Restricted CSPs and F-free Digraph Algorithmics*

Santiago Guzmán-Pro†[1] and Barnaby Martin‡[2]

[1]Institut für Algebra, TU Dresden
[2]Durham University, UK

Abstract

In recent years, much attention has been placed on the complexity of graph homomorphism problems when the input is restricted to $\mathbb{P}_k$-free and $\mathbb{P}_k$-subgraph-free graphs. We consider the directed version of this research line, by addressing the questions *is it true that digraph homomorphism problems* CSP($\mathbb{H}$) *have a* P *versus* NP*-complete dichotomy when the input is restricted to* $\vec{\mathbb{P}}_k$*-free (resp.* $\vec{\mathbb{P}}_k$*-subgraph-free) digraphs?* Our main contribution in this direction shows that if CSP($\mathbb{H}$) is NP-complete, then there is a positive integer N such that CSP($\mathbb{H}$) remains NP-hard even for $\vec{\mathbb{P}}_N$-subgraph-free digraphs. Moreover, it remains NP-hard for *acyclic* $\vec{\mathbb{P}}_N$-subgraph-free digraphs, and becomes polynomial-time solvable for $\vec{\mathbb{P}}_{N-1}$-subgraph-free *acyclic* digraphs. We then verify the questions above for digraphs on three vertices and a family of smooth tournaments. We prove these results by establishing a connection between $\mathbb{F}$-(subgraph)-free algorithmics and constraint satisfaction theory. On the way, we introduce *restricted CSPs*, i.e., problems of the form CSP($\mathbb{H}$) restricted to yes-instances of CSP($\mathbb{H}'$) — these were called restricted homomorphism problems by Hell and Nešetřil. Another main result of this paper presents a P versus NP-complete dichotomy for these problems. Moreover, this complexity dichotomy is accompanied by an algebraic dichotomy in the spirit of the finite domain CSP dichotomy.

As little as a few years ago, most graph theorists, while passively aware of a few classical results on graph homomorphisms, would not include homomorphisms among the topics of central interest in graph theory. We believe that this perception is changing, principally because of the usefulness of the homomorphism perspective... At the same time, the homomorphism framework strengthens the link between graph theory and other parts of mathematics, making graph theory more attractive, and understandable, to other mathematicians. Pavol Hell and Jaroslav Nesetril, *Graphs and Homomorphisms*, 2004 [31].

*Barnaby Martin has been supported by EPSRC grant EP/X03190X/1. Santiago Guzmán-Pro has been funded by the European Research Council (Project POCOCOP, ERC Synergy Grant 101071674). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

†santiago.guzman_pro@tu-dresden.de
‡barnabymartin@gmail.com

1 Introduction

Story of the main questions

The Hell-Nešetřil theorem asserts that if $\mathbb{H}$ is a finite undirected graph, then $\mathrm{CSP}(\mathbb{H})$ is polynomial-time solvable whenever $\mathbb{H}$ is either a bipartite graph or contains a loop, and otherwise $\mathrm{CSP}(\mathbb{H})$ is NP-complete. In recent years, much attention has been placed on the complexity of graph homomorphism problems when the input is restricted to $\mathbb{F}$-free and $\mathbb{F}$-subgraph-free graphs, i.e., to graphs avoiding $\mathbb{F}$ as an induced subgraph, and as a subgraph, respectively. In [26], the authors show that $\mathrm{CSP}(\mathbb{C}_5)$ is polynomial-time solvable when the input is restricted to $\mathbb{P}_8$-free graphs. Moreover, they show that there are finitely many $\mathbb{P}_8$-free obstructions to $\mathrm{CSP}(\mathbb{C}_5)$. Also, in [15] the authors prove that $\mathrm{CSP}(\mathbb{K}_3)$ becomes tractable if the input is restricted to $\mathbb{P}_7$-free graphs, while Huang [32] proved that $\mathrm{CSP}(\mathbb{K}_4)$ remains NP-hard even for $\mathbb{P}_7$-free graphs. In general, several complexity classifications are known for $\mathbb{K}_n$-Colouring with input restriction to $\mathbb{P}_k$-free graphs (see, e.g., [27, Theorem 7]), however, a complete complexity classification of graph colouring problems with input restriction fo $\mathbb{F}$-(subgraph)-free graphs remains wide open.

The finite domain CSP dichotomy [38] (announced independently by Bulatov [18] and Zhuk [39]) asserts that digraph colouring problems also exhibit a P versus NP-complete dichotomy. In this paper we consider the directed version of the research line described above, where we consider the following to be the long-term question of this variant: Is there a P versus NP-complete dichotomy of $\mathrm{CSP}(\mathbb{H})$ where the input is restricted (1) to $\mathbb{F}$-free digraphs? and (2) to $\mathbb{F}$-subgraph-free digraphs?[1]

The Sparse Incomparability Lemma asserts that if $\mathrm{CSP}(\mathbb{H})$ is NP-hard, then it remains NP-hard even for high-girth digraphs. Hence, both questions above have a positive answer whenever $\mathbb{F}$ is not an oriented forest. Motivated by the literature on graph colouring problems restricted to $\mathbb{P}_k$-free and $\mathbb{P}_k$-subgraph-free digraphs, we consider the restriction of these questions to the case when F is a directed path $\vec{\mathbb{P}}_k$.

Question 1. *Is there a* P *versus* NP*-complete dichotomy of* $\mathrm{CSP}(\mathbb{H})$ *where the input is restricted*

1. *to $\vec{\mathbb{P}}_k$-free digraphs?*
2. *to $\vec{\mathbb{P}}_k$-subgraph-free digraphs?*

In our effort to settle Question 1 we stumble into three more questions which we also address in this paper. Allow us to elaborate. Clearly, if $\mathrm{CSP}(\mathbb{H})$ if NP-hard even for $\vec{\mathbb{P}}_k$-subgraph-free digraphs, then $\mathrm{CSP}(\mathbb{H})$ is NP-hard for $\vec{\mathbb{P}}_k$-free digraphs. In turn, if $\mathrm{CSP}(\mathbb{H})$ restricted to $\vec{\mathbb{P}}_k$-homomorphism-free digraphs is NP-hard, then $\mathrm{CSP}(\mathbb{H})$ restricted to $\vec{\mathbb{P}}_k$-subgraph-free digraphs. It is well-known that a digraph $\mathbb{D}$ is $\vec{\mathbb{P}}_k$-homomorphism-free if and only if $\mathbb{D}$ homomorphically maps to the transitive tournament in $k-1$ vertices $\mathbb{TT}_{k-1}$ (see, e.g, Observation 4). Hence, a simple way to find complexity upperbounds to the problems in Question 1 is to consider the complexity of $\mathrm{CSP}(\mathbb{H})$ restricted to $\mathrm{CSP}(\mathbb{TT}_k)$.

The last problems have the following natural generalization: decide $\mathrm{CSP}(\mathbb{H})$ restricted to input digraphs $\mathbb{D}$ in $\mathrm{CSP}(\mathbb{H}')$. We denote this problem by $\mathrm{RCSP}(\mathbb{H},\mathbb{H}')$. These problems have been called restricted homomorphism problems [16, 17] and it was conjectured by Hell and Nešetřil that

[1]Note that a negative answer to any of these questions implies that, if $\mathrm{P} \neq \mathrm{NP}$, then there are finite digraphs $\mathbb{F}$ and $\mathbb{H}$ such that $\mathrm{CSP}(\mathbb{H})$ restricted to $\mathbb{F}$-(subgraph)-free digraphs is NP-intermediate — which we believe to be a more natural problem than the current NP-intermediate problems constructed in the literature. At this point we conjecture neither a positive nor a negative answer to the previous questions.

when $\mathbb{H}$ and $\mathbb{H}'$ are undirected graphs, then $\mathbb{H}' \to \mathbb{H}$, or $\mathbb{H}$ is bipartite, and in these cases CSP($\mathbb{H}$) restricted to CSP($\mathbb{H}'$) is in P; and otherwise it is NP-hard. This was later confirmed by Brewster and Graves [16], where they actually propose a hardness condition for a broader family of digraphs $\mathbb{H}'$ (see Theorem 3 below). It is then natural to ask our second main question.

Question 2. *Is there a* P *versus* NP*-hard dichotomy of problems* RCSP($\mathbb{H}, \mathbb{H}'$) *parametrized by finite digraphs (structures)* $\mathbb{H}$ *and* $\mathbb{H}'$*?*

Clearly, every digraph $\mathbb{D}$ that admits a homomorphism to some transitive tournament must be an acyclic digraph. A natural question of a digraph $\mathbb{H}$, for which CSP($\mathbb{H}$) is NP-complete, is to ask if this problem remains NP-complete on acyclic instances. Notice that this is the same problem as RCSP($\mathbb{H}, (\mathbb{Q}, <)$). If $\mathbb{H}$ is an undirected graph, then of course this will be true (choose any total order of the vertex set and orient the edges according to this ordering). However, when $\mathbb{H}$ is not undirected, the situation is not a priori clear. Indeed, it is addressed for some small digraphs by Hell and Mishra in [30]. They prove, for example, that CSP($\vec{\mathbb{C}}_3^+$) – where $\vec{\mathbb{C}}_3^+$ is drawn in the forthcoming Figure 4 – does indeed remain NP-complete on acyclic inputs. The general question is not posed in [30], but it is perfectly natural, and we address it here.

Question 3. *If* CSP($\mathbb{H}$) *is* NP*-hard for a finite graph* $\mathbb{H}$*, does* CSP($\mathbb{H}$) *remain* NP*-hard for acyclic instances? Equivalently, is* RCSP($\mathbb{H}, (\mathbb{Q}, <)$) NP*-hard whenever* CSP($\mathbb{H}$) *is* NP*-hard?*

Some readers might have already noticed that there is a third natural question motivated by the previous paragraph: is there a P versus NP-complete dichotomy of CSP($\mathbb{H}$) restricted to $\vec{\mathbb{P}}_k$-homomorphism-free digraphs? Or more generally: is there a P versus NP-hard dichotomy of CSP($\mathbb{H}$) where the input is restricted to $\mathcal{F}$-homomorphism-free digraphs? (where $\mathcal{F}$ is a fixed finite set of digraphs). This question can already be settled from results in the literature: every such problem can be coded into *monotone monadic strict NP* (MMSNP), and Feder and Vardi [23] proved that if finite domain CSPs have a P versus NP-complete dichotomy, then MMSNP exhibits the same dichotomy. However, we provide an alternative prove of this fact in Section 5.

Story of the paper

Our paper splits into two parts. The first part focuses on relational structures, of which digraphs are somewhat canonical examples, while the second part focuses on digraphs specifically. In Figure 1 we depict the flow of ideas and results of this paper.

Within the first part (Sections 3– 5), we begin by introducing *restricted* CSPs (RCSPs) which have been studied as restricted homomorphism problems in [16, 17]. In particular, we elaborate a connection with promise CSPs (PCSPs) that enables one to view RCSPs as PCSPs.

Our first main result shows that, for every pair of finite digraphs (structures) $\mathbb{A}$ and $\mathbb{B}$, the problem CSP($\mathbb{B}$) with input restricted to CSP($\mathbb{A}$) is either in P or NP-hard (settling Question 2). Moreover, this complexity classification is accompanied with an algebraic dichotomy (Theorem 21) which stems from the finite domain CSP dichotomy result. We then push the previous complexity dichotomy to finite domain CSPs with restrictions in GMSNP (Theorem 23). Another contribution of this work builds again on the finite domain CSP dichotomy to present a new proof of the complexity dichotomy for finite domain CSP restricted to $\mathcal{F}$-homomorphism-free digraphs (structures): we avoid going via MMSNP to the infinite domain CSP setting, and stay in the finite domain world via Lemma 27 (Section 5).

In the second part of the paper (Sections 6–9), we leverage results from the first part, in order to study digraph CSPs, where the ultimate focus will be on $\mathbb{F}$-free and $\mathbb{F}$-subgraph-free algorithmics.

Our second main result shows that, if $\mathbb{H}$ is a digraph and CSP($\mathbb{H}$) is NP-complete, then there is a positive integer N such that CSP($\mathbb{H}$) remains NP-complete even for $\vec{\mathbb{P}}_N$-subgraph-free acyclic inputs (settling Question 3). Moreover, N can be chosen so that CSP($\mathbb{H}$) is polynomial-time solvable for $\vec{\mathbb{P}}_{N-1}$-subgraph-free acyclic inputs. This also yields a partial answer to Question 1: for every digraph $\mathbb{H}$ there is a positive integer $N \leq 4^{|H|}$ such that Question 1 has a positive answer restricted to $k \geq N$. We complement this general partial answer by settling Question 1 for digraphs $\mathbb{H}$ on three vertices (Theorems 42 and 46), and for a family of smooth tournaments $\mathbb{TC}_n$ (Theorems 49 and Theorem 53). We note that eventual hardness on $\vec{\mathbb{P}}_N$-subgraph-free instances does not hold in general for CSP($\mathbb{H}$), if $\mathbb{H}$ is an infinite digraph. We provide a counterexample (Example 36) which is otherwise well-behaved (for example, by being ω-categorical). As byproducts of our work we see that there are finitely many minimal $\vec{\mathbb{P}}_3$-obstructions to CSP($\mathbb{TC}_n$) for each positive integer n (Theorem 55), and if $\mathbb{F}$ is not an oriented path, then CSP($\mathbb{TC}_n$) (and CSP($\vec{\mathbb{C}}_3^+$)) is NP-hard even for $\mathbb{F}$-subgraph-free instances (Theorem 57).

2 Preliminaries

2.1 Relational structures and digraphs

For a finite relational signature $\tau = (R_1, \ldots, R_k)$, a *relational ($\tau$-)structure* $\mathbb{A}$ on domain A consists of k relations $R_1 \subseteq A^{a_1}, \ldots, R_k \subseteq A^{a_k}$, where a_i is the arity of R_i. We denote the cardinality of A as $|A|$. We tend to conflate the relation symbol and actual relation since this will not introduce confusion.

For some signature τ, the *loop* $\mathbb{L}$ is the structure on one element a all of whose relations are maximally full, that is, contains one tuple $(a, \ldots, a)$ — the structure $\mathbb{L}$ clearly depends on the signature τ, but in this work τ will always be clear from context.

Directed graphs (digraphs) $\mathbb{D}$ are relational structures on the signature $\{E\}$ where E is a binary relation. A digraph is a *graph* if E is symmetric, i.e. $(x, y) \in E$ iff $(y, x) \in E$. We will use the same blackboard font notation for digraphs as we do for relational structures.

Given a positive integer k we denote by $\vec{\mathbb{C}}_k$ the directed cycle on k vertices, by $\vec{\mathbb{P}}_k$ the directed path on k vertices, by $\mathbb{T}_k$ the transitive tournament on k vertices. Similarly we denote by K_k the complete graph on k vertices, and we think of it as a digraph with edges (i, j) for every $i \neq j$.

A *(directed) walk* on a (directed) graph $\mathbb{D}$ is a sequence of vertices $x_1, \ldots, x_k$ such that for every $i \in [k-1]$ there is a (directed) edge $(x_i, x_{i+1}) \in E$. An *oriented graph* is a digraph with no pair of symmetric edges, i.e,. a loopless $\mathbb{K}_2$-free digraph. A *tree* is an oriented digraph whose underlying graph has no cycle (equivalently, is an orientation of a traditional undirected tree). It follows that trees are $\mathbb{K}_2$-free. A *forest* is a disjoint union of trees.

2.2 Constraint satisfaction problems

Given a pair of digraphs (structures) $\mathbb{D}$ and $\mathbb{H}$ a *homomorphism* $f\colon \mathbb{D} \to \mathbb{H}$ is a vertex mapping such that, for every (u, v) that is an edge of $\mathbb{D}$, the image $(f(u), f(v))$ is an edge of $\mathbb{H}$. If such a homomorphism exists, we write $\mathbb{D} \to \mathbb{H}$, and otherwise $\mathbb{D} \not\to \mathbb{H}$. We follow notation of constraint satisfaction theory, and denote by CSP($\mathbb{H}$) the class of finite digraphs such that $\mathbb{D} \to \mathbb{H}$. We also denote by CSP($\mathbb{H}$) the computational problem of deciding if an input digraph $\mathbb{D}$ belongs to CSP($\mathbb{H}$).

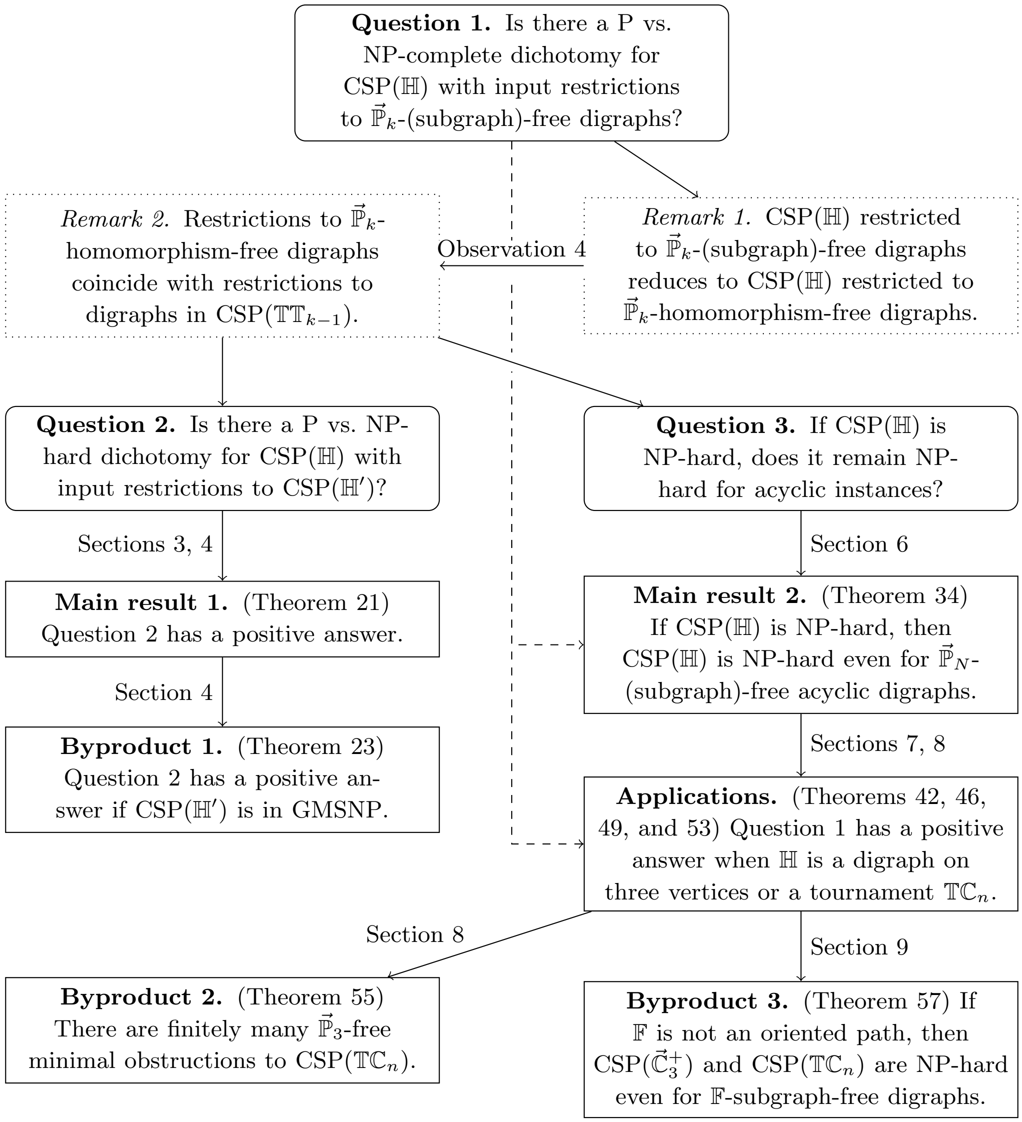


Figure 1: A depiction of the flow between main results and main questions addressed in this paper. Rectangles with rounded corners mark the main questions considered here. Dotted squares indicate the (simple) remarks connecting the $\mathcal{F}$-subgraph-free CSPs to the theory of digraph homomorphisms. Solid rectangles with straight corners indicate the main results of the present paper. A solid (resp. dashed) edge between a question and a result indicates that the result provides an answer (resp. a partial answer) to the corresponding question. Finally, edges between results represent that the result at the head is proved using tools introduced while proving the result at the tail of the corresponding edge.

An *endomorphism* is a homomorphism $f\colon \mathbb{H} \to \mathbb{H}$. A digraph (structure) $\mathbb{H}$ is a *core* if all of its endomorphisms are self-embeddings. If $\mathbb{H}$ is finite then this is equivalent to all endomorphisms being automorphisms.

2.3 Smooth digraphs

A digraph $\mathbb{D}$ is *smooth* if it has no sources nor sinks. The following statement was conjectured in [1] and proved in [7].

Theorem 1 (Conjecture 6.1 in [1] proved in [7])**.** *For every smooth digraph $\mathbb{H}$ one of the following holds.*

- *The core of $\mathbb{H}$ is a disjoint union of cycles, and in this case* $\mathrm{CSP}(\mathbb{H})$ *is polynomial-time solvable.*
- *Otherwise,* $\mathrm{CSP}(\mathbb{H})$ *is* NP*-complete.*

A digraph $\mathbb{H}$ is *hereditarily hard* if $\mathrm{CSP}(\mathbb{H}')$ is NP-complete for every loopless digraph $\mathbb{H}'$ such that $\mathbb{H} \to \mathbb{H}'$. Bang-Jensen, Hell, and Niven conjectured that a smooth digraph $\mathbb{H}$ is hereditarily hard whenever $\mathbb{H}$ does not homomorphically map to a disjoint union of directed cycles. Moreover, they showed that this conjecture is implied by the statement in Theorem 1 (which was only a conjecture at that time).

Theorem 2 (Conjecture 2.5 in [3] proved in [7])**.** *A digraph $\mathbb{H}$ is hereditarily hard whenever the digraph $R(\mathbb{H})$ obtained from $\mathbb{H}$ by iteratively removing sources and sinks does not admit a homomorphism to a disjoint unions of directed cycles.*

As far as we are aware, the most general result regarding hardness of restricted CSP problems is Theorem 3 in [16].

Theorem 3 (Theorem 3 in [16])**.** *If $\mathbb{H}$ is a hereditarily hard digraph and $\mathbb{H}'$ is a finite digraph such that $\mathbb{H}' \not\to \mathbb{H}$ then* $\mathrm{RCSP}(\mathbb{H}, \mathbb{H}')$ *is* NP*-hard.*

2.4 Duality pairs

A pair of digraphs (relational structure) $(\mathbb{T}, \mathbb{D}_{\mathbb{T}})$ are called a *duality pair* if for every digraph it is the case that $\mathbb{D} \to \mathbb{D}_{\mathbb{T}}$ if and only if $\mathbb{T} \not\to \mathbb{D}$. It was proved in [35] that for every tree $\mathbb{T}$ there is a digraph $\mathbb{D}_{\mathbb{T}}$ such that $(\mathbb{T}, \mathbb{D}_{\mathbb{T}})$ is a duality pair. Moreover, they also proved that if $(\mathbb{T}, \mathbb{D}_{\mathbb{T}})$ is a duality pair, then $\mathbb{T}$ is homomorphically equivalent to a tree. A well-known example of a family of duality pairs is the following one.

