Treewidth-Pliability and PTAS for Max-CSPs*

Miguel Romero[†] Marcin Wrochna[‡] Stanislav Živný[§]

Abstract

We identify a sufficient condition, *treewidth-pliability*, that gives a polynomial-time approximation scheme (PTAS) for a large class of Max-2-CSPs parametrised by the class of allowed constraint graphs (with arbitrary constraints on an unbounded alphabet). Our result applies more generally to the maximum homomorphism problem between two rational-valued structures.

The condition unifies the two main approaches for designing PTASes. One is Baker's layering technique, which applies to sparse graphs such as planar or excluded-minor graphs. The other is based on Szemerédi's regularity lemma and applies to dense graphs. We extend the applicability of both techniques to new classes of Max-CSPs.

Treewidth-pliability turns out to be a robust notion that can be defined in several equivalent ways, including characterisations via size, treedepth, or the Hadwiger number. We show connections to the notions of fractional-treewidth-fragility from structural graph theory, hyperfiniteness from the area of property testing, and regularity partitions from the theory of dense graph limits. These may be of independent interest. In particular we show that a monotone class of graphs is hyperfinite if and only if it is fractionally-treewidth-fragile and has bounded degree.

The full version [59] containing detailed proofs is available at https://arxiv.org/abs/1911.03204.

1 Introduction

The problem of finding a *maximum cut* in a graph (Max-Cut) is one of the most studied problems from Karp's original list of 21 NP-complete problems [43]. While Max-Cut is NP-hard to solve optimally, there is a trivial 0.5-approximation algorithm [60] and the celebrated 0.878-approximation algorithm of Goemans and Williamson [31]. Papadimitriou and Yannakakis established that Max-Cut is Max-SNP-hard [57]. By the work of Arora, Lund, Motwani, Sudan, and Szegedy [5] this implies that, unless P=NP, there is no polynomial-time approximation scheme (PTAS) for Max-Cut in general graphs. However, non-trivial results exist for important special cases. On the one hand, Max-Cut is solvable exactly in planar graphs, as shown by Hadlock [38], and more generally, Max-Cut admits a PTAS on graph classes excluding a fixed minor, as shown by Demaine, Hajiaghayi, and Kawarabayashi [20]. On the other hand, Arora, Karger, and Karpinski showed a PTAS for Max-Cut in dense graphs [4], where a graph class is dense if every graph in it contains at least a constant fraction of all possible edges.

Max-Cut is an example of *maximum constraint satisfaction problem* (Max-CSP), although a very special one (the alphabet size is 2, in particular constant, and every constraint uses the same symmetric predicate "$x \neq y$" of arity 2). Another well-known example is Max-r-SAT, with alphabet size 2 and r-ary clauses. Motivated by results on planar, excluded-minor, and dense graph classes, our goal in this paper is to understand the following question:

What structure allows for the existence of a PTAS for Max-CSPs?

We focus on two computational problems. First, we study the general Max-2-CSP($\mathcal{G}$) problem parameterised by the class of underlying constraint graphs (a.k.a. *primal* or *Gaifman* graphs). The input is a graph $G \in \mathcal{G}$, an alphabet Σ_v for each vertex, and for each edge uv a valued constraint $f_{uv}\colon \Sigma_u \times \Sigma_v \to \mathbb{Q}_{\geq 0}$. The goal is to find an assignment $h(v) \in \Sigma_v$ maximising $\sum_{uv} f_{uv}(h(u), h(v))$. Similarly, in Max-$r$-CSP($\mathcal{G}$) a constraint may appear on any r-clique in G. The constraints are arbitrary (non-negative) and the alphabets are not fixed, making the problem very expressive.[1]

Second, we consider a more general framework called the *maximum homomorphism problem* (Max-Hom) of computing the maximum value of any

*Stanislav Živný was supported by a Royal Society University Research Fellowship. Work done while Miguel Romero was at the University of Oxford. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 714532). The paper reflects only the authors' views and not the views of the ERC or the European Commission. The European Union is not liable for any use that may be made of the information contained therein.

†Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Santiago, Chile. miguel.romero.o@uai.cl

‡Department of Computer Science, University of Oxford, UK. marcin.wrochna@cs.ox.ac.uk

§Department of Computer Science, University of Oxford, UK. standa.zivny@cs.ox.ac.uk

[1] One could attempt to generalise counting problems by maximizing $\prod_{uv} f_{uv}(h(u), h(v))$ instead, or equivalently its logarithm $\sum_{uv} \log f_{uv}(h(u), h(v))$. However, the requirement $f_{uv} \geq 0$ and the approximation ratio change. This changes the complexity: for example, approximating the number of 3-colourings requires deciding whether there is at least one in polynomial time, which is NP-hard already in 4-regular planar graphs [13].

map between two given $\mathbb{Q}_{\geq 0}$-valued structures $\mathbb{A}$ and $\mathbb{B}$; the value will be denoted by $\text{opt}(\mathbb{A}, \mathbb{B})$ (see Section 2 for precise definitions). Intuitively, the left-hand-side structure describes the (weighted) scopes of the constraints and the right-hand-side structure describes the different types of constraints. Following Grohe's notation [36], for a class of structures $\mathcal{A}$ we denote by Max-Hom($\mathcal{A}, -$) the restriction of Max-Hom to instances $(\mathbb{A}, \mathbb{B})$ with $\mathbb{A} \in \mathcal{A}$ and $\mathbb{B}$ arbitrary. This framework captures the Max-r-CSP($\mathcal{G}$) problem as a particular case: it is equivalent to Max-Hom($\mathcal{A}_{\mathcal{G}}^{(r)}, -$), where by $\mathcal{A}_{\mathcal{G}}^{(r)}$ we denote the class of all valued structures with an underlying graph in $\mathcal{G}$ and arity r. Another example is the case of *graph* Max-CSP, by which we mean a Max-2-CSP that uses the same symmetric predicate in all constraints (as in Max-Cut or Max-q-Cut); this case is equivalent to Max-Hom($\mathcal{A}, -$) where the structures in $\mathcal{A}$ are graphs.

The question of what structure allows to solve Max-CSPs *exactly* in polynomial time is well understood. A standard dynamic approach works for Max-r-CSP($\mathcal{G}$) when $\mathcal{G}$ is a class of graphs of bounded treewidth. Grohe, Schwentick, and Segoufin [37] in fact proved the converse: if $\mathcal{G}$ has unbounded treewidth then Max-r-CSP($\mathcal{G}$), in fact already deciding the existence of a solution satisfying all constraints, cannot be solved in polynomial time (assuming FPT$\neq$W[1]). Grohe's theorem [36] then extended it to the more general framework: for a class of relational (or $\{0, 1\}$-valued) structures $\mathcal{A}$ of bounded arity, the decision problem Hom($\mathcal{A}, -$) can be solved in polynomial time if and only if the *cores* of structures in $\mathcal{A}$ have bounded treewidth. (The core is the smallest homomorphically equivalent substructure; for example, bipartite graphs all have the single edge graph K_2 as a core, so Hom($\mathcal{A}, -$) is easy when $\mathcal{A}$ is a class of bipartite graphs). This was recently extended further to optimisation with valued structures by Carbonnel, Romero, and Živný [9, 10].

Max-r-CSPs do not admit a PTAS in general, since already Max-Cut does not. On the other hand, the techniques that give PTASes for Max-Cut on sparse and dense graphs apply more generally (in fact to a variety of problems beyond Max-CSPs). Our main contribution is a unifying condition, *treewidth-pliability*, that captures all known PTASes for Max-r-CSP($\mathcal{G}$) and Max-Hom($\mathcal{A}, -$) problems.

We call a class of structures $\mathcal{A}$ tw-pliable if it is uniformly close to structures of bounded treewidth. More formally, for any $\varepsilon > 0$ there is a $k = k(\varepsilon)$ such that every structure in $\mathcal{A}$ has an ε-close structure with treewidth at most k. Here we consider two structures $\mathbb{A}$ and $\mathbb{B}$ to be ε-close if $\text{opt}(\mathbb{A}, \mathbb{C})$ is ε-close to $\text{opt}(\mathbb{B}, \mathbb{C})$ for all $\mathbb{C}$ (details in Section 2.3; this notion of distance, which we also characterise combinatorially, may be of independent interest). While the structure of bounded treewidth is not known and cannot be efficiently computed, we show that the Sherali-Adams LP relaxation gives a PTAS for Max-Hom($\mathcal{A}, -$).

THEOREM 1.1. *If $\mathcal{A}$ is a* tw*-pliable class of structures of bounded arity, then Max-Hom($\mathcal{A}, -$) admits a PTAS.*

We emphasise the generality of Theorem 1.1. Firstly, the computational problem (Max-Hom) captures many fundamental problems. Secondly, the notion of pliability captures many previously discovered cases of structures that admit a PTAS. In particular, we now discuss how Theorem 1.1 extends the applicability of the two main approaches for obtaining PTASes.

