SIAM J. COMPUT.
Vol. 48, No. 4, pp. 1224–1264

CONSTRAINT SATISFACTION PROBLEMS FOR REDUCTS OF HOMOGENEOUS GRAPHS*

MANUEL BODIRSKY†, BARNABY MARTIN‡, