2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)

# Counting Small Induced Subgraphs Satisfying Monotone Properties

Marc Roth
*Merton College*
*University of Oxford*
*Oxford, United Kingdom*
*marc.roth@merton.ox.ac.uk*

Johannes Schmitt
*Mathematical Institute*
*University of Bonn*
*Bonn, Germany*
*schmitt@math.uni-bonn.de*

Philip Wellnitz
*Max Planck Institute for Informatics*
*Saarland Informatics Campus*
*Saarbrücken, Germany*
*wellnitz@mpi-inf.mpg.de*

***Abstract*—Given a graph property $\Phi$, we study the problem $\#\textsc{IndSub}(\Phi)$ which asks, on input a graph $G$ and a positive integer $k$, to compute the number $\#\mathsf{IndSub}(\Phi, k \to G)$ of induced subgraphs of size $k$ in $G$ that satisfy $\Phi$. The search for *explicit* criteria on $\Phi$ ensuring that $\#\textsc{IndSub}(\Phi)$ is hard was initiated by Jerrum and Meeks [J. Comput. Syst. Sci. 15] and is part of the major line of research on counting small patterns in graphs. However, apart from an implicit result due to Curticapean, Dell and Marx [STOC 17] proving that a full classification into "easy" and "hard" properties is possible and some partial results on edge-monotone properties due to Meeks [Discret. Appl. Math. 16] and Dörfler et al. [MFCS 19], not much is known.**

**In this work, we fully answer and explicitly classify the case of monotone, that is subgraph-closed, properties: We show that for any non-trivial monotone property $\Phi$, the problem $\#\textsc{IndSub}(\Phi)$ cannot be solved in time $f(k) \cdot |V(G)|^{o(k/\log^{1/2}(k))}$ for any function $f$, unless the Exponential Time Hypothesis fails. By this, we establish that any significant improvement over the brute-force approach is unlikely; in the language of parameterized complexity, we also obtain a $\#\mathsf{W}[1]$-completeness result.**

***Keywords*-Counting complexity, fine-grained complexity, graph homomorphisms, induced subgraphs, parameterized complexity**

## I. Introduction

Detection, enumeration and counting of patterns in graphs are among the most well-studied computational problems in theoretical computer science with a plethora of applications in diverse disciplines, including biology [51], [29], statistical physics [53], [36], [37], neural and social networks [44] and database theory [30], to name but a few. At the same time, those problems subsume in their unrestricted forms some of the most infamous NP-hard problems such as Hamiltonicity, the clique problem, or, more generally, the subgraph isomorphism problem [16], [56]. In the modern-day era of "big data", where even quadratic-time algorithms may count as inefficient, it is hence crucial to find relaxations of hard computational problems that allow for tractable instances.

A full version of this paper is available at arxiv.org/abs/2004.06595.

A very successful approach for a more fine-grained understanding of hard computational problems is a multivariate analysis of the complexity of the problem: Instead of establishing upper and (conditional) lower bounds only depending on the input size, we aim to find additional parameters that, in the best case, are small in real-world instances and allow for efficient algorithms if assumed to be bounded. In case of detection and counting of patterns in graphs, it turns out that the size of the pattern is often significantly smaller than the size of the graph: Consider as an example the evaluation of database queries. While a classical analysis of this problem requires considering instances where the size of the query is as large as the database, a multivariate analysis allows us to impose the restriction of the query being much smaller than the database, which is the case for real-world instances. More concretely, suppose we are given a query $\varphi$ of size $k$ and a database $B$ of size $n$, and we wish to evaluate the query $\varphi$ on $B$. Assume further, that we are given two algorithms for the problem: One has a running time of $O(n^k)$, and the other one has a running time of $O(2^k \cdot n)$. While, classically, both algorithms are inefficient in the sense that their running times are not bounded by a polynomial in the input size $n + k$, the second algorithm is significantly better than the first one for real-world instances and can even be considered efficient.

In this work, we focus on *counting* of small patterns in large graphs. The field of counting complexity was founded by Valiant's seminal result on the complexity of computing the permanent [57], [58], where it was shown that computing the number of perfect matchings in a graph is $\#$P-complete, and thus harder than every problem in the polynomial-time hierarchy PH [55]. This is in sharp contrast to the fact that *finding* a perfect matching in a graph can be done in polynomial-time [25]. Hence, a perfect matching is a pattern that allows for efficient detection but is unlikely to admit efficient counting. Initiated by Valiant, computational counting evolved into a well-studied subfield of theoretical computer science. In particular, it turns out that counting problems are closely related to computing partition functions in statistical physics [53], [37], [28], [15], [3].

2575-8454/20/$31.00 ©2020 IEEE
DOI 10.1109/FOCS46700.2020.00128

Indeed, one of the first algorithmic result in the field of computational counting is the famous FKT-Algorithm by the statistical physicists Fisher, Kasteleyn and Temperley [53], [37] that computes the partition function of the so-called dimer model on planar structures, which is essentially equivalent to computing the number of perfect matchings in a planar graph. The FKT-Algorithm is the foundation of the framework of holographic algorithms, which, among others, have been used to identify the tractable cases of a variety of complexity classifications for counting constraint satisfaction problems [59], [7], [5], [6], [31], [8], [9], [2]. Unfortunately, the intractable cases of those classifications indicate that, except for rare examples, counting is incredibly hard (from a complexity theory point of view). In particular, many efficiently solvable combinatorial decision problems turn out to be intractable in their counting versions, such as counting of satisfying assignments of monotone 2-CNFs [58], counting of independent sets in bipartite graphs [46] or counting of $s$-$t$-paths [58], to name but a few. For this reason, we follow the multivariate approach as outlined previously and restrict ourselves in this work on counting of *small* patterns in large graphs. Among others, problems of this kind find applications in neural and social networks [44], computational biology [1], [49], and database theory [24], [11], [12], [22].

Formally, we follow the approach of Jerrum and Meeks [33] and study the family of problems $\#\textsc{IndSub}(\Phi)$: Given a graph property $\Phi$, the problem $\#\textsc{IndSub}(\Phi)$ asks, on input a graph $G$ and a positive integer $k$, to compute the number of induced subgraphs of size $k$ in $G$ that satisfy $\Phi$.[1] As observed by Jerrum and Meeks, the generality of the definition allows to express counting of almost arbitrary patterns of size $k$ in a graph, subsuming counting of $k$-cliques and $k$-independent sets as very special cases.

Assuming that $\Phi$ is computable, we note that the problem $\#\textsc{IndSub}(\Phi)$ can be solved by brute-force in time $O(f(k)\cdot |V(G)|^k)$ for some function $f$ only depending on $\Phi$. The corresponding algorithm enumerates all subsets of $k$ vertices of $G$ and counts how many of those subsets satisfy $\Phi$. As we consider $k$ to be significantly smaller than $|V(G)|$, we are interested in the dependence of the exponent on $k$. More precisely, the goal is to find the best possible $g(k)$ such that $\#\textsc{IndSub}(\Phi)$ can be solved in time

$$O(f(k)\cdot |V(G)|^{g(k)}) \qquad (1)$$

for some function $f$ such that $f$ and $g$ only depend on $\Phi$. Readers familiar with parameterized complexity theory will identify the case of $g(k)\in O(1)$ as fixed-parameter tractability (FPT) results. We first provide some background and elaborate on the existing results on $\#\textsc{IndSub}(\Phi)$ before we present the contributions of this paper.

[1] Note that $\#\textsc{IndSub}(\Phi)$ is identical to $p$-$\#\textsc{UnlabelledInducedSubgraphWithProperty}(\Phi)$ as defined in [33].

### A. Prior Work

So far, the problem $\#\textsc{IndSub}(\Phi)$ has been investigated using primarily the framework of parameterized complexity theory. As indicated before, $\#\textsc{IndSub}(\Phi)$ is in FPT if the function $g$ in Eq. (1) is bounded by a constant (independent of $k$), and the problem is $\#\mathsf{W}[1]$-complete, if it is at least as hard as the parameterized clique problem; here $\#\mathsf{W}[1]$ should be considered a parameterized counting equivalent of NP and we provide the formal details in Section II. In particular, the so-called Exponential Time Hypothesis (ETH) implies that $\#\mathsf{W}[1]$-complete problems are not in FPT; again, this is made formal in Section II.

The problem $\#\textsc{IndSub}(\Phi)$ was first studied by Jerrum and Meeks [33]. They introduced the problem and proved that $\#\textsc{IndSub}(\Phi)$ is $\#\mathsf{W}[1]$-complete if $\Phi$ is the property of being connected. Implicitly, their proof also rules out the function $g$ of Eq. (1) being in $o(k)$, unless ETH fails, which establishes a tight conditional lower bound. In a subsequent line of research [34], [43], [35], Jerrum and Meeks proved $\#\textsc{IndSub}(\Phi)$ to be $\#\mathsf{W}[1]$-complete if at least one of the following is true:

1) The property $\Phi$ has *low edge-densities*; this is true for instance for all sparse properties such as planarity. This is made formal in the full version.
2) The property $\Phi$ holds for a graph $H$ if and only if the number of edges of $H$ is even/odd.
3) The property $\Phi$ is closed under the addition of edges, and the minimal elements have large treewidth.

