CONSTRAINT SATISFACTION WITH COUNTING QUANTIFIERS

BARNABY MARTIN [*†], FLORENT MADELAINE [‡§], AND JURAJ STACHO [¶‖]

Abstract. We initiate the study of *constraint satisfaction problems* (CSPs) in the presence of counting quantifiers $\exists^{\geq j}$ which assert the existence of at least j elements such that the ensuing property holds. These are natural variants of CSPs in the mould of *quantified CSPs* (QCSPs). Namely, $\exists^{\geq 1} := \exists$ and $\exists^{\geq n} := \forall$ (for the domain of size n)

We observe that a single counting quantifier $\exists^{\geq j}$ strictly between $\exists$ and $\forall$ already affords the maximal possible complexity of QCSPs (which have both $\exists$ and $\forall$), namely being Pspace-complete for a suitably chosen template. Therefore, to better understand the complexity of this problem, we focus on restricted cases for which we derive the following results.

Firstly, for all subsets of counting quantifiers on clique and cycle templates, we give a full trichotomy – all such problems are in P, NP-complete or Pspace-complete.

Secondly, we consider the problem with exactly two quantifiers: $\exists^{\geq 1} := \exists$ and $\exists^{\geq j}$ $(j \neq 1)$. Such a CSP is already NP-hard on non-bipartite graph templates. We explore the situation of this generalized CSP on graph templates, giving various conditions for both tractability and hardness.

For quantifiers $\exists^{\geq 1}$ and $\exists^{\geq 2}$, we give a dichotomy for all graphs; namely, the problem is NP-hard if the graph contains a triangle or has girth at least 5, and is in P otherwise. We strengthen this result in the following two ways. For bipartite graphs, the problem is in P for forests and graphs of girth 4, and is Pspace-hard otherwise. For complete multipartite graphs, the problem is in L, NP-complete or Pspace-complete.

Finally, using counting quantifiers we solve the complexity of a concrete QCSP whose complexity was previously open.

Key words. Quantified Constraint Satisfaction, Counting Quantifiers, Retraction, Computational Complexity.

AMS subject classifications. 03B70, 68Q25, 05C85

1. Introduction.

The *constraint satisfaction problem* CSP($\mathbf{B}$), much studied in artificial intelligence, is known to admit several equivalent formulations, two of the best known of which are the query evaluation of primitive positive (pp) sentences – those involving only existential quantification and conjunction – on $\mathbf{B}$, and the homomorphism problem to $\mathbf{B}$ (see, e.g., [23]). The problem CSP($\mathbf{B}$) is NP-complete in general, and a great deal of effort has been expended in classifying its complexity for certain restricted cases. Notably it is conjectured [20, 8] that for all fixed $\mathbf{B}$, the problem CSP($\mathbf{B}$) is in P or NP-complete. While this has not been settled in general, a number of partial results are known – e.g. over structures of size at most three [34, 5] and over smooth digraphs [22, 1].

A popular generalisation of the CSP involves considering the query evaluation problem for *positive Horn* logic – involving only the two quantifiers, $\exists$ and $\forall$, together with conjunction. The resulting *Quantified Constraint Satisfaction Problem* QCSP($\mathbf{B}$) allows for a broader class, used in artificial intelligence to capture non-monotonic reasoning, whose complexities rise to Pspace-completeness.

[*]School of Science and Technology, Middlesex University, The Burroughs, Hendon, London, NW4 4BT, UK. (barnabymartin@gmail.com).

[†]*Supported by EPSRC grant EP/L005654/1.*

[‡]GREYC, CNRS UMR 6072, Université de Caen Basse Normandie, France (florent.madelaine@unicaen.fr).

[§]*Supported by ANR grant ALCOCLAN*

[¶]IEOR Department, Columbia University, 500 West 120th Street, New York, NY 10027, United States (stacho@cs.toronto.edu).

[‖]*Partially supported by NSF grant DMS-1265803, and by EPSRC grant EP/I01795X/1*

In this paper, we study counting quantifiers of the form $\exists^{\geq j}$, which allow one to assert the existence of at least j elements such that the ensuing property holds. Thus on a structure $\mathbf{B}$ with domain of size n, the quantifiers $\exists^{\geq 1}$ and $\exists^{\geq n}$ are precisely $\exists$ and $\forall$, respectively. Counting quantifiers have been extensively studied in finite model theory (see [15, 31]), where the focus is on supplementing the descriptive power of various logics. Of more general interest is the majority quantifier $\exists^{\geq n/2}$ (on a structure of domain size n), which sits broadly midway between $\exists$ and $\forall$. Majority quantifiers are studied across diverse fields of logic and have various practical applications, e.g. in cognitive appraisal and voting theory [14]. They have also been studied in computational complexity, e.g. in [24].

We study variants of CSP($\mathbf{B}$) in which the input sentence to be evaluated on $\mathbf{B}$ (of size $|B|$) remains positive conjunctive in its quantifier-free part, but is quantified by various counting quantifiers from some non-empty set. For $X \subseteq \{1, \ldots, |B|\}$, $X \neq \emptyset$, the X-CSP($\mathbf{B}$) takes as input a sentence given by a conjunction of atoms quantified by quantifiers of the form $\exists^{\geq j}$ for $j \in X$. It then asks whether this sentence is true on $\mathbf{B}$. In this fashion, the X-CSP may be seen as a natural generalisation of the CSP and QCSP.

To briefly introduce the reader to this concept, consider the following sentence:

$$\Phi_1 := \exists^{\geq 4} x_1 \, \exists^{\geq 3} x_2 \; E(x_1, x_2)$$

When E is the edge relation of a graph, this sentence asserts the existence of at least 4 vertices (x_1) of degree at least 3 (having at least 3 distinct neighbors x_2). A similar, more complicated example is as follows.

$$\Phi_2 := \exists^{\geq 1} x_1 \, \exists^{\geq 2} x_2 \, \exists^{\geq 2} x_3 \; E(x_1, x_2) \wedge E(x_1, x_3) \wedge E(x_2, x_3)$$

Here, the formula asks for a pair adjacent edges ($x_1 x_2$) that each belong to at least 2 triangles (x_1, x_2, x_3). Note that the sentence Φ_2 uses quantifiers $\exists^{\geq 1}$ and $\exists^{\geq 2}$. Thus Φ_2 is an input of $\{1,2\}$-CSP($\mathbf{B}$) while Φ_1 is an input of $\{3,4\}$-CSP($\mathbf{B}$), or of any X-CSP($\mathbf{B}$) where $\{3,4\} \subseteq X$. Simple formulas like this are, of course, easy to evaluate. The most challenging task is when both the formula and the structure are given as input, without any restrictions. Not surprisingly deciding the truth value of an arbitrary formula on an arbitrary structure is difficult, short of trying all possible assignments. The standard approach to this is to assume that either the formula or the structure is fixed. Here, we consider the latter.

The counting quantifiers allow us to explore an exponentially-sized tree of assignments with a linear number of variables. In this manner, they dynamically model fault-tolerance because the $\exists^{\geq j}$ quantifier asserts that even if $j-1$ elements are faulty then still there is one that works.

1.1. Summary of results. Throughout the paper, we study the X-CSP($\mathbf{B}$) problem for various restrictions on the set X and the structure $\mathbf{B}$. We often concentrate on graph structures.

First, in §3, we consider the power of a single quantifier $\exists^{\geq j}$, for some j. We prove that for each $n \geq 3$, there is a template $\mathbf{B}_n$ of size n, such that $\exists^{\geq j}$ ($1 < j < n$) already has the full complexity of QCSP, i.e., $\{j\}$-CSP($\mathbf{B}_n$) is Pspace-complete.

Then, in §4 and §5, we study the complexity of all possible subsets of counting quantifiers on clique and cycle templates, $\mathbf{K}_n$ and $\mathbf{C}_n$, respectively. We derive the following classification theorems.

Theorem 1.1. *For $n \in \mathbb{N}$ and $X \subseteq \{1, \ldots, n\}$, the problem X-CSP($\mathbf{K}_n$) is*

(i) *in* L *if $n \leq 2$ or $X \subseteq \{\lfloor n/2 \rfloor + 1, \ldots, n\}$,*
(ii) *in* P *if $n = 4$ and $X = \{2\}$,*
(iii) NP-*complete if $n \geq 3$ and $X = \{1\}$,*

(iv) Pspace-*complete in all other cases.*

THEOREM 1.2. *For* $n \geq 3$ *and* $X \subseteq \{1, \dots, n\}$*, the problem* X-CSP($\mathbf{C}_n$) *is*
(i) *in* L *if* $n = 4$*, or* $1 \notin X$*, or* n *is even and* $X \cap \{2, \dots, n/2\} = \emptyset$*,*
(ii) NP-*complete if* n *is odd and* $X = \{1\}$*,*
(iii) Pspace-*complete in all other cases.*

Next, in §6, we consider $\{1, j\}$-CSP($\mathbf{H}$), for $j \neq 1$ on graphs. The CSP($\mathbf{H}$) (in our language $\{1\}$-CSP($\mathbf{H}$)) is already NP-hard for non-bipartite graphs $\mathbf{H}$ [22]. We therefore explore the complexity of $\{1, j\}$-CSP($\mathbf{H}$) on bipartite graph templates, giving various conditions for both tractability and hardness, using and extending results of §4 and §5. We are most interested here in the distinction between P and NP-hard. This choice seems unavoidable – to understand precisely which cases are Pspace-complete would include as a subclassification the Pspace-complete cases of QCSP($\mathbf{H}$), a challenging question which has been open for some time [29].

We first easily observe that $\{1, 2\}$-CSP($\mathbf{H}$) is in L if $\mathbf{H}$ is a bipartite graph containing a 4-cycle; in fact, in this case the problem coincides with $\{1\}$-CSP($\mathbf{H}$). For bipartite graphs of larger girth, we extend Theorem 1.2 to prove Pspace-hardness of the problem. The remaining cases are when $\mathbf{H}$ contains no cycle, i.e., $\mathbf{H}$ is a forest. For trees, we describe a polynomial-time algorithm. Notably, it turns out in §7 that an algorithm for trees is directly linked to an algorithm for paths in the following sense.

COROLLARY 1.3. *Let* $\mathbf{T}$ *be a tree, and let* $\mathbf{P}$ *be a longest path in* $\mathbf{T}$*. Then* Ψ *is a yes-instance of* $\{1, 2\}$-CSP($\mathbf{T}$) *if and only if* Ψ *is a yes-instance of* $\{1, 2\}$-CSP($\mathbf{P}$)*.*

Our algorithm for trees is based on a characterization of yes-instances. We describe two particular obstructions, both of which take the form of a special walk, where the presence or absence of this walk determines the answer to the problem.

This naturally extends to forests which completes the classification. Combined with [22] and §8, this leads to the following dichotomy theorem which can be seen as a companion to the celebrated result of Hell and Nešetřil.

THEOREM 1.4. *Let* $\mathbf{H}$ *be a graph. Then* $\{1, 2\}$-CSP($\mathbf{H}$) *is in* P *if* $\mathbf{H}$ *is a forest or is a bipartite graph with a* 4*-cycle; the problem is* NP-*hard in all other cases.*

Let $\mathbf{H}$ *be a bipartite graph. Then* $\{1, 2\}$-CSP($\mathbf{H}$) *is in* P *if* $\mathbf{H}$ *is a forest or has a* 4*-cycle; the problem is* Pspace-*complete in all other cases.*

Note that the second part of this result can not be reproduced for non-bipartite graphs in general, since there exist NP-complete cases, such as when $\mathbf{H}$ is the octahedron $\mathbf{K}_{2,2,2}$. The situation regarding exactly which cases are NP-complete is less clear. We investigate this issue and present some further partial results in this direction in §9. For complete multipartite graphs, we give a full trichotomy that can be seen as a final *coup de grâce*.

THEOREM 1.5. *For integers* $1 \leq a_1 \leq \ldots \leq a_n$*, let* $\mathbf{K}_{a_1,\ldots,a_n}$ *denote the complete multipartite graph where parts have sizes* $a_1, \dots, a_n$*. Then* $\{1, 2\}$-CSP($\mathbf{K}_{a_1,\ldots,a_n}$) *is*
(i) *in* L *if* $n = 2$*,*
(ii) NP-*complete if* $n \geq 3$ *and* $a_2 \geq 2$*,*
(iii) Pspace-*complete in all other cases.*

We close the paper in §10 where we use counting quantifiers to solve the complexity of QCSP for the reflexive 4-cycle. The complexity of this problem was previously open. Some closing remarks and open problems are then discussed in §11.

1.2. Related work. This paper is the full version of the conference reports [25] and [26] in both of which most proofs were omitted. In addition, Theorem 1.5 is completely new. The idea to study $\{1,\ldots,|B|\}$-CSP($\mathbf{B}$) comes from Andrei Krokhin, but similar motivations are behind the paper [6] of which we were not aware when we wrote [25]. The paper [6] should be seen as complementary to this and makes inquiry into counting quantifiers that is orthogonal to ours. We will return to discuss its algebraic ideas in the conclusion. Many of our NP- and Pspace-hardness proofs are elaborate re-workings of those NP-hardness proofs given for the retraction problem by Feder, Hell and Huang in [17, 18]. The Pspace-hardness for Theorem 1.5 borrows from the Pspace-hardness proof of [3].

2. Preliminaries. Let $\mathbf{B}$ be a finite structure over a finite signature σ whose domain B is of cardinality $|B|$. In the case of graphs $\mathbf{G}$ we will use the more common domain (vertex) notation $V(\mathbf{G})$, as well as $E(\mathbf{G})$ for the edge set. In general, sets are denoted in upper case and elements of sets in lower case. Boldface is reserved for structures (walks and paths in graphs will mostly not be designated in boldface unless we wish to emphasize them as graphs in their own right).

For $1 \leq j \leq |B|$, the formula $\exists^{\geq j}x\ \phi(x)$ with *counting quantifier* should be interpreted on $\mathbf{B}$ as stating that there exist at least j distinct elements $b \in B$ such that $\mathbf{B} \models \phi(b)$. Counting quantifiers generalize existential ($\exists := \exists^{\geq 1}$), universal ($\forall := \exists^{\geq |B|}$) and (weak) majority ($\exists^{\geq |B|/2}$) quantifiers. Counting quantifiers do not in general commute with themselves, viz $\exists^{\geq j}x\exists^{\geq j}y \neq \exists^{\geq j}y\exists^{\geq j}x$. For an example of this consider $\exists^{\geq 2}x\exists^{\geq 2}y\exists^{\geq 2}z\ E(x,y)\wedge E(x,z)$ and $\exists^{\geq 2}y\exists^{\geq 2}z\exists^{\geq 2}x\ E(x,y)\wedge E(x,z)$ on the disjoint union of two paths $\mathbf{P}_3$ of length 2. By contrast, $\exists$ and $\forall$ do commute with themselves, if not with one another. The counting quantifiers are first-order definable in $\exists$, but not in primitive positive logic.

For $\emptyset \neq X \subseteq \{1,\ldots,|B|\}$, the X-CSP($\mathbf{B}$) takes as input a sentence of the form $\Phi := Q_1x_1Q_2x_2\ldots Q_mx_m\ \phi(x_1,x_2,\ldots,x_m)$, where ϕ is a conjunction of positive atoms of σ and each Q_i is of the form $\exists^{\geq j}$ for some $j \in X$. The set of such sentences forms the logic X-pp (recall the pp is primitive positive). The yes-instances are those for which $\mathbf{B} \models \Phi$. Note that all problems X-CSP($\mathbf{B}$) are trivially in Pspace, by cycling through all possible evaluations for the variables. The problem $\{1\}$-CSP($\mathbf{B}$) is better-known as just CSP($\mathbf{B}$), and $\{1,|B|\}$-CSP($\mathbf{B}$) is better-known as QCSP($\mathbf{B}$).

A homomorphism from $\mathbf{A}$ to $\mathbf{B}$, both σ-structures, is a function $h : A \to B$ such that $(a_1,\ldots,a_r) \in R^{\mathbf{A}}$ implies $(h(a_1),\ldots,h(a_r)) \in R^{\mathbf{B}}$, for all relations R of σ. A frequent role will be played by the *retraction* problem Ret($\mathbf{B}$) in which one is given a structure $\mathbf{A}$ containing $\mathbf{B}$, and one is asked if there is a homomorphism from $\mathbf{A}$ to $\mathbf{A}$ that is the identity on $\mathbf{B}$. It is well-known that retraction problems are special instances of CSPs in which the constants of the template are all named [17].

In line with convention we consider the notion of hardness reduction in proofs to be polynomial many-to-one (though logspace is sufficient for our results).

2.1. Game characterization. There is a simple game characterization for the truth of sentences of the logic X-pp on a structure $\mathbf{B}$. Given a sentence Ψ of X-pp, and a structure $\mathbf{B}$, we define the following game $\mathscr{G}(\Psi,\mathbf{B})$. Let $\Psi := Q_1x_1Q_2x_2\ldots Q_mx_m$ $\psi(x_1,x_2,\ldots,x_m)$. Working from the outside in, coming to a quantified variable $\exists^{\geq j}x$, the *Prover* (female) picks a subset B_x of j elements of B as witnesses for x, and an *Adversary* (male) chooses one of these, say b_x, to be the value of x. Prover wins iff $\mathbf{B} \models \psi(b_{x_1},b_{x_2},\ldots,b_{x_m})$. The following comes immediately from the definitions.

Lemma 2.1. *Prover has a winning strategy in the game* $\mathscr{G}(\Psi,\mathbf{B})$ *iff* $\mathbf{B} \models \Psi$.

We will often move seemlessly between the two characterizations of Lemma 2.1.

One may alternatively view the game in the language of homomorphisms. There is an obvious bijection between σ-structures with domain $\{1,\dots,m\}$ and conjunctions of positive atoms in variables $\{v_1,\dots,v_m\}$. From a structure $\mathbf{B}$ build the conjunction $\phi_{\mathbf{B}}$ listing the tuples that hold on $\mathbf{B}$ in which element i corresponds to variable v_i. Likewise, for a conjunction of positive atoms ψ, let $\mathbf{D}_\psi$ be the structure whose relation tuples are listed by ψ, where variable v_i corresponds to element i. The relationship of $\mathbf{B}$ to $\phi_{\mathbf{B}}$ and ψ to $\mathbf{D}_\psi$ is very similar to that of *canonical query* and *canonical database* (see [23]), except there we consider the conjunctions of atoms to be existentially quantified. For example, $\mathbf{K}_3$ on domain $\{1,2,3\}$ gives rise to

$$\phi_{\mathbf{K}_3} := \exists v_1, v_2, v_3\ E(v_1,v_2)\wedge E(v_2,v_1)\wedge E(v_2,v_3)\wedge E(v_3,v_2)\wedge E(v_3,v_1)\wedge E(v_3,v_1).$$

The Prover-Adversary game $\mathscr{G}(\Psi,\mathbf{B})$ may be seen as Prover giving j potential maps for element x in $\mathbf{D}_\psi$ (ψ is quantifier-free part of Ψ) and Adversary choosing one of them. The winning condition for Prover is now that the map given from $\mathbf{D}_\psi$ to $\mathbf{B}$ is a homomorphism. We denote by $\prec$ the total order of variables of Ψ as they are quantified in the formula (from left to right).

In the case of QCSP, i.e., $\{1,|B|\}$-pp, each move of a game $\mathscr{G}(\Psi,\mathbf{B})$ is trivial for one of the players. For $\exists^{\geq 1}$ quantifiers, Prover gives a singleton set, so Adversary's choice is forced. In the case of $\exists^{\geq |B|}$ quantifiers, Prover must advance all of B. Thus, essentially, Prover alone plays $\exists^{\geq 1}$ quantifiers and Adversary alone plays $\exists^{\geq |B|}$ quantifiers.

3. Complexity of a single quantifier. In this section we consider the complexity of evaluating X-pp sentences when X is a singleton, i.e., we have at our disposal only a single quantifier.

THEOREM 3.1.

(i) *$\{1\}$-CSP($\mathbf{B}$) is in* NP *for all $\mathbf{B}$. For each $n\geq 2$, there exists a template $\mathbf{B}_n$ of size n such that $\{1\}$-CSP($\mathbf{B}_n$) is* NP*-complete.*

(ii) *$\{|B|\}$-CSP($\mathbf{B}$) is in* L *for all $\mathbf{B}$.*

(iii) *For each $n\geq 3$, there exists a template $\mathbf{B}_n$ of size n such that $\{j\}$-CSP($\mathbf{B}_n$) is* Pspace*-complete for all $1<j<n$.*

Proof. Parts (i) and (ii) are well-known (see [32], resp. [27]). For (iii), let $\mathbf{B}_{\mathrm{NAE}}$ be the Boolean structure on domain $\{0,1\}$ with a single ternary not-all-equal relation $R_{\mathrm{NAE}} := \{0,1\}^3\setminus\{(0,0,0),(1,1,1)\}$. To show Pspace-completeness, we reduce from QCSP($\mathbf{B}_{\mathrm{NAE}}$), the *quantified not-all-equal-3-satisfiability* (see [32]).

We distinguish two cases.

Case I: $j\leq\lfloor n/2\rfloor$. Define $\mathbf{B}_n$ on domain $\{0,\dots,n-1\}$ with a single unary relation U and a single ternary relation R. Set $U:=\{0,\dots,j-1\}$ and set

$$R:=\{0,\dots,n-1\}^3\setminus\{(a,b,c):\ a,b,c \text{ either all odd or all even}\}.$$

The even numbers will play the role of false 0 and odd numbers the role of true 1.

Case II: $j>\lfloor n/2\rfloor$. Define $\mathbf{B}_n$ on domain $\{0,\dots,n-1\}$ with a single unary relation U and a single ternary relation R. Set $U:=\{0,\dots,j-1\}$ and set

$$R:=\{0,\dots,n-1\}^3\setminus\{(a,b,c):\ a,b,c\leq n-j \text{ and either all odd or all even}\}.$$

In this case even numbers $\leq n-j$ play the role of false 0 and odd numbers $\leq n-j$ play the role of true 1. The $j-1$ numbers $n-j+1,\dots,n-1$ are somehow universal and will always satisfy any R relation.

The reduction we use is the same for Cases I and II. We reduce QCSP($\mathbf{B}_{\mathrm{NAE}}$) to $\{j\}$-CSP($\mathbf{B}_n$). Given an input $\Psi := Q_1x_1Q_2x_2\dots Q_mx_m\ \psi(x_1,x_2,\dots,x_m)$ to the former (i.e., each Q_i is $\exists$ or $\forall$) we build an instance Ψ' for the latter. From the outside in, we convert quantifiers $\exists x$ to $\exists^{\geq j} x$. For quantifiers $\forall x$, we convert also to $\exists^{\geq j} x$, but we add the conjunct $U(x)$ to the quantifier-free part ψ.

We claim $\mathbf{B}_{\mathrm{NAE}} \models \Psi$ iff $\mathbf{B}_n \models \Psi'$. For the $\exists$ variables of Ψ, we can see that any j witnesses from the domain B_n for $\exists^{\geq j}$ must include some element playing the role of either false 0 or true 1 (and the other $j-1$ may always be found somewhere). For the $\forall$ variables of Ψ, the relation U forces us to choose both 0 and 1 among the $\exists^{\geq j}$ (and the other $j-2$ will come from $2, \ldots, j-1$). The result follows. ∎

4. Cliques: proof of Theorem 1.1. Recall that $\mathbf{K}_n$ is the complete irreflexive graph on n vertices. We will sometimes refer to edges $E(i,j)$ as simply ij.

We discuss the cases of Theorem 1.1 individually. For $n \leq 2$, the results follow from [29], and (iii) is proved in [22]. The remainder of (i) will be proved as Proposition 4.7 in §4.3 while (iv) will be given as Theorem 4.3 in §4.1, and Corollary 4.5 and Proposition 4.6 in §4.2. Finally, (ii) will be proved in §4.4 as Theorem 4.9.

4.1. Pspace-completeness of $\{n\}$-CSP($\mathbf{K}_{2n}$) when $n \geq 3$. The proof is by reduction from QCSP($\mathbf{K}_n$), the *quantified n-coloring* problem, which is known to be Pspace-complete [3, 4]. The template $\mathbf{K}_n$ consists of vertices $\{1, 2, \ldots, n\}$ and all possible edges between distinct vertices. We shall call these vertices *colors*. We describe a reduction from QCSP($\mathbf{K}_n$)=$\{1,n\}$-CSP($\mathbf{K}_n$) to $\{n\}$-CSP($\mathbf{K}_{2n}$). Consider an instance of QCSP($\mathbf{K}_n$), namely a formula Ψ of the following form

$$\Psi = \exists^{\geq b_1} v_1 \, \exists^{\geq b_2} v_2 \, \ldots \, \exists^{\geq b_N} v_N \, \psi$$

where each $b_i \in \{1, n\}$. We let $\mathbf{G}$ denote the graph $\mathbf{D}_\psi$ with vertex set $\{v_1, \ldots, v_N\}$ and edge set $\{v_i v_j \mid E(v_i, v_j) \text{ appears in } \psi\}$.

We construct an instance Φ of $\{n\}$-CSP($\mathbf{K}_{2n}$) with the property that Ψ is a yes-instance of QCSP($\mathbf{K}_n$) if and only if Φ is a yes-instance of $\{n\}$-CSP($\mathbf{K}_{2n}$).

In short, we model the n-coloring using $2n-1$ colors, $n-1$ of which will be treated as *don't care* colors (vertices colored using any of such colors will be ignored). We make sure that the colorings where no vertex is assigned a don't-care color precisely model all colorings that we need to check to verify that Ψ is a yes-instance.

We describe Φ by giving a graph $\mathbf{H}$ together with a total order of its vertices with the usual interpretation that the vertices are the variables of Φ, the total order is the order of quantification of the variables, and the edges of $\mathbf{H}$ define the conjunction of predicates $E(\cdot,\cdot)$ which forms the quantifier-free part ϕ of Φ.

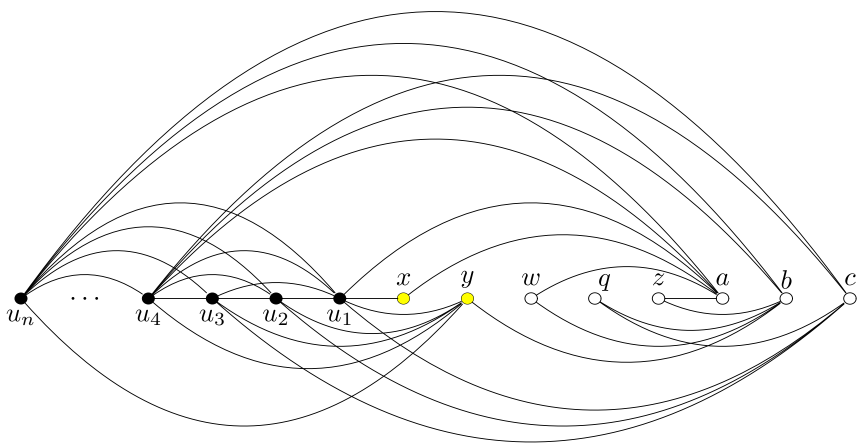


FIG. 4.1. *The edge gadget (here, as an example, x is an $\exists$ vertex while y is a $\forall$ vertex).*

We start constructing $\mathbf{H}$ by adding the vertices $v_1, v_2, \dots, v_N$ and no edges. Then we add new vertices $u_1, u_2, \dots, u_n$ and make them pairwise adjacent.

We make each v_i adjacent to u_1, and if $b_i = n$ (i.e., if v_i was quantified $\forall$), then we also make v_i adjacent to $u_2, u_3, \dots, u_n$.

We complete $\mathbf{H}$ by introducing for each edge $xy \in E(\mathbf{G})$, a gadget consisting of new vertices w, q, z, a, b, c with edges wa, wb, qb, qc, za, zb, and we connect this gadget to the rest of the graph as follows: we make x adjacent to a, make y adjacent to b, make a adjacent to u_1, make c adjacent to u_1, u_2, u_3, and make each of a, b, c adjacent to $u_4, \dots, u_n$. We refer to Figure 4.1 for an illustration.

The total order of $V(\mathbf{H})$ first lists $u_1, u_2, \dots, u_n$, then $v_1, v_2, \dots, v_N$ (exactly in the same order as quantified in Ψ), and then lists the remaining vertices of each gadget, in turn, as depicted in Figure 4.1 (listing w, q, z, a, b, c in this order).

We consider the game $\mathscr{G}(\Phi, \mathbf{K}_{2n})$ of Prover and Adversary played on Φ where Prover and Adversary take turns, for each variable in Φ in the order of quantification, respectively providing a set of n colors and choosing a color from the set. Prover wins if this process leads to a proper $2n$-coloring of $\mathbf{H}$ (no adjacent vertices receive the same color), otherwise Prover loses and Adversary wins. The formula Φ is a yes-instance if and only if Prover has a winning strategy.

Without loss of generality (up to renaming colors), we may assume that the vertices $u_1, u_2, \dots, u_n$ get assigned colors $n+1, n+2, \dots, 2n$, respectively, i.e., each u_i gets color $n+i$. (The edges between these vertices make sure that Prover must offer distinct colors while Adversary has no way of forcing a conflict, since there are $2n$ colors available.)

Lemma 4.1. *If Adversary is allowed to choose for the vertices x, y in the edge gadget (Figure 4.1) the same color from $\{1, 2, \dots, n\}$, then Adversary wins.*

If Adversary is allowed to choose $n+1$ for x or y, then Adversary also wins.

In all other cases, Prover wins.

Proof. If Prover offers $n+1$ for x or y, then Adversary can choose this color (for x or y) and Prover immediately loses, since both x and y are adjacent to u_1 which is assumed to be assigned the color $n+1$. (Prover loses since the coloring is not proper.)

Assume that x and y are assigned the same color i from $\{1, 2, \dots, n\}$. We describe a winning strategy for Adversary. Consider the set of n colors Prover offers for w. Since the colors are distinct and there is n of them, at least one of the colors, denote it k, is different from i and each of $n+1, n+4, n+5, \dots, 2n$. Adversary chooses the color k for w.

Then consider the n colors Prover offers for q. If any of the n colors, denote it j, is from $\{1, 2, \dots, n\}$, then Adversary chooses the color j for q, which makes Prover lose when considering the vertex c where Prover must offer n values different from $n+1, n+2, \dots, 2n$ and from $j \in \{1, 2, \dots, n\}$, which is impossible. (Note that c is adjacent to q as well as $u_1, u_2, \dots, u_n$.)

Therefore we may assume that Prover offers the set $\{n+1, n+2, \dots, 2n\}$ for q. Let ℓ be any color in the set $\{n+2, n+3\} \setminus \{k\}$. By definition, ℓ is different from k, and clearly also different from i, since $i \in \{1, 2, \dots, n\}$ while $\ell \in \{n+2, n+3\}$. We make Adversary choose the color ℓ for q.

