The Descriptive Complexity of Subgraph Isomorphism Without Numerics

Oleg Verbitsky¹ · Maksim Zhukovskii²

Abstract Let \( F \) be a connected graph with \( \ell \) vertices. The existence of a subgraph isomorphic to \( F \) can be defined in first-order logic with quantifier depth no better than \( \ell \), simply because no first-order formula of smaller quantifier depth can distinguish between the complete graphs \( K_\ell \) and \( K_{\ell-1} \). We show that, for some \( F \), the existence of an \( F \) subgraph in sufficiently large connected graphs is definable with quantifier depth \( \ell - 3 \). On the other hand, this is never possible with quantifier depth better than \( \ell / 2 \). If we, however, consider definitions over connected graphs with sufficiently large treewidth, the quantifier depth can for some \( F \) be arbitrarily small comparing to \( \ell \) but never smaller than the treewidth of \( F \). Moreover, the definitions over highly connected graphs require quantifier depth strictly more than the density of \( F \). Finally, we determine the exact values of these descriptive complexity parameters for all connected pattern graphs \( F \) on 4 vertices.

Keywords The subgraph isomorphism problem · Descriptive complexity · Encoding-independent computation · First-order logic · Quantifier depth

This article is part of the Topical Collection on Computer Science Symposium in Russia

O. Verbitsky is supported by DFG grants VE 652/1–2 and KO 1053/8–1. M. Zhukovskii is supported by grants No. 15-01-03530 and 16-31-60052 of Russian Foundation for Basic Research.

Oleg Verbitsky is on leave from the IAPMM, Lviv, Ukraine.

Oleg Verbitsky
verbitsk@informatik.hu-berlin.de

¹ Institut für Informatik, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany

² Laboratory of Advanced Combinatorics and Network Applications, Moscow Institute of Physics and Technology, Moscow, Russia
1 Introduction

For a fixed graph $F$ on $\ell$ vertices, let $S(F)$ denote the class of all graphs containing a subgraph isomorphic to $F$. The decision problem for $S(F)$ is known as SUBGRAPH ISOMORPHISM problem. It is solvable in time $O(n^\ell)$ on $n$-vertex input graphs by exhaustive search. Nešetřil and Poljak [23] showed that $S(F)$ can be recognized in time $O(n^{(\omega/3)^\ell+2})$, where $\omega<2.373$ is the exponent of fast square matrix multiplication [14]. Moreover, the color-coding method by Alon, Yuster and Zwick [2] yields the time bound

$$2^{O(\ell)} \cdot n^{tw(F) + 1} \log n,$$

where $tw(F)$ denotes the treewidth of $F$. On the other hand, the decision problem for $S(K_\ell)$, that is, the problem of deciding if an input graph contains a clique of $\ell$ vertices, cannot be solved in time $n^{o(\ell)}$ unless the Exponential Time Hypothesis fails [8].

We here are interested in the descriptive complexity of SUBGRAPH ISOMORPHISM. A sentence $\Phi$ defines a class of graphs $C$ if

$$G \models \Phi \iff G \in C,$$

where $G \models \Phi$ means that $\Phi$ is true on $G$. For a logic $\mathcal{L}$, we let $D_\mathcal{L}(C)$ (resp. $W_\mathcal{L}(C)$) denote the minimum quantifier depth (resp. variable width) of $\Phi \in \mathcal{L}$ defining $C$. Note that $W_\mathcal{L}(C) \leq D_\mathcal{L}(C)$. We simplify notation by writing

$$W_\mathcal{L}(F) = W_\mathcal{L}(S(F)) \text{ and } D_\mathcal{L}(F) = D_\mathcal{L}(S(F)).$$

We are primarily interested in the first-order logic of graphs with relation symbols for adjacency and equality of vertices, that will be denoted by FO. We suppose that the vertex set of any $n$-vertex graph is $\{1, \ldots, n\}$. Seeking the adequate logical formalism for various models of computation, descriptive complexity theory considers also more expressive logics involving numerical relations over the integers. Given a set $\mathcal{N}$ of such relations, FO[$\mathcal{N}$] is used to denote the extension of FO whose language contains symbols for each relation in $\mathcal{N}$. Of special interest are FO[$\prec$], FO[+ , ×], and FO[Arb], where Arb indicates that arbitrary relations are allowed. It is known [16, 22] that FO[Arb] and FO[+, ×] capture (non-uniform) AC$^0$ and DLOGTIME-uniform AC$^0$ respectively.

We will simplify the notation (2) further by writing $D(F) = D_{FO}(F)$ and $W(F) = W_{FO}(F)$. Dropping FO in the subscript, we also use notation like $D_{\prec}(F)$ or $W_{Arb}(F)$. In this way we obtain two hierarchies of width and depth parameters. In particular,

$$W_{Arb}(F) \leq W_{\prec}(F) \leq W(F) \text{ and } D_{Arb}(F) \leq D_{\prec}(F) \leq D(F).$$

The relation of FO[Arb] to circuit complexity implies that $S(F)$ is recognizable on $n$-vertex graphs by bounded-depth unbounded-fan-in circuits of size $O(n^{W_{Arb}(F)})$; see [16, 27]. The interplay between the two areas has been studied in [19–21, 27, 28]. Noteworthy, the parameters $W_{Arb}(F)$ and $D_{Arb}(F)$ admit combinatorial upper bounds

$$W_{Arb}(F) \leq tw(F) + 3 \text{ and } D_{Arb}(F) \leq td(F) + 2$$

\(\square\) Springer
in terms of the treewidth and treedepth of $F$; see [29].

The focus of our paper is on FO without any background arithmetical relations. Our interest in this, weakest setting is motivated by the prominent problem on the power of encoding-independent computations; see, e.g., [15]. It is a long-standing open question in finite model theory as to whether there exists a logic capturing polynomial time on finite relational structures. The existence of a natural logic capturing polynomial time would mean that any polynomial-time computation could be made, in a sense, independent of the input encoding. If this is true, are the encoding-independent computations necessarily slower than the standard ones? This question admits the following natural variation. Suppose that a decision problem a priori admits an encoding-independent polynomial-time algorithm, say, being definable in FO, like \textsc{Subgraph Isomorphism} for a fixed pattern graph $F$. Is it always true that the running time of this algorithm can be improved in the standard encoding-dependent Turing model of computation?

A straightforward conversion of an FO sentence defining $S(F)$ into an algorithm recognizing $S(F)$ results in the time bound $O(n^{D(F)})$ for \textsc{Subgraph Isomorphism}, which can actually be improved to $O(n^{W(F)})$; see [22, Prop. 6.6]. The same applies to FO[$\prec$]. The last logic is especially interesting in the context of order-invariant definitions. It is well known [22, 30] that there are properties of (unordered) finite structures that can be defined in FO[$\prec$] but not in FO. Even if a property, like $S(F)$, is definable in FO, one can expect that in FO[$\prec$] it can be defined much more succinctly. As a simple example, take $F$ to be the star graph $K_{1,s}$ and observe that $D_{\prec}(K_{1,s}) \leq \log_2 s + 3$ and $W_<(K_{1,s}) \leq 3$ while $W(K_{1,s}) = s + 1$.

The main goal we pose in this paper is examining abilities and limitations of the “pure” FO in succinctly defining \textsc{Subgraph Isomorphism}. Actually, if a pattern graph $F$ has $\ell$ vertices, then the trivial upper bound $D(F) \leq \ell$ cannot be improved. We have $W(F) = \ell$ simply because no first-order formula with less than $\ell$ variables can distinguish between the complete graphs $K_\ell$ and $K_{\ell-1}$. Is this, however, the only reason preventing more succinct definitions of $S(F)$? How succinctly can $S(F)$ be defined on large enough graphs? The question can be formalized as follows. We say that a sentence $\Phi$ defines $S(F)$ on sufficiently large connected graphs if there is $k$ such that the equivalence (1) with $C = S(F)$ is true for all connected $G$ with at least $k$ vertices. Let $W_v(F)$ (resp. $D_v(F)$) denote the minimum variable width (resp. quantifier depth) of such $\Phi$.

Throughout the paper, we assume that the fixed pattern graph $F$ is connected. Therefore, $F$ is contained in a host graph $G$ if and only if it is contained in a connected component of $G$. By this reason, the decision problem for $S(F)$ efficiently reduces to its restriction to connected input graphs. Since it suffices to solve the problem only on all sufficiently large inputs, $S(F)$ is still recognizable in time $O(n^{W_v(F)})$, while $W_v(F) \leq W(F)$.