1.1 Sparse structures: Baker's technique and fragility Perhaps the best known technique for solving problems on planar graphs is Lipton and Tarjan's planar separator theorem [48] and the divide & conquer approach it enables [49]. It can be used to give a PTAS for Max-CSPs with fixed alphabet size on planar graphs (this extends to excluded-minor graphs [2] and more [28]) of bounded degree.

This approach was superseded by Baker's technique [6], which provides better running times and is easily applied to general Max-r-CSPs on arbitrary planar graphs (see e.g. [44]). The idea is very elegant: we partition a planar graph into Breadth-First-Search layers, remove every ℓ-th layer, and show that the remaining components of $\ell - 1$ consecutive layers have bounded treewidth (and so can be solved exactly). By trying different starting layers we can ensure that the removed layers intersect an unknown optimal solution at most $\mathcal{O}(\frac{1}{\ell})$ times, giving a $1 \pm \mathcal{O}(\frac{1}{\ell})$ approximation.

From planar graphs this was extended to graphs of bounded genus by Eppstein [29] and later to all graph classes excluding a fixed minor by Grohe [35] and Demaine et al. [20]. The structural property needed for this approach, originally proved for excluded-minor graphs by DeVos, Ding, Oporowski, Sanders, Reed, Seymour, and Vertigan [21], is tw-*fragility*: they can be partitioned into any constant number of parts such that removing any one part leaves a graph of bounded treewidth. As shown by Hunt, Marathe, and Stearns [53, 41] (see also [42]) as well as Grigoriev and Bodlaender [34], the same property applies to some geometrically-defined graph classes that do not exclude any minor. One example is intersection graphs of unit disks whose centers are at least some constant apart (capturing some applications of the closely related shifting technique of Hochbaum and Maass [40] for geometric packing and covering problems). Another example is 1-planar graphs, or more generally graphs drawn on a fixed surface with a bounded number of intersections per edge.

An important generalisation, *fractional*-tw-*fragility*, was introduced by Dvořák [24]: it suffices that the parts whose removal results in a graph of bounded treewidth are nearly-disjoint (Definition 4.2). This applies to d-dimensional variants of the geometric classes mentioned above (for any constant d), in particular to d-dimensional grids, which are not tw-fragile [8]. For such concrete examples of fragile classes, known proofs show that the nearly-disjoint parts can be computed efficiently. A PTAS can then easily be designed from the definition [24].

We show that the assumption about efficient construction is not needed. We do this by proving that if $\mathcal{G}$ is *any* fractionally-tw-fragile class of graphs (intuitively, any class where a Baker-like technique is known to work), then the class $\mathcal{A}_{\mathcal{G}}^{(r)}$ of all possible structures with Gaifman graph in $\mathcal{G}$ and bounded arity r is tw-pliable.

THEOREM 1.2. *Let $\mathcal{G}$ be a fractionally-*tw*-fragile class of graphs. Then $\mathcal{A}_{\mathcal{G}}^{(r)}$ is* tw*-pliable for every r. Consequently, Max-r-CSP($\mathcal{G}$) admits a PTAS.*

This captures all graph classes $\mathcal{G}$ where a PTAS for Max-r-CSP($\mathcal{G}$) is known.

1.2 Dense structures: the regularity lemma

It is perhaps more surprising that dense structures admit a PTAS. Here a class is *dense* if a constant factor of all possible constraints is present in every structure in the class, e.g. graphs with $\Omega(n^2)$ edges. Arora, Karger, and Karpinski [4] showed that Max-r-CSPs admit a PTAS in the dense regime if the alphabet size is constant (in fact Boolean); de la Vega [15] independently gave a PTAS for dense Max-Cut. Frieze and Kannan [30] proved that these results are essentially possible because of Szemerédi's regularity lemma [63]: intuitively, every graph can be approximated to within an additive $\pm\varepsilon n^2$ error by a random graph (with a constant number of parts, depending on ε only, so that the edges between two parts form a uniformly random graph of some density). For dense graphs, the additive error translates to a relative error, giving a PTAS. They also showed a variant of the regularity lemma that is still applicable to Max-r-CSPs with constant alphabet size, yet avoids its infamous tower-type dependency on ε.

Goldreich, Goldwasser, and Ron [33] connected these results to the area of *property testing*, spawning an entirely new direction of research. They gave constant-time algorithms estimating the optimum value of some graph Max-CSPs. In fact, Alon, de la Vega, Kannan, and Karpinski [1] (see also Andersson and Engebretsen [3]) showed that Max-r-CSPs with a fixed alphabet can be approximated with accuracy $\pm\varepsilon n^r$ by sampling a constant number of vertices (polynomial in $\frac{1}{\varepsilon}$) and finding the optimum on the resulting (constant-size) induced substructure.

None of these results apply to any Max-r-CSP($\mathcal{G}$) and Max-Hom($\mathcal{A}, -$) problem, that is, to unbounded alphabets. We give the first such example: undirected graphs with $\Omega(n^2)$ edges.

THEOREM 1.3. *Let $c > 0$ and let $\mathcal{A}$ be a class of graphs with at least cn^2 edges. Then $\mathcal{A}$ is* tw*-pliable. Consequently, Max-Hom($\mathcal{A}, -$) admits a PTAS.*

(Note here the graphs in $\mathcal{A}$ are input structures, not just Gaifman graphs of input structures). We also show that this cannot be extended to non-graph CSPs: already for the class of tournaments—that is, orientations of complete graphs—a PTAS is impossible, assuming Gap-ETH [59, Corollary E.5].

1.3 Robustness of pliability

The notion of treewidth-pliability not only unifies the different existing algorithmic techniques but it is also quite robust: treewidth-pliability captures a valued analogue of "homomorphic equivalence" (e.g. bipartite graphs, or 3-colourable graphs where each edge is contained in exactly one triangle, cf. [59, Examples D.5 and D.6] as well as small edits: if $\mathcal{A}$ is an pliable class of graphs, say, then the class of graphs obtained by adding or removing $o(m)$ edges from m-edge graphs in $\mathcal{A}$ is again pliable [59, Corollary D.4]. However, this generality comes at a price. First, we show that even for fixed alphabet size, although the approximate optimum value can be found, an approximate solution cannot be constructed (unless P = NP, cf. [59, Example D.7]). Second, unlike in some of the previous results for more restricted classes, our result does not give an EPTAS (i.e., with the degree of the polynomial time bound independent of ε) for fixed alphabet size (cf. [59, Question D.8]). Finally, the use of strong versions of the regularity lemma yields tower-type dependencies on the approximation ratio ε in the dense case.

In the definition of treewidth-pliability we approximate structures by comparing their opt() values and we ask them to be close to structures where the problem can be solved exactly. This is a non-constructive and very general definition. In fact, it is not inconceivable that this captures all tractable cases, i.e., that Max-Hom($\mathcal{A}, -$) has a PTAS if and only if $\mathcal{A}$ is tw-pliable. Nevertheless, we show a variety of equivalent combinatorial definitions, which allow us place a fairly tight bound on what pliability is, structurally.

For classes of the form $\mathcal{A}_{\mathcal{G}}^{(r)}$, that is, if we only restrict the underlying Gaifman graphs, we show that pliability collapses to fractional fragility. In this sense we understand the "sparse" setting exactly.

LEMMA 1.1. *Let $\mathcal{G}$ be a class of graphs. The following are equivalent, for any $r \geq 2$:*

- *$\mathcal{G}$ is fractionally-*tw*-fragile;*
- *$\mathcal{A}_{\mathcal{G}}^{(r)}$ is* tw*-pliable.*

In general, we can replace treewidth with other parameters of the Gaifman graph: size (number of vertices), treedepth td, Hadwiger number (maximum clique minor size), or maximum connected component size, which we denote by cc.

THEOREM 1.4. *Let $\mathcal{A}$ be any class of structures. The following are equivalent:*

- *$\mathcal{A}$ is* td*-pliable;*
- *$\mathcal{A}$ is* tw*-pliable;*
- *$\mathcal{A}$ is Hadwiger-pliable.*

If structures in $\mathcal{A}$ have bounded signatures, then the following are equivalent to the above as well:

- *$\mathcal{A}$ is* size*-pliable;*
- *$\mathcal{A}$ is* cc*-pliable.*

Classes of structures with bounded signatures (see Section 2 for precise definitions) correspond to Max-CSP instances with a bounded number of constraint types; e.g. maximum graph homomorphism. For example any class of dense graphs as in Theorem 1.3 is in fact size-pliable. An example of a class with unbounded signatures is any class of the form $\mathcal{A}_{\mathcal{G}}^{(r)}$. Theorem 1.4 allows us to give concrete and general examples of classes that are *not* tw-pliable: the class of orientations of graphs in $\mathcal{G}$, where $\mathcal{G}$ is *any* class of unbounded average degree, or any class of 3-regular graphs with unbounded girth [59, Lemmas 10.3 and 10.8].