Now if Prover offers for z a color, denote it r, different from i, k and each of $n+1$, $n+4$, $n+5$, $\dots$, $2n$, then Adversary chooses this color and Prover loses at a when she has to provide n colors distinct from $i, k, r, n+1, n+4, n+5, \dots, 2n$, which is impossible. Similarly, Prover loses, this time at b, if she offers for z a color different from i, k, ℓ, $n+4$, $n+5$, $\dots$, $2n$. Notice that there is no set of n values that excludes both these situations, since i, k, ℓ, $n+1$, $n+4$, $\ldots$, $2n$ are distinct values. This shows that Adversary wins no matter what Prover does.

Now, for the second part of the claim, assume that x and y are either given distinct colors different from $n+1$, or same colors from $\{n+2, n+3, \ldots, 2n\}$. This time Prover wins no matter what Adversary does.

First, assume that x and y have distinct colors i and j, respectively where $i, j \neq n+1$. We consider three cases.

Case 1: assume that $j \in \{n+4, n+5, \ldots, 2n\}$. Then we have Prover offer for the vertices w and q the set $\{n+1, n+2, \ldots, 2n\}$. Let k be the color chosen by Adversary for w, and let ℓ be the color chosen for q.

If $\{i, k, n+1, n+4, n+5, \ldots, 2n\}$ contains n distinct elements, we have Prover offer this set for z; otherwise Prover offers any set of n distinct elements for z. Let r be the color chosen for z. Now Prover offers for a any set of n colors disjoint from $\{i, k, r, n+1, n+4, \ldots, 2n\}$, which is possible since $r \in \{i, k, n+1, n+4, \ldots, 2n\}$. For b Prover offers any set of n colors disjoint from $\{j, k, \ell, r, n+4, \ldots, 2n\}$, which is again possible because $j \in \{n+4, \ldots, 2n\}$. Finally, Prover offers $\{1, 2, \ldots, n\}$ for c. It is now easy to see that any choice of Adversary yields a proper coloring and so Prover wins, as claimed.

Case 2: assume that $i \in \{n+4, n+5, \ldots, 2n\}$. We similarly have Prover offer $\{n+1, n+2, \ldots, 2n\}$ for both w and q, and let k and ℓ be the colors chosen by Adversary for the two vertices. If $\{j, k, \ell, n+4, n+5, \ldots, 2n\}$ contains n distinct elements, Prover offers this set for z; otherwise Prover offers any set of n distinct elements. Just like in Case 1.1, this now allows us to choose n distinct colors for each of a, b, c so that none of the colors appears on their neighbors. So again, for any choice of Adversary, Prover wins as required.

Case 3: assume that $i, j \notin \{n+4, n+5, \ldots, 2n\}$. Recall that $i, j \neq i+1$ and $i \neq j$. Thus we have Prover offer for w the set $\{n+1, i, j, n+4, \ldots, 2n\}$ and for q the set $\{n+1, \ldots, 2n\}$. Let k be the color chosen by Adversary for w, and let ℓ be the color chosen for q.

Suppose first that $k \in \{i, n+1, n+4, \ldots, 2n\}$. If $\{j, k, \ell, n+4, \ldots, 2n\}$ contains n distinct elements, Prover offers this set for z; otherwise, she offers any set for z. Again, for each of a, b, c, there are at most n colors used on their neighbors and so Prover can offer each of a, b, c a set of n colors distinct from their neighbors to get a proper coloring for any choice of Adversary.

So we may assume that $k = j$. In this case, we have Prover offer the set $\{i, n+1, n+4, \ldots, 2n\}$ for z. Again, for each of a, b, c we have n colors distinct from their neighbors and we can thus complete a proper coloring regardless of Adversary's choices. Thus Prover wins in any situation.

This exhausts all possibilities for when x, y have distinct colors different from $n+1$. To finish the proof, it remains to consider the case when x, y have the same color i from $\{n+2, n+3, \ldots, 2n\}$ In this case, Prover offers for the vertices w, q, z the set $\{n+1, n+2, \ldots, 2n\}$, while for a, b, c Prover offers the set $\{1, 2, \ldots, n\}$. It is easy to see that any choice of Adversary yields a proper coloring. Thus Prover wins as required. This concludes the proof. ◻

LEMMA 4.2. *Φ is a yes-instance of $\{n\}$-CSP($\mathbf{K}_{2n}$) if and only if Ψ is a yes-instance of* QCSP($\mathbf{K}_n$).

Proof. We treat the colors $n+2, n+3, \ldots, 2n$ as *don't care* colors, while $1, 2, \ldots, n$ will be the actual colors used for coloring $\mathbf{G}$. By Lemma 4.1, the edge gadget makes sure that vertices x, y do not receive the same colors unless at least one of the colors is from $\{n+2, n+3, \ldots, 2n\}$ (the don't-care colors). This implies that Φ correctly simulates Ψ whereby Prover offers $\{1, 2, \ldots, n\}$ for each $\forall$ variable of Ψ, and offers $\{i, n+2, n+3, \ldots, 2n\}$ for each $\exists$ variable of Ψ where $i \in \{1, 2, \ldots, n\}$. Note that the construction forces Prover to offer $\{1, 2, \ldots, n\}$ for each $\forall$ variable, while for each $\exists$ variable Prover must offer n values excluding the value $n+1$. In the latter case we may assume that the set of offered values is of the form $\{i, n+2, n+3, \ldots, 2n\}$, where $i \in \{1, 2, \ldots, n\}$, since offering more values from $\{1, 2, \ldots, n\}$ makes it even easier for Adversary to win (has more choices to force a monochromatic edge).

Thus this shows that Φ indeed correctly simulates Ψ as required. ∎

THEOREM 4.3. *$\{n\}$-CSP($\mathbf{K}_{2n}$) is Pspace-complete for all $n \geq 3$.*

Proof. The claim follows from the previous two lemmas. We finish the proof by remarking that the construction of Φ is polynomial in the size of Ψ (in fact the reduction is in L). Thus, since QCSP($\mathbf{K}_n$) is Pspace-hard, so is $\{n\}$-CSP($\mathbf{K}_{2n}$). ∎

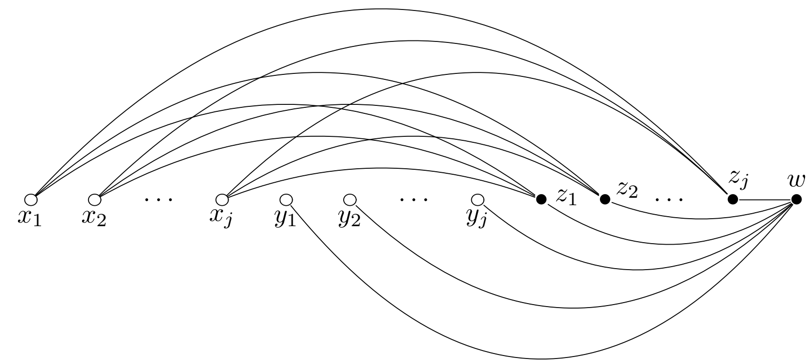


FIG. 4.2. *The gadget $\mathbf{G}_j$.*

4.2. Other Pspace-complete cases. We now discuss the remaining Pspace-complete cases of Theorem 1.1. Similar to Theorem 4.3, the reductions will also be from the quantified coloring problem.

PROPOSITION 4.4. *If $1 < j$, then $\{j\}$-CSP($\mathbf{K}_{2j+1}$) is* Pspace-*complete.*

Proof. The proof is by reduction from QCSP$\left(\mathbf{K}_{\binom{2j+1}{j}}\right)$. The key part of our proof involves the gadget $\mathbf{G}_j$ (Figure 4.2) having vertices $x_1, \ldots, x_j$, $y_1, \ldots, y_j$, $z_1, \ldots, z_j, w$ and all possible edges between $\{x_1, \ldots, x_j\}$ and $\{z_1, \ldots, z_j\}$, and between w and $\{y_1, \ldots, y_j, z_1, \ldots, z_j\}$. The leftmost $2j$ vertices represent free variables $x_1, \ldots, x_j$, $y_1, \ldots, y_j$. Observe that $\exists^{\geq j} z_1, \ldots, z_j, w\ \phi_{\mathbf{G}_j}$ is true on $\mathbf{K}_{2j+1}$ iff $|\{x_1, \ldots, x_j\} \cap \{y_1, \ldots, y_j\}| < j$. If $|\{x_1, \ldots, x_j\}| = |\{y_1, \ldots, y_j\}| = j$, then this is equivalent to $\{x_1 \ldots x_j\} \neq \{y_1 \ldots y_j\}$. We use this gadget to encode the edge relation of $\mathbf{K}_{\binom{2j+1}{j}}$ by representing vertices as sets $\{a_1, \ldots, a_j\} \subset \{1, \ldots, 2j+1\}$ with $|\{a_1, \ldots, a_j\}| = j$.

Consider an instance Ψ of QCSP$\left(\mathbf{K}_{\binom{2j+1}{j}}\right)$. We construct the instance Ψ' of $\{j\}$-CSP($\mathbf{K}_{2j+1}$) as follows. From the graph $\mathbf{D}_\psi$, build $\mathbf{D}_{\psi'}$ by transforming each vertex v into an independent set of j vertices $\{v^1, \ldots, v^j\}$, and each edge uv of D_ψ we transform to an instance of the gadget $\mathbf{G}_j$ in which the $2j$ free variables correspond to $u^1, \ldots, u^j, v^1, \ldots, v^j$. The other variables of the gadget $\{z_1, \ldots, z_j, w\}$ are unique to each edge and are quantified innermost in Ψ' in the order $z_1, \ldots, z_j, w$.

It remains to explain the quantification of the variables of the form $v^1,\dots,v^j$. We follow the quantifier order of Ψ. Existentially quantified variables $\exists v$ of Ψ are quantified as $\exists^{\geq j} v^1,\dots,v^j$ in Ψ'. Universally quantified variables $\forall v$ of Ψ are also quantified $\exists^{\geq j} v^1,\dots,v^j$ in Ψ', but we introduce additional variables $v^{1,1},\dots,v^{1,j+1}$, $\dots$, $v^{j,1},\dots,v^{j,j+1}$ before $v^1,\dots,v^j$ in the quantifier order of Ψ'. For each $i \in \{1,\dots,j\}$, we join $v^{i,1},\dots,v^{i,j+1}$ into a clique with v^i.

We show that $\mathbf{K}_{\binom{2j+1}{j}} \models \Psi$ iff $\mathbf{K}_{2j+1} \models \Psi'$. Observe there is a natural bijection π from subsets of j elements of $\mathbf{K}_n$ to vertices of $\mathbf{K}_{\binom{2j+1}{j}}$. In the simulation of $\mathrm{QCSP}\left(\mathbf{K}_{\binom{2j+1}{j}}\right)$ in $\{j\}$-CSP($\mathbf{K}_{2j+1}$), Adversary may be seen to take on the role of denying $\mathbf{K}_{\binom{2j+1}{j}} \models \Psi$ while Prover is asserting that it is true. Thus, Adversary may always be assumed to play variables $v^1,\dots,v^j$ such that $|\{v^1,\dots,v^j\}| = j$, because otherwise he is simply making the job of Prover easier (by the properties of the gadget $\mathbf{G}_j$). The behavior of existential quantification in the simulation is easy to see, but we will consider more carefully the behavior of universal quantification. The additional $v^{1,1},\dots,v^{1,j+1},\dots,v^{j,1},\dots,v^{j,j+1}$ cause that every possible subset $\{a_1,\dots,a_j\} \subset \{1,\dots,2j+1\}$ can be forced by Adversary on $v^1,\dots,v^j$. Indeed, Adversary may force any single element on v^i by avoiding it in $v^{i,1},\dots,v^{i,j+1}$.

($\Rightarrow$) Assume $\mathbf{K}_{\binom{2j+1}{j}} \models \Psi$. However Prover plays the variables in Ψ' corresponding to universal variables of Ψ, she will be able to find a witness set $\pi^{-1}(a)$ for the variables in Ψ' corresponding to an existential variable x in Ψ, precisely because that existential variable has some witness $a \in V(\mathbf{K}_{\binom{2j+1}{j}})$.

($\Leftarrow$) Assume $\mathbf{K}_{2j+1} \models \Psi'$. No matter how Prover plays to win $\mathscr{G}(\Psi',\mathbf{K}_{2j+1})$, she will have possible witnesses sets $\{a_1,\dots,a_j\}$ for variables $\{v^1,\dots,v^j\}$ in Ψ' corresponding to an existential variable v of Ψ, for all sets $\{b_1,\dots,b_j\} \subset \{1,\dots,2j+1\}$ corresponding to universal variables $\{u_1,\dots,v_j\}$ of Ψ (because of the behavior of the universal variable simulation). Thus the existential witness $\pi(\{a_1,\dots,a_j\})$ may be used in Ψ for v, and the result follows. ◻

Corollary 4.5. *If $1 < j < n/2$, then $\{j\}$-CSP($\mathbf{K}_n$) is* Pspace-*complete.*

Proof. We reduce from $\{j\}$-CSP($\mathbf{K}_{2j+1}$) and appeal to Proposition 4.4. Given an input Ψ for $\{j\}$-CSP($\mathbf{K}_{2j+1}$), we build an instance Ψ' for $\{j\}$-CSP($\mathbf{K}_n$) by adding an $(n-2j-1)$-clique on new variables, quantified outermost in Ψ', and link by an edge each variable of this clique to every other variable. Adversary chooses $n-2j-1$ elements of the domain for this clique, effectively reducing the domain size to $2j+1$ for the rest. Thus $\mathbf{K}_n \models \Psi'$ iff $\mathbf{K}_{2j+1} \models \Psi$ follows. ◻

Proposition 4.6. *If $1 < j \leq n$, then $\{1,j\}$-CSP($\mathbf{K}_n$) is* Pspace-*complete.*

Proof. By reduction from QCSP($\mathbf{K}_n$). We simulate existential quantification $\exists v$ by itself, and universal quantification $\forall v$ by the introduction of $(n-j+1)$ new variables $v^1,\dots,v^{n-j}$, joined in a clique with v, and quantified by $\exists^{\geq j}$ before v (which is also quantified by $\exists^{\geq j}$). The argument follows as in Proposition 4.4. ◻

4.3. Logspace cases. Define the n-star $\mathbf{K}_{1,n}$ to be the graph with vertex set $\{0,1,\dots,n\}$ and edge set $\{(0,j),\,(j,0)\ :\ j \geq 1\}$ where 0 is called the *hub* and the remaining vertices are called *leaves*.

Proposition 4.7. *If $X \cap \{1,\dots,\lfloor n/2\rfloor\} = \emptyset$, then X-CSP($\mathbf{K}_n$) is in* L.

Proof. Instance Ψ of X-CSP($\mathbf{K}_n$) of the form $\exists^{\geq\lambda_1}x_1\dots\exists^{\geq\lambda_m}x_m\ \psi(x_1,\dots,x_m)$ induces the graph $\mathbf{D}_\psi$, which we may consider totally ordered (the order $\prec$ is given left-to-right ascending by the quantifiers). We claim that $\mathbf{K}_n \models \Psi$ iff $\mathbf{D}_\psi$ does not contain as a subgraph (not necessarily induced) a $(n-\lambda_i+1)$-star in which the

$n-\lambda_i+1$ leaves all come before the hub x_i in the ordering.

($\Rightarrow$) If $\mathbf{D}_\psi$ contains such a star, then Ψ is a no-instance, as we may give a winning strategy for Adversary in the game $\mathscr{G}(\Psi, \mathbf{K}_n)$. Adversary should choose distinct values for the variables associated with the $n-\lambda_i+1$ leaves of the star (can always be done as each of the possible quantifiers assert existence of $> n/2$ elements and $n-\lambda_i < n/2$), whereupon there is no possibility for Prover to choose λ_i witnesses to the variable x_i associated with the hub.

($\Leftarrow$) If $\mathbf{D}_\psi$ does not contain such a star, then we give the following winning strategy for Prover in the game $\mathscr{G}(\Psi, \mathbf{K}_n)$. Whenever a new variable comes up, its corresponding vertex in $\mathbf{D}_\psi$ has $l < n-\lambda_i+1$ adjacent predecessors, which were answered with $b_1, \ldots, b_l$. Prover suggests any set of size λ_i from $B \setminus \{b_1, \ldots, b_l\}$ (which always exists) and the result follows. ◻

4.4. Polynomial-time algorithm for $\{2\}$-CSP($\mathbf{K}_4$). In this section, we analyze the last remaining case of Theorem 1.1. In particular, we describe a polynomial-time algorithm for the problem $\{2\}$-CSP($\mathbf{K}_4$).

The template $\mathbf{K}_4$ has vertices $\{1,2,3,4\}$ and all possible edges between distinct vertices. As before, we consider an instance $\Psi := \exists^{\geq 2} v_1 \ldots \exists^{\geq 2} v_N \psi$ of $\{2\}$-CSP($\mathbf{K}_4$) as a graph $\mathbf{G} = \mathbf{D}_\psi$ with vertex set $V(\mathbf{G}) = \{v_1, \ldots, v_N\}$ and edge set $E(\mathbf{G}) = \{v_i v_j \mid E(v_i, v_j) \text{ appears in } \psi\}$, and with a linear order $\prec$ where $v_1 \prec v_2 \prec \ldots \prec v_N$.

We iteratively construct the following three sets: R^+, R^-, and F. The set F will be a collection of unordered pairs of vertices of $\mathbf{G}$, while R^+ and R^- will consist of unordered triples of vertices. (For simplicity we write $xy \in F$ in place of $\{x,y\} \in F$, and write $xyz \in R^+$ or R^- in place of $\{x,y,z\} \in R^+$ or R^-.)

The meaning of these sets is as follows. A pair $xy \in F$ where $x \prec y$ indicates that Prover, in order not to lose $\mathscr{G}(\Psi, \mathbf{K}_4)$, must offer values so that the value $f(x)$ chosen by Adversary for x is different from the value $f(y)$ chosen for y. A triple $xyz \in R^+$ where $x \prec y \prec z$ indicates that if Adversary chose $f(x) \neq f(y)$, then Prover must offer one (or both) of $f(x), f(y)$ for z. A triple $xyz \in R^-$ where $x \prec y \prec z$ tells us that Prover must offer for z values different from $f(x), f(y)$ if $f(x) \neq f(y)$.

With this, we describe how to iteratively compute the three sets F, R^+, R^-. We start by initializing the sets as follows: $F = E(\mathbf{G})$ and $R^+ = R^- = \emptyset$. Then we perform the following rules as long as possible:

- (X1) If there are distinct $x,y,z \in V(\mathbf{G})$ such that $\{x,y\} \prec z$ where $xz, yz \in F$, then add xyz into R^-.
- (X2) If there are distinct vertices $x,y,w,z \in V(\mathbf{G})$ such that $\{x,y,w\} \prec z$ with $wz \in F$ and $xyz \in R^-$, then add xyw into R^+.
- (X3) If there are distinct $x,y,w,z \in V(\mathbf{G})$ such that $\{x,y,w\} \prec z$ with $wz \in F$ and $xyz \in R^+$, then if $\{x,y\} \prec w$, then add xyw into R^- else add xw and yw into F.
- (X4) If there are distinct vertices $x,y,w,z \in V(\mathbf{G})$ such that $\{x,w\} \prec y \prec z$ with $xyz \in R^+$ and $wyz \in R^-$, then add xw into F, and add xwy into R^+.
- (X5) If there are distinct vertices $x,y,w,z \in V(\mathbf{G})$ such that $\{x,y,w\} \prec z$ where either $xyz, wyz \in R^+$, or $xyz, wyz \in R^-$, then add xyw into R^+.
- (X6) If there are distinct vertices $x,y,q,w,z \in V(\mathbf{G})$ such that $\{x,y,w\} \prec q \prec z$ where either $xyz, wqz \in R^+$, or $xyz, wqz \in R^-$, then add xyw and xyq into R^+.

(X7) If there are distinct vertices $x, y, q, w, z \in V(\mathbf{G})$ such that $\{x, y, w\} \prec q \prec z$ where either $xyz \in R^+$ and $wqz \in R^-$, or $xyz \in R^-$ and $wqz \in R^+$, then add xyq into R^-, and if $\{x, y\} \prec w$, also add xyw into R^-,
else add xw and yw into F.

Leting f denote the mapping produced during an execution of the game $\mathscr{G}(\Psi, \mathbf{K}_4)$, We uncover the following theorem.

THEOREM 4.8. *The following are equivalent:*

(i) $\mathbf{K}_4 \models \Psi$

(ii) *Prover has a winning strategy in* $\mathscr{G}(\Psi, \mathbf{K}_4)$.

(iii) *Prover can play so that in every execution of the game, the resulting mapping* $f : V(\mathbf{G}) \to \{1, 2, 3, 4\}$ *satisfies the following properties:*

(S1) *For every* $xy \in F$, *we have:* $f(x) \neq f(y)$.

(S2) *For every* $xyz \in R^+$ *such that* $x \prec y \prec z$:
if $f(x) \neq f(y)$, *then* $f(z) \in \{f(x), f(y)\}$.

(S3) *For every* $xyz \in R^-$ *such that* $x \prec y \prec z$:
if $f(x) \neq f(y)$, *then* $f(z) \notin \{f(x), f(y)\}$.

(iv) *there is no triple* xyz *in* R^+ *such that* $x \prec y \prec z$ *and (see Figure 4.3)*
– $xz \in F$ *or* $yz \in F$,
– *or* $xwz \in R^-$ *for some* $w \prec z$ *(possibly* $w = y$*),*
– *or* $ywz \in R^-$ *for some* $y \prec w \prec z$.

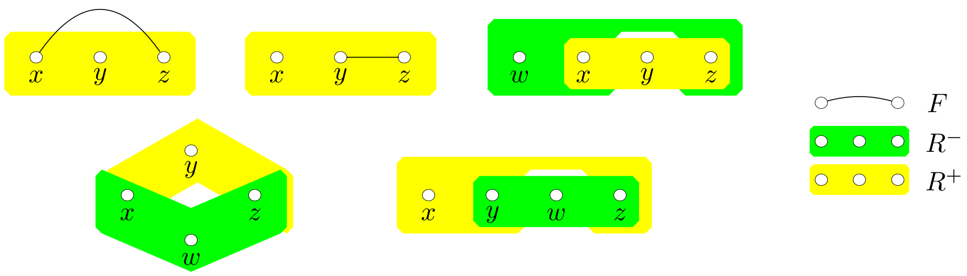


FIG. 4.3. *Pictorial representation of the five forbidden configurations of Theorem 4.8.*

Proof. **(i)** $\iff$ **(ii)** is by definition. **(iii)⇒(ii)** is implied by the fact that $F \supseteq E(\mathbf{G})$, and that by (iii) Prover can play to satisfy (S1). Thus in every execution of the game the mapping f is a homomorphism of $\mathbf{G}$ to $\mathbf{K}_4 \Rightarrow$ (ii).

(ii)⇒(iii): Suppose that Prover plays a winning strategy in the game $\mathscr{G}(\Psi, \mathbf{K}_4)$ but (iii) fails. We show that this is impossible. Namely, we show how Adversary can play to win. Consider an execution of the game producing a mapping f.

We say that **(S1) fails at a vertex** $\boldsymbol{v}$ if there exists $a \in V(\mathbf{G})$ with $a \prec v$ such that $av \in F$ and $f(a) = f(v)$. We say that **(S2) fails at** $\boldsymbol{v}$ if there exist $a, b \in V(\mathbf{G})$ with $a \prec b \prec v$ such that $abv \in R^+$ while $f(a) \neq f(b)$ and $f(v) \notin \{f(a), f(b)\}$. We say that **(S3) fails at** $\boldsymbol{v}$ if there exist $a, b \in V(\mathbf{G})$ with $a \prec b \prec v$ such that $abv \in R^-$ while $f(a) \neq f(b)$ and $f(v) \in \{f(a), f(b)\}$.

Since (iii) fails, there is an execution of the game producing a mapping f that fails (S1)-(S3) at some vertex v. Among all such executions, pick one for which v is largest possible with respect to the order $\prec$. We will show that this is impossible, namely we will produce a (possibly) different execution violating the maximality of this choice. Note that, since we assume that Prover plays a winning strategy, the mapping f is a homomorphism of $\mathbf{G}$ to $\mathbf{K}_4$.

Case 1: Suppose that (S1) fails at v. Then there is $a \prec v$ such that $av \in F$ and $f(a) = f(v)$. If $av \in E(\mathbf{G})$, then the mapping f is not a homomorphism of G to $\mathbf{K}_4$. Thus Adversary wins, which contradicts (ii). So we may assume that $av \notin E(\mathbf{G})$. This implies that av was added to F using one of the rules (X3), (X4), (X7).

Case 1.1: Suppose that av was added to F using (X3). Then there exist vertices x, y, w, z where $\{x, y, w\} \prec z$ and $\{x, y\} \not\prec w$ such that $wz \in F$ and $xyz \in R^+$, and either $a \in \{x, y\}$ and $v = w$, or $v \in \{x, y\}$ and $a = w$. In particular, since $f(a) = f(v)$, we deduce that $f(x) = f(w)$ or $f(y) = f(w)$.

We may assume by symmetry that $x \prec y$. Recall that $\{x, y\} \not\prec w$. Thus $\{x, w\} \prec y$. Consider the point of the execution of the game producing f when Prover offers values for y. From this point on, we have Adversary play as follows: for y, if $f(x) = f(w)$, choose any value that is different from $f(w)$; if $f(x) \neq f(w)$, choose $f(y)$ for y. Let β denote the value chosen for y. Observe that the choice is always possible, since Prover offers for y two distinct values, one of which is $f(y)$. Moreover, the choice guarantees that $f(x) \neq \beta$, since either $f(x) = f(w) \neq \beta$ or $f(x) \neq f(w) = f(y) = \beta$. For this, recall that $f(x) = f(w)$ or $f(y) = f(w)$. Then for z, if $f(w)$ is offered by Prover, we let Adversary choose $f(w)$ for z; otherwise Adversary chooses any value different from $f(x)$ and β. Let α denote the value chosen for z. Again, note that the choice is always possible, in particular in the latter case where Prover offers for z two distinct values, neither of which is $f(w)$, while $f(w) = f(x)$ or $f(w) = f(y) = \beta$. For the remaining vertices, we let Adversary play any choices. This produces a (possibly) different execution of the game. It follows that this execution fails (S1) or (S2) at z. Namely, if $f(w)$ was offered for z, then $\alpha = f(w)$ and (S1) fails at z, since $wz \in F$. If $f(w)$ was not offered for z, then $\alpha \notin \{f(x), \beta\}$, in which case (S2) fails at z, because $xyz \in R^+$. However, since $v \prec z$, this contradicts our choice of v.

Case 1.2: Suppose that av was added to F using (X4). Then there exist vertices x, y, w, z where $\{x, w\} \prec y \prec z$ such that $xyz \in R^+$ and $wyz \in R^-$, and where $\{x, w\} = \{a, v\}$. In particular, we deduce that $f(x) = f(w)$.

We consider the point when Prover offers values for y and have Adversary play as follows: for y, choose any value different from $f(x)$; let β denote this value. Note that $\beta \neq f(x)$. Then for z, choose any value, denote it α, and then play any choices for the remaining vertices. The execution this produces fails (S2) or (S3) at z. Namely, if $\alpha \notin \{f(x), \beta\}$, then (S2) fails, since $xyz \in R^+$, while if $\alpha \in \{f(x), \beta\}$, then (S3) fails, since $f(x) = f(w)$ and $wyz \in R^-$. This again contradicts our choice of v, since $v \prec z$.

Case 1.3: Suppose that av was added to F using (X7). Then there exist vertices x, y, q, w, z where $\{x, y, w\} \prec q \prec z$ and $\{x, y\} \not\prec w$ such that $xyz \in R^+$, $wqz \in R^-$, or $xyz \in R^-$, $wqz \in R^+$, and such that $a \in \{x, y\}$ and $v = w$, or $a = w$ and $v \in \{x, y\}$. Since $f(a) = f(v)$, we deduce that $f(x) = f(w)$ or $f(y) = f(w)$.

By symmetry, assume $x \prec y$. Thus $\{x, w\} \prec y \prec q \prec z$ because $\{x, y\} \not\prec w$. We have Adversary play from y as follows: for y, if $f(x) = f(w)$, choose any value different from $f(w)$; if $f(x) \neq f(w)$, choose $f(y)$ for y. Let β denote the value chosen for y. Note that $f(x) \neq \beta$ and $f(w) \in \{f(x), \beta\}$, since either $f(x) = f(w) \neq \beta$ or $f(x) \neq f(w) = f(y) = \beta$. For this, recall that $f(x) = f(w)$ or $f(y) = f(w)$. Next, for q, we let Adversary choose any value different from $f(w)$, and denote it γ. Note that $\gamma \neq f(w)$. Finally, for z, if $xyz \in R^+$ and $wqz \in R^-$, Adversary chooses $f(w)$ if offered by Prover, and if not, he chooses any value different from $f(x)$ and β. Similarly, if $xyz \in R^-$ and $wqz \in R^+$, Adversary chooses $f(x)$ or β if offered by Prover, and otherwise he chooses any value different from $f(w)$ and γ. Let

α denote the value chosen for z. Again, we see that this choice is always possible, since $f(w) \in \{f(x), \beta\}$. Adversary plays any choices for the rest. We claim that the execution this produces fails (S2) or (S3) at z. Namely, if $xyz \in R^+$ and $wqz \in R^-$, then (S2) fails at z if $\alpha \notin \{f(x), \beta\}$, since $xyz \in R^+$, while if $\alpha \in \{f(x), \beta\}$, then (S3) fails at z, since in that case we must have $\alpha = f(w) \neq \gamma$ by the choice of α, while $wqz \in R^-$. Similarly, if $xyz \in R^-$ and $wqz \in R^+$, then either $\alpha \in \{f(x), \beta\}$ and (S3) fails at z, since $xyz \in R^-$, or $\alpha \notin \{f(x), \beta\}$ in which case $\alpha \notin \{f(w), \gamma\}$ and (S2) fails, since $wqz \in R^+$. Since $v \prec z$, this contradicts our choice of v.

From now on, we may assume that (S1) does not fail at v.

Case 2: Suppose that (S2) fails at v. Then there are vertices a, b with $a \prec b \prec v$ such that $abv \in R^+$ while $f(a) \neq f(b)$ and $f(v) \notin \{f(a), f(b)\}$. Since the set R^+ is initially empty, the triple abv was added to R^+ by one of the rules (X2), (X4), (X5), or (X6).

Case 2.1: Suppose that abv was added to R^+ using (X2). Then there exist vertices x, y, w, z where $\{x, y, w\} \prec z$ such that $wz \in F$ and $xyz \in R^-$, and where $\{a, b, v\} = \{x, y, w\}$. We deduce that $f(x), f(y), f(w)$ are three distinct values, since $f(a), f(b), f(v)$ are. We have Adversary play from z as follows: for z, from the two offered values, one will be in $\{f(x), f(y), f(w)\}$; choose this value α. For the rest, play any choices. It follows that this execution fails (S1) or (S3) at z. Namely, if $\alpha = f(w)$, then (S1) fails, since $wz \in F$, while if $\alpha \in \{f(x), f(y)\}$, then (S3) fails, since $xyz \in R^-$. This contradicts our choice of v, since $v \prec z$.