Unfortunately, only the second of the previous results establishes a conditional lower bound that comes close to the upper bound given by the brute-force algorithm. This is particularly true due to the application of Ramsey's Theorem in the proofs of some of the prior results: Ramsey's Theorem states that there is a function $R(k) = 2^{\Theta(k)}$ such that every graph with at least $R(k)$ vertices contains either a $k$-independent set or a $k$-clique [47], [52], [26]. Relying on this result for a reduction from finding or counting $k$-independent sets or $k$-cliques, the best implicit conditional lower bounds achieved only rule out an algorithm running in time $f(k)\cdot |V(G)|^{o(\log k)}$ for any function $f$.

Moreover, the previous results only apply to a very specific set of properties. In particular, Jerrum and Meeks posed the following open problem concerning a generalization of the second result (2); we say that $\Phi$ is $k$-trivial, if it is either true or false for all graphs with $k$ vertices.

**Conjecture 1** ([34], [35])**.** *Let $\Phi$ be a graph property that only depends on the number of edges of a graph. If for infinitely many $k$ the property $\Phi$ is not $k$-trivial, then $\#\textsc{IndSub}(\Phi)$ is $\#\mathsf{W}[1]$-complete.*

Note that the condition of $\Phi$ not being $k$-trivial for infinitely many $k$ is necessary for hardness, as otherwise, the problem becomes trivial if $k$ exceeds a constant depending only on $\Phi$.

The first major breakthrough towards a complete understanding of the complexity of $\#\text{IndSub}(\Phi)$ is the following implicit classification due to Curticapean, Dell and Marx [18]:

**Theorem 2** ([18])**.** *Let $\Phi$ denote a graph property. Then the problem $\#\text{IndSub}(\Phi)$ is either* FPT *or* $\#\mathsf{W}[1]$*-complete.*

While the previous classification provides a very strong result for the structural complexity of $\#\text{IndSub}(\Phi)$, it leaves open the question of the precise bound on the function $g$. Furthermore, it is implicit in the sense that it does not reveal the complexity of $\#\text{IndSub}(\Phi)$ if a concrete property $\Phi$ is given. Nevertheless, the technique introduced by Curticapean, Dell and Marx, which is now called *Complexity Monotonicity*, turned out to be the right approach for the treatment of $\#\text{IndSub}(\Phi)$. In particular, the subsequent results on $\#\text{IndSub}(\Phi)$, including the classification in this work, have been obtained by strong refinements of *Complexity Monotonicity*; we provide a brief introduction when we discuss the techniques used in this paper. More concretely, a superset of the authors established the following classifications for edge-monotone properties in recent years [48], [23]: The problem $\#\text{IndSub}(\Phi)$ is $\#\mathsf{W}[1]$-complete and, assuming ETH, cannot be solved in time $f(k)\cdot|V(G)|^{o(k)}$ for any function $f$, if at least one of the following is true:[2]

- The property $\Phi$ is non-trivial, closed under the removal of edges and false on odd cycles.
- The property $\Phi$ is non-trivial on bipartite graphs and closed under the removal of edges.

While the second result completely answers the case of edge-monotone properties on bipartite graphs, a general classification of edge-monotone properties is still unknown.

### B. Our Results

We begin with monotone properties, that is, properties that are closed under taking subgraphs. We classify those properties completely and explicitly; the following theorem establishes hardness and almost tight conditional lower bounds.

**Main Theorem 1.** *Let $\Phi$ denote a monotone graph property. Suppose that for infinitely many $k$ the property $\Phi$ is not $k$-trivial. Then $\#\text{IndSub}(\Phi)$ is $\#\mathsf{W}[1]$-complete and cannot be solved in time $f(k)\cdot|V(G)|^{o(k/\sqrt{\log k})}$ for any function $f$, unless ETH fails.*

In fact, we obtain a tight bound, that is, we can drop the factor of $1/\sqrt{\log k}$ in the exponent, assuming the conjecture that "you cannot beat treewidth" [42], which we formally introduce and discuss later in the paper.

[2]We provide simplified statements; the formal and more general results can be found in [48], [23].

A concrete example of a property that is classified by Main Theorem 1, but that was not classified before, is the (monotone) property of being 3-colorable.

We remark further, that $\#\mathsf{W}[1]$-completeness in Main Theorem 1 is not surprising, as the decision version of $\#\text{IndSub}(\Phi)$ was implicitly shown to be $\mathsf{W}[1]$-complete by Khot and Raman [38]; $\mathsf{W}[1]$ is the decision version of $\#\mathsf{W}[1]$ and should be considered a parameterized equivalent of NP. However, their reduction is not parsimonious. Also, their proof uses Ramsey's Theorem and thus only yields an implicit conditional lower bound of $f(k)\cdot|V(G)|^{o(\log k)}$, whereas our lower bound is almost tight.

Our second result establishes an almost tight lower bound for sparse properties, that is, properties $\Phi$ that admit a constant $s$ such that every graph $H$ for which $\Phi$ holds has at most $s\cdot|V(H)|$ many edges. Furthermore, the bound can be made tight if the set $\mathcal{K}(\Phi)$ of positive integers $k$ for which $\Phi$ is not $k$-trivial is additionally *dense*. By this we mean that there is a constant $\ell$ such that for every positive integer $n$, there exists $n \le k \le \ell n$ such that $\Phi$ is not $k$-trivial. Note that density rules out artificial properties such as $\Phi(H)=1$ if and only if $H$ is an independent set and has precisely $2\uparrow n$ vertices for some positive integer $n$, with $2\uparrow n$ the $n$-fold exponential tower with base 2. In particular, for every property $\Phi$ that is $k$-trivial only for finitely many $k$, the set $\mathcal{K}(\Phi)$ is dense.

**Main Theorem 2.** *Let $\Phi$ denote a sparse graph property such that $\Phi$ is not $k$-trivial for infinitely many $k$. Then, $\#\text{IndSub}(\Phi)$ is $\#\mathsf{W}[1]$-complete and cannot be solved in time $f(k)\cdot|V(G)|^{o(k/\log k)}$ for any function $f$, unless ETH fails. If $\mathcal{K}(\Phi)$ is additionally dense, then $\#\text{IndSub}(\Phi)$ cannot be solved in time $f(k)\cdot|V(G)|^{o(k)}$ for any function $f$, unless ETH fails.*

Our third result solves the open problem posed by Jerrum and Meeks by proving that (a strengthened version of) Conjecture 1 is true.

**Main Theorem 3.** *Let $\Phi$ denote a computable graph property that only depends on the number of edges of a graph. If $\Phi$ is not $k$-trivial for infinitely many $k$, then $\#\text{IndSub}(\Phi)$ is $\#\mathsf{W}[1]$-complete and cannot be solved in time $f(k)\cdot|V(G)|^{o(k/\log k)}$ for any function $f$, unless ETH fails. If $\mathcal{K}(\Phi)$ is additionally dense, then $\#\text{IndSub}(\Phi)$ cannot be solved in time $f(k)\cdot|V(G)|^{o(k/\sqrt{\log k})}$ for any function $f$, unless ETH fails.*

Note that, similar to Main Theorem 1, the conditional lower bounds in the previous two theorems become tight, if "you cannot beat treewidth" [42]; in particular, the condition of being dense can be removed in that case.
Finally, we consider properties that are *hereditary*, that is, closed under taking *induced* subgraphs. We obtain a criterion on such graph properties that, if satisfied, yields a tight conditional lower bound for the complexity of $\#\text{IndSub}(\Phi)$.

While the statement of the criterion is deferred to the technical discussion, we can see that every hereditary property that is defined by a single forbidden induced subgraph satisfies the criterion.

**Main Theorem 4.** *Let $H$ be a graph with at least 2 vertices and let $\Phi$ denote the property of being $H$-free, that is, a graph satisfies $\Phi$ if and only if it does not contain $H$ as an induced subgraph. Then, $\#\textsc{IndSub}(\Phi)$ is $\#\mathsf{W}[1]$-complete and cannot be solved in time $f(k) \cdot |V(G)|^{o(k)}$ for any function $f$, unless ETH fails.*

Note that the case of $H$ being the graph with one vertex, which is excluded above, yields the property $\Phi$ which is false on all graphs $G$ with at least one vertex, for which $\#\textsc{IndSub}(\Phi)$ is the constant zero-function and thus trivially solvable. Hence, Main Theorem 4 establishes indeed a complete classification for all properties $\Phi$="$H$-free".

### C. Technical Overview

We rely on *Complexity Monotonicity* of computing linear combinations of homomorphism counts [18]. More precisely, it is known that for every computable graph property $\Phi$ and positive integer $k$, there exists a unique computable function $a$ from graphs to rational numbers such that for all graphs $G$

$$\#\mathsf{IndSub}(\Phi, k \to G) = \sum_H a(H) \cdot \#\mathsf{Hom}(H \to G)\,, \quad (2)$$

where $\#\mathsf{IndSub}(\Phi, k \to G)$ denotes the number of induced subgraphs of size $k$ in $G$ that satisfy $\Phi$, and $\#\mathsf{Hom}(H \to G)$ denotes the number of graph homomorphisms from $H$ to $G$. It is known that the function $a$ has finite support, that is, there is only a finite number of graphs $H$ for which $a(H) \neq 0$.