Finally, as a side result, we connect *hyperfiniteness* to fragility. A class of graphs $\mathcal{G}$ is called *hyperfinite* if for every $\varepsilon > 0$ there is a $k = k(\varepsilon)$ such that in every $G \in \mathcal{G}$ one can remove an at-most-ε fraction of edges to obtain a graph with connected components of size at most k. For a monotone class of graphs (closed under taking subgraphs), hyperfiniteness easily implies bounded degree. It is an important notion in property testing: many results in sparse graphs were generalised by the statement that every property of hyperfinite graphs is testable [56]. The idea, originating in the work of Benjamini, Schramm, and Shapira [7] and Kassidim, Kelner, Nguyen, and Onak [39], is that following the approach of Lipton and Tarjan, graphs with sufficiently sublinear separators, such as planar or excluded-minor graphs [2], can be recursively partitioned into bounded-size components, which for bounded-degree graphs gives hyperfiniteness (see e.g. [12, Cor. 3.2] for a slightly stronger property, cf. [54]). This allows, analogously as in the dense case, to give a constant-size approximate description of such graphs by sampling constant-radius balls in them [56]. See [32] for a book on property testing and [46] for a recent improvement for excluded-minor graphs.

We show that a monotone class $\mathcal{G}$ is hyperfinite if and only if it is fractionally-tw-fragile and has bounded degree. In fact, replacing the parameter treewidth by the maximum size of a connected component in a graph, we have:

THEOREM 1.5. *Let $\mathcal{G}$ be a monotone class of graphs. The following are equivalent, for any $r \geq 2$:*

- *$\mathcal{G}$ is hyperfinite;*
- *$\mathcal{G}$ is fractionally-*tw*-fragile and has bounded degree;*
- *$\mathcal{G}$ is fractionally-*cc*-fragile;*
- *$\mathcal{A}_{\mathcal{G}}^{(r)}$ is* cc*-pliable.*

The equivalence of the second and third bullet points was shown by Dvořák [24, Observation 15, Corollary 20], while for the third and fourth the proof is established by the same proof as Lemma 1.1.

1.4 Related work While this paper focuses on Max-r-CSPs, Baker's technique and the regularity lemma apply to many more problems. In fact Khanna and Motwani [44] argued that most known PTAS algorithms can be derived from three canonical optimisation problems on planar graphs, the first being Max-CSP and the latter two being so-called Max-Ones and Min-Ones CSPs (also solvable with Baker's technique). One of the very few results that did not fit their framework was the PTAS for dense Max-Cut.

Generic frameworks extending Baker's technique include the bidimensionality theory of Demaine, Fomin, Hajiaghayi, and Thilikos [18] and its application in the design of PTASes by Demaine and Hajiaghayi [19] (which is however limited to minor-closed graph classes); monotone FO problems on minor-closed graph classes by Dawar, Grohe, Kreutzer, and Schweikardt [14]; and the very recent idea of Baker games, introduced by Dvořák [25] (see also [27]). The latter gives conditions stronger than fractional-tw-fragility, but useful for problems beyond Max-CSPs, and achievable for all examples known to be fractionally fragile.

De la Vega and Karpinski [16, 17] extended the dense approach to subdense cases ($\Omega(\frac{n^2}{\log n})$ edges) for specific problems such as MaxCut and Max-2-SAT. In contrast, they show that Max-Cut on graphs with $\Omega(n^{2-\delta})$ edges is hard to approximate, for any $\delta > 0$.

The best known approximation algorithm for general Max-2-CSPs is due to Charikar, Hajiaghayi, and Karloff [11] and achieves an approximation factor of $\mathcal{O}((nq)^{1/3})$, where n is the number of variables and q is the alphabet size. On the hardness side, Dinur, Fischer, Kindler, Raz, and Safra [23] showed that $\mathcal{O}(2^{\log^{1-\delta}(nq)})$-approximation of Max-2-CSPs is NP-hard. Manurangsi and Moshkovitz [51] gave approximation algorithms for *dense* Max-2-CSPs with large alphabet size (but not PTASes). Manurangsi and Raghavendra [52] establish a tight trade-off between running time and approximation ratio for dense Max-r-CSPs for $r > 2$.

CSPs have also been extensively studied for fixed constraint types, i.e., Max-Hom$(-, \mathbb{B})$ problems for fixed $\mathbb{B}$. Raghavendra showed that the best approximation ratio is always achieved by the basic SDP relaxation [58], assuming Khot's unique games conjecture [45]. The exactly solvable cases were characterised by Thapper and Živný [64]. The approximation factor of graph Max-CSPs was studied by Langberg, Rabani, and Swamy [47].

1.5 Overview

In Section 2, we give formal definitions and present our basic tool: two structures $\mathbb{A}, \mathbb{B}$ have similar values of opt$(-, \mathbb{C})$ if and only if there is a certain fractional cover, which we call an *overcast*, from $\mathbb{A}$ to $\mathbb{B}$ and from $\mathbb{B}$ to $\mathbb{A}$. To prove that treewidth-pliability leads to a PTAS (Theorem 1.1) the main idea is that an overcast allows to show that the values of opt$(-, \mathbb{C})$ are still similar when we look at linear programming relaxations. We delay the details to Section 5.

Section 3 sketches our approach to dense graphs and to Theorem 1.3. In Section 4, we introduce equivalent definitions of fractional fragility and prove Theorem 1.2 by showing how the definition implies suitable overcasts. This also allows us to conclude half of Lemma 1.1 and Theorem 1.4, and to outline the remainder of their proofs.

Sections 5 defines the Sherali-Adams linear programming relaxation and gives the proof of Theorem 1.1. The rest of the proofs and future directions can be found in the full version of this paper [59].

2 Preliminaries

2.1 Structures

A *signature* is a finite set σ of (function) symbols f, each with a specified arity ar(f). We denote by $|\sigma|$ the number of symbols in the signature σ. A *structure* $\mathbb{A}$ over a signature σ (or σ-structure $\mathbb{A}$, for short) is a finite domain A together with a function $f^{\mathbb{A}} : A^{\mathrm{ar}(f)} \to \mathbb{Q}_{\geq 0}$ for each symbol $f \in \sigma$.

We denote by $A, B, C, \ldots$ the domains of structures $\mathbb{A}, \mathbb{B}, \mathbb{C}, \ldots$. For sets A and B, we denote by B^A the set of all mappings from A to B. We define tup$(\mathbb{A})$ to be the set of all pairs $(f, \mathbf{x})$ such that $f \in \sigma$ and $\mathbf{x} \in A^{\mathrm{ar}(f)}$, and by tup$(\mathbb{A})_{>0}$ the pairs $(f, \mathbf{x}) \in \mathrm{tup}(\mathbb{A})$ with $f^{\mathbb{A}}(\mathbf{x}) > 0$.

We denote $\|\mathbb{A}\|_\infty := \max_{(f,\mathbf{x})\in\mathrm{tup}(\mathbb{A})} f^{\mathbb{A}}(x)$ and $\|\mathbb{A}\|_1 := \sum_{(f,\mathbf{x})\in\mathrm{tup}(\mathbb{A})} f^{\mathbb{A}}(x)$. For $\lambda \geq 0$ we write $\lambda\mathbb{A}$ for the *rescaled* σ-structure with domain A and $f^{\lambda\mathbb{A}}(\mathbf{x}) := \lambda f^{\mathbb{A}}(\mathbf{x})$, for $(f, \mathbf{x}) \in \mathrm{tup}(\mathbb{A})$.

Given a σ-structure $\mathbb{A}$, the *Gaifman graph* (or *primal graph*), denoted by G$(\mathbb{A})$, is the graph whose vertex set is the domain A, and whose edges are the pairs $\{u, v\}$ for which there is a tuple $\mathbf{x}$ and a symbol $f \in \sigma$ such that u, v appear in $\mathbf{x}$ and $f^{\mathbb{A}}(\mathbf{x}) > 0$. For $r \geq 2$ and a class of graphs $\mathcal{G}$, we denote by $\mathcal{A}^{(r)}_{\mathcal{G}}$ the class of σ-structures $\mathbb{A}$ with G$(\mathbb{A}) \in \mathcal{G}$ and ar$(f) \leq r$ for every $f \in \sigma$.