Case 2.2: Suppose that abv was added to R^+ using (X4). Then there are vertices x, y, w, z where $\{x, w\} \prec y \prec z$ such that $xyz \in R^+$ and $wyz \in R^-$, and where $\{a, b, v\} = \{x, y, w\}$. Again, we deduce that $f(x), f(y), f(w)$ are pairwise distinct. Adversary plays from z as follows: for z, if $f(y)$ is offered, choose $f(y)$; otherwise, choose any value different from $f(x)$. For the rest, play any choices. Let α be the value chosen for z. We claim that this execution fails (S2) or (S3) at z. Namely, if $\alpha = f(y)$, then (S3) fails, since $f(w) \neq f(y)$ and $wyz \in R^-$, while if $\alpha \neq f(y)$, then $\alpha \notin \{f(x), f(y)\}$, in which case (S2) fails, since $f(x) \neq f(y)$ and $xyz \in R^+$. This contradicts our choice of v, since $v \prec z$.

Case 2.3: Suppose that abv was added to R^+ using (X5). Then there exist vertices x, y, w, z where $\{x, y, w\} \prec z$ such that either $xyz, wyz \in R^+$ or $xyz, wyz \in R^-$, and where $\{a, b, v\} = \{x, y, w\}$. Hence, we deduce that $f(x), f(y), f(w)$ are pairwise distinct. Adversary plays from z as follows: for z, if $xyz, wyz \in R^+$, choose any value different from $f(y)$; if $xyz, wyz \in R^-$, choose any value in $\{f(x), f(y), f(w)\}$. For the rest, play any choices. Let α denote the value chosen for z. We claim that this execution fails (S2) or (S3) at z. Namely, if $xyz, wyz \in R^+$, then $\alpha \neq f(y)$ and also $\alpha \neq f(x)$ or $\alpha \neq f(w)$, since $f(x) \neq f(w)$. Thus (S2) fails, since either $\alpha \notin \{f(x), f(y)\}$ and $xyz \in R^+$, or $\alpha \notin \{f(w), f(y)\}$ and $wyz \in R^+$. Similarly, if $xyz, wyz \in R^-$, then $\alpha \in \{f(x), f(y), f(w)\}$. Thus if $\alpha \in \{f(x), f(y)\}$, then (S3) fails, since $xyz \in R^-$, while if $\alpha = f(w)$, then (S3) fails, since $wyz \in R^-$. This contradicts the choice of v, since $v \prec z$.

Case 2.4: Suppose that abv was added to R^+ using (X6). Then there are vertices x, y, q, w, z where $\{x, y, w\} \prec q \prec z$ such that either $xyz, wqz \in R^+$ or $xyz, wqz \in R^-$, and where either $\{a, b, v\} = \{x, y, w\}$ or $\{a, b, v\} = \{x, y, q\}$. In either case, we have $f(x) \neq f(y)$. Adversary plays from q as follows: if $f(w) \in \{f(x), f(y)\}$, then choose $f(q)$ for q; otherwise choose any value different from $f(w)$. Let γ denote the value chosen for q. Then for z, if $xyz, wqz \in R^-$, choose any value in $\{f(x), f(y), f(w), \gamma\}$;

if $xyz, wqz \in R^+$ and $f(w) \in \{f(x), f(y)\}$, choose any value different from $f(w)$; otherwise choose any value different from γ. Let α denote the value chosen for z. Note that $\gamma \neq f(w)$. Indeed, if $f(w) \notin \{f(x), f(y)\}$, then $\gamma \neq f(w)$ by our choice. If $f(w) \in \{f(x), f(y)\}$, then $\{a, b, v\} \neq \{x, y, w\}$, since $f(a), f(b), f(v)$ are pairwise distinct. Thus $\{a, b, v\} = \{x, y, q\}$ implying that $f(x), f(y), f(q)$ are pairwise distinct; so $f(w) \neq \gamma = f(q)$, since $f(w) \in \{f(x), f(y)\}$.

We claim that this execution fails (S2) or (S3) at z. Namely, if $xyz, wqz \in R^-$, then $\alpha \in \{f(x), f(y), f(w), \gamma\}$ and so (S3) either fails because $\alpha \in \{f(x), f(y)\}$ while $xyz \in R^-$, or it fails because $\alpha \in \{f(w), \gamma\}$ and $f(w) \neq \gamma$ while $wqz \in R^-$. If $xyz, wqz \in R^+$ and $f(w) \in \{f(x), f(y)\}$, then $\alpha \neq f(w)$ and $\gamma = f(q) \notin \{f(x), f(y)\}$; thus (S2) fails either because $\alpha \notin \{f(x), f(y)\}$ while $xyz \in R^+$, or $\{\alpha, f(w)\} = \{f(x), f(y)\}$ and $\gamma \notin \{f(x), f(y)\}$ while $wqz \in R^+$. Finally, if $xyz, wqz \in R^+$ and $f(w) \notin \{f(x), f(y)\}$, then $\alpha \neq \gamma$, and (S2) fails either because $\alpha = f(w)$ while $xyz \in R^+$, or because $\alpha \neq f(w)$ and $f(w) \neq \gamma \neq \alpha$, while $wqz \in R^+$. This contradicts the choice of v, since $v \prec z$.

Case 3: Suppose that (S3) fails at v. Then there are vertices a, b with $a \prec b \prec v$ such that $abv \in R^-$ while $f(a) \neq f(b)$ and $f(v) \in \{f(a), f(b)\}$. Since the set R^- is initially empty, the triple abv was added to R^- by one of the rules (X1), (X3), or (X7).

Case 3.1: Suppose that abv was added to R^- using (X1). Then $av, bv \in F$ and it follows that (S1) fails at v. Namely, if $f(v) = f(a)$, then (S1) fails, since $av \in F$, while if $f(v) = f(b)$, then (S1) fails, since $bv \in F$. However, we assume that (S1) does not fail at v (as this leads to Case 1), a contradiction.

Case 3.2: Suppose that abv was added to R^- using (X3). Then there exist vertices x, y, w, z where $\{x, y, w\} \prec z$ and $\{x, y\} \prec w$ such that $wz \in F$ and $xyz \in R^+$, and where $\{a, b, v\} = \{x, y, w\}$. Since $a \prec b \prec v$, we have $\{a, b\} = \{x, y\}$ and $v = w$. In particular, we deduce that $f(w) \in \{f(x), f(y)\}$. Adversary plays from z as follows: for z, if $f(w)$ offered, choose this value; otherwise, choose any value different from $f(x)$ and $f(y)$. Let α denote the value chosen for z. Note that this choice is always possible, since Prover offers for z two distinct values; if neither is $f(w)$, then at least one of them is distinct from both $f(x)$ and $f(y)$, since $f(w) \in \{f(x), f(y)\}$. For the rest, Adversary play any choices. We claim that this execution fails (S1) or (S2) at z. Namely, if $\alpha = f(w)$, then (S1) fails at z, since $wz \in F$, while if $\alpha \neq f(w)$, then $\alpha \notin \{f(x), f(y)\}$, in which case (S2) fails, since $f(x) \neq f(y)$ and $xyz \in R^+$. This contradicts the choice of v, since $v \prec z$.

Case 3.3: Suppose that abv was added to R^- using (X7). Then there are vertices x, y, q, w, z where $\{x, y, w\} \prec q \prec z$ such that either $xyz \in R^+$ and $wqz \in R^-$, or $xyz \in R^-$ and $wqz \in R^+$, and where either $\{a, b, v\} = \{x, y, q\}$, or where $\{x, y\} \prec w$ and $\{a, b, v\} = \{x, y, w\}$. Since $a \prec b \prec v$, we deduce that $\{a, b\} = \{x, y\}$ and $v \in \{w, q\}$. In particular, $f(x) \neq f(y)$ and either $f(w) \in \{f(x), f(y)\}$ or $f(q) \in \{f(x), f(y)\}$. Adversary plays from q as follows: if $f(w) \notin \{f(x), f(y)\}$, then choose $f(q)$ for q; otherwise, choose any value different from $f(w)$. Let γ denote the value chosen for q. Then for z, if $xyz \in R^+$ and $wqz \in R^-$, choose $f(w)$ or γ if offered, else choose any value distinct from $f(x)$ and $f(y)$. Note that this choice is always possible, since in the latter case Prover offers two distinct values, neither of which is $f(w)$, γ, while either $f(w) \in \{f(x), f(y)\}$ or $\gamma = f(q) \in \{f(x), f(y)\}$. Similarly, if $xyz \in R^-$ and $wqz \in R^+$, we have Adversary choose $f(x)$ or $f(y)$ if offered, and else choose any value distinct from $f(w)$ and γ. Again, this choice is always possible, since $\{f(w), \gamma\} \cap \{f(x), f(y)\} \neq \emptyset$. Let α denote the value chosen for z. For the rest, Adversary plays any choices. Note that $\gamma \neq f(w)$. Indeed, if $f(w) \in \{f(x), f(y)\}$, then $\gamma \neq f(w)$ by our choice. If $f(w) \notin \{f(x), f(y)\}$, then $f(q) \in \{f(x), f(y)\}$ and $\gamma = f(q)$; thus $\gamma \neq f(w)$, since γ is in $\{f(x), f(y)\}$ while $f(w)$ is not.

We claim that this execution fails (S2) or (S3) at z. Namely, if $xyz \in R^+$ and $wqz \in R^-$, then either (S3) fails, since $\alpha \in \{f(w), \gamma\}$ while $wqz \in R^-$, or (S2) fails, since $\alpha \notin \{f(x), f(y)\}$ while $xyz \in R^+$. Similarly, if $xyz \in R^-$ and $wqz \in R^+$, then either (S3) fails, since $\alpha \in \{f(x), f(y)\}$ while $xyz \in R^-$, or (S2) fails, since $\alpha \notin \{f(w), \gamma\}$ while $wqz \in R^+$. This contradicts the choice of v, since $v \prec z$.

This exhausts all possibilities. Therefore no such execution of the game exists, which proves (ii)$\Rightarrow$(iii).

(iii)$\Rightarrow$(iv): Assume that Prover has a strategy as described in (iii), but (iv) fails, i.e., there exists a triple $xyz \in R^+$ such that $x \prec y \prec z$ and either $xz \in F$, or $yz \in F$, or $xwz \in R^-$ for some $w \prec z$, or $ywz \in R^-$ for some $y \prec w \prec z$. We show that this is impossible. Namely, we show that there is a way that Adversary can play to violate the conditions of (iii). Again, we let f denote the mapping produced during an execution of the game.

Suppose first that $xz \in F$ or $yz \in F$. Adversary plays as follows: until the game reaches y, Adversary plays any choices. When the game reaches y, Prover offers two values for y; from the two, Adversary chooses, as the value $f(y)$, any offered value that is different from $f(x)$. Then Adversary again plays any choices until the game reaches z when Prover offers two distinct values for z. If any of the two values is not in $\{f(x), f(y)\}$, then Adversary chooses this value to be the value $f(z)$. Otherwise, he chooses $f(x)$ if $xz \in F$, and chooses $f(y)$ if $yz \in F$. For the rest, Adversary plays any choices. It follows that the mapping f fails to satisfy (S1) or (S2). Namely, if $f(z) \notin \{f(x), f(y)\}$, then (S2) fails, since $f(x) \neq f(y)$ and $xyz \in R^+$. If $f(z) \in \{f(x), f(y)\}$, then either $f(z) = f(x)$ in case $xz \in F$ and so (S1) fails, or $f(z) = f(y)$ in case $yz \in F$ and so (S1) fails again. This contradicts our assumption (iii).

Now, assume that $ywz \in R^-$ for some $y \prec w \prec z$. Adversary again chooses $f(x)$ and $f(y)$ to be distinct, and then chooses $f(w)$ to be distinct from $f(y)$. When z is reached, Adversary chooses $f(y)$ or $f(w)$ if offered by Prover, and else he chooses any value distinct from $f(x)$. For the rest, Adversary plays any choices. It follows that (S2) or (S3) fails for f. Namely, if $f(z) \in \{f(y), f(w)\}$, then (S3) fails, since $f(y) \neq f(w)$ and $ywz \in R^-$. If $f(z) \notin \{f(y), f(w)\}$, then $f(z) \neq f(x)$ and (S2) fails, since $f(x) \neq f(y)$ and $xyz \in R^+$. This contradicts (iii).

Lastly, assume that $xwz \in R^-$ where $w \prec z$ (possibly $w = y$). Adversary chooses $f(x)$ and $f(w)$ to be distinct and also chooses $f(y)$ so that $f(x)$ and $f(y)$ are distinct (possibly $y = w$). When z is reached, Adversary chooses $f(x)$ or $f(w)$ if offered by Prover, and else he chooses any value distinct from $f(y)$. For the rest, Adversary plays any choices. Again, we have that (S2) or (S3) fails for f. Namely, if $f(z) \in \{f(x), f(w)\}$, then (S3) fails, since $f(x) \neq f(w)$ and $xwz \in R^-$. If $f(z) \notin \{f(x), f(w)\}$, then also $f(z) \neq f(y)$ in which case (S2) fails, since $f(x) \neq f(y)$ and $f(z) \notin \{f(x), f(y)\}$, while $xyz \in R^+$. This again contradicts (iii).

This concludes the proof of (iii)$\Rightarrow$(iv).

(iv)$\Rightarrow$(iii): Assume (iv). We describe a strategy for Prover that will satisfy (iii). As before, let f denote the mapping produced during the game. When asked to offer values for z, Prover offers values as follows.

(1) If there exist $x, y \in V(\mathbf{G})$ where $x \prec y \prec z$, $xyz \in R^+$, $f(x) \neq f(y)$, then Prover offers $\{f(x), f(y)\}$.

(2) Else if there exist $x, y \in V(\mathbf{G})$ where $x \prec y \prec z$, $xyz \in R^-$, $f(x) \neq f(y)$, then Prover offers $\{1,2,3,4\} \setminus \{f(x), f(y)\}$.
(3) Else if there exists $x \in V(\mathbf{G})$ with $x \prec z$ and $xz \in F$, then Prover offers any two values different from $f(x)$.
(4) Else Prover offers any two values.

We prove that this strategy satisfies the conditions of (iii). For contradiction, suppose that Adversary can play against this strategy so that the resulting mapping f fails one of the conditions (S1)-(S3).

Consider the first point of the game when the value $f(z)$ was assigned to z causing one of (S1)-(S3) to fail. Recall that we assume (iv). We examine the three possibilities as follows.

Case 1: Suppose that (S1) fails when the value is chosen for z. Namely, suppose that there is $a \in V(\mathbf{G})$ with $a \prec z$ where $az \in F$ and $f(a) = f(z)$. This means that $f(a)$ was one of the values offered by Prover for z. Recall that Prover offered values for z in steps (1)-(4) in that order.

Case 1.1: Suppose that Prover offered values for z in step (1). Then there exist vertices x, y where $x \prec y \prec z$ such that $xyz \in R^+$ and $f(x) \neq f(y)$, and Prover offered $\{f(x), f(y)\}$ for z. Thus $f(a) \in \{f(x), f(y)\}$, since $f(a) = f(z)$. It follows that $a \notin \{x, y\}$, since otherwise we contradict (iv). If $\{x, y\} \prec a$, then we have $xya \in R^-$ by (X3). But then (S3) is violated at a, since $f(a) \in \{f(x), f(y)\}$. Similarly, if $\{x, y\} \not\prec a$, then we have $xa, ya \in F$ by (X3), and (S1) is violated either at y if $f(a) = f(y)$, or else at a or x if $f(a) = f(x)$. This contradicts our choice of z, since $\{x, y, a\} \prec z$.

Case 1.2: Suppose that Prover offered values for z in step (2). Then there exist vertices x, y where $x \prec y \prec z$ such that $xyz \in R^-$ and $f(x) \neq f(y)$, and Prover offered for z the set $\{1,2,3,4\} \setminus \{f(x), f(y)\}$. Since $f(z)$ was chosen from this set, we have $f(z) \notin \{f(x), f(y)\}$. Recall that $f(a) = f(z)$. Thus also $f(a) \notin \{f(x), f(y)\}$ and hence $a \notin \{x, y\}$. This implies by (X2) that $xya \in R^+$. But since $f(a) \notin \{f(x), f(y)\}$ and $f(x) \neq f(y)$, we notice that $f(a), f(x), f(y)$ are pairwise distinct, and hence, (S2) is violated at either y or a, since $xya \in R^+$. This contradicts our choice of z, since $\{y, a\} \prec z$.

Case 1.3: Suppose that Prover offered values for z in step (3). Then there exists a vertex x where $x \prec z$ such that $xz \in F$ and Prover offered for z a set of two distinct values, neither of which was $f(x)$. Since $f(z)$ was chosen from this set, we have $f(z) \neq f(x)$. Recall that $f(a) = f(z)$. Thus $f(a) \neq f(x)$ and $a \neq x$. From this we deduce using (X1) that $xaz \in R^-$. Consequently, Prover should have offered values in step (2), never reaching step (3), since $f(a) \neq f(x)$ and $xaz \in R^-$. Thus Prover never reached step (3), a contradiction.

Case 1.4: Suppose that Prover offered values for z in step (4). Since step (4) was reached, there is no x such that $x \prec z$ and $xz \in F$. Thus it is impossible that (S1) failed when the value for z was chosen, a contradiction.

Case 2: Suppose that (S2) fails when the value is chosen for z. Namely, suppose that there exist vertices a, b where $a \prec b \prec z$ and $abz \in R^+$ such that $f(a) \neq f(b)$ and $f(z) \notin \{f(a), f(b)\}$. Note that this implies that Prover offered values for z in step (1), since we may always take $x = a$ and $y = b$ to satisfy the conditions of step (1). Thus we only need to consider this possibility. Namely, we have that there exist vertices x, y where $x \prec y \prec z$ such that $xyz \in R^+$ and $f(x) \neq f(y)$, and Prover offered $\{f(x), f(y)\}$ for z. Thus $f(z) \in \{f(x), f(y)\}$. Recall that $f(z) \notin \{f(a), f(b)\}$. Hence $\{a, b\} \neq$

$\{x, y\}$. Moreover, since $f(a) \neq f(b)$ and $f(x) \neq f(y)$, it follows that $\{f(a), f(b)\} \neq \{f(x), f(y)\}$, since $\{f(x), f(y)\}$ contains $f(z)$, and $\{f(a), f(b)\}$ does not.

Assume first that $\{x, y\}$ is disjoint from $\{a, b\}$. If $y \prec b$, then we deduce using (X6) that $xya, xyb \in R^+$. This means that (S2) fails at b or at one of y, a, since $f(x) \neq f(y)$ and $\{f(a),\ f(b)\} \neq \{f(x), f(y)\}$. Similarly, if $b \prec y$, we deduce using (X6) that $abx, aby \in R^+$, and (S2) fails at y or at one of b, x, since $f(a) \neq f(b)$ and $\{f(x), f(y)\} \neq \{f(a), f(b)\}$. This contradicts our choice of z, since $\{x, y, a, b\} \prec z$.

So we may assume that $\{x, y\}$ intersects $\{a, b\}$. If $y \in \{a, b\}$, then $xab \in R^+$ by (X5). Recall that $\{f(a), f(b)\} \neq \{f(x), f(y)\}$. Since $y \in \{a, b\}$, we deduce that $f(x) \notin \{f(a), f(b)\}$. This implies that $f(x), f(a), f(b)$ are pairwise distinct, since also $f(a) \neq f(b)$. Thus (S2) fails at b, since $xab \in R^+$. Similarly, if $x \in \{a, b\}$, then $aby \in R^+$ by (X5) and we have $f(y) \notin \{f(a), f(b)\}$. Hence, $f(a), f(b), f(y)$ are pairwise distinct and so (S2) fails at y or b, since $aby \in R^+$. This contradicts our choice of z, since $\{y, b\} \prec z$.

Case 3: Suppose that (S3) fails when the value is chosen for z. Namely, suppose that there exist vertices a, b where $a \prec b \prec z$ and $abz \in R^-$ such that $f(a) \neq f(b)$ and $f(z) \in \{f(a), f(b)\}$. Note that this implies that Prover offered values for z in either step (1) or step (2), since we may always take $x = a$ and $y = b$ to satisfy the conditions of step (2). Thus we only need to consider the steps (1) and (2).

Case 3.1: Suppose that Prover offered values for z in step (1). Then there exist vertices x, y where $x \prec y \prec z$ such that $xyz \in R^+$ and $f(x) \neq f(y)$, and Prover offered $\{f(x), f(y)\}$ for z. Thus $f(z) \in \{f(x), f(y)\}$. Recall that $f(z) \in \{f(a), f(b)\}$. We deduce that $\{f(x), f(y)\} \cap \{f(a), f(b)\} \neq \emptyset$.

Assume first that $\{a, b\}$ and $\{x, y\}$ are disjoint. If $y \prec b$, then we deduce using (X7) that $xyb \in R^-$ and either $xya \in R^-$ if $\{x, y\} \prec a$, or else $xa, ya \in F$. Thus if $f(b) \in \{f(x), f(y)\}$, then (S3) fails at b, since $f(x) \neq f(y)$ and $xyb \in R^-$. So we may assume that $f(b) \notin \{f(x), f(y)\}$ which yields that $f(a) \in \{f(x), f(y)\}$, since $\{f(a), f(b)\} \cap \{f(x), f(y)\} \neq \emptyset$. Thus if $\{x, y\} \prec a$, we have $xya \in R^-$ and so (S3) fails at a, since $f(x) \neq f(y)$. If $\{x, y\} \not\prec a$, we have $xa, ya \in F$ in which case (S1) fails at either y or one of x, a. Similarly, if $b \prec y$, we have by (X7) that $aby \in R^-$ and either $abx \in R^-$ if $\{a, b\} \prec x$, or $ax, bx \in F$ if otherwise. Thus either (S3) fails at y if $f(y) \in \{f(a), f(b)\}$, or we have $f(x) \in \{f(a), f(b)\}$ in which case either (S3) fails at x if $\{a, b\} \prec x$, or (S1) fails at a or b or x if $\{a, b\} \not\prec x$. This contradicts our choice of z, since $\{x, y, a, b\} \prec z$.

Thus we may assume that $\{a, b\}$ intersects $\{x, y\}$. Recall that $a \prec b$ and $x \prec y$. We observe that if $x \in \{a, b\}$ or $y = a$, then we contradict (iv), the second or third condition thereof, respectively. Thus it follows that $x \notin \{a, b\}$ and $y = b$. From this we deduce using (X4) that $ax \in F$ and $axy \in R^+$. We recall that $f(x) \neq f(y)$ and $f(a) \neq f(b)$. Since $y = b$, we deduce that $f(y) \notin \{f(a), f(x)\}$. Thus either $f(a) = f(x)$ and (S1) fails at one of a, x, since $ax \in F$, or we have $f(a) \neq f(x)$ in which case (S2) fails at y, since $axy \in R^+$ and $f(y) \notin \{f(a), f(x)\}$. This again contradicts our choice of z, since $\{a, x, y\} \prec z$.

Case 3.2: Suppose that Prover offered values for z in step (2). Then there exist vertices x, y where $x \prec y \prec z$ such that $xyz \in R^-$ and $f(x) \neq f(y)$, and Prover offered for z the set $\{1, 2, 3, 4\} \setminus \{f(x), f(y)\}$. Since $f(z)$ was chosen from this set, we have $f(z) \notin \{f(x), f(y)\}$. Recall that $f(z) \in \{f(a), f(b)\}$ and $f(a) \neq f(b)$. We deduce that $\{f(a), f(b)\} \neq \{f(x), f(y)\}$ and so $\{a, b\} \neq \{x, y\}$. Now we proceed exactly as in Case 2.

If $\{x, y\}$ is disjoint from $\{a, b\}$, we consider two cases: $y \prec b$ or $b \prec y$. If $y \prec b$, then $xya, xyb \in R^+$ by (X6), while if $b \prec y$, we have $abx, aby \in R^+$. In either case, we deduce that (S2) fails at one of x, y, a, b, since $\{f(a), f(b)\} \neq \{f(x), f(y)\}$. If $\{x, y\} \cap \{a, b\} \neq \emptyset$, then we again have two cases: $y \in \{a, b\}$ or $x \in \{a, b\}$. If $y \in \{a, b\}$, we have $xab \in R^+$ by (X5) and we deduce that $f(x)$, $f(a)$, $f(b)$ are pairwise distinct. Thus (S2) fails at b, since $xab \in R^+$. If $x \in \{a, b\}$, then $aby \in R^+$ by (X5) and $f(y)$, $f(a)$, $f(b)$ are pairwise distinct. Thus (S2) fails at b or y, since $aby \in R^+$. This contradicts our choice of z, since $\{y, b\} \prec z$.

This exhausts all possibilities. Thus we conclude that no such vertex z exists which proves that the strategy for Prover described in steps (1)-(4) is indeed a strategy satisfying the conditions of (iii). Therefore (iv)⇒(iii). This completes the proof. ◻

We now have all pieces to prove the main theorem of this section.

THEOREM 4.9. *$\{2\}$-CSP($\mathbf{K}_4$) is decidable in polynomial time.*

Proof. By Theorem 4.8, it suffices to construct the sets F, R^+, and R^-, and check the conditions of item (iv) of the said theorem. This can clearly be accomplished in polynomial time, since each of the three sets contains at most n^3 elements, where n is the number of variables in the input formula, and elements are only added (never removed) from the sets. Thus either a new pair (triple) needs to be added as follows from one of the rules (X1)-(X7), or we can stop and output the resulting sets. ◻

5. Cycles: proof of Theorem 1.2. Similarly to §4, we discuss the cases of Theorem 1.2 individually. The case (i) will be proved as Proposition 5.2, and the case (ii) follows from [22]. Finally, the case (iii) will be proved as Proposition 5.3.

Recall that we denote by $\mathbf{C}_n$ the irreflexive symmetric cycle on n vertices. We consider $\mathbf{C}_n$ to have vertex set $\{0, 1, \ldots, n-1\}$ and edge set $\big\{(i, j) \ : \ |i - j| \in \{1, n-1\}\big\}$.

In the forthcoming proof, we use the following elementary observation from additive combinatorics. Let $n \geq 2$, $j \geq 1$, and A, B be sets of integers. Define:

- $A +_n B = \big\{(a + b) \bmod n \ \big| \ a \in A,\, b \in B\big\}$
- $j \times_n A = \underbrace{A +_n \ldots +_n A}_{j \text{ times}}$

LEMMA 5.1. *Let $n \geq 3$ and $2 \leq j < n$. Then*

$$\Big| j \times_n \{-1, +1\} \Big| = \begin{cases} j+1 & n \text{ is odd} \\ \min\big\{j+1, n/2\big\} & n \text{ is even} \end{cases}$$

$$\Big| n \times_n \{-1, +1\} \Big| = \Big| n \times_n \{-2, 0, +2\} \Big| = \begin{cases} n & n \text{ is odd} \\ n/2 & n \text{ is even} \end{cases}$$

PROPOSITION 5.2. *If $n \geq 3$, then X-CSP($\mathbf{C}_n$) is in L if $n = 4$, or $1 \notin X$, or n is even and $X \cap \{2, 3 \ldots, n/2\} = \emptyset$,*

Proof. Let Ψ be an instance of X-CSP($\mathbf{C}_n$). Recall that we denote by $\mathbf{D}_\psi$ the graph corresponding to the quantifier-free part of Ψ, and write $x \prec y$ if x is quantified before y in Ψ. For an edge xy of $\mathbf{D}_\psi$ where $x \prec y$, we say that x is a *predecessor* of y. Note that a vertex can have several predecessors.

The following claims restrict the yes-instances of X-CSP($\mathbf{C}_n$).

Let x be a vertex of $\mathbf{D}_\psi$ quantified in Ψ by $\exists^{\geq j}$ for some j. If $\mathbf{C}_n \models \Psi$ then

(a) *if $j \geq 3$, then x has no predecessors,*

(b) *if n is even and $j > n/2$, then x is the first vertex (w.r.t. $\prec$) of some connected component of $\mathbf{D}_\psi$, and*

(c) *if $n \neq 4$ and $j = 2$, then all predecessors of x except for its first predecessor (w.r.t. $\prec$) are quantified by $\exists^{\geq 1}$.*

For (a), let y be a predecessor of x. Then for the value i chosen by Adversary for y, Prover must offer a set of at least three vertices of $\mathbf{C}_n$ that are adjacent to i in $\mathbf{C}_n$. Since there are only two such vertices, Adversary can always choose for x a vertex non-adjacent to i at which point Prover loses.

For (b), let y be the first vertex of the connected component of $\mathbf{D}_\psi$ that contains x. Assume $y \neq x$ and consider the path $\mathbf{P}$ between x and y in $\mathbf{D}_\psi$. Without loss of generality, we may assume that the value i chosen by Adversary for y is even. Note that, because n is even, if the length of $\mathbf{P}$ is also even, then Adversary must choose an even value for x, while if the length is odd, he must choose an odd value (otherwise Prover loses). However, as $j > n/2$, the set provided by Prover for x contains both an even and an odd number. Thus Adversary is allowed to choose for x the wrong parity and Prover loses.

For (c), suppose that y and z with $y \prec z$ are predecessors of x where z is quantified by $\exists^{\geq j'}$ for some $j' \geq 2$. If i is the value chosen by Adversary for y, then Prover must offer for z a set of $j' \geq 2$ values which hence must contain at least one value different from i. Adversary then chooses this value i' after which Prover must offer for x two distinct vertices i'', i''' of $\mathbf{C}_n$ adjacent to both i and i'. But then i, i', i'', i''' yield a 4-cycle in $\mathbf{C}_n$, impossible if $n \neq 4$.

Using these claims, we prove the proposition. First, we consider the case $n = 4$. We show that $\{1,2,3,4\}$-CSP($\mathbf{C}_4$) is in L. This will imply that X-CSP($\mathbf{C}_4$) is in L for every X. Observe that if $\mathbf{D}_\psi$ contains a vertex x quantified by $\exists^{\geq 3}$ or $\exists^{\geq 4}$, then by (b) this vertex is the first in its component (if Ψ is not a trivial no-instance). Thus replacing its quantification by $\exists^{\geq 1}$ does not change the truth of Ψ. So we may assume that Ψ is an instance of $\{1,2\}$-CSP($\mathbf{C}_4$). We now claim that $\mathbf{C}_4 \models \Psi$ if and only if $\mathbf{D}_\psi$ is bipartite. Clearly, if $\mathbf{D}_\psi$ is not bipartite, it has no homomorphism to $\mathbf{C}_4$ and hence $\mathbf{C}_4 \not\models \Psi$. Conversely, assume that $\mathbf{D}_\psi$ is bipartite with bipartition (A, B). Our strategy for Prover offers the set $\{0, 2\}$ or its subsets for the vertices in A and offers $\{1, 3\}$ or its subsets for every vertex in B. It is easy to verify that this is a winning strategy for Prover. Thus $\mathbf{C}_4 \models \Psi$. The complexity now follows as checking (b) and checking if a graph is bipartite is in L by [33].