\footnote{In his presentation [29], Benjamin Rossman states upper bounds $W_{FO}(F) \leq tw(F) + 1$ and $D_{FO}(F) \leq td(F)$ for the colorful version of \textsc{Subgraph Isomorphism} studied in [21]. It is not hard to observe that the auxiliary color predicates can be defined in FO[Arb] at the cost of two extra quantified variables by the color-coding method developed in [2]; see also [3, Thm. 4.2].}
A further relaxation is motivated by Courcelle’s theorem [9] saying that every graph property definable by a sentence in monadic second-order logic can be efficiently decided on graphs of bounded treewidth. More precisely, for SUBGRAPH ISOMORPHISM Courcelle’s theorem implies that \( S(F) \) is decidable in time \( f(\ell, \text{tw}(G)) \cdot n \), which means linear time for any class of input graphs having bounded treewidth.

Now, we say that a sentence \( \Phi \) defines \( S(F) \) on connected graphs with sufficiently large treewidth if there is \( k \) such that the equivalence (1) with \( C = S(F) \) is true for all connected \( G \) with treewidth at least \( k \). Denote the minimum variable width (resp. quantifier depth) of such \( \Phi \) by \( W_{tw}(F) \) (resp. \( D_{tw}(F) \)). Fix \( k \) that ensures the minimum value \( W_{tw}(F) \) and recall that, by Courcelle’s theorem, the subgraph isomorphism problem is solvable on graphs with treewidth less than \( k \) in linear time. Note that, for a fixed \( k \), whether or not \( \text{tw}(G) < k \) is also decidable in linear time [4]. It follows that \( S(F) \) is recognizable even in time \( O(n W_{tw}(F)) \), while \( W_{tw}(F) \leq W_v(F) \). Moreover, the last inequality can be strict, for example, as a consequence of the fact that some subgraphs are just forced by large treewidth; see Section 4 for details.

The above discussion shows that the parameters \( W_v(F) \), \( D_v(F) \), \( W_{tw}(F) \), and \( D_{tw}(F) \) have clear algorithmic meaning. Analyzing this setting, we obtain the following results.

- We demonstrate that non-trivial definitions over sufficiently large graphs are possible by showing that \( D_v(F) \leq v(F) - 3 \) for some \( F \), where \( v(F) \) denotes the number of vertices in \( F \). On the other hand, we show limitations of this approach by proving that \( W_v(F) \geq (v(F) - 1)/2 \) for all \( F \).
- The last barrier (as well as any lower bound in terms of \( v(F) \)) can be overcome by definitions over graphs with sufficiently large treewidth. Specifically, for every \( \ell \) and \( a \leq \ell \) there is an \( \ell \)-vertex \( F \) such that \( D_{tw}(F) \leq a \) and, moreover, \( \text{tw}(F) = a - 1 \). On the other hand, \( W_{tw}(F) \geq \text{tw}(F) \) for all \( F \). Note that, along with (3), this implies that \( W_{Arb}(F) \leq W_{tw}(F) + 3 \).

Furthermore, we also consider definitions of \( S(F) \) over graphs of sufficiently large connectedness. Denote the corresponding quantifier depth parameter by \( D_k(F) \) and note that \( D_k(F) \leq D_{tw}(F) \leq D_v(F) \) (see Section 2 for details), which motivates our interest in lower bound for \( D_k(F) \). Let \( e(F) \) denote the number of edges in \( F \). For every pattern graph \( F \) with \( e(F) > v(F) \), we prove that \( D_k(F) \geq \frac{e(F)}{v(F)} + 2 \).

Finally, we determine the exact values of the parameters \( D_k(F) \), \( D_{tw}(F) \), and \( D_v(F) \) for all connected pattern graphs \( F \) on 4 vertices.

Related work In an accompanying paper [33], we address the descriptive complexity of the INDUCED SUBGRAPH ISOMORPHISM problem. Let \( I(F) \) denote the class of all graphs containing an induced subgraph isomorphic to \( F \). The state-of-the-art of the algorithmics for INDUCED SUBGRAPH ISOMORPHISM is different from SUBGRAPH ISOMORPHISM. Floderus et al. [13] collected evidence in favor of the conjecture that \( I(F) \) for \( F \) with \( \ell \) vertices cannot be recognized faster than \( I(K_{c\ell}) \), where \( c < 1 \) is a constant. Similarly to \( D(F) \), we use the notation \( D[F] = D(I(F)) \) and \( W[F] = W(I(F)) \), where the square brackets indicate that the case of induced subgraphs is considered. The trivial argument showing that \( W(F) = v(F) \) does not
work anymore unless $F$ is a complete graph. Proving or disproving that $D[F] = W[F] = v(F)$ seems to be a subtle problem. An example of a pattern graph $F$ for which $D[F] < v(F)$ is given by considering the paw graph, as a consequence of Olariu’s characterization of the class of paw-free graphs [24]. In [33], we prove a general lower bound $W[F] \geq (1/2 - o(1))v(F)$.

**Organization of the paper** We introduce our setting formally in Section 2, which also contains necessary logical and graph-theoretic preliminaries. The first-order definitions of $S(F)$ over sufficiently large connected graphs (the parameters $D_v(F)$ and $W_v(F)$) are addressed in Section 3. Sections 4 and 5 are devoted to the definitions over graphs of sufficiently large treewidth ($D_{tw}(F)$ and $W_{tw}(F)$) and of sufficiently large connectedness ($D_{\kappa}(F)$) respectively. The exact values of $D_v(F)$, $D_{tw}(F)$, and $D_{\kappa}(F)$ are determined for all connected $F$ on 4 vertices in Section 6. The width parameters are also determined with the exception of $W_{\kappa}(F)$ for $F$ being the diamond graph and the 4-cycle. We conclude with discussing further questions in Section 7.

A preliminary version of this paper appeared in [32].

## 2 Preliminaries

### 2.1 First-order Complexity of Graph Properties

We consider first-order sentences about graphs in the language containing the adjacency and the equality relations. Let $C$ be a first-order definable class of graphs and $\pi$ be a graph parameter. Let $D^{k}_\pi(C)$ denote the minimum quantifier depth of a first-order sentence $\Phi$ such that, for every connected graph $G$ with $\pi(G) \geq k$, $\Phi$ is true on $G$ exactly when $G$ belongs to $C$. Note that $D^{k}_\pi(C) \geq D^{k+1}_\pi(C)$, and define $D_\pi(C) = \min_k D^{k}_\pi(C)$. In other words, $D_\pi(C)$ is the minimum quantifier depth of a first-order sentence defining $C$ over connected graphs with sufficiently large values of $\pi$.

The **variable width** of a first-order sentence $\Phi$ is the number of first-order variables used to build $\Phi$; different occurrences of the same variable do not count. Similarly to the above, by $W_\pi(C)$ we denote the minimum variable width of $\Phi$ defining $C$ over connected graphs with sufficiently large $\pi$. Note that

$$W_\pi(C) \leq D_\pi(C).$$

Recall that a graph is $k$-connected if it has more than $k$ vertices, is connected, and remains connected after removal of any $k - 1$ vertices. The connectivity $\kappa(G)$ of $G$ is equal to the maximum $k$ such that $G$ is $k$-connected. We will consider the depth $D_\pi(C)$ and the width $W_\pi(C)$ for three parameters $\pi$ of a graph $G$, namely the number of vertices $v(G)$, the treewidth $tw(G)$, and the connectivity $\kappa(G)$. Note that $tw(G) < v(G)$. Note also that any graph $G$ with $v(G) > k$ and $tw(G) < k$ can be disconnected by removing fewer than $k$ vertices. Therefore, every $k$-connected graph has treewidth at least $k$. It follows that

$$D^{k+1}_v(C) \geq D^k_{tw}(C) \geq D^k_{\kappa}(C)$$
and, hence,\[ D_v(C) \geq D_{tw}(C) \geq D_k(C). \]

Similarly,\[ W_v(C) \geq W_{tw}(C) \geq W_k(C). \]

As it was discussed in Section 1, the values of \( D_v(C) \) and \( D_{tw}(C) \), as well as \( W_v(C) \) and \( W_{tw}(C) \), are related to the time complexity of the decision problem for \( C \). Consideration of \( D_k(C) \) and \( W_k(C) \) is motivated by the fact that some lower bounds we are able to show for \( D_v(C) \) and \( D_{tw}(C) \) actually hold for \( D_k(C) \) or even for \( W_k(C) \), and it is natural to present them in this stronger form.