Intuitively, Complexity Monotonicity states that computing a linear combination of homomorphism counts is *precisely* as hard as computing its hardest term [18]. Furthermore, the complexity of computing the number of homomorphisms from a small graph $H$ to a large graph $G$ is (almost) precisely understood [21], [42]: Roughly speaking, it is possible to compute $\#\mathsf{Hom}(H \to G)$ efficiently if and only if $H$ has small treewidth. As a consequence, the complexity of computing $\#\mathsf{IndSub}(\Phi, k \to G)$ is precisely determined by the support of the function $a$. Unfortunately, determining the latter turned out to be an incredibly hard task: It was shown in [48] and [23] that the function $a$ subsumes a variety of algebraic and even topological invariants. As a concrete example, a subset of the authors showed that for edge-monotone properties $\Phi$, the coefficient $a(K_k)$ of the complete graph in Eq. (2) is, up to a factor of $k!$, equal to the reduced Euler characteristic of what is called the simplicial graph complex of $\Phi$ and $k$ [48].

The previous example illustrated that identifying the support of the function $a$ in Eq. (2) is a hard task, but using the framework of Complexity Monotonicity requires us to solve this task. In this work, we present a solution for properties whose $f$-vectors (see below) have low Hamming weight: Given a property $\Phi$ and a positive integer $k$, we define a $\binom{k}{2}+1$ dimensional vector $f^{\Phi,k}$ by setting $f_i^{\Phi,k}$ to be the number of edge-subsets of size $i$ of the complete graph with $k$ vertices such that the induced graph satisfies $\Phi$, that is,

$$f_i^{\Phi,k} := \#\{A \subseteq E(K_k) \mid \#A = i \wedge \Phi(K_k[A]) = 1\}$$

for all $i = 0, \dots, \binom{k}{2}$. By this, we lift the notion of $f$-vectors from abstract simplicial complexes to graph properties; readers familiar with the latter will observe that the $f$-vector of an edge-monotone property $\Phi$ equals the $f$-vector of its associated graph complex (see for instance [4]). Similarly, we introduce the notions of $h$-vectors $h^{\Phi,k}$ and $f$-polynomials $\mathtt{f}_{\Phi,k}$ of graph properties, defined as follows; we set $d = \binom{k}{2}$.

$$h_\ell^{\Phi,k} := \sum_{i=0}^{\ell} (-1)^{\ell-i} \cdot \binom{d-i}{\ell-i} \cdot f_i^{\Phi,k}, \text{ where } \ell \in \{0,\dots,d\};$$

$$\text{and} \quad \mathtt{f}_{\Phi,k}(x) := \sum_{i=0}^{d} f_i^{\Phi,k} \cdot x^{d-i}.$$

Our main combinatorial insight relates the function $a$ in Eq. (2) to the $h$-vector of $\Phi$. For the formal statement, we let $\mathcal{H}(\Phi,k,i)$ denote the set of all graphs $H$ with $k$ vertices and $i$ edges that satisfy $\Phi$. We then show that for all $i = 0,\dots,d$, we have

$$k! \sum_{H \in \mathcal{H}(\Phi,k,i)} a(H) = h_i^{\Phi,k}.$$

In particular, the previous equation shows that there is a graph $H$ with $i$ edges that survives with a non-zero coefficient $a(H)$ in Eq. (2) whenever the $i$-th entry of the $h$-vector $h^{\Phi,k}$ is non-zero. As graphs with many edges have high treewidth, we can thus establish hardness of computing $\#\mathsf{IndSub}(\Phi, k \to G)$ by proving that there is a non-zero entry with a high index in $h^{\Phi,k}$. To this end, we relate $h^{\Phi,k}$ and $f^{\Phi,k}$ by observing that their entries are evaluations of the derivatives of the $f$-polynomial $\mathtt{f}_{\Phi,k}(x)$. More concretely, our goal is to show that a large amount of high-indexed zero entries of $h^{\Phi,k}$ yields that the only polynomial of degree at most $d$ that satisfies the constrains given by the evaluations of the derivatives is the zero polynomial. However, the latter can only be true if $\Phi$ is trivially false on $k$-vertex graphs. Using Hermite-Birkhoff interpolation and Pólya's Theorem we are able to achieve this goal whenever the Hamming weight of $f^{\Phi,k}$ is small. Our meta-theorem thus classifies the complexity of $\#\textsc{IndSub}(\Phi)$ in terms of the Hamming weight of the $f$-vectors of $\Phi$.

**Main Theorem 5.** *Let $\Phi$ denote a computable graph property and suppose that $\Phi$ is not $k$-trivial for infinitely many $k$. Let $\beta : \mathcal{K}(\Phi) \to \mathbb{Z}_{\geq 0}$ denote the function that maps $k$ to $\binom{k}{2} - \mathsf{hw}(f^{\Phi,k})$. If $\beta(k) \in \omega(k)$ then the problem $\#\textsc{IndSub}(\Phi)$ is $\#\mathsf{W}[1]$-complete and cannot be solved in time*

$$g(k) \cdot n^{o((\beta(k)/k)/(\log(\beta(k)/k)))}$$

*for any function $g$, unless ETH fails.*

For the refined conditional lower bounds in case of monotone properties and properties for which the set $\mathcal{K}(\Phi)$ is dense (see Main Theorems 1 to 3), we furthermore rely on a consequence of the Kostochka-Thomason-Theorem [39], [54] that establishes a lower bound on the size of the smallest clique-minors of graphs with many edges.

In contrast to the previous families of properties, we do not rely on the general meta-theorem (Main Theorem 5) for our treatment of hereditary properties. Instead, we carefully construct a reduction from counting $k$-independent sets in bipartite graphs. Due to space constraints, the discussion of hereditary properties is deferred to the full version.

## II. Preliminaries

For a positive integer $r$, we set $[r] := \{1, \ldots, r\}$. Given a finite set $S$, we write $\#S$ and $|S|$ for the cardinality of $S$. For a vector $f \in \mathbb{Q}^n$, we define its *Hamming weight* $\mathsf{hw}(f)$ as the number of non-zero entries of $f$.

Graphs in this work are simple and do not contain self-loops. Given a graph $G$, we write $V(G)$ for the vertices of $G$ and $E(G)$ for the edges of $G$. We write $\mathcal{G}$ for the set of all (isomorphism classes of) graphs. For a graph $G$, we define its *complement* $\overline{G}$ by $\overline{G} := (V(G), \overline{E}(G) \setminus \{\{v,v\} \mid v \in V(G)\})$. Given a subset $\hat{E}$ of edges of a graph $G$, we write $G[\hat{E}]$ for the graph with vertices $V(G)$ and edges $\hat{E}$. Given a subset $\hat{V}$ of vertices of a graph $G$, we write $G[\hat{V}]$ for the graph with vertices $\hat{V}$ and edges $E(G) \cap \hat{V}^2$. In particular, we say that $G[\hat{V}]$ is an *induced subgraph* of $G$. Given graphs $H$ and $G$, we write $\mathsf{IndSub}(H \to G)$ for the set of all induced subgraphs of $G$ that are isomorphic to $H$.

Given graphs $H$ and $G$, a *homomorphism* from $H$ to $G$ is a function $\varphi : V(H) \to V(G)$ such that $\{\varphi(u), \varphi(v)\} \in E(G)$ whenever $\{u, v\} \in E(H)$. We write $\mathsf{Hom}(H \to G)$ for the set of all homomorphisms from $H$ to $G$. In particular, we write $\#\mathsf{Hom}(H \to \star)$ for the function that maps a graph $G$ to $\#\mathsf{Hom}(H \to G)$. A bijective homomorphism from a graph $H$ to itself is an *automorphism* and we write $\mathsf{Aut}(H)$ to denote the set of all automorphisms of $H$.

For a graph $H$, we define the *average degree* of $H$ as

$$d(H) := \frac{1}{|V(H)|} \cdot \sum_{v \in V(H)} \mathsf{deg}(v).$$

Further, we rely on the *treewidth* of a graph, which is a graph parameter $\mathsf{tw} : \mathcal{G} \to \mathbb{N}$. As we only work with the treewidth in a black-box manner, we omit the definition here and refer the interested reader to the literature, (see for instance [20, Chapter 7]). Specifically, we use the following well-known result from extremal graph theory, which relates the treewidth of a graph $H$ to its average degree.

**Lemma 3** (Folklore, see for instance [10, Corollary 1])**.** *Any graph $H$ with average degree at least $d$ satisfies $\mathsf{tw}(H) \geq \frac{d}{2}$.*

Finally, we also rely on the following celebrated result from extremal graph theory:

**Theorem 4** (Turán's Theorem, see for instance [41, Section 2.1])**.** *Any graph $H$ with more than $\left(1 - \frac{1}{r}\right) \cdot \frac{1}{2}|V(H)|^2$ edges contains the clique $K_{r+1}$ as a subgraph.*

*Graph Properties:* A *graph property* $\Phi$ is a function from graphs to $\{0,1\}$ such that $\Phi(H) = \Phi(G)$ whenever $H$ and $G$ are isomorphic. We say that a graph $H$ *satisfies* $\Phi$ if $\Phi(H) = 1$. Given a positive integer $k$ and a graph property $\Phi$, we write $\Phi_k$ for the set of all (isomorphism classes of) graphs with $k$ vertices that satisfy $\Phi$. Given a graph $G$, we write $\mathsf{IndSub}(\Phi, k \to G)$ for the set of all induced subgraphs with $k$ vertices of $G$ that satisfy $\Phi$. In particular, we write $\#\mathsf{IndSub}(\Phi, k \to \star)$ for the function that maps a graph $G$ to $\#\mathsf{IndSub}(\Phi, k \to G)$.