Now, we may assume $n \neq 4$, and next we consider the case $1 \notin X$. If also $2 \notin X$, then by (a) the graph $\mathbf{D}_\psi$ contains no edges (otherwise Ψ is a trivial no-instance). This is clearly easy to check in L. Thus $2 \in X$. We claim that if we satisfy (a) and (c), then $\mathbf{C}_n \models \Psi$. We provide a winning strategy for Prover. Namely, for a vertex x, if x has no predecessors, offer any set for x. If x has a unique predecessor y for which the value i was chosen, then x is quantified by $\exists^{\geq 2}$ (or $\exists$) by (a) and we offer $\{i-1, i+1\}$ (mod n) for x . There are no other cases by (a) and (c). It follows that Prover always wins with this strategy. In terms of complexity, it suffices to check (a) and (c) which is in L.

Finally, suppose that n is even and $X \cap \{2, \ldots, n/2\} = \emptyset$. Note that every vertex of $\mathbf{D}_\psi$ is either quantified by $\exists^{\geq 1}$ or by $\exists^{\geq j}$ where $j > n/2$. Thus, using (b), unless Ψ is a trivial no-instance, we can again replace every $\exists^{\geq j}$ in Ψ by $\exists^{\geq 1}$ without changing the truth of Ψ. Hence, we may assume that Ψ is an instance of $\{1\}$-CSP($\mathbf{C}_n$). Thus, as n is even, $\mathbf{C}_n \models \Psi$ if and only if $\mathbf{D}_\psi$ is bipartite. The complexity again follows from [33]. This concludes the proof. □

Proposition 5.3. *Let $n \geq 3$. Then X-CSP($\mathbf{C}_n$) is* Pspace*-complete if $n \neq 4$ and $\{1, j\} \subseteq X$: where $j \in \{2, \ldots, n\}$ if n is odd and $j \in \{2, \ldots, n/2\}$ if n is even.*

Proof. The proof is by reduction from QCSP($\mathbf{C}_n$) for odd n, and from QCSP($\mathbf{K}_{n/2}$) for even n. Both problems are known to be Pspace-hard [4].

First, consider the case of odd n. Let Ψ be an instance of QCSP($\mathbf{C}_n$). In other words, Ψ is an instance of $\{1,n\}$-CSP($\mathbf{C}_n$). Clearly, $j < n$ otherwise we are done.

We modify Ψ by replacing each universally-quantified variable x of Ψ by a path. Namely, let π_x denote the pp-formula that encodes that

$$x_1^1, x_2^1, \ldots, x_{j-1}^1,\ x_1^2, x_2^2, \ldots, x_{j-1}^2, \quad \ldots \quad, x_1^n, x_2^n, \ldots, x_{j-1}^n,\ x$$

is a path in that order (all but x are new variables). We replace $\forall x$ by

$$Q_x \ = \ \exists^{\geq j} x_1^1\ \exists^{\geq j} x_1^2 \ldots\ \exists^{\geq j} x_1^n\ \exists^{\geq j} x\ \exists^{\geq 1} x_2^1 \ldots \exists^{\geq 1} x_{j-1}^1\ \ldots\ \exists^{\geq 1} x_2^n \ldots \exists^{\geq 1} x_{j-1}^n$$

and append π_x to the quantifier-free part of the formula. Let Ψ' denote the final formula after considering all universally quantified variables. Note that Ψ' is an instance of $\{1,j\}$-CSP($\mathbf{C}_n$).

We argue that $\mathbf{C}_n \models \Psi$ if and only if $\mathbf{C}_n \models \Psi'$. To see this, it suffices to show that Ψ' correctly simulates the universal quantifiers of Ψ. Namely, it suffices to prove that $\mathbf{C}_n \models Q_x\pi_x$, and for each $\ell \in \{0,\ldots,n-1\}$, Adversary has a strategy on $Q_x\pi_x$ that evaluates x to ℓ.

For the first part, we provide a strategy for Prover. We treat x as x_1^{n+1}. For x_1^1, Prover offers any set. For x_1^k where $k \geq 2$, let i be the value chosen by Adversary for x_1^{k-1}. By Lemma 5.1, we observe that there are exactly j vertices in $\mathbf{C}_n$ having a walk to i of length $j-1$. Prover offers this set for x_1^k. This allows her to choose values for $x_2^{k-1} \ldots x_{j-1}^{k-1}$ as the path $x_1^{k-1}, \ldots x_{j-1}^{k-1}, x_1^k$ encodes precisely the fact that there exists a walk of length $j-1$ between the values chosen for x_1^{k-1} and x_1^k. Thus $\mathbf{C}_n \models Q_x\pi_x$.

For the second part, consider any $\ell \in \{0,\ldots,n-1\}$. We explain a strategy for Adversary that allows him to choose ℓ for x. First, Adversary chooses any value for x_1^1. Let i_0 be this value, and by the second part of Lemma 5.1, choose a sequence of n numbers $i_1, i_2, \ldots, i_n$ either all from $\{-1,+1\}$ if j is odd, or all from $\{-2,0,+2\}$ if j is even, such that $i_0 + i_1 + i_2 + \ldots + i_n = \ell$. After this, consider inductively $k \geq 2$ and let i be the value chosen by Adversary for x_1^{k-1}. By Lemma 5.1, there are exactly j possible values that Prover can offer if she does not want to lose. Thus Prover is forced to offer all these values. In particular, if j is even, this set contains values $i+1$ and $i-1 \pmod n$ while if j is odd, the set contains values $i+2$, i, and $i-2 \pmod n$. Thus Adversary is allowed to choose the value $i + i_{k-1} \pmod n$ for x_1^k. This shows that Adversary is allowed to choose the value $i_0 + i_1 + \ldots + i_n = \ell$ for $x_1^{n+1} = x$.

Thus, this proves that Ψ' correctly simulates the universal quantifiers of Ψ, and consequently $\mathbf{C}_n \models \Psi$ if and only if $\mathbf{C}_n \models \Psi'$. For odd n, this completes the proof of the claim that $\{1,j\}$-CSP($\mathbf{C}_n$) is Pspace-hard.

It remains to investigate the case of even n. Recall that $n \geq 6$ and $j \leq n/2$. We show a reduction from QCSP($\mathbf{K}_{n/2}$) to $\{1,j\}$-QCSP($\mathbf{C}_n$). The reduction is a variant of the construction from [18] for the problem of retraction to even cycles.

Let Ψ be an instance of QCSP($\mathbf{K}_{n/2}$), and define $r = (-n/2-2) \bmod (j-1)$ (recalling anything mod 1 is 0). We construct a formula Ψ' from Ψ as follows. First, we modify Ψ by replacing universal quantifiers exactly as in the case of odd n. Namely, we define Q_x and π_x as before, replace each $\forall x$ by Q_x, and append π_x to the quantifier-free part of the formula. After this, we append to the formula a cycle on n vertices $v_0, v_1, \ldots, v_{n-1}$ with a path on $r+1$ vertices $w_0, w_1, \ldots, w_r$. (See the black vertices in Figure 5.1.) Then, for each edge xy of $\mathbf{D}_\psi$, we replace $E(x,y)$ in Ψ by the gadget depicted in Figure 5.1 (consisting of the cartesian product of $\mathbf{C}_n$ and a path on $3n/2$ vertices together with two attached paths on $n/2-2$, resp. $r+1$ vertices). The vertices x and y represent the variables x and y while all other white vertices are new

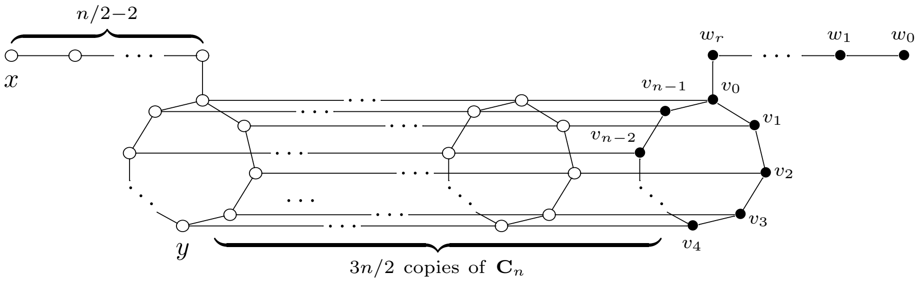


FIG. 5.1. *The gadget for the case of even n where $r = (-n/2 - 2) \bmod (j-1)$.*

variables, and the black vertices are identified with $v_0, \ldots, v_{n-1}, w_0, \ldots, w_r$ introduced in the previous step.

Finally, we prepend the following quantification to the formula:

$$\exists^{\geq 1} w_0 \, \exists^{\geq j} v_{j-r-2} \, \exists^{\geq j} v_{2j-r-3} \ldots \exists^{\geq j} v_{(k \cdot j-r-k-1)} \ldots \exists^{\geq j} v_{n/2+1}$$

followed by $\exists^{\geq 1}$ quantification of all the remaining variables of the gadgets.

We prove that $\mathbf{K}_{n/2} \models \Psi$ if and only if $\mathbf{C}_n \models \Psi'$. First, we show that Ψ' correctly simulates the universal quantification of Ψ. The argument for this is essentially the same as in the case of odd n. Next, we analyse possible assignments to the vertices $v_0, \ldots, v_{n-1}$. For clarity, we define $\alpha_k = kj - r - k - 1$ and note that $n/2 + 1 = \alpha_k$ for $k = \lceil \frac{n+4}{2(j-1)} \rceil$. By the symmetry of $\mathbf{C}_n$, we assume that Adversary chooses for w_0 the value $n - r - 1$. The next quantified vertex is $v_{j-r-2} = v_{\alpha_1}$ in distance $j - 1$ from w_0. Thus, by Lemma 5.1, there are exactly j values that Prover can and must offer. Among them, we find $j - r - 2 = \alpha_1$. Similarly, for $2 \leq k \leq \lceil \frac{n+4}{2(j-1)} \rceil$, the vertex $v_{\alpha_{k-1}}$ is in distance $j - 1$ from v_{α_k}, and hence, Prover is forced to offer a set of j values only depending on the value chosen for $v_{\alpha_{k-1}}$. In particular, if α_{k-1} was chosen for $v_{\alpha_{k-1}}$, then Adversary can choose α_k for v_{α_k}. This argument also shows that if Prover acts as we describe, then in every possible case she can complete the homomorphism for the path $w_0, w_1, \ldots, w_r, v_0 \ldots, v_{n/2+1}$. Further, she also has a way of assigning the values to $v_{n/2+2}, \ldots, v_{n-1}$. This can be seen as follows. First, note that the distance between $v_{n/2+1}$ and v_0 is $n/2 - 1$. Thus, if $n/2$ is odd, then the values assigned to v_0 and $v_{n/2+1}$ have the same parity because $n/2 + 1$ is even and we observe that between any two vertices of the same parity in $\mathbf{C}_n$ there exists a walk of length $n/2 - 1$. Similarly, if $n/2$ is even, the values chosen for $v_0, v_{n/2+1}$ have different parity and between any two vertices of $\mathbf{C}_n$ of different parity there is a walk of length $n/2 - 1$.

This concludes the argument for the vertices $v_0, \ldots, v_{n-1}$. It implies two possible types of outcomes: either the values chosen for $v_0, \ldots, v_{n-1}$ are all distinct, or not. To obtain the former case, for each $1 \leq k \leq \lceil \frac{n+4}{2(j-1)} \rceil$, Adversary chooses α_k for v_{α_k}. This forces assigning i to v_i for all $i \in \{0, \ldots, n/2 + 1\}$ and thus consequently also for all the other v_i's. We shall assume this situation first. For all other (degenerate) cases we use a different argument explained later.

Thus assuming that $v_0, \ldots, v_{n-1}$ get assigned values $0, \ldots, n - 1$ in that order and the values for the original variables of Ψ are chosen, we argue that Prover can finish the homomorphism if and only if the assignment to the variables of Ψ is a proper coloring for $\mathbf{D}_\psi$. This follows exactly as in [18]. Namely, in every gadget, each copy of $\mathbf{C}_n$ is forced to copy the assignment from the adjacent copy of $\mathbf{C}_n$, shifted by $+1$ or by -1 (mod n). In particular, if i is the value assigned to y, the vertex z opposite y in the last copy of $\mathbf{C}_n$ is necessarily assigned value $n/2 + i$ (mod n). This implies that the value assigned to x is different from i as the path from z to x is too short (of length less than $n/2$). On the other hand, this path is long enough so that any value of the same parity as i but different from i can be chosen for x such that the homomorphism can be completed. This precisely simulates the edge predicate of Ψ. Finally, we observe that Prover can choose whether consecutive copies of $\mathbf{C}_n$ are shifted by $+1$ or -1 and there are exactly $3n/2$ copies of $\mathbf{C}_n$. Thus, by Lemma 5.1, every possible odd number from $\{0, \ldots, n-1\}$ can be chosen for y by a particular series of shifts. It follows that $\{1, 3, \ldots, n-1\}$ is precisely the set colors we use to simulate QCSP$(\mathbf{K}_{n/2})$.

Now, we discuss the degenerate cases. Namely, we show that, regardless of the assignment to $v_0, \ldots, v_{n-1}$, for each copy of the gadget (in Figure 5.1) there is a way to complete the homomorphism (by assigning the values to the white vertices) in such a way that if ℓ is the value assigned to y, then the vertex opposite y in the last copy of $\mathbf{C}_n$ is assigned value $\ell + n/2 \pmod n$. As this is exactly what happens in the non-degenerate case, the rest will follow. Note that consideration of these degenerate cases is the reason we use a chain of $3n/2$ copies of $\mathbf{C}_n$ in the gadget of Figure 5.1, instead of the $n/2$ used in the like gadget in [17].

Since we assume that the vertices $v_0, \ldots, v_{n-1}$ are assigned a proper subset of $\{0, \ldots, n-1\}$, it can be seen that they are assigned a circularly consecutive subset of these numbers, and this subset is of size at most $n/2+1$ as otherwise the assignment cannot be a homomorphism. (Recall that in the proof we argue that we can assume that the assignment to $v_0, \ldots, v_{n-1}$ is a homomorphism).

For simplicity, let λ denote the assignment constructed so far, i.e., a mapping from the assigned vertices to their assigned values. We explain how to complete this assignment for the gadget so that it becomes a homomorphism to $\mathbf{C}_n$.

The gadget contains $3n/2$ copies of $\mathbf{C}_n$. We consider them from the right to left, namely $\{v_0, \ldots, v_{n-1}\}$ is the 1st copy, and the $3n/2$-th copy is the one containing y. With this in mind, we denote by $v_0^i, \ldots, v_{n-1}^i$ the respective copies of $v_0, \ldots, v_{n-1}$ in the i-th copy of $\mathbf{C}_n$. In particular, y is the vertex $v_{n/2}^{3n/2}$.

We describe the assignment to the copies of $\mathbf{C}_n$ in three phases. In the first phase, we assign values to the first $n/2$ copies. Consider $1 \leq i < n/2$, and assume that the vertices $v_0^i, \ldots, v_{n-1}^i$ are assigned values between a and b (inclusive) in the clock-wise order. Then the assignment to the $(i+1)$-st copy of $\mathbf{C}_n$ is as follows. For $k \in \{0, \ldots, n-1\}$, if $\lambda(v_k^i) = a$, then we set $\lambda(v_k^{i+1}) = a+1 \pmod n$, otherwise we set $\lambda(v_k^{i+1}) = \lambda(v_k^i) - 1 \pmod n$. It is easy to verify that this constitutes a homomorphism to $\mathbf{C}_n$. It follows that $\left|\lambda\left(\{v_0^{n/2}, \ldots, v_{n-1}^{n/2}\}\right)\right| = 2$.

Next, we explain the assignment to the second $n/2$ copies of $\mathbf{C}_n$. Let ℓ be the value assigned to y. We choose the values for the second $n/2$ copies in such a way that consecutive copies of $\mathbf{C}_n$ are just shifted by $+1$ or -1. We can choose an appropriate sequence of $+1, -1$ shifts so that the value assigned to $v_{n/2}^n$ is exactly $\ell + n/2 \pmod n$. (The argument about the parity of these values is the same as in the non-degenerate case.) We further conclude that $\left|\lambda\left(\{v_0^n, \ldots, v_{n-1}^n\}\right)\right| = 2$.

The assignment to the final $n/2$ copies is as follows. For $n \leq i < 3n/2$, again assume that the vertices $v_0^i, \ldots, v_{n-1}^i$ are assigned values between a and b in the clockwise order. Then for $k \in \{0, \ldots, n-1\} \setminus \{n/2\}$, if $\lambda(v_k^i) = a+1 \pmod n$ and $\lambda\left(v_{(k-1) \bmod n}^i\right) = \lambda\left(v_{(k+1) \bmod n}^i\right) = a$, then we set $\lambda(v_k^{i+1}) = a$, and otherwise we set $\lambda(v_k^{i+1}) = \lambda(v_k^i) + 1$. Again, we conclude that this constitutes a homomorphism, and it follows that $\left|\lambda\left(\{v_0^{3n/2}, \ldots, v_{n-1}^{3n/2}\}\right)\right| = n/2 + 1$. In particular, we observe that

$\lambda\big(y = v^{3n/2}_{n/2}\big) = \ell$ and $\lambda\big(v^{3n/2}_0\big) = \ell + n/2 \pmod n$.

This concludes the proof. □

6. Extensions of the CSP: simple cases. In this section we consider single-quantifier extensions of the classical CSP($\mathbf{B}$), namely the problem $\{1, j\}$-CSP($\mathbf{B}$) for some $1 < j \leq |B|$. The results will follow as outlined in the introduction.

6.1. Logspace cases of bipartite graphs. In the case of (irreflexive, undirected) graphs, it is known that $\{1\}$-CSP($\mathbf{H}$) = CSP($\mathbf{H}$) is in L if $\mathbf{H}$ is bipartite and is NP-complete otherwise [22] (for membership in L, one needs also [33]).

It is also known that something similar holds for $\{1, |H|\}$-CSP($\mathbf{H}$) = QCSP($\mathbf{H}$) – this problem is in L if $\mathbf{H}$ is bipartite and is NP-hard otherwise [29]. Of course, the fact that $\{1, j\}$-CSP($\mathbf{H}$) is hard on non-bipartite $\mathbf{H}$ is clear, but we will see that it is not always easy on bipartite $\mathbf{H}$. In the rest of this section, we look at particular cases where the problem $\{1, j\}$-CSP($\mathbf{H}$) is in L.

Proposition 6.1. *Let $\mathbf{K}_{k,\ell}$ be the complete bipartite graph with partite sets of size k and ℓ. Then $\{1, \ldots, k + \ell\}$-CSP($\mathbf{K}_{k,\ell}$) is in* L.

Proof. We reduce to QCSP($\mathbf{K}^1_2$), where $\mathbf{K}^1_2$ indicates $\mathbf{K}_2$ with one vertex named by a constant, say 1. QCSP($\mathbf{K}^1_2$) is equivalent to QCSP($\mathbf{K}_2$) (identify instances of 1 to a single vertex) and both are well-known to be in L (see, e.g., [29]). Let Ψ be input to $\{1, \ldots, k + \ell\}$-CSP($\mathbf{K}_{k,\ell}$). Produce Ψ' by substituting quantifiers $\exists^{\geq j}$ with $\exists$, if $j \leq \min\{k, \ell\}$, or with $\forall$, if $j > \max\{k, \ell\}$. Variables quantified by $\exists^{\geq j}$ for $\min\{k, \ell\} < j \leq \max\{k, \ell\}$ should be replaced by the constant 1. It is easy to see that $\mathbf{K}_{k,\ell} \models \Psi$ iff $\mathbf{K}_2 \models \Psi'$, and the result follows. □

Proposition 6.2. *Let $\mathbf{H}$ be a bipartite graph with vertices properly colored black and white such that no j vertices of the same color belong to the same connected component of $\mathbf{H}$. Then $\{1, j\}$-CSP(H) is in* L.

Proof. We will consider an input Ψ to $\{1, j\}$-CSP(H) of the form $Q_1x_1\ Q_2x_2 \ldots Q_mx_m\ \psi(x_1, x_2, \ldots, x_m)$. An $\exists^{\geq j}$ variable is called *trivial* if it has neither a path to another (distinct) $\exists^{\geq j}$ variable in $\mathbf{D}_\psi$, nor a path to an $\exists$ variable that precedes it in the quantifier order $\prec$. The key observation here is that any non-trivial $\exists^{\geq j}$ variable *must* be evaluated on vertices of both colors of some connected component. If in Ψ there is a non-trivial $\exists^{\geq j}$ variable, then Ψ must be a no-instance (as $\exists^{\geq j}$s must be evaluated on vertices of both colors of a connected component, and a path cannot be both even and odd in length). All other instances are readily seen to be satisfiable. Testing if Ψ contains a non-trivial $\exists^{\geq j}$ variable is in L by [33], and the result follows. □

Proposition 6.3. *If $\mathbf{H}$ is bipartite and contains $\mathbf{C}_4$, then $\Psi \in \{1, 2\}$-CSP($\mathbf{C}_4$) iff the underlying graph $\mathbf{D}_\psi$ of Ψ is bipartite. In particular, $\{1, 2\}$-CSP($\mathbf{H}$) is in* L.

Proof. Necessity is clear; sufficiency follows by the canonical evaluation of $\exists^{\geq 1}$ and $\exists^{\geq 2}$ on a fixed copy of $\mathbf{C}_4$ in $\mathbf{H}$. Membership in L follows from [33]. □

The path $\mathbf{P}_n$ has vertices $\{1, 2, \ldots, n\}$ and edges $\{ij\ :\ |i - j| = 1\}$.

Proposition 6.4. *For $n \leq 5$, the problem $\{1, 2\}$-CSP($\mathbf{P}_n$) is in* L.

Proof. Let Ψ be an instance of $\{1, 2\}$-CSP($\mathbf{P}_n$). As usual, let $\mathbf{G}$ be the graph $\mathbf{D}_\psi$ corresponding to Ψ, and let $\prec$ be the corresponding total order of $V(\mathbf{G})$. If $\mathbf{G}$ is not bipartite, then Ψ is a no-instance, and we reject immediately. This can be tested in L by [33]. We may therefore assume that the vertices of $\mathbf{G}$ are properly colored using colors black and white. Also we may assume that $\mathbf{G}$ is connected, otherwise we test each connected component individually. Let q be the first vertex in the ordering $\prec$. Then

(i) $\mathbf{P}_1 \models \Psi \iff \mathbf{G}$ is the single $\exists^{\geq 1}$ vertex q.

(ii) $\mathbf{P}_2 \models \Psi \iff \mathbf{G}$ does not contain an $\exists^{\geq 2}$ vertex except possibly for q.
(iii) $\mathbf{P}_3 \models \Psi \iff$ all $\exists^{\geq 2}$ vertices in $\mathbf{G}$ have the same color.
(iv) $\mathbf{P}_4 \models \Psi \iff$ all $\exists^{\geq 2}$ vertices in $\mathbf{G}$ are pairwise non-adjacent except possibly for q.
(v) $\mathbf{P}_5 \models \Psi \iff$ there is a color C (black or white) such that each edge xy between two $\exists^{\geq 2}$ vertices where $x \prec y$ is such that x has color C, and there is no neighbor of y before x in $\prec$.

The claim (i) is clear. For (ii), any $\exists^{\geq 2}$ vertex other than q must be offered both 1 and 2, one of which will violate the parity with respect to q (since $\mathbf{G}$ is connected). In all other cases, $\mathbf{P}_2 \models \Psi$ because $\mathbf{D}_\psi$ is bipartite.

A similar argument works for (iii), if there are two $\exists^{\geq 2}$ vertices u, v of different color where $u \prec v$, then Prover must offer 1 or 3 among the values for u and Adversary chooses it. She also must offer 1 or 3 among the values for v and again, Adversary chooses it. This now violates the parity, since u and v are in different sides of the bipartition of $\mathbf{G}$. Conversely, if all $\exists^{\geq 2}$ vertices are, say black, then Prover offers $\{1,3\}$ to all black $\exists^{\geq 2}$ vertices, 1 to all black $\exists^{\geq 1}$ vertices, and 2 to all white vertices. This will allow Prover to win.

For (iv), if $\mathbf{G}$ contains adjacent $\exists^{\geq 2}$ vertices u, v distinct from q where $u \prec v$, then Prover must offer $\{1,3\}$ or $\{2,4\}$ for u because of the parity with respect to q. Adversary chooses either 1 or 4 and Prover subsequently loses at v. Conversely, if no such vertices u, v exist, Prover first offers $\{2,3\}$ for q. By symmetry of the path $\mathbf{P}_4$, we may assume that Advesary chooses 2 for q, and that q is black. Prover subsequently offers 2 for each black $\exists^{\geq 1}$ vertex, and offers 3 for each white $\exists^{\geq 1}$ vertex. She offers $\{2,4\}$ for each $\exists^{\geq 2}$ black vertex and offers $\{1,3\}$ for each white $\exists^{\geq 2}$ vertex. Since no two $\exists^{\geq 2}$ vertices are adjacent, any Adversary's choices lead to a homomorphism, and so Prover always wins.

Finally, for (v), by symmetry, assume that q is black. Suppose that $\mathbf{G}$ contains edges xy and wz where x, y, w, z are $\exists^{\geq 2}$ vertices with $x \prec y$ and $w \prec z$, and where x is black and w is white (possibly $q = x$). If 2 or 4 is chosen for q, then $\{1,3\}$ or $\{1,5\}$ or $\{3,5\}$ is offered for w because of the parity (q is black and w is white). This allows Adversary to choose 1 or 5 for w and Prover loses at z. If one of $1,3,5$ is chosen for q, then 1 or 5 is among values offered for x because of parity. So Adversary can choose 1 or 5 for x and Prover loses at y. For the second part, suppose that x, y are adjacent $\exists^{\geq 2}$ vertices where $x \prec y$, and there exists a neighbor z of y with $z \prec x$. Since x is an $\exists^{\geq 2}$ vertex, Adversary can choose for x a value different from that chosen for z. Prover then loses at y because there is no 4-cycle in $\mathbf{P}_5$.

Conversely, suppose first that every edge xy where x, y are $\exists^{\geq 2}$ vertices is such that x is black (recall that q is black), and there is no neighbor of y before x in $\prec$. Prover offers $\{2,4\}$ for every black $\exists^{\geq 2}$ vertex. She offers 3 for each white $\exists^{\geq 1}$ vertex. For each black $\exists^{\geq 1}$ vertex z, Prover offers 2 if there exist vertices x, y where $x \prec z \prec y$ such that y is an $\exists^{\geq 2}$ vertex adjacent to both x and z, and the value 2 was chosen for x. Similarly, Prover chooses 4 if the value 4 was chosen for x. On the other hand, if $w \prec z$ with w and z adjacent and w (being an $\exists^{\geq 2}$ vertex) was chosen as 5, then z is chosen as 4. Similarly, if w was chosen as 1 then z is chosen as 2. In any other case, Prover chooses 2 or 4 arbitrarily. Lastly, each white $\exists^{\geq 2}$ vertex is offered $\{1,3\}$ if 2 was chosen for all its preceding neighbors, or $\{3,5\}$ if 4 was chosen on all its preceding neighbors. One of these two possibilities must happen as guaranteed by our rules for black vertices. This shows that Prover indeed always wins.

Thus we may assume that every edge xy where x, y are $\exists^{\geq 2}$ vertices is such that

x is white. Here Prover offers 3 or $\{3,5\}$ for q. Then we proceed similarly to before, switching black and white. ◻

A similar argument gives the following result.

PROPOSITION 6.5. $\{1,3\}$-CSP($\mathbf{P}_5$) *is in* L.

Proof. Let Ψ be an instance of $\{1,3\}$-CSP($\mathbf{P}_n$), considered as the graph $\mathbf{G} = \mathbf{D}_\psi$ together with the quantifier order $\prec$. Again, $\mathbf{G}$ must be bipartite otherwise it is a no-instance. (This can be determined in L.) Thus we may assume that the vertices of $\mathbf{G}$ are properly colored black and write. Further, we may assume that $\mathbf{G}$ is connected (the argument can be applied on each connected component individually). Let q be the first vertex in $\prec$. If there are no $\exists^{\geq 3}$ vertices, then this is a yes-instance. If q is the only $\exists^{\geq 3}$ variable, Prover offers $\{2,3,4\}$ for q, and then depending on the choice of Adversary, she offers 2 and 3 for the black and white vertices, respectively, or she offers 3 and 2 for the two colors, respectively. Prover wins with this strategy.

For the remaining cases, we have the following claim.

(vi) $\mathbf{P}_5 \models \Psi \iff$ each edge xy where $x \prec y$ is such that y is an $\exists^{\geq 1}$ variable, and all $\exists^{\geq 3}$ vertices have the same color, and any two are at distance at least 3.

For the forward direction, recall that we assume that $\mathbf{G}$ is connected. This implies that Prover in order to win must offer values of the same parity for vertices of the same color (including q, since there are other $\exists^{\geq 3}$ vertices). This implies that for any $\exists^{\geq 3}$ variable Prover must offer the set $\{1,3,5\}$, the only set of 3 distinct values of the same parity. Thus if two $\exists^{\geq 3}$ variables have diffent color, then Prover cannot offer $\{1,3,5\}$ to both and consequently loses. This shows that any such variables must be at an even distance. In particular, if they are at distance 2, then Adversary chooses 1 for one of them and 5 for the other, and Prover loses, since there is no value that can be offered for their common neighbor (irrespective of the order of the vertices).

Now, consider an edge xy where $x \prec y$ and y is an $\exists^{\geq 3}$ variable. If one of $1,3,5$ was chosen for x, then Prover must offer values 2 or 4 for y, but there are only two such values available. So Prover loses at y. If one of $2,4$ was chosen for x, then Prover must offer $\{1,3,5\}$ for y. Then Adversary chooses 1 or 5 for y if 4 or 2 respectively was chosen for x. Again Prover loses. This proves the forward direction.

Conversely, assume that every edge xy with $x \prec y$ is such that y is an $\exists^{\geq 1}$ variable, and all $\exists^{\geq 3}$ vertices have the same color, and any two are at distance at least 3.