The *maximum homomorphism problem* (Max-Hom) is the following computational problem. An instance of Max-Hom consists of two structures $\mathbb{A}$ and $\mathbb{B}$ over the same signature. For a mapping $h : A \to B$, we define value$(h) = \sum_{(f,\mathbf{x})\in\mathrm{tup}(\mathbb{A})} f^{\mathbb{A}}(\mathbf{x}) f^{\mathbb{B}}(h(\mathbf{x}))$. The goal is to find the maximum value over all possible mappings $h : A \to B$.[2] We denote this value by opt$(\mathbb{A}, \mathbb{B})$. Note that when seen as a Max-CSP instance, the domain of the left-hand side structure $\mathbb{A}$ is the variable set, while the domain of the right-hand side structure $\mathbb{B}$ is the *alphabet*.

Given a class $\mathcal{A}$ of structures, Max-Hom$(\mathcal{A}, -)$ is the problem restricted to instances $(\mathbb{A}, \mathbb{B})$ of Max-Hom with $\mathbb{A} \in \mathcal{A}$ (it is a promise problem: algorithms are allowed to do anything when $\mathbb{A} \notin \mathcal{A}$). Recall that for a class of graphs $\mathcal{G}$, the problem Max-r-CSP$(\mathcal{G})$ is equivalent to Max-Hom$(\mathcal{A}^{(r)}_{\mathcal{G}}, -)$.[3]

2.2 Overcasts

Before we define pliability formally, it is useful to consider the following relation. The starting

[2]While called maximum homomorphism, we note that the maximisation is over all possible maps, not only homomorphisms, i.e. those that map non-zero tuples into non-zero tuples.

[3]Note that Max-Hom$(\mathcal{A}^{(r)}_{\mathcal{G}}, -)$ is different from the maximum graph homomorphism problem Max-Hom$(\mathcal{G}, -)$. Indeed, graphs are also structures over the signature $\{e\}$ with one symbol of arity 2 (where $e^G(u, v) = [uv$ is an edge of $G]$, if the graph is not weighted). To avoid confusion, we use $\mathcal{G}$ for a class of Gaifman graphs of some structures and $\mathcal{A}$ for a class of graphs that are themselves used as input structures.

point of all our results is the equivalence of this relation to a more combinatorial notion: the existence of a certain fractional cover, which we shall call an *overcast*.

DEFINITION 2.1. *Let $\mathbb{A}$ and $\mathbb{B}$ be σ-structures. We say that $\mathbb{A}$* overcasts *$\mathbb{B}$, denoted $\mathbb{A} \succeq \mathbb{B}$ if, for all σ-structures $\mathbb{C}$, we have that* $\mathrm{opt}(\mathbb{A}, \mathbb{C}) \geq \mathrm{opt}(\mathbb{B}, \mathbb{C})$.

A distribution over a finite set U is a function $\pi\colon U \to \mathbb{Q}_{\geq 0}$ such that $\sum_{x \in U} \pi(x) = 1$. The support of π is the set $\mathrm{supp}(\pi) := \{x \in U\colon \pi(x) > 0\}$. We write $\mathbb{E}_{x\sim\pi} f(x)$ for $\sum_{x\in U} \pi(x) \cdot f(x)$ and $\Pr_{x\sim\pi}[\phi(x)]$ for $\mathbb{E}_{x\sim\pi}[\phi(x)]$, where $[\phi(x)]$ is 1 if x satisfies the predicate ϕ and 0 otherwise.

DEFINITION 2.2. *Let $\mathbb{A}$ and $\mathbb{B}$ be σ-structures. An* overcast *from $\mathbb{A}$ to $\mathbb{B}$ is a distribution ω over B^A such that for each $(f, \mathbf{x}) \in tup(\mathbb{B})$ we have that*

$$\mathop{\mathbb{E}}_{g\sim\omega} f^{\mathbb{A}}(g^{-1}(\mathbf{x})) \;\geq\; f^{\mathbb{B}}(\mathbf{x}).$$

Here $f^{\mathbb{A}}(g^{-1}(\mathbf{x}))$ denotes the sum of $f^{\mathbb{A}}(\mathbf{y})$ over $\mathbf{y} \in g^{-1}(\mathbf{x}) \subseteq A^{\mathrm{ar}(f)}$.

The following is a consequence of Farkas' Lemma (or LP duality), as shown in [59, Appendix B]. [4]

PROPOSITION 2.1. $\mathbb{A} \succeq \mathbb{B}$ *if and only if there is an overcast from $\mathbb{A}$ to $\mathbb{B}$.*

2.3 Pliability Our definition of pliability involves a notion of distance which may be of independent interest. It quantifies the *relative* difference between two structures (as measured from the right by weighted multicut densities, in the language of Lovász' book on graph limits [50, Ch. 12]).

DEFINITION 2.3. *The* opt-distance *between two structures with the same signature is defined as:*

$$\mathrm{d_{opt}}(\mathbb{A}, \mathbb{B}) := \sup_{\mathbb{C}} |\ln \mathrm{opt}(\mathbb{A}, \mathbb{C}) - \ln \mathrm{opt}(\mathbb{B}, \mathbb{C})| \,.$$

Here $\ln 0 = -\infty$ and $|\ln 0 - \ln 0| = 0$. Equivalently, we can compare rescaled structures; by definition of $\succeq$ and the fact that $\mathrm{opt}(\lambda\mathbb{A}, \mathbb{C}) = \lambda \mathrm{opt}(\mathbb{A}, \mathbb{C})$, we have:

$$\mathrm{d_{opt}}(\mathbb{A}, \mathbb{B}) = \inf\left\{\varepsilon \;\middle|\; \mathbb{A} \succeq e^{-\varepsilon}\,\mathbb{B} \;\text{ and }\; \mathbb{B} \succeq e^{-\varepsilon}\,\mathbb{A}\right\}.$$

One may think of $e^{\pm\varepsilon}$ as close to $1 \pm \varepsilon$. Formally $1 - \varepsilon \leq e^{-\varepsilon} \leq \frac{1}{1+\varepsilon} = 1 - \varepsilon + \mathcal{O}(\varepsilon^2)$ for $\varepsilon \geq 0$.

Finally, a class is treewidth-pliable if it is uniformly close to structures of bounded treewidth:

DEFINITION 2.4. *A class of structures $\mathcal{A}$ is* p-pliable *with respect to a parameter* p *if for every $\varepsilon > 0$, there is $k = k(\varepsilon)$ such that for every σ-structure $\mathbb{A} \in \mathcal{A}$ there is a σ-structure $\mathbb{B}$ with* $\mathrm{p}(\mathbb{B}) \leq k$ *and* $\mathrm{d_{opt}}(\mathbb{A}, \mathbb{B}) \leq \varepsilon$.

Thus to show tw-pliability of various classes, we will construct overcasts from structures $\mathbb{A}$ in the class to $(1-\varepsilon)\mathbb{B}$, for some $\mathbb{B}$ of bounded treewidth, and from $\mathbb{B}$ back to $(1-\varepsilon)\mathbb{A}$.

In this paper we only consider graph parameters: $\mathrm{size}(\mathbb{A}) = |V(\mathrm{G}(\mathbb{A}))| = |A|$, $\mathrm{cc}(\mathbb{A})$ – the maximum size of a connected component of $\mathrm{G}(\mathbb{A})$, treedepth $\mathrm{td}(\mathbb{A})$ as defined in [59, Section 7.3], treewidth $\mathrm{tw}(\mathbb{A})$, and finally the Hadwiger number $\mathrm{Hadwiger}(\mathbb{A})$, which is the maximum K_k minor of $\mathrm{G}(\mathbb{A})$. The treewidth of a structure $\mathbb{A}$ is the treewidth of its Gaifman graph: $\mathrm{tw}(\mathbb{A}) = \mathrm{tw}(\mathrm{G}(\mathbb{A}))$, similarly for other graph parameters. We refer to [22] for definitions of treewidth and minors.

3 Dense graphs: sketch of Theorem 1.3

We start with simple examples of dense graphs. Observe that large cliques can be arbitrarily well approximated by cliques of constant size $\lceil \frac{2}{\varepsilon} \rceil$ (up to normalising total edge weights).