Observation 4. *For every positive integer k the directed path $\vec{\mathbb{P}}_{k+1}$ together with the transitive tournament $\mathbb{T}_k$ are duality pair.*

Given a set of digraphs $\mathcal{F}$, we denote by $\mathrm{Forb}(\mathcal{F})$ the set of digraphs $\mathbb{D}$ such that $\mathbb{F} \not\to \mathbb{D}$ for every $\mathbb{F} \in \mathcal{F}$. It is straightforward to observe that for any such set $\mathcal{F}$, there is a (possibly infinite) digraph $\mathbb{D}$ such that $\mathrm{Forb}(\mathcal{F}) = \mathrm{CSP}(\mathbb{D})$. A *generalized duality* is a pair $(\mathcal{F}, \mathcal{D})$ of finite sets of digraphs such that

$$\mathrm{Forb}(\mathcal{F}) = \bigcup_{\mathbb{D} \in \mathcal{D}} \mathrm{CSP}(\mathbb{D}).$$

In particular, when $\mathcal{D} = \{\mathbb{D}\}$ we simply write $(\mathcal{F}, \mathbb{D})$. Generalized dualities have a similar characterization to duality pairs.

Theorem 5 (Theorems 2 and 11 in [25])**.** *For every finite set of forests $\mathcal{F}$ there is a finite set of digraphs $\mathcal{D}$ such that $(\mathcal{D}, \mathcal{F})$ is a generalized duality pair. Moreover, if $\mathcal{F}$ is a finite set of trees, then there is a digraph $\mathbb{D}$ such that $(\mathcal{F}, \mathbb{D})$ is a generalized duality.*

2.5 Large girth

A well-known result from Erdős [22] about k-colourabilty states that for every pair of positive integers l, k there is a graph G with girth strictly larger than l and such that G does not admit a proper k-colouring. This result generalizes to arbitrary relational structures, and it is known as the Sparse Incomprability Lemma [34] — in order to stay within the scope of this paper, we state it for digraphs.

Given a digraph (structure) $\mathbb{D}$, the *incidence graph* of $\mathbb{D}$ is the undirected bipartite graph $\mathbb{I}(\mathbb{D})$ with vertex set $V \cup E$, for $v \in D$ a vertex of D and $e = (x, y) \in E$ an edge of D. There is an (undirected) edge (v, e) in $\mathbb{I}(\mathbb{D})$ if and only if $v \in \{x, y\}$. The *girth* of $\mathbb{D}$ is half the length of the shortest cycle in $\mathbb{I}(\mathbb{D})$. Notice that if $\mathbb{D}$ has a pair of symmetric arcs, its girth is 2, and otherwise, it is the graph theoretic girth of the underlying graph of $\mathbb{D}$.

Theorem 6 (Sparse Incomparability Lemma [34])**.** *For every digraph $\mathbb{D}$ and every pair of positive integers k and ℓ there is a digraph $\mathbb{D}'$ with the following properties:*

- $\mathbb{D}' \to \mathbb{D}$,
- *the girth of $\mathbb{D}'$ is larger than ℓ,*
- $\mathbb{D} \to \mathbb{H}$ *if and only if* $\mathbb{D}' \to \mathbb{H}$ *for every digraph $\mathbb{H}$ on at most k vertices,*
- $\mathbb{D}'$ *can be constructed in polynomial time (from $\mathbb{D}$).*

Corollary 7. *For ever finite digraph $\mathbb{H}$ and every positive integer ℓ,* CSP($\mathbb{H}$) *is polynomial-time equivalent to* CSP($\mathbb{H}$) *restricted to input digraphs of girth strictly larger that ℓ.*

3 Restricted constraint satisfaction problems

Promise problems (not to be confused with Promise CSPs) can be thought as decision problems with input restrictions. Formally [37], a *promise problem* is a pair $(\mathcal{P}, \mathcal{C})$ of decidable sets. A solution to $(\mathcal{P}, \mathcal{C})$ is a decidable set $\mathcal{S}$ such that $\mathcal{S} \cap \mathcal{P} = \mathcal{C} \cap \mathcal{P}$. We say that the promise problem $(\mathcal{P}, \mathcal{C})$ is polynomial-time solvable if it has a solution in P, and if every solution is NP-hard, we say that $(\mathcal{P}, \mathcal{C})$ is NP-hard.

Given a pair of (possibly infinite) structures $\mathbb{A}$ and $\mathbb{B}$ (with the same finite signature) whose CSPs are decidable, the *restricted CSP* RCSP($\mathbb{A}, \mathbb{B}$) is the promise problem (CSP($\mathbb{B}$), CSP($\mathbb{A}$)). In this case, we call $(\mathbb{A}, \mathbb{B})$ the *template* of the restricted CSP, $\mathbb{A}$ is called the *domain* and $\mathbb{B}$ the *restriction.* In particular, if $\mathbb{A}$ is finite we say that RCSP($\mathbb{A}, \mathbb{B}$) is a finite domain RCSP, and if $\mathbb{B}$ is finite, we say that RCSP($\mathbb{A}, \mathbb{B}$) is an RCSP with finite restriction.

Informally, the promise problem RCSP($\mathbb{A}, \mathbb{B}$) is CSP($\mathbb{A}$) where the input is promised to belong to CSP($\mathbb{B}$). For instance, RCSP($\mathbb{K}_3, \mathbb{K}_4$) is essentially the problem of deciding whether an input 4-colourable graph is 3-colourable.

Notice that for any digraph (structure) $\mathbb{A}$ the problems CSP($\mathbb{A}$) and RCSP($\mathbb{A}, \mathbb{L}$) are the same problems where $\mathbb{L}$ is the loop. So every decidable CSP is captured by an RCSP with finite restriction. One of the main results of this work shows that every RCSP with finite restriction is log-space equivalent to a CSP. It follows from the proof that every finite domain RCSP with finite restriction is log-space equivalent to a finite domain CSP, and thus, finite domain restricted CSPs with finite restrictions have a P versus NP-hard dichotomy. Moreover, the reduction mentioned in this paragraph are obtained by *restricted primitive positive construction (rpp-constructions)*, which are the natural cousins of pp-constructions for CSPs [9] and for PCSPs [5].

3.1 Restricted primitive positive constructions

Several reductions in graph algorithmics arise from gadget reductions. For instance, a standard way of proving that CSP($\mathbb{C}_5$) is NP-complete can be done with the following gadget reduction from CSP($\mathbb{K}_5$). On input $\mathbb{G}$ to CSP($\mathbb{K}_5$), consider the graph obtained from $\mathbb{G}$ by replacing every edge $e := xy$ by a path x, u_e, v_e, y (see, e.g., [31] where this is called an indicator construction). So in this case, the path on four vertices is the gadget associated to this reduction). The algebraic approach to CSPs proposes a general framework encompassing gadget reductions between constraint satisfaction problems. In this section we introduce primitive positive constructions, and *restricted* primitive positive constructions.

Primitive positive constructions

Given a pair of finite relational signature τ and σ, a *primitive positive definition* (of τ in σ)[2] of dimension $d \in \mathbb{Z}^+$ is a finite set Δ of primitive positive formulas $\delta_R(\bar{x})$ indexed by σ, and for each $R \in \sigma$ of arity r the formula $\delta_R(\bar{x})$ has $r \cdot d$ free variables.

For every primitive positive definition of σ in τ we associate a mapping Π_Δ from τ-structures to σ-structures as follows. Given a τ-structure $\mathbb{A}$ the *pp power* $\Pi_\Delta(\mathbb{A})$ of $\mathbb{A}$ is the structure

- with domain of $\Pi_\Delta(\mathbb{A})$ is A^d, and
- for each $R \in \sigma$ of arity r the interpretation of R in $\Pi_\Delta(\mathbb{A})$ consists of the tuples $(\bar{a}_1, \ldots, \bar{a}_r) \in (A^d)^r$ such that $\mathbb{A} \models \delta_R(\bar{a}_1, \ldots, \bar{a}_r)$.

We say that a structure $\mathbb{A}$ *pp-constructs* a structure $\mathbb{B}$ is there is a primitive positive definition Δ such that $\Pi_\Delta(\mathbb{A}) \to \mathbb{B} \to \Pi_\Delta(\mathbb{A})$, i.e., $\Pi_\Delta(\mathbb{A})$ is homomorphically equivalent to $\mathbb{B}$. For instance, if Δ is the 1-dimensional primitive positive definition consisting of

$$\delta_E(x, y) := \exists z_1, z_2.\ E(x, z_1) \wedge E(z_1, z_2) \wedge E(z_2, y),$$

then the pp power $\Pi_\Delta(\mathbb{C}_5)$ of the 5-cycle is the complete graph $\mathbb{K}_5$ (see Figure 2).

Remark 8. It is well-known that for every fixed primitive positive definition Δ, the pp-power Π_Δ is a monotone construction with respect to the homomorphism order, i.e., if $\mathbb{A} \to \mathbb{B}$, then $\Pi_\Delta(\mathbb{A}) \to \Pi_\Delta(\mathbb{B})$ (this follows from the fact that existential positive formulas are preserved by homomorphism [11, Theorem 2.5.2]).

Lemma 9. *For every primitive positive definitions Δ_1 of σ in τ, and Δ_3 of τ in ρ, there is a primitive positive definition of Δ_3 of σ in ρ such that for every ρ-structure $\mathbb{A}$*

$$\Pi_{\Delta_3}(\mathbb{A}) = \Pi_{\Delta_1}(\Pi_{\Delta_2}(\mathbb{A})).$$

[2] When the signatures are clear from context we will simply say talk about a primitive positive definition.

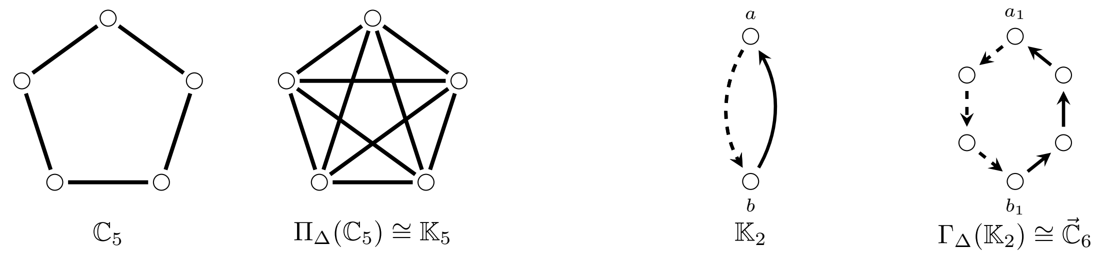


Figure 2: Let Δ be the primitive positive definition (of $\{E\}$ in $\{E\}$) where $\delta_E(x,y) := \exists z_1, z_2.\ E(x,z_1) \wedge E(z_1,z_2) \wedge E(z_2,y)$. On the left, we depict $\mathbb{C}_5$ and its pp power $\Pi_\Delta(\mathbb{C}_5) \cong \mathbb{K}_5$ (an undirected edge xy represents (x,y) and (y,x)), and on the right, we depict $\mathbb{K}_2$ and its gadget replacement $\Gamma_\Delta(\mathbb{K}_2) \cong \vec{\mathbb{C}}_6$ (dashed edges and solid edges indicate the respective edge replacements).

Proof. By substituting each relation symbol $R \in \tau$ in a formula $\delta^1_S \in \Delta_1$ by the existential positive formula $\delta^2_R \in \Delta_2$, we obtain a set Δ_3 of existential positive formulas δ^3_S indexed by $S \in \sigma$. It is not hard to observe that this primitive positive definition Δ_3 satisfies the claim of this lemma. For further details we refer the reader to [11, Section 4.1.2]. □

Primitive positive constructions yield a canonical class of log-space reductions between constraint satisfaction problems.

Lemma 10 (Corollary 3.5 in [9])**.** *Let $\mathbb{A}$ and $\mathbb{B}$ be (possibly infinite) structures with finite relational signature. If $\mathbb{A}$ pp-constructs $\mathbb{B}$, then* CSP($\mathbb{B}$) *reduces in logarithmic space to* CSP($\mathbb{A}$).

Gadget replacements

For every primitive positive formula $\delta(\bar{x})$ without equalities[3] we construct a structure $\mathbb{D}_\delta(\bar{d})$ with a distinguished tuples of vertices d as follows.

- The domain D of $\mathbb{D}$ consists of a vertex v_y for every (free or bounded) variable y of δ.
- The distinguished vertices $d_1, \dots, d_m$ are the vertices $v_{x_1}, \dots, v_{x_m}$.
- For every relation symbol R there is a tuple $\bar{v} \in R^{\mathbb{D}}$ if and only if δ contains the conjunct $R(\bar{y})$ where $\bar{v}$ is the tuple of vertices corresponding to the tuple of variables $\bar{y}$.

The structure $\mathbb{D}(\bar{d})$ is sometimes called the *canonical database* of δ. For instance, if δ_E is the formula considered above, the the canonical database of δ_E is the directed with vertices $1, 2, 3, 4$, and the distinguished vertices are 1 and 4.

The canonical database $\mathbb{D}_\delta(\bar{d})$ and the primitive positive formula δ are closely related: for every structure $\mathbb{A}$ and a tuple $\bar{a}$ of elements of A

$$\mathbb{A} \models \delta(\bar{a}) \text{ if and only if there is a homomorphism } f \colon \mathbb{D} \to \mathbb{A} \text{ such that } f(\bar{d}) = f(\bar{a})$$

(see, e.g., [11, Proposition 1.2.5]).

[3] If a primitive positive formula contains a conjunct $x = y$, we obtain an equivalent formula δ' by deleting the conjunct $x = y$ and replacing each occurrence of the variable y by the variable x.

Building on canonical databases, for every d-dimensional primitive positive definition Δ of σ in τ we associate a mapping Γ_Δ from σ-structures to τ-structures as follows. Given a σ-structure $\mathbb{B}$ the *gadget replacement* $\Gamma_\phi(\mathbb{B})$ is the τ-structure obtained from $\mathbb{B}$ where

- for each vertex $b \in B$ we introduce d vertices $b_1, \dots, b_d$, and
- for each $R \in \sigma$ of arity r, and every tuple $(b^1, \dots, b^r) \in R^{\mathbb{B}}$ we introduce a fresh copy of $\mathbb{D}_{\delta_R}(\bar{d}^1, \dots, \bar{d}^r)$ and we identify each d-tuple $\bar{d}^i$ with $(b^i_1, \dots, b^i_d)$.

Going back to our on going example $\Delta := \{\delta_E\}$ and taking $\mathbb{B} := \mathbb{K}_2$, the gadget replacement $\Gamma_\Delta(\mathbb{B})$ is isomorphic to the directed 6-cycle $\vec{\mathbb{C}}_6$ (see Figure 2 for a depiction).

It is straightforward to observe that, in general, for every primitive positive definition Δ, the gadget replacement $\Gamma_\Delta(\mathbb{B})$ can be constructed in logarithmic space from $\mathbb{B}$. This fact and the following observation can be used to prove Lemma 10.

Observation 11 (Observation 4.4 in [33])**.** *The following statement holds for every primitive positive formula ϕ, and every pair of digraphs (structures) $\mathbb{A}$ and $\mathbb{A}'$*

$$\mathbb{A}' \to \Pi_\Delta(\mathbb{A}) \textit{ if and only if } \Gamma_\Delta(\mathbb{A}') \to \mathbb{A}.$$

Log-space reductions and rpp-constructions

We say that a (not necessarily finite) restricted CSP template $(\mathbb{A}, \mathbb{B})$ *rpp-constructs* a restricted CSP template $(\mathbb{A}', \mathbb{B}')$ if there is a primitive positive definition Δ such that

$$(\Pi_\Delta(\mathbb{A}) \times \mathbb{B}') \leftrightarrow (\mathbb{A}' \times \mathbb{B}') \text{ and } \mathbb{B}' \to \Pi_\Delta(\mathbb{B}).$$

Remark 12. Every primitive positive formula is satisfiable in some finite structure, e.g., in the loop $\mathbb{L}$. This implies that for every primitive positive definition of σ in τ the pp-power $\Pi_\Delta(\mathbb{L}_\tau)$ of the (τ-)loop is homomorphically equivalent to the (σ-)loop $\mathbb{L}_\sigma$. In particular, this implies that the following statement are equivalent for every τ-structure $\mathbb{A}$ and every σ-structure $\mathbb{A}'$.

- $\mathbb{A}$ pp-constructs $\mathbb{A}'$.
- $(\mathbb{A}, \mathbb{L}_\tau)$ rpp-constructs $(\mathbb{A}', \mathbb{L}_\sigma)$.
- $(\mathbb{A}, \mathbb{L}_\tau)$ rpp-constructs $(\mathbb{A}', \mathbb{B}')$ for ever structure $\mathbb{B}'$.

It is well-known that pp-constructions compose, and building on this fact, it is straightforward to observe that rpp-constructions compose.

Lemma 13. *Consider three restricted CSP templates* $(\mathbb{A}_1, \mathbb{B}_1)$*,* $(\mathbb{A}_2, \mathbb{B}_2)$*, and* $(\mathbb{A}_3, \mathbb{B}_3)$*. If* $(\mathbb{A}_1, \mathbb{B}_1)$ *rpp-constructs* $(\mathbb{A}_2, \mathbb{A}_2)$*, and* $(\mathbb{A}_2, \mathbb{B}_2)$ *rpp-constructs* $(\mathbb{A}_3, \mathbb{B}_3)$*, then* $(\mathbb{A}_1, \mathbb{B}_1)$ *rpp-constructs* $(\mathbb{A}_3, \mathbb{B}_3)$*.*

Proof. For $i \in \{1, 2\}$, let Δ_i be a primitive positive definition witnessing that $(\mathbb{A}_i, \mathbb{B}_i)$ rpp-constructs $(\mathbb{A}_{i+1}, \mathbb{B}_{i+1})$. Let Δ_3 by a primitive positive definition such that $\Pi_{\Delta_3} = \Pi_{\Delta_2} \circ \Pi_{\Delta_1}$ (Lemma 9). It is straightforward to observe that $\mathbb{B}_3 \to \Pi_{\Delta_3}(\mathbb{B}_1)$: since $\mathbb{B}_3 \to \Pi_{\Delta_2}(\mathbb{B}_2)$, and $\mathbb{B}_2 \to \Pi_1(\mathbb{B}_1)$, it follows (via Remark 8) that

$$\mathbb{B}_3 \to \Pi_{\Delta_2}(\mathbb{B}_2) \to \Pi_{\Delta_2}(\Pi_{\Delta_1}(\mathbb{B}_1)) = \Pi_{\Delta_3}(\mathbb{B}_1).$$

It is also not hard to notice that $\Pi_{\Delta_3}(\mathbb{A}_1) \times \mathbb{B}_3$ is homomorphically equivalent to $\mathbb{A}_3 \times \mathbb{B}_3$: by the choice of Δ_3 we know that $\Pi_{\Delta_3}(\mathbb{A}_1) = \Pi_{\Delta_2}(\Pi_{\Delta_1}(\mathbb{A}_1)) = \Pi_{\Delta_2}(\mathbb{A}_2)$, and since $\Pi_{\Delta_2}(\mathbb{A}_2) \times \mathbb{B}_3$ is homomorphically equivalent to $\mathbb{A}_3 \times \mathbb{B}_3$, we conclude that

$$(\Pi_{\Delta_3}(\mathbb{A}_1) \times \mathbb{B}_3) \leftrightarrow (\mathbb{A}_3 \times \mathbb{B}_3) \text{ and } \mathbb{B}_3 \to \Pi_{\Delta_3}(\mathbb{B}_1).$$

Therefore, Δ_3 is a primitive positive definition witnessing that $(\mathbb{A}_1, \mathbb{B}_1)$ rpp-constructs $(\mathbb{A}_3, \mathbb{B}_3)$. □

Lemma 14. *Consider two (not necessarily finite) restricted CSP templates* $(\mathbb{A}, \mathbb{B})$ *and* $(\mathbb{A}', \mathbb{B}')$*. If* $(\mathbb{A}, \mathbb{B})$ *rpp-constructs* $(\mathbb{A}', \mathbb{B}')$*, then there is a log-space reduction from* $\mathrm{RCSP}(\mathbb{A}', \mathbb{B}')$ *to* $\mathrm{RCSP}(\mathbb{A}, \mathbb{B})$*.*

Proof. Let Δ be a primitive positive definition witnessing that $(\mathbb{A}, \mathbb{B})$ rpp-constructs $(\mathbb{A}', \mathbb{B}')$, and let $\mathbb{C}$ be an instance to $\mathrm{RCSP}(\mathbb{A}', \mathbb{B}')$. Since $\mathbb{C} \to \mathbb{B}'$ and $\mathbb{B}' \to \Pi_\Delta(\mathbb{B})$, we know that $\mathbb{C} \to \Pi_\Delta(\mathbb{B})$. Hence, by Observation 11, it follows that $\Gamma_\Delta(\mathbb{C}) \to \mathbb{B}$, so $\Gamma_\Delta(\mathbb{C})$ is a valid instance to $\mathrm{RCSP}(\mathbb{A}, \mathbb{B})$. Now we show that $\mathbb{C}$ is a yes-instance to $(\mathbb{A}', \mathbb{B}')$ if and only if $\Gamma_\Delta(\mathbb{C})$ is a yes-instance to $\mathrm{RCSP}(\mathbb{A}, \mathbb{B})$. Since $\mathbb{C} \to \mathbb{B}$, we know that $\mathbb{C} \to \mathbb{A}'$ if and only if $\mathbb{C} \to \mathbb{A}' \times \mathbb{B}'$, which in turn is the case if and only if $\mathbb{C} \to \Pi_\Delta(\mathbb{A}) \times \mathbb{B}'$. Again, using our assumption that $\mathbb{C} \to \mathbb{B}'$ we see that $\mathbb{C} \to \Pi_\Delta(\mathbb{A}) \times \mathbb{B}'$ if and only if $\mathbb{C} \to \Pi_\Delta(\mathbb{A})$. Now, by Observation 11, $\mathbb{C} \to \Pi_\Delta(\mathbb{A})$ if and only if $\Gamma_\Delta(\mathbb{C}) \to \mathbb{A}$, so putting together all these equivalences we conclude that $\mathbb{C} \to \mathbb{A}'$ if and only if $\Gamma_\Delta(\mathbb{C}) \to \mathbb{A}$. Therefore, $\mathrm{RCSP}(\mathbb{A}', \mathbb{B}')$ reduces in logarithmic space to $\mathrm{RCSP}(\mathbb{A}, \mathbb{B})$. □

Example 15. Observe that the restricted CSP template $(\mathbb{C}_5, \mathbb{K}_3)$ rpp-constructs the template $(\mathbb{K}_5, \mathbb{L})$ via the formula

$$\delta_E(x, y) := \exists z_1, z_2.\ E(x, z_1) \wedge E(z_1, z_2) \wedge E(z_2, y).$$

Indeed, we already noticed that the pp power $\Pi_\Delta(\mathbb{C}_5)$ is isomorphic to $\mathbb{K}_5$. It is also not hard to observe that the pp power $\Pi_\Delta(\mathbb{K}_5)$ is homomorphically equivalent to the loop $\mathbb{L}$. Hence, by Lemma 14 we conclude that $\mathrm{RCSP}(\mathbb{C}_5, \mathbb{K}_3)$ is NP-hard, i.e., deciding if a 3-colourable graph $\mathbb{G}$ admits a homomorphism to $\mathbb{C}_5$ is NP-hard.