By symmetry, we may assume that all $\exists^{\geq 3}$ vertices are black. Prover offers $\{1,3,5\}$ for every $\exists^{\geq 3}$ variable. For every black $\exists^{\geq 1}$ variable, Prover offers 3. For every white $\exists^{\geq 1}$ variable x, Prover offers 2 if 5 was not chosen for any neighbor of x preceding it. She offers 4 if 1 was not chosen for any neighbor of x preceding it. We argue that one of these two is always possible. If not, then x has neighbors $\{y,z\} \prec x$ such that 1 was chosen for y and 5 for z. By our rules, this means that y and z are $\exists^{\geq 3}$ variables. But they are now at distance 2 as witnessed by the path y,x,z, and we assume that no such vertices exist. Thus this is a winning strategy for Prover. ◻

We remark in passing the following proposition.

PROPOSITION 6.6. *If* $j \in \{2,...,n-3\}$ *then one may exhibit a bipartite graph* $\mathbf{H}_j$ *of size* n *such that* $\{1,j\}$-CSP($\mathbf{H}_j$) *is* Pspace-*complete.*

Proof. The case $j = 2$ follows from Theorem 1.2; assume $j \geq 3$. Take the graph $\mathbf{C}_6$ and construct $\mathbf{H}_j$ as follows. Augment $\mathbf{C}_6$ with $j-3$ independent vertices each with an edge to vertices 1, 3 and 5 of $\mathbf{C}_6$. Apply the proof of Theorem 1.2 with $\mathbf{H}_j$. ◻

We now show the above statement is tight, in the following sense.

Proposition 6.7. *Let* $\mathbf{H}$ *be a bipartite graph. Then* $\{1,j\}$-CSP($\mathbf{H}$) *is in* L *if* $j \in \{1, |H|-2, |H|-1, |H|\}$.

Proof. When $j = |H|$ we have QCSP($\mathbf{H}$) and we refer for the result to [29]. For $j = |H| - 1$, we argue as in the proof of Proposition 6.2 unless $\mathbf{H}$ is the complete bipartite (star) $\mathbf{K}_{1,\ell}$ (for some ℓ), in which case we appeal to Proposition 6.1. The case $j = |H| - 2$ is not much more complicated. If we do not fall as in the proof of Proposition 6.2 or under Proposition 6.1, then we are equivalent to $\{1,3\}$-CSP($\mathbf{P}_5$) and the result follows from Proposition 6.5. ☐

After these introductory results, we move on to another major result of the paper.

7. Algorithm for $\{1,2\}$**-CSP(T) for trees T.** Let $\mathbf{T}$ be a fixed tree. In this section, we describe a polynomial-time algorithm for the problem $\{1,2\}$-CSP($\mathbf{T}$). The algorithm will be based around metric properties [9] of the tree $\mathbf{T}$, and will follow from a characterization of yes-instances. First we look at some useful properties of the distance metric of $\mathbf{T}$.

7.1. Metric properties of trees. For vertices $x, y \in V(\mathbf{T})$, we write $\mathsf{dist}(x,y)$ to denote the distance between x and y in $\mathbf{T}$, i.e., the length of a shortest path in $\mathbf{T}$ between x and y. Note that since $\mathbf{T}$ is a tree, there is a unique path between x and y for all $x, y \in V(\mathbf{T})$. We write $\mathsf{ecc}(x)$ for the *eccentricity* of vertex x in $\mathbf{T}$, defined as $\mathsf{ecc}(x) = \max\{\mathsf{dist}(x,y) \mid y \in V(\mathbf{T})\}$. We write $\mathsf{rad}(\mathbf{T})$ and $\mathsf{diam}(\mathbf{T})$ for the *radius* and *diameter* of $\mathbf{T}$, respectively, where $\mathsf{rad}(\mathbf{T}) = \min\{\mathsf{ecc}(x) \mid x \in V(\mathbf{T})\}$, and $\mathsf{diam}(\mathbf{T}) = \max\{\mathsf{ecc}(x) \mid x \in V(\mathbf{T})\}$. In [9], it is proved that vertices of a tree satisfy the following so-called *4-point condition*.

Lemma 7.1. [9] *Let* x, y, z, w *be distinct vertices of* $\mathbf{T}$. *Then*

$$\mathsf{dist}(x,y) + \mathsf{dist}(z,w) \leq \max \left\{ \begin{array}{c} \mathsf{dist}(x,z) + \mathsf{dist}(y,w) \\ \mathsf{dist}(x,w) + \mathsf{dist}(z,y) \end{array} \right\}$$

A collection of sets $S_1, S_2, \ldots, S_t$ has the *strong Helly property* [21] if there exist indices i, j such that $(S_1 \cap \ldots \cap S_t) = (S_i \cap S_j)$. Note that any collection of subpaths of a path has the strong Helly property [21].

7.2. Definitions. Recall that an instance to $\{1,2\}$-CSP($\mathbf{T}$) is a sentence $\Psi := \exists^{\geq \beta_1} v_1 \; \exists^{\geq \beta_2} v_2 \; \cdots \; \exists^{\geq \beta_n} v_m \; \psi$, where ψ is a conjunction of atoms $E(v_i, v_j)$ for some i, j. As usual we think of this instance as the graph $\mathbf{G} = \mathbf{D}_\psi$ where $V(\mathbf{G}) = \{v_1, \ldots, v_m\}$ and $E(\mathbf{G}) = \{v_i v_j \mid E(v_i, v_j) \text{ appears in } \psi\}$, together with a linear order $\prec$ where $v_1 \prec v_2 \prec \ldots \prec v_m$. We write $X \prec Y$ if $x \prec y$ for each $x \in X$ and each $y \in Y$. Also, we write $x \prec Y$ in place of $\{x\} \prec Y$. We write $\beta(v_i)$ for β_i. Note that β is a function $\beta : V(\mathbf{G}) \to \{1,2\}$.

A *walk* Q of $\mathbf{G}$ consists of a sequence $x_1, x_2, \ldots, x_r$ of vertices of $\mathbf{G}$ where $x_i x_{i+1} \in E(\mathbf{G})$ for all $i \in \{1, \ldots, r-1\}$. A walk $x_1, \ldots, x_r$ is a *closed walk* if $x_1 = x_r$. Write $|Q|$ to denote the *length* of the walk Q (number of edges on Q).

Definition 7.2. *If* $Q = x_1, \ldots, x_r$ *is a walk of* $\mathbf{G}$*, we define* $\lambda(Q)$ *as follows:*

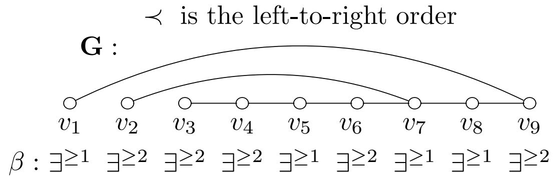
$$\lambda(Q) = |Q| - 2 \sum_{i=2}^{r-1} \big(\beta(x_i) - 1\big)$$

Put differently, we assign weights to the vertices of $\mathbf{G}$, with weight $+1$ assigned to each $\exists^{\geq 1}$ node, and weight -1 to each $\exists^{\geq 2}$ node; the value $\lambda(Q)$ is then simply one plus the total weight of all inner nodes in the walk Q.

Definition 7.3. *A walk* $x_1, \ldots, x_r$ *of* $\mathbf{G}$ *is a* <u>looping walk</u> *if* $x_1 \neq x_r$ *and if* $r \geq 3$

(i) $\{x_1, x_r\} \prec \{x_2, \ldots, x_{r-1}\}$, *and*

(ii) *there is* $\ell \notin \{1, r\}$ *such that both* $x_1, \ldots, x_\ell$ *and* $x_\ell, \ldots, x_r$ *are looping walks.*



Example looping walks:

$Q^* = v_1, v_9, v_8, v_7, v_2$ $\quad |Q^*| = 4 \quad \lambda(Q^*) = 4 - 2 \cdot 1 = 2$

$Q = v_1, v_9, v_8, v_7, v_6, v_5, v_4, v_3, v_4, v_5, v_6, v_7, v_2$ $\quad |Q| = 12$

$\{v_1, v_2\} \prec \{v_3, \ldots, v_9\}$ $\quad \lambda(Q) = 12 - 2 \cdot 6 = 0$

We decompose Q into looping walks:

$Q_1 = v_1, v_9, v_8, v_7, v_6, v_5, v_4, v_3$ $\quad \lambda(Q_1) = 7 - 2 \cdot 3 = 1$

$Q_2 = v_2, v_7, v_6, v_5, v_4, v_3$ $\quad \lambda(Q_2) = 5 - 2 \cdot 2 = 1$

$\{v_1, v_2\} \prec v_3 \prec \{v_4, \ldots, v_9\}$

Note that Q is a bad walk, while neither Q^* nor Q_1 nor Q_2 is.

FIG. 7.1. *Examples of looping walks.*

The above is a recursive definition. Note that endpoints of a looping walk are distinct and never appear in the interior of the walk. Other vertices, however, may appear on the walk multiple times as long as the walk obeys (ii). Notably, it is possible that the same vertex is one of $x_2, \ldots, x_{\ell-1}$ as well as one of $x_{\ell-1}, \ldots, x_{r-1}$ where ℓ is as defined in (ii). See Figure 7.1 for examples. Using looping walks, we define a notion of "distance" in $\mathbf{G}$ that will guide Prover in the game.

DEFINITION 7.4. *For vertices $u, v \in V(\mathbf{G})$, define $\delta(u, v)$ as follows:*

$$\delta(u,v) = \min\Big\{\lambda(Q) \;\Big|\; Q = x_1, \ldots, x_r \text{ is a looping walk of } \mathbf{G} \text{ with } x_1 = u \text{ and } x_r = v\Big\}.$$

If no looping walk between u and v exists, define $\delta(u, v) = \infty$.

In other words, $\delta(u, v)$ denotes the smallest λ-value of a looping walk between u and v. Note that $\delta(u, v) = \delta(v, u)$, since the definition of a looping walk does not prescribe the order of the endpoints of the walk.

The main structural obstruction in our characterization is the following.

DEFINITION 7.5. *A bad walk of $\mathbf{G}$ is a looping walk $Q = x_1, \ldots, x_r$ of $\mathbf{G}$ such that $x_1 \prec x_r$ and $\lambda(Q) \leq \beta(x_r) - 2$.*

We also define two values $\gamma(v)$ and $\gamma'(v)$ that will allow us to keep track of the distance from the center edge or vertex of $\mathbf{T}$, respectively.

DEFINITION 7.6. *For each vertex v we define $\gamma(v)$ recursively as follows:*

$$\gamma(v) = 0 \qquad \text{if } v \text{ is first in the ordering } \prec$$

$$\text{else} \quad \gamma(v) = \beta(v) - 1 + \max\Big\{0, \max_{u \prec v}\Big(\gamma(u) - \delta(u,v) + \beta(v) - 1\Big)\Big\}$$

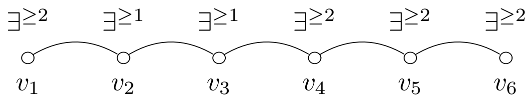
For each vertex v we define $\gamma'(v)$ recursively as follows:

$$\gamma'(v) = \beta(v) - 1 + \max\Big\{0, \max_{u \prec v}\Big(\gamma'(u) - \delta(u,v) + \beta(v) - 1\Big)\Big\}$$

It follows that $\gamma'(v) = \beta(v) - 1$ if v is first in the ordering $\prec$.

7.3. Example. Before the proofs, let us pause for a brief moment to illustrate complications that may arise in instances, especially in the case of even diameter.

Consider the 7-vertex path $\mathbf{P}_7$ with vertex set $\{1, \ldots, 7\}$ and with edges between consecutive numbers. The example instance Ψ of $\{1, 2\}$-CSP($\mathbf{P}_7$), depicted in



$$\Psi := \exists^{\geq 2} v_1 \, \exists^{\geq 1} v_2, v_3 \, \exists^{\geq 2} v_4, v_5, v_6 \; E(v_1, v_2) \wedge E(v_2, v_3) \wedge E(v_3, v_4) \\ \wedge E(v_4, v_5) \wedge E(v_5, v_6)$$

FIG. 7.2. *Example instance of* $\{1,2\}$-*CSP*($\mathbf{P}_7$) *and its graphical representation.*

Figure 7.2, is a yes-instance of this problem. To see this, let Prover offer $\{2,4\}$ for v_1, then 3 for v_2, 4 for v_3, and $\{3,5\}$ for v_4. Afterwards let Prover offer $\{2,4\}$ or $\{4,6\}$ for v_5, depending on Adversary's choice. Finally, let Prover offer $\{1,3\}$ or $\{3,5\}$ or $\{5,7\}$ for v_6, again based on Adversary's choice.

Observe that it is paramount that Prover is allowed to offer $\{3,5\}$ for the vertex v_4. If this were not the case, Adversary can make Prover lose by playing numbers closest to any end of the path. Note that Prover must honor the parity and so only even or only odd numbers can be offered for v_4. But in order for Prover to offer $\{3,5\}$ for v_4, she must start by offering $\{2,4\}$ for v_1, which could be seen as counterintuitive (one would expect that best strategy simply offers values closest to the center).

This illustrates the type of complications we shall pay attention to, and also shows one type of obstruction, the path v_4, v_5, v_6, that we will be looking to identify in instances (the other being a bad walk). The example also shows that the set offered to the first vertex in the instance is crucial for later success, and a particular attention will be paid to this. We hope that this discussion will help the reader to gain some insight into the definitions of γ and γ', before proceeding to proofs.

7.4. Characterization. Now we are ready to state the main theorem of this section.

THEOREM 7.7. *Suppose that* $\mathbf{G}$ *is connected and bipartite, with a fixed bipartition into black and white vertices. Similarly, assume that vertices of the tree* $\mathbf{T}$ *are properly colored black and white. Assume that* $\mathsf{diam}(\mathbf{T}) \geq 3$.

Then the following statements are equivalent.

- (I) $\mathbf{T} \models \Psi$
- (II) *Prover has a winning strategy in* $\mathscr{G}(\Psi, \mathbf{T})$.
- (III) *Prover can play* $\mathscr{G}(\Psi, \mathbf{T})$ *so that in every execution of the game, the resulting mapping* f *satisfies the following for all* $u, v \in V(\mathbf{G})$ *with* $\delta(u,v) < \infty$*:*
  - (IIIa) $\mathsf{dist}(f(u), f(v)) \leq \delta(u,v)$,
  - (IIIb) $f(u)$ *and* $f(v)$ *have the same color* $\iff$ $\delta(u,v)$ *is even.*
- (IV) *The following conditions hold:*
  - (IVa) *There are no* $u, v \in V(\mathbf{G})$ *where* $u \prec v$ *such that* $\delta(u,v) \leq \beta(v) - 2$.
  - (IVb) *If* $\mathsf{diam}(\mathbf{T})$ *is odd, then there is no vertex* v *such that* $\gamma(v) \geq \mathsf{rad}(\mathbf{T})$. *If* $\mathsf{diam}(\mathbf{T})$ *is even, then there are no vertices* u, v *such that* u *is black and* v *is white, and such that* $\gamma'(u) \geq \mathsf{rad}(\mathbf{T})$ *and* $\gamma'(v) \geq \mathsf{rad}(\mathbf{T})$.

7.5. Proof of Theorem 7.7. We prove the claim by considering individual implications. Note that the condition (IVa) states that $\mathbf{G}$ has no bad walk.

The equivalence (I)$\Leftrightarrow$(II) is proved as Lemma 2.1. The other implications are proved as follows. For (III)$\Rightarrow$(II), we show that Prover's strategy described in (III) is a winning strategy. For (II)$\Rightarrow$(III), we show that every winning strategy must satisfy the conditions of (III). For (III)$\Rightarrow$(IV), we show that violating (IVa) or (IVb) allows Adversary to play along a specific walk and win. Finally, for (IV)$\Rightarrow$(III), assuming

(IV), we describe a strategy for Prover that satisfies (III). Before the proofs of the individual implications, we make several useful observations as follows.

LEMMA 7.8. *Every path in* $\mathbf{G}$ *whose internal vertices appear after its endpoints in* $\prec$ *is a looping walk.*

Proof. Consider a smallest counterexample, namely a path $P = x_1, x_2, \ldots, x_r$ where $\{x_1, x_r\} \prec \{x_2, \ldots, x_{r-1}\}$ but where P is not a looping walk. Thus $r \geq 3$, as otherwise P is trivially a looping walk. So there exists an index ℓ such that $\{x_1, x_r\} \prec x_\ell \prec \{x_2, \ldots, x_{\ell-1}, x_{\ell+1}, \ldots, x_{r-1}\}$. Note that both $P' = x_1, \ldots, x_\ell$ and $P'' = x_\ell, \ldots, x_r$ are paths whose internal vertices appear after their endpoints in $\prec$. By the minimality of P, we deduce that both P' and P'' are looping walks. But then so is P, since we now fullfil both (i) and (ii) of the definition, a contradiction. □

LEMMA 7.9. *If* v *is not first in* $\prec$*, then there exists* $u \prec v$ *with* $\delta(u, v) < \infty$*.*

Proof. Since v is not first in $\prec$, there exists $w \prec v$. Since $\mathbf{G}$ is connected, consider a shortest path P from w to v in $\mathbf{G}$. Since this path starts before v and ends at v, let u be the last vertex on P such that $u \prec v$. Let Q denote the subpath of P from u to v. Note that by the maximality of u, all internal vertices of Q appear after their endpoints u, v. Thus by Lemma 7.8, we conclude that Q is a looping walk. Therefore $\delta(u, v) \leq \lambda(Q) < \infty$ as required. □

LEMMA 7.10. *If* Q *is a looping walk of* $\mathbf{G}$ *between* u *and* v*, then* $\lambda(Q) + \delta(u, v)$ *is an even number.*

Proof. Let $Q = x_1, \ldots, x_r$ be a looping walk of $\mathbf{G}$ with $x_1 = u$ and $x_r = v$. By definition $\delta(u, v) \leq \lambda(Q) < \infty$. So there exists a looping walk $Q' = x'_1, \ldots, x'_s$ with $x'_1 = u$ and $x'_s = v$ such that $\delta(u, v) = \lambda(Q')$. We calculate:

$$\lambda(Q) + \delta(u, v) = \lambda(Q) + \lambda(Q') = |Q| - 2\sum_{i=2}^{r-1}\big(\beta(x_i) - 1\big) + |Q'| - 2\sum_{i=2}^{s-1}\big(\beta(x'_i) - 1\big)$$

It follows that $\lambda(Q)+\delta(u, v)$ is even if and only if $|Q|+|Q'|$ is. Recall that $x_1 = x'_1 = u$ and $x_r = x'_s = v$. Thus $Q'' = x_1, x_2, \ldots, x_{r-1}, x'_s, x'_{s-1}, \ldots, x'_1$ is a closed walk of $\mathbf{G}$ of length $|Q| + |Q'|$. Since $\mathbf{G}$ is bipartite, it contains no closed walk of odd length. Thus $|Q| + |Q'|$ is even, and so is $\lambda(Q) + \delta(u, v)$. □

The following can be viewed as a triangle inequality for $\delta(\cdot, \cdot)$.

LEMMA 7.11. *If* $\{u, w\} \prec v$*, then* $\delta(u, w) \leq \delta(u, v)+\delta(w, v)-2\beta(v)+2$*. Moreover,* $\delta(u, v) + \delta(w, v) + \delta(u, w)$ *is an even number or* ∞*.*

Proof. If $u = w$, the result is trivial since there are no bad walks, so we assume $u \neq w$. If $\delta(u, v) = \infty$ or $\delta(w, v) = \infty$, the claim is clearly true. So we may assume that $\delta(u, v) < \infty$ and $\delta(w, v) < \infty$. This means that there exists a looping walk $Q = x_1, \ldots, x_r$ with $x_1 = u$ and $x_r = v$ where $\lambda(Q) = \delta(u, v)$, and also a looping walk $Q' = x'_1, \ldots, x'_s$ with $x'_1 = w$ and $x'_s = v$ where $\delta(w, v) = \lambda(Q')$. Note that $x_r = x'_s = v$, and define $Q'' = x_1, x_2, \ldots, x_{r-1}, x'_s, x'_{s-1}, \ldots, x'_1$. We calculate:

$$\begin{aligned}
\lambda(Q) + \lambda(Q') &= |Q| - 2\sum_{i=2}^{r-1}\big(\beta(x_i) - 1\big) + |Q'| - 2\sum_{i=2}^{s-1}\big(\beta(x'_i) - 1\big) \\
&= |Q| + |Q'| - 2\sum_{i=2}^{r-1}\big(\beta(x_i) - 1\big) - 2(\beta(v) - 1) - 2\sum_{i=2}^{s-1}\big(\beta(x'_i) - 1\big) + 2(\beta(v) - 1) \\
&= \lambda(Q'') + 2\beta(v) - 2
\end{aligned}$$

Observe that Q'' is a walk of $\mathbf{G}$ whose endpoints are u and w. We verify that Q'' is in fact a looping walk of $\mathbf{G}$. We need to check the conditions (i)-(ii) of the definition. First, we recall the assumption $u \neq w$. For (i), we note that $\{x_1, x_r\} \prec \{x_2, \ldots, x_{r-1}\}$

since Q is a looping walk. Similarly, $\{x'_1, x'_s\} \prec \{x'_2, \ldots, x'_{s-1}\}$ since Q' is a looping walk. We also recall the assumptions $\{u, w\} \prec v$. Thus, since $v = x_r = x'_s$, we conclude that $\{u, w\} \prec v \prec \{x_2, \ldots, x_{r-1}, x'_2, \ldots, x'_{s-1}\}$ which shows (i). Finally, (ii) follows from the fact that both Q and Q' are looping walks. This verifies that Q'' indeed is a looping walk of $\mathbf{G}$ between u and v.

So we have $\delta(u, w) \leq \lambda(Q'')$ by definition, and we can calculate:

$$\delta(u, v) + \delta(w, v) = \lambda(Q) + \lambda(Q') = \lambda(Q'') + 2\beta(v) - 2 \geq \delta(u, w) + 2\beta(v) - 2$$

Thus $\delta(u, w) \leq \delta(u, v) + \delta(w, v) - 2\beta(v) + 2$ as claimed.

Now, recall that $\lambda(Q) + \lambda(Q') = \lambda(Q'') + 2\beta(v) - 2$. This implies that $\lambda(Q) + \lambda(Q') + \lambda(Q'')$ is an even number. By Lemma 7.10, also $\lambda(Q) + \delta(u, v)$ and $\lambda(Q') + \delta(w, v)$ are even. For the same reason $\lambda(Q'') + \delta(u, w)$ is even. So since $\lambda(Q) + \lambda(Q') + \lambda(Q'')$ is even, it follows that $\delta(u, w) + \delta(u, v) + \delta(w, v)$ is even, as claimed. ☐

7.5.1. Proof of (III)⇒(II). Assume (III), namely that Prover can play so as to satisfy (IIIa) and (IIIb) in every execution of the game. We show that this is a winning strategy for Prover, thus proving (II).

To this end, we need to verify that in every execution of the game the resulting mapping f is a homomorphism of $\mathbf{G}$ to $\mathbf{T}$. Namely, we verify that for every edge $uv \in E(\mathbf{G})$, the mapping f satisfies $(f(u), f(v)) \in E(\mathbf{T})$.

Consider an edge $uv \in E(\mathbf{G})$. Observe that $Q = u, v$ is a looping walk of $\mathbf{G}$ with $\lambda(Q) = 1$. Thus $\delta(u, v) \leq \lambda(Q) = 1$ by the definition of $\delta(u, v)$. Using Lemma 7.10, we deduce that $\lambda(Q) + \delta(u, v) = 1 + \delta(u, v)$ is even, i.e., $\delta(u, v)$ is odd. Thus, by (IIIb), we conclude that $f(u)$ and $f(v)$ are from different color classes of T. In particular, $f(u) \neq f(v)$. Further, by (IIIa), we observe that $\mathsf{dist}(f(u), f(v)) \leq \delta(u, v)$. Thus $\mathsf{dist}(f(u), f(v)) \leq 1$, since $\delta(u, v) \leq \lambda(Q) = 1$. In fact, $\mathsf{dist}(f(u), f(v)) = 1$, since $f(u) \neq f(v)$. Thus together we conclude $(f(u), f(v)) \in E(T)$ as required.

7.5.2. Proof of (II)⇒(III). Assume (II), i.e., Prover has a winning strategy. Assume that Prover plays this strategy. Then no matter how Adversary plays, Prover always wins. We show that this strategy satisfies the conditions of (III).

For contradiction, suppose that there is an execution of the game for which the conditions of (III) do not hold. Let g denote the resulting mapping produced by this execution. Let us play the game again to produce a new mapping f by making Adversary play according to the following rules:

1. *When a vertex v is considered, examine the set $S_v \subseteq V(\mathbf{T})$ that Prover offers for v. Check if there exists $\mu \in S_v$ and $u \prec v$ with $\delta(u, v) < \infty$ such that at least one of the following holds:*
   - $\mathsf{dist}(f(u), \mu) > \delta(u, v)$, *or*
   - $\delta(u, v)$ *is odd and* $f(u), \mu$ *have the same color, or*
   - $\delta(u, v)$ *is even and* $f(u), \mu$ *have different colors.*
2. *If such a μ exists, choose $f(v) = \mu$. If such a μ does not exist and $f(w) = g(w)$ for all $w \prec v$, then choose $f(v) = g(v)$. If neither is possible, then choose any value from S_v for $f(v)$.*

We now argue that f does not satisfy the conditions of (III) much like g. Indeed, if for some v, we find in step 1 that μ exists, then setting $f(v) = \mu$ in step 2 makes (IIIa) or (IIIb) fail for the pair u, v. Thus f fails the conditions of (III). Otherwise, if for every v the value μ in step 1 is not found, then we conclude that $f = g$, and hence, f again fails the conditions of (III), since g does. (For this to hold, it is important to note that this is only possible because Prover plays the same deterministic strategy in both executions of the game; thus Prover will offer the same values for $f(v)$ as she did for $g(v)$ as long as Adversary makes the same choices for f as he did for g; i.e., as long as $f(w) = g(w)$ for all vertices $w \prec v$.)

Since Prover plays a winning strategy, we must conclude that f is a homomorphism of $\mathbf{G}$ to $\mathbf{T}$. Namely, we have that each $uv \in E(\mathbf{G})$ satisfies $\mathsf{dist}(f(u), f(v)) = 1$. We show that this leads to a contradiction. From this we will conclude that g does not exist, and hence, Prover's winning strategy satisfies (III) as we advertized earlier.

We say that a vertex v is *good* if for all $u \prec v$ with $\delta(u, v) < \infty$ the conditions (IIIa) and (IIIb) hold. Otherwise, we say that v is *bad*. The following is a restatement of Adversary's strategy as described above.

(+) *A vertex v is good iff every $\mu \in S_v$ and every $u \prec v$ with $\delta(u, v) < \infty$ satisfy*
- $\mathsf{dist}(f(v), \mu) \leq \delta(u, v)$, *and*
- $f(v)$ *and* μ *have the same color* $\iff$ $\delta(u, v)$ *is even.*

To see this, observe that if v is bad, then one of the above two conditions fails for $\mu = f(v) \in S_v$, since one of (IIIa), (IIIb) fails. Conversely, suppose that for some $\mu \in S_v$ there is $u \prec v$ with $\delta(u, v) < \infty$ such that $\mathsf{dist}(f(v), \mu) > \delta(u, v)$ or such that either $\delta(u, v)$ is odd and $f(v), \mu$ have the same color, or $\delta(u, v)$ is even and $f(v), \mu$ have different colors. Then, by our strategy in steps 1 and 2, we deduce that Adversary chose $f(v) = \mu$ in step 2. Thus v is not a good vertex. This proves (+).

Since f fails the conditions of (III), there exists at least one bad vertex. Among all bad vertices, choose v to be the bad vertex that is largest with respect to $\prec$.

Since v is bad, there exists $u \prec v$ with $\delta(u, v) < \infty$ such that

(i) $\mathsf{dist}(f(u), f(v)) > \delta(u, v)$, or

(ii) $f(u), f(v)$ have the same color $\iff$ $\delta(u, v)$ is odd.

In particular, since $\delta(u, v) < \infty$, there exists a looping walk $Q = x_1, \ldots, x_r$ with $x_1 = u$ and $x_r = v$ such that $\lambda(Q) = \delta(u, v)$.

Suppose first that $r = 2$. Then $Q = u, v$ and $\lambda(Q) = 1$. In particular, $uv \in E(\mathbf{G})$ and $\delta(u, v) = \lambda(Q) = 1$. Recall that we assume that one of (i), (ii) holds. If $\mathsf{dist}(f(u), f(v)) > \delta(u, v)$, then $\mathsf{dist}(f(u), f(v)) > 1$, since $\delta(u, v) = 1$. But then, since $uv \in E(\mathbf{G})$, we conclude that f is not a homomorphism, a contradiction. So by (i) and (ii), we deduce that $f(u)$, $f(v)$ have the same color, since $\delta(u, v) = 1$ is odd. But then $\mathsf{dist}(f(u), f(v)) \neq 1$ again contradicting our assumption that f is a homomorphism.

This excludes the case $r = 2$. Thus we may assume $r \geq 3$. Since Q is a looping walk, this implies that there exists $\ell \in \{2, \ldots, r-1\}$ such that both $Q_1 = x_1, \ldots, x_\ell$ and $Q_2 = x_\ell, \ldots, x_r$ are looping walks of G.

Let us denote $w = x_\ell$. So Q_1 is a looping walk from u to w, while Q_2 is a looping walk from w to v. This implies that $\delta(u, w) \leq \lambda(Q_1)$ and $\delta(v, w) \leq \lambda(Q_2)$. Note that $u \prec v \prec w$, since $\{u, v\} \prec \{x_2, \ldots, x_{r-1}\}$ because Q is a looping walk. Thus by the maximality of v, we deduce that w is a good vertex. We calculate:

$$\begin{aligned}
\lambda(Q) &= |Q| - 2\sum_{i=2}^{r-1} \big(\beta(x_i) - 1\big) \\
&= |Q_1| + |Q_2| - 2\sum_{i=2}^{\ell-1} \big(\beta(x_i) - 1\big) - 2(\beta(w) - 1) - 2\sum_{i=\ell+1}^{r-1} \big(\beta(x_i) - 1\big) \\
&= \lambda(Q_1) + \lambda(Q_2) - 2\beta(w) + 2 \\
&\geq \delta(u, w) + \delta(v, w) - 2\beta(w) + 2 \geq \delta(u, v) = \lambda(Q)
\end{aligned}$$

The last inequality is by Lemma 7.11. Therefore, $\delta(u, v) = \delta(u, w) + \delta(v, w) - 2\beta(w) + 2$.

Now, since w is a good vertex, we have by (IIIb) that both $f(u), f(w)$ have the same color iff $\delta(u, w)$ is even, and also $f(v), f(w)$ have the same color iff $\delta(v, w)$ is even. If $\delta(u, v)$ is even, then both $\delta(u, w)$ and $\delta(v, w)$ have the same parity, since $\delta(u, v) = \delta(u, w) + \delta(v, w) - 2\beta(w) + 2$. It follows that $f(u)$ and $f(v)$ have the same color. If $\delta(u, v)$ is odd, then one of $\delta(u, w)$, $\delta(v, w)$ is even, and one is odd. Thus either $f(u), f(w)$ have the same color different from $f(v)$, or $f(v), f(w)$ have the same color different from $f(u)$. In either case, $f(u)$, $f(v)$ have different colors.