DEFINITION 3.1. *Let $0 < \varepsilon < 1$ and let $n, k \geq \frac{2}{\varepsilon}$. Then $\mathrm{d_{opt}}(K_n, \lambda K_k) \leq \varepsilon$, for $\lambda = \binom{n}{2}/\binom{k}{2}$.*

Proof. For $n, k \geq 2$, define an overcast ω by taking a random function $V(K_n) \to V(K_k)$ (each vertex is placed independently uniformly at random). Then for each $e \in E(K_k)$,

$$\mathop{\mathbb{E}}_{g\sim\omega} |g^{-1}(e)| = \sum_{e' \in E(K_n)} \mathop{\mathbb{E}}_{g\sim\omega} [g(e') = e]$$
$$= \binom{n}{2}\frac{2}{k^2} = \lambda \cdot \binom{k}{2} \cdot \frac{2}{k^2} = (1 - \frac{1}{k})\lambda.$$

Therefore $K_n \succeq (1-\frac{1}{k})\lambda K_k$. Symmetrically $\lambda K_k \succeq (1-\frac{1}{n})K_n$. Since $1 - x \geq e^{-2x}$ for $0 \leq x \leq \frac{1}{2}$, this means $\mathrm{d_{opt}}(K_n, \lambda K_k) \leq \frac{2}{\min(n,k)}$. Consequently if $n, k \geq \frac{2}{\varepsilon}$, then $\mathrm{d_{opt}}(K_n, \lambda K_k) \leq \varepsilon$. ∎

In particular, this means the class $\mathcal{A}$ consisting of all clique graphs is size-pliable. This corresponds to an easy PTAS for graph Max-Hom($\mathcal{A}$, $-$): the maximum graph homomorphism from K_n to G is well approximated

[4] The definitions of the $\succeq$ relation and of an overcast are analogous to the "improvement" relation and "inverse fractional homomorphisms" from [10]. Here, however, opt() is maximising, not minimising, so inequalities in definitions are swapped. This has consequences such as the fact that mappings in the support of an overcast are in general not homomorphisms (mapping non-zero tuples to non-zero tuples), unlike for inverse fractional homomorphisms. The proof of Proposition 2.1 nevertheless is identical to the proof of [10, Proposition 6].

by finding the maximum graph homomorphism from a constant size K_k to G and mapping K_n randomly to the resulting $\leq k$ vertices in G. The situation is very different for Densest Subgraph problems, because they disallow choosing two equal vertices in G (see [59, Observation E.2]).

As another important example, consider Erdős-Rényi random graphs $G(n,p)$ (for constant $p \in (0,1)$; each pair in $\binom{n}{2}$ becomes an edge independently with probability p). They are similar to each other (and in fact to pK_n, as well as to λK_k for constant k and suitable λ):

Definition 3.2. *Let $p, \varepsilon > 0$ be constants. Let G_1, G_2 be independent Erdős-Rényi random graphs $G(n,p)$. Then $\Pr[\mathrm{d_{opt}}(G_1, G_2) < \varepsilon] \to 1$ as $n \to \infty$.*

Proof. [Proof sketch] Let k be a sufficiently large constant depending on ε only. It is sufficient to prove that $\Pr[\mathrm{d_{opt}}(G_1, \lambda K_k) < \frac{\varepsilon}{2}] \to 1$ as $n \to \infty$. The rescaling factor here is $\lambda := p\binom{n}{2}/\binom{k}{2}$. The number of edges of $G(n,p)$ is concentrated around $p\binom{n}{2}$, so just as before a random function gives $G(n,p) \succeq (1 - \frac{1}{k})\lambda K_k \succeq e^{-\varepsilon/2}\lambda K_k$ with high probability (tending to 1 as $n \to \infty$).

For the other direction, we use the fact that the number of k-cliques in $G(n,p)$ is concentrated around the mean $\binom{n}{k}p^{\binom{k}{2}}$ and, more strongly, the number of k-cliques containing any given edge of $G(n,p)$ (conditioned on it being an edge) is concentrated around the mean $\binom{n-2}{k-2}p^{\binom{k}{2}-1}$. The concentration is good enough that with high probability, *every* edge of $G(n,p)$ is contained in $(1 \pm \frac{\varepsilon}{4})\binom{n-2}{k-2}p^{\binom{k}{2}-1}$ k-cliques (see e.g. [62]). Thus if we take ω by mapping λK_k injectively to a random k-clique in $G(n,p)$, then w.h.p. for each edge e of $G(n,p)$ we have

$$\mathop{\mathbb{E}}_{g\sim\omega} |g^{-1}(e)| \geq (1 - \frac{\varepsilon}{4})\binom{n-2}{k-2}p^{\binom{k}{2}-1} / \binom{n}{k}p^{\binom{k}{2}}$$
$$= (1 - \frac{\varepsilon}{4})\frac{k(k-1)}{n(n-1)}p^{-1} = (1 - \frac{\varepsilon}{4})\lambda^{-1}.$$

Thus $\lambda K_k \succeq e^{-\varepsilon/2}G_1$ and consequently $\mathrm{d_{opt}}(G_1, \lambda K_k) \leq \frac{\varepsilon}{2}$ w.h.p. ∎

To show Theorem 1.3, we extend the above informal proof to any class of dense graphs. This is possible because of the *Szemerédi's regularity lemma* [63], which, very roughly speaking, guarantees that all such graphs are random-like. This allows to provide similar bounds on the number of k-cliques containing any given edge, a fact known as the *extension lemma*, though we prove a variant that is somewhat tighter than usual. More details and the full proof can be found in [59].

Note that the above proof sketch does not work for random tournaments (orientations of cliques): if we try to approximate them by the small graph $\frac{1}{2}\overset{\leftrightarrow}{K}_k$ (each arc taken with weight $\frac{1}{2}$), then every overcast from it to a tournament will always lose at least half of the total weight. If instead we tried to take a small random tournament, no overcast to it from the big random tournament would work. Indeed, [59, Lemma 10.3] shows the class of tournaments is not pliable (neither are "random tournaments", i.e. the proof can be adapted to show that any class which contains a random tournament with constant probability cannot be pliable) and in fact the problem Max-Hom($\mathcal{A}, -$) for the class of tournaments $\mathcal{A}$ is hard to approximate, as we show in [59, Lemma E.4]. This is why, even though variants of the regularity lemma exist for directed graphs and even more general structures, we limit our discussion to undirected graphs (the proofs do extend to $[0,1]$-weighted undirected graphs, however).

4 Fractional fragility: proof of Theorem 1.2

To give Dvořák's definition of fractional fragility [24] we first define ε-thin distributions.

Definition 4.1. *Let $\mathcal{F}$ be a family of subsets of a set V and $\varepsilon > 0$. We say a distribution π over $\mathcal{F}$ is* ε-thin *if $\Pr_{X\sim\pi}[v \in X] \leq \varepsilon$ for all $v \in V$.*

Definition 4.2. *For a graph parameter* p *and a number k, we define a* (p $\leq k$)-modulator *of a graph G to be a set $X \subseteq V(G)$ such that $\mathrm{p}(G - X) \leq k$. A* fractional (p $\leq k$)-modulator *is a distribution π of such modulators X. We say that a class of graphs $\mathcal{G}$ is* fractionally-p-fragile *if for every $\varepsilon > 0$ there is a k such that every $G \in \mathcal{G}$ has an ε-thin fractional* (p $\leq k$)-*modulator. We can analogously define* (p $\leq k$)-*edge-modulators $F \subseteq E(G)$ and* fractionally-p-edge-fragility.

One crucial property of fractional fragility is that it allows a dual definition by a variant of Farkas' Lemma (cf. [59, Appendix A] for details); this is already implicit in [26, Lemma 5].

Lemma 4.1. *Let $\mathcal{F}$ be a family of subsets of a set V. The following are equivalent:*

- *there is an ε-thin distribution π of sets in $\mathcal{F}$;*
- *for all non-negative weights $(w(v))_{v\in V}$, there is an $X \in \mathcal{F}$ such that $w(X) \leq \varepsilon \cdot w(V)$.*

Thus a class of graphs $\mathcal{G}$ is fractionally-tw-fragile if and only if for every $\varepsilon > 0$ there is a k such that for every graph $G \in \mathcal{G}$ and every vertex-weight function w, one can remove a set of vertices of weight at most $\varepsilon \cdot w(V)$ to obtain a graph with tw $\leq k$. Here $w(X) := \sum_{x\in X} w(x)$.

Another useful property is that the edge version is equivalent to the vertex version, for most parameters of interest. A parameter is *monotone* if $\mathrm{p}(H) \leq \mathrm{p}(G)$ for H a subgraph of G. The proof of the following is in [59, Section 6].