Recall that \( S(F) \) denotes the class of graphs containing a subgraph isomorphic to \( F \). Simplifying the notation, we write \( D_v(F) = D_v(S(F)) \), \( W_v(F) = W_v(S(F)) \), etc.

Given two non-isomorphic graphs \( G \) and \( H \), let \( D(G, H) \) (resp. \( W(G, H) \)) denote the minimum quantifier depth (resp. variable width) of a sentence that is true on one of the graphs and false on the other. It may be useful to note that \( D(G, H) > d \) if and only if \( G \equiv_d H \), where the \( \equiv_d \)-equivalence relation means that \( G \) and \( H \) satisfy the same first-order sentences of quantifier depth \( d \). Similarly, \( W(G, H) > d \) if and only if \( G \equiv_d^d H \), that is, \( G \) and \( H \) are indistinguishable in \( d \)-variable logic; see [12].

**Lemma 2.1**

1. \( D_\pi(C) \geq d \) if there are connected graphs \( G \in C \) and \( H \notin C \) with arbitrarily large values of \( \pi(G) \) and \( \pi(H) \) such that \( D(G, H) \geq d \).
2. \( W_\pi(C) \geq d \) if there are connected graphs \( G \in C \) and \( H \notin C \) with arbitrarily large values of \( \pi(G) \) and \( \pi(H) \) such that \( W(G, H) \geq d \).
3. \( D_\pi(C) \leq d \) if \( D(G, H) \leq d \) for all connected graphs \( G \in C \) and \( H \notin C \) with sufficiently large values of \( \pi(G) \) and \( \pi(H) \).

**Proof** Parts 1 and 2 follow directly from the definitions as any sentence defining \( C \) on connected graphs with sufficiently large \( \pi \) distinguishes between any two graphs \( G \in C \) and \( H \notin C \) with sufficiently large \( \pi \). Let us prove Part 3. By assumption, any two connected graphs \( G \in C \) and \( H \notin C \) with sufficiently large \( \pi \) (say, with \( \pi(G) \geq k \) and \( \pi(H) \geq k \)) are distinguished by a sentence \( \Phi_{G,H} \) of quantifier depth at most \( d \) (that is true on \( G \) and false on \( H \)). For a connected graph \( G \in C \) with \( \pi(G) \geq k \), consider the sentence \( \Phi_G \overset{\text{def}}{=} \bigwedge_H \Phi_{G,H} \), where the conjunction is over all connected \( H \notin C \) with \( \pi(H) \geq k \). This sentence distinguishes \( G \) from all \( H \notin C \) with \( \pi(H) \geq k \) and has quantifier depth at most \( d \). The only problem with it is that the conjunction over \( H \) is actually infinite. Luckily, there are only finitely many pairwise inequivalent first-order sentences about graphs of quantifier depth \( d \); see, e.g., [25, Theorem 2.4]. Removing all but one formula \( \Phi_{G,H} \) from each equivalence class, we make \( \Phi_G \) a legitimate finite sentence. Now, consider \( \Phi \overset{\text{def}}{=} \bigvee_G \Phi_G \), where the disjunction is over all connected \( G \in C \) with \( \pi(G) \geq k \). It can be made finite in the same way. The sentence \( \Phi \) defines \( C \) over connected graphs with \( \pi(G) \geq k \) and has quantifier depth \( d \). Therefore, \( D_\pi(C) \leq D^*_k(C) \leq d \).
Lemma 2.1 reduces estimating $D_\pi(C)$ to estimating $D(G, H)$ over connected $G \in C$ and $H \not\in C$ with large values of $\pi$. Also, proving lower bounds for $W_\pi(C)$ reduces to proving lower bounds for $W(G, H)$. For estimating $D(G, H)$ and $W(G, H)$ there is a remarkable tool.

In the $k$-pebble Ehrenfeucht-Fraïssé game, the board consists of two vertex-disjoint graphs $G$ and $H$. Two players, Spoiler and Duplicator (or he and she) have equal sets of $k$ pairwise different pebbles. In each round, Spoiler takes a pebble and puts it on a vertex in $G$ or in $H$; then Duplicator has to put her copy of this pebble on a vertex of the other graph. Duplicator’s objective is to ensure that the pebbling determines a partial isomorphism between $G$ and $H$ after each round; when she fails, she immediately loses. The proof of the following facts can be found in [16]:

1. $D(G, H)$ is equal to the minimum $k$ such that Spoiler has a winning strategy in the $k$-round $k$-pebble game on $G$ and $H$.
2. $W(G, H)$ is equal to the minimum $k$ such that, for some $d$, Spoiler has a winning strategy in the $d$-round $k$-pebble game on $G$ and $H$.

2.2 Graph-theoretic Preliminaries

Recall that $v(G)$ denotes the number of vertices in a graph $G$. The treewidth of $G$ is denoted by $tw(G)$. The neighborhood $N(v)$ of a vertex $v$ consists of all vertices adjacent to $v$. The number $\deg v = |N(v)|$ is called the degree of $v$. A vertex of degree 1 is called pendant.

We use the standard notation $K_n$ for complete graphs, $P_n$ for paths, and $C_n$ for cycles on $n$ vertices. Furthermore, $K_{a,b}$ denotes the complete bipartite graph whose vertex classes have $a$ and $b$ vertices. In particular, $K_{1,n-1}$ is the star graph on $n$ vertices. The subscript in the name of a graph will almost always denote the number of vertices. If a graph is indexed by two parameters, their sum is typically equal to the total number of vertices in the graph.

The following definitions are illustrated in Fig. 1. Let $a \geq 3$ and $b \geq 1$. The lollipop graph $L_{a,b}$ is obtained from $K_a$ and $P_b$ by adding an edge between an end vertex of $P_b$ and a vertex of $K_a$. We also make a natural convention that $L_{a,0} = K_a$. Furthermore, the sparkler graph $S_{a,b}$ is obtained from $K_{1,a-1}$ and $P_b$ by adding an edge between an end vertex of $P_b$ and the central vertex of $K_{1,a-1}$. The jellyfish

![Special graph families: Lollipops, sparklers, jellyfishes, and megastars](image)

Fig. 1 Special graph families: Lollipops, sparklers, jellyfishes, and megastars
graph $J_{a,b}$ is the result of attaching $b$ pendant vertices to a vertex of $K_a$. Finally, the megastar graph $M_{s,t}$ is obtained from the star $K_{1,s}$ by subdividing each edge into $P_{t+1}$; thus $v(M_{s,t}) = st + 1$.

3 Definitions Over Sufficiently Large Graphs

Our first goal is to demonstrate that non-trivial definitions over large connected graphs are really possible. The lollipop graphs $L_{a,1}$ give simple examples of pattern graphs $F$ with $D_v(F) \leq v(F) - 1$. Though not so easily, the same can be shown for the path graphs $P_\ell$. We are able to show better upper bounds using sparkler graphs.

Theorem 3.1 There is a graph $F$ with $D_v(F) \leq v(F) - 3$. Specifically, $D_v(S_{4,4}) = 5$.

For the proof we need two technical lemmas.

Lemma 3.2 Suppose that a connected graph $H$ contains the 4-star $K_{1,4}$ as a subgraph but does not contain any subgraph $S_{4,4}$. Then $H$ contains a vertex of degree more than $(v(H)/2)^{1/7}$.

Proof $H$ cannot contain $P_{15}$ because, together with $K_{1,4}$, it would give an $S_{4,4}$ subgraph. Indeed, assume that $X$ is a copy of $P_{15}$ in $H$. Let $Z$ be a copy of $K_{1,4}$ in $H$ consisting of the central vertex $v$ and its neighbors $v_1, v_2, v_3, v_4$. Suppose that $v$ does not belong to $X$. If $X$ does not contain also any of $v_1, v_2, v_3, v_4$, then the star $Z$ can be extended to an $S_{4,4}$ along a shortest path from $v$ to $X$ and, if needed, further along $X$. If some of the vertices $v_1, v_2, v_3, v_4$ lie on $X$, they split $X$ into at most 5 parts. One of these parts must contain at least 3 vertices, which can be used for extending $Z$ to a copy of $S_{4,4}$ in $H$. If $v$ lies strictly inside of $X$, we can suppose without loss of generality that $v_1, v, v_2$ are successive vertices in the path $X$. Then the segment $v_1, v, v_2$ and, possibly, the vertices $v_3$ and $v_4$ split $X$ into at most four parts. Again, one of them must contain at least 3 vertices, which can be used to extend the star $Z$ to a sparkler $S_{4,4}$. If $v$ is an end vertex of $X$, we can suppose that $v_1$ is the next vertex of $X$. The initial segment $v, v_1$ and, possibly, the vertices $v_2, v_3, v_4$ split $X$ into at most four parts, and the same argument applies.