For a graph property $\Phi$, we set $\neg\Phi(H) = 1 :\Leftrightarrow \Phi(H) = 0$; we call $\neg\Phi(H)$ the *negation* of $\Phi$. We set $\overline{\Phi}(H) = 1 :\Leftrightarrow \Phi(\overline{H}) = 1$; we call $\overline{\Phi}(H)$ the *inverse* of $\Phi$.

**Fact 5.** *For every graph property $\Phi$, graph $G$ and positive integer $k$, we have the equalities*

$$\#\mathsf{IndSub}(\neg\Phi, k \to G) = \binom{|V(G)|}{k} - \#\mathsf{IndSub}(\Phi, k \to G),$$

$$\text{and} \quad \#\mathsf{IndSub}(\overline{\Phi}, k \to G) = \#\mathsf{IndSub}(\Phi, k \to \overline{G})\,.$$

*Proof:* The first identity is immediate. For the second identity, using $\#\mathsf{IndSub}(H \to G) = \#\mathsf{IndSub}(\overline{H} \to \overline{G})$ from [41, Section 5.2.3], we observe

$$\begin{aligned} &\#\mathsf{IndSub}(\overline{\Phi}, k \to G) \\ &= \sum_{H \in \overline{\Phi}_k} \#\mathsf{IndSub}(H \to G) = \sum_{\overline{H} \in \Phi_k} \#\mathsf{IndSub}(H \to G) \\ &= \sum_{\overline{H} \in \Phi_k} \#\mathsf{IndSub}(\overline{H} \to \overline{G}) = \#\mathsf{IndSub}(\Phi, k \to \overline{G}). \end{aligned}$$

■

*Fine-Grained and Parameterized Complexity Theory:* Let $\Phi$ denote a computable graph property. In the problem $\#\textsc{IndSub}(\Phi)$ the task for a given graph $G$ with $n$ vertices and a positive integer $k$ is to compute the number $\#\mathsf{IndSub}(\Phi, k \to G)$ of induced subgraphs of size $k$ in $G$ that satisfy $\Phi$. Note that we can solve $\#\textsc{IndSub}(\Phi)$ by brute-force in time $f(k) \cdot O(n^k)$.

As elaborated in the introduction, our goal is to understand the complexity of $\#\textsc{IndSub}(\Phi)$ for instances with small $k$ and large $n$. More precisely, we wish to identify the best possible exponent of $n$ in the running time. To this end, we rely on the frameworks of fine-grained and parameterized complexity theory. Regarding the former, we prove conditional lower bounds based on the *Exponential Time Hypothesis* due to Impagliazzo and Paturi [32]:

**Conjecture 6** (Exponential Time Hypothesis (ETH))**.** *The problem* 3-SAT *cannot be solved in time* $\mathsf{exp}(o(m))$*, where* $m$ *is the number of clauses of the input formula.*

Assuming ETH, we are able to prove that the exponent ($k$) of the brute-force algorithm for $\#\textsc{IndSub}(\Phi)$ cannot be improved significantly for non-trivial monotone properties by establishing that no algorithm with a running time of $f(k) \cdot |V(G)|^{o(k/\sqrt{\log k})}$ for any function $f$ exists.

In the language of parameterized complexity theory, our reductions also yield $\#\mathsf{W}[1]$-completeness results, where $\#\mathsf{W}[1]$ should be considered as the parameterized counting equivalent of NP; we provide a rough introduction in what follows and refer the interested reader to references like [20] and [27] for a detailed treatment.

A *parameterized counting problem* is a pair of a function $P : \Sigma^* \to \mathbb{N}$ and a computable parameterization $\kappa : \Sigma^* \to \mathbb{N}$. Examples include the problems $\#\textsc{VertexCover}$ and $\#\textsc{Clique}$ which ask, given a graph $G$ and a positive integer $k$, to compute the number $P(G, k)$ of vertex covers or cliques, respectively, of size $k$. Both problems are parameterized by the solution size, that is $\kappa(G, k) := k$. Similarly, the problem $\#\textsc{IndSub}(\Phi)$ can be viewed as a parameterized counting problem when parameterized by $\kappa(G, k) := k$; we implicitly assume this parameterization of $\#\textsc{IndSub}(\Phi)$ in the remainder of this paper.

We say that a parameterized counting problem $P$ is *fixed-parameter tractable* (FPT) if there is a computable function $f$ such that $P$ can be solved in time $f(\kappa(x)) \cdot |x|^{O(1)}$, where $|x|$ is the input size. Given two parameterized counting problems $(P, \kappa)$ and $(\hat{P}, \hat{\kappa})$, a *parameterized Turing-reduction* from $(P, \kappa)$ to $(\hat{P}, \hat{\kappa})$ is an algorithm $\mathbb{A}$ that is given oracle access to $\hat{P}$ and, on input $x$, computes $P(x)$ in time $f(\kappa(x)) \cdot |x|^{O(1)}$ for some computable function $f$; the parameter $\kappa(y)$ of every oracle query posed by $\mathbb{A}$ must be bounded by $g(\kappa(x))$ for some computable function $g$.

While $\#\textsc{VertexCover}$ is known to be fixed-parameter tractable [27], $\#\textsc{Clique}$ is not fixed-parameter tractable, unless ETH fails [13], [14]. Moreover, $\#\textsc{Clique}$ is the canonical complete problem for the parameterized complexity class $\#\mathsf{W}[1]$, see [27]; in particular, we use the following definition of $\#\mathsf{W}[1]$-completeness in this work.

**Definition 7.** *A parameterized counting problem is* $\#\mathsf{W}[1]$-complete *if it is interreducible with* $\#\textsc{Clique}$ *with respect to parameterized Turing-reductions.*

Note that the absence of an FPT algorithm for $\#\textsc{Clique}$ under ETH and the definition of parameterized Turing-reductions yield that $\#\mathsf{W}[1]$-complete problems are not fixed-parameter tractable, unless ETH fails. This legitimizes the notion of $\#\mathsf{W}[1]$-completeness as evidence for (fixed-parameter) intractability. Jerrum and Meeks [33] have shown that $\#\textsc{IndSub}(\Phi)$ reduces to $\#\textsc{Clique}$ for every computable property $\Phi$ with respect to parameterized Turing-reductions. Thus we will only treat the "hardness part" of the $\#\mathsf{W}[1]$-completeness results in this paper.

The fine-grained and parameterized complexity of the homomorphism counting problem are the foundation of the lower bounds established in this work: Given a class of graphs $\mathcal{H}$, the problem $\#\textsc{Hom}(\mathcal{H})$ asks, on input a graph $H \in \mathcal{H}$ and an arbitrary graph $G$, to compute $\#\mathsf{Hom}(H \to G)$; the parameter is $|V(H)|$. Roughly speaking, the complexity of $\#\textsc{Hom}(\mathcal{H})$ is determined by the treewidth of the graphs in $\mathcal{H}$:

**Theorem 8** ([21], [42])**.** *Let* $\mathcal{H}$ *denote a recursively enumerable class of graphs. If the treewidth of* $\mathcal{H}$ *is bounded by a constant, then* $\#\textsc{Hom}(\mathcal{H})$ *is solvable in polynomial time. Otherwise, the problem is* $\#\mathsf{W}[1]$*-complete and cannot be solved in time* $f(|V(H)|) \cdot |V(G)|^{o(\mathsf{tw}(H)/\log \mathsf{tw}(H))}$ *for any function* $f$*, unless ETH fails.*

Note that the classification of $\#\textsc{Hom}(\mathcal{H})$ into polynomial-time and $\#\mathsf{W}[1]$-complete cases is explicitly stated and proved in the work of Dalmau and Jonson [21]. However, the conditional lower bound follows only implicitly by a result of Marx [42]. The question whether the lower bound from Theorem 8 can be strengthened to $f(|V(H)|) \cdot |V(G)|^{o(\mathsf{tw}(H))}$ is known as "Can you beat treewidth?" and constitutes a major open problem in parameterized complexity theory and an obstruction for tight conditional lower bounds on the complexity of a variety of (parameterized) problems, see for instance [40], [17], [19], [18].

As described in the introduction, the complexity of computing a finite linear combination of homomorphism counts is precisely determined by the complexity of computing the non-vanishing terms. The formal statement is provided subsequently.

**Theorem 9** (Complexity Monotonicity [12], [18])**.** *Let* $a : \mathcal{G} \to \mathbb{Q}$ *denote a function of finite support and let* $F$ *denote a graph such that* $a(F) \neq 0$*. There are a computable function* $g$ *and a deterministic algorithm* $\mathbb{A}$ *with oracle access to the function*

$$G \mapsto \sum_{H \in \mathcal{G}} a(H) \cdot \#\mathsf{Hom}(H \to G),$$

*and which, given a graph* $G$ *with* $n$ *vertices, computes* $\#\mathsf{Hom}(F \to G)$ *in time* $g(a) \cdot n^c$*, where* $c$ *is a constant independent of* $a$*. Furthermore, each queried graph has at most* $g(a) \cdot n$ *vertices.*

As observed by Curticapean, Dell and Marx [18], counting induced subgraphs of size $k$ that satisfy $\Phi$ is equivalent to computing a finite linear combination of homomorphism counts. Thus, the previous results yield an *implicit* dichotomy for $\#\text{IndSub}(\Phi)$.