LEMMA 4.2. *Let* p *be a monotone graph parameter such that the average degree* $\frac{2|E(G)|}{|V(G)|}$ *of a graph is bounded by a function of* $\mathrm{p}(G)$*. Let* $\mathcal{G}$ *be a class of graphs. Then the following are equivalent:*

- $\mathcal{G}$ *is fractionally-*p*-fragile;*
- $\mathcal{G}$ *is fractionally-*p*-edge-fragile;*
- $\forall_{\varepsilon>0}\exists_k\forall_{G\in\mathcal{G}}\forall_{w\colon V(G)\to\mathbb{Q}_{\geq 0}}\exists_{X\subseteq V(G)}\quad w(X) \leq \varepsilon w(V(G))$ *and* $\mathrm{p}(G-X)\leq k$*;*
- $\forall_{\varepsilon>0}\exists_k\forall_{G\in\mathcal{G}}\forall_{w\colon E(G)\to\mathbb{Q}_{\geq 0}}\exists_{F\subseteq E(G)}\quad w(F) \leq \varepsilon w(E(G))$ *and* $\mathrm{p}(G-F)\leq k$*.*

Dvořák and Sereni [26, Theorem 28] showed that graphs of bounded treewidth are fractionally-td-fragile. It follows from a result of DeVos et al. [21, Theorem 1.2] that for every graph H, H-minor-free graphs are fractionally-tw-fragile. In fact, as shown by Dvořák [25], a proof of van den Heuvel et al. [65, Lemma 4.1] can be adapted to show this without the Graph Minors Structure Theorem.

THEOREM 4.1. ([26, 21]) *For every* H*, the class of* H*-minor-free graphs is fractionally-*tw*-fragile. For every* k*, the class of graphs of treewidth at most* k *is fractionally-*td*-fragile.*

Consequently (cf. [24, Lemma 12]), the following are equivalent for a class of graphs $\mathcal{G}$*:*

- $\mathcal{G}$ *is fractionally-*td*-fragile;*
- $\mathcal{G}$ *is fractionally-*tw*-fragile;*
- $\mathcal{G}$ *is fractionally-Hadwiger-fragile.*

4.1 Fragility implies pliability We denote by $G \uplus H$ the disjoint union of graphs G and H.

LEMMA 4.3. *Let* p *be a graph parameter such that* $\mathrm{p}(G \uplus H) = \max(\mathrm{p}(G), \mathrm{p}(H))$ *for all* G, H*. Let* $\mathcal{A}$ *be a class of structures of bounded arity* r *such that the class* $\mathcal{G}$ *of their Gaifman graphs is fractionally-*p*-fragile. Then* $\mathcal{A}$ *is* p*-pliable.*

Proof. By definition of fractional-p-fragility, $\forall_{\varepsilon>0}\exists_{k(\varepsilon)}\forall_{G\in\mathcal{G}}\ G$ has an ε-thin fractional $(\mathrm{p} \leq k)$-modulator, for some function $k(\varepsilon)$. For $\varepsilon > 0$, let $\varepsilon' := \frac{\varepsilon}{1+\varepsilon}\cdot\frac{1}{r}$ and let $k := k(\varepsilon')$. Let $\mathbb{A} \in \mathcal{A}$ be a structure with Gaifman graph $G \in \mathcal{G}$. By assumption, G has a fractional $(\mathrm{p} \leq k)$-modulator π such that for every $v \in V(G)$, $\Pr_{X\sim\pi}[v \in X] \leq \varepsilon'$. For $X \subseteq V(G) = A$ in the support of π, let $\mathbb{B}_X$ be the rescaling of $\mathbb{A} - X$ by a factor of $\pi(X)$; let $\mathbb{B}$ be the disjoint union of all $\mathbb{B}_X$. Since each X in the support of π is a $(\mathrm{p} \leq k)$-modulator and p is closed under disjoint union, $\mathrm{p}(\mathrm{G}(\mathbb{B})) \leq k$.

We define overcasts $\omega\colon \mathbb{A} \to \mathbb{B}$ and $\omega'\colon \mathbb{B} \to (1 - r\varepsilon')\mathbb{A}$. The first, ω, maps $\mathbb{A}$ identically to each component $\mathbb{B}_X$ of $\mathbb{B}$ with probability $\pi(X)$ (vertices of $\mathbb{A}$ in X are mapped arbitrarily). The second, ω', deterministically maps each component $\mathbb{B}_X$ of $\mathbb{B}$ identically to $\mathbb{A}$. To check that ω' is indeed an overcast, consider a tuple $(f, \mathbf{x}) \in \mathrm{tup}(\mathbb{A})$. The tuple is covered by its copies in $\mathbb{B}_X$ with weight $\pi(X)\cdot f^{\mathbb{A}}(X)$ for all X which do not intersect $\mathbf{x}$. In total, the fraction of $f^{\mathbb{A}}(\mathbf{x})$ lost is hence exactly $\Pr_{X\sim\pi}[X \cap \mathbf{x} \neq \emptyset]$, which is (by union bound and by the assumption $|\mathbf{x}| \leq r$) at most $\varepsilon' r$. Since $1 - \varepsilon' r = \frac{1}{1+\varepsilon} \geq e^{-\varepsilon}$, we have $\mathbb{A} \succeq \mathbb{B} \succeq (1 - \varepsilon' r)\mathbb{A} \succeq e^{-\varepsilon}\mathbb{A}$, which means $\mathbb{B}$ is a structure at opt-distance $\leq \varepsilon$ from $\mathbb{A}$. ∎

This concludes Theorem 1.2: structures on fractionally-tw-fragile graphs are tw-pliable.

For Lemma 1.1, we need the other direction: that if all structures on Gaifman graphs in $\mathcal{G}$ are tw-pliable, then $\mathcal{G}$ is fractionally-tw-fragile. To do this, we consider, for a graph $G \in \mathcal{G}$, a structure $\mathbb{A}$ where each edge is used by a different symbol of a signature. If we have a structure $\mathbb{B}$ (of bounded treewidth) close to $\mathbb{A}$ in opt-distance, this implies overcasts from $\mathbb{A}$ to $e^{-\varepsilon}\mathbb{B}$ and from $\mathbb{B}$ to $e^{-\varepsilon}\mathbb{A}$; composing the two gives an overcast from $e^{+\varepsilon}\mathbb{A}$ to $e^{-\varepsilon}\mathbb{A}$ in which (since each edge is used by a different symbol) an edge can only be covered by itself. This shows that the overcasts are mostly injective and that $\mathbb{B}$, sandwiched between $e^{+\varepsilon}\mathbb{A}$ and $e^{-\varepsilon}\mathbb{A}$, must be close in edit distance. The bounded treewidth of $\mathbb{B}$ then implies that the graph G underlying $\mathbb{A}$ is in fact fractionally-tw-edge-fragile, which by Lemma 4.2 concludes the proof. Details can be found in [59, Section 6].

The first half of Theorem 1.4 already follows easily as a corollary of Theorem 4.1, Lemma 4.3 and the following simple observation. Details can be found in [59, Section 7].

DEFINITION 4.1. (TRANSITIVITY OF PLIABILITY) *Let* $\mathcal{A}$ *be a class of structures with signatures from a set* Σ*. Suppose* $\mathcal{A}$ *is* p*-pliable and that for each* k*,* $\{\mathbb{A}\colon \mathrm{p}(\mathbb{A}) \leq k\}$ *is* p′*-pliable, where* $\mathbb{A}$ *runs over all structures with signatures in* Σ*. Then* $\mathcal{A}$ *is* p′*-pliable.*

The second half of Theorem 1.4 similarly reduces to showing that structures of bounded treedepth with a

$$\max \sum_{(f,\mathbf{x})\in\text{tup}(\mathbb{A}),\, s\colon \text{Set}(\mathbf{x})\to B} \lambda(\text{Set}(\mathbf{x}),s) f^{\mathbb{A}}(\mathbf{x}) f^{\mathbb{B}}(s(\mathbf{x}))$$

$$\lambda(X,s) = \sum_{r\colon Y\to B,\, r|_X = s} \lambda(Y,r) \qquad \text{for } X\subseteq Y\in \tbinom{A}{\le k} \text{ and } s\colon X\to B$$

$$\sum_{s\colon X\to B} \lambda(X,s) = 1 \qquad \text{for } X\in \tbinom{A}{\le k}$$

$$\lambda(X,s)\ge 0 \qquad \text{for } X\in \tbinom{A}{\le k} \text{ and } s\colon X\to B$$

Figure 1: The Sherali-Adams relaxation of level k of $(\mathbb{A},\mathbb{B})$.

bounded signature are size-pliable. The strategy for the proof is similar to a proof of Nešetřil and Ossona de Mendez [55, Corollary 3.3] that relational structures of bounded treedepth have bounded cores. However the argument is much more intricate due to the fact that we consider valued structures: the statement that there are only finitely many structures of size at most C, for every C, is not true anymore. The main difficulty is proving an approximate version of it: we do this in the full version of this paper [59].