Consider an arbitrary spanning tree $T$ in $H$ and denote its maximum vertex degree by $d$ and its radius by $r$. Note that $v(T) \leq 1 + d + d(d - 1) + \ldots + d(d - 1)^{r-1}$. Since $T$ contains no $P_{15}$, we have $r \leq 7$. It follows that $v(H) = v(T) < 2d^3$.

Let $\sim$ denote the adjacency relation and recall that $N(v)$ denotes the neighborhood of a vertex $v$.

Lemma 3.3 Let $y_0 \in V(H)$ and assume that

- $H$ is a sufficiently large connected graph,
- $H$ does not contain $S_{4,4}$,
- $\deg y_0 \geq 4$,
- $y_0 y_1 y_2 y_3 y_4$ is a path in $H$. 
Then (see Fig. 2)

1. $\text{deg } y_0 = 4$,
2. $y_0 \sim y_2$, $y_0 \sim y_3$, $y_0 \sim y_4$,
3. if $N(y_0) = \{y_1, y_2, y', y''\}$, then $y_1 \sim y'$ and $y_1 \sim y''$.

Proof By Lemma 3.2 we know that $H$ must contain a vertex $z$ of large degree, namely $\text{deg } z \geq 7$. We have $y_0 \nsim y_4$ for else $H$ would contain a cycle $C_5$ and, together with $z$, this would give us a subgraph $S_{4,4}$ (because, by connectedness of $H$, we would have a path $P_5$ emanating from $z$). Therefore, $y_0$ has a neighbor $y' \notin \{y_1, y_2, y_3, y_4\}$. Furthermore, $y_0 \sim y_3$ for else, considering a path from $z$ to one of the vertices $y'$, $y_0$, $y_1$, $y_2$, $y_3$, $y_4$, we get a $P_5$ emanating from $z$ and, hence, an $S_{4,4}$. Therefore, $y_0$ has another neighbor $y'' \notin \{y', y_1, y_2, y_3, y_4\}$. Furthermore, $y_0 \sim y_2$ for else $y_0$ would have three neighbors $y'$, $y''$, $y'''$ different from $y_1$, $y_2$, $y_3$, $y_4$, which would give $S_{4,4}$. By the same reason, $y_0$ has no other neighbors, that is, $N(y_0) = \{y_1, y_2, y', y''\}$ and $\text{deg } y_0 = 4$. Note that $z \in \{y_0, y_1, y_2, y_3, y_4\}$ for else we easily get an $S_{4,4}$ by considering a path from $z$ to one of these vertices. It is also easy to see that $z \neq y_0, y_4, y_3, y_1$ (for example, if $\text{deg } y_1 \geq 7$, then it would give an $S_{4,4}$ with tail $y_1y_0y_2y_3y_4$). Therefore, $z = y_2$. If $y_1 \sim y'$ or $y_1 \sim y''$, we would have an $S_{4,4}$ with tails $y_2y_1y'y_0y''$ or $y_2y_1y'y_0y'$ respectively.

Proof of Theorem 3.1 We are now ready to prove the upper bound $D_v(S_{4,4}) \leq 5$. Consider sufficiently large connected graphs $G$ and $H$ and suppose that $G$ contains an $S_{4,4}$, whose vertices are labeled as in Fig. 2, and $H$ contains no copy of $S_{4,4}$. We describe a winning strategy for Spoiler in the game on $G$ and $H$. 

Fig. 2 Proof of Theorem 3.1
1st round.Spoiler pebbles $x_0$. Denote the response of Duplicator in $H$ by $y_0$. Assume that $\deg y_0 \geq 4$ for else Spoiler wins in the next 4 moves. Assume that $x_0 \sim x_2$ for else Spoiler wins by pebbling $x_1, x_2, x_3, x_4$ (if Duplicator responds with a path $y_0y_1y_2y_3y_4$, she loses by Condition 2 in Lemma 3.3).

2nd round. Spoiler pebbles $x_1$. Denote the response of Duplicator in $H$ by $y_1$. Assume that there is a path $y_0y_1y_2y_3y_4$ for else Spoiler wins in the next 3 moves.

Case 1: $x_1$ is adjacent to any of the vertices $x', x'', x'''$, say, to $x'$. Spoiler pebbles $x_2$ and $x'$ and wins. Indeed, Duplicator has to respond with two vertices in $H$ both in $N(y_0) \cap N(y_1)$, which is impossible by Conditions 1 and 3 of Lemma 3.3.

Case 2: $x_1 \sim x', x_1 \sim x'', x_1 \sim x'''$. Spoiler wins by pebbling $x', x'', x'''$. Duplicator has to respond with three vertices in $N(y_0) \setminus N(y_1)$, which is impossible by Conditions 1 and 2 of Lemma 3.3.

This completes the proof of the upper bound. On the other hand, we have $D_v(S_{4,4}) > 4$ by considering the jellyfish graphs $G = J_{5,n}$ and $H = J_{4,n}$. □

With more technical effort, we can show that $D_v(F) \leq v(F) - 3$ for infinitely many $F$, namely for all $F = S_{a,a}$ (the proof of this fact is rather involved and will appear elsewhere).

We now show general lower bounds for $D_v(F)$ and $W_v(F)$. For this, we need some definitions. Let $v_0v_1 \ldots v_t$ be an induced path in a graph $G$. We call it pendant if $\deg v_0 \neq 2$, $\deg v_i = 1$ and $\deg v_i = 2$ for all $1 \leq i < t$. Furthermore, let $S$ be an induced star $K_{1,s}$ in $G$ with the central vertex $v_0$. We call $S$ pendant if all its pendant vertices are pendant also in $G$, and in $G$ there are no more than $s$ pendant vertices adjacent to $v_0$. The definition ensures that a pendant path (or star) cannot be contained in a larger pendant path (or star). As an example, note that the sparkler graph $S_{s+1,t}$ has a pendant $P_{t+1}$ and a pendant $K_{1,s}$.

Let $p(F)$ denote the maximum $t$ such that $F$ has a pendant path $P_{t+1}$. Similarly, let $s(F)$ denote the maximum $s$ such that $F$ has a pendant star $K_{1,s}$. If $F$ has no pendant vertex, then we set $p(F) = 0$ and $s(F) = 0$.

**Theorem 3.4** $D_v(F) \geq (v(F) + 1)/2$ and $W_v(F) \geq (v(F) - 1)/2$ for every connected $F$ unless $F = P_2$ or $F = P_3$.

**Proof** Denote

$$\ell = v(F), \ t = p(F) \text{ and } s = s(F).$$

We begin with noticing that

$$D_v(F) \geq \ell - t \text{ and } W_v(F) \geq \ell - t - 1. \quad (4)$$

Indeed, this is obvious if $F$ is a path, that is, $F = P_{t+1}$. If $F$ is not a path, we consider lollipop graphs $G = L_{\ell-t,n}$ and $H = L_{\ell-t-1,n}$ for each $n \geq t$ (note that $\ell \geq t + 3$ and, if $\ell = t + 3$, then $H = L_{2,n} = P_{n+2}$). Obviously, $G$ contains $F$, and $H$ does not. It remains to note that $D(G, H) \geq \ell - t$ and $W(G, H) \geq \ell - t - 1$. □
We also claim that
\[ D_v(F) \geq \ell - s \quad \text{and} \quad W_v(F) \geq \ell - s - 1. \] (5)
This is obvious if \( F \) is a star, that is, \( F = K_{1,s} \). If \( F \) is not a star, we consider jellyfish graphs \( G = J_{\ell-s,n} \) and \( H = J_{\ell-s-1,n} \) for each \( n \geq s \) (note that \( \ell \geq s + 3 \) and, if \( \ell = s + 3 \), then \( H = J_{2,n} = K_{1,n+1} \)). Clearly, \( G \) contains \( F \), and \( H \) does not. It remains to observe that \( D(G,H) \geq \ell - s \) and \( W(G,H) \geq \ell - s - 1 \).