This shows that (ii) fails. Thus since we assume that at least one of (i), (ii) holds, we deduce that $\mathsf{dist}(f(u), f(v)) > \delta(u, v)$. In fact, since $f(u), f(v)$ have the same color iff $\delta(u, v)$ is even, it follows that $\mathsf{dist}(f(u), f(v)) \geq \delta(u, v)+2$. For this, observe that no two vertices of the same color in $\mathbf{T}$ are at an odd distance, since the coloring is proper.

Now, recall again that w is a good vertex. By (+), every $\mu \in S_w$ satisfies $\mathsf{dist}(f(u), \mu) \leq \delta(u, w)$ and $\mathsf{dist}(f(v), \mu) \leq \delta(v, w)$. Using this we can calculate:

$$\begin{aligned} \mathsf{dist}(f(u), f(v)) &\geq \delta(u, v) + 2 = \delta(u, w) + \delta(v, w) - 2\beta(w) + 2 + 2 \\ &\geq \mathsf{dist}(f(u), \mu) + \mathsf{dist}(f(v), \mu) - 2\beta(w) + 4 \\ &\geq \mathsf{dist}(f(u), f(v)) - 2\beta(w) + 4 \end{aligned}$$

For the last inequality, note that dist satisfies the triangle-inequality.

This implies that $\beta(w) = 2$ and all the above inequalities are in fact equalities. Namely, the set S_w contains distinct values μ_1, μ_2 such that for $i = 1, 2$:

$$\mathsf{dist}(f(u), f(v)) = \mathsf{dist}(f(u), \mu_i) + \mathsf{dist}(f(v), \mu_i)$$

This means that concatenating a shortest path from $f(u)$ to μ_i with a shortest path from μ_i to $f(v)$ yields a shortest path P_i from $f(u)$ to $f(v)$. Since $\mu_1 \neq \mu_2$, the paths P_1, P_2 are different. However, this is impossible, since in $\mathbf{T}$, because it is a tree, there is a unique shortest path between any two vertices, a contradiction.

The proof is now complete.

7.5.3. Proof of (III)⇒(IV). Assume that (III) holds. Namely, Prover can play so as to always satisfy the conditions (IIIa) and (IIIb) for all u, v with $\delta(u, v) < \infty$.

Assume first that (IVa) fails. Namely, suppose that there are $u, v \in V(\mathbf{G})$ with $u \prec v$ such that $\delta(u, v) \leq \beta(v) - 2$. We show that Adversary can play to violate (IIIa) for u, v. If $\delta(u, v) \leq -1$, then (IIIa) can never be satisfied, since $\mathsf{dist}(f(u), f(v))$ is always non-negative. In that case (III) fails no matter how Adversary plays. So we may assume that $\delta(u, v) \geq 0$. This implies that $\delta(u, v) = 0$ and $\beta(v) = 2$, since we assume $\delta(u, v) \leq \beta(v) - 2$. So Prover must offer to Adversary two distinct values for v. Since the values are different, at least one of them must be different from $f(u)$. Adversary chooses this value for $f(v)$. This yields $\mathsf{dist}(f(u), f(v)) \geq 1$ which violates (IIIa), since $\delta(u, v) = 0$. This shows that Prover cannot play to always satisfy the conditions of (III), and hence (III) fails, a contradiction.

Next, assume that (IVb) fails. We distinguish the case when $\mathsf{diam}(\mathbf{T})$ is even and when $\mathsf{diam}(\mathbf{T})$ is odd. A *center* of $\mathbf{T}$ is a vertex α with $\mathsf{ecc}(\alpha) = \mathsf{rad}(\mathbf{T})$.

We first observe the following useful property.

LEMMA 7.12. *If μ_1, μ_2 are two distinct centers of $\mathbf{T}$, then they are adjacent.*

Proof. For contradiction, suppose that $\mathsf{ecc}(\mu_1) = \mathsf{ecc}(\mu_2) = \mathsf{rad}(\mathbf{T})$, but μ_1 and μ_2 are not adjacent. Consider any internal vertex ρ on a shortest path from μ_1 to μ_2 in $\mathbf{T}$. Let η be the vertex that maximizes $\mathsf{ecc}(\rho)$, i.e., $\mathsf{ecc}(\rho) = \mathsf{dist}(\rho, \eta)$. If for some $i \in \{1, 2\}$, we have $\mathsf{dist}(\mu_i, \eta) > \mathsf{dist}(\rho, \eta)$, then we obtain a contradiction as follows.

$$\mathsf{rad}(\mathbf{T}) \leq \mathsf{ecc}(\rho) = \mathsf{dist}(\rho, \eta) < \mathsf{dist}(\mu_i, \eta) \leq \mathsf{ecc}(\mu_i) = \mathsf{rad}(\mathbf{T})$$

Now consider three shortest paths, path P_1 between μ_1 and η, path P_2 between η and μ_2, and path P between μ_1 and μ_2. Observe that for each $i \in \{1,2\}$, the path P_i does not contain ρ, since $\rho \neq \mu_i$ and $\mathsf{dist}(\mu_i,\eta) \leq \mathsf{dist}(\rho,\eta)$. This implies that the union of paths P_1 and P_2 contains a path P' from μ_1 to μ_2, and this path does not contain ρ. Recall that P is another path from μ_1 to μ_2, but P contains ρ. This is impossible, since in $\mathbf{T}$ there is a unique path between any two vertices. ◻

The case of odd diameter follows from the next lemma.

Lemma 7.13. *Adversary can play so that the resulting mapping f satisfies* $\mathsf{ecc}(f(v)) - \mathsf{rad}(\mathbf{T}) \geq \gamma(v)$ *for every vertex v.*

Proof. Recall that we assume that Prover plays so as to satisfy (IIIa) and (IIIb). Adversary when given choice between two values will choose the value of higher eccentricity (ties broken arbitrarily). Note that all $\alpha \in V(\mathbf{T})$ satisfy $\mathsf{ecc}(\alpha) \geq \mathsf{rad}(\mathbf{T})$.

We prove the claim by induction on the number of steps. Consider some step in the game, and let v denote the vertex considered in this step. If v is the first vertex in $\prec$, then $\gamma(v) = 0$ and $\mathsf{ecc}(f(v)) - \mathsf{rad}(\mathbf{T}) \geq 0 \geq \gamma(v)$ holds. So v is not first in $\prec$.

Next, suppose that $\gamma(v) = \beta(v) - 1$. If $\beta(v) = 1$, then $\gamma(v) = \beta(v) - 1 = 0$ and so again $\mathsf{ecc}(f(v)) - \mathsf{rad}(\mathbf{T}) \geq 0 = \gamma(v)$. Thus $\beta(v) = 2$ in which case Prover offers two distinct values μ_1, μ_2. Recall that v is not first in $\prec$, and $\mathbf{G}$ is connected. Since the values that Prover offers must satisfy (IIIb) when used as $f(v)$, we deduce that μ_1, μ_2 must have the same color. Indeed, since $\mathbf{G}$ is connected and v is not first in $\prec$, we conclude by Lemma 7.9 that there exists a looping walk from some u to v where $u \prec v$. Then (IIIb) applied to u and v implies that the values μ_1, μ_2 have the same color, either different from or same as that of $f(u)$. In particular, μ_1, μ_2 are not adjacent. Thus by Lemma 7.12, at least one of μ_1, μ_2 is not a center of $\mathbf{T}$. Let μ_i be this vertex. Then $\mathsf{ecc}(\mu_i) - \mathsf{rad}(\mathbf{T}) \geq 1$ and Adversary chooses $f(v) = \mu_i$. We conclude $\mathsf{ecc}(f(v)) - \mathsf{rad}(\mathbf{T}) \geq 1 = \beta(v) - 1 = \gamma(v)$ as required.

It therefore remains to consider the case when there is a vertex $u \prec v$ that maximizes $\gamma(v)$, namely $\gamma(v) = 2(\beta(v) - 1) + \gamma(u) - \delta(u,v)$. From the inductive hypothesis, we know that $\mathsf{ecc}(f(u)) - \mathsf{rad}(\mathbf{T}) \geq \gamma(u)$. From this we conclude

$$(+_1)\quad \mathsf{ecc}(f(u)) - \mathsf{rad}(\mathbf{T}) \geq \gamma(u) = \gamma(v) + \delta(u,v) - 2(\beta(v) - 1)$$

Recall that Prover offers values that satisfy (IIIa) when chosen for $f(v)$. There are two cases to examine.

Case 1: suppose that $\beta(v) = 1$. Then Prover offers one value that becomes $f(v)$ where $\mathsf{dist}(f(u), f(v)) \leq \delta(u,v)$ by (IIIa). By $(+_1)$, we deduce that $\mathsf{ecc}(f(u)) - \delta(u,v) \geq \gamma(v) + \mathsf{rad}(\mathbf{T})$, since $\beta(v) = 1$. Let η be the vertex that maximizes $\mathsf{ecc}(f(u))$, i.e., $\mathsf{ecc}(f(u)) = \mathsf{dist}(f(u), \eta)$. Then we can calculate

$$\mathsf{ecc}(f(u)) = \mathsf{dist}(f(u),\eta) \leq \mathsf{dist}(f(v),\eta) + \mathsf{dist}(f(u),f(v)) \leq \mathsf{ecc}(f(v)) + \delta(u,v)$$

From this, we conclude that $\mathsf{ecc}(f(v)) - \mathsf{rad}(\mathbf{T}) \geq \gamma(v)$ as follows.

$$\mathsf{ecc}(f(v)) \geq \mathsf{ecc}(f(u)) - \delta(u,v) \geq \gamma(v) + \mathsf{rad}(\mathbf{T})$$

Case 2: suppose that $\beta(v) = 2$. Then Prover offers two values μ_1, μ_2 where $\mu_1 \neq \mu_2$, and both satisfy (IIIa) and (IIIb) in place of $f(v)$. Namely, for each $i \in \{1,2\}$, we have $\mathsf{dist}(f(u), \mu_i) \leq \delta(u,v)$. From (IIIb) we know that μ_1, μ_2 have the same color.

We now prove the following claim.

$(+_2)$ *There exists $i \in \{1,2\}$ such that* $\mathsf{ecc}(f(u)) \leq \delta(u,v) + \mathsf{ecc}(\mu_i) - 2$

Let η be a vertex that maximizes $\mathsf{ecc}(f(u))$, i.e., $\mathsf{ecc}(f(u)) = \mathsf{dist}(f(u),\eta)$. Recall that $\mathsf{dist}(f(u), \mu_i) \leq \delta(u,v)$ for all $i \in \{1,2\}$.

Let P be a shortest path from $f(u)$ to η. Suppose first that μ_i does not appear on P for some $i \in \{1,2\}$. Let P_i be a shortest path from $f(u)$ to μ_i, and let P_i' be a shortest path from μ_i to η. Concatenating P_i and P_i' gives us a walk Q from $f(u)$ to η. If this walk has the same length as P, then $Q = P$ because in $\mathbf{T}$ there is a unique path between any two vertices. But since μ_i is on Q but not on P, this is impossible. So the walk Q is longer than P. In fact, Q has at least two more edges that P, since otherwise the union of P and Q yields a closed walk of odd length, and $\mathbf{T}$ contains no such a walk, since $\mathbf{T}$ is bipartite. Thus we conclude the following

$$\mathsf{ecc}(f(u)) = \mathsf{dist}(f(u),\eta) \le \mathsf{dist}(f(u),\mu_i) + \mathsf{dist}(\mu_i,\eta) - 2 \le \delta(u,v) + \mathsf{ecc}(\mu_i) - 2$$

This yields the required $i \in \{1,2\}$ for the claim. Thus we may assume that both μ_1 and μ_2 appear on P. Recall that $\mathsf{dist}(f(u),\mu_i) \le \delta(u,v)$ for all $i \in \{1,2\}$. Since both μ_1 and μ_2 belong to P, the subpath of P from $f(u)$ to μ_i is a shortest path from $f(u)$ to μ_i. Since $\mu_1 \neq \mu_2$, we may assume by symmetry μ_1 appears before μ_2 on P. Moreover, since μ_1, μ_2 have the same color, they are not consecutive on P. Thus we deduce that $\mathsf{dist}(f(u),\mu_1) + 2 \le \mathsf{dist}(f(u),\mu_2) \le \delta(u,v)$. From this we calculate

$$\mathsf{ecc}(f(u)) = \mathsf{dist}(f(u),\eta) = \mathsf{dist}(f(u),\mu_1) + \mathsf{dist}(\mu_1,\eta) \le \delta(u,v) + \mathsf{ecc}(\mu_1) - 2$$

This proves $(+_2)$.

In view of $(+_2)$, there exists $i \in \{1,2\}$ such that $\mathsf{ecc}(f(u)) \le \delta(u,v) + \mathsf{ecc}(\mu_i) - 2$. Adversary chooses $f(v) = \mu_i$. Note that by $(+_1)$, we have $\mathsf{ecc}(f(u)) - \delta(u,v) \ge \gamma(v) + \mathsf{rad}(\mathbf{T}) - 2$, since $\beta(v) = 2$. Thus we calculate

$$\mathsf{ecc}(f(v)) = \mathsf{ecc}(\mu_i) \ge \mathsf{ecc}(f(u)) - \delta(u,v) + 2 \ge \gamma(v) + \mathsf{rad}(\mathbf{T}) - 2 + 2 = \gamma(v) + \mathsf{rad}(\mathbf{T})$$

Therefore $\mathsf{ecc}(f(v)) - \mathsf{rad}(\mathbf{T}) \ge \gamma(v)$ as required. This completes the proof. ◻

The case of even diameter similarly follows from this lemma.

Lemma 7.14. *Suppose that* $\mathsf{diam}(\mathbf{T})$ *is even. Then Adversary can play so that the resulting mapping* f *satisfies* $\mathsf{ecc}(f(v)) - \mathsf{rad}(\mathbf{T}) \ge \gamma'(v)$ *for every vertex* v.

Proof. The proof is by induction on the number of steps. For the base case, we consider v such that $\gamma'(v) = \beta(v) - 1$. If $\beta(v) = 1$, then $\gamma'(v) = 0$ and $\mathsf{ecc}(f(v)) - \mathsf{rad}(\mathbf{T}) \ge 0 = \gamma'(v)$, as required. If $\beta(v) = 2$, Prover offers two distinct values for v, and at least one of them is not a center of $\mathbf{T}$, since $\mathsf{diam}(\mathbf{T})$ is even. Adversary chooses this value as $f(v)$. Then $\mathsf{ecc}(f(v)) - \mathsf{rad}(T) \ge 1 = \gamma'(v)$, as required.

The rest of the proof (the inductive case) is exactly as in the proof of Lemma 7.13 (with γ replaced by γ'). This completes the proof. ◻

Now we are ready to finish the proof of (III)⇒(IV). Recall that we assume that (IVb) fails. Suppose first that $\mathsf{diam}(\mathbf{T})$ is odd. Then by (IVb), there exists v such that $\gamma(v) \ge \mathsf{rad}(\mathbf{T})$. By Lemma 7.13, Adversary can play so that the resulting mapping f satisfies $\mathsf{ecc}(f(v)) - \mathsf{rad}(\mathbf{T}) \ge \gamma(v)$. Note that since $\mathbf{T}$ is a tree of odd diameter, we have $\mathsf{diam}(\mathbf{T}) = 2\mathsf{rad}(\mathbf{T}) - 1$. From this we obtain a contradiction as follows.

$$\mathsf{diam}(\mathbf{T}) \ge \mathsf{ecc}(f(v)) \ge \gamma(v) + \mathsf{rad}(\mathbf{T}) \ge 2\mathsf{rad}(\mathbf{T}) > \mathsf{diam}(\mathbf{T}).$$

Now assume that $\mathsf{diam}(\mathbf{T})$ is even. By (IVb), there exist vertices u, v of different colors such that $\gamma'(u) \ge \mathsf{rad}(\mathbf{T})$ and $\gamma'(v) \ge \mathsf{rad}(\mathbf{T})$. By Lemma 7.14, Adversary can play so that the resulting mapping f satisfies $\mathsf{ecc}(f(u)) - \mathsf{rad}(\mathbf{T}) \ge \gamma'(u)$ and $\mathsf{ecc}(f(v)) - \mathsf{rad}(\mathbf{T}) \ge \gamma'(v)$. Note that $\mathsf{diam}(\mathbf{T}) = 2\mathsf{rad}(\mathbf{T})$, since $\mathbf{T}$ is a tree of even diameter. Thus we conclude that

$$\mathsf{diam}(\mathbf{T}) \ge \mathsf{ecc}(f(u)) \ge \gamma'(u) + \mathsf{rad}(\mathbf{T}) \ge 2\mathsf{rad}(\mathbf{T}) = \mathsf{diam}(\mathbf{T})$$

and similarly $\mathsf{ecc}(f(v)) = \mathsf{diam}(\mathbf{T})$. Let μ be a vertex that maximizes $\mathsf{ecc}(f(u))$, i.e., $\mathsf{ecc}(f(u)) = \mathsf{dist}(f(u),\mu)$. Similarly, let η be a vertex such that $\mathsf{ecc}(f(v)) =$

$\mathsf{dist}(f(v),\eta)$. From implication (III)⇒(II) we deduce that f is a homomorphism of $\mathbf{G}$ to $\mathbf{T}$, since we assume (III) (and we have already proved (III)⇒(II)). Thus, since both $\mathbf{G}$ and $\mathbf{T}$ are bipartite and since u,v have different colors, it follows that $f(u)$ and $f(v)$ also have different colors. By symmetry, we may assume that $f(u)$ is black and $f(v)$ is white. Observe that μ is also black, since the shortest path between $f(u)$ and μ has even length $\mathsf{diam}(\mathbf{T})$. By the same token, η is white. Since $\mathbf{T}$ is connected, this implies that the vertices $f(u), f(v), \mu, \eta$ are pairwise distinct. Moreover, $\mathsf{dist}(f(u),\eta) < \mathsf{diam}(\mathbf{T})$ since $f(u),\eta$ have different colors and $\mathsf{diam}(\mathbf{T})$ is even. Likewise, each of $\mathsf{dist}(f(v),\mu)$, $\mathsf{dist}(f(u),f(v))$, $\mathsf{dist}(\mu,\eta)$ is less than $\mathsf{diam}(\mathbf{T})$.

By Lemma 7.1, we must have

$$\mathsf{dist}(f(u),\mu)+\mathsf{dist}(f(v),\eta)\leq \max\left\{\begin{array}{l}\mathsf{dist}(f(u),\eta)+\mathsf{dist}(f(v),\mu)\\ \mathsf{dist}(f(u),f(v))+\mathsf{dist}(\mu,\eta)\end{array}\right\}$$

However, the expression on the left is $2\mathsf{diam}(\mathbf{T})$, while both expressions on the right are at most $2\mathsf{diam}(\mathbf{T})-2$, a contradiction.

Therefore, such vertices u,v cannot exist which proves (III)⇒(IV).

7.5.4. Proof of (IV)⇒(III). Assume (IV), namely that both (IVa) and (IVb) hold. We explain how Prover can play so as to satisfy the conditions of (III). In particular, we explain how Prover can play exclusively on a longest path of $\mathbf{T}$ and satisfy (IIIa) and (IIIb). In the following argument, we consider the two cases of odd and even diameter together. In case $\mathsf{diam}(\mathbf{T})$ is even, we make the following additional assumption. By (IVb), we have that either no black vertex v has $\gamma'(v)\geq \mathsf{rad}(\mathbf{T})$, or no white vertex v has $\gamma'(v)\geq\mathsf{rad}(\mathbf{T})$. By symmetry, we shall assume:

($+_0$) *If $\mathsf{diam}(\mathbf{T})$ is even, then every white vertex v is such that $\gamma'(v)\leq\mathsf{rad}(\mathbf{T})-1$.*

Now, let P be a fixed longest path of $\mathbf{T}$, i.e., a path of length $\mathsf{diam}(\mathbf{T})$. Let α_1,α_2 be the endpoints of P. Define $\mathsf{ecc}_P(x)$ as $\max\{\mathsf{dist}(x,\alpha_1),\mathsf{dist}(x,\alpha_2)\}$. Clearly $\mathsf{ecc}(x)\geq\mathsf{ecc}_P(x)$. For $x\in P$, we observe the following.

($+_1$) $\mathsf{ecc}(x)=\mathsf{ecc}_P(x)$ *for all $x\in P$.*

To see this, note that for each $i\in\{1,2\}$, the subpath of P between x and α_i is a shortest path, since there is a unique path between any two vertices in $\mathbf{T}$. Therefore $\mathsf{dist}(x,\alpha_1)+\mathsf{dist}(x,\alpha_2)=\mathsf{diam}(\mathbf{T})$, since P has length $\mathsf{diam}(\mathbf{T})$.

For contradiction, assume $\mathsf{ecc}(x)>\mathsf{ecc}_P(x)$. Let η be a vertex maximizing $\mathsf{ecc}(x)$, i.e., $\mathsf{ecc}(x)=\mathsf{dist}(x,\eta)$. Then $\mathsf{dist}(x,\alpha_1)<\mathsf{dist}(x,\eta)$ and $\mathsf{dist}(x,\alpha_2)<\mathsf{dist}(x,\eta)$, since $\mathsf{ecc}_P(x)<\mathsf{ecc}(x)$. Note that $\mathsf{dist}(x,\alpha_1)+\mathsf{dist}(\alpha_2,\eta)<\mathsf{dist}(x,\eta)+\mathsf{dist}(\alpha_1,\alpha_2)$ since $\mathsf{dist}(x,\alpha_1)<\mathsf{dist}(x,\eta)$ and $\mathsf{dist}(\alpha_2,\eta)\leq\mathsf{diam}(\mathbf{T})=\mathsf{dist}(\alpha_1,\alpha_2)$. Similarly, $\mathsf{dist}(x,\alpha_2)+\mathsf{dist}(\alpha_1,\eta)<\mathsf{dist}(x,\eta)+\mathsf{dist}(\alpha_1,\alpha_2)$. Thus $\mathsf{dist}(x,\eta)+\mathsf{dist}(\alpha_1,\alpha_2)>\max\{\mathsf{dist}(x,\alpha_1)+\mathsf{dist}(\alpha_2,\eta),\mathsf{dist}(x,\alpha_2)+\mathsf{dist}(\alpha_1,\eta)\}$ contradicting Lemma 7.1.

This proves ($+_1$).

The next lemma explains the desired strategy for Prover.

LEMMA 7.15. *Prover can play so that in every execution of the game the resulting mapping f satisfies (IIIa), (IIIb) and for every vertex v, we have $f(v)\in P$, and*

(i) *if $\mathsf{diam}(\mathbf{T})$ is odd, then $\mathsf{ecc}(f(v))-\mathsf{rad}(\mathbf{T})\leq\gamma(v)$,*

(ii) *if $\mathsf{diam}(\mathbf{T})$ is even, then $\mathsf{ecc}(f(v))-\mathsf{rad}(\mathbf{T})\leq\gamma'(v)+1$,*
and $\mathsf{ecc}(f(v))$ is odd $\iff$ v is white.

Proof. In the desired strategy, Prover offers vertices of smallest eccentricity. We construct Prover's strategy by induction on the number of processed vertices during the game. As usual let f denote the (partial) assignment constructed by Adversary.

If $\mathsf{diam}(\mathbf{T})$ is odd, the path P contains an edge m_1m_2 where m_1 and m_2 are centers of $\mathbf{T}$, namely $\mathsf{ecc}(m_1) = \mathsf{ecc}(m_2) = \mathsf{rad}(\mathbf{T})$. If $\mathsf{diam}(\mathbf{T})$ is even, P contain a unique center m. Observe that no two consecutive vertices on P have the same eccentricity except for m_1, m_2 if $\mathsf{diam}(\mathbf{T})$ is odd.

For the base case of the induction, consider v to be the first vertex in $\prec$. Then $\gamma(v) = 0$ and $\gamma'(v) = \beta(v) - 1$. If $\mathsf{diam}(\mathbf{T})$ is odd, Prover offers for v the vertex m_1 if $\beta(v) = 1$, and offers the set $\{m_1, m_2\}$ if $\beta(v) = 2$. Adversary chooses $f(v) = m_i$ for some $i \in \{1, 2\}$ and we have $\mathsf{ecc}(f(v)) - \mathsf{rad}(\mathbf{T}) = \mathsf{ecc}(m_i) - \mathsf{rad}(\mathbf{T}) = 0 = \gamma(v)$ as required. This shows (i) for v. Conditions (IIIa) and (IIIb) are satisfied vacuously.

Similarly, if $\mathsf{diam}(\mathbf{T})$ is even. Recall that $\gamma'(v) = \beta(v) - 1$. Let μ_1, μ_2 be the two neighbors of m on P. Let μ be a distance-2 vertex from m on P. Note that μ exists, since $\mathsf{diam}(\mathbf{T}) \geq 3$. For later, note that $\mathsf{ecc}(\mu) = \mathsf{rad}(\mathbf{T}) - 2$ and $\mathsf{ecc}(\mu_i) = \mathsf{rad}(\mathbf{T}) - 1$ for $i = 1, 2$, since $\mathsf{dist}(\mu_i, m) = 1$, $\mathsf{dist}(\mu, m) = 2$, and $\mathsf{ecc}(m) = \mathsf{rad}(\mathbf{T})$.

If v is white and $\mathsf{rad}(\mathbf{T})$ is odd, or if v is black and $\mathsf{rad}(\mathbf{T})$ is even, Prover offers $f(v) = m$ if $\beta(v) = 1$, and offers $f(v) \in \{m, \mu\}$ if $\beta(v) = 2$. In all other cases, Prover offers $f(v) = \mu_1$ if $\beta(v) = 1$, and offers $f(v) \in \{\mu_1, \mu_2\}$ if $\beta(v) = 2$. Irrespective of the choice of Adversary, we conclude that $\mathsf{ecc}(f(v)) - \mathsf{rad}(\mathbf{T}) \leq \beta(v) = \gamma'(v) + 1$, as required. Moreover, $\mathsf{ecc}(f(v))$ is odd $\iff$ v is white. Therefore (ii) holds for v.

We may therefore assume that v is not the first vertex in $\prec$. Let U denote the set of vertices $u \prec v$ with $\delta(u, v) < \infty$. For each $u \in U$, write I_u to denote the set $I_u = \{\mu \in P \mid \mathsf{dist}(f(u), \mu) \leq \delta(u, v)\}$. Note that since v is not first in $\prec$, the set U is non-empty by Lemma 7.9. Let L denote the intersection of the sets $\{I_u\}_{u \in U}$. Observe that each I_u is a subpath of the path P. Since subpaths of a path have the strong Helly property [21] and since $U \neq \emptyset$, there exist $x, y \in U$ (possibly $x = y$) such that $L = I_x \cap I_y$. Thus, L is also a subpath of P.

Define c to be the color (black or white) such that $f(x)$ has color c if and only if $\delta(x, v)$ is even. We have the following observations.

($+_2$) *Let $\mu \in L$ have color c. Then $f(v) = \mu$ satisfies (IIIa) and (IIIb) for all $u \prec v$.*

Since $\mu \in L$, setting $f(v) = \mu$ satisfies (IIIa), since $\mu \in I_u$ for all $u \in U$ (and thus $\mathsf{dist}(f(u), \mu) \leq \delta(u, v)$ for all $u \in U$). For contradiction, assume that (IIIb) fails for some $u \in U$. Recall that μ has color c. Since u fails (IIIb) for $f(v) = \mu$, we have that $f(u)$ has color c if and only if $\delta(u, v)$ is odd. Recall that $f(x)$ has color c if and only if $\delta(x, v)$ is even. This implies that $x \neq u$, and also that $f(u)$ and $f(x)$ have different colors if and only if $\delta(u, v) + \delta(x, v)$ is even. Since $x \neq u$, we may apply Lemma 7.11 to deduce that $\delta(u, x) + \delta(x, v) + \delta(u, v)$ is an even number. Since both x and u appear before v in $\prec$, we have by the inductive hypothesis that x, u satisfy (IIIb). Thus $f(u)$ and $f(x)$ have the same color if and only if $\delta(u, x)$ is even. Put together, we deduce that $f(u)$ and $f(x)$ have the same color $\iff$ $\delta(u, x)$ is even $\iff$ $\delta(x, v) + \delta(u, v)$ is even $\iff$ $f(u)$ and $f(x)$ have different colors, a contradiction. This proves ($+_2$).

($+_3$) *Suppose that $\mu \in L$ has a neighbor $\eta \in P \setminus L$. Then μ has color c, and there exists $u \in U$ such that $\mathsf{dist}(f(u), \mu) = \delta(u, v) < \mathsf{dist}(f(u), \eta)$.*

Recall that $\mu \in L = I_x \cap I_y$. Thus since $\eta \notin L$, we have either $\eta \notin I_x$ or $\eta \notin I_y$.

Suppose first that $\eta \notin I_x$. Then $\mathsf{dist}(f(x), \eta) > \delta(x, v)$ and so

$$\mathsf{dist}(f(x), \mu) \leq \delta(x, v) < \mathsf{dist}(f(x), \eta) \leq \mathsf{dist}(f(x), \mu) + \mathsf{dist}(\mu, \eta) = \mathsf{dist}(f(x), \mu) + 1$$

This proves that $\mathsf{dist}(f(x), \mu) = \delta(x, v)$. Since the colors black and white alternate on paths in $\mathbf{T}$, it follows that $f(x)$ and μ have the same color $\iff$ $\delta(x, v)$ is even. Recall that $f(x)$ has color c $\iff$ $\delta(x, v)$ is even. This implies that μ has color c, and that we may take $u := x$ for the existential statement.

We may therefore assume that $\eta \in I_x$ and $\eta \notin I_y$. In particular, this implies $x \neq y$. Similarly to the above, we deduce that $\mathsf{dist}(f(y), \mu) = \delta(y, v)$, and hence, $f(y)$ and μ have the same color $\iff$ $\delta(y, v)$ is even. Since $x \neq y$ and $\{x, y\} \prec v$, the inductive hypothesis applies to x and y, and so by (IIIb), we have that $f(x)$ and $f(y)$ have the same color $\iff$ $\delta(x, y)$ is even. Therefore $f(x)$ and μ have the same color $\iff$ $\delta(x, y) + \delta(y, v)$ is even. Moreover, by Lemma 7.11, $\delta(x, y) + \delta(x, v) + \delta(y, v)$ is even. Finally, recall that $f(x)$ has color c $\iff$ $\delta(x, v)$ is even. Putting all together, we deduce that $f(x)$ has color c $\iff$ $\delta(x, v)$ is even $\iff$ $\delta(x, y) + \delta(y, v)$ is even $\iff$ $f(x)$ and μ have the same color. This implies that μ has color c, and that we may take $u := y$ for the existential statement.