**Theorem 10** ([18])**.** *Let $\Phi$ denote a computable graph property and let $k$ denote a positive integer. There is a unique and computable function $a : \mathcal{G} \to \mathbb{Q}$ of finite support such that*

$$\#\mathsf{IndSub}(\Phi, k \to \star) = \sum_{H \in \mathcal{G}} a(H) \cdot \#\mathsf{Hom}(H \to \star).$$

*Furthermore, the problem $\#\text{IndSub}(\Phi)$ is either fixed-parameter tractable or $\#\mathsf{W}[1]$-complete.*

Note that the result on $\#\text{IndSub}(\Phi)$ in the previous theorem does not concern the fine-grained complexity of the problem. To reveal the latter, it is necessary to understand the support of the function $a$; we tackle this task in detail in Section III.

***f*-Vectors and *h*-Vectors:** It was observed in [48] that there is a close connection between the structure of the simplicial graph complex of edge-monotone properties $\Phi$ and the complexity of $\#\text{IndSub}(\Phi)$. In this work, we generalize two important topological invariants of simplicial complexes to arbitrary graph properties: The $f$-vector and the $h$-vector.

**Definition 11.** *Let $\Phi$ denote a graph property, let $k$ denote a positive integer and set $d = \binom{k}{2}$. The* $f$-vector $f^{\Phi,k} = (f_i^{\Phi,k})_{i=0}^{d}$ *of $\Phi$ and $k$ is defined by*

$$f_i^{\Phi,k} := \#\{A \subseteq E(K_k) \mid \#A = i \wedge \Phi(K_k[A]) = 1\},$$

*where $i \in \{0, \dots, d\}$. That is, $f_i^{\Phi,k}$ is the number of edge-subsets of size $i$ of $K_k$ such that the induced graph satisfies $\Phi$.*

*The* $h$-vector $h^{\Phi,k} = (h_\ell^{\Phi,k})_{\ell=0}^{d}$ *is defined by*

$$h_\ell^{\Phi,k} := \sum_{i=0}^{\ell} (-1)^{\ell-i} \cdot \binom{d-i}{\ell-i} \cdot f_i^{\Phi,k}, \text{ where } \ell \in \{0, \dots, d\}.$$

As mentioned before, note that those notions of $f$ and $h$-vectors correspond to the eponymous notions for simplicial (graph) complexes.[3] We omit the definition of the latter as we are only concerned with the generalized notions and refer the interested reader for instance to [4].

It turns out that the *non-vanishing* of suitable entries $h_\ell^{\Phi,k}$ of the $h$-vector implies hardness for $\#\text{IndSub}(\Phi)$. The result in [48] can be considered as a very restricted special case as it shows that the non-vanishing of the reduced Euler characteristic of the complex (which is equal to the entry $h_d^{\Phi,k}$) implies hardness. On the other hand, for many graph properties it is easy to deduce information about the $f$-vector (for instance that $f_\ell^{\Phi,k} = 0$ for sufficiently large $\ell$ with respect to $k$). We observe that the $f$ and $h$-vectors of a graph property are related by the so-called the $f$-polynomial which is again a generalization of the epynomous notion for simplicial complexes:

**Definition 12.** *Let $\Phi$ denote a graph property, let $k$ denote a positive integer and set $d = \binom{k}{2}$. The* $f$-polynomial *of $\Phi$ and $k$ is a univariate polynomial of degree at most $d$ defined as follows:*

$$\mathtt{f}_{\Phi,k}(x) := \sum_{i=0}^{d} f_i^{\Phi,k} \cdot x^{d-i}.$$

As we see in the proof of Lemma 16, the entries of the $f$ and $h$-vectors are given up to combinatorial factors by derivatives of the $f$-polynomial at 0 and $-1$. Intuitively, we apply Hermite-Birkhoff interpolation (which we introduce below) on $\mathtt{f}_{\Phi,k}$ and its derivatives to prove that specific entries of $h^{\Phi,k}$ cannot vanish in case a sufficient number of entries of $f^{\Phi,k}$ do, unless $\Phi$ is trivially false on $k$-vertex graphs.

*Hermite-Birkhoff Interpolation and Pólya's Theorem:* While a univariate polynomial of degree $d$ is uniquely determined by $d+1$ evaluations in pairwise different points, the problem of *Hermite-Birkhoff interpolation* asks under which conditions we can uniquely recover the polynomial if we instead impose conditions on the derivatives of the polynomial at $m$ distinct points. Following the notation of Schoenberg [50], the problem is formally expressed as follows. Given a matrix $E = (\varepsilon_{ij}) \in \{0,1\}^{m \times d+1}$ where $i \in \{1, \dots, m\}$ and $j \in \{0, \dots d\}$, as well as reals $x_1 < \dots < x_m$, the goal is to find a polynomial $\mathtt{f}$ of degree at most $d$ such that for all $i$ and $j$ with $\varepsilon_{ij} = 1$ we have

$$\mathtt{f}^{(j)}(x_i) = 0$$

Here, $\mathtt{f}^{(j)}$ denotes the $j$-th derivative of $\mathtt{f}$. In particular, we are interested under which conditions on the matrix $E$, the zero polynomial is the *unique* solution. In this case, $E$ is called *poised*. It turns out that the case $m = 2$ is sufficient for our purposes; fortunately, this case was fully solved by Pólya:

**Theorem 13** (Pólya's Theorem [45], [50])**.** *Let $E$ be defined as above with $m = 2$. Suppose that $\sum_{i,j} \varepsilon_{ij} = d + 1$ and for every $j \in \{0, \dots, d\}$ set*

$$M_j := \sum_{i=0}^{j} \varepsilon_{1,i} + \varepsilon_{2,i}.$$

*Then, $E$ is poised if and only if $M_j \geq j + 1$ holds true for all $j \in \{0, \dots, d-1\}$.*

[3] In some parts of the literature, the $f$-vector comes with an index shift of $-1$ due to the topological interpretation of simplicial complexes.

## III. Homomorphism Vectors of Graph Properties

In this section we present our main technical result:

**Theorem 14.** *Let $\Phi$ denote a computable graph property, let $k$ denote a positive integer, and let $w$ denote the Hamming weight of the $f$-vector $f^{\Phi,k}$. If $\Phi$ is not trivially false on $k$-vertex graphs, then there is a unique and computable function $a : \mathcal{G} \to \mathbb{Q}$ of finite support such that*

$$\#\mathsf{IndSub}(\Phi, k \to \star) = \sum_{H \in \mathcal{G}} a(H) \cdot \#\mathsf{Hom}(H \to \star),$$

*satisfying that there is a graph $K$ on $k$ vertices and at least $\binom{k}{2} - w + 1$ edges such that $a(K) \neq 0$.*

First, recall from Theorem 10 that for any computable graph property $\Phi$ and positive integer $k$, there is a unique computable function $a : \mathcal{G} \to \mathbb{Q}$ (with finite support) satisfying

$$\#\mathsf{IndSub}(\Phi, k \to \star) = \sum_{H \in \mathcal{G}} a(H) \cdot \#\mathsf{Hom}(H \to \star). \quad (3)$$

Now, for the remainder of the section, fix a (computable) graph property $\Phi$ and a positive integer $k$ (and thus the function $a$). This allows us to simplify the notation for the $f$ and $h$-vectors, as well as for the $f$-polynomial: We write $f := f^{\Phi,k}$, $h := h^{\Phi,k}$, and $\mathtt{f} := \mathtt{f}_{\Phi,k}$. Furthermore, we set $d := \binom{k}{2}$ and we write $\mathcal{H}_i$ for the set of all graphs on $k$ vertices and with $i$ edges.

Next, we define the vector $\tilde{h}_i$ as

$$\tilde{h}_i := \sum_{K \in \mathcal{H}_i} a(K)\,, \text{ where } i \in \{0, \dots, d\},$$

that is, the $i$-th entry of $\tilde{h}$ is the sum of the coefficients of graphs with $k$ vertices and $i$ edges in Eq. (3). Now we present the aforementioned connection between the coefficients of Eq. (3) and the $h$-vector of the property $\Phi$; due to space constraints, we defer the proof to the full version.

**Lemma 15.** *We have $k! \cdot \tilde{h} = h$.*

Note that as a consequence, the $h$-vector of a simplicial graph complex is determined by the coefficients of its associated linear combination of homomorphisms. In the next step, we use Pólya's Theorem to prove that the Hamming weight of the $f$-vector determines an index $\beta$ of the $h$-vector such that at least one entry of $h$ with index at least $\beta$ is non-zero. By Lemma 15 the same then follows for $\tilde{h}$.