4 Finite domain restrictions

In this section we show that for every restricted CSP template $(\mathbb{A}, \mathbb{B})$ with finite restriction, there is a structure $\mathbb{C}$ such that $\mathrm{RCSP}(\mathbb{A}, \mathbb{B})$ and $\mathrm{CSP}(\mathbb{C})$ are log-space equivalent. Moreover, if $\mathbb{A}$ is also a finite structure, then $\mathbb{C}$ is a finite structure. It thus follows from the finite domain dichotomy [18, 39] that restricted CSPs with finite domain and finite restriction, have a P versus NP-complete dichotomy.

4.1 Exponential structures

Given a pair of τ-structure $\mathbb{A}$ and $\mathbb{B}$ the *exponential* $\mathbb{A}^\mathbb{B}$ is the τ-structure

- with vertex set all functions $f \colon B \to A$, and
- for each $R \in \tau$ of arity r there is an r-tuple $(f_1, \ldots, f_r)$ belongs to the interpretation of R in $\mathbb{A}^\mathbb{B}$ if and only if $(f_1(b_1), \ldots, f_r(b_r)) \in R^\mathbb{A}$ whenever $(b_1, \ldots, b_r) \in R^\mathbb{B}$.

In Figure 3 we depict an exponential digraph construction. This image is also found in [31], were the reader can also find further discussion and properties of exponential digraphs. For this paper we state the following properties of general exponential structures.

Lemma 16. *The following statements hold for all τ-structures $\mathbb{A}, \mathbb{B}$ and $\mathbb{C}$.*

- $\mathbb{A}^{\mathbb{L}}$ *is isomorphic to* $\mathbb{A}$.
- $\mathbb{C} \to \mathbb{A}^{\mathbb{B}}$ *if and only if* $\mathbb{C} \times \mathbb{B} \to \mathbb{A}$.

Proof. The first statement follows by noticing that the mapping $a \mapsto f_a$ where $f_a \colon L \to A$ is the function mapping the unique element $l \in L$ to a, defines an isomorphism from $\mathbb{A}$ to $\mathbb{A}^{\mathbb{L}}$. For the second statement we refer the reader to [24, Corollary 1.5.12]. □

Corollary 17. *For every pair of (possibly infinite) τ-structures $\mathbb{A}$ and $\mathbb{B}$, the restricted CSP template $(\mathbb{A}^{\mathbb{B}}, \mathbb{L})$ rpp-constructs the template* $\mathrm{RCSP}(\mathbb{A}, \mathbb{B})$.

Proof. Let Δ be the trivial primitive positive definition of τ in τ, i.e., for each $R \in \tau$ or arity r let $\delta_R(x_1, \dots, x_r) := R(x_1, \dots, x_r)$. So for every structure $\mathbb{A}$ the equality $\Pi_\Delta(\mathbb{A}) = \mathbb{A}$ holds. Clearly, $\mathbb{B} \to \mathbb{L} = \Pi_\Delta(\mathbb{L})$, and notice that $\mathbb{A}^{\mathbb{B}} \times \mathbb{B}$ is homomorphically equivalent to $\mathbb{A} \times \mathbb{B}$: since $\mathbb{A}^{\mathbb{B}} \to \mathbb{A}^{\mathbb{B}}$, it follows from the second item in Lemma 16 that $\mathbb{A}^{\mathbb{B}} \times \mathbb{B} \to \mathbb{A}$, and clearly $\mathbb{A}^{\mathbb{B}} \times \mathbb{B} \to \mathbb{B}$, so $\mathbb{A}^{\mathbb{B}} \times \mathbb{B} \to \mathbb{A} \times \mathbb{B}$; conversely, $(\mathbb{A} \times \mathbb{B}) \times \mathbb{B} \to \mathbb{A}$, so by Lemma 16, $\mathbb{A} \times \mathbb{B} \to \mathbb{A}^{\mathbb{B}}$, and since $\mathbb{A} \times \mathbb{B} \to \mathbb{B}$, it follows that $\mathbb{A} \times \mathbb{B} \to \mathbb{A}^{\mathbb{B}} \times \mathbb{B}$. Therefore, Δ is a primitive positive definition witnessing that the claim of the lemma holds. □

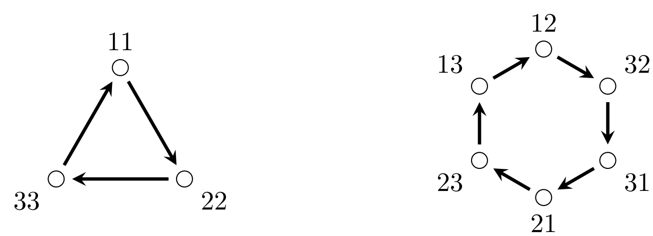


Figure 3: The exponential $\vec{\mathbb{C}}_3^{\mathbb{K}_2}$ where a label ij represents the function $f \colon \{1, 2\} \to \{1, 2, 3\}$ defined by $1 \mapsto i$ and $2 \mapsto j$.

Lemma 18. *Let τ be a finite relational signature, and $\mathbb{A}$ be a (possible infinite) τ-structure. If $\mathbb{B}$ is a finite τ-structure, then the restricted* RCSP *template* $(\mathbb{A}, \mathbb{B})$ *rpp-constructs the* RCSP *template* $(\mathbb{A}^{\mathbb{B}}, \mathbb{L})$.

Proof. We first show that $(\mathbb{A}, \mathbb{B})$ rpp-constructs $(\mathbb{A}^{\mathbb{B}}, \mathbb{L})$, and we consider the case of digraphs but it naturally generalizes to τ-structures for finite τ. Let $B = \{b_1, \dots, b_m\}$ and consider the following m-dimensional primitive positive definition $\Delta := \{\delta_E\}$, where

$$\delta_E(x_1, \dots, x_m, y_1, \dots, y_m) := \bigwedge_{(b_i, b_j) \in E^{\mathbb{B}}} E(x_i, y_j).$$

Notice that for any structure $\mathbb{A}'$ the pp-power Π_Δ is isomorphic to $\mathbb{A}'^{\mathbb{B}}$. Indeed, identify a tuple $(a_1, \dots, a_m)$ of A with the function defined by $b_i \mapsto a_i$, it now follows from the definition of the exponential structure and of δ_E that this mapping is an isomorphism. In particular, notice that $\Pi_\Delta(\mathbb{B}) = \mathbb{B}^{\mathbb{B}}$ contains the loop $\mathbb{L}$ (the identity function $I\colon B \to B$ is a loop on $\mathbb{B}^{\mathbb{B}}$). Hence,

$$\Pi_\Delta(\mathbb{A}) \cong \mathbb{A}^{\mathbb{B}} \text{ and } \mathbb{L} \to \Pi_\Delta(\mathbb{B}),$$

i.e., Δ is a witness to the fact that $(\mathbb{A}, \mathbb{B})$ rpp-constructs $(\mathbb{A}^{\mathbb{B}}, \mathbb{L})$. □

We discuss two applications of this lemma. The first one being that rpp-constructions between RCSP templates with finite restrictions are captured by (standard) pp-constructions.

Theorem 19. *The following statements are equivalent for every pair of restricted* CSP *templates* $(\mathbb{A}_1, \mathbb{B}_1)$ *and* $(\mathbb{A}_2, \mathbb{B}_2)$ *with finite restrictions.*

- $(\mathbb{A}_1, \mathbb{B}_1)$ *rpp-constructs* $(\mathbb{A}_2, \mathbb{B}_2)$ *and,*
- $\mathbb{A}_1^{\mathbb{B}_1}$ *pp-constructs* $\mathbb{A}_2^{\mathbb{B}_2}$.

In particular, $(\mathbb{A}_1, \mathbb{B}_1)$ *and* $(\mathbb{A}_1^{\mathbb{B}_1}, \mathbb{L})$ *are mutually rpp-constructible (and so* $\mathrm{RCSP}(\mathbb{A}_1, \mathbb{B}_1)$ *and* $\mathrm{CSP}(\mathbb{A}_1^{\mathbb{B}_2})$ *are log-space Turing-equivalent).*

Proof. Suppose that $\mathbb{A}_1^{\mathbb{B}_1}$ pp-constructs $\mathbb{A}_2^{\mathbb{B}_2}$, so by Remark 12 the template $(\mathbb{A}_1^{\mathbb{B}_1}, \mathbb{L})$ rpp-constructs $(\mathbb{A}_2^{\mathbb{B}_2}, \mathbb{L})$. By Lemma 18, $(\mathbb{A}_1, \mathbb{B}_1)$ rpp-constructs $(\mathbb{A}_1^{\mathbb{B}_1}, \mathbb{L})$, and by Corollary 17 $(\mathbb{A}_2^{\mathbb{B}_2}, \mathbb{L})$ rpp-constructs $(\mathbb{A}_2, \mathbb{B}_2)$. Hence, by composing rpp-constructions (Lemma 13), we conclude that $(\mathbb{A}_1, \mathbb{B}_1)$ rpp-constructs $(\mathbb{A}_2, \mathbb{B}_2)$. The converse implication follows with similar arguments: $(\mathbb{A}_2, \mathbb{B}_2)$ rpp-constructs $(\mathbb{A}_2^{\mathbb{B}_2}, \mathbb{L})$ (Lemma 18), and $(\mathbb{A}_1^{\mathbb{B}_1}, \mathbb{L})$ rpp-constructs $(\mathbb{A}_1, \mathbb{B}_1)$ (Corollary 17); by composing rpp-constructions we see that $(\mathbb{A}_1^{\mathbb{B}_1}, \mathbb{L})$ rpp-constructs $(\mathbb{A}_2^{\mathbb{B}_2}, \mathbb{L})$, and it thus follows that $\mathbb{A}_1^{\mathbb{B}_1}$ pp-constructs $\mathbb{A}_2^{\mathbb{B}_2}$ (Remark 12). □

The second application of Lemma 18 is the following statement which is analogous to the case of CSPs and pp-constructions (see, e.g., [11, Theorem 3.2.2]).

Theorem 20. *The following statements are equivalent for every (possibly infinite) structure* $\mathbb{A}$ *and a finite structure* $\mathbb{B}$ *with a finite signature.*

- *The restricted* CSP *template* $(\mathbb{A}, \mathbb{B})$ *rpp-constructs* $(\mathbb{K}_3, \mathbb{L})$.
- *The restricted* CSP *template* $(\mathbb{A}, \mathbb{B})$ *rpp-constructs every restricted* CSP *template* $(\mathbb{A}', \mathbb{B}')$ *where* $\mathbb{A}'$ *is a finite structure (and* $\mathbb{B}'$ *a possibly infinite structure).*
- *The structure* $\mathbb{A}^{\mathbb{B}}$ *pp-constructs* $\mathbb{K}_3$.

If any of these equivalent statement hold, then $\mathrm{RCSP}(\mathbb{A}, \mathbb{B})$ *is* NP*-hard.*

Proof. By the first item of Lemma 16 we know that $\mathbb{K}_3^{\mathbb{L}}$ is isomorphic to $\mathbb{K}_3$, and so $\mathbb{A}^{\mathbb{B}}$ pp-constructs $\mathbb{K}_3$ if and only if $\mathbb{A}^{\mathbb{B}}$ pp-constructs $\mathbb{K}_3^{\mathbb{L}}$. Hence, the equivalence between the first and third itemized statement follows from Theorem 19. Also, second statement clearly implies the third one. Finally, we show that the first item implies the second one. It is well-known that that $\mathbb{K}_3$ pp-constructs every finite structure $\mathbb{A}'$ (see, e.g., [11, Corollary 3.2.1]). So, by Remark 12 $(\mathbb{K}_3, \mathbb{L})$ rpp-constructs every finite domain RCSP template $(\mathbb{A}', \mathbb{B}')$ (where $\mathbb{B}'$ is a possibly infinite structure). Again, by composing rpp-constructions, we conclude that $(\mathbb{A}, \mathbb{B})$ rpp-constructs $(\mathbb{A}', \mathbb{B}')$. The equivalence between the three itemized statement is now settled. The final statement holds via Lemma 10 because $\mathrm{RCSP}(\mathbb{K}_3, \mathbb{L}) = \mathrm{CSP}(\mathbb{K}_3)$ is NP-hard. □

4.2 A dichotomy for finite domain RCSPs with finite restrictions

The dichotomy for finite domain CSPs asserts that if $\mathbb{A}$ is a finite structure, then either $\mathbb{A}$ pp-constructs $\mathbb{K}_3$, and in this case CSP($\mathbb{A}$) is NP-complete; or otherwise, CSP($\mathbb{A}$) is polynomial-time solvable. We apply the results of this section to obtain an analogous (actually, equivalent) statement for finite domain CSPs with finite restrictions.

Theorem 21. *For every pair of finite structures $\mathbb{A}, \mathbb{B}$ one of the following statement hold.*

- *Either* $(\mathbb{A}, \mathbb{B})$ *rpp-constructs* $(\mathbb{K}_3, \mathbb{L})$ *(and consequently,* RCSP$(\mathbb{A}, \mathbb{B})$ *is* NP*-hard), or*
- RCSP$(\mathbb{A}, \mathbb{B})$ *is polynomial-time solvable.*

Proof. Suppose that the first itemized statement does not hold, and assume P $\neq$ NP (otherwise, CSP($\mathbb{A}$) is in P, and thus RCSP$(\mathbb{A}, \mathbb{B})$ is polynomial-time solvable). Since $(\mathbb{A}, \mathbb{B})$ does not rpp-construct $(\mathbb{K}_3, \mathbb{L})$, it follows from Theorem 20 that $\mathbb{A}^{\mathbb{B}}$ does not pp-construct $\mathbb{K}_3$. Hence, the finite domain dichotomy implies that CSP($\mathbb{A}^{\mathbb{B}}$) is in P. By the "in particular" statement of Theorem 19 we know that CSP($\mathbb{A}^{\mathbb{B}}$) and RCSP$(\mathbb{A}, \mathbb{B})$ are polynomial-time equivalent, and we thus conclude that RCSP$(\mathbb{A}, \mathbb{B})$ is polynomial-time solvable. □

4.3 Finite-domain restrictions up to high girth

Given a structure $\mathbb{B}$ and a positive integer ℓ we denote by CSP$_{>\ell}(\mathbb{B})$ the subclass of CSP($\mathbb{B}$) consisting of structure $\mathbb{A} \in$ CSP($\mathbb{B}$) with girth strictly larger than ℓ. The CSP of a structure $\mathbb{B}$ is *finite-domain up to hight girth* if there is a positive integer ℓ and a finite structure $\mathbb{B}'$ such that CSP$_{>\ell}(\mathbb{B})$ = CSP$_{>\ell}(\mathbb{B}')$. It was proved in [28] that for every such structure $\mathbb{B}$, there is a finite structure $\mathbb{B}_S$ (unique up to homomorphic equivalence) such that for every finite structure $\mathbb{A}$ there is a homomorphism $\mathbb{B} \to \mathbb{A}$ if and only if $\mathbb{B}_S \to \mathbb{A}$. We call $\mathbb{B}_S$ the *smallest finite factor* of $\mathbb{B}$ — notice that in particular, $\mathbb{B} \to \mathbb{B}_S$. Moreover, if $\mathbb{B}$ is finite domain up to high girth and $\mathbb{B}_S$ is its smallest finite factor, then there is a positive integer ℓ such that CSP$_{>\ell}(\mathbb{B})$ = CSP$_{>\ell}(\mathbb{B}_S)$ [28, Corollary 10].

Lemma 22. *Let $\mathbb{B}$ be a structure such that* CSP($\mathbb{B}$) *is finite-domain up to high girth. If $\mathbb{B}_S$ is the smallest finite factor of $\mathbb{B}$, then* RCSP$(\mathbb{A}, \mathbb{B})$ *and* RCSP$(\mathbb{A}, \mathbb{B}_S)$ *are polynomial-time equivalent.*

Proof. Since $\mathbb{B} \to \mathbb{B}_S$, it follows that RCSP$(\mathbb{A}, \mathbb{B}_S)$ is at least as hard as RCSP$(\mathbb{A}, \mathbb{B})$. Fo the converse reduction, let $k := \max\{|A|, |B_S|\}$, and ℓ a positive integer such that CSP$_{>\ell}(\mathbb{B})$ = CSP$_{>\ell}(\mathbb{B}_S)$. let $\mathbb{C}$ be an input to RCSP$(\mathbb{A}, \mathbb{B}_S)$. By the Sparse Incomparability Lemma (applied to ℓ and k), there is a structure $\mathbb{C}'$ of girth larger than ℓ, and such that $\mathbb{C}' \to \mathbb{B}_S$, and $\mathbb{C} \to \mathbb{A}$ if and only if $\mathbb{C}' \to \mathbb{A}$. Since $\mathbb{C}'$ has girth larger than ℓ and CSP$_{>\ell}(\mathbb{B})$ = CSP$_{>\ell}(\mathbb{B}_S)$, it follows that $\mathbb{C}' \to \mathbb{B}$. Hence $\mathbb{C}'$ is a valid input to RCSP$(\mathbb{A}, \mathbb{B})$, and since $\mathbb{C}'$ is constructible in polynomial-time from $\mathbb{C}$ (Theorem 6) and $\mathbb{C}' \to \mathbb{A}$ if and only if $\mathbb{C} \to \mathbb{A}$, we conclude that RCSP$(\mathbb{A}, \mathbb{B}_S)$ reduces in polynomial-time to RCSP$(\mathbb{A}, \mathbb{B})$. □

4.4 A dichotomy for finite domain RCSPs with GMSNP restrictions

Forbidden pattern problems (for graphs) are parametrized by a finite set $\mathcal{F}$ of vertex and edge coloured connected graphs, and the task is to decide is an input graph $\mathbb{G}$ admits a colouring $\mathbb{G}'$ (with the same colours used in $\mathcal{F}$) such that $\mathbb{G}' \in$ Forb($\mathcal{F}$), i.e., there is no homomorphism $\mathbb{F} \to \mathbb{G}'$ for any $\mathbb{F} \in \mathcal{F}$. A standard example of the problem of deciding if an input graph $\mathbb{G}$ admits a 2-edge-colouring with no monochromatic triangles.

There is a natural fragment of existential second order logic called *guarded monotone strict NP* (GMSNP) such that CSPs expressible in GMSNP capture forbidden pattern problems. We refer the reader to [4] for further background on GMSNP.

Lemma 12 in [28] shows that every CSP expressible in GMSNP is finite-domain up to high girth. The following statement is an immediate consequence of this fact, and of Lemma 22.

Theorem 23. *For every finite structure* $\mathbb{A}$*, and every structure* $\mathbb{B}$ *such that* $\mathrm{CSP}(\mathbb{B})$ *is expressible in* GMSNP *one of the following statement holds.*

- *Either* $(\mathbb{A}, \mathbb{B}_S)$ *rpp-constructs* $(\mathbb{K}_3, \mathbb{L})$*, where* $\mathbb{B}_S$ *is the smallest finite factor of* $\mathbb{B}$*, (and consequently,* $\mathrm{RCSP}(\mathbb{A}, \mathbb{B})$ *is* NP*-hard), or*
- $\mathrm{RCSP}(\mathbb{A}, \mathbb{B})$ *is polynomial-time solvable.*

4.5 The tractability conjecture and RCSPs

The tractability conjecture is a generalization of the Feder-Vardi conjecture to a broad class of "well-behaved" infinite structures, namely, to *reducts* of *finitely bounded homogeneous* structure. A structure $\mathbb{B}$ is homogeneous if for every isomorphism $f\colon \mathbb{A} \to \mathbb{A}'$ between finite substructures of $\mathbb{B}$, there is an automorphism $f'\colon \mathbb{B} \to \mathbb{B}$ that extends f, i.e., $f'(a) = f(a)$ for every $a \in A$. A structure $\mathbb{B}$ is called finitely bounded if there exists a finite set $\mathcal{F}$ of finite structures such that a finite structure $\mathbb{A}$ embeds into $\mathbb{B}$ if and only if no structure from $\mathcal{F}$ embeds into $\mathbb{A}$. A reduct of a structure $\mathbb{B}$ is a structure $\mathbb{A}$ obtained from $\mathbb{B}$ by forgetting some relations.

Conjecture 1 (Conjecture 3.7.1 in [11])**.** *Let* $\mathbb{A}$ *be a reduct of a finitely bounded homogeneous structure. If* $\mathbb{A}$ *does not pp-construct* $\mathbb{K}_3$*, then* $\mathrm{CSP}(\mathbb{A})$ *is polynomial-time solvable.*

This is a wide-open conjecture yielding a very active research line (e.g., [6, 9–13, 36]). Here we show that this conjecture is equivalent to the following conjecture for restricted CSPs templates with finite restriction and whose domain is a reduct of a finitely bounded homogeneous structure.

Conjecture 2. *Let* $\mathbb{A}$ *be a reduct of a finitely bounded homogeneous structure, and* $\mathbb{B}$ *be a finite structure. If the restricted* CSP*-template does not rpp-construct* $(\mathbb{K}_3, \mathbb{L})$*, then* $\mathrm{RCSP}(\mathbb{A}, \mathbb{B})$ *is polynomial-time solvable.*

Theorem 24. *Conjecture 1 and Conjecture 2 are equivalent.*

Proof. The fact that Conjecture 2 implies Conjecture 1 follows from Remark 12: if $\mathbb{A}$ does not pp-construct $\mathbb{K}_3$, then $(\mathbb{A}, \mathbb{L})$ does not rpp-constructs $(\mathbb{K}_3, \mathbb{L})$; so, Conjecture 2 implies that $\mathrm{RCSP}(\mathbb{A}, \mathbb{L})$ is polynomial-time solvable, and thus, $\mathrm{CSP}(\mathbb{A}) = \mathrm{RCSP}(\mathbb{A}, \mathbb{L})$ is in P. For the converse implication, suppose that $(\mathbb{A}, \mathbb{B})$ does not rpp-construct $(\mathbb{K}_3, \mathbb{L})$. By the equivalence between the second and third statement of Theorem 20, it must be the case that $\mathbb{A}^{\mathbb{B}}$ does not pp-construct $\mathbb{K}_3$. Moreover, in the proof of Lemma 18 we propose a primitive positive definition Δ such that $\Pi_\Delta(\mathbb{A}) \cong \mathbb{A}^{\mathbb{B}}$. Since reducts of finitely bounded homogeneous structures are closed under pp-powers (because reducts of finitely bounded homogeneous structures are closed under products and pp-definitions), it follows that $\mathbb{A}^{\mathbb{B}}$ is a reduct of a finitely bounded homogeneous structure, so Conjecture 1 implies that $\mathrm{CSP}(\mathbb{A}^{\mathbb{B}})$ is in P. This implies that $\mathrm{RCSP}(\mathbb{A}, \mathbb{B})$ is polynomial-time solvable (see, e.g., the last sentence of Theorem 19). $\square$

4.6 A small remark regarding PCSPs

Theorem 19 asserts in particular that $\mathrm{RCSP}(\mathbb{A},\mathbb{B})$ is log-space equivalent to $\mathrm{CSP}(\mathbb{A}^{\mathbb{B}})$. The following statement strengthens this connection by showing that these two problems are further log-space equivalent to $\mathrm{PCSP}(\mathbb{A},\mathbb{A}^{\mathbb{B}})$.