This proves $(+_3)$.

Recall that by (IVa), we have $\delta(u, v) \geq \beta(v) - 1$ for all $u \in U$.

$(+_4)$ *Suppose that there exists $u \in U$ such that $\delta(u, v) = \beta(v) - 1$ and $f(u)$ is an endpoint of P. Then $\beta(v) = 1$.*

First, suppose that $\mathsf{diam}(\mathbf{T})$ is odd. Since $u \prec v$, applying the inductive hypothesis to u yields $\mathsf{ecc}(f(u)) - \mathsf{rad}(\mathbf{T}) \leq \gamma(u)$. Since $f(u)$ is an endpoint of P which is a longest path of $\mathbf{T}$, we have $\mathsf{ecc}(f(u)) = \mathsf{diam}(\mathbf{T}) = 2\mathsf{rad}(\mathbf{T}) - 1$. This implies $\gamma(u) \geq \mathsf{rad}(\mathbf{T}) - 1$. Observe, by the definition of $\gamma(v)$, that $\gamma(v) \geq 2\beta(v) - 2 + \gamma(u) - \delta(u, v)$. Recall that $\delta(u, v) = \beta(v) - 1$, and that we assume (IVb). Thus $\gamma(v) \leq \mathsf{rad}(T) - 1$ and so

$$\mathsf{rad}(\mathbf{T}) - 1 \geq \gamma(v) \geq 2\beta(v) - 2 + \gamma(u) - \delta(u, v) \geq \beta(v) - 1 + \mathsf{rad}(\mathbf{T}) - 1$$

Therefore the above inequalities are, in fact, equalities and $\beta(v) = 1$, as claimed.

So we may assume, for contradiction, that $\mathsf{diam}(\mathbf{T})$ is even and $\beta(v) = 2$. Recall that we assume that $\delta(u, v) = \beta(v) - 1 = 1$. Let Q be the walk minimizing $\delta(u, v)$, i.e., $\delta(u, v) = \lambda(Q) = 1$. By definition, $|Q|$ and $\lambda(Q)$ have the same parity. Thus since $\lambda(Q) = 1$ is odd, the walk Q has odd length. This shows that the vertices u, v have different colors.

There are two cases to consider. Suppose first that u is black and v is white. Since $u \prec v$, the condition (ii) holds for u, namely $\mathsf{ecc}(f(u)) - \mathsf{rad}(\mathbf{T}) \leq \gamma'(u) + 1$. Recall that $f(u)$ is an endpoint of P, and so $\mathsf{ecc}(f(u)) = \mathsf{diam}(\mathbf{T}) = 2\mathsf{rad}(\mathbf{T})$. Therefore $\gamma'(u) \geq \mathsf{rad}(\mathbf{T}) - 1$. Observe by the definition of $\gamma'(v)$ that $\gamma'(v) \geq 2\beta(v) - 2 + \gamma'(u) - \delta(u, v)$. Since $\delta(u, v) = 1$, $\beta(v) = 2$, and $\gamma'(u) \geq \mathsf{rad}(\mathbf{T}) - 1$, we can calculate

$$\gamma'(v) \geq 2\beta(v) - 2 + \gamma'(u) - \delta(u, v) \geq \gamma'(u) + 1 \geq \mathsf{rad}(\mathbf{T})$$

This shows that $\gamma'(v) \geq \mathsf{rad}(\mathbf{T})$, but v is white, contradicting our assumption $(+_0)$.

Similarly, suppose that u is white and v is black. By (ii), we have that $\mathsf{ecc}(f(u))$ is odd, since u is white. However, $f(u)$ is an endpoint of $\mathbf{T}$ and hence $\mathsf{ecc}(f(u)) = \mathsf{diam}(\mathbf{T})$ which is even, a contradiction. This proves $(+_4)$.

$(+_5)$ $|L| \geq 2\beta(v) - 1$

To see this, note that $f(x) \in P$ and $f(y) \in P$ by the inductive hypothesis. We consider the following possibilities.

Suppose first that $\mathsf{dist}(f(x), f(y)) < \delta(y, v)$. By the definition of I_y, this implies $f(x) \in I_y$ and so $f(x) \in L = I_x \cap I_y$. We recall that $\delta(x, v) \geq \beta(v) - 1$. If $\beta(v) = 1$, we deduce $\{f(x)\} \subseteq L$ and so $|L| \geq 2\beta(v) - 1$, as claimed. So we may assume $\beta(v) = 2$. If $f(x)$ is not an endpoint of P, then both its neighbors belong to L, since $\delta(x, v) \geq \beta(v) - 1$ and $\mathsf{dist}(f(x), f(y)) < \delta(y, v)$. So $f(x)$ together with those two neighbors implies $|L| \geq 2\beta(v) - 1$. We may therefore assume that $f(x)$ is an endpoint of P. By $(+_4)$, we deduce that $\delta(x, v) \geq \beta(v)$. Recall that $f(y) \in P$. If $\mathsf{dist}(f(x), f(y)) \geq 2$, then $f(x)$ together with next two vertices on P belong to I_y

and thus to L, since $\mathsf{dist}(f(x), f(y)) < \delta(y, v)$. If $\mathsf{dist}(f(x), f(y)) = 1$, then since $\delta(y, v) \geq \beta(v) - 1 = 1$, we again have that the $f(x)$ and next two vertices on P are in L. If $f(x) = f(y)$, then $f(y)$ is an endpoint of P and so $\delta(y, v) \geq \beta(v)$ by $(+_4)$. Thus again $f(x)$ and next two vertices on P are in L. In all cases, $|L| \geq 2\beta(v) - 1$.

Therefore, we may assume that $\mathsf{dist}(f(x), f(y)) \geq \delta(y, v)$, and symmetrically, $\mathsf{dist}(f(x), f(y)) \geq \delta(x, v)$. Since $\{x, y\} \prec v$, the inductive hypothesis applies to x, y which yields $\mathsf{dist}(f(x), f(y)) \leq \delta(x, y)$ by (IIIa). Recall that $f(x) \in P$ and $f(y) \in P$. Let P' denote the subpath of P between $f(x)$ and $f(y)$. Note that P' is a shortest path betweeen x and y, since there is a unique path between any two vertices in $\mathbf{T}$. By Lemma 7.11, we have $\delta(x, y) \leq \delta(x, v) + \delta(y, v) - 2\beta(v) + 2$. Together we deduce $\mathsf{dist}(f(x), f(y)) \leq \delta(x, v) + \delta(y, v) - 2\beta(v) + 2$. Since $\mathsf{dist}(f(x), f(y)) \leq \delta(x, v) + \delta(y, v) - 2\beta(v) + 2$, the path P' contains at most $\delta(x, v) + \delta(y, v) - 2\beta(v) + 3$ vertices, where the first $\delta(x, v) + 1$ belong to I_x, since $\mathsf{dist}(f(x), f(y)) \geq \delta(x, v)$, and the last $\delta(y, v) + 1$ belong to I_y, since $\mathsf{dist}(f(x), f(y)) \geq \delta(y, v)$. Thus by inclusion-exclusion,

$$|I_x \cap I_y| \geq \delta(x, v) + 1 + \delta(v, y) + 1 - \big(\delta(x, v) + \delta(y, v) - 2\beta(v) + 3\big) = 2\beta(v) - 1$$

This proves that $|L| \geq 2\beta(v) - 1$ as claimed. This proves $(+_5)$.

After these preliminary observations, we are ready to analyze cases.

Case 1a: assume $\mathsf{diam}(\mathbf{T})$ is odd and $m_1, m_2 \in L$. Note that $\gamma(v) \geq \beta(v) - 1$, since v is not first in $\prec$. Since m_1, m_2 are adjacent, they have different colors. By symmetry, we may assume that m_1 has color c. If $\beta(v) = 1$, Prover offers for v the vertex m_1 which becomes $f(v)$. Since m_1 has color c, we have by $(+_2)$ that setting $f(v) = m_1$ satisfies (IIIa), (IIIb), and $f(v) \in P$. Since $\mathsf{ecc}(m_1) = \mathsf{rad}(\mathbf{T})$, we have $\mathsf{ecc}(f(v)) - \mathsf{rad}(\mathbf{T}) = 0 \leq \beta(v) - 1 \leq \gamma(v)$, as required.

Thus assume $\beta(v) = 2$. Since $\mathsf{diam}(\mathbf{T}) \geq 3$, the vertex m_2 is not an endpoint of P. Thus m_2 has a neighbor $\mu \neq m_1$ on P. By $(+_3)$, we conclude that $\mu \in L$, since m_2 does not have color c. This also implies that μ has color c. Observe that $\mathsf{ecc}(\mu) \leq \mathsf{rad}(\mathbf{T}) + 1$, since μ is adjacent to m_2 which is a center of $\mathbf{T}$. Prover therefore offers the set $\{m_1, \mu\}$. Adversary chooses a value from this set to become $f(v)$. It follows that $f(v)$ satisfies (IIIa) and (IIIb), by $(+_2)$, since both m_1 and μ are in L and have color c. Moreover, $f(v) \in P$ and since $\mathsf{ecc}(\mu) \leq \mathsf{rad}(\mathbf{T}) + 1$ and $\mathsf{ecc}(m_1) = \mathsf{rad}(\mathbf{T})$, we have $\mathsf{ecc}(f(v)) - \mathsf{rad}(\mathbf{T}) \leq 1 = \beta(v) - 1 \leq \gamma(v)$, as required.

Case 1b: assume that $\mathsf{diam}(\mathbf{T})$ is even and both m and its two neighbors μ_1, μ_2 on P are in L. Observe that μ_1, μ_2 either both have color c or neither has. If both have color c, then Prover offers μ_1 or $\{\mu_1, \mu_2\}$ if $\beta(v) = 1$ or $\beta(v) = 2$, respectively. It follows by $(+_2)$ that setting $f(v) = \mu_i$ satisfies both (IIIa) and (IIIb) and $f(v) \in P$. Moreover, since $\mathsf{ecc}(\mu_i) \leq \mathsf{rad}(\mathbf{T}) + 1$, we deduce that $\mathsf{ecc}(f(v)) - \mathsf{rad}(\mathbf{T}) \leq 1 \leq \beta(v) = \gamma'(v) + 1$, as required. Similarly, if m has color c, then neither μ_1 nor μ_2 has color c and so by $(+_3)$ both neighbors of μ_1 and μ_2 on P are in L. Let μ be one such neighbor different from m. Prover offers m or $\{m, \mu\}$ if $\beta(v) = 1$ or $\beta(v) = 2$, respectively. Note that μ also has color c. Thus by $(+_2)$, any choice that Adversary makes for $f(v)$ satisfies (IIIa) and (IIIb). Moreover, we have $\mathsf{ecc}(f(v)) = \mathsf{rad}(\mathbf{T})$ if $\beta(v) = 1$, and $\mathsf{ecc}(f(v)) \leq \mathsf{rad}(\mathbf{T}) + 2$ if $\beta(v) = 2$. Thus $\mathsf{ecc}(f(v)) - \mathsf{rad}(\mathbf{T}) \leq \beta(v) = \gamma'(v) + 1$, as required.

Case 2: all other possibilities. Recall that L is a subpath of P, and no two consecutive vertices on P have the same eccentricity, except for m_1, m_2 if $\mathsf{diam}(\mathbf{T})$ is odd but in that case, we assume that not both m_1, m_2 are in L. This implies that L lies on P on one side of the centers (possibly including one center). We therefore conclude that there exists $\mu \in L$ and $i \in \{1, 2\}$ such that $\mu \neq \alpha_i$ (recall α_i are the endpoints of P), and $\mathsf{ecc}_P(\mu) = \mathsf{dist}(\mu, \alpha_i)$, and such that μ is the only vertex in L on the subpath P' of P between μ and α_i.

In particular, μ has a neighbor $\eta \in P'$. Thus $\eta \notin L$ and by $(+_3)$, we deduce that μ has color c and $\mathsf{dist}(f(u), \mu) = \delta(u, v) < \mathsf{dist}(f(u), \eta)$ for some $u \in U$.

We have $f(u) \in P$ by the inductive hypothesis. We observe that $f(u)$ does not belong to P', except if $f(u) = \mu$, since $\mathsf{dist}(f(u), \mu) < \mathsf{dist}(f(u), \eta)$. Therefore concatenating to P' the subpath of P between $f(u)$ and μ yields a shortest path from $f(u)$ to α_i. By $(+_1)$, we note that $\mathsf{ecc}_P(\mu) = \mathsf{ecc}(\mu) = \mathsf{dist}(\alpha_i, \mu)$. Put together,

$$\mathsf{ecc}(\mu) + \delta(u, v) = \mathsf{dist}(\alpha_i, \mu) + \mathsf{dist}(f(u), \mu) = \mathsf{dist}(f(u), \alpha_i) \leq \mathsf{ecc}(f(u))$$

If $\beta(v) = 1$, then Prover offers for v the vertex μ which becomes $f(v)$. Since μ is in L and has color c, we deduce by $(+_2)$ that (IIIa), (IIIb) hold for $f(v) = \mu$. We also clearly have $f(v) \in P$. If $\mathsf{diam}(\mathbf{T})$ is odd, we deduce from the definition of $\gamma(v)$ that $\gamma(v) \geq 2\beta(v) - 2 + \gamma(u) - \delta(u, v)$. Thus $\gamma(v) \geq \gamma(u) - \delta(u, v)$, since $\beta(v) = 1$. By (i), we have $\mathsf{ecc}(f(u)) - \mathsf{rad}(\mathbf{T}) \leq \gamma(u)$. Recall that $\mathsf{ecc}(\mu) + \delta(u, v) \leq \mathsf{ecc}(f(u))$. When put together, we have

$$\mathsf{ecc}(f(v)) = \mathsf{ecc}(\mu) \leq \mathsf{ecc}(f(u)) - \delta(u, v) \leq \gamma(u) - \delta(u, v) + \mathsf{rad}(\mathbf{T}) \leq \gamma(v) + \mathsf{rad}(\mathbf{T})$$

We therefore conclude that $\mathsf{ecc}(f(v)) - \mathsf{rad}(\mathbf{T}) \leq \gamma(v)$ as required for (i).

If $\mathsf{diam}(\mathbf{T})$ is even, the process is similar. From (ii), we have $\mathsf{ecc}(f(u)) - \mathsf{rad}(\mathbf{T}) \leq \gamma'(u) + 1$, and by the definition of $\gamma'(v)$, we have $\gamma'(v) \geq \gamma'(u) - \delta(u, v)$, since $\beta(v) = 1$. The facts $\mathsf{ecc}(\mu) + \delta(u, v) \leq \mathsf{ecc}(f(u))$ and $f(v) = \mu$ now yield $\mathsf{ecc}(f(v)) - \mathsf{rad}(\mathbf{T}) \leq \gamma'(v) + 1$ as required. It remains to show that $\mathsf{ecc}(f(v))$ is odd $\iff$ v is white.

Let Q be the walk that minimizes $\delta(u, v)$, i.e., $\delta(u, v) = \lambda(Q)$. By definition, $\lambda(Q)$ and $|Q|$ have the same parity. This implies that u, v have the same color $\iff$ $\delta(u, v)$ is even. Now recall that $\mathsf{dist}(f(u), \mu) = \delta(u, v)$. Since both $\mu = f(v)$ and $f(u)$ belong to P, we deduce by $(+_1)$ that $\mathsf{ecc}(f(u))$, $\mathsf{ecc}(f(v))$ have the same parity $\iff$ $\delta(u, v)$ is even. Put together, we conclude that u, v have the same color $\iff$ $\mathsf{ecc}(f(u))$, $\mathsf{ecc}(f(v)$ have the same parity. Thus since by (ii), we have $\mathsf{ecc}(f(u))$ is odd $\iff$ u is white, it now follows that $\mathsf{ecc}(f(v))$ is odd $\iff$ v is white, as required for (ii).

It remains to consider $\beta(v) = 2$. In this case, we conclude by $(+_5)$ that the set L has at least $2\beta(v) - 1 = 3$ elements. In particular, since L is a subpath of P and μ is one of the endpoints of L, the set L contain another vertex μ' at distance 2 from μ. Since μ has color c, so does μ'. Moreover, by $(+_1)$, we deduce $\mathsf{ecc}(\mu') = \mathsf{ecc}(\mu) + 2$, since $\mathsf{dist}(\mu, \mu') = 2$ and $\mu, \mu' \in P$. Thus $\mathsf{ecc}(\mu') \leq \mathsf{ecc}(f(u)) - \delta(u, v) + 2$, since $\mathsf{ecc}(\mu) \leq \mathsf{ecc}(f(u)) - \delta(u, v)$. Prover therefore offers the set $\{\mu, \mu'\}$. Adversary chooses $f(v)$ from this set. Since both μ, μ' are in L and have color c, we satisfy (IIIa), (IIIb) by $(+_2)$. Also $f(v) \in P$. We again distinguish at the case of odd and even diameter.

If $\mathsf{diam}(\mathbf{T})$ is odd, we have by the definition of $\gamma(v)$ that $\gamma(v) \geq \gamma(u) - \delta(u, v) + 2$, since $\beta(v) = 2$. By (i), we have $\mathsf{ecc}(f(u)) \leq \gamma(u) + \mathsf{rad}(\mathbf{T})$. Since $f(v) \in \{\mu, \mu'\}$, we have $\mathsf{ecc}(f(v)) \leq \mathsf{ecc}(f(u)) - \delta(u, v) + 2$. So put together

$$\mathsf{ecc}(f(v)) \leq \mathsf{ecc}(f(u)) - \delta(u, v) + 2 \leq \gamma(u) + \mathsf{rad}(\mathbf{T}) - \delta(u, v) + 2 \leq \gamma(v) + \mathsf{rad}(\mathbf{T})$$

We therefore conclude that $\mathsf{ecc}(f(v)) - \mathsf{rad}(\mathbf{T}) \leq \gamma(v)$, as required for (i).

If $\mathsf{diam}(\mathbf{T})$ is even, then by (ii), we have $\mathsf{ecc}(f(u)) - \mathsf{rad}(\mathbf{T}) \leq \gamma'(u) + 1$ and the same calculation yields $\mathsf{ecc}(f(v)) - \mathsf{rad}(\mathbf{T}) \leq \gamma'(v) + 1$. We recall that in the case $\beta(v) = 1$ we proved that have $\mathsf{ecc}(\mu)$ is odd $\iff$ v is white (in fact this holds even when $\beta(v) \neq 1$). Thus $\mathsf{ecc}(\mu')$ is odd $\iff$ v is white, since $\mathsf{ecc}(\mu') = \mathsf{ecc}(\mu) + 2$ as argued earlier. Since $f(v) \in \{\mu, \mu'\}$, we therefore conclude that $\mathsf{ecc}(f(v))$ is odd $\iff$ v is white, proving that (ii) holds for v.

This completes the proof (of Lemma 7.15 and with it (IV)$\Rightarrow$(III)). $\square$

7.6. Main theorem. With this characterization, the main theorem of this section is straightforward.

THEOREM 7.16. $\{1,2\}$-CSP($\mathbf{T}$) *is decidable in polynomial time.*

Proof. If $\mathsf{diam}(\mathbf{T}) \leq 2$, then $\mathbf{T}$ is a star $\mathbf{K}_{1,\ell}$ and $\{1,2\}$-CSP($\mathbf{T}$) is in L by Proposition 6.1. Thus we may assume $\mathsf{diam}(\mathbf{T}) \geq 3$. If the input formula Ψ has disconnected graph $\mathbf{G} = \mathbf{D}_\psi$, we apply Theorem 7.7 on each connected component of $\mathbf{G}$ individually. Clearly, Ψ is a yes-instance if and only if every connected component of $\mathbf{G}$ yields a yes-instance. We may therefore assume that $\mathbf{G}$ is connected. To use Theorem 7.7, we observe that the values $\delta(u,v)$, $\gamma(v)$, and $\gamma'(v)$ can be easily computed in polynomial time by dynamic programming. This allows us to test conditions of Theorem 7.7 and thus decide $\{1,2\}$-CSP($\mathbf{T}$) in polynomial time as claimed. ☐

An immediate question arises from the above proof. Is it possible to test the conditions of Theorem 7.7 in L? If the walks that we need to examine in order to do so were of bounded length, then the answer would be yes. This may be unlikely. Consider a long path P on $\exists^{\geq 1}$ and $\exists^{\geq 2}$ vertices, where for every subpath Q of P, we have $k \leq \lambda(Q) \leq \ell$ for some constants k, $\ell > 0$. In particular, $\lambda(P) \geq k$. Let u be the first vertex on P, and assume the last vertex on P is $\exists^{\geq 2}$. Then attach to the end of P another path P' with $\lambda(P)$ vertices, each quantified $\exists^{\geq 2}$. In the quantifier order $\prec$, put u first, then the vertices of P', and then the rest of P. It follows that $P'' = P \cup P'$ is a bad walk, however, we cannot tell this unless we examine the entirety of P''. Moreover, the path P yields a yes-instance of $\{1,2\}$-CSP($\mathbf{P}_{2\ell}$) and therefore for a sufficiently large tree, we cannot discard this instance based on the values γ or γ'. We return to this question in the conclusion of the paper.

Now note that the proof of Theorem 7.7 implies the following interesting corollary.

COROLLARY 7.17. *Let $\mathbf{T}$ be a tree. Let $\mathbf{P}$ be a longest path that is a substructure of $\mathbf{T}$. Then Ψ is a yes-instance of $\{1,2\}$-CSP($\mathbf{T}$) if and only if Ψ is a yes-instance of $\{1,2\}$-CSP($\mathbf{P}$).*

This implies the following main case of Theorem 1.4.

COROLLARY 7.18. $\{1,2\}$-CSP($\mathbf{H}$) *is polynomial-time solvable, for each forest* $\mathbf{H}$.

Proof. Let Ψ be a given instance to $\{1,2\}$-CSP($\mathbf{H}$) and let $\mathbf{G} = \mathbf{D}_\psi$ be the corresponding graph. Just like in the proof of Theorem 7.16, we may assume that $\mathbf{G}$ is connected. The problem reduces to the case when $\mathbf{H}$ is a tree as follows.

Let Ψ' be obtained from Ψ by making the first quantified variable (vertex of $\mathbf{G}$) an $\exists^{\geq 1}$ vertex (possibly $\Psi' = \Psi$). The theorem will follow from this observation.

(+) $\mathbf{H} \models \Psi$ *if and only if at least one of the following,*
- (i) $\mathbf{K} \models \Psi$ *for some connected component* $\mathbf{K}$ *of* $\mathbf{H}$,
- (ii) $\mathbf{K} \models \Psi'$ *and* $\mathbf{K}' \models \Psi'$ *for some distinct components* $\mathbf{K}$, $\mathbf{K}'$ *of* $\mathbf{H}$.

If $\mathbf{K} \models \Psi$ for some connected component $\mathbf{K}$ of $\mathbf{H}$, then clearly also $\mathbf{H} \models \Psi$. If $\mathbf{K} \models \Psi'$ and $\mathbf{K}' \models \Psi'$ for distinct components $\mathbf{K}$, $\mathbf{K}'$ of $\mathbf{H}$, we may assume $\Psi \neq \Psi'$. Prover then plays the following strategy. For the first variable, Prover offers the pair of initial vertices from $\mathbf{K}$ and $\mathbf{K}'$ from the winning strategies for the games $\mathscr{G}(\Psi', \mathbf{K})$ and $\mathscr{G}(\Psi', \mathbf{K}')$. Then based on the choice of Adversary, Prover plays the respective strategy for Ψ' on $\mathbf{K}$ or $\mathbf{K}'$. This shows that $\mathbf{H} \models \Psi$.

For the converse, suppose that $\mathbf{H} \models \Psi$. Then Prover has a winning strategy in the game $\mathscr{G}(\Psi, \mathbf{H})$. Let us examine the values that Prover offers for the first variable in Ψ in this strategy. If there is only one value from a connected component $\mathbf{K}$, or two values from $\mathbf{K}$, then the remaining values also belong to $\mathbf{K}$, because $\mathbf{G}$ is connected.

Thus $\mathbf{K} \models \Psi$ and we obtain (i). If there are two values from distinct connected components $\mathbf{K}, \mathbf{K}'$ of $\mathbf{H}$, then any choice from these values fixes the remaining vertices to the component $\mathbf{K}$ or $\mathbf{K}'$. In particular, each choice yields a winning strategy for Ψ', namely one strategy for the game $\mathscr{G}(\Psi', \mathbf{K})$ and one for the game $\mathscr{G}(\Psi', \mathbf{K}')$. This shows that $\mathbf{K} \models \Psi'$ and $\mathbf{K}' \models \Psi'$. We thus obtain (ii). This proves (+).

Thus in view of (+) and Corollary 7.17, we let $\mathbf{P}$ be a longest path in $\mathbf{H}$, and see that $\mathbf{H} \models \Psi$ if and only if either $\mathbf{P} \models \Psi$, or $\mathbf{P} \models \Psi'$ and at least two connected components of $\mathbf{H}$ contain a longest path. Both can be tested in polynomial time by Theorem 7.16. This concludes the proof. ∎

8. Proof of Theorem 1.4. For the proof, the last remaining piece of the puzzle is the following proposition.

PROPOSITION 8.1. *If* $\mathbf{H}$ *is a bipartite graph whose smallest cycle is* $\mathbf{C}_{2j}$ *for* $j \geq 3$, *then* $\{1,2\}$-CSP($\mathbf{H}$) *is* Pspace-*complete.*

Proof. We reuse one of the reductions used to prove case (iii) of Theorem 1.2. We briefly discuss the key steps. The reduction is from QCSP($\mathbf{K}_j$). Let Ψ be an input formula for QCSP($\mathbf{K}_j$). We begin by considering the graph $\mathbf{D}_\psi$ to which we add a disjoint copy $V = \{v_1, \ldots, w_{2j}\}$ of $\mathbf{C}_{2j}$. Then we replace every edge $(x, y) \in \mathbf{D}_\psi$ with a gadget shown in Figure 5.1, where the black vertices are identified with V. Finally, for $\forall$ variables x of Ψ, we add a new path $z_1, z_2, \ldots, z_j$ where $z_j = x$.

The resulting graph defines the quantifier-free part of θ of our desired formula Θ. The quantification in Θ is as follows. The outermost quantification is $\exists^{\geq 2} v_1, \ldots, v_{j+1}$ $\exists^{\geq 1} v_{j+2}, \ldots, v_{2j}$. Then we move inwards through the quantifier order of Ψ; when we encounter an existential variable x, we apply $\exists^{\geq 1}$ to it in Θ. When we encounter a $\forall$ variable x, we apply $\exists^{\geq 2}$ to the path $z_1, z_2, \ldots, z_j$ constructed for x, in that order. All the remaining variables are then quantified $\exists^{\geq 1}$.

As proved in Theorem 1.2, the cycle $\mathbf{C}_{2j}$ models Θ if and only if $\mathbf{K}_j$ models Ψ. We now adjust this to the bipartite graph $\mathbf{H}$. There are three difficulties arising from simply using the above construction as it is.

Firstly, assume the variables $v_1, \ldots, v_{2j}$ are mapped to a fixed copy C of $\mathbf{C}_{2j}$ in $\mathbf{H}$. We need to ensure that variables x, y, derived from the original instance Ψ and constrained by $E(x, y)$, are also mapped to C. For y variables in our gadget one can check this must be true – the successive cycles in the edge gadget may never deviate from C, since $\mathbf{H}$ contains no cycle smaller than $2j$. For x variables off on the pendant this might not be true. To fix this, we insist that Ψ contains an atom $E(x, y)$ iff it also contains $E(y, x)$; QCSP($\mathbf{K}_j$) remains Pspace-complete on such instances [3].

Secondly, we need to check that Adversary has freedom to assign any value from C to each $\forall$ variable x. Consider $z_1, \ldots, z_j$, the path associated with x. As long as Prover offers values for $z_1, \ldots, z_j$ from C, Adversary has freedom to chose any value for $x = z_j$. If on the other hand Prover offers for one of $z_1, \ldots, z_j$, say for z_i, a value not on C, then Adversary can choose all subsequent $z_{i+1}, \ldots, z_j$ to also be mapped outside C, since $\mathbf{H}$ has no cycle shorter than $\mathbf{C}_{2j}$. Thus $x = z_j$ ends up mapped outside C, which we ensured is not possible, i.e., Prover would lose if she had done this.

Finally, we discuss how to ensure that V is mapped to a copy of $\mathbf{C}_{2j}$. Since the first $j + 1$ vertices of V are quantified $\exists^{\geq 2}$, Adversary can force this by always choosing a value not seen already when going through each of $v_1, \ldots, v_{j+1}$ in turn. If this is not possible (both offered values have been seen), this gives rise to a cycle in $\mathbf{H}$ shorter than $\mathbf{C}_{2j}$, which does not exist. In conclusion, if Adversary maps V to a cycle, then Prover must play exclusively on this cycle, thus solving QCSP($\mathbf{K}_j$).

If Adversary maps V to a subpath of $\mathbf{C}_{2j}$, then Prover can play to win (certainly if Φ is a yes-instance and possibly even if not). So the situation is just like with $\{1,2\}$-CSP($\mathbf{C}_{2j}$). $\square$

Finally, we are ready to prove a dichotomy for $\{1,2\}$-CSP($\mathbf{H}$) where $\mathbf{H}$ is a graph (and a stronger dichotomy when $\mathbf{H}$ is a bipartite graph).

Proof of Theorem 1.4. If $\mathbf{H}$ is not bipartite, then $\{1\}$-CSP($\mathbf{H}$) is NP-hard by [22]; thus $\{1,2\}$-CSP($\mathbf{H}$) is also NP-hard. So we may assume that $\mathbf{H}$ is a bipartite graph. If $\mathbf{H}$ contains a $\mathbf{C}_4$, then $\{1,2\}$-CSP($\mathbf{H}$) is in L as shown in Proposition 6.3.

We may therefore assume that G contains no $\mathbf{C}_4$. If $\mathbf{H}$ contains any cycles at all, then the smallest cycle in $\mathbf{H}$ is $\mathbf{C}_{2j}$ where $j \geq 3$, and the problem is Pspace-complete by Proposition 8.1. Thus we may assume that $\mathbf{H}$ contains no cycles and so it is a forest. In this case, $\{1,2\}$-CSP($\mathbf{H}$) is polynomial-time solvable as Corollary 7.18 shows.

This completes the proof. $\square$

9. Finer classifications.

9.1. Partially reflexive graphs. In this section, we briefly list some results for graphs allowing self-loops on some vertices (so-called *partially reflexive* graphs). Our understanding of these cases is rather limited and some recent results [30, 28] suggest that a simple dichotomy is very unlikely. Nonetheless, some cases might still be of further interest. First, we consider the class of undirected graphs with a single dominating vertex w which is also a self-loop. Let $\mathbf{H} \setminus \{w\}$ be the subgraph of $\mathbf{H}$ induces by the set $V(\mathbf{H}) \setminus \{w\}$.