Let \( F = K_{1,\ell-1} \) or \( F = P_\ell \), where \( \ell \geq 4 \). Using (4) and (5) respectively, we get \( D_v(F) \geq \ell - 1 \geq \frac{\ell+1}{2} \) and, similarly, \( W_v(F) \geq \ell - 2 \geq \frac{\ell-1}{2} \). Assume, therefore, that \( F \) is neither a star nor a path. In this case we claim that
\[ t + s < \ell. \] (6)
This is obviously true if \( F \) has no pendant vertex, that is, \( t = s = 0 \). Suppose that \( F \) has a pendant vertex and, therefore, both \( t > 0 \) and \( s > 0 \). Consider an arbitrary spanning tree \( T \) of \( F \) and note that \( T \) contains all pendant paths and stars of \( F \). Fix a longest pendant path \( P \) and a largest pendant star \( S \) in \( F \). If \( P \) and \( S \) share at most one common vertex, we readily get (6). If they share two vertices, then \( S = K_{1,1} \), i.e., \( s = 1 \), and \( t + 1 < \ell \) follows from the assumption that \( F \) is not a path.

The theorem readily follows from (4), (5), and (6).

4 Definitions Over Graphs of Sufficiently Large Treewidth

Theorem 3.4 poses limitations on the succinctness of definitions over sufficiently large connected graphs. We now show that there are no such limitations for definitions over connected graphs with sufficiently large treewidth.

The Grid Minor Theorem says that every graph of large treewidth contains a large grid minor; see [11]. The strongest version of this result belongs to Chekuri and Chuzhoy [7] who proved that, for some \( \epsilon > 0 \), every graph \( G \) of treewidth \( k \) contains the \( m \times m \) grid as a minor with \( m = \Omega(k^\epsilon) \). If \( m > 2b \), then \( G \) must contain \( M_{3,b} \) as a subgraph. This applies also to all subgraphs of \( M_{3,b} \). The following result is based on the fact that a graph of large treewidth contains a long path.

**Theorem 4.1** For all \( a \) and \( \ell \) such that \( 3 \leq a \leq \ell \) there is a graph \( F \) with \( v(F) = \ell \) and \( tw(F) = a - 1 \) such that \( D_{tw}(F) \leq a \). Specifically, \( D_{tw}(La,b) = W_{\kappa}(La,b) = a \) if \( a \geq 3 \) and \( b \geq 0 \).

Note for comparison that \( W_v(La,b) \geq a + b - 2 \), as follows from the bound (5) in the proof of Theorem 3.4.

**Proof** We first prove the upper bound \( D_{tw}(La,b) \leq a \). If a connected graph \( H \) of large treewidth does not contain \( L_{a,b} \), it cannot contain even \( K_a \), for else \( K_a \) could be combined with a long path to give \( L_{a,b} \). Therefore, Spoiler wins on \( G \in S(La,b) \) and such \( H \) in \( a \) moves.

For the lower bound \( W_{\kappa}(La,b) \geq a \), consider \( G = K(a,n) \) and \( H = K(a-1,n) \), where \( K(a,n) \) denotes the complete \( a \)-partite graph with each part having \( n \) vertices.
Note that the graph $K(a, n)$ is $(a - 1)n$-connected. If $n > b$, then $G$ contains $L_{a, b}$, while $H$ for any $n$ does not contain even $K_a$. It remains to note that $W(G, H) \geq a$ if $n \geq a - 1$.

We now prove a general lower bound for $W_{tw}(F)$ in terms of the treewidth $tw(F)$.

Using the terminology of [17, Chap. 5], we define the core $F_0$ of $F$ to be the graph obtained from $F$ by removing, consecutively and as long as possible, vertices of degree at most 1. If $F$ is not a forest, then $F_0$ is nonempty; it consists of all cycles of $F$ and the paths between them.

We will use the well-known fact that there are cubic graphs of arbitrary large treewidth. This fact dates back to Pinsker [26] who showed that a random cubic graph with high probability has good expansion properties, implying linear treewidth. An explicit example, pointed out to us by an anonymous referee, consists of a sequence of $k$ cycles $A_0, \ldots, A_{k-1}$, each of length $k$, where every vertex in $A_i$ is adjacent to the corresponding vertex in $A_{i+1}$ or in $A_{i-1}$, alternatingly (arithmetic operations in the subscript are modulo $k$, and $k$ is supposed to be even). This graph contains the $k$-wall graph as a subgraph, which is a well-known graph of maximum degree 3 and treewidth $\Omega(k)$; see, e.g., [18].

**Theorem 4.2**

1. $W_{tw}(F) \geq v(F_0)$ for every $F$.
2. $W_{tw}(F) \geq tw(F) + 1$ for every connected $F$ except for the case that $F$ is a subtree of some 3-megastar $M_{3,b}$.

Note that the bound in part 2 of Theorem 4.2 is tight by Theorem 4.1.

**Proof** 1. Denote $v(F) = \ell$ and $v(F_0) = \ell_0$. If $F$ is a forest, then $\ell_0 = 0$, and the claim is trivial. Suppose, therefore, that $F$ is not a forest. In this case, $\ell_0 \geq 3$.

We begin with a cubic graph $B$ of as large treewidth $tw(B)$ as desired. Let $(B)_\ell$ denote the graph obtained from $B$ by subdividing each edge by $\ell$ new vertices. Since $B$ is a minor of $(B)_\ell$, we have $tw((B)_\ell) \geq tw(B)$; see [11].

Next, we construct a gadget graph $A$ as follows. By a $k$-uniform tree we mean a tree of even diameter where every non-leaf vertex has degree $k$ and all distances between a leaf and the central vertex are equal. The graph $A$ is obtained by taking the $\ell$-uniform tree of radius $\ell$ and attaching $(B)_\ell$ to a leaf of this tree. More precisely, the tree and the graph $(B)_\ell$ are merged by identifying an arbitrary leaf of the tree and an arbitrary vertex of $(B)_\ell$.

We now construct $G$ by attaching a copy of $A$ to each vertex of $K_{\ell_0}$. Specifically, a copy $A_u$ of $A$ is created for each vertex $u$ of $K_{\ell_0}$, and $u$ is identified with the central vertex of (the tree part of) $A_u$. Let $H$ be obtained from $G$ by shrinking its clique part to $K_{\ell_0-1}$. Since both $G$ and $H$ contain copies of $(B)_\ell$, these two graphs have treewidth at least as large as $tw(B)$.

The clique part of $G$ is large enough to host the core $F_0$, and the remaining tree shoots of $F$ fit into the $A$-parts of $G$. Therefore, $G$ contains $F$ as a subgraph. On the other hand, the clique part of $H$ is too small for hosting $F_0$, and no cycle of $F$ fits
into any $A$-part because $A$ has larger girth than $F$. Therefore, $H$ does not contain $F$. It remains to notice that $W(G, H) \geq \ell_0$.

2. Suppose first that a connected graph $F$ is not a tree. By part 1, we then have

$$W_{tw}(F) \geq v(F_0) \geq tw(F_0) + 1 = tw(F) + 1.$$ 

If $F$ is a tree and is not contained in any 3-megastar, that is, has a vertex of degree more than 3 or at least two vertices of degree 3, then there are connected graphs of arbitrarily large treewidth that do not contain $F$ as a subgraph (for example, consider $(B)_\ell$ for a connected cubic graph $B$ as in part 1). Trivially, there are also connected graphs of arbitrarily large treewidth that contain $F$ as a subgraph. Since one pebble is not enough for Spoiler to distinguish the latter from the former, we have $W_{tw}(F) \geq 2 = tw(F) + 1$ in this case.

Theorem 4.2 implies that $W_{tw}(F) \geq tw(F)$ for all $F$. Combining it with the bound $W_{Arb}(F) \leq tw(F) + 3$ mentioned in Section 1, we obtain the following relation.

**Corollary 4.2.1** $W_{Arb}(F) \leq W_{tw}(F) + 3$.

Note that $W_{Arb}(F)$ and $W_{tw}(F)$ are within a constant factor from each other for infinitely many $F$. This is so for $F = K_\ell$ as $W_{Arb}(K_\ell) > \ell/4$ (Rossman [27]). On the other hand, a gap between the two parameters can be large. For example, part 1 of Theorem 4.2 gives $W_{tw}(C_\ell) = \ell$ whereas $W_{Arb}(C_\ell) \leq 5$.

### 5 Definitions Over Highly Connected Graphs

In this section we prove a lower bound for $D_\kappa(F)$ in terms of the density of $F$. The proof will use known facts about random graphs in the Erdős-Rényi model $G(n, p)$, collected below.

The density of a graph $K$ is defined to be the ratio $\rho(K) = e(K)/v(K)$. The maximum $\rho(K)$ over all subgraphs $K$ of a graph $F$ will be denoted by $\rho^*(F)$. The following fact from the random graph theory was used also in [21] for proving average-case lower bounds on the $\text{AC}^0$ complexity of SUBGRAPH ISOMORPHISM. With high probability means the probability approaching 1 as $n \to \infty$.

