. D. Sullivan. Structural sparsity of complex networks: Bounded expansion in random models and real-world graphs. *CoRR*, 2014.
- P. G. Drange, M. S. Dregi, F. V. Fomin, S. Kreutzer, D. Lokshtanov, M. Pilipczuk, M. Pilipczuk, F. Reidl, S. Saurabh, F. Sánchez Villaamil, and S. Sikdar. Kernelization and sparseness: the case of dominating set. *CoRR*, abs/1411.4575, 2014.