Lemma 25. *For every possibly infinite structure $\mathbb{A}$ and a (possibly infinite) structure and $\mathbb{B}$ a finite structure, the following problems are polynomial-time equivalent.*

- *the constraint satisfaction problem* $\mathrm{CSP}(\mathbb{A}^{\mathbb{B}})$*,*
- *the promise constraint satisfaction problem* $\mathrm{PCSP}(\mathbb{A},\mathbb{A}^{\mathbb{B}})$*, and*
- *the restricted constraint satisfaction problem* $\mathrm{RCSP}(\mathbb{A},\mathbb{B})$*.*

Proof. The first and third itemized problems are polynomial-time equivalent by Theorem 19. The promise CSP form the second item clearly reduces to the CSP in the first item. Now we see that the RCSP reduces to the PCSP. Let $\mathbb{C}$ be an input to $\mathrm{RCSP}(\mathbb{A},\mathbb{B})$, so $\mathbb{C}\to\mathbb{B}$. We claim that $\mathbb{C}$ is a valid input to $\mathrm{PCSP}(\mathbb{A},\mathbb{A}^{\mathbb{B}})$: if $\mathbb{C}\to\mathbb{A}$, then $\mathbb{C}$ is clearly a valid input to the PCSP; otherwise, notice that $\mathbb{C}\times\mathbb{B}$ is homomorphically equivalent to $\mathbb{C}$ because $\mathbb{C}\to\mathbb{B}$, and so if $\mathbb{C}\not\to\mathbb{A}$, then $\mathbb{C}\times\mathbb{B}\leftrightarrow\mathbb{C}\not\to\mathbb{A}^{\mathbb{B}}$, and so by Lemma 16, we see that $\mathbb{C}\not\to\mathbb{A}^{\mathbb{B}}$. Therefore, $\mathbb{C}$ is a valid input to $\mathrm{PCSP}(\mathbb{A},\mathbb{B})$, and clearly the reduction is complete and correct. $\square$

Corollary 26. *For every non-bipartite graph $\mathbb{H}$ and every finite digraph $\mathbb{D}$ one of the following holds.*

- *Either $\mathbb{D}\to\mathbb{H}$, and in this case* $\mathrm{PCSP}(\mathbb{H},\mathbb{H}^{\mathbb{D}})$ *is polynomial-time solvable,*
- *or $\mathbb{D}\not\to\mathbb{H}$, and in this case* $\mathrm{PCSP}(\mathbb{H},\mathbb{H}^{\mathbb{D}})$ *is* NP*-hard.*

Proof. The first item holds because if $\mathbb{D}\to\mathbb{H}$, then $\mathbb{H}^{\mathbb{D}}$ has a loop, and hence $\mathrm{PCSP}(\mathbb{H},\mathbb{H}^{\mathbb{D}})$ is trivial. Now suppose that $\mathbb{D}\not\to\mathbb{H}$. In this case, Theorem 3 in [16] implies that $\mathrm{RCSP}(\mathbb{H},\mathbb{D})$ is NP-hard, and we conclude that $\mathrm{PCSP}(\mathbb{H},\mathbb{H}^{\mathbb{D}})$ is NP-hard (Lemma 25). $\square$

5 Homomorphism-free restrictions

As mentioned above, the reader familiar with monotone monadic strict NP (MMSNP) can notice that digraph (finite domain) CSPs with input restriction to $\mathcal{F}$-homomorphism-free digraphs (structures) can be solved by infinite domain CSPs expressible in MMSNP. Thus, these problems exhibit a P versus NP-complete dichotomy (see, e.g., [14, 23]). Feder and Vardi reduce problems in MMSNP to finite domain CSPs. Their reduction changes the input signature, and then uses the Sparse Incomparability Lemma. Here, we prove the following lemma which allows us to preserve the signature by using duality pairs, and then, as Feder and Vardi, we proceed via Sparse Incomparability.

Lemma 27. *For every finite set of structures $\mathcal{F}$ and every finite digraph (structure) $\mathbb{B}$, there is a finite set of structures $\mathcal{C}$ such that the following problems are polynomial-time equivalent:*

- *deciding if an input structure $\mathbb{A}$ belongs to* $\mathrm{CSP}(\mathbb{C})$ *for some $\mathbb{C}\in\mathcal{C}$, and*
- $\mathrm{CSP}(\mathbb{B})$ *restricted to $\mathcal{F}$-homomorphism-free structures.*

Proof. We first consider the case when for every $\mathbb{F} \in \mathcal{F}$ there is no forest $\mathbb{T}$ such that $\mathbb{F} \to \mathbb{T}$. In this case, let $\mathcal{C} = \{\mathbb{B}\}$. Clearly, CSP($\mathbb{B}$) is at least as hard as CSP($\mathbb{B}$) restricted to $\mathcal{F}$-homomorphism-free inputs. For the converse reduction, let ℓ be the maximum number of vertices of a structure $\mathbb{F} \in \mathcal{F}$, and k the number of vertices of $\mathbb{B}$. On input $\mathbb{A}$ to CSP($\mathbb{B}$), let $\mathbb{A}'$ be the structure of girth larger than ℓ obtained via Theorem 6. In particular, $\mathbb{A}'$ can be constructed in polynomial-time from $\mathbb{A}$ and $\mathbb{A}' \to \mathbb{B}$ if and only if $\mathbb{A} \to \mathbb{B}$. Finally, since the girth of $\mathbb{A}'$ is strictly larger than the number of vertices of any structure $\mathbb{F} \in \mathcal{F}$, and no $\mathbb{F} \in \mathcal{F}$ maps to a forest, then $\mathbb{A}' \in \mathrm{Forb}(\mathcal{F})$. Hence, $\mathbb{A}'$ is a valid input to the second itemized problem, and thus CSP($\mathbb{B}$) is polynomial-time equivalent to CSP($\mathbb{B}$) restricted to $\mathcal{F}$-homomorphism-free structures.

Otherwise, let $\mathcal{T}$ be the non-empty set of forests $\mathbb{T}$ for which there is an $\mathbb{F} \in \mathcal{F}$ and a surjective homomorphism $f \colon \mathbb{F} \to \mathbb{T}$. By Theorem 5, there is a finite set of structures $\mathcal{D}$ such that $(\mathcal{T}, \mathcal{D})$ is a generalized duality pair. Let $\mathcal{C}$ be the set $\{\mathbb{B} \times \mathbb{D} \colon \mathbb{D} \in \mathcal{D}\}$. We now prove that the claim of this lemma holds for $\mathcal{C}$ and $\mathbb{B}$.

Consider an input structure $\mathbb{A}$ to the second itemized problem. Since $\mathbb{A} \in \mathrm{Forb}(\mathcal{F})$, it must also be the case that $\mathbb{A} \in \mathrm{Forb}(\mathcal{T})$, and so there is some $\mathbb{D} \in \mathcal{D}$ such that $\mathbb{A} \to \mathbb{D}$. Hence, $\mathbb{A} \to \mathbb{B}$ if and only if $\mathbb{A} \to \mathbb{D} \times \mathbb{B}$, and so $\mathbb{A} \to \mathbb{B}$ implies that $\mathbb{A} \in \mathrm{CSP}(\mathbb{C})$ for some $\mathbb{C} \in \mathcal{C}$. Conversely, if $\mathbb{A} \in \mathrm{CSP}(\mathbb{C})$ for some $\mathbb{C} \in \mathcal{C}$, then $\mathbb{A} \to \mathbb{B} \times \mathbb{D}$ for some $\mathbb{D} \in \mathcal{D}$, and therefore $\mathbb{A} \to \mathbb{B}$. We thus conclude that the problem in the second item reduces in polynomial time (via the trivial reduction) to deciding if an input structure $\mathbb{A}$ belongs to $\bigcup_{\mathbb{C} \in \mathcal{C}} \mathrm{CSP}(\mathbb{C})$.

For the reduction back, we consider two possible cases. First, suppose that $\mathbb{D} \to \mathbb{B}$ for every $\mathbb{D} \in \mathcal{D}$, and notice that in this case $\mathbb{B} \times \mathbb{D}$ is homomorphically equivalent to $\mathbb{D}$ for every $\mathbb{D} \in \mathcal{D}$. Hence, CSP($\mathbb{C}$) is polynomial-time solvable for every for every $\mathbb{C} \in \mathcal{C}$, and so the problem in the first item is polynomial-time solvable as well. Notice that if we show that in this case, CSP($\mathbb{B}$) restricted to $\mathcal{F}$-homomorphism-free structures is polynomial-time solvable, the both itemized problems are clearly polynomial-time equivalent. Similarly as above, if an input structure $\mathbb{A}$ to CSP($\mathbb{B}$) has no homomorphic image from any $\mathbb{F} \in \mathcal{F}$, then $\mathbb{A} \in \mathrm{Forb}(\mathcal{T})$ and so $\mathbb{A} \to \mathbb{D}_0$ for some $\mathbb{D}_0 \in \mathcal{D}$. Since $\mathbb{D} \to \mathbb{B}$ for every $\mathbb{D} \in \mathcal{D}$, we see that $\mathbb{A} \to \mathbb{D}_0 \to \mathbb{B}$, i.e., every input $\mathbb{A}$ to CSP($\mathbb{B}$) with no homomorphic image of any $\mathbb{F} \in \mathcal{F}$ is a yes-instance to CSP($\mathbb{B}$). Therefore, if $\mathbb{D} \to \mathbb{B}$ for every $\mathbb{D} \in \mathcal{D}$, both itemized problems are polynomial-time solvable and thus, polynomial-time equivalent.

Now, suppose that $\mathbb{D}_0 \not\to \mathbb{B}$ for some $\mathbb{D}_0 \in \mathcal{D}$, and recall that there is no forest $\mathbb{T} \in \mathcal{T}$ such that $\mathbb{T} \to \mathbb{D}_0$. With a similar trick as before, we can use the Sparse Incomparability Lemma to find a structure $\mathbb{D}_0'$ such that $\mathbb{D}_0' \not\to \mathbb{B}$ and $\mathbb{D}_0$ has no homomorphic image from any $\mathbb{F} \in \mathcal{F}$. The reduction from the first itemized problem for $\mathcal{C}$ to CSP($\mathbb{B}$) to structures with no homomorphic image of any $\mathbb{F} \in \mathcal{F}$ is now simple: on input $\mathbb{A}$ we check if $\mathbb{A} \in \mathrm{Forb}(\mathcal{T})$, if not, we return $\mathbb{D}_0'$, and if yes we return $\mathbb{A}'$ obtained from the Sparse Incomparability Lemma applied to $\mathbb{A}$, to ℓ the maximum number of vertices of a digraph in $\mathcal{F}$, and $k := |V(\mathbb{B})|$. The fact that this reduction is consistent and correct follows by the assumption that $\mathbb{D}_0' \not\to \mathbb{B}$, and with similar arguments as in the first paragraph, one can notice that both $\mathbb{A}'$ and $\mathbb{D}_0$ are valid inputs to the problem in the second item. $\square$

The following statement is an immediate consequence of Lemma 27 and finite domain CSP dichotomy [18, 38].

Theorem 28. *For every finite digraph (structure) $\mathbb{H}$ and every finite set of digraphs (structures) $\mathcal{F}$,* CSP($\mathbb{H}$) *restricted to $\mathcal{F}$-homomorphism-free digraphs (structures) is either in* P *or it is* NP-*complete.*

Proof. Let $\mathcal{C}$ be the finite set of structures from Lemma 27 applied to CSP($\mathbb{A}$) and to $\mathcal{F}$. If CSP($\mathbb{C}$) is polynomial-time solvable for every $\mathbb{C} \in \mathcal{C}$, then deciding if an input structure belongs to $\bigcup_{\mathbb{C}\in\mathcal{C}} \mathrm{CSP}(\mathbb{C})$ is polynomial-time solvable, and so $\mathrm{CSP}(\mathbb{B})$ is polynomial-time solvable for $\mathcal{F}$-homomorphism-free structures. Otherwise, it follows from the dichotomy for finite domain CSPs [18, 38] than $\mathrm{CSP}(\mathbb{C}_0)$ is NP-complete for some $\mathbb{C}_0 \in \mathcal{C}$. It straightforward to observe that in this case deciding if an input structure belongs to $\bigcup_{\mathbb{C}\in\mathcal{C}} \mathrm{CSP}(\mathbb{C})$ is NP-complete as well (see, e.g., [25, Theorem 25]), and hence $\mathrm{CSP}(\mathbb{B})$ is NP-hard even for $\mathcal{F}$-homomorphism-free structures. The claim now follows because $\mathrm{Forb}(\mathcal{F}) = \mathrm{CSP}(\mathbb{B})$. $\square$

6 Acyclic digraphs and bounded paths

In this section we apply our results above, and the theory of constraint satisfaction to obtain results in the context of $\mathcal{F}$-(subgraph)-free algorithmics. The main result of this section settles Question 3 (Theorem 34). Moreover, we show that if $\mathrm{CSP}(\mathbb{H})$ is NP-complete, then there is a positive integer N such that $\mathrm{CSP}(\mathbb{H})$ remains NP-complete for acyclic digraphs with no directed path on N vertices, and $\mathrm{CSP}(\mathbb{H})$ can be solved in polynomial time if the input is an acyclic digraph with no directed path on $N-1$ vertices. We begin by stating the following corollary of (the proof of) Lemma 27.

Corollary 29. *For every finite set of trees $\mathcal{F}$ with dual $\mathbb{D}$, and every digraph $\mathbb{H}$ the following problems are polynomial-time equivalent.*

- $\mathrm{RCSP}(\mathbb{H}, \mathbb{D})$.
- $\mathrm{CSP}(\mathbb{H} \times \mathbb{D})$.
- $\mathrm{CSP}(\mathbb{H})$ *restricted to $\mathcal{F}$-homomorphism-free digraphs.*

A *polymorphism* of a structure $\mathbb{A}$ is a homomorphism $f\colon \mathbb{A}^n \to \mathbb{A}$, and in this case we say that f is an *n-ary* polymorphism. Under composition, polymorphisms form an algebraic structure called a *clone*, and this structure captures the computational complexity of $\mathrm{CSP}(\mathbb{A})$ (see, e.g., [8]). A 4-ary polymorphism $f\colon \mathbb{A}^4 \to \mathbb{A}$ satisfies the *Sigger's identity* if

$$f(x_1, x_2, x_3, x_1) = f(x_2, x_1, x_2, x_3) \text{ for every } x_1, x_2, x_3, x_4 \in A.$$

The Sigger's identity is one of several identities that equivalently describe the tractability frontier for finite domain CSPs.

Theorem 30 (Equivalent to Theorem 1.4 in [38])**.** *For every finite structure $\mathbb{A}$ one of the following holds.*

- *Either $\mathbb{A}$ has a polymorphism $f\colon \mathbb{A}^4 \to \mathbb{A}$ that satisfies the Sigger's identity, and in this case* $\mathrm{CSP}(\mathbb{A})$ *is polynomial-time solvable, or*
- *otherwise, $\mathbb{A}$ pp-constructs $\mathbb{K}_3$, and in this case* $\mathrm{CSP}(\mathbb{A})$ *is* NP*-complete.*

Consider a finite digraph $\mathbb{H}$, and let k be a positive integer. For a given function $f\colon (H\times[k])^n \to (H \times [k])$ and for each $i \in [k]$ we define a function

$$f_i\colon H^n \to H \text{ by } (h_1, \dots, h_n) \mapsto \pi_H(f((h_1, i), \dots, (h_n, i)),$$

where π_H is the projection of $H \times [k]$ onto H.

Lemma 31. *Let $\mathbb{H}$ be a digraph, k a positive integer, and $f\colon (\mathbb{H}\times\mathbb{T}_k)^n \to (\mathbb{H}\times\mathbb{T}_k)$ and n-ary polymorphism of $\mathbb{H}\times\mathbb{T}_k$. If there are $i,j\in[k]$ with $i<j$ such that $f_i = f_j$, then f_i is an n-ary polymorphism of $\mathbb{H}$, and if f satisfies the equalities*

$$f(x_{\sigma(1)},\dots,x_{\sigma(n)}) = f(x_{\rho(1)},\dots,x_{\rho(n)}) \text{ for all } x_1,\dots,x_m \in H\times[k]$$

for some $\sigma,\rho\colon [n]\to[m]$, then f_i satisfies the equalities

$$f_i(h_{\sigma(1)},\dots,h_{\sigma(n)}) = f_i(h_{\rho(1)},\dots,h_{\rho(n)}) \text{ for all } h_1,\dots,h_n \in H.$$

Proof. We first see that f_i is an n-ary polymorphism of $\mathbb{H}$. Let $((h_1,\dots,h_n),(h'_1,\dots,h'_n))$ be an edge of $\mathbb{H}^n$, i.e., (h_m,h'_m) is an edge for $\mathbb{H}$ for each $m\in[n]$. Since $i<j$, it follows that $((h_1,i),\dots,(h_n,i)),((h'_1,j),\dots,(h'_n,j))$ is an edge of $\mathbb{H}\times\mathbb{T}_n$. Since f is a polymorphism, it must be the case that $(f((h_1,i),\dots,(h_n,i)),f((h'_1,j),\dots,(h'_n,j)))$ is an edge of $\mathbb{H}\times\mathbb{T}_n$. Recall that the projection π_H is a homomorphism from $\mathbb{H}\times\mathbb{T}_n$ onto $\mathbb{H}$, and so

$$(f_i(h_1,\dots,h_n),f_j(h'_1,\dots,h'_n)) = (\pi_H f((h_1,i),\dots,(h_n,i)),\pi_H f((h'_1,j),\dots,(h'_n,j))) \in E(\mathbb{H}).$$

Finally, since $f_i = f_j$, it follows that $(f_i(h_1,\dots,h_n),f_j(h_1,\dots,h_n))$ is an edge of $\mathbb{H}$, and thus $f_i\colon \mathbb{H}^n\to\mathbb{H}$ is an n-ary polymorphism of $\mathbb{H}$.

Now, let $h_1,\dots,h_m\in H$, and for each $l\in[n]$ let $\bar h_l := (h_l,i)$, so

$$f_i(h_{\sigma(1)},\dots,h_{\sigma(n)}) = \pi_H(f(\bar h_{\sigma(1)},\dots,\bar h_{\sigma(n)}).$$

Since f satisfies the loop condition for σ and ρ, it follows that

$$\pi_H(f(\bar h_{\sigma(1)},\dots,\bar h_{\sigma(n)}) = \pi_H(f(\bar h_{\rho(1)},\dots,\bar h_{\rho(n)}) = f_i(h_{\rho(1)},\dots,h_{\rho(n)}).$$

And thus, $f_i(h_{\sigma(1)},\dots,h_{\sigma(n)}) = f_i(h_{\rho(1)},\dots,h_{\rho(n)})$, and since the choice of $h_1,\dots,h_m\in H$ was arbitrary, the claim of this lemma follows. □

Building on this lemma and Theorem 30 together with Corollary 29, we prove the main result of this section.

Theorem 32. *For every finite digraph $\mathbb{H}$ one of the following statements holds.*

- *Either $\mathbb{H}$ has a polymorphism satisfying the Sigger's identity, and* $\mathrm{RCSP}(\mathbb{H},\mathbb{T}_k)$ *is in* P *for every positive integer k, or*
- *otherwise, there is a positive integer N such that* $\mathrm{RCSP}(\mathbb{H},\mathbb{T}_k)$ *is* NP*-hard for every $k\geq N$.*

Proof. The first item holds by Theorem 30, and because if $\mathrm{CSP}(\mathbb{H})$ is in P, then $\mathrm{RCSP}(\mathbb{H},\mathbb{T}_k)$ is in P for ever positive integer k. Now, suppose that $\mathbb{H}$ does not have a polymorphism satisfying the Sigger's identity. It follows from Corollary 29 that $\mathrm{RCSP}(\mathbb{H},\mathbb{T}_k)$ and $\mathrm{CSP}(\mathbb{H}\times\mathbb{T}_k)$ are polynomial-time equivalent. Let $M = |H|^{4|H|}+1$, and consider a polymorphism $f\colon (\mathbb{H}\times\mathbb{T}_N)^{\to}\mathbb{H}$. Clearly, there are $M-1$ functions from H^4 to H, and since for each $i\in[M]$ every f_i is a 4-ary function of H, it follows from the choice of M that there are $i<j\leq M$ such that $f_i = f_j$. So, by Lemma 31 if f satisfies the Sigger's identity, then f_i satisfies the Sigger's identity. Since $\mathbb{H}$ does not have such a polymorphism, then $\mathbb{H}\times\mathbb{T}_M$ does not have such a polymorphism, and hence $\mathrm{CSP}(\mathbb{H}\times\mathbb{T}_k)$ is NP-hard (see, e.g., Theorem 30). Moreover, since $\mathrm{CSP}(\mathbb{H}\times\mathbb{T}_M)$ is polynomial-time equivalent to $\mathrm{RCSP}(\mathbb{H},\mathbb{T}_M)$, it follows that $\mathrm{RCSP}(\mathbb{H},\mathbb{T}_M)$. Finally, it is straightforward to observe that if $\mathrm{RCSP}(\mathbb{H},\mathbb{T}_n)$ is at least as hard as $\mathrm{RCSP}(\mathbb{H},\mathbb{T}_{n+1})$, and so, if N is the smallest integer such that $\mathbb{H}\times\mathbb{T}_N$ does not have a Sigger's polymorphism, then N witnesses that the second itemized statement holds. □

The following is an immediate application of this theorem, of its proof, and of Corollary 29.

Corollary 33. *For every finite digraph $\mathbb{H}$, there is a positive integer $N \leq |H|^{4|H|} + 1$ such that* CSP($\mathbb{H}$) *is polynomial-time equivalent to* CSP($\mathbb{H}$) *restricted to acyclic digraphs with no directed walk on N vertices.*

As promised, we apply the framework of RCSPs to the context of $\mathcal{F}$-(subgraph)-free algorithmics.

Theorem 34. *For every digraph $\mathbb{H}$ such that* CSP($\mathbb{H}$) *is* NP*-hard, there is a positive integer N such that* CSP($\mathbb{H}$) *remains* NP*-complete even for $\vec{\mathbb{P}}_N$-subgraph-free acyclic digraphs. Moreover, there is such an N such that* CSP($\mathbb{H}$) *is polynomial-time solvable when the input is a $\vec{\mathbb{P}}_{N-1}$-subgraph-free digraph.*

Proof. If P = NP the claim is trivial, so we assume that P $\neq$ NP. Notice that a digraph $\mathbb{D}$ is acyclic and has no directed path on N verties if and only if there is no homomorphism $\vec{\mathbb{P}}_N \to \mathbb{D}$. Also, since CSP($\mathbb{H}$) is NP-complete and we are assuming that P $\neq$ NP, it follows from Theorem 30 that $\mathbb{H}$ has no Sigger's polymorphism. The claim of this theorem now follows from these simple arguments and Theorem 32. □

Corollary 35. *Let $\mathbb{H}$ be a digraph such that* CSP($\mathbb{H}$) *is* NP*-hard. Then, there are positive integers N and M such that*

- CSP($\mathbb{H}$) *is* NP*-hard for $\vec{\mathbb{P}}_k$-free digraphs whenever $k \geq N$, and*
- CSP($\mathbb{H}$) *is* NP*-hard for $\vec{\mathbb{P}}_k$-subgraph-free digraphs whenever $k \geq M$.*

We conclude this section with a simple example showing that Theorem 34 does not hold when $\mathbb{H}$ is infinite: there are infinite graphs $\mathbb{H}$ such that CSP($\mathbb{H}$) is NP-complete, but CSP($\mathbb{H}$) becomes tractable on acyclic instances. Moreover, $\mathbb{H}$ can be chosen to be *ω-categorical*, i.e., for every positive integer k the automorphism group of $\mathbb{H}$ defines finitely many orbits of k-tuples.