**Lemma 16.** *Let $w$ denote the Hamming weight of $f$ and set $\beta = d - w$. If $\Phi$ is not trivially false on $k$-vertex graphs then at least one of the values $h_d, \dots, h_{\beta+1}$ is non-zero.*

*Proof:* Recall the definition of the $f$-polynomial

$$\mathtt{f}(x) = \sum_{i=0}^{d} f_i \cdot x^{d-i}\,.$$

Now observe that

$$\mathtt{f}^{(j)}(x) = \sum_{i=0}^{d-j} f_i \cdot \frac{(d-i)!}{(d-i-j)!} \cdot x^{d-j-i}.$$

At $x = 0$ we obtain $\mathtt{f}^{(j)}(0) = f_{d-j} \cdot j!$. Therefore, by assumption, we have $\mathtt{f}^{(j)}(0) = 0$ for $\beta + 1$ many indices $j$.

Furthermore, we see that

$$\begin{aligned}\mathtt{f}^{(j)}(-1) &= \sum_{i=0}^{d-j} f_i \cdot \frac{(d-i)!}{(d-i-j)!} \cdot (-1)^{d-j-i} \\ &= j! \cdot \sum_{i=0}^{d-j} f_i \cdot \binom{d-i}{j} \cdot (-1)^{d-j-i} \\ &= j! \cdot \sum_{i=0}^{d-j} f_i \cdot \binom{d-i}{(d-j)-i} \cdot (-1)^{d-j-i} = j! \cdot h_{d-j}.\end{aligned}$$

Now assume for the sake of contradiction that each of the values $h_d, \dots, h_{\beta+1}$ is zero. Consequently, $\mathtt{f}^{(j)}(-1) = 0$ for $j = 0, \dots, w-1$. Interpreting those evaluations of the derivatives of the $f$-polynomial as an instance of Hermite-Birkhoff interpolation, the corresponding matrix $E$ looks as follows:[4]

$$\begin{array}{cccccccc} 0 & 1 & 2 & \dots & w-1 & w & \dots & d \end{array}$$
$$\begin{pmatrix} 1 & 1 & 1 & \dots & 1 & 0 & \dots & 0 \\ \varepsilon_{20} & \varepsilon_{21} & \varepsilon_{22} & \dots & \varepsilon_{2(w-1)} & \varepsilon_{2w} & \dots & \varepsilon_{2d} \end{pmatrix}$$

In particular, at least $\beta + 1 = d + 1 - w$ of the values $\varepsilon_{2j}$ are 1; As $\beta + 1$ and $w$ sum up to $d + 1$, we can easily verify that the conditions of Pólya's Theorem (Theorem 13) are satisfied: Let us modify $E$ by arbitrarily choosing *precisely* $\beta + 1$ of the $\varepsilon_{2,j}$ that are 1 and set the others to 0, and call the resulting matrix $\hat{E}$. We then have both $M_j \geq j + 1$ (for all $j \in \{0, \dots, d-1\}$) and the first and second row of $\hat{E}$ sum up to *precisely* $d + 1$. Hence the matrix $\hat{E}$ is poised, that is, the only polynomial of degree at most $d$ that satisfies the corresponding instance of Hermite-Birkhoff interpolation is the zero polynomial. As we obtained $\hat{E}$ from $E$ just by ignoring some vanishing conditions, the same conclusion is true for $E$ and thus $\mathtt{f} = 0$ is the unique solution. This, however, contradicts the fact that the property $\Phi$ is not trivially false on $k$-vertex graphs, completing the proof. ∎

[4] Recall that an entry 1 in the matrix $E$ represents an evaluation $\mathtt{f}^{(j)}(-1) = 0$ in the first row and an evaluation $\mathtt{f}^{(j)}(0) = 0$ in the second row.

Combining Lemmas 15 and 16 yields our main technical result.

*Proof of Theorem 14:* Set $d = \binom{k}{2}$ and $\beta = d - w$. By Eq. (3) the function $a$ exists and is computable and has a finite support. Now, Lemma 16 implies that at least one of the values $h_d^{\Phi,k}, \dots, h_{\beta+1}^{\Phi,k}$ is non-zero and thus, by Lemma 15, at least one of the values $\tilde{h}_d, \dots, \tilde{h}_{\beta+1}$ is non-zero as well. Next, observe that $\tilde{h}_i = \sum_{K \in \mathcal{H}_i} a(K)$ for all $i \in \{0, \dots, d\}$, where $\mathcal{H}_i$ is the set of all graphs on $k$ vertices and $i$ edges. In particular, $\tilde{h}_i \neq 0$ implies that $a(K) \neq 0$ for at least one $K \in \mathcal{H}_i$, yielding the claim. ■

## IV. A Classification of $\#\text{IndSub}(\Phi)$ by the Hamming Weight of the $f$-Vectors

In this section, we derive a general hardness result for $\#\text{IndSub}(\Phi)$ based on the Hamming weight of the $f$-vector. In a sense, we "black-box" Theorem 14; using the resulting classification, we establish first hardness results and almost tight conditional lower bounds for a variety of families of graph properties.

However, note that taking a closer look at the number of edges of the graphs with non-vanishing coefficients (as provided by Theorem 14) often yields improved, sometimes even matching conditional lower bounds; we defer the treatment of the refined analysis to the full version.

In what follows, we write $\mathcal{K}(\Phi)$ for the set of all $k$ such that $\Phi_k$ is non-empty.

**Theorem 17.** *Let $\Phi$ denote a computable graph property and suppose that the set $\mathcal{K}(\Phi)$ is infinite. Let $\beta : \mathcal{K}(\Phi) \to \mathbb{Z}_{\geq 0}$ denote the function that maps $k$ to $\binom{k}{2} - \mathsf{hw}(f^{\Phi,k})$. If $\beta(k) \in \omega(k)$ then the problem $\#\text{IndSub}(\Phi)$ is $\#\mathsf{W}[1]$-complete and cannot be solved in time*

$$g(k) \cdot |V(G)|^{o((\beta(k)/k)/\log(\beta(k)/k))}$$

*for any function $g$, unless ETH fails. The same statement holds for $\#\text{IndSub}(\overline{\Phi})$ and $\#\text{IndSub}(\neg\Phi)$.*

Note that the condition of $\mathcal{K}(\Phi)$ being infinite is necessary for hardness: Otherwise there is a constant $c$ such that we can output 0 whenever $k \geq c$ and solve the problem by brute-force if $k < c$, yielding an algorithm with a polynomial running time. Note further that the $(\log(\beta(k)/k))^{-1}$-factor in the exponent is related to the question of whether it is possible to "beat treewidth" [42].[5] In particular, if the factor of $(\log \mathsf{tw}(H))^{-1}$ in Theorem 8 can be dropped, then all further results in this section can be strengthened to yield tight conditional lower bounds under ETH.

*Proof:* By Theorem 14, for each $k \in \mathcal{K}(\Phi)$ we obtain a graph $H_k$ with $k$ vertices and at least $\beta(k)$ edges such that $a(H_k) \neq 0$, where $a$ is the function in Eq. (3). The average degree of $H_k$ satisfies

$$d(H_k) = \frac{1}{k} \cdot \sum_{v \in V(H_k)} \mathsf{deg}(v) = \frac{2|E(H_k)|}{k} \geq \frac{2\beta(k)}{k},$$

where the second equality is due to the Handshaking Lemma. By Lemma 3, we thus obtain that $\mathsf{tw}(H_k) \geq \frac{\beta(k)}{k}$, which is unbounded as $\beta(k) \in \omega(k)$ by assumption.

Now let $\mathcal{H}$ denote the set of all graphs $H_k$ for $k \in \mathcal{K}(\Phi)$. By Theorem 8, we obtain that $\#\text{Hom}(\mathcal{H})$ is $\#\mathsf{W}[1]$-complete and cannot be solved in time

$$g(k) \cdot |V(G)|^{o((\beta(k)/k)/\log(\beta(k)/k))}$$

for any function $g$, unless ETH fails. Further, by Complexity Monotonicity (Theorem 9), the same is true for $\#\text{IndSub}(\Phi)$ as well. Finally, we use Fact 5 to obtain the same result for $\#\text{IndSub}(\overline{\Phi})$ and $\#\text{IndSub}(\neg\Phi)$; completing the proof. ■

We conclude this paper by demonstrating the applicability of Theorem 17: As a direct consequence, we obtain hardness for properties that are monotone, or only depend on the number of edges of a graph; the proofs of our main theorems require a slightly more detailed analysis and can be found in the full version.

### A. Monotone Graph Properties

Recall that a property $\Phi$ is called monotone if it is closed under taking subgraphs. The decision version of $\#\text{IndSub}(\Phi)$, that is, deciding whether there is an induced subgraph of size $k$ that satisfies $\Phi$ is known to be $\mathsf{W}[1]$-complete if $\Phi$ is monotone. Further, $\Phi$ is non-trivial and $\mathcal{K}(\Phi)$ is infinite (this follows implicitly by a result of Khot and Raman [38]). However, as the reduction of Khot and Raman is not parsimonious, the reduction does not yield $\#\mathsf{W}[1]$-completeness of the counting version. More importantly, the proof of Khot and Raman uses Ramsey's Theorem and thus only implies a conditional lower bound of $g(k) \cdot |V(G)|^{o(\log k)}$. Using our main result, we achieve a much stronger and almost tight lower bound under ETH.