**Lemma 5.1** (Subgraph Threshold, see [17, Chap. 3]) If $\alpha = 1/\rho^*(F)$, then the probability that $G(n, n^{-\alpha})$ contains $F$ as a subgraph converges to a limit different from 0 and 1 as $n \to \infty$.

Let $\alpha > 0$. Given a graph $S$ and its subgraph $K$, we define $f_\alpha(S, K) = v(S) - v(K) - \alpha(e(S) - e(K))$.

**Lemma 5.2** (Generic Extension, see [1, Chap. 10]) Let $T$ be a graph with vertices $v_1, \ldots, v_t$ and $K$ be the induced subgraph of $T$ with vertices $v_1, \ldots, v_k$. Assume that $f_\alpha(S, K) > 0$ for every induced subgraph $S$ of $T$ containing $K$ as a proper subgraph. Then with high probability every sequence of pairwise distinct vertices $x_1, \ldots, x_k$ in
$G(n, n^{-\alpha})$ can be extended with pairwise distinct $x_{k+1}, \ldots, x_t$ such that, for all $i \leq t$ and $k < j \leq t$, the vertices $x_i$ and $x_j$ are adjacent in $G(n, n^{-\alpha})$ if and only if $v_i$ and $v_j$ are adjacent in $T$.

**Lemma 5.3** (Zero-One $d$-Law [36]) Let $0 < \alpha < \frac{1}{d-2}$, and $\Psi$ be a first-order statement of quantifier depth $d$. Then the probability that $\Psi$ is true on $G(n, n^{-\alpha})$ converges either to 0 or to 1 as $n \to \infty$.

We are now ready to prove our result.

**Theorem 5.4** If $e(F) > v(F)$, then $D_k(F) \geq \frac{e(F)}{v(F)} + 2$.

**Proof** Set $\alpha = 1/\rho^*(F)$ and denote $\mathbb{G}_n = G(n, n^{-\alpha})$. By Lemma 5.1, the limit probability $\lim_{n \to \infty} P[\mathbb{G}_n \in S(F)]$ exists and equals neither 0 nor 1. Assume that a first-order sentence $\Phi_1$ of quantifier depth $d$ defines $S(F)$ over $k$-connected graphs for all $k \geq k_0$. We have to prove that $d \geq \frac{e(F)}{v(F)} + 2$, whatever $k_0$.

By the assumption of the theorem, $\rho^*(F) \geq \rho(F) > 1$. Fix $k$ such that $1 + 1/k < \rho(F)$ and $k \geq k_0$. Lemma 5.2 implies that with high probability every two vertices in $\mathbb{G}_n$ can be connected by $k$ vertex-disjoint paths. To see this, consider $T$ formed by $k$ paths $P_{k+2}$ that share two terminal vertices $v_1$ and $v_2$. Let $K$ be the subgraph of $T$ consisting of only the vertices $v_1$ and $v_2$. Given a subgraph $S$ of $T$ that properly contains $K$, let $S_i$ denote the subgraph of $S$ contained in the $i$-th of the $k$ path components of $T$. Note that every $S_i$ contains $K$ and

$$f_{\alpha}(S, K) = \sum_{i=1}^{k} f_{\alpha}(S_i, K).$$

If $S_i = K$, that is, $S$ contains only the end vertices $v_1$ and $v_2$ of the $i$-th path of $T$, then $f_{\alpha}(S_i, K) = 0$. If $S_i$ contains $K$ properly, that is, $S$ contains at least one inner vertex of the $i$-th path, then it is easy to see that $f_{\alpha}(S_i, K) \geq k - \alpha(k+1)$. Since there must be at least one $S_i$ of the latter kind, we have

$$f_{\alpha}(S, K) \geq k - \alpha(k+1) > 0,$$

where the positiveness follows from the inequalities $\alpha \leq \frac{1}{\rho(F)} < \frac{k}{k+1}$. We conclude that $\mathbb{G}_n$ is $k$-connected with high probability.

Since $\Phi$ correctly decides the existence of a subgraph $F$ on all $k$-connected graphs,

$$P[\mathbb{G}_n \models \Phi] = P[\mathbb{G}_n \in S(F)] + o(1).$$

Therefore, $P[\mathbb{G}_n \models \Phi]$ converges to the same limit as $P[\mathbb{G}_n \in S(F)]$, which is different from 0 and 1. By Lemma 5.3, this implies that $\alpha \geq \frac{1}{d-2}$. From here we conclude that

$$d \geq \rho^*(F) + 2 \geq \frac{e(F)}{v(F)} + 2,$$

as required. $\square$
Table 1: Results on 4-vertex subgraphs

|   | $D_c(F)$ | $D_{tw}(F)$ | $D_e(F)$ |
|---|----------|-------------|----------|
|   | $W_c(F)$ | $W_{tw}(F)$ | $W_e(F)$ |
| (path) | $P_4$ | 1 | 2 | 3 |
| (claw) | $K_{1,3}$ | 1 | 3 | 4 |
| (paw) | $L_{3,1}$ | 3 | | 3 |
| (cycle) | $C_4$ | $\geq 3$ | 4 | 4 |
| (diamond) | $K_4 \setminus e$ | $\geq 3$ | 4 | 4 |
| (complete) | $K_4$ | 4 | | |

6 Small Pattern Graphs

Now we aim at determining exact values of the depth and the width parameters for small connected pattern graphs. There are only two connected graphs with 3 vertices, the path graph $P_3$ and the complete graph $K_3$. Since every connected graph with at least 3 vertices contains a subgraph $P_3$, we have $D_v(P_3) = 1$. Theorem 4.1 in the case of $L_{a,0} = K_a$ gives us

$$W_v(K_a) = a$$  \hspace{1cm} (7)$$

for every $a \geq 3$. In particular, $W_v(K_3) = 3$.

In this section, we consider the six connected graphs with 4 vertices, which are shown in Table 1. Each pattern graph $F$ is presented in this table by a row consisting of two layers, the upper for the depth parameters and the lower for the width parameters. Note that the values within each row are monotonically non-decreasing in the right and right upward directions. To improve visual clarity of the table, we remove all entries whose values are implied by monotonicity.

We begin with several simple observations. First, if a connected graph $H$ has $n > 3$ vertices and does not contain $P_4$, then $H = K_{1,n-1}$. Second, if a connected graph $H$ has $n$ vertices and does not contain $K_{1,3}$, that is, the maximum vertex degree of $H$ does not exceed 2, then $H = P_n$ or $H = C_n$. In each case, $H$ has treewidth at most 2. It readily follows that $D_{tw}(P_4) = 1$ and $D_{tw}(K_{1,3}) = 1$.

Moreover, the simple structure of connected $K_{1,3}$-free graphs easily implies that $W_v(K_{1,3}) = 3$ and $D_v(K_{1,3}) = 4$. The lower bounds follow here from the inequalities $W(M_{3,b}, P_n) > 2$ and $D(M_{3,b}, P_n) > 3$ that are true for all $b \geq 2$ and $n \geq 5$. To see the upper bound $W_v(K_{1,3}) \leq 3$, consider the 3-pebble Ehrenfeucht-Fraïssé game on $G$ with a vertex of degree at least 3 and $H = P_n$ or $H = C_n$ with $n > 6$. In the

2Note that our language does not contain the truth constants $\top$ and $\bot$; otherwise we would have $D_v(P_3) = 0$. 

© Springer
first three rounds, Spoiler pebbles three vertices in \( G \) having a common neighbor. If Duplicator is still alive then, whatever she responds, two of her vertices in \( H \) are at distance more than 2. This allows Spoiler to win in the next round.

By (7), we have \( W_\kappa(K_4) = 4 \). As another direct consequence of Theorem 4.1, \( W_\kappa(L_{3,1}) = D_v(L_{3,1}) = 3 \). Here, the upper bound \( D_v(L_{3,1}) \leq 3 \) follows from the observation that a connected graph with at least 4 vertices contains a subgraph \( L_{3,1} \) if and only if it contains a subgraph \( K_3 \).

If \( F \) is 2-connected, part 1 of Theorem 4.2 implies that \( W_{tw}(F) = v(F) \). This applies to the 4-cycle and the diamond graph, and we have the equalities \( W_{tw}(C_4) = 4 \) and \( W_{tw}(K_4 \setminus e) = 4 \). Furthermore, \( D_\kappa(K_4 \setminus e) = 4 \) as a consequence of Theorem 5.4. The lower bound \( W_\kappa(K_4 \setminus e) \geq 3 \) can be seen by considering, like in the proof of Theorem 4.1, the complete multipartite graphs \( G = K(3,n) \) and \( H = K(2,n) \).