Example 36. Let $\mathbb{H}$ be the disjoint union of $\mathbb{K}_3$ and the rational number with the strict linear order. It is straightforward to observe that 3-Colouring reduces to CSP($\mathbb{H}$), and that every acyclic digraph $\mathbb{D}$ is a yes-instance of CSP($\mathbb{H}$). Hence, CSP($\mathbb{H}$) is NP-complete, but it is polynomial-time solvable on acyclic instance. The fact that $\mathbb{H}$ is ω-categorical follows from its definition — it is the disjoin union of two ω-categorical digraphs.

7 Digraphs on three vertices

In this section we answer Question 1 for digraphs on three vertices. A loopless digraph on three vertices either contains two directed cycles and its CSP if NP-complete, or otherwise its CSP is polynomial-time solvable. Hence, we focus on the three loopless digraphs on three vertices: $\vec{\mathbb{C}}_3^+$ (obtained from the directed cycle by adding one edge), $\vec{\mathbb{C}}_3^{++}$ (obtained from the directed cycle by adding two edges), and $\mathbb{K}_3$ — see also Figure 4.

Also note that these three digraphs are hereditarily hard (see Theorem 2), and it thus follows from Theorem 3 that RCSP($\mathbb{H}, \mathbb{T}_4$) is NP-hard for $\mathbb{H} \in \{\vec{\mathbb{C}}_3^+, \vec{\mathbb{C}}_3^{++}, \mathbb{K}_3\}$. In turn, this implies that for such digraphs $\mathbb{H}$ the problem CSP($\mathbb{H}$) is NP-hard even restricted to $\vec{\mathbb{P}}_5$-(subgraph)-free digraphs. We use this remark in both subsections below.

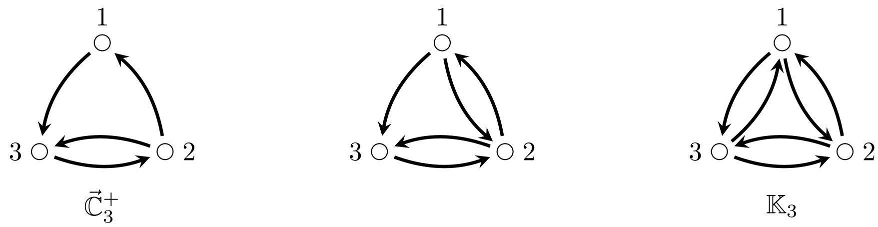


Figure 4: The three digraphs on three vertices with at least two directed cycles, and whose CSP is NP-complete.

7.1 $\vec{\mathbb{P}}_k$-subgraph-free digraphs

It is straightforward to observe that any orientation of an odd cycle contains $\vec{\mathbb{P}}_3$ as a subgraph. Hence, if $\mathbb{D}$ is a $\vec{\mathbb{P}}_3$-subgraph-free digraph, then it is bipartite, i.e., $\mathbb{D} \to \mathbb{K}_2$. In particular, this implies that CSP($\vec{\mathbb{C}}_3^+$), CSP($\vec{\mathbb{C}}_3^{++}$), and CSP(K_3) are polynomial-time solvable for $\vec{\mathbb{P}}_3$-subgraph-free digraphs. As previously mentioned, each of these CSPs is NP-hard for $\vec{\mathbb{P}}_5$-subgraph-free digraphs. In this section, we study the complexity CSP($\vec{\mathbb{C}}_3^+$), CSP($\vec{\mathbb{C}}_3^{++}$), and CSP($\mathbb{K}_3$) restricted to $\vec{\mathbb{P}}_4$-subgraph-free digraphs. We begin with a remark which we use a couple times in the remaining of the paper.

Remark 37*.* If a $\vec{\mathbb{P}}_4$-subgraph-free weakly connected digraph $\mathbb{D}$ contains (as a subgraph) a directed 3-cycle v_1, v_2, v_3, then $D = \{v_1, v_2, v_3\}$. Indeed, since $\mathbb{D}$ is $\vec{\mathbb{P}}_4$-subgraph-free and $(v_1, v_2), (v_2, v_3), (v_3, v_1) \in E$, it must be the case that the out- and in-neighbourhood of each v_i is a subset of $\{v_1, v_2, v_3\}$ (and so, the claim follows because $\mathbb{D}$ is weakly connected).

Observation 38. *Every loopless $\vec{\mathbb{P}}_4$-subgraph-free digraph is* 3*-colourable.*

Proof. We claim that it suffices to consider oriented graphs. Indeed, if $\mathbb{D}$ is not an oriented digraph, let $\mathbb{D}'$ be any spanning subdigraph of $\mathbb{D}$ obtained by removing exactly one edge of each symmetric pair $(x, y), (y, x) \in E(\mathbb{D})$. Clearly, $\mathbb{D}$ is 3-colourable if and only if $\mathbb{D}'$ is 3-colourable, and moreover, if $\mathbb{D}$ is a loopless $\vec{\mathbb{P}}_4$-subgraph-free digraph, then $\mathbb{D}'$ is loopless $\vec{\mathbb{P}}_4$-subgraph-free oriented graph. So assume that $\mathbb{D}$ is a loopless $\vec{\mathbb{P}}_4$-subgraph free oriented digraph, and without loss of generality consider the case when $\mathbb{D}$ is weakly connected. If $\mathbb{D}$ contains a directed 3-cycle x_1, x_2, x_3, then $D = \{x_1, x_2, x_3\}$ (Remark 37), and hence $\mathbb{D} \to \mathbb{K}_3$. Otherwise, assume that $\mathbb{D}$ contains no 3-cycle, and notice that in this case it follows that $\vec{\mathbb{P}}_4 \not\to \mathbb{D}$. Hence, by Observation 4 we conclude that $\mathbb{D} \to \mathbb{T}_3 \to \mathbb{K}_3$. □

Clearly, Observation 38 does not hold if we replace CSP($\mathbb{K}_3$) (i.e., 3-colourable) by CSP($\vec{\mathbb{C}}_3^{++}$): $\mathbb{K}_3$ is a simple counterexample. However, it turns out that $\mathbb{K}_3$ is the unique minimal counterexample. In the following proof we use the notion of a *leaf* of a digraph $\mathbb{D}$, i.e., a vertex $x \in \mathbb{D}$ with $|(N^+(x) \cup N^-(x)) \setminus \{x\}| = 1$.

Lemma 39. *Every loopless $\{\mathbb{K}_3, \vec{\mathbb{P}}_4\}$-subgraph-free digraph $\mathbb{D}$ admits a homomorphism to $\vec{\mathbb{C}}_3^{++}$.*

Proof. Without loss of generality assume that $\mathbb{D}$ is a weakly connected digraph. If $\mathbb{D}$ contains $\vec{\mathbb{C}}_3$ as a subgraph, then the claim follows from Remark 37. Also, notice that if $\mathbb{D}$ contains a symmetric path on three vertices x_1, x_2, x_3, i.e., $x_1 \neq x_3$ and $(x_1, x_2), (x_2, x_1), (x_2, x_3), (x_3, x_2)$, then x_2 is the unique neighbour of x_1 and also the unique neighbour of x_3 (because $\mathbb{D}$ is $\vec{\mathbb{P}}_4$-free). In other words, x_1 and x_3 are leaves of $\mathbb{D}$ with the same neighbourhood, and so, the digraph $\mathbb{D}$ homomorphically maps to $\mathbb{D} - x_3$, and thus $\mathbb{D} \to \vec{\mathbb{C}}_3^{++}$ if and only if $(\mathbb{D} - x_3) \to \vec{\mathbb{C}}_3^{++}$. Hence, it suffices to prove the claim for weakly connected $\{\vec{\mathbb{C}}_3, \mathbb{P}_3, \vec{\mathbb{P}}_4\}$-subgraph-free digraphs, and since every vertex of $\vec{\mathbb{C}}_3^{++}$ is incident to a symmetric edge, we may further assume that $\mathbb{D}$ has no leaves.

Let $\mathbb{D}^*$ be any digraph obtained from $\mathbb{D}$ after removing exactly one edge from every symmetric pair of edges of $\mathbb{D}$. Since $\mathbb{D}$ is $\{\vec{\mathbb{C}}_3, \vec{\mathbb{P}}_4\}$-subgraph-free, $\mathbb{D}^*$ has no directed walk on four vertices, and so there is a homomorphism $f\colon \mathbb{D}^* \to \mathbb{TT}_3$ (by Observation 4). We claim that the same function f defines a homomorphism $f\colon \mathbb{D} \to \vec{\mathbb{C}}_3^{++}$. To do so, it suffices to show that there is no edge (x, y) such that $f(x) = 3$ and $f(y) = 1$. Notice that if this were the case, since $f\colon \mathbb{D}^* \to \mathbb{TT}_3$ is a homomorphism, then x and y would induce a symmetric pair of edges in $\mathbb{D}$, i.e., $(x, y), (y, x) \in E(\mathbb{D})$. Thus, it suffices to show that if $(x, y), (y, x)$ is a symmetric pair of edges, then $f(x) = 2$ or $f(y) = 2$. Recall that we assume that $\mathbb{D}$ has no leaves, hence x and y have some neighbours x' and y' respectively. Without loss of generality assume that $(x, y) \in E(\mathbb{D}^*)$, and since $\mathbb{D}$ is $\vec{\mathbb{P}}_4$-subgraph-free, either $(x, x'), (y, y') \in E(\mathbb{D})$ or $(x', x), (y', y) \in E(\mathbb{D})$. Notice that in the former case, y is the middle vertex of a directed path of length 2 in $\mathbb{D}^*$. Hence, any homomorphism from $\mathbb{D}^*$ to $\mathbb{TT}_3$ maps y to 2, so in particular, $f(y) = 2$. With symmetric arguments it follows that if $(x', x), (y', y) \in E(\mathbb{D})$, then $f(x) = 2$. This proves that the function $f\colon D \to \{1, 2, 3\}$ defines a homomorphism $f\colon \mathbb{D} \to \vec{\mathbb{C}}_3^{*}$. □

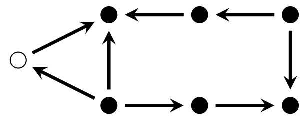
Now we show that $\mathrm{CSP}(\vec{\mathbb{C}}_3^{+})$ is polynomial-time solvable when the input is restricted to $\vec{\mathbb{P}}_4$-subgraph-free digraphs. To do so, we consider a unary predicate U, and we reduce the problem mentioned above to the CSP of the $\{E, U\}$-structure depicted in Figure 5. We first show that the underlying graph has a *conservative majority* polymorphism, i.e., a polymorphism $f\colon \mathbb{G}^3 \to \mathbb{G}$ such that $f(x, x, x) = f(x, x, y) = f(x, y, x) = f(y, x, x) = x$ and $f(x, y, z) \in \{x, y, z\}$.

Lemma 40. *The digraph $\mathbb{G}$ from Figure 5 has a conservative majority polymorphism. In particular,* $\mathrm{CSP}(\mathbb{G}, U)$ *can be solved in polynomial time.*

Proof. We define a conservative majority function $f\colon G^3 \to G$, and we then argue that it is a polymorphism. We denote by π_1 the projection onto the first coordinate, and by $\mathrm{maj}(x, y, z)$ the majority operation when $|\{x, y, z\}| \leq 2$. Also, we simplify our writing by implicitly assuming that in the n-th itemized case of the following definition, neither of the first $n - 1$ cases holds.

$$f(x,y,z) = \begin{cases} \mathrm{maj}(x,y,z) & \text{if } |\{x,y,z\}| \leq 2, \\ 1 & \text{if } 1 \in \{x,y,z\}, \text{ and } \{3,7\} \cap \{x,y,z\} \neq \varnothing, \\ 2 & \text{if } 2 \in \{x,y,z\}, \text{ and } \{6,7\} \cap \{x,y,z\} \neq \varnothing, \\ 3 & \text{if } 3 \in \{x,y,z\}, \text{ and } \{5,7\} \cap \{x,y,z\} \neq \varnothing, \\ 4 & \text{if } 4, 6 \in \{x,y,z\}, \\ 6 & \text{if } \{x,y,z\} = \{5,6,7\}, \\ \pi_1(x,y,z) & \text{otherwise.} \end{cases}$$

It is clear that f is a symmetric conservative majority function. It is routine verifying that f is indeed a polymorphism. The fact that $\mathrm{CSP}(\mathbb{G}, U)$ can be solved in polynomial time now follows from [20] because the same function f defines a majority polymorphism of $(\mathbb{G}, U)$. □



$(\mathbb{G}, U)$

Figure 5: A $\{E, U\}$-structure where E is a binary relation represented by arcs, and U is a unary relation represented by black vertices.

Lemma 41. $\mathrm{CSP}(\vec{\mathbb{C}}_3^+)$ *can be solved in polynomial-time when the input is restricted to* $\vec{\mathbb{P}}_4$*-subgraph-free digraphs.*

Proof. Consider a loopless weakly connected digraph $\mathbb{D}$ with no directed 3-cycles. Let B be the subset of vertices of $\mathbb{D}$ incident to some symmetric edge, and let $\mathbb{D}^*$ be the digraph obtained from $\mathbb{D}$ after removing exactly one edge from each symmetric pair of edges. It is straightforward to observe there is no homomorphism from the directed path on four vertices to $\mathbb{D}^*$. Hence, by Observation 4, $(\mathbb{D}^*, B)$ homomorphically maps to $(\mathbb{TT}_3, \{1,2,3\})$ considered as a $\{E, U\}$-structure where E is a binary predicate and U a unary predicate (this simply means that there is a homomorphism $\mathbb{D}^* \to \mathbb{TT}_3$ such that every vertex in B is mapped to a vertex in $\{1,2,3\}$). Also note that $\mathbb{D} \to \vec{\mathbb{C}}_3^+$ if and only if there is a homomorphism $f\colon \mathbb{D}^* \to \vec{\mathbb{C}}_3^+$ where $f(B) \subseteq \{2,3\}$, i.e., $\mathbb{D} \to \vec{\mathbb{C}}_3^+$ if and only if $(\mathbb{D}^*, B) \to (\vec{\mathbb{C}}_3^+, \{2,3\})$ (as $\{E, U\}$-structures). It follows from these arguments that $\mathrm{CSP}(\vec{\mathbb{C}}_3^+)$ restricted to $\{\vec{\mathbb{C}}_3, \vec{\mathbb{P}}_4\}$-subgraph-free digraphs reduces in polynomial time to $\mathrm{CSP}((\vec{\mathbb{C}}_3^+, \{2,3\}) \times (\mathbb{TT}_3, \{1,2,3\}))$. It is not hard to observe that the core of $(\vec{\mathbb{C}}_3^+, \{2,3\}) \times (\mathbb{TT}_3, \{1,2,3\})$ is the structure $(\mathbb{G}, U)$ depicted in Figure 5. It follows from Lemma 40 that $\mathrm{CSP}(\mathbb{G}, U)$ can be solved in polynomial time.

An algorithm that solves $\mathrm{CSP}(\vec{\mathbb{C}}_3^+)$ in polynomial time when the input is restricted to $\vec{\mathbb{P}}_4$-subgraph-free digraphs processes each weakly connected component of the input $\mathbb{D}$ and accepts if and only if the following subroutine accepts each weakly connected component. On a weakly connected component $\mathbb{D}'$ of $\mathbb{D}$, the subroutine distinguishes whether the $\mathbb{D}'$ is $\vec{\mathbb{C}}_3$-subgraph-free: if not, it accepts if $\mathbb{D}'$ is either $\vec{\mathbb{C}}_3$ or $\vec{\mathbb{C}}_3^+$, and rejects otherwise — this step is consistent and correct because $|D'| = 3$ (see Remark 37); if $\mathbb{D}'$ is $\vec{\mathbb{C}}_3$-free, then the subroutine accepts or rejects $\mathbb{D}'$ according to the reduction to $\mathrm{CSP}(\mathbb{G}, U)$ explained above. Given the arguments in the previous paragraph, it is clear that this algorithm is consistent, correct, and runs in polynomial time. $\square$

Theorem 42. *The following statements hold for every positive integer k.*

- *If $k \leq 4$, then* $\mathrm{CSP}(\vec{\mathbb{C}}_3^+)$, $\mathrm{CSP}(\vec{\mathbb{C}}_3^{++})$*, and* $\mathrm{CSP}(\mathbb{K}_3)$ *are solvable in polynomial-time when the input is restricted to* $\vec{\mathbb{P}}_k$*-subgraph-free digraphs.*
- *If $k \geq 5$, then* $\mathrm{CSP}(\vec{\mathbb{C}}_3^+)$, $\mathrm{CSP}(\vec{\mathbb{C}}_3^{++})$*, and* $\mathrm{CSP}(\mathbb{K}_3)$ *are* NP*-hard even if the input is restricted to* $\vec{\mathbb{P}}_k$*-subgraph-free digraphs.*

Proof. The first itemized statement follows from Observation 38, from Lemma 39 and from Lemma 41. The second itemized claim follows from the discussion in the second paragraph of Section 7. $\square$

7.2 $\vec{\mathbb{P}}_k$-free digraphs

Notice that a digraph $\mathbb{D}$ is $\vec{\mathbb{P}}_2$-free if and only if it is a symmetric (undirected) graph. Hence, CSP($\mathbb{K}_3$) is NP-hard for $\vec{\mathbb{P}}_2$-free digraphs. Also note that a symmetric graph $\mathbb{D}$ maps to $\vec{\mathbb{C}}_3^+$ if and only if $\mathbb{D}$ is bipartite, and the same statement holds for $\vec{\mathbb{C}}_3^{++}$. Hence, in this section we study the complexity of CSP($\vec{\mathbb{C}}_3^+$) and CSP($\vec{\mathbb{C}}_3^{++}$) with input restrictions to $\vec{\mathbb{P}}_3$-free and to $\vec{\mathbb{P}}_4$-free digraphs.

Lemma 43. CSP($\vec{\mathbb{C}}_3^+$) *is* NP*-hard even when the input* $\mathbb{D}$ *satisfies the following conditions, where* $\mathbb{F}$ *is any given orientation of the claw* $\mathbb{K}_{1,3}$*:*

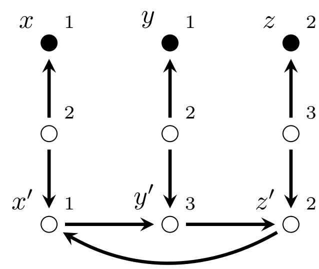
- $\mathbb{D}$ *is* $\{\mathbb{F}, \vec{\mathbb{P}}_5, \mathbb{P}_5^{\leftarrow\leftarrow\rightarrow\rightarrow}, \mathbb{P}_5^{\rightarrow\rightarrow\leftarrow\leftarrow}\}$*-subgraph-free,*
- $\mathbb{D}$ *is* $\vec{\mathbb{P}}_4$*-free, and*
- $d^+(v) + d^-(v) \leq 3$ *for every* $v \in D$*.*

Proof. We propose a polynomial-time reduction from positive 1-IN-3 SAT to CSP($\vec{\mathbb{C}}_3^+$) mapping an instance $\mathbb{I}$ to a digraph $\mathbb{D}$ satisfying the statements above. Let $\mathbb{I}$ be an instance $(x_1 \vee y_1 \vee z_1) \wedge \cdots \wedge (x_m \vee y_m \vee z_m)$ where all variables are positive. Construct $\mathbb{D}$ as follows:

- For each clause $(x_i \vee y_i \vee z_i)$ introduce a fresh copy $\mathbb{G}_i$ of $\mathbb{G}$ (depicted in Figure 6) with distinguished vertices x_i, y_i, z_i, and
- for each variable v that occurs n_v times, construct an undirected cycle with exactly n_v vertices. Now substitute each edge pq in this cycle for a gadget of two edges on three vertices given by the back-and-forward formation $(r_{p,q}, p), (r_{p,q}, q)$, where $r_{p,q}$ is a new vertex. Notice that the resulting digraph is an oriented cycle $\mathbb{C}_v$ where the n_v vertices from the original undirected cycle correspond to sinks in $\mathbb{C}_v$. Now, when variable v appears in the i-th clause, identify the vertex v in the clause gadget $\mathbb{G}_i$ with a unique sink the oriented cycle $\mathbb{C}_v$.

Notice that $\exists w E(w,x) \wedge E(w,y)$ defines an equivalence relation in $\vec{\mathbb{C}}_3^+$ with two classes $\{1,3\}$ and $\{2\}$. Hence, for each $a \in \{1,2,3\}$, any homomorphism $f\colon \mathbb{G} \to \vec{\mathbb{C}}_3^+$ satisfies that $f(a') = 2$ if and only if $f(a) = 2$. Moreover, it follows from the same argument that a sink in the cycle $\mathbb{C}_v$ is mapped to 2 if and only if all sinks in $\mathbb{C}_v$ are mapped to 2. Also notice that any homomorphism from the directed 3-cycle to $\vec{\mathbb{C}}_3^+$ must map exactly one vertex to 2. Hence, any homomorphism $g\colon \mathbb{D} \to \vec{\mathbb{C}}_3^+$ satisfies that exactly one of the vertices x_i, y_i, z_i is mapped to 2. This yields a solution to the positive 1-IN-3 SAT instance $\mathbb{I}$ by assigning for each $i \in [m]$ and $a \in \{x,y,z\}$ the value 1 if $g(a_i) = 2$, and $a_i := 0$ otherwise. The converse implication (if $\mathbb{I}$ is a yes-instance then $\mathbb{D} \to \vec{\mathbb{C}}_3^+$) follows similarly by noticing that for each $a \in \{x,y,z\}$ there is a homomorphism $f\colon \mathbb{G} \to \vec{\mathbb{C}}_3^+$ mapping a to 2 and $f(b) = 1$ for $b \in \{x,y,z\} \setminus \{a\}$ (in Figure 6 we describe such a homomorphism for $a = x$).

Finally, it is straightforward to notice that every digraph $\mathbb{D}$ in the image of the reduction contains no path on five vertices, and no induced path on four vertices. It is also clear that $\mathbb{D}$ is $\{\mathbb{P}_5^{\leftarrow\leftarrow\rightarrow\rightarrow}, \mathbb{P}_5^{\rightarrow\rightarrow\leftarrow\leftarrow}\}$-subgraph-free, and that $d^+(u) + d^-(u) \leq 3$ for every $u \in D$. Now, notice that if $d^+(u)+d^-(u) = 3$, then u belongs to some directed 3-cycle from a gadget $\mathbb{G}_i$ or is a source in some cycle $\mathbb{C}_v$. In the former case $d^-(v) = 2$ and in the latter $d^-(v) = 0$, and thus $\mathbb{D}$ is $\mathbb{F}$-subgraph-free whenever $\mathbb{F}$ is the orientation of $\mathbb{K}_{1,3}$ where the center vertex is a sink, or when it has out-degree two. The cases when $\mathbb{F}$ is on of the remaining two possible orientations of $\mathbb{K}_{1,3}$ simply follows by considering the reduction that maps a digraph $\mathbb{D}$ to the digraph obtained from $\mathbb{D}$ by reversing the orientation of each edge of $\mathbb{D}$. ◻



$\mathbb{G}(x, y, z)$

Figure 6: A depiction of the gadget reduction $\mathbb{I} \mapsto \mathbb{D}$ from positive 1-IN-3 SAT to CSP($\vec{\mathbb{C}}_3^+$) applied to a clause $(x \vee y \vee z)$ of the instance $\mathbb{I}$ to 1-IN-3 SAT. The numbers indicate a function that defined a homomorphism $f \colon \mathbb{G} \to \vec{\mathbb{C}}_3^+$.