**Theorem 18.** *Let $\Phi$ denote a computable graph property that is monotone and non-trivial. Suppose that $\mathcal{K}(\Phi)$ is infinite. Then $\#\text{IndSub}(\Phi)$ is $\#\mathsf{W}[1]$-complete and cannot be solved in time $g(k) \cdot |V(G)|^{o(k/\log k)}$ for any function $g$, unless ETH fails. The same is true for $\#\text{IndSub}(\overline{\Phi})$ and $\#\text{IndSub}(\neg\Phi)$.*

*Proof:* As $\Phi$ is non-trivial, there is a graph $F$ such that $\Phi(H)$ is false for every $H$ that contains $F$ as a (not necessarily induced) subgraph. Set $r = |V(F)|$ and fix $k \in \mathcal{K}(\Phi)$. By Turán's Theorem (Theorem 4) we have that every graph $H$ on $k$ vertices with more than $\left(1 - \frac{1}{r}\right) \cdot \frac{k^2}{2}$ edges contains the clique $K_{r+1}$ and thus $F$ as a subgraph. Consequently, $\Phi$ is false on every graph with $k$ vertices and

[5] See Theorem 8 and its discussion.

more than $\left(1-\frac{1}{r}\right)\cdot\frac{k^2}{2}$ edges. Therefore, we have

$$\beta(k)=\binom{k}{2}-\mathsf{hw}(f^{\Phi,k})\geq\binom{k}{2}-\left(1-\frac{1}{r}\right)\cdot\frac{k^2}{2}\in\Omega(k^2).$$

Thus $\beta(k)\in\Theta(k^2)$ and we conclude that

$$o\left(\frac{\beta(k)/k}{\log(\beta(k)/k)}\right)=o\left(k/\log k\right).$$

The claim hence follows by Theorem 17. ∎

*B. Graph Properties Depending Only on the Number of Edges*

Jerrum and Meeks [34], [35] asked whether $\#\textsc{IndSub}(\Phi)$ is $\#\mathsf{W}[1]$-complete whenever $\Phi$ is non-trivial infinitely often and only depends on the number of edges of a graph, that is,

$$\forall H_1,H_2:|E(H_1)|=|E(H_2)|\Rightarrow\Phi(H_1)=\Phi(H_2).$$

We answer this question affirmatively, even for properties that can depend both on the number of edges and vertices of the graph, and additionally provide an almost tight conditional lower bound:

**Theorem 19.** *Let $\Phi$ denote a computable graph property that only depends on the number of edges and the number of vertices of a graph. If $\Phi_k$ is non-trivial only for finitely many $k$ then $\#\textsc{IndSub}(\Phi)$ is fixed-parameter tractable. Otherwise, $\#\textsc{IndSub}(\Phi)$ is $\#\mathsf{W}[1]$-complete and cannot be solved in time*

$$g(k)\cdot|V(G)|^{o(k/\log k)}$$

*for any function $g$, unless ETH fails.*

Note that Theorem 19 is also true for $\#\textsc{IndSub}(\overline{\Phi})$ and $\#\textsc{IndSub}(\neg\Phi)$, as $\neg\Phi$ and $\overline{\Phi}$ depend only on the number of edges and vertices of a graph if and only if $\Phi$ does.

*Proof:* First, assume that $\Phi_k$ is non-trivial only for finitely many $k$. Then, there is a constant $c$ such that for every $k>c$, the property $\Phi_k$ is either trivially true or trivially false. Hence, given as input a graph $G$ and an integer $k$, we check whether $k\leq c$. If this is the case, we solve the problem by brute-force. Otherwise, we check whether $\Phi_k$ is trivially false or trivially true.[6] If $\Phi_k$ is false, we output 0; otherwise we output $\binom{n}{k}$. It is immediate that this algorithm yields fixed-parameter tractability.

Now assume that $\Phi_k$ is non-trivial for infinitely many $k$. Since for $\Phi_k$ we fix the number of vertices to be $k$, by assumption $\Phi_k$ only depends on the number of edges of a graph. Thus, we have

$$\mathsf{hw}(f^{\neg\Phi,k})=\binom{k}{2}-\mathsf{hw}(f^{\Phi,k}).\tag{4}$$

[6]This step is the reason why we only get fixed-parameter tractability and not necessarily polynomial-time tractability.

Hence, set

$$\hat{\Phi}_k:=\begin{cases}\Phi_k & \text{if } \mathsf{hw}(f^{\Phi,k})\leq\frac{1}{2}\binom{k}{2}\\ \neg\Phi_k & \text{if } \mathsf{hw}(f^{\Phi,k})>\frac{1}{2}\binom{k}{2}.\end{cases}$$

We observe that, by assumption, $\mathcal{K}(\hat{\Phi})$ is infinite, and by Fact 5 the problems $\#\textsc{IndSub}(\Phi)$ and $\#\textsc{IndSub}(\hat{\Phi})$ are equivalent. By definition and by Eq. (4), we see that $\mathsf{hw}(f^{\hat{\Phi},k})\leq\frac{1}{2}\binom{k}{2}$ and therefore $\beta(k)=\binom{k}{2}-\mathsf{hw}(f^{\hat{\Phi},k})\in\Theta(k^2)$. Thus, we have

$$o\left(\frac{\beta(k)/k}{\log(\beta(k)/k)}\right)=o\left(k/\log k\right).$$

The claim hence follows by Theorem 17. ∎


## Acknowledgments

We thank Dániel Marx and Jacob Focke for helpful comments.


## References


[1] N. Alon, P. Dao, I. Hajirasouliha, F. Hormozdiari, and S. C. Sahinalp, "Biomolecular network motif counting and discovery by color coding," *Bioinformatics*, vol. 24, no. 13, pp. i241–i249, 07 2008.

[2] M. Backens, "A Complete Dichotomy for Complex-Valued Holant$^c$," in *45th International Colloquium on Automata, Languages, and Programming, ICALP 2018, July 9-13, 2018, Prague, Czech Republic*, 2018, pp. 12:1–12:14.

[3] M. Backens, A. Bulatov, L. A. Goldberg, C. McQuillan, and S. Zivny, "Boolean approximate counting CSPs with weak conservativity, and implications for ferromagnetic two-spin," *J. Comput. Syst. Sci.*, vol. 109, pp. 95–125, 2020. [Online]. Available: https://doi.org/10.1016/j.jcss.2019.12.003

[4] L. Billera and A. Björner, "Face Numbers of Polytopes and Complexes," in *Handbook of discrete and computational geometry*. CRC Press, Inc., 1997, pp. 291–310.

[5] J. Cai, Z. Fu, H. Guo, and T. Williams, "A Holant Dichotomy: Is the FKT Algorithm Universal?" in *IEEE 56th Annual Symposium on Foundations of Computer Science, FOCS 2015, Berkeley, CA, USA, 17-20 October, 2015*, 2015, pp. 1259–1276.

[6] J. Cai, S. Huang, and P. Lu, "From Holant to #CSP and Back: Dichotomy for Holant$^c$ Problems," *Algorithmica*, vol. 64, no. 3, pp. 511–533, 2012.

[7] J. Cai and P. Lu, "Holographic algorithms: From art to science," *J. Comput. Syst. Sci.*, vol. 77, no. 1, pp. 41–61, 2011.

[8] J. Cai, P. Lu, and M. Xia, "Holographic Algorithms with Matchgates Capture Precisely Tractable Planar #CSP," *SIAM J. Comput.*, vol. 46, no. 3, pp. 853–889, 2017.

[9] ——, "Dichotomy for Real Holant$^c$ Problems," in *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018, New Orleans, LA, USA, January 7-10, 2018*, 2018, pp. 1802–1821.

[10] L. S. Chandran and C. R. Subramanian, "Girth and treewidth," *J. Comb. Theory, Ser. B*, vol. 93, no. 1, pp. 23–32, 2005.

[11] H. Chen and S. Mengel, "A Trichotomy in the Complexity of Counting Answers to Conjunctive Queries," in *18th International Conference on Database Theory, ICDT 2015, March 23-27, 2015, Brussels, Belgium*, 2015, pp. 110–126.

[12] ——, "Counting Answers to Existential Positive Queries: A Complexity Classification," in *Proceedings of the 35th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems, PODS 2016, San Francisco, CA, USA, June 26 - July 01, 2016*, 2016, pp. 315–326.

[13] J. Chen, B. Chor, M. Fellows, X. Huang, D. W. Juedes, I. A. Kanj, and G. Xia, "Tight lower bounds for certain parameterized NP-hard problems," *Inf. Comput.*, vol. 201, no. 2, pp. 216–231, 2005.

[14] J. Chen, X. Huang, I. A. Kanj, and G. Xia, "Strong computational lower bounds via parameterized complexity," *J. Comput. Syst. Sci.*, vol. 72, no. 8, pp. 1346–1367, 2006.

[15] Z. Chen, A. Galanis, L. A. Goldberg, W. Perkins, J. Stewart, and E. Vigoda, "Fast Algorithms at Low Temperatures via Markov Chains," in *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2019, September 20-22, 2019, Massachusetts Institute of Technology, Cambridge, MA, USA*, ser. LIPIcs, D. Achlioptas and L. A. Végh, Eds., vol. 145. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019, pp. 41:1–41:14.