Proposition 9.1. *If $\mathbf{H}$ has a reflexive dominating vertex w and $\mathbf{H} \setminus \{w\}$ contains a loop or is irreflexive bipartite, then $\{1,2\}$-CSP($\mathbf{H}$) is in* P.

Proof. If $\mathbf{H} \setminus \{w\}$ contains a loop then $\mathbf{H}$ contains adjacent looped vertices and $\{1,2\}$-CSP($\mathbf{H}$) is trivial (all instances are yes-instances). Assume $\mathbf{H} \setminus \{w\}$ is an irreflexive bipartite graph and consider an input Ψ for $\{1,2\}$-CSP($\mathbf{H}$). All variables quantified by $\exists^{\geq 1}$ can be evaluated as w and can be safely removed while preserving satisfaction. So, let Ψ' be the subinstance of Ψ induced by the variables quantified by $\exists^{\geq 2}$ and let ψ' be the associated quantifier-free part. If $\mathbf{D}_{\psi'}$ is bipartite, the instance is a yes-instance, otherwise it is a no-instance. $\square$

Proposition 9.2. *If $\mathbf{H}$ has a reflexive dominating vertex w and $\mathbf{H} \setminus \{w\}$ is an irreflexive non-bipartite graph, then $\{1,2\}$-CSP($\mathbf{H}$) is* NP*-complete.*

Proof. For membership of NP we note the following. Let U be a unary predicate defining the set $V(\mathbf{H}) \setminus \{w\}$. From an input Ψ for $\{1,2\}$-CSP($\mathbf{H}$) we will build an instance Ψ' for CSP($\mathbf{H}; U$) so that $\mathbf{H} \models \Psi$ iff $(\mathbf{H}; U) \models \Psi'$. The latter is clearly in NP, so the result follows. To build Ψ' we take Ψ and add $U(v)$ to the quantifier-free part for all $\exists^{\geq 2}$ quantified variables v, before converting in the quantification $\exists^{\geq 2} v$ to $\exists^{\geq 1} v$.

For NP-hardness we reduce from CSP($\mathbf{H} \setminus \{w\}$) which is NP-hard by [22]. For an input Ψ to this, we build an input Ψ' for $\{1,2\}$-CSP($\mathbf{H}$) by converting each $\exists$ to $\exists^{\geq 2}$. It is easy to see that $(\mathbf{H} \setminus \{w\}) \models \Psi$ iff $\mathbf{H} \models \Psi'$ and the result follows. $\square$

We sum up the previous two propositions in the following corollary.

Corollary 9.3. *If $\mathbf{H}$ has a reflexive dominating vertex, then $\{1,2\}$-CSP($\mathbf{H}$) is either in* P *or is* NP*-complete.*

9.2. Small graphs. It follows from Proposition 9.2 that there is a partially reflexive graph on four vertices, $\mathbf{K}_4$ with a single reflexive vertex, so that the corresponding $\{1,2\}$-CSP is NP-complete. We can argue this phenomenon is not visible on smaller graphs.

PROPOSITION 9.4. *Let* $\mathbf{H}$ *be a (partially reflexive) graph on at most three vertices, then either* $\{1,2\}$-CSP($\mathbf{H}$) *is in* P *or it is* Pspace-*complete.*

Proof. The Pspace-complete cases are $\mathbf{K}_3$ (see [25]) and $\mathbf{P}^*_{101}$, which is the path of length two whose internal vertex is loopless while the end vertices are looped. It is known that QCSP($\mathbf{P}^*_{101}$) is Pspace-complete [28]. One can reduce this problem to $\{1,2\}$-CSP($\mathbf{P}^*_{101}$) by substituting $\forall x$ in the former by $\exists^{\geq 2} x, x'\ E(x,x')$ in the latter (where x' is a newly introduced variable). We now address the tractable cases.

Case 1. If $\mathbf{H}$ contains a reflexive clique of size 2, then $\{1,2\}$-CSP($\mathbf{H}$) is trivial.

Case 2. If $\mathbf{H}$ is an irreflexive bipartite graph, i.e., $\mathbf{H}$ is a forest, then $\{1,2\}$-CSP($\mathbf{H}$) is in P according to Corollary 7.18.

Case 3. When $\mathbf{H}$ contains just isolated loops and non-loops then it is easy to give a tailored algorithm, as follows. If $\mathbf{H}$ contains at least two loops then: any input with a subinstance $Qx\exists^{\geq 2}x'\ E(x,x')$, where Q is any quantifier, and $x \neq x'$, is false; all other inputs are true. If $\mathbf{H}$ contains only one loop: any input with a subinstance $Qx\exists^{\geq 2}x'\ E(x,x')$ where Q is any quantifier, and possibly $x = x'$, is false; all other inputs are true. If $\mathbf{H}$ contains just isolated loops then it is bipartite.

Case 4. We are now left with one remaining 2-vertex graph, $\mathbf{P}^*_{10}$, the 2-vertex path with one loop. For this problem, any input with a subinstance $\exists^{\geq 2}x, x'\ E(x,x')$ where possibly $x = x'$, is false; all other inputs are true.

Case 5. We continue with graphs on exactly three vertices.

Case 5a. Among the remaining possibilities where $\mathbf{H}$ has exactly two loops is only $\mathbf{P}^*_{10} \uplus \mathbf{K}^*_1$, the disjoint union of $\mathbf{P}^*_{10}$ and a single vertex with a loop $\mathbf{K}^*_1$. For this problem, any input with a subinstance $\exists^{\geq 2}x, x'\ E(x,x')$ where $x \neq x'$, is false; all other inputs are true.

Case 5b. We now address the case in which there is precisely one loop. If the vertex with a loop dominates, then we have tractability by Proposition 9.1. If it is isolated, then the remaining case is $\mathbf{K}^*_1 \uplus \mathbf{K}_2$, the disjoint union of $\mathbf{K}^*_1$ and $\mathbf{K}_2$. For this, any input Ψ with subinstance $\exists^{\geq 2}x\ E(x,x)$ or an $\exists^{\geq 2}v$ at the beginning of a sequence (connected component of v in $\mathbf{D}_\psi$) that is non-bipartite is false; all other inputs are true. The remaining possibilities are $\mathbf{P}^*_{100}$, the path on three vertices with loop at one endpoint, and $\mathbf{P}^*_{10} \uplus \mathbf{K}_1$, the disjoint union of $\mathbf{P}^*_{10}$ and $\mathbf{K}_1$. For the latter, we have the same $\{1,2\}$-CSP as for $\mathbf{P}^*_{10}$, which has already been resolved. $\{1,2\}$-CSP($\mathbf{P}^*_{100}$) requires some subtlety and appears as its own result in Proposition 9.5.

Case 5c. All other 3-vertex cases are irreflexive and are hence already resolved. ◻

PROPOSITION 9.5. $\{1,2\}$-CSP($\mathbf{P}^*_{100}$) *is in* P.

Proof. Let Φ be an instance of $\{1,2\}$-CSP($\mathbf{P}^*_{100}$) in which we assume without loss of generality that the quantifier-free part ϕ is symmetrically closed, i.e., contains an atom $E(v,v')$ only if it contains $E(v',v)$. The following four types of subinstance in Ψ result in it being false (including the symmetric closure of the quantifier-free parts).

(i) $\exists^{\geq 2}x_1, x_2, x_3\ E(x_1,x_2) \wedge E(x_2,x_3)$

(ii) $Qx_1\exists^{\geq 2}x_2, x_3\ E(x_1,x_3) \wedge E(x_2,x_3)$ (Q is any quantifier)

(iii) $\exists^{\geq 2}x_1, x_2, x_3\ E(x_2,x_3) \wedge E(x_1,y) \wedge E(x_3,y)$
(y is quantified existentially, and x_2 is quantified before x_3)

(iv) $\exists^{\geq 2}x_1, x_2, x_3, x_4\ E(x_1,x_2) \wedge E(x_3,x_4) \wedge E(x_2,y_1) \wedge E(x_4,y_2) \wedge E(y_1,y_2)$
(y_1, y_2 are quantified existentially, x_1 is quantified before x_2, and x_3 before x_4)

We claim that all other inputs are yes-instances. We give the following strategy for Prover. Consider $\mathbf{P}^*_{100}$ to have vertex set $\{0,1,2\}$ with edges 01, 12, and loop at 0.

For $\exists^{\geq 2}$ variables Prover offers $\{0,1\}$, unless constrained by adjacency of a variable already played to 1, in which case she offers $\{0,2\}$. For $\exists$ variables Prover offers $\{0\}$, unless constrained by adjacency of a variable already played to 2, in which case she offers $\{1\}$. We argue this strategy must be winning. This is tantamount to saying that Prover is never offered (a) an $\exists^{\geq 2}$ variable that is simultaneously adjacent to 0 and 1; (b) an $\exists^{\geq 2}$ variable that is adjacent to 2; and (c) an $\exists$ variable that is simultaneously adjacent to both 2 and 1. We observe that (b) follows from Rule (i) and (a) follows from Rule (ii), while (c) follows from (iii) and (iv). ∎

9.3. Multipartite graphs trichotomy. For $1 \leq a_1 \leq \ldots \leq a_n$, let $\mathbf{K}_{a_1,\ldots,a_n}$ be the complete multipartite graph with respective parts of size $a_1, \ldots, a_n$. We refer to a part $a_i = 1$ as *singleton*. We write $\mathbf{K}_n(x_1, \ldots, x_n)$ to denote $\bigwedge_{i \neq j \in [n]} x_i \neq x_j$.

PROPOSITION 9.6. $\{1,2\}$-CSP$(\mathbf{K}_{1,1,2})$ *is* Pspace-*complete.*

Proof. In $\mathbf{K}_{1,1,2}$ let us label the vertices r, b, g, g' in the obvious fashion. If we evaluate $\exists^{\geq 1} p, q \exists^{\geq 2} x\ \mathbf{K}_3(x,p,q)$ then it must be that $\{p,q\}$ is evaluated as $\{r,b\}$ and x is offered to be both $\{g,g'\}$. It follows that $\exists^{\geq 2} w \exists^{\geq 1} p, q \exists^{\geq 2} x\ \mathbf{K}_3(x,p,q) \wedge \mathbf{K}_2(x,w)$ enforces the quantification "w must hold for both r and b". Armed with this it is simple to adapt the Pspace-hardness proof of [3] to reduce *Quantified not-all-equal* 3-*satisfiability* QCSP$(\mathbf{B}_{\text{NAE}})$ to $\{1,2\}$-CSP$(\mathbf{K}_{1,1,2})$, thus proving Pspace-completeness of the latter. Naturally, r and b play the role of true and false. We outline the reduction here (see [3, Proposition 5.1] and [3, Figure 2 on Page 993] for further details.)

For an input Φ of QCSP$(\mathbf{B}_{\text{NAE}})$ we build an input Ψ for $\{1,2\}$-CSP$(\mathbf{K}_{1,1,2})$, via its quantifier-free part ψ, so that the Φ is a yes-instance of QCSP$(\mathbf{B}_{\text{NAE}})$ iff Ψ is a yes-instance of $\{1,2\}$-CSP$(\mathbf{K}_{1,1,2})$. Each variable of Φ becoms a triangle $\mathbf{K}_3$ on $i, \neg v, v$ in $\mathbf{D}_\psi$, where we will identify the i vertex of all of the variable gadgets and ensure it is evaluated as either (or both) g, g' (which we already explained that we can do). Each clause also becomes a triangle $\mathbf{K}_3$ on ℓ, ℓ', ℓ'' in $\mathbf{D}_\psi$ and is connected by a single edge to the respective variable. That is, if we have the clause $(v_1 \vee \neg v_2 \vee v_3)$ then we have edges ℓv_1, $\ell' \neg v_2$, $\ell'' v_3$. Finally we need a single edge variable gadget from a new vertex $v_\forall$ attached to the vertex $\neg v$ for each universal variable v.

The quantification goes as follows. Outermost, i is quantified so as to force it to be $\{g, g'\}$. Then we move in according to the variable order of Φ. For existential variables v, we existentially quantify v and $\neg v$; while for universal variables v we quantify $v_\forall$ so that it ranges over both r and b, then existentially quantify v and $\neg v$. All remaining variables are quantified existentially innermost.

It is not hard to see the reduction stays true to [3] and the result follows. ∎

COROLLARY 9.7. *If* $a_1 = a_2 = 1$, *then* $\{1,2\}$-CSP$(\mathbf{K}_{a_1,\ldots,a_n})$ *is* Pspace-*complete.*

Proof. The proof is similar to that of Proposition 9.6, we just need to isolate two vertices each in singleton parts. Suppose $a_1 = a_2 = \ldots = a_m = 1$ and the remaining $n - m$ are > 1. Parts corresponding to $a_1, \ldots, a_m$ are called singletons. Use

$$(\exists^{\geq 1} y_1, \ldots, y_{m-2}) \exists^{\geq 2} w \exists^{\geq 1} p, q \exists^{\geq 2} x_1, \ldots, x_{n-m} \psi$$

$$\psi := \mathbf{K}_n(x_1, \ldots, x_{n-m}, y_1, \ldots, y_{m-2}, p, q) \wedge \mathbf{K}_{n-1}(x_1, \ldots, x_{n-m}, y_1, \ldots, y_{m-2}, w)$$

to enforce the quantification "w must hold for both of two singletons". In the reduction, $(\exists^{\geq 1} y_1, \ldots, y_{m-2})$ can sit outermost in the quantification and is only written once, whereas the subsequent piece may be repeated whenever the quantification "w must hold for both of two singletons" is needed; $y_1, \ldots, y_{m-1}$ will have to have been evaluated on singletons, so precisely two will remain. ∎

PROPOSITION 9.8. *If* $1 = a_1 < a_2 \leq \ldots \leq a_n$ *then* $\{1,2\}$-CSP$(\mathbf{K}_{a_1,\ldots,a_n})$ *is in* NP.

Proof. Let Φ be an input to CSP($\mathbf{K}_{a_1,\ldots,a_n}$). For $2 \leq i \leq n$, let x_i, y_i be two vertices of $\mathbf{K}_{a_1,\ldots,a_n}$ corresponding to the part of size $a_i > 1$. Let x_1 be the single vertex in the part corresponding to $a_1 = 1$. We claim that the following are equivalent.

(i) Φ is true on $\mathbf{K}_{a_1,\ldots,a_n}$, and
(ii) $\mathbf{D}_\phi$ has a proper n-coloring in which no $\exists^{\geq 2}$ vertex has color 1.

(ii)$\Rightarrow$(i): Using the n-coloring, Prover offers x_i or $\{x_i, y_i\}$ for any vertex of color i. Prover wins, since color 1 is never used on an $\exists^{\geq 2}$ vertex of $\mathbf{D}_\psi$.

(i)$\Rightarrow$(ii): We build an n-coloring for (ii) from Prover's winning strategy for Φ in the following fashion. Each time we meet an $\exists^{\geq 2}$ variable v, Prover must offer at least one choice z that is not x_1, and Adversary chooses z. The vertex z belongs to a part of $\mathbf{K}_{a_1,\ldots,a_n}$ corresponding to a_i for some i. The variable v is then assigned the color i and we continue. This builds the required n-coloring.

It now follows that $\{1,2\}$-CSP($\mathbf{K}_{a_1,\ldots,a_n}$) is in NP by condition (ii). ∎

We are now ready to prove our trichotomy theorem.

Proof of Theorem 1.5. Let $1 \leq a_1 \leq \ldots \leq a_n$. Part (i) appears as Proposition 6.1. For Part (ii), if $a_1 \geq 2$, then each $\exists^{\geq 2}$ may be evaluated as $\exists^{\geq 1}$ and we collapse to NP. The remaing case of Part 2 is Proposition 9.8. Note that NP-hardness in (ii) follows from Hell and Nešetřil [22]. Finally, Part (iii) is proved as Corollary 9.7. ∎

10. The complexity of QCSP($\mathbf{C}_4^*$)**.** Let $\mathbf{C}_4^*$ be the reflexive 4-cycle, i.e., the 4-cycle with loop at every vertex. The complexities of Ret($\mathbf{C}_6$) and Ret($\mathbf{C}_4^*$) are both high (NP-complete) [18, 17], and retraction is recognized to be a "cousin" of QCSP (see [2]). The problem QCSP($\mathbf{C}_6$) is known to be in L (see [29]), but the complexity of QCSP($\mathbf{C}_4^*$) was hitherto unknown. Perhaps surprisingly, we show that is is markedly different from that of QCSP($\mathbf{C}_6$), namely being Pspace-complete.

Proposition 10.1. *$\{1,2,3,4\}$-CSP($\mathbf{C}_4^*$) is Pspace-complete.*

Proof. We will reduce from the problem QCSP($\mathbf{K}_4$), known to be Pspace-complete from, e.g., [3]. We will borrow heavily from the reduction of CSP($\mathbf{K}_4$) to Ret($\mathbf{C}_4^*$) in [17]. The reduction has a very similar flavour to that used in case (iii) of Theorem 1.2, but borrows from [17] instead of [18].

For an input $\Psi := Q_1 x_1 Q_2 x_2 \ldots Q_m x_m\ \psi(x_1, x_2, \ldots, x_m)$ to QCSP($\mathbf{K}_4$), we build an input Ψ' to $\{1,2,3,4\}$-CSP($\mathbf{C}_4^*$). We begin by considering the graph $\mathbf{G} := \mathbf{D}_\psi$, from which we build $\mathbf{G}'$ by taking the disjoint union of $\mathbf{G}$ and a copy of $\mathbf{C}_4^*$. We then build $\mathbf{G}''$ from $\mathbf{G}'$ by replacing every edge $(x, y) \in E(\mathbf{G})$ with the following gadget. In this gadget, the dark vertices z_1, z_2, z_3, z_4 are identified with the fixed copy of $\mathbf{C}_4^*$ in $\mathbf{G}'$.

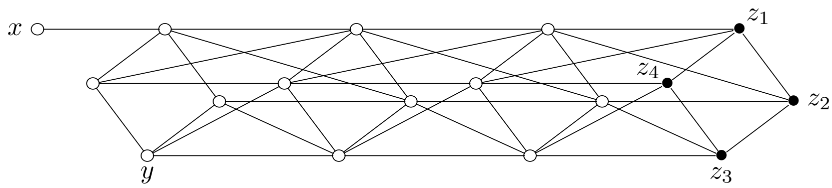


Fig. 10.1. *Reduction gadget for QCSP*($\mathbf{C}_4^*$).

The formula $\phi_{\mathbf{G}''}$ will form the quantifier-free part of Ψ'; we now explain the structure of the quantifiers. Let z_1, z_2, z_3, z_4 be the the fixed copy of $\mathbf{C}_4^*$ in $\mathbf{G}''$ as before. The formula Ψ' begins $\exists z_1 \exists^{\geq 2} z_2 \exists^{\geq 3} z_3 \exists^{\geq 2} z_4$ (z_1 could equally be quantified with $\exists^{\geq j}$, $j > 1$). Now we continue in the quantifier order of Ψ. When we meet an $\exists$

quantifier, we quantify as $\exists$. When we meet a $\forall$ quantifier, we quantify with $\forall = \exists^{\geq 4}$. Finally, we quantify with $\exists$ all remaining variables, corresponding to vertices we added in gadgets in $\mathbf{G}''$. We claim that $\mathbf{K}_4 \models \Psi$ iff $\mathbf{C}_4^* \models \Psi'$. The proof of this proceeds as with Theorem 1.2 (though there are several more degenerate cases to consider). ∎

Corollary 10.2. *QCSP($\mathbf{C}_4^*$) is* Pspace*-complete.*

Proof. We give a reduction from $\{1,2,3,4\}$-CSP($\mathbf{C}_4^*$) to QCSP($\mathbf{C}_4^*$), using the following shorthands (x', x'' are new vertices which appear nowhere else in ϕ).

$\exists^{\geq 1} x \ \phi(x) := \exists x \ \phi(x)$
$\exists^{\geq 2} x \ \phi(x) := \forall x' \exists x \ E(x', x) \wedge \phi(x)$
$\exists^{\geq 3} x \ \phi(x) := \forall x'' \forall x' \exists x \ E(x'', x) \wedge E(x', x) \wedge \phi(x)$
$\exists^{\geq 4} x \ \phi(x) := \forall x \ \phi(x)$

On $\mathbf{C}_4^*$, it is easy to verify that, for each $i \in \{1,2,3,4\}$, $\exists^{\geq i} x \ \phi(x)$ holds iff there exist at least i elements x satisfying ϕ. The result follows easily (note that each use of shorthand substitution involves new variables corresponding to x and x' above).

That concludes the proof. ∎

While QCSP($\mathbf{C}_4^*$) has different complexity from QCSP($\mathbf{C}_6$), we remark that the better analog of the retraction complexities is perhaps that $\{1, |C_4^*|\}$-CSP($\mathbf{C}_4^*$) and $\{1, |C_6|/2\}$-CSP($\mathbf{C}_6$) do have the same complexities (recall the reductions to Ret($\mathbf{C}_4^*$) and Ret($\mathbf{C}_6$) involved CSP($\mathbf{K}_{|C_4^*|}$) and CSP($\mathbf{K}_{|C_6|/2}$), respectively).

11. Conclusion. We have taken first important steps to understanding the complexity of CSPs with counting quantifiers. We mainly focused on graph templates. As we uncovered, this already provides a rich landscape of problems and challenging questions. In the case of cliques and cycles, we have managed a complete classification. For quantifiers $\exists^{\geq 1}$ and $\exists^{\geq 2}$, we have proved a dichotomy for all graphs (and a stronger one for bipartite graphs). In addition, we explored a number of cases of special graphs (with loops, complete multipartite graphs). Our view of the general situation so far does not go beyond basic observations. One important aspect that we have not fully utilized is the connection to algebraic methods as we now discuss.

The algebraic method has been very potent in understanding the complexity of CSPs and QCSPs [8, 5, 1, 3, 10]. Recently, we have become aware of an algebraic theory tailored to counting quantifiers [6] (early version was [7]).

A *polymorphism* of a structure $\mathbf{B}$ is a homomorphism from $\mathbf{B}^k$ to $\mathbf{B}$, for some k. Call a function $f : B^k \to B$ *j-expanding* if, for all $X_1, \ldots, X_k \subseteq B$ such that $|X_1| = \ldots = |X_k| = j$, we have $|f(X_1, \ldots, X_k)| \geq j$. This condition at $j = 1$ is trivial (it says that f is a function) and at $j = |B|$ asserts surjectivity. For $X \subseteq \{1, \ldots, |B|\}$, we say that f is *X-expanding* if it is j-expanding for all $j \in X$. Now, the relations that are X-pp-definable over $\mathbf{B}$ are exactly those that are preserved by the X-expanding polymorphisms of $\mathbf{B}$ [6]. In the case of $\{1\}$-pp and $\{1, |B|\}$-pp, this includes the celebrated Galois connections previously known.

The link between polymorphisms and algorithms has been very useful in the study of CSPs and QCSPs. For counting quantifiers, this leads to the following question.

Question 1. *What expanding polymorphisms might be responsible for tractability of X-CSP($\mathbf{B}$)?*

For example, if $\mathbf{B}$ has an expanding majority polymorphism (expanding for each j) might that bestow tractability as it does for the $\{1, |B|\}$-CSP($\mathbf{B}$) in [3]?

For more combinatorial questions, recall that we have proved that $\{1,2\}$-CSP($\mathbf{H}$) displays dichotomy on graphs between those cases in P and those that are NP-hard. On bipartite graphs, this dichotomy is between P and Pspace-complete. For non-bipartite graphs, it is not difficult to exhibit template graphs $\mathbf{H}$ such that the problem is NP-complete, such as those furnished by Theorem 1.5. However, this is far from a complete answer. Is it possible to extend the dichotomy of Theorem 1.4 to a trichotomy? In other words,

QUESTION 2. *Is it possible to classify all* $\{1,2\}$-CSP($\mathbf{H}$) *problems as either in* P, NP*-complete, or* Pspace*-complete, depending on* $\mathbf{H}$*?*

We conjecture this might be difficult, though we doubt there exist cases of intermediate complexity.

Another direction concerns space complexity. While not the usual focus, a recent work [16] on space complexity of the CSP on graphs suggests a finer classification than the usual P/NP-hard. A natural question is to investigate how this applies to counting quantifiers. In particular, the following seem like reasonable questions.

QUESTION 3. *Is* $\{2\}$-CSP($\mathbf{K}_4$) *in* L*, or is it* NL*-complete, or perhaps* P*-complete?*

QUESTION 4. *Let* $\mathbf{T}$ *be a fixed tree. Is* $\{1,2\}$-CSP($\mathbf{T}$) *in* L*, or is it* NL*-complete, or perhaps* P*-complete?*

Finally, can we uncover a P/ NP-hard dichotomy for X-CSP($\mathbf{H}$) for graphs with any given subset X? This may be achievable, though to then separate the NP-hard from the Pspace-complete would incorporate the open question for QCSP($\mathbf{H}$), and should be considered challenging.

Acknowledgements. We owe a debt of gratitude to reviewers at CSR 2012, CSR 2014 and SIAM Journal of Discrete Mathematics. The third author also wishes to thank Department of Quantitative Methods and Economic Informatics of the University of Žilina where some of the work on this paper was done.

REFERENCES

[1] BARTO, L., KOZIK, M., AND NIVEN, T. 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), 1782–1802.

[2] BODIRSKY, M., KÁRA, J., AND MARTIN, B. The complexity of surjective homomorphism problems – a survey. *Discrete Applied Mathematics 160*, 12 (2012), 1680–1690.

[3] BÖRNER, F., BULATOV, A. A., CHEN, H., JEAVONS, P., AND KROKHIN, A. A. The complexity of constraint satisfaction games and QCSP. *Information and Computation 207*, 9 (2009), 923–944.

[4] BÖRNER, F., KROKHIN, A., BULATOV, A., AND JEAVONS, P. Quantified constraints and surjective polymorphisms. Tech. Rep. PRG-RR-02-11, Oxford University, 2002.

[5] BULATOV, A. A dichotomy theorem for constraint satisfaction problems on a 3-element set. *Journal of the ACM 53*, 1 (2006), 66–120.

[6] BULATOV, A. AND HEDAYATY, A., Counting Predicates, Subset Surjective Functions, and Counting CSPs. *42nd IEEE International Symposium on Multiple-Valued Logic, ISMVL 2012*, (2012), 331–336.

[7] BULATOV, A. AND HEDAYATY, A., Galois correspondence for counting quantifiers. *CoRR abs/1210.3344*, (2012).

[8] BULATOV, A., KROKHIN, A., AND JEAVONS, P. G. Classifying the complexity of constraints using finite algebras. *SIAM Journal on Computing 34* (2005), 720–742.

[9] BUNEMAN P. A note on the metric properties of trees. *Journal of Combinatorial Theory Series B 1* (1974), 48–50.

[10] CHEN, H. The complexity of quantified constraint satisfaction: Collapsibility, sink algebras, and the three-element case. *SIAM Journal on Computing. 37*, 5 (2008), 1674–1701.

[11] CHEN, H. Existentially restricted quantified constraint satisfaction. *Information and Computation 207*, 3 (2009), 369–388.

[12] CHEN, H. Quantified Constraint Satisfaction and the Polynomially Generated Powers Property. *Algebra Universalis*, 65 (2011), 213–241

[13] CHEPOI, V., DRAGAN, F., ESTELLON, B., HABIB, M., VAXES, Y. Diameters, centers, and approximating trees of δ-hyperbolic geodesic spaces and graphs. In: *Symposium on Computational Geometry.* (2008), pp. 59–68.

[14] CLARK, R., AND GROSSMAN, M. Number sense and quantifier interpretation. *Topoi 26*, 1 (2007), 51–62.

[15] EBBINGHAUS, H.-D., AND FLUM, J. *Finite Model Theory.* Springer, 1999. 2nd edition.

[16] Egri, L. , Hell, P., Larose, B., Rafiey, A. Space complexity of list H-colouring: a dichotomy. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA*, (2014), pp. 349–365.

[17] Feder, T., and Hell, P. List homomorphisms to reflexive graphs. *Journal of Combinatorial Theory, Series B 72*, 2 (1998), 236–250.

[18] Feder, T., Hell, P., and Huang, J. List homomorphisms and circular arc graphs. *Combinatorica 19*, 4 (1999), 487–505.

[19] Feder, T., Hell, P., Jonsson, P., Krokhin, A. A., and Nordh, G. Retractions to pseudoforests. *SIAM Journal Discrete Math. 24*, 1 (2010), 101–112.

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

[21] Golumbic, M. C., Jamison, R. E. The edge intersection graphs of paths in a tree, *Journal of Combinatorial Theory, Series B 38*, 1 (1985), 8–22

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

[23] Kolaitis, P. G., and Vardi, M. Y. *Finite Model Theory and Its Applications (Texts in Theoretical Computer Science. An EATCS Series).* Springer-Verlag New York, Inc., 2005, ch. A logical Approach to Constraint Satisfaction.

[24] Lange, K.-J. Some results on majority quantifiers over words. In *Proceedings of the 19th IEEE Annual Conference on Computational Complexity, CCC 2004.* (2004), pp. 123 – 129.

[25] Madelaine, F., Martin, B., and Stacho, J. Constraint Satisfaction with Counting Quantifiers In *Computer Science - Theory and Applications - 7th International Computer Science Symposium in Russia, CSR 2012* (2012), 253-265.

[26] Martin, B., and Stacho, J. Constraint Satisfaction with Counting Quantifiers 2 In *Computer Science - Theory and Applications - 9th International Computer Science Symposium in Russia, CSR 2014* (2014).

[27] Martin, B. First order model checking problems parameterized by the model. In *4th Conf. on Computatibility in Europe, LNCS 5028* (2008), pp. 417–427.

[28] Martin, B. QCSP on partially reflexive forests. In *Principles and Practice of Constraint Programming - 17th International Conference, CP 2011* (2011).

[29] Martin, B., and Madelaine, F. Towards a trichotomy for quantified H-coloring. In *2nd Conf. on Computatibility in Europe, LNCS 3988* (2006), pp. 342–352.

[30] Martin, B., and Madelaine, F. QCSP on partially reflexive cycles – the wavy line of tractability. In *Computer Science - Theory and Applications - 9th International Computer Science Symposium in Russia, CSR 2013* (2013).

[31] Otto, M. *Bounded variable logics and counting – A study in finite models*, vol. 9. Springer-Verlag, 1997. IX+183 pages.

[32] Papadimitriou, C. H. *Computational Complexity.* Addison-Wesley, 1994.

[33] Reingold, O. Undirected connectivity in log-space. *Journal of the ACM 55*, 4 (2008), 1–24.

[34] Schaefer, T. J. The complexity of satisfiability problems. In *Proceedings of the 10th Annual ACM Symposium on Theory of Computing (STOC)* (1978), pp. 216–226.