Now we show that CSP($\vec{\mathbb{C}}_3^+$) can be solved in polynomial time when the input digraph $\mathbb{D}$ contains no induced directed path on three vertices. On a given $\vec{\mathbb{P}}_3$-free digraph $\mathbb{D}$, the algorithm works as follows. At each step there are three sets $X_1, X_2, X_3 \subseteq D$ such that the mapping $x \mapsto i$ if $x \in X_i$ defines a partial homomorphism from $\mathbb{D}$ to $\vec{\mathbb{C}}_3^+$. We extend these sets to X_1', X_2', X_3' in such a way that the partial homomorphism defined by X_1, X_2, X_3 extends to a homomorphism from $\mathbb{D} \to \vec{\mathbb{C}}_3^+$ if and only if the partial homomorphism defined by X_1', X_2', X_3' extends to a homomorphism from $\mathbb{D} \to \vec{\mathbb{C}}_3^+$.

Lemma 44. CSP($\vec{\mathbb{C}}_3^+$) *can be solved in quadratic time when the input is restricted to* $\vec{\mathbb{P}}_3$*-free digraphs.*

Proof. We assume without loss of generality that $\mathbb{D}$ is a $\vec{\mathbb{C}}_3$-free weakly connected digraph. First notice that $\mathbb{D} \to \vec{\mathbb{C}}_3^+$ if and only if there is a homomorphism $f \colon \mathbb{D} \to \vec{\mathbb{C}}_3$ such that $f^{-1}(2) \neq \varnothing$. As previously mentioned, the idea of the algorithm is to deterministically extend a partial homomorphism f to a partial homomorphism f' in such a way that f can be extended to a homomorphism $\mathbb{D} \to \vec{\mathbb{C}}_3^+$ if an only if f' can be extended as well. Hence, by the previous observation it sufficed to run this routine over all partial homomorphisms defined on one vertex v where $v \mapsto 2$, and if some homomorphism is found, then we accept $\mathbb{D}$, and otherwise we reject $\mathbb{D}$.

Let $\mathbb{D}$ be a finite digraph, X_1, X_2, X_3 three disjoint vertex subsets of D such that $X_1 \cup X_2 \cup X_3 \neq \varnothing$, and $x \mapsto i$ if $x \in X_i$ defines a partial homomorphism $f \colon \mathbb{D}[X_1 \cup X_2 \cup X_3] \to \vec{\mathbb{C}}_3^+$. Now, consider any vertex $v \in D \setminus (X_1 \cup X_2 \cup X_3)$ that belongs to the in- or out-neighbourhood of $X_1 \cup X_2 \cup X_3$. We define X_1', X_2', X_3' according to the following case distinction.

1. If $v \in N^+(X_1)$ or $v \in N^-(X_2)$, then $X_1' := X_1$, $X_2' := X_2$, and $X_3' := X_3 \cup \{v\}$;
2. If $v \in N^-(X_1)$ or $v \in N^+(X_3)$, then $X_1' := X_1$, $X_2' := X_2 \cup \{v\}$, and $X_3' := X_3$;
3. If $v \in N^+(X_2)$, we consider the following subcases,
   (a) if there is a symmetric edge incident in v, then $X_1' := X_1$, $X_2' := X_2$, and $X_3' := X_3 \cup \{v\}$;
   (b) otherwise, if v has an out-neighbour, then $X_1' := X_1 \cup \{v\}$, $X_2' := X_2$, and $X_3' := X_3$;
   (c) if neither of the above hold, then $X_1' := X_1$, $X_2' := X_2$, and $X_3' := X_3 \cup \{v\}$;

4. If $v \in N^-(X_3)$, we consider the following subcases,

   (a) if there is a symmetric edge incident in v, then $X_1' := X_1$, $X_2' := X_2 \cup \{v\}$, and $X_3' := X_3$;

   (b) otherwise, if v has an in-neighbour, then $X_1' := X_1 \cup \{v\}$, $X_2' := X_2$, and $X_3' := X_3$;

   (c) if neither of the above hold, then $X_1' := X_1$, $X_2' := X_2 \cup \{v\}$, and $X_3' := X_3$.

Let $f' \colon \mathbb{D}[X_1' \cup X_2' \cup X_3'] \to \vec{\mathbb{C}}_3^+$ be the partial mapping $x \mapsto i$ if $x \in X_i'$. We prove that if f extends to a homomorphism from $\mathbb{D}$ to $\vec{\mathbb{C}}_3^+$, then so does f'. This is clearly the case in cases 1 and 2, and also in 3a and 4a. We now consider the case 3b. Let $x \in X_2$ and $y \in D$ such that $(x,v),(v,y) \in E(\mathbb{D})$. Since v is not incident in any symmetric edge, then x and y must be adjacent vertices in $\mathbb{D}$, i.e., $(x,y) \in E(\mathbb{D})$ or $(y,x) \in E(\mathbb{D})$ (or both). In either case, there is a unique way of extending the partial homomorphism $x \mapsto 2$ to the subbgraph with vertex set $\{x,y,v\}$, and in this unique extension it is the case that v is mapped to 1. Hence, f extends to a homomorphism if and only if f' extends to a homomorphism $\mathbb{D} \to \vec{\mathbb{C}}_3^+$. Now, consider the case 3c and let $x \in X_2$ be an in-neighbour of v. Since v no symmetric edge is incident in v, and v has no out-neighbours, it follows that for any vertex y adjacent to v, there is a walk of the form $\to\leftarrow$ connecting y and x. Notice that if $i \in \{1,2,3\}$ is connected to 2 by such a walk in $\vec{\mathbb{C}}_3^+$, then $i \in \{1,2\}$. This implies that if some homomorphism $g \colon \mathbb{D} \to \vec{\mathbb{C}}_3^+$ mapping x to 2, then the image of the neighbourhood of v is contained in $\{1,2\}$. Since v has in-neighbours but no out-neighbours, and 1 and 2 are in-neighbours of 3 in $\vec{\mathbb{C}}_3^+$, it follows that the mapping $g' \colon \mathbb{D} \to \vec{\mathbb{C}}_3^+$ defined by $g'(v) = 3$ and $g'(u) = g(u)$ if $u \neq v$ is a also a homomorphism from $\mathbb{D} \to \vec{\mathbb{C}}_3^+$. Therefore, if f extends to a homomorphism $g \colon \mathbb{D} \to \vec{\mathbb{C}}_3^+$, then f' extends to a homomorphism $g' \colon \mathbb{D} \to \vec{\mathbb{C}}_3^+$. The cases 4b and 4c follows with symmetric arguments.

Hence, on a given weakly connected digraph $\mathbb{D}$, a quadratic time algorithm works as follows. For a vertex $v \in D$ define $X_1 := \varnothing$, $X_2 := \{v\}$, and $X_3 := \varnothing$, and perform the subroutine above until either $X_1' \cup X_2' \cup X_3'$ does not define a partial homomorphism, and in this case choose a new vertex $u \in D$; or $X_1 \cup X_2 \cup X_3 = D$, and in this case accept $\mathbb{D}$. If after running the subroutine on every vertex $v \in D$ we arrive to some sets X_1', X_2', X_3' that do not define a partial homomorphism, reject $\mathbb{D}$. $\square$

Lemma 45. CSP($\vec{\mathbb{C}}_3^{++}$) *is* NP*-hard even for a* $\{\vec{\mathbb{P}}_3, \vec{\mathbb{P}}_3^{\leftarrow\to}, \vec{\mathbb{P}}_3^{\to\leftarrow}\}$*-free digraphs.*

Proof. We follow the proof of Lemma 43, noting that $\exists w. E(w,x) \wedge E(x,w) \wedge E(w,y) \wedge E(y,w)$ defines an equivalence relation with two classes $\{1,2\}$ and $\{3\}$. The gadget we use is depicted in Figure 7, with the same reduction from 1-IN-3 SAT. $\square$

The following statement now follows from Lemmas 43, 44, and 45.

Theorem 46. *The following statements hold for every positive integer* $k \geq 2$.

- CSP($\vec{\mathbb{C}}_3^+$) *is in solvable in quadratic time for* $\vec{\mathbb{P}}_k$*-free digraphs if* $k \leq 3$*, and* NP*-hard even for* $\vec{\mathbb{P}}_k$*-free digraphs if* $k \geq 4$,
- CSP($\vec{\mathbb{C}}_3^{++}$) *is solvable in linear time for* $\vec{\mathbb{P}}_k$*-free digraphs if* $k = 2$*, and* NP*-hard even for* $\vec{\mathbb{P}}_k$*-free digraphs if* $k \geq 3$*, and*
- CSP($\mathbb{K}_3$) *is* NP*-hard even for* $\vec{\mathbb{P}}_k$*-free digraphs.*

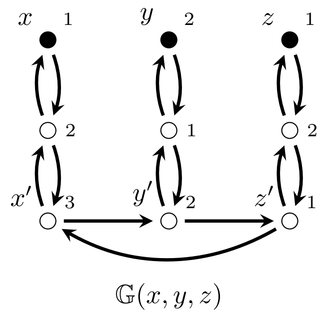


Figure 7: A depiction of the gadget reduction $\mathbb{I} \mapsto \mathbb{D}$ from positive 1-IN-3 SAT to CSP($\vec{\mathbb{C}}_3^{++}$) applied to a clause $(x \vee y \vee z)$ of the instance $\mathbb{I}$ to 1-IN-3 SAT. The numbers indicate a function that defined a homomorphism $f\colon \mathbb{G} \to \vec{\mathbb{C}}_3^{++}$.

8 A family of smooth tournaments

In this section we answer Question 1 for a natural family of tournaments smooth tournaments $\mathbb{TC}_n$. Given a positive integer n, we denote by $\mathbb{TC}_n$ the tournament obtained from $\mathbb{T}_n$ be reversing the edge from the source to the sink (see Figure 8). In particular, $\mathbb{TC}_2 \cong \mathbb{T}_2$, and $\mathbb{TC}_3 \cong \vec{\mathbb{C}}_3$, so CSP($\mathbb{TC}_n$) is polynomial-time solvable for $n \leq 3$, and NP-complete for $n \geq 4$ (see, e.g., [2]).

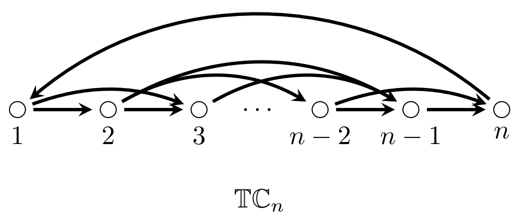


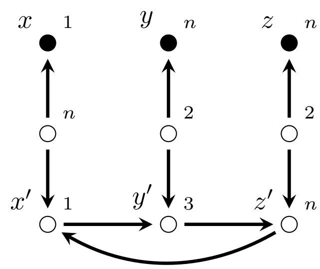
Figure 8: A depiction of the $\mathbb{TC}_n$ — for a cleaner picture we omit the edges $(1, n-2)$, $(1, n-1)$, $(2, n)$, and $(3, n)$.

Since $\mathbb{TC}_n$ is a hereditary hard digraph (Theorem 2) and $\mathbb{T}_n \not\to \mathbb{TC}_n$, it follows that RCSP($\mathbb{TC}_n, \mathbb{T}_n$) is NP-hard (Theorem 3). Equivalently, CSP($\mathbb{TC}_n$) is NP-hard for digraphs with no directed walk on $n+1$ vertices, and since $\mathbb{T}_{n-1} \to \mathbb{TC}_n$, CSP($\mathbb{TC}_n$) is polynomial-time solvable for digraphs with no directed walk on n vertices. However, is we only forbid $\vec{\mathbb{P}}_k$ as a subgraph (and not homomorphically), it turns out that CSP($\mathbb{TC}_n$) is NP-hard even for $\vec{\mathbb{P}}_5$-subgraph-free digraphs.

8.1 $\vec{\mathbb{P}}_k$-subgraph-free digraphs

For these hardness results we consider the gadget reduction depicted in Figure 9 and described in the proof of the following lemma.

Lemma 47. *For every positive integer $n \geq 4$, CSP($\mathbb{TC}_n$) is NP-hard even when the input $\mathbb{D}$ satisfies the following conditions:*



$\mathbb{G}(x, y, z)$

Figure 9: A depiction of the gadget reduction $\mathbb{I} \mapsto \mathbb{D}$ from positive 1-IN-3 SAT to CSP($\mathbb{TC}_n$) applied to a clause $(x \vee y \vee z)$ of the instance $\mathbb{I}$ to 1-IN-3 SAT. The numbers indicate a function that defined a homomorphism $f\colon \mathbb{G} \to \mathbb{TC}_n$ whenever $n \geq 4$.

- $\mathbb{D}$ *is* $\{\mathbb{F}, \vec{\mathbb{P}}_5, \mathbb{P}_5^{\leftarrow\leftarrow\rightarrow\rightarrow}, \mathbb{P}_5^{\rightarrow\rightarrow\leftarrow\leftarrow}\}$*-subgraph-free,*
- $\mathbb{D}$ *is* $\vec{\mathbb{P}}_4$*-free, and*
- $d^+(v) + d^-(v) \leq 3$ *for every* $v \in D$*.*

Proof. We use the same reduction from 1-IN-3-SAT to CSP($\vec{\mathbb{C}}_3^+$) from Lemma 43:

- For each clause $(x_i \vee y_i \vee z_i)$ introduce a fresh copy $\mathbb{G}_i$ of $\mathbb{G}$ (depicted in Figure 9) with distinguished vertices x_i, y_i, z_i, and
- for each variable v that occurs n_v times, construct an undirected cycle with exactly n_v vertices. Now substitute each edge pq in this cycle for a gadget of two edges on three vertices given by the back-and-forward formation $(r_{p,q}, p), (r_{p,q}, q)$, where $r_{p,q}$ is a new vertex. Notice that the resulting digraph is an oriented cycle $\mathbb{C}_v$ where the n_v vertices from the original undirected cycle correspond to sinks in $\mathbb{C}_v$. Now, when variable v appears in the i-th clause, identify the vertex v in the clause gadget $\mathbb{G}_i$ with a unique sink the oriented cycle $\mathbb{C}_v$.

Notice that any homomorphism $f\colon \mathbb{G} \to \mathbb{TC}_n$ satisfies that $f(a') = 1$, if and only if $f(a) = 1$ for each $a \in \{x, y, z\}$. Indeed, this is independent of the cycle in this gadget, it applies already to the back-and-forward formation $(r_{a,a'}, a), (r_{a,a'}, a')$. It thus follows that a sink in the cycle $\mathbb{C}_v$ is mapped to 1 if and only if all sinks in $\mathbb{C}_v$ are mapped to 1. Also notice that any homomorphism from the directed 3-cycle to $\mathbb{TC}_n$ must map exactly one vertex to 1. Hence, any homomorphism $g\colon \mathbb{D} \to \mathbb{TC}_n$ satisfies that exactly one of the vertices x_i, y_i, z_i is mapped to 1. This yields a solution to the positive 1-IN-3 SAT instance $\mathbb{I}$ by assigning for each $i \in [m]$ and $a \in \{x, y, z\}$ the value 1 if $g(a_i) = 1$, and $a_i := 0$ otherwise. The converse implication (if $\mathbb{I}$ is a yes-instance then $\mathbb{D} \to \mathbb{TC}_n$) follows similarly by noticing that for each $a \in \{x, y, z\}$ there is a homomorphism $f\colon \mathbb{G} \to \mathbb{TC}_n$ mapping a to 1 and $f(b) = n$ for $b \in \{x, y, z\} \setminus \{a\}$ (in Figure 9 we describe such a homomorphism for $a = x$).

Finally, the fact that $\mathbb{D}$ satisfies the structural restrictions follows with the same arguments as in the proof of Lemma 43. □

We now argue that CSP($\mathbb{TC}_n$) is polynomial-time solvable when the input is restricted to $\vec{\mathbb{P}}_4$-subgraph-free digraphs, and together with Lemma 47 we obtain a complexity classification for these CSPs restricted to $\vec{\mathbb{P}}_k$-subgraph-free digraphs.

Lemma 48. *Consider a (possibly infinite) digraph $\mathbb{H}$. If $\mathbb{T}_3 \to \mathbb{H}$, then* CSP($\mathbb{H}$) *is in* P *for $\vec{\mathbb{P}}_4$-subgraph-free oriented graphs. In particular, if $\mathbb{H}$ is an oriented graph and $\mathbb{T}_3 \to \mathbb{H}$, then* CSP($\mathbb{H}$) *is in* P *for $\vec{\mathbb{P}}_4$-subgraph-free digraphs.*

Proof. The claim is trivial when $\mathbb{H}$ contains a loop, so we assume that $\mathbb{H}$ is a loopless digraph. Let $\mathbb{D}$ be an oriented graph and without loss of generality assume it is a weakly connected digraph. First check if $\mathbb{D}$ contains a directed 3-cycle, and if yes, then by Remark 37 we know that D has exactly three vertices (at this step the algorithm would say yes if $\mathbb{D}$ is one the subdigraphs of $\mathbb{H}$ on three vertices). If not, notice that $\vec{\mathbb{P}}_4 \not\to \mathbb{D}$ if and only if $\mathbb{D}$ contains no loops. Hence, the algorithm accepts whenever $\mathbb{D}$ contains no loops, and rejects otherwise. The "in particular" statement follows because if the input $\mathbb{D}$ is not an oriented graph but $\mathbb{H}$ is, then $\mathbb{D} \not\to \mathbb{H}$, hence it suffices to solve CSP($\mathbb{H}$) for $\vec{\mathbb{P}}_4$-subgraph-free oriented graphs. □

Theorem 49. *For every pair of positive integers n and k the following statements hold.*

- *If $n \leq 3$ or $k \leq 4$, then* CSP($\mathbb{TC}_n$) *is in* P *for $\vec{\mathbb{P}}_k$-subgraph-free digraphs.*
- *If $n \geq 4$ and $k \geq 5$, then* CSP($\mathbb{TC}_n$) *is* NP*-hard even for $\vec{\mathbb{P}}_k$-subgraph-free digraphs.*

Proof. The second statement follows via Lemma 47, and the case $n \geq 3$ holds because $\mathbb{TC}_2$ and $\mathbb{TC}_3$ are transitive tournaments. Finally, the case $k \leq 4$ and $n \geq 4$ follows from Lemma 48 because each $\mathbb{T}_n$ is an oriented graph that contains a transitive tournament three vertices whenever $n \geq 3$. □

8.2 $\vec{\mathbb{P}}_k$-free digraphs

In this subsection we prove a structural result (Theorem 55) asserting that there is a finite set of digraphs $\mathcal{F}$ such that a $\vec{\mathbb{P}}_3$-free digraphs $\mathbb{D}$ belongs to CSP($\mathbb{TC}_n$) if and only if $\mathbb{D}$ is $\mathcal{F}$-free. We then use this result to propose a complexity classification for these CSPs restricted to $\vec{\mathbb{P}}_k$-free digraphs.

We begin by showing that there are finitely many minimal $\vec{\mathbb{P}}_3$-free digraphs that do not homomorphically map to CSP($\mathbb{TC}_4$). We depict the four non-isomorphic tournaments on four vertices in Figure 10.

Lemma 50. *The following statements are equivalent for a $\vec{\mathbb{P}}_3$-free digraph $\mathbb{D}$.*

- *$\mathbb{D} \to \mathbb{TC}_4$, and*
- *$\mathbb{D}$ is a $\{\mathbb{T}_4, \mathbb{T}_4^a, \mathbb{T}_4^b\}$-free loopless oriented graph with no tournament on five vertices.*

Proof. The first itemized statement clearly implies the second one. To prove the converse implication we proceed by case distinction over the largest tournament in $\mathbb{D}$, and without loss of generality we assume that $\mathbb{D}$ is a weakly connected digraph.

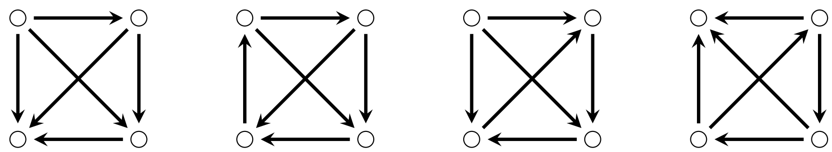
- If $\mathbb{D}$ is a $\vec{\mathbb{P}}_3$-free loopless oriented graph with **no tournament on three vertices**, then $\vec{\mathbb{P}}_3 \not\to \mathbb{D}$ so $\mathbb{D} \to \mathbb{T}_2$ (Observation 4), and thus $\mathbb{D} \to \mathbb{TC}_4$.
- If $\mathbb{D}$ contains **no tournament on four vertices and no** $\vec{\mathbb{C}}_3$, then there is no homomorphism $\vec{\mathbb{P}}_4 \to \mathbb{D}$, so $\mathbb{D} \to \mathbb{T}_3$ (Observation 4), which implies that $\mathbb{D} \to \mathbb{TC}_4$.

- Now, suppose that $\mathbb{D}$ contains **no tournament on four vertices and a** $\vec{\mathbb{C}}_3$ with vertices d_1, d_2, d_3 and edges $(d_1, d_2), (d_2, d_3), (d_3, d_1)$. Notice that if d is an out-neighbour of d_2, then d is an in-neighbour of d_1. Indeed, since $\mathbb{D}$ is $\vec{\mathbb{P}}_3$-free, it must be the case that d is a neighbour of d_1, and if $(d_1, d) \in E(\mathbb{D})$ it would also be the case that d is a neighbour of d_3 contradicting the choice of $\mathbb{D}$. Hence, every out-neighbour of d_2 is an in-neighbour of d_1, and symmetrically, every in-neighbour of d_2 is an out-neighbour of d_3. Therefore, the mapping f defined by $d \mapsto 1$ if $d \in N^-(d_2)$, $d \mapsto 2$ if $d \in N^-(d_3)$, and $d \mapsto 3$ if $d \in N^-(d_1)$ is a partial homomorphism from $\mathbb{D}$ to $\vec{\mathbb{C}}_3$. The fact that f is a homomorphism also follows from the previous observation and the assumption that $\mathbb{D}$ is weakly connected.