[16] S. A. Cook, "The Complexity of Theorem-Proving Procedures," in *Proceedings of the 3rd Annual ACM Symposium on Theory of Computing, May 3-5, 1971, Shaker Heights, Ohio, USA*, 1971, pp. 151–158.

[17] R. Curticapean, "The simple, little and slow things count: On parameterized counting complexity," Ph.D. dissertation, Saarland University, 2015. [Online]. Available: http://scidok.sulb.uni-saarland.de/volltexte/2015/6217/

[18] R. Curticapean, H. Dell, and D. Marx, "Homomorphisms are a good basis for counting small subgraphs," in *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, 2017, pp. 210–223.

[19] R. Curticapean and D. Marx, "Complexity of Counting Subgraphs: Only the Boundedness of the Vertex-Cover Number Counts," in *55th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2014, Philadelphia, PA, USA, October 18-21, 2014*, 2014, pp. 130–139.

[20] M. Cygan, F. V. Fomin, L. Kowalik, D. Lokshtanov, D. Marx, M. Pilipczuk, M. Pilipczuk, and S. Saurabh, *Parameterized Algorithms*. Springer, 2015.

[21] V. Dalmau and P. Jonsson, "The complexity of counting homomorphisms seen from the other side," *Theor. Comput. Sci.*, vol. 329, no. 1-3, pp. 315–323, 2004.

[22] H. Dell, M. Roth, and P. Wellnitz, "Counting Answers to Existential Questions," in *46th International Colloquium on Automata, Languages, and Programming, ICALP 2019, July 9-12, 2019, Patras, Greece*, 2019, pp. 113:1–113:15.

[23] J. Dörfler, M. Roth, J. Schmitt, and P. Wellnitz, "Counting Induced Subgraphs: An Algebraic Approach to #W[1]-hardness," in *44th International Symposium on Mathematical Foundations of Computer Science, MFCS 2019, August 26-30, 2019, Aachen, Germany*, ser. LIPIcs, P. Rossmanith, P. Heggernes, and J. Katoen, Eds., vol. 138. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019, pp. 26:1–26:14.

[24] A. Durand and S. Mengel, "Structural Tractability of Counting of Solutions to Conjunctive Queries," *Theory Comput. Syst.*, vol. 57, no. 4, pp. 1202–1249, 2015.

[25] J. Edmonds, "Paths, Trees, and Flowers," *Canadian Journal of Mathematics*, vol. 17, pp. 449–467, 1965.

[26] P. Erdös and G. Szckeres, "A combinatorial problem in geometry," in *Classic Papers in Combinatorics*, I. Gessel and G.-C. Rota, Eds. Boston, MA: Birkhäuser Boston, 1987, pp. 49–56.

[27] J. Flum and M. Grohe, "The Parameterized Complexity of Counting Problems," *SIAM J. Comput.*, vol. 33, no. 4, pp. 892–922, 2004.

[28] L. A. Goldberg and M. Jerrum, "Approximating Pairwise Correlations in the Ising Model," *TOCT*, vol. 11, no. 4, pp. 23:1–23:20, 2019.

[29] J. A. Grochow and M. Kellis, "Network Motif Discovery Using Subgraph Enumeration and Symmetry-Breaking," in *Research in Computational Molecular Biology*, T. Speed and H. Huang, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 92–106.

[30] M. Grohe, T. Schwentick, and L. Segoufin, "When is the evaluation of conjunctive queries tractable?" in *Proceedings on 33rd Annual ACM Symposium on Theory of Computing, July 6-8, 2001, Heraklion, Crete, Greece*, 2001, pp. 657–666.

[31] S. Huang and P. Lu, "A Dichotomy for Real Weighted Holant Problems," *Computational Complexity*, vol. 25, no. 1, pp. 255–304, 2016.

[32] R. Impagliazzo and R. Paturi, "On the Complexity of k-SAT," *J. Comput. Syst. Sci.*, vol. 62, no. 2, pp. 367–375, 2001.

[33] M. Jerrum and K. Meeks, "The parameterised complexity of counting connected subgraphs and graph motifs," *J. Comput. Syst. Sci.*, vol. 81, no. 4, pp. 702–716, 2015.

[34] ——, "Some Hard Families of Parameterized Counting Problems," *TOCT*, vol. 7, no. 3, pp. 11:1–11:18, 2015.

[35] ——, "The parameterised complexity of counting even and odd induced subgraphs," *Combinatorica*, vol. 37, no. 5, pp. 965–990, 2017.

[36] P. W. Kasteleyn, "The statistics of dimers on a lattice: I. The number of dimer arrangements on a quadratic lattice," *Physica*, vol. 27, no. 12, pp. 1209–1225, 1961.

[37] ——, "Dimer Statistics and Phase Transitions," *Journal of Mathematical Physics*, vol. 4, no. 2, pp. 287–293, 1963.

[38] S. Khot and V. Raman, "Parameterized complexity of finding subgraphs with hereditary properties," *Theor. Comput. Sci.*, vol. 289, no. 2, pp. 997–1008, 2002.

[39] A. V. Kostochka, "Lower bound of the Hadwiger number of graphs by their average degree," *Combinatorica*, vol. 4, no. 4, pp. 307–316, 1984.

[40] D. Lokshtanov, D. Marx, and S. Saurabh, "Lower bounds based on the Exponential Time Hypothesis," *Bulletin of the EATCS*, vol. 105, pp. 41–72, 2011. [Online]. Available: http://eatcs.org/beatcs/index.php/beatcs/article/view/92

[41] L. Lovász, *Large Networks and Graph Limits*, ser. Colloquium Publications. American Mathematical Society, 2012, vol. 60. [Online]. Available: http://www.ams.org/bookstore-getitem/item=COLL-60

[42] D. Marx, "Can You Beat Treewidth?" *Theory of Computing*, vol. 6, no. 1, pp. 85–112, 2010.

[43] K. Meeks, "The challenges of unbounded treewidth in parameterised subgraph counting problems," *Discrete Applied Mathematics*, vol. 198, pp. 170–194, 2016.

[44] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, "Network Motifs: Simple Building Blocks of Complex Networks," *Science*, vol. 298, no. 5594, pp. 824–827, 2002.

[45] G. Pólya, "Bemerkung zur Interpolation und zur Näherungstheorie der Balkenbiegung," *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, vol. 11, no. 6, pp. 445–449, 1931.

[46] J. S. Provan and M. O. Ball, "The Complexity of Counting Cuts and of Computing the Probability that a Graph is Connected," *SIAM J. Comput.*, vol. 12, no. 4, pp. 777–788, 1983.

[47] F. P. Ramsey, "On a Problem of Formal Logic," *Proceedings of the London Mathematical Society*, vol. s2-30, no. 1, pp. 264–286, 1930.

[48] M. Roth and J. Schmitt, "Counting induced subgraphs: A Topological Approach to #W[1]-hardness," in *13th International Symposium on Parameterized and Exact Computation, IPEC 2018, August 20-24, 2018, Helsinki, Finland*, 2018, pp. 24:1–24:14.

[49] B. Schiller, S. Jager, K. Hamacher, and T. Strufe, "StreaM - A Stream-Based Algorithm for Counting Motifs in Dynamic Graphs," in *Algorithms for Computational Biology*, A.-H. Dediu, F. Hernández-Quiroz, C. Martín-Vide, and D. A. Rosenblueth, Eds. Cham: Springer International Publishing, 2015, pp. 53–67.

[50] I. J. Schoenberg, "On Hermite-Birkhoff Interpolation," *Journal of Mathematical Analysis and Applications*, vol. 16, no. 3, pp. 538 – 543, 1966.

[51] F. Schreiber and H. Schwöbbermeyer, "Frequency Concepts and Pattern Detection for the Analysis of Motifs in Networks," in *Transactions on Computational Systems Biology III*, C. Priami, E. Merelli, P. Gonzalez, and A. Omicini, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 89–104.

[52] J. Spencer, "Ramsey's theorem—a new lower bound," *Journal of Combinatorial Theory, Series A*, vol. 18, no. 1, pp. 108 – 115, 1975.

[53] H. N. V. Temperley and M. E. Fisher, "Dimer problem in statistical mechanics-an exact result," *The Philosophical Magazine: A Journal of Theoretical Experimental and Applied Physics*, vol. 6, no. 68, pp. 1061–1063, 1961.

[54] A. Thomason, "The Extremal Function for Complete Minors," *J. Comb. Theory, Ser. B*, vol. 81, no. 2, pp. 318–338, 2001.

[55] S. Toda, "PP is as Hard as the Polynomial-Time Hierarchy," *SIAM J. Comput.*, vol. 20, no. 5, pp. 865–877, 1991.

[56] J. R. Ullmann, "An algorithm for subgraph isomorphism," *J. ACM*, vol. 23, no. 1, pp. 31–42, 1976. [Online]. Available: http://doi.acm.org/10.1145/321921.321925

[57] L. G. Valiant, "The Complexity of Computing the Permanent," *Theor. Comput. Sci.*, vol. 8, pp. 189–201, 1979.

[58] ——, "The Complexity of Enumeration and Reliability Problems," *SIAM J. Comput.*, vol. 8, no. 3, pp. 410–421, 1979.

[59] ——, "Holographic Algorithms," *SIAM J. Comput.*, vol. 37, no. 5, pp. 1565–1594, 2008.