5 Pliable structures admit a PTAS: proof of Theorem 1.1

We first define the Sherali-Adams LP hierarchy [61] for Max-Hom. Let $(\mathbb{A},\mathbb{B})$ be an instance of Max-Hom over a signature σ and let $k\ge \max_{f\in\sigma}\text{ar}(f)$. For a tuple $\mathbf{x}$, we denote by $\text{Set}(\mathbf{x})$ the set of elements appearing in $\mathbf{x}$. We write $\binom{A}{\le k}$ for the set of subsets of A with at most k elements. The *Sherali-Adams relaxation of level k* [61] of $(\mathbb{A},\mathbb{B})$ is the linear program given in Figure 1, which has one variable $\lambda(X,s)$ for each $X\in\binom{A}{\le k}$ and each $s\colon X\to B$.

We denote by $\text{opt}_k(\mathbb{A},\mathbb{B})$ the optimum value of this linear program.

Definition 5.1. *Let $\mathbb{A}$ be a σ-structure, $\lambda\ge 0$ and $k\ge \max_{f\in\sigma}\text{ar}(f)$. Then for all σ-structures $\mathbb{C}$, we have* $\text{opt}(\lambda\mathbb{A},\mathbb{C}) = \lambda\text{opt}(\mathbb{A},\mathbb{C})$ *and* $\text{opt}_k(\lambda\mathbb{A},\mathbb{C}) = \lambda\text{opt}_k(\mathbb{A},\mathbb{C})$.

Definition 5.1. *Let $\mathbb{A}$ and $\mathbb{B}$ be σ-structures and $k\ge \max_{f\in\sigma}\text{ar}(f)$. We write $\mathbb{A}\succeq_k\mathbb{B}$ if, for all σ-structures $\mathbb{C}$, we have* $\text{opt}_k(\mathbb{A},\mathbb{C})\ge\text{opt}_k(\mathbb{B},\mathbb{C})$.

The proof of the following is analogous to the proof of [10, Proposition 27] and can be found in [59, Appendix C].

Proposition 5.1. *Let $\mathbb{A}$ and $\mathbb{B}$ be σ-structures and $k\ge \max_{f\in\sigma}\text{ar}(f)$. If there is an overcast from $\mathbb{A}$ to $\mathbb{B}$ then $\mathbb{A}\succeq_k\mathbb{B}$.*

Using Observation 5.1 and Proposition 5.1, we are ready to prove the following.

Proposition 5.2. *Let $\mathbb{A}$ be a σ-structure. Let $\varepsilon\ge 0$ and $k\ge\max_{f\in\sigma}\text{ar}(f)$. Suppose that there exists a σ-structure $\mathbb{B}$ such that $\text{d}_{\text{opt}}(\mathbb{A},\mathbb{B})\le\varepsilon$ and $\text{tw}(\mathbb{B})\le k$. Then, for every σ-structure $\mathbb{C}$, we have that*

$$\text{opt}(\mathbb{A},\mathbb{C}) \;\le\; \text{opt}_k(\mathbb{A},\mathbb{C}) \;\le\; (1+\mathcal{O}(\varepsilon))\text{opt}(\mathbb{A},\mathbb{C}).$$

Proof. The left-hand side inequality is from the definition of Sherali-Adams. For the right-hand side inequality, observe first that, by definition of d_{opt}, $\mathbb{A}\succeq e^{-\varepsilon}\mathbb{B}$ and $\mathbb{B}\succeq e^{-\varepsilon}\mathbb{A}$. By Proposition 2.1, there is an overcast from $\mathbb{B}$ to $e^{-\varepsilon}\mathbb{A}$, so by Proposition 5.1, it follows that $\mathbb{B}\succeq_k e^{-\varepsilon}\mathbb{A}$. By Observation 5.1, we have that $\text{opt}_k(\mathbb{B},\mathbb{C})\ge e^{-\varepsilon}\text{opt}_k(\mathbb{A},\mathbb{C})$. Since $\text{tw}(\mathbb{B})\le k$, we have $\text{opt}_k(\mathbb{B},\mathbb{C}) = \text{opt}(\mathbb{B},\mathbb{C})$ – this follows, for example, from [10, Theorem 33].[5] Since moreover $\mathbb{A}\succeq e^{-\varepsilon}\mathbb{B}$, by Observation 5.1, it follows that $\text{opt}(\mathbb{A},\mathbb{C})\ge e^{-\varepsilon}\text{opt}(\mathbb{B},\mathbb{C})$. Together, $\text{opt}(\mathbb{A},\mathbb{C})\ge e^{-\varepsilon}\text{opt}(\mathbb{B},\mathbb{C}) = e^{-\varepsilon}\text{opt}_k(\mathbb{B},\mathbb{C})\ge e^{-2\varepsilon}\text{opt}_k(\mathbb{A},\mathbb{C})$. Hence $\text{opt}_k(\mathbb{A},\mathbb{C})\le e^{2\varepsilon}\text{opt}(\mathbb{A},\mathbb{C})$. □

Since $\text{opt}_k(\mathbb{A},\mathbb{C})$ can be computed in time $(|A|\cdot|C|)^{\mathcal{O}(k)}$, this concludes the proof of Theorem 1.1.

References

[1] N. Alon, W. F. de la Vega, R. Kannan, and M. Karpinski. Random sampling and approximation of MAX-CSPs. *J. Comput. Syst. Sci.*, 67(2):212–243, 2003.

[2] N. Alon, P. D. Seymour, and R. Thomas. A separator theorem for graphs with an excluded minor and its

[5] Our definition of the LP slightly differs from [9, 10], where there are additional variables $\lambda(f,\mathbf{x},s)$ associated with tuples $(f,\mathbf{x})$ with $f^{\mathbb{A}}(\mathbf{x})>0$. However, since we are assuming without loss of generality that $k\ge\max_{f\in\sigma}\text{ar}(f)$, the two definitions are equivalent.

applications. In *Proc. 22nd Annual ACM Symposium on Theory of Computing (STOC'90)*, pages 293–299. ACM, 1990.

[3] G. Andersson and L. Engebretsen. Property testers for dense constraint satisfaction programs on finite domains. *Random Struct. Algorithms*, 21(1):14–32, 2002.

[4] S. Arora, D. R. Karger, and M. Karpinski. Polynomial Time Approximation Schemes for Dense Instances of NP-Hard Problems. *J. Comput. Syst. Sci.*, 58(1):193–210, 1999.

[5] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. Proof verification and the hardness of approximation problems. *J. ACM*, 45(3):501–555, 1998.

[6] B. S. Baker. Approximation Algorithms for NP-Complete Problems on Planar Graphs. *J. ACM*, 41(1):153–180, 1994.

[7] I. Benjamini, O. Schramm, and A. Shapira. Every minor-closed property of sparse graphs is testable. In *Proc. 40th Annual ACM Symposium on Theory of Computing (STOC'08)*, pages 393–402. ACM, 2008.

[8] E. Berger, Z. Dvořák, and S. Norin. Treewidth of grid subsets. *Combinatorica*, 38(6):1337–1352, 2018, arXiv:1512.06441.

[9] C. Carbonnel, M. Romero, and S. Živný. The complexity of general-valued CSPs seen from the other side. In *Proc. 59th IEEE Annual Symposium on Foundations of Computer Science (FOCS'18)*, pages 236–246. 2018.

[10] C. Carbonnel, M. Romero, and S. Živný. The complexity of general-valued CSPs seen from the other side. Technical report, 8 2018, arXiv:1710.03148.

[11] M. Charikar, M. Hajiaghayi, and H. J. Karloff. Improved approximation algorithms for label cover problems. *Algorithmica*, 61(1):190–206, 2011.

[12] A. Czumaj, A. Shapira, and C. Sohler. Testing hereditary properties of nonexpanding bounded-degree graphs. *SIAM J. Comput.*, 38(6):2499–2510, 2009.

[13] D. P. Dailey. Uniqueness of colorability and colorability of planar 4-regular graphs are NP-complete. *Discret. Math.*, 30(3):289–293, 1980.

[14] A. Dawar, M. Grohe, S. Kreutzer, and N. Schweikardt. Approximation Schemes for First-Order Definable Optimisation Problems. In *Proc. 21st IEEE Symposium on Logic in Computer Science (LICS'06)*, pages 411–420. IEEE Computer Society, 2006.

[15] W. F. de la Vega. MAX-CUT has a randomized approximation scheme in dense graphs. *Random Struct. Algorithms*, 8(3):187–198, 1996.

[16] W. F. de la Vega and M. Karpinski. A polynomial time approximation scheme for subdense MAX-CUT. *Electronic Colloquium on Computational Complexity (ECCC)*, (044), 2002.

[17] W. F. de la Vega and M. Karpinski. Approximation complexity of nondense instances of MAX-CUT. *Electronic Colloquium on Computational Complexity (ECCC)*, 13(101), 2006.