The other entries of Table 1 are not so obvious.

**Theorem 6.1**

1. \( W_v(P_4) = 2 \) and \( D_v(P_4) = 3 \).
2. \( D_\kappa(C_4) = 4 \) and \( W_\kappa(C_4) \geq 3 \).

The proof of Theorem 6.1 takes the rest of this section.

**The path subgraph \( (P_4) \)** We restate part 1 of Theorem 6.1 as two lemmas.

**Lemma 6.2** \( D_v(P_4) = 3 \).

**Proof** We first show that \( D_\kappa^4(P_4) \leq 3 \). Indeed, the list of connected graphs not containing \( P_4 \) consists of \( K_1, K_3 \), and all stars \( K_{1,n-1} \). It suffices, therefore, to define the class of all star graphs with at least 3 vertices by a sentence of quantifier depth 3. This can be done as follows:

\[
\exists x \exists y \exists z \ (x \neq y \land x \neq z \land y \neq z) \land \\
\exists x \forall y \forall z \ (x = y \lor x = z \lor y = z \lor (y \sim x \land z \sim x \land y \sim z)),
\]

where \( \sim \) denotes the adjacency relation.

It remains to prove the lower bound. Consider \( G = J_{3,n} \) and \( H = J_{2,n} = K_{1,n+1} \). These graphs have \( n + 3 \) and \( n + 2 \) vertices respectively, and the parameter \( n \) can be chosen arbitrarily large. Note that \( G \) contains \( P_4 \) as a subgraph, while \( H \) does not. As easily seen, \( D(G, H) \geq 3 \).

**Lemma 6.3** \( W_v(P_4) = 2 \).

**Proof** Suppose that \( G \) contains a (not necessarily induced) subgraph \( P_4 \) and \( H \) does not. Let \( H \) have more than 3 vertices; then \( H = K_{1,n} \). We have to show that Spoiler wins the 2-pebble game on \( G \) and \( H \) making a bounded number of moves (that does not depend on how large \( G \) and \( H \) are). In the first round he pebbles the central vertex in \( H \). Suppose that \( G \) has a universal vertex for else Spoiler wins in the next round.
Not to lose in the next round, Duplicator pebbles a universal vertex $u$ in $G$. If $G$ has yet another universal vertex, Spoiler pebbles it and wins in the 3rd round by reusing the first pebble. Assume, therefore, that $u$ is the only universal vertex in $G$. Since $G$ is not a star graph, it contains a vertex $v \neq u$ having a neighbor $w \neq u$. Spoiler pebbles $v$ in the second round. Duplicator responds with leaf in $H$. In the third round Spoiler moves the pebble from $u$ to $w$. Duplicator is forced to respond with pebbling the central vertex in $H$. Spoiler wins in the fourth round by moving the pebble from $v$ to a vertex non-adjacent with $w$.

The cycle subgraph ($C_4$) To prove part 3 of Theorem 6.1, we need some properties of random regular graphs, established by Bollobás [5] and Wormald [35]. We collect them in the following lemma.

**Lemma 6.4** Let $d, g \geq 3$ be fixed, and $dn$ be even. Let $G_{n,d}$ denote a random $d$-regular graph on $n$ vertices.

1. $G_{n,d}$ is $d$-connected with high probability (see [5, Section 7.6] or [35, Section 2.6]).
2. $G_{n,d}$ has girth $g$ with probability that converges to a limit different from 0 and from 1 (see [5, Corollary 2.19] or [35, Theorem 2.5]).
3. $G_{n,d}$ has no non-trivial automorphism with high probability (see [5, Theorem 9.10]).

**Lemma 6.5** $D(C_4) = 4$.

**Proof** Fix $k$ as large as desired. Lemma 6.4 provides us with $k$-regular graphs $G$ and $H$ such that

1. $G$ has girth exactly 4, and $H$ has girth strictly more than 4;
2. both $G$ and $H$ are $k$-connected;
3. $G$ has no non-trivial automorphism;
4. both $G$ and $H$ have no less than $k^2 + 2$ vertices.

It suffices to show that $D(G, H) > 3$. To this end, we describe a strategy allowing Duplicator to win the 3-round game.

As usually, we denote the vertices pebbled in $G$ and $H$ in the $i$-th round by $x_i$ and $y_i$ respectively. We describe Duplicator’s move assuming that Spoiler has moved in $G$; the other case is symmetric with only one exception, which will occur in the end of our analysis.

1st round. Duplicator pebbles an arbitrary vertex $y_1$ in $H$.

2nd round. Duplicator pebbles $y_2$ such that $d(y_1, y_2) = d(x_1, x_2)$ if $d(x_1, x_2) \leq 2$ and $d(y_1, y_2) \geq 3$ if $d(x_1, x_2) \geq 3$. If $d(x_1, x_2) = 2$, such a vertex exists by the assumptions on the girth. If $d(x_1, x_2) \geq 3$, it exists because $H$ has more than $1 + k + k(k - 1)$ vertices.

3rd round. If $x_3$ is adjacent neither to $x_1$ nor to $x_2$, then Duplicator pebbles $y_3$ adjacent neither to $y_1$ nor to $y_2$, which exists because there are more than $2k + 2$
vertices in the graph. Assume, therefore, that $x_3$ is adjacent to at least one of $x_1$ and $x_2$.

**Case 1:** $d(x_1, x_2) = 1$ or $d(x_1, x_2) \geq 3$. Duplicator wins by pebbling $y_3$ adjacent either to $y_1$ or to $y_2$ depending on the neighborhood of $x_3$. Note that $x_3$ cannot be adjacent to both $x_1$ and $x_2$. If $d(x_1, x_2) = 1$, this follows from the assumption on the girth.

**Case 2:** $d(x_1, x_2) = 2$. If $x_3$ is adjacent to both $x_1$ and $x_2$, Duplicator wins by pebbling $y_3$ adjacent to both $y_1$ and $y_2$. It remains to consider the subcase when $x_3$ is adjacent to exactly one of $x_1$ and $x_2$, say, to $x_2$. Then Duplicator pebbles an arbitrary vertex $y_3$ adjacent to $y_2$. Note that $y_3$ and $y_1$ are not adjacent because $H$ has girth larger than 4. However, this argument does not work when Spoiler plays the 3rd round in $H$.

Thus, the following case takes more care: $d(x_1, x_2) = d(y_1, y_2) = 2$ and Spoiler pebbles $y_3$ adjacent to $y_2$ and not adjacent to $y_1$. Fortunately, Duplicator anyway has a choice of $x_3$ adjacent to $x_2$ and not adjacent to $x_1$. Indeed, if all neighbors of $x_2$ were adjacent also to $x_1$, then the transposition of $x_2$ and $x_1$ would be an automorphism of $G$, contradicting the assumption.

**Lemma 6.6** $W_k(C_4) \geq 3$.

**Proof** Fix an arbitrarily large $d$. Consider a random $d$-regular graph $G_{n,d}$ with sufficiently large number of vertices $n$. By Lemma 6.4, this graph is $d$-connected with high probability. Also, $G_{n,d}$ has girth 4 with nonzero probability, and as well it has girth 5 with nonzero probability. This yields us two $d$-connected graphs $G$ and $H$ of girth 4 and 5 respectively. Note that both graphs have neither universal nor isolated vertices. This readily implies that $W(G, H) > 2$.

The proof of Theorem 6.1 is complete.

### 7 Concluding Remarks and Questions

1. Is the bound $D_v(F) > v(F)/2$, given by Theorem 3.4, tight? On the other hand, currently we cannot disprove even that $D_v(F) \geq v(F) - O(1)$.

2. The relations $W_k(F) \leq W_{tw}(F) \leq W_v(F)$ and $D_k(F) \leq D_{tw}(F) \leq D_v(F)$ motivate an interest in lower bounds on the parameters $W_k(F)$ and $D_k(F)$. For instance, one can show that $D_k(C_\ell) \geq \log_3 \ell$. On the other hand, we currently cannot disprove even that $W_k(C_\ell) = O(1)$. The best what we can show in this direction is that $\ell$ first-order variables are necessary to define $S(C_\ell)$ over 2-connected graphs of large treewidth. For small graphs, we do not know whether or not $W_k(C_4) = 4$ and $W_k(K_4 \setminus e) = 4$; cf. Table 1.

---

3 Very recently [34], we were able to settle this problem by showing that $D_v(F) \leq \frac{3}{2} v(F) + 1$ for infinitely many pattern graphs $F$ and $W_v(F) > \frac{3}{2} v(F) - 2$ for all $F$. 