- Finally, suppose that $\mathbb{D}$ **contains a tournament on four vertices** d_1, d_2, d_3, d_4. Since $\mathbb{D}$ is $\{\mathbb{T}_4, \mathbb{T}_4^a, \mathbb{T}_4^b\}$-free, we assume without loss of generality that the mapping $i \mapsto d_i$ is an embedding of $\mathbb{TC}_4$ into $\mathbb{D}$. Let d be an out-neighbour of d_4. Since $(d_2, d_4), (d_3, d_4) \in E(\mathbb{D})$, it follows that d is a neighbour of d_2 and of d_3. Notice that if $(d_2, d) \in E(\mathbb{D})$ or $(d_3, d) \in E(\mathbb{D})$, then the $\vec{\mathbb{P}}_3$-freeness of $\mathbb{D}$ implies that d is also a neighbour of d_1 contradicting the assumption that $\mathbb{D}$ contains no tournament on five vertices. Hence, every out-neighbour of d is an in-neighbour of d_2 and of d_3, and it is not adjacent to d_1. With symmetric arguments it follows that every in-neighbour of d_1 is an out-neighbour of d_2 and of d_3, and it is not adjacent to d_4. We define $X_1 := N^+(d_4)$, $X_4 := N^-(d_1)$, and $Y := N^+(d_1)$. We now claim that $Y = N^+(d_1) = N^-(d_4)$ and that (X_1, Y, X_4) is a partition of the vertex set D. Again, using the fact that $\mathbb{D}$ is $\vec{\mathbb{P}}_3$-free one can notice that every $d \in N^+(d_1)$ is a neighbour of d_4 (because $(d_4, d_1) \in E(\mathbb{D})$), and since every out-neighbour of d_4 is not adjacent to d_1 (by the previous arguments), it follows that $d \in N^-(d_4)$. Hence, $N^+(d_1) \subseteq N^-(d_4)$, and with symmetric arguments we conclude that $N^+(d_1) = N^-(d_4)$. Since $\mathbb{D}$ is an oriented graph, the sets X_1, $N^+(d_1)$, and X_4 are disjoint, and – anticipating a contradiction – suppose there is a vertex $d \in D$ not adjacent to d_1 nor d_4. Moreover, since $\mathbb{D}$ is weakly connected, we choose d so that it does not belong to $X_1 \cup N^+(d_1) \cup X_4$ and it is adjacent to some vertex $c \in X_1 \cup N^+(d_1) \cup X_4$. To arrive to a contradiction it suffices to notice that d is a neighbour of d_1 or of d_4 — this implies that $d \in N^+(d_1) \cup N^-(d_1) \cup N^+(d_4) \cup N^-(d_4) = X_1 \cup N^+(d_1) \cup X_4$. To do so one can proceed with a case distinction over which set c belongs to (from X_1, $N^+(d_1)$, or X_4)), and whether $(d, c) \in E(\mathbb{D})$ or $(c, d) \in E(\mathbb{D})$. We give a sketch and leave the missing details to the reader: in case of $c \in N^+(d_1)$, d ends up adjacent to d_4 if $(d, c) \in E(\mathbb{D})$, and adjacent to d_1 if $(c, d) \in E(\mathbb{D})$; in case of $c \in X_1$, d ends up adjacent to d_4 if $(c, d) \in E(\mathbb{D})$, and adjacent to some $c' \in N^+(d_1)$, if $(c, d) \in E(\mathbb{D})$, hence we fall in the first case; finally, the case when $c \in X_4$ follows with symmetric arguments to the case of $c \in X_1$.

  Now, notice that $N^+(d_4)$ cannot contain three vertices c_1, c_2, c_3 that induces either a directed three cycle, or a transitive tournament; otherwise, d_1, c_1, c_2, c_3 induce a tournament on four vertices not isomorphic to $\mathbb{TC}_4$ contradicting the choice of $\mathbb{D}$. From this observation together with the fact that $\mathbb{D}$ is $\vec{\mathbb{P}}_3$-free we conclude that the subdigraph of $\mathbb{D}$ induced by $N^+(d_1)$ does not contain an oriented walk on three vertices. Hence, by Observation 4, there is a homomorphism $h\colon \mathbb{D}[N^+(d_1)] \to \mathbb{TT}_2$ from the subdigraph of $\mathbb{D}$ induced by $N^+(d_1)$ to the transitive tournament on two vertices, i.e., there is a partition (X_2, X_3) of $N^+(d_1)$ such that for $c, d \in X_2, X_3$ if $(c, d) \in E(\mathbb{D})$, then $c \in X_2$ and $d \in X_3$. We finally conclude that the function f defined by $d \mapsto i$ if $d \in X_i$ for $i \in \{1, 2, 3, 4\}$ defines a homomorphism $f\colon \mathbb{D} \to \mathbb{TC}_4$. added a sentence here

$\square$



$\mathbb{T}_4$ $\mathbb{TC}_4$ $\mathbb{T}_4^a$ $\mathbb{T}_4^b$

Figure 10: The four non-isomorphic oriented tournaments on 4 vertices

For the remaining of this section, for each positive integer $n \geq 4$ we fix the set $\mathcal{F}_n$ to be the (finite) set of tournaments on at most $n+1$ vertices that do not embed into $\mathbb{TC}_n$ (up to isomorphism).

Remark 51. Since every (induced subgraph) tournament on $n-1$ vertices in $\mathbb{TC}_n$ is either $\mathbb{TT}_{n-1}$ or $\mathbb{TC}_{n-1}$, the following equality holds

$$\mathcal{F}_{n-1} \setminus \mathcal{F}_n = \{\mathbb{TT}_{n-1}, \mathbb{TC}_n\}.$$

Building on Lemma 50 we prove the following statement.

Lemma 52. *For every $\vec{\mathbb{P}}_3$-free digraph $\mathbb{D}$ the following statements are equivalent*

- $\mathbb{D} \to \mathbb{TC}_n$*, and*
- $\mathbb{D}$ *is a $\mathcal{F}_n$-free loopless oriented graph.*

Proof. The first item clearly implies the second one. We prove the converse implication by induction over n, and the base case $n = 4$ follows from Lemma 50. Let $\mathbb{D}$ be a $\vec{\mathbb{P}}_3$-free digraph which is also $\mathcal{F}_n$-free for some positive integer $n \geq 5$, and without loss of generality assume that $\mathbb{D}$ is weakly connected. Clearly, if $\mathbb{D}$ is also $\mathcal{F}_{n-1}$-free, then, by induction, we know that $\mathbb{D} \to \mathbb{TC}_{n-1}$ and so, $\mathbb{D} \to \mathbb{TC}_n$. Now assume that $\mathbb{D}$ is not $\mathcal{F}_{n-1}$-free, and let $\mathbb{F} \in \mathcal{F}_{n-1}$ be a witness of this fact. Since $\mathbb{D}$ is $\mathcal{F}_n$-free, it follows from Remark 51 that either $\mathbb{F} \cong \mathbb{TT}_{n-1}$ or $\mathbb{F} \cong \mathbb{TC}_n$. We conclude the proof by distinguishing whether $\mathbb{D}$ contains an induced copy of $\mathbb{TC}_n$.

- Suppose that $\mathbb{D}$ **does not contain** $\mathbb{TC}_n$ as an induced subgraph. In this case $\mathbb{D}$ contains a transitive tournament on $n-1$ vertices (as witnessed by $\mathbb{F}$), and $\mathbb{D}$ contains no tournament on n vertices (because $\mathbb{D}$ is $\mathcal{F}_n$-free and does not contain $\mathbb{TC}_n$). We now show that $\mathbb{D}$ is an acyclic digraph. Proceeding by contradiction assume that $\mathbb{D}$ contains a directed cycle, and using the fact that $\mathbb{D}$ is $\vec{\mathbb{P}}_3$-free notice that the shortest directed cycle in $\mathbb{D}$ is a 3-cycle $c_1c_2c_3$. Let $f_1, \ldots, f_{n-1}$ be the vertices of $\mathbb{F}$ and assume that $(f_i, f_j) \in E(\mathbb{D})$ if and only if $i < j$. We argue that there is a directed 3-cycle $\mathbb{C}$ whose vertex set intersects the vertex set of $\mathbb{F}$. If $\{c_1, c_2, c_3\} \cap \{f_1, \ldots, f_{n-1}\} \neq \varnothing$ there is nothing left to prove. Since $\mathbb{D}$ is weakly connected, there is an oriented c_1f_i-path for some $i \in [n-1]$. Consider the shortest oriented c_1f_i-path $d_1 = c_1, d_2 \ldots, d_k = f_i$ and without loss of generality assume that $d_2 \notin \{c_2, c_3\}$. We argue by finite induction that every vertex of this path belongs to a directed 3-cycle, and thus, f_i belongs to a directed 3-cycle. Suppose that there are vertices $u, v \in D$ such that $(d_i, u), (u, v), (v, d_i) \in E(\mathbb{D})$. If $(d_i, d_{i+1}) \in E(\mathbb{D})$, then $v = d_{i+1}$ or v and d_{i+1} must be adjacent since $\mathbb{D}$ is $\vec{\mathbb{P}}_3$-free. If $v = d_{i+1}$ or $(d_{i+1}, v) \in E(\mathbb{D})$, then d_i, d_{i+1}, v witness that d_{i+1} belongs to a directed 3-cycle. Otherwise, $(v, d_{i+1}) \in E(\mathbb{D})$, and since $\mathbb{D}$ is $\mathcal{F}_n$-free and $\mathbb{T}_4^b \in \mathcal{F}_n$ there is no edge (u, d_{i+1}) in $\mathbb{D}$. This implies that either $d_{i+1} = u$ or $(d_{i+1}, u) \in E(\mathbb{D})$ (because $\mathbb{D}$ is $\vec{\mathbb{P}}_3$-free), and in both cases d_{i+1} belongs to a directed 3-cycle. The case when $(d_{i+1}, d_i) \in E(\mathbb{D})$ follows with symmetric arguments. We thus conclude by finite induction that there is some vertex $f_i \in F$ that belongs to a directed 3-cycle. With similar arguments as above, one can use such a directed 3-cycle together with $\mathbb{F}$ and the fact that $\mathbb{D}$ is $\vec{\mathbb{P}}_3$-free to find a tournament on n vertices in $\mathbb{D}$. This contradicts the assumption that $\mathbb{D}$ is a $\mathcal{F}_n$-free oriented graph that does not contain $\mathbb{TC}_n$. Finally, using the observation that $\mathbb{D}$ is a $\vec{\mathbb{P}}_3$-free acyclic digraph with no tournament on n vertices we conclude that there is no homomorphism $f\colon \vec{\mathbb{P}}_n \to \mathbb{D}$. Hence, $\mathbb{D} \to \mathbb{TT}_{n-1}$ (Observation 4), and thus $\mathbb{D} \to \mathbb{TC}_n$.

- Suppose that $\mathbb{D}$ **contains a copy of** $\mathbb{TC}_n$, and let $d_1, \ldots, d_n$ be vertices such that $i \mapsto d_i$ defines an embedding of $\mathbb{TC}_n$ into $\mathbb{D}$. Let $X_1 := N^+(d_n)$, $X_n := N^-(d_1)$, and $Y := N^+(d_1)$. With analogous arguments as in the fourth itemized case of the proof of Lemma 50 it follows that $Y = N^+(d_1) = N^-(d_n)$, that every vertex in X_1 (resp. in X_n) has the same in- and out-neighbourhood as d_1 (resp. as d_n), that (X_1, Y, X_2) is a partition of the vertex set of $\mathbb{D}$, and that the subdigraph of $\mathbb{D}$ induced by Y homomorphically maps to $\mathbb{TT}_{n-2}$. Hence, by homomorphically mapping the subdigraph $\mathbb{D}[Y]$ of $\mathbb{D}$ to the subdigraph $\mathbb{TC}_n[\{2, \ldots, n-1\}]$ of $\mathbb{D}$, all vertices in X_1 to 1, and all vertices in X_n to n, we obtain a homomorphism $f\colon \mathbb{D} \to \mathbb{TC}_n$.

□

Lemma 52 implies that for every positive integer n there is a polynomial-time algorithm that solves $\mathrm{CSP}(\mathbb{TC}_n)$ when the input is restricted to $\vec{\mathbb{P}}_3$-free digraphs. Hence, the following classification follows from this lemma and from Lemma 47.

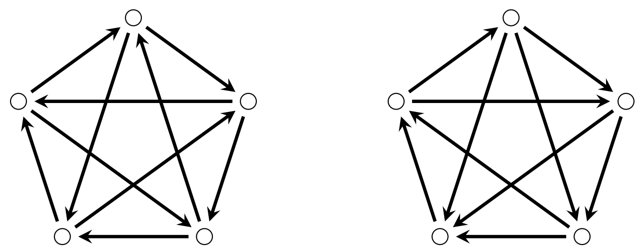
Theorem 53. *For every positive integer k and $n \geq 4$ one of the following holds*

- *$k \leq 3$ and in this case $\mathrm{CSP}(\mathbb{TC}_n)$ is tractable where the input is restricted to $\vec{\mathbb{P}}_k$-free digraphs, or*
- *$k \geq 4$ and in this case $\mathrm{CSP}(\mathbb{TC}_n)$ is NP-complete where the input is restricted to $\vec{\mathbb{P}}_k$-free digraphs.*

In the rest of this section we improve Lemma 52 by listing the (finitely many) *minimal* $\vec{\mathbb{P}}_3$-free obstructions to $\mathrm{CSP}(\mathbb{TC}_n)$. To do so, we introduce two new tournaments on five vertices depicted in Figure 11, and we prove the following lemma.

Lemma 54. *Let $\mathbb{D}$ be a $\{\mathbb{T}_4^a, \mathbb{T}_4^b\}$-free tournament. If $\mathbb{D}$ contains a pair of directed triangles with no common edge, then $\mathbb{D}$ contains a subtournament $\mathbb{D}'$ on five vertices that also contains a pair of directed triangles with no common edge.*

Proof. Let $c_1c_2c_3$ and $d_1d_2d_3$ be a pair of directed 3-cycles in $\mathbb{D}$ without a common edge. If $\{c_1, c_2, c_3\} \cap \{d_1, d_2, d_3\} \neq \varnothing$ there is nothing left to prove, so suppose that $|\{c_1, c_2, c_3, d_1, d_2, d_3\}| = 6$. Now notice that since $\mathbb{D}$ is $\{\mathbb{T}_4^a, \mathbb{T}_4^b\}$-free, c_i has at least one out-neighbour and one in-neighbour in $\{d_1, d_2, d_3\}$, and symmetrically each d_i has at least one out-neighbour and one in-neighbour in $\{c_1, c_2, c_3\}$. Without loss of generality assume that $(d_1, c_1), (d_1, c_2), (c_3, d_1) \in E(\mathbb{D})$. So, if $(d_2, c_3) \in E(\mathbb{D})$, then $d_1d_2c_3$ and $c_1c_2c_3$ are a pair of directed 3-cycles with no common edge and spanning five vertices. Otherwise, $(c_3, d_2) \in E(\mathbb{D})$, and since $(c_3, d_1) \in E(\mathbb{D})$ and $\mathbb{D}$ is $\mathbb{T}_4^a$-free, it must be the case that (d_3, c_3) is an edge of $\mathbb{D}$. In this case $c_3d_2d_3$ and $c_1c_2c_3$ are a pair of directed 3-cycles in $\mathbb{D}$ with no common edge and induction a tournament on five vertices. The claim of the lemma follows. $\square$



$\mathbb{RT}_5$ $\qquad$ $\mathbb{T}_5^c$

Figure 11: The unique $\{\mathbb{T}_4^a, \mathbb{T}_4^b\}$-free tournaments on five vertices that contain a pair of directed triangles with no a common edge (up to isomorphism).

Theorem 55. *The following statements are equivalent for every positive integer $n \geq 4$ and every $\vec{\mathbb{P}}_3$-free digraph $\mathbb{D}$.*

- $\mathbb{D} \to \mathbb{TC}_n$*, and*
- $\mathbb{D}$ *is a* $\{\mathbb{T}_4^a, \mathbb{T}_4^b, \mathbb{RT}_5, \mathbb{T}_5^c, \mathbb{TT}_n, \mathbb{TC}_{n+1}\}$*-free loopless oriented graph.*

Proof. The second itemized statement is clearly a necessary condition for the first itemized statement to hold. We show that it is also sufficient by proving the contrapositive statement: if $\mathbb{D}$ does not homomorphically map to $\mathbb{TC}_n$, then $\mathbb{D}$ is not a $\{\mathbb{T}_4^a, \mathbb{T}_4^b, \mathbb{RT}_5, \mathbb{T}_5^c, \mathbb{TT}_n, \mathbb{TC}_{n+1}\}$-free loopless oriented graph. Suppose that $\mathbb{D} \not\to \mathbb{TC}_n$, and that $\mathbb{D}$ is a loopless oriented graph. It follows from Lemma 52 that $\mathbb{D}$ is not $\mathcal{F}_n$-free — recall that $\mathcal{F}_n$ is the set of tournaments on at most $n+1$ vertices that do not embed into $\mathbb{TC}_n$. Hence, it suffices to show that if $\mathbb{D}$ is a tournament on at most $n+1$ vertices that does not embed into $\mathbb{TC}_n$, then $\mathbb{D}$ contains some tournament from $\{\mathbb{T}_4^a, \mathbb{T}_4^b, \mathbb{RT}_5, \mathbb{T}_5^c, \mathbb{TT}_n, \mathbb{TC}_{n+1}\}$. Since both tournaments on three vertices embed into $\mathbb{TC}_n$ for every $n \geq 4$, it follows that $\mathbb{D}$ has at least 4 vertices. If $\mathbb{D}$ contains no directed triangle, then $\mathbb{D}$ is a transitive tournament, and since every transitive tournament on at most $n-1$ vertices embeds into $\mathbb{TC}_n$, it must be the case that $\mathbb{D}$ contains a $\mathbb{TT}_n$. Now suppose that $\mathbb{D}$ contains a directed triangle, and an edge (u, v) such that all directed 3-cycles of $\mathbb{D}$ contain the edge (u, v). It is not hard to notice that in this case $\mathbb{D}$ is isomorphic to $\mathbb{TC}_m$ for some positive integer $m \geq 4$. Since $\mathbb{TC}_k$ is an induced subtournament of $\mathbb{TC}_n$ whenever $k \leq n$, it follows that $m \geq n+1$, and thus $\mathbb{D} \cong \mathbb{TC}_{n+1}$ (because $|D| \leq n+1$). Finally, suppose that $\mathbb{D}$ contains a pair of directed 3-cycles with no common edge. Also assume that $\mathbb{D}$ is $\{\mathbb{T}_4^a, \mathbb{T}_4^b\}$-free (otherwise there is nothing left to prove). Hence, by Lemma 54, $\mathbb{D}$ contains a subtournament $\mathbb{D}'$ on five vertices that contains two directed triangles with no common edge. The claim now follows because the only two such tournament on five vertices which are also $\{\mathbb{T}_4^a, \mathbb{T}_4^b\}$-free are $\mathbb{RT}_5$ and $\mathbb{T}_5^c$. $\square$

9 Omitting a single digraph

As mentioned in the introduction, the long-term question of the research line introduced in this paper is the following.

Question 4. *Is there a* P *versus* NP*-complete dichotomy of* CSP($\mathbb{H}$) *where the input is restricted*

1. *to* $\mathbb{F}$*-free digraphs?*
2. *to* $\mathbb{F}$*-subgraph-free digraphs?*

Having settled Question 1 for the digraphs on three vertices and for the family of tournaments $\mathbb{TC}_n$, a natural next step is tackling Question 4 for these digraphs. Regarding digraphs $\mathbb{H}$ on three vertices, we leave Question 4 (2) open for $\mathbb{K}_3$ and $\vec{\mathbb{C}}_3^{++}$, and notice that (1) has a simple solution in these cases (and $\mathbb{F}$ connected).

Corollary 56. *For every connected digraph* $\mathbb{F}$ *the following statements hold.*

- *Either* $\mathbb{F} \cong \mathbb{TT}_2$*, then* CSP($\vec{\mathbb{C}}_3^{++}$) *is polynomial-time solvable for* $\mathbb{F}$*-free digraphs, and* CSP($\mathbb{K}_3$) *is* NP*-hard for* $\mathbb{F}$*-free digraphs, or*
- *otherwise,* CSP($\vec{\mathbb{C}}_3^{++}$) *and* CSP($\mathbb{K}_3$) *are* NP*-hard for* $\mathbb{F}$*-free digraphs.*

Proof. The case of CSP($\mathbb{K}_3$) is straightforward to observe. The first itemized claim for $\vec{\mathbb{C}}_3^{++}$ follows from Theorem 46, and the second one by Sparse Incomparability for $\mathbb{F}$ not an oriented forest, and by Lemma 45 when $\mathbb{F}$ is an oriented forest. □

In the remaining of this section we see that some of our proof already yield the first steps for settling Question 4 for $\vec{\mathbb{C}}_3^{+}$ and the family of tournaments. The main result in this direction being the following one.

Theorem 57. *For every positive integer* n *and every digraph* $\mathbb{F}$ *which is not a disjoint union of oriented paths the following statements hold.*

- CSP($\vec{\mathbb{C}}_3^{+}$) *is* NP*-hard even when the input is restricted to* $\mathbb{F}$*-subgraph-free digraphs.*
- CSP($\mathbb{TC}_n$) *is* NP*-hard even when the input is restricted to* $\mathbb{F}$*-subgraph-free digraphs.*

Proof. If $\mathbb{F}$ contains some vertex v such that $d^+(v) + d^-(v) \geq 3$, then $\mathbb{F}$ contains an orientation of $\mathbb{K}_{1,3}$, and so the itemized claims follow from Lemma 43 and Lemma 47, respectively. In particular, this proves both claims for oriented forests. If $\mathbb{F}$ is not an oriented forest, i.e., it contains an oriented cycle, then both itemized statements follows from Sparse Incomparability Lemma (Corollary 6). □

The following statement asserts that the CSPs of $\mathbb{TC}_n$ and of $\vec{\mathbb{C}}_3^{+}$ remain NP-hard when restricted $\mathbb{P}$-subgraph-free digraphs whenever $\mathbb{P}$ is a path that contains two pairs of consecutive edges oriented in the same direction.

Proposition 58. *The following statements hold for every connected digraph* $\mathbb{F}$ *that contains* $\vec{\mathbb{P}}_3 + \vec{\mathbb{P}}_3$ *as a subgraph.*

- CSP($\vec{\mathbb{C}}_3^{+}$) *is* NP*-hard even when the input is restricted to* $\mathbb{F}$*-subgraph-free digraphs.*

- $\mathrm{CSP}(\mathbb{TC}_n)$ *is* NP*-hard even when the input is restricted to* $\mathbb{F}$*-subgraph-free digraphs.*

Proof. We make a reduction from 1-IN-3-SAT similar to how we did in the proofs of Lemmas 43 and 47. We use the gadget $\exists w E(w,x) \wedge E(w,y)$, which defines an equivalence relation on both these digraphs, to ensure that the $\vec{\mathbb{P}}_3$s (which only exist in the clause gadgts) are further apart than they are in $\mathbb{F}$. We do this by placing $\exists w E(w,x) \wedge E(w,y)$ gadgets end-to-end as they run into the clause gadgets. If the graph underlying $\mathbb{F}$ is of diameter d, then it will suffice to add d copies of this gadget in series at the points at which the variable gadgets meet the clause gadgets. $\square$

Consider a word $w \in \{\leftarrow, \rightarrow\}^*$, i.e., a sequence $w := w_1 \dots w_n$ where $w_i \in \{\leftarrow, \rightarrow\}$ for each $i \in [n]$. We denote by $\mathbb{P}^w_{n+1}$ the oriented path with vertex set $[n+1]$ where there is an edge $(i, i+i)$ if $w_i = \rightarrow$, and there is an edge $(i+1, i)$ if $w_i = \leftarrow$. In particular, $\vec{\mathbb{P}}_n = \mathbb{P}^w_n$ where w is the constant word on $n-1$ letters $\rightarrow$.