[18] E. D. Demaine, F. V. Fomin, M. T. Hajiaghayi, and D. M. Thilikos. Subexponential parameterized algorithms on bounded-genus graphs and H-minor-free graphs. *J. ACM*, 52(6):866–893, 2005.

[19] E. D. Demaine and M. T. Hajiaghayi. Bidimensionality: new connections between FPT algorithms and ptass. In *Proc. 16th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'05)*, pages 590–601. SIAM, 2005.

[20] E. D. Demaine, M. T. Hajiaghayi, and K. Kawarabayashi. Algorithmic Graph Minor Theory: Decomposition, Approximation, and Coloring. In *Proc. 46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)*, pages 637–646. IEEE Computer Society, 2005.

[21] M. DeVos, G. Ding, B. Oporowski, D. P. Sanders, B. Reed, P. Seymour, and D. Vertigan. Excluding any graph as a minor allows a low tree-width 2-coloring. *J. Comb. Theory B*, 91(1):25–41, 2004.

[22] R. Diestel. *Graph Theory*. Springer, fourth edition, 2010.

[23] I. Dinur, E. Fischer, G. Kindler, R. Raz, and S. Safra. PCP characterizations of NP: toward a polynomially-small error-probability. *Computational Complexity*, 20(3):413–504, 2011.

[24] Z. Dvořák. Sublinear separators, fragility and subexponential expansion. *Eur. J. Comb.*, 52:103–119, 2016, arXiv:1404.7219.

[25] Z. Dvořák. Baker game and polynomial-time approximation schemes. In *Proc. 31st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'20)*, 2020, arXiv:1901.01797. To appear.

[26] Z. Dvořák and J.-S. Sereni. On fractional fragility rates of graph classes. Technical report, 7 2019, arXiv:1907.12634.

[27] Z. Dvořák. Thin graph classes and polynomial-time approximation schemes. In *Proc. 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'18)*, pages 1685–1701. SIAM, 2018.

[28] Z. Dvořák and S. Norin. Strongly sublinear separators and polynomial expansion. *SIAM J. Discrete Math.*, 30(2):1095–1101, 2016.

[29] D. Eppstein. Diameter and treewidth in minor-closed graph families. *Algorithmica*, 27(3-4):275–291, 2000.

[30] A. M. Frieze and R. Kannan. The regularity lemma and approximation schemes for dense problems. In *Proc. 37th Annual Symposium on Foundations of Computer Science (FOCS'96)*, pages 12–20. 1996.

[31] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. ACM*, 42(6):1115–1145, 1995.

[32] O. Goldreich. *Introduction to property testing*. Cambridge University Press, 2017.

[33] O. Goldreich, S. Goldwasser, and D. Ron. Property testing and its connection to learning and approximation. *J. ACM*, 45(4):653–750, 1998.

[34] A. Grigoriev and H. L. Bodlaender. Algorithms for graphs embeddable with few crossings per edge. *Algo-*

rithmica, 49(1):1–11, 9 2007.
[35] M. Grohe. Local tree-width, excluded minors, and approximation algorithms. *Combinatorica*, 23(4):613–632, 2003.
[36] M. Grohe. The complexity of homomorphism and constraint satisfaction problems seen from the other side. *J. ACM*, 54(1):1–24, 2007.
[37] M. Grohe, T. Schwentick, and L. Segoufin. When is the evaluation of conjunctive queries tractable? In *Proc. 33rd Annual ACM Symposium on Theory of Computing (STOC'01)*, pages 657–666. ACM, 2001.
[38] F. Hadlock. Finding a maximum cut of a planar graph in polynomial time. *SIAM J. Comput.*, 4(3):221–225, 1975.
[39] A. Hassidim, J. A. Kelner, H. N. Nguyen, and K. Onak. Local graph partitions for approximation and testing. In *Proc. 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS'09)*, pages 22–31. IEEE Computer Society, 2009.
[40] D. S. Hochbaum and W. Maass. Approximation schemes for covering and packing problems in image processing and vlsi. *J. ACM*, 32(1):130–136, 1 1985.
[41] H. B. Hunt III, R. Jacob, M. V. Marathe, and R. E. Stearns. Towards syntactic characterizations of approximation schemes via predicate and graph decompositions. In *Proc. IEEE Symposium on Foundations of Computer Science (FOCS'98)*, 12 1998.
[42] H. B. Hunt III, M. V. Marathe, V. Radhakrishnan, S. Ravi, D. J. Rosenkrantz, and R. E. Stearns. NC-approximation schemes for NP- and PSPACE-hard problems for geometric graphs. *J. Algorithms*, 26(2):238 – 274, 1998.
[43] R. M. Karp. Reducibility Among Combinatorial Problems. In *Complexity of Computer Computations*, pages 85–103, 1972.
[44] S. Khanna and R. Motwani. Towards a syntactic characterization of PTAS. In *Proc. 28th Annual ACM Symposium on the Theory of Computing (STOC'96)*, pages 329–337, 1996.
[45] S. Khot. On the power of unique 2-prover 1-round games. In *Proc. 34th Annual ACM Symposium on Theory of Computing (STOC'02)*, pages 767–775. ACM, 2002.
[46] A. Kumar, C. Seshadhri, and A. Stolman. Random walks and forbidden minors II: a poly($d\epsilon^{-1}$)-query tester for minor-closed properties of bounded degree graphs. In M. Charikar and E. Cohen, editors, *Proc. 51st Annual ACM SIGACT Symposium on Theory of Computing (STOC'19)*, pages 559–567. ACM, 2019.
[47] M. Langberg, Y. Rabani, and C. Swamy. Approximation algorithms for graph homomorphism problems. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2006*, volume 4110 of LNCS, pages 176–187. Springer, 2006.
[48] R. J. Lipton and R. E. Tarjan. A separator theorem for planar graphs. *SIAM J. Appl. Math.*, 36(2):177–189, 1979.
[49] R. J. Lipton and R. E. Tarjan. Applications of a planar separator theorem. *SIAM J. Comput.*, 9(3):615–627, 1980.
[50] L. Lovász. *Large Networks and Graph Limits*, volume 60 of *Colloquium Publications*. American Mathematical Society, 2012.
[51] P. Manurangsi and D. Moshkovitz. Approximating dense max 2-CSPs. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2015*, pages 396–415. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2015.
[52] P. Manurangsi and P. Raghavendra. A Birthday Repetition Theorem and Complexity of Approximating Dense CSPs. In *Proc. 44th International Colloquium on Automata, Languages, and Programming (ICALP'17)*, volume 80 of *LIPIcs*, pages 78:1–78:15. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2017.
[53] M. V. Marathe, H. B. Hunt III, and R. E. Stearns. Level-treewidth property, exact algorithms and approximation schemes. In *Proc. ACM Symposium on Theory of Computing (STOC'97)*, 6 1997.
[54] G. Moshkovitz and A. Shapira. Decomposing a graph into expanding subgraphs. *Random Struct. Algorithms*, 52(1):158–178, 2018.
[55] J. Nešetřil and P. O. de Mendez. Tree-depth, subgraph coloring and homomorphism bounds. *Eur. J. Comb.*, 27(6):1022–1041, 2006.
[56] I. Newman and C. Sohler. Every property of hyperfinite graphs is testable. *SIAM J. Comput.*, 42(3):1095–1112, 2013.
[57] C. H. Papadimitriou and M. Yannakakis. Optimization, Approximation, and Complexity Classes. *J. Comput. Syst. Sci.*, 43(3):425–440, 1991.
[58] P. Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *Proc. 40th Annual ACM Symposium on Theory of Computing (STOC'08)*, pages 245–254, 2008.
[59] M. Romero, M. Wrochna, and S. Živný. Treewidth-pliability and PTAS for Max-CSPs. Technical report, 11 2019, arXiv:1911.03204.
[60] S. Sahni and T. F. Gonzalez. P-Complete Approximation Problems. *J. ACM*, 23(3):555–565, 1976.
[61] H. D. Sherali and W. P. Adams. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM J. Discrete Math.*, 3(3):411–430, 1990.
[62] J. Spencer. Counting extensions. *J. Comb. Theory, Ser. A*, 55(2):247–255, 1990.
[63] E. Szemérédi. Regular partitions of graphs. In *Proc. Colloq. Inter. CNRS*, pages 399–401, 1978.
[64] J. Thapper and S. Živný. The complexity of finite-valued CSPs. *J. ACM*, 63(4), 2016. Article No. 37.
[65] J. van den Heuvel, P. O. de Mendez, D. Quiroz, R. Rabinovich, and S. Siebertz. On the generalised colouring numbers of graphs that exclude a fixed minor. *Eur. J. Comb.*, 66:129–144, 2017, arXiv:1602.09052.