© Springer
3. It is known [6] that $tw(F) \geq e(F)/v(F)$. Can one improve Theorem 5.4 to $D_κ(F) \geq tw(F)$?

4. The parameters $D_L(C)$ and $W_L(C)$ have been studied in various contexts also for other graph properties $C$ and other logics $L$. We refer an interested reader to [10, 31]. In many cases, it would be interesting to compare $D_L(C)$ and $D_{L'}(C)$ (or $W_L(C)$ and $W_{L'}(C)$) for various different logics $L$ and $L'$ and the same property $C$.

Acknowledgements We would like to thank Tobias Müller for his kind hospitality during the Workshop on Logic and Random Graphs in the Lorentz Center (August 31 – September 4, 2015), where this work was originated. We are also thankful to the anonymous referees for a number of useful suggestions, in particular, for a simplification in the proof of Lemma 6.2.

References

1. Alon, N., Spencer, J.H.: The probabilistic method. Wiley, Hoboken (2016)
2. Alon, N., Yuster, R., Zwick, U.: Color-coding. J. ACM 42(4), 844–856 (1995). https://doi.org/10.1145/210332.210337
3. Amano, K.: $k$-Subgraph isomorphism on $\mathbb{AC}^0$ circuits. Comput. Complex. 19(2), 183–210 (2010). https://doi.org/10.1007/s00037-010-0288-y
4. Bodlaender, H.L.: A linear-time algorithm for finding tree-decompositions of small treewidth. SIAM J. Comput. 25(6), 1305–1317 (1996). https://doi.org/10.1137/S0097539793251219
5. Bollobás, B.: Random graphs Cambridge Studies in Advanced Mathematics, 2nd edn., vol. 73. Cambridge University Press, Cambridge (2001). https://doi.org/10.1017/CBO9780511814068
6. Chandran, L.S., Subramanian, C.R.: Girth and treewidth. J. Comb. Theory, Ser. B 93(1), 23–32 (2005). https://doi.org/10.1016/j.jctb.2004.05.004
7. Chekuri, C., Chuzhoy, J.: Polynomial Bounds for the Grid-Minor Theorem. In: Proceedings of the 46Th ACM Symposium on Theory of Computing (STOC’14), pp. 66–69 (2014). https://doi.org/10.1145/2591796.2591813
8. Chen, J., Huang, X., Kanj, I.A., Xia, G.: Strong computational lower bounds via parameterized complexity. J. Comput. Syst. Sci. 72(8), 1346–1367 (2006). https://doi.org/10.1016/j.jcss.2006.04.007
9. Courcelle, B.: The monadic second-order logic of graphs I. Recognizable sets of finite graphs. Inf. Comput. 85(1), 12–75 (1990). https://doi.org/10.1016/0890-5401(90)90043-H
10. Dawar, A.: A restricted second order logic for finite structures. Inf. Comput. 143(2), 154–174 (1998). https://doi.org/10.10106/inco.1998.2703
11. Diestel, R.: Graph theory. Springer, New York (2000)
12. Ebbinghaus, H.D., Flum, J.: Finite Model Theory, Enlarged Edn. Springer Monographs in Mathematics. Springer-Verlag, Berlin (2006)
13. Floderus, P., Kowaluk, M., Lingas, A., Lundell, E.: Induced Subgraph Isomorphism: Are some patterns substantially easier than others? Theor. Comput. Sci. 605, 119–128 (2015). https://doi.org/10.1016/j.tcs.2015.09.001
14. Gall, F.L.: Powers of tensors and fast matrix multiplication. In: Proc. of the Int. Symposium on Symbolic and Algebraic Computation (ISSAC’14), pp. 296–303. ACM (2014). https://doi.org/10.1145/2608628.2608664
15. Grädel, E., Grohe, M.: Is polynomial time choiceless? In: Fields of Logic and Computation II, Essays Dedicated to Yuri Gurevich on the Occasion of His 75th Birthday, Lecture Notes in Computer Science, vol. 9300, pp. 193–209. Springer (2015). https://doi.org/10.1007/978-3-319-23534-9_11
16. Immernan, N.: Descriptive complexity. Springer, New York (1999). https://doi.org/10.1007/978-1-4612-0539-5
17. Janson, S., Łuczak, T., Ruciński, A.: Random Graphs. Wiley, New York (2000)
18. Kawarabayashi, K., Mohar, B.: Some recent progress and applications in graph minor theory. Graphs Comb. 23(1), 1–46 (2007). https://doi.org/10.1007/s00373-006-0684-x
19. Kawarabayashi, K., Rossman, B.: An excluded-minor approximation of tree-depth. Manuscript (2016)
20. Koucký, M., Lautemann, C., Poloczek, S., Thérien, D.: Circuit Lower Bounds via Ehrenfeucht-Fraïssé Games. In: Proceedings of the 21ST Ann. IEEE Conference on Computational Complexity (CCC’06), pp. 190–201 (2006). https://doi.org/10.1109/CCC.2006.12
21. Li, Y., Razborov, A.A., Rossman, B.: On the AC^0 complexity of Subgraph Isomorphism. SIAM J. Comput. 46(3), 936–971 (2017). https://doi.org/10.1137/14099721X
22. Libkin, L.: Elements of finite model theory. Texts in theoretical computer science. An EATCS series. Springer, Berlin (2004). https://doi.org/10.1007/978-3-662-07003-1
23. Nešetřil, J., Poljak, S.: On the complexity of the subgraph problem. Commentat. Math. Univ. Carol. 26, 415–419 (1985)
24. Olariu, S.: Paw-free graphs. Inf. Process. Lett. 28, 53–54 (1988). https://doi.org/10.1016/0020-0190(88)90143-3
25. Pikhurko, O., Verbitsky, O.: Logical Complexity of Graphs: a Survey. In: Grohe, M., Makowsky, J. (eds.) Model Theoretic Methods in Finite Combinatorics, Contemporary Mathematics, vol. 558, pp. 129–179. American Mathematical Society (AMS), Providence (2011)
26. Pinsker, M.S.: On the complexity of a concentrator. In: Annual 7th International Teletraffic Conference, pp. 318/1–318/4 (1973)
27. Rossman, B.: On the Constant-Depth Complexity of K-Clique. In: Proceedings of the 40th Ann. ACM Symposium on Theory of Computing (STOC’08), pp. 721–730. ACM (2008). https://doi.org/10.1145/1374376.1374480
28. Rossman, B.: An improved homomorphism preservation theorem from lower bounds in circuit complexity. SIGLOG News 3(4), 33–46 (2016). http://dl.acm.org/citation.cfm?id=3026746
29. Rossman, B.: Lower bounds for Subgraph Isomorphism and consequences in first-order logic. Talk in the Workshop on Symmetry, Logic. Computation at the Simons Institute, Berkeley (2016). https://simons.berkeley.edu/talks/benjamin-rossman-11-08-2016
30. Schweikardt, N.: A short tutorial on order-invariant first-order logic. In: Proceedings of the 8th Int. Computer Science Symposium in Russia (CSR’13), Lecture Notes in Computer Science, vol. 7913, pp. 112–126. Springer (2013). https://doi.org/10.1007/978-3-642-38536-0_10
31. Turán, G.: On the definability of properties of finite graphs. Discret. Math. 49(3), 291–302 (1984). 10.1016/0012-365X(84)90166-3
32. Verbitsky, O., Zhukovskii, M.: The descriptive complexity of Subgraph Isomorphism without numerics. In: Weil, P. (ed.) Proceedings of the 12th Int. Computer Science Symposium in Russia (CSR’17), Lecture Notes in Computer Science, vol. 10304, pp. 308–322. Springer (2017). https://doi.org/10.1007/978-3-319-58747-9_27
33. Verbitsky, O., Zhukovskii, M.: On the first-order complexity of Induced Subgraph Isomorphism. In: Proceedings of the 26th EACSL Annual Conference on Computer Science Logic (CSL 2017), LIPics—Leibniz International Proceedings in Informatics, vol. 82, pp. 40:1–40:16 (2017)
34. Verbitsky, O., Zhukovskii, M.: Tight bounds on the asymptotic descriptive complexity of Subgraph Isomorphism. E-print: arXiv:1802.02143 (2018)
35. Wormald, N.: Models of Random Regular Graphs. In: Surveys in Combinatorics, pp. 239–298. Cambridge University Press (1999)
36. Zhukovskii, M.: Zero-one k-law. Discret. Math. 312, 1670–1688 (2012)