Corollary 59. *The following statements hold for every digraph* $\mathbb{F}$ *that contains (as a subgraph) the oriented path* $\mathbb{P}_7^{(\leftarrow\rightarrow)^3}$*, the path* $\mathbb{P}_7^{(\rightarrow\leftarrow)^3}$*.*

- $\mathrm{CSP}(\vec{\mathbb{C}}_3^+)$ *is* NP*-hard even when the input is restricted to* $\mathbb{F}$*-subgraph-free digraphs.*
- $\mathrm{CSP}(\mathbb{TC}_n)$ *is* NP*-hard even when the input is restricted to* $\mathbb{F}$*-subgraph-free digraphs for every* $n \geq 4$*.*

Proof. We consider the case when $\mathbb{F}$ contains $\mathbb{P}_7^{(\leftarrow\rightarrow)^3}$ as a subgraph, en the remaining one follows dually. For (1) (resp. for (2)) We use the same proof as for Lemma 43 (resp. for Lemma 47) except that we change the variable gadget such that it is no longer built from an undirected cycle of length n_v but rather identify all vertices that correspond to the same variable. $\square$

Corollary 60. *The following statements hold for every connected digraph* $\mathbb{F}$ *on at least* 12 *vertices.*

- $\mathrm{CSP}(\vec{\mathbb{C}}_3^+)$ *is* NP*-hard even when the input is restricted to* $\mathbb{F}$*-subgraph-free digraphs.*
- $\mathrm{CSP}(\mathbb{TC}_n)$ *is* NP*-hard even when the input is restricted to* $\mathbb{F}$*-subgraph-free digraphs for every* $n \geq 4$*.*

Proof. By Theorem 57 and Proposition 58, it suffices to prove the claim for $2\vec{\mathbb{P}}_3$-subgraph-free paths, and by Lemmas 43 and 47 we also assume that $\mathbb{F}$ is $\vec{\mathbb{P}}_5$-subgraph-free. Notice that, up to isomorphism, such a path is a subpath of $\mathbb{P}_{2n+2m+4}^{(\rightarrow\leftarrow)^n\rightarrow\rightarrow(\leftarrow\rightarrow)^m}$, or of $\mathbb{P}_{2n+2m+3}^{(\rightarrow\leftarrow)^n\rightarrow\rightarrow(\leftarrow\rightarrow)^m}$. Finally, any such path on at least 12 vertices contains either $\mathbb{P}_7^{(\leftarrow\rightarrow)^3}$ or $\mathbb{P}_7^{(\rightarrow\leftarrow)^3}$, so we conclude via Corollary 59. $\square$

Forbidden paths on three vertices

Notice that if $\mathbb{D}$ is an oriented graph with no directed path on three vertices, then $\mathbb{D} \to \mathbb{TT}_2$ (Observation 4). Also, if $\mathbb{D}$ contains a symmetric pair of edges $(u,v), (v,u)$, then the subgraph with vertices u, v is a connected component of $\mathbb{D}$. With these simple arguments one can notice that for any digraph $\mathbb{H}$, the problem $\mathrm{CSP}(\mathbb{H})$ is in P when the input is restricted to $\vec{\mathbb{P}}_3$-subgraph-free digraphs.

Observation 61. *For every digraph $\mathbb{H}$,* CSP($\mathbb{H}$) *is in* P *for* $\vec{\mathbb{P}}_3$*-subgraph-free digraphs.*

Hell and Mishra [30] proved that, for any digraph $\mathbb{H}$, the problem CSP($\mathbb{H}$) is polynomial-time solvable where the input is a $\vec{\mathbb{P}}_3^{\leftarrow\rightarrow}$-subgraph-free or a $\vec{\mathbb{P}}_3^{\rightarrow\leftarrow}$-subgraph-free digraphs. Here we briefly argue that if $\mathbb{H}$ is an oriented graph, then CSP($\mathbb{H}$) is polynomial-time solvable when the input is a $\vec{\mathbb{P}}_3^{\leftarrow\rightarrow}$-free digraph.

A *tree decomposition* for a digraph $\mathbb{G} = (G, E)$ is a pair $(\mathbb{T}, X)$ where $\mathbb{T}$ is a tree and X consists of subsets of vertices from G which we call bags. Each node of $\mathbb{T}$ corresponds to a single bag of X. For each vertex $v \in G$ the nodes of $\mathbb{T}$ containing v must induce a non-empty connected subgraph of $\mathbb{T}$ and for each edge $(u, v) \in E$, there must be at least one bag containing both u and v. We can then define the *width* of $(\mathbb{T}, X)$ to be one less than the size of the largest bag. From this, the *treewidth* of a digraph, $tw(G)$, is the minimum width of any *tree decomposition*.

Lemma 62. *If $\mathbb{H}$ is a finite oriented graph, then* CSP($\mathbb{H}$) *is in* P *for both the class of $\vec{\mathbb{P}}_3^{\leftarrow\rightarrow}$-free digraphs and the class of $\vec{\mathbb{P}}_3^{\rightarrow\leftarrow}$-free digraphs.*

Proof. We make the argument for $\vec{\mathbb{P}}_3^{\leftarrow\rightarrow}$-free graphs, the other case is dual. We also assume that the input digraph $\mathbb{D}$ is an oriented graph (otherwise, we immediately reject because $\mathbb{H}$ is an oriented graph). Let $|H| = m$, let $\mathcal{C}_{m+1}$ be the class of digraphs that are $\vec{\mathbb{P}}_3^{\leftarrow\rightarrow}$-free and further do not contain a tournament of size $m + 1$.

Let $\mathbb{G}$ be the symmetric closure of $\mathbb{D}$, equivalently the undirected graph underlying $\mathbb{D}$. Clearly, $\mathbb{G}$ admits an $\vec{\mathbb{P}}_3^{\leftarrow\rightarrow}$-free orientation (namely, $\mathbb{D}$), which is exactly the definition of 1-perfectly orientable graphs from [29]. According to Theorem 6.3 in [29], this class of graphs is $\mathbb{K}_{2,3}$-induced-minor-free. If $\mathcal{B}$ is a graph class that is $\mathbb{K}_{2,3}$-induced-minor-free, then, for each n, there is $f(n)$ so that every element of $\mathcal{B}$ either contains an n-clique or has treewidth bounded by $f(n)$. This is part of Corollary 4.12 in [19]. It follows that $\mathcal{C}_{m+1}$ has bounded treewidth, and we may plainly assume our input $\mathbb{D}$ is in $\mathcal{C}_{m+1}$ by preprocessing out some no-instances that contain a large oriented clique. That CSP($\mathbb{H}$) can be solved in polynomial time on instances of bounded treewidth is known from [21]. $\square$

10 Conclusion and outlook

In this paper we have brought together homomorphisms, digraphs and $\mathbb{H}$-(subgraph-)free algorithmics. In doing so, we have uncovered a series of results concerning not only restricted CSPs, but also hardness of digraph CSPs under natural restrictions such as acyclicity. Our work raises numerous open problems, which we believe deserve attention in the future. Besides Questions 1 and 4, we ask the following.

- Is it true that for every finite structure $\mathbb{A}$ and every (possibly infinite) $\mathbb{B}$, if $(\mathbb{A}, \mathbb{B})$ does not rpp-construct $(\mathbb{K}_3, \mathbb{L})$, then RCSP($\mathbb{A}, \mathbb{B}$) is polynomial-time solvable? (Compare to Theorem 21).
- Is it true that for every finite structure $\mathbb{A}$ and every (possibly infinite) $\mathbb{B}$ the problem RCSP($\mathbb{A}, \mathbb{B}$) is either in P or NP-hard? (Compare to Theorem 21).
- Let $\mathbb{A}$ be a finite structure and $\mathbb{B}$ a structure whose CSP is in GMSNP. Is it true that if $(\mathbb{A}, \mathbb{B})$ does not rpp-construct $(\mathbb{K}_3, \mathbb{L})$, then RCSP($\mathbb{A}, \mathbb{B}$) is polynomial-time solvable? (Compare to Theorem 23).

Constraint (e.g. $\mathbb{F} :=$)	$\mathrm{CSP}(\mathbb{TC}_n)$ on $\mathbb{F}$-subgraph-free	$\mathrm{CSP}(\mathbb{TC}_n)$ on $\mathbb{F}$-free
$\mathbb{F}$ is not an oriented tree	NP-complete Sparse Incomparability (Corollary 7)	
$\mathbb{F}$ is not an oriented path	NP-complete (Theorem 57)	
$\mathbb{F}$ contains $\vec{\mathbb{P}}_3 + \vec{\mathbb{P}}_3$ as a subgraph	NP-complete (Proposition 58)	
$\lvert F\rvert \geq 12$	NP-complete (Corollary 60)	
$\leftrightarrow$	NP-complete [2]	NP-complete [2]
$\leftarrow\rightarrow$	P [30, Lemma 1]	P (Lemma 62)
$\rightarrow\leftarrow$	P [30, Lemma 1]	P (Lemma 62)
$\rightarrow\rightarrow$	P (Theorem 49)	P (Theorem 55)
$\rightarrow\rightarrow\rightarrow$	P (Theorem 49)	NP-complete (Theorem 55)
$\rightarrow\rightarrow\rightarrow\rightarrow$	NP-complete (Theorem 49)	NP-complete (Theorem 55)
$\rightarrow\rightarrow\leftarrow\leftarrow$	NP-complete (Lemma 47)	
$\leftarrow\leftarrow\rightarrow\rightarrow$	NP-complete (Lemma 47)	
$(\leftarrow\rightarrow)^3$	NP-complete (Corollary 59)	
$(\rightarrow\leftarrow)^3$	NP-complete (Corollary 59)	

Table 1: Complexity landscape for $\mathrm{CSP}(\mathbb{TC}_n)$ under the omission of single connected subgraph or induced connected subgraph.

- Since $\mathrm{CSP}(\mathbb{H})$ reduces to $\mathrm{CSP}(\mathbb{H})$ restricted to acyclic digraphs (Theorem 34), and the latter is polynomial-time equivalent to $\mathrm{CSP}(\mathbb{H} \times \mathbb{Q})$ we ask: is it true that for every finite digraph $\mathbb{H}$ the (infinite) digraph $\mathbb{Q} \times \mathbb{H}$ pp-constructs $\mathbb{H}$?
- Is is true that for every oriented graph $\mathbb{H}$ there a finitely many $\vec{\mathbb{P}}_3$-free minimal obstructions to $\mathrm{CSP}(\mathbb{H})$? (Compare to Theorem 55).
- Is is true that for every tournament $\mathbb{T}$ there a finitely many $\vec{\mathbb{P}}_3$-free minimal obstructions to $\mathrm{CSP}(\mathbb{T})$? (Compare to Theorem 55).
- Settle Question 4 for digraphs on three vertices (see also Table 2).
- Settle Question 4 for tournaments, in particular, for the family of tournaments $\mathbb{TC}_n$ (see also Table 1).

Persistent structures

A natural question arising from restricted CSPs is if there are structures $\mathbb{A}$ such that $\mathrm{RCSP}(\mathbb{A}, \mathbb{B})$ is NP-hard whenever $\mathbb{B} \not\to \mathbb{A}$. We say that a structure $\mathbb{A}$ *persistently constructs* $\mathbb{K}_3$ if for every

Constraint (e.g. $\mathbb{F}$:=)	CSP($\vec{\mathbb{C}}_3^+$) on $\mathbb{F}$-subgraph-free	CSP($\vec{\mathbb{C}}_3^+$) on $\mathbb{F}$-free
$\mathbb{F}$ is not an oriented tree	NP-complete Sparse Incomparability (Corollary 7)	
$\mathbb{F}$ is not an oriented path	NP-complete (Theorem 57)	
$\mathbb{F}$ contains $\vec{\mathbb{P}}_3 + \vec{\mathbb{P}}_3$ as a subgraph	NP-complete (Proposition 58)	
$\|F\| \geq 12$	NP-complete (Corollary 60)	
$\leftarrow\rightarrow$	P [30, Lemma 1]	Open
$\rightarrow\leftarrow$	P [30, Lemma 1]	Open
$\rightarrow\rightarrow$	P (Theorem 42)	P (Theorem 46)
$\rightarrow\rightarrow\rightarrow$	P (Theorem 42)	NP-complete (Theorem 46)
$\rightarrow\rightarrow\rightarrow\rightarrow$	NP-complete (Theorem 42)	NP-complete (Theorem 46)
$\rightarrow\rightarrow\leftarrow\leftarrow$	NP-complete (Lemma 43)	
$\leftarrow\leftarrow\rightarrow\rightarrow$	NP-complete (Lemma 43)	
$(\leftarrow\rightarrow)^3$	NP-complete (Corollary 59)	
$(\leftarrow\rightarrow)^3$	NP-complete (Corollary 59)	

Table 2: Complexity landscape for CSP($\vec{\mathbb{C}}_3^+$) under the omission of single connected subgraph or induced connected subgraph.

(possibly infinite) $\mathbb{B}$ such that $\mathbb{B} \not\to \mathbb{A}$ the restricted CSP template $(\mathbb{A}, \mathbb{B})$ rpp-constructs $(\mathbb{K}_3, \mathbb{L})$. Notice that in this case RCSP$(\mathbb{A}, \mathbb{B})$ is NP-hard whenever $\mathbb{B} \not\to \mathbb{A}$.

Observation 63. *For a finite structure $\mathbb{A}$ the following statements are equivalent.*

- $(\mathbb{A}, \mathbb{B})$ *rpp-constructs* $(\mathbb{K}_3, \mathbb{L})$ *for every finite structure* $\mathbb{B} \not\to \mathbb{A}$.
- $(\mathbb{A}, \mathbb{B})$ *rpp-constructs* $(\mathbb{K}_3, \mathbb{L})$ *for every (possibly infinite) structure* $\mathbb{B} \not\to \mathbb{A}$.
- RCSP$(\mathbb{A}, \mathbb{B})$ *is* NP*-hard for every finite structure* $\mathbb{B} \to \mathbb{A}$ *(assuming* P $\neq$ NP*).*

Proof. The equivalence between the first and third statement follows from the assumption that P $\neq$ NP, and from Theorem 21. The second item clearly implies the first one. Finally, to see that the first one implies the second one, suppose $\mathbb{B} \not\to \mathbb{A}$. By compactness there is a finite substructure $\mathbb{B}'$ of $\mathbb{B}$ such that $\mathbb{B}' \not\to \mathbb{A}$. Hence, $(\mathbb{A}, \mathbb{B}')$ rpp-constructs $(\mathbb{K}_3, \mathbb{L})$, and clearly $(\mathbb{A}, \mathbb{B})$ rpp-constructs $(\mathbb{A}, \mathbb{B}')$. Since rpp-constructions compose, we conclude that $(\mathbb{A}, \mathbb{B})$ rpp-constructs $(\mathbb{K}_3, \mathbb{L})$. $\square$

Remark 64. A similar compactness as in the proof of Observation 63 applies to *ω-categorical structures* (see, e.g., [11, Lemma 4.1.7]), i.e., to structures $\mathbb{A}$ whose automorphism group has finitely many orbits of k-tuples for each positive integer k. Hence, if $\mathbb{A}$ is an ω-categorical structure, then $\mathbb{A}$ persistently constructs $\mathbb{K}_3$ if and only if $(\mathbb{A}, \mathbb{B})$ rpp-constructs $(\mathbb{K}_3, \mathbb{L})$ for every finite structure $\mathbb{B} \not\to \mathbb{A}$.

Problem 5. *Characterize the class of finite digraphs (structures) that persistently construct* $\mathbb{K}_3$.

Theorem 3 asserts that for every hereditarily hard digraph $\mathrm{RCSP}(\mathbb{H}, \mathbb{H}')$ is NP-hard whenever $\mathbb{H}'$ is a finite digraph and $\mathbb{H}' \not\to \mathbb{H}$. Hence, assuming $\mathrm{P} \neq \mathrm{NP}$, this implies that every hereditarily hard digraph persistently constructs $\mathbb{K}_3$ (Observation 63).

Problem 6. *Characterize the class of finite digraphs (structures)* $\mathbb{H}$ *such that* $\mathrm{RCSP}(\mathbb{H}, \mathbb{H}')$ *is* NP*-hard whenever* $\mathbb{H}' \not\to \mathbb{H}$ *(assuming* $\mathrm{P} \neq \mathrm{NP}$*).*

Acknowledgments

We are grateful to Martin Milanič for the proof of Lemma 62.

References

[1] J. Bang-Jensen and P. Hell. The effect of two cycles on the complexity of colourings by directed graphs. *Discrete Applied Mathematics*, 26(1):1–23, 1990.

[2] J. Bang-Jensen, P. Hell, and G. MacGillivray. The complexity of colouring by semicomplete digraphs. *SIAM Journal on Discrete Mathematics*, 1(3):281–298, 1988.

[3] J. Bang-Jensen, P. Hell, and G. MacGillivray. Hereditarily hard H-colouring problems. *Discrete Mathematics*, 138:75–92, 1995.

[4] A. Barsukov. On dichotomy above Feder and Vardi's logic, 2022. PhD thesis, École Doctorale Science pour l'Engénieur, Clermont-Ferrand.

[5] L. Barto, J. Bulín, A. A. Krokhin, and J. Opršal. Algebraic approach to promise constraint satisfaction. *J. ACM*, 68(4):28:1–28:66, 2021.

[6] L. Barto, M. Kompatscher, M. Olšák, T. V. Pham, and M. Pinsker. The equivalence of two dichotomy conjectures for infinite domain constraint satisfaction problems. In *Proceedings of the 32nd Annual ACM/IEEE Symposium on Logic in Computer Science – LICS'17*, 2017. Preprint arXiv:1612.07551.

[7] L. Barto, M. Kozik, and T. Niven. The CSP dichotomy holds for digraphs with no sources and no sinks (a positive answer to a conjecture of Bang-Jensen and Hell). *SIAM Journal on Computing*, 38(5), 2009.

[8] L. Barto, A. A. Krokhin, and R. Willard. Polymorphisms, and how to use them. In *The Constraint Satisfaction Problem: Complexity and Approximability*, pages 1–44. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2017.

[9] L. Barto, J. Opršal, and M. Pinsker. The wonderland of reflections. *Israel Journal of Mathematics*, 223(1):363–398, 2018.

[10] L. Barto and M. Pinsker. The algebraic dichotomy conjecture for infinite domain constraint satisfaction problems. In *Proceedings of the 31th Annual IEEE Symposium on Logic in Computer Science – LICS'16*, pages 615–622, 2016. Preprint arXiv:1602.04353.

[11] M. Bodirsky. *Complexity of Infinite-Domain Constraint Satisfaction*. Lecture Notes in Logic (52). Cambridge University Press, Cambridge, United Kingdom; New York, NY, 2021.

[12] M. Bodirsky and B. Bodor. Canonical polymorphisms of Ramsey structures and the unique interpolation property. In *Proceedings of the Symposium on Logic in Computer Science (LICS)*, 2021.

[13] M. Bodirsky and V. Dalmau. Datalog and constraint satisfaction with infinite templates. *Journal on Computer and System Sciences*, 79:79–100, 2013. A preliminary version appeared in the proceedings of the Symposium on Theoretical Aspects of Computer Science (STACS'05).

[14] M. Bodirsky, F. R. Madelaine, and A. Mottet. A proof of the algebraic tractability conjecture for monotone monadic SNP. *SIAM J. Comput.*, 50(4):1359–1409, 2021.

[15] F. Bonomo, M. Chudnovsky, P. Maceli, O. Schaudt, M. Stein, and M. Zhong. Three-Coloring and List Three-Coloring of Graphs Without Induced Paths on Seven Vertices. *Combinatorica*, 38:779–801, 2018.

[16] R. C. Brewster and T. Graves. On the restricted homomorphism problem. *Discrete Applied Mathematics*, 156(14):2823–2826, 2008.

[17] R. C. Brewster, P. Hell, and G. MacGillivray. The complexity of restricted graph homomorphisms. *Discrete Mathematics*, 167-168:145–154, 1997. Selected Papers 15th British Combinatorial Conference.

[18] A. A. Bulatov. A dichotomy theorem for nonuniform CSPs. In *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17*, pages 319–330, 2017.

[19] C. Dallard, M. Milanič, and K. Štorgel. Treewidth versus clique number. i. graph classes with a forbidden structure. *SIAM J. Discret. Math.*, 35(4):2618–2646, Jan. 2021.

[20] V. Dalmau and A. A. Krokhin. Majority constraints have bounded pathwidth duality. *Eur. J. Comb.*, 29(4):821–837, 2008.

[21] R. Dechter. From local to global consistency. *Artificial Intelligence*, 55(1):87–108, 1992.

[22] P. Erdős. Graph theory and probability. *Canadian Journal of Mathematics*, 11:34–38, 1959.

[23] T. Feder and M. Y. Vardi. The computational structure of monotone monadic SNP and constraint satisfaction: a study through Datalog and group theory. *SIAM Journal on Computing*, 28:57–104, 1999.

[24] J. Foniok. Homomorphisms and structural properties of relational systems. Doctoral Thesis, Charles University Prague, Faculty of Mathematics and Physics, 2007.

[25] J. Foniok, J. Nešetřil, and C. Tardif. Generalised dualities and maximal finite antichains in the homomorphism order of relational structures. *European J. Combin.*, 29(4):881–899, 2008.

[26] J. Goedgebeur, J. Jooken, K. Okrasa, P. Rzażewski, and O. Schaudt. Minimal Obstructions to C_5-Coloring in Hereditary Graph Classes. In R. Královič and A. Kučera, editors, *49th International Symposium on Mathematical Foundations of Computer Science (MFCS 2024)*, volume 306 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 55:1–55:15, Dagstuhl, Germany, 2024. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[27] P. A. Golovach, M. Johnson, D. Paulusma, and J. Song. A survey on the computational complexity of coloring graphs with forbidden subgraphs. *J. Graph Theory*, 84(4):331–363, 2017.

[28] S. Guzmán-Pro. GMSNP and finite structures. *arXiv:2406.13529*, 2024.

[29] T. R. Hartinger and M. Milanič. Partial characterizations of 1-perfectly orientable graphs. *Journal of Graph Theory*, 85(2):378–394, 2017.

[30] P. Hell and A. Mishra. H-coloring degree-bounded (acyclic) digraphs. *Theoretical Computer Science*, 554:40–49, 2014.

[31] P. Hell and J. Nešetřil. On the complexity of H-coloring. *Journal of Combinatorial Theory, Series B*, 48:92–110, 1990.

[32] S. Huang. Improved complexity results on k-coloring. *European J. Combin.*, 51:336–346, 2016.

[33] A. Krokhin, J. Opršal, M. Wrochna, and S. Živný. Topology and adjunction in promise constraint satisfaction. *SIAM J. Comput.*, 52(1):38–79, 2023.

[34] G. Kun. Constraints, MMSNP, and expander relational structures. *Combinatorica*, 33(3):335–347, 2013.

[35] J. Nešetřil and C. Tardif. Duality theorems for finite structures. *Journal of Combinatorial Theory, Series B*, 80:80–97, 2000.

[36] M. Pinsker. Current challenges in infinite-domain constraint satisfaction: Dilemmas of the infinite sheep. In *52nd IEEE International Symposium on Multiple-Valued Logic, ISMVL 2022, Dallas, TX, USA, May 18-20, 2022*, pages 80–87, 2022.

[37] A. L. Selman. Promise problems for complexity classes. *Information and Computation*, 78:87–98, 1988.

[38] D. Zhuk. A proof of the CSP dichotomy conjecture. *J. ACM*, 67(5):30:1–30:78, 2020.

[39] D. N. Zhuk. A proof of CSP dichotomy conjecture. In *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17*, pages 331–342, 2017. https://arxiv.org/abs/1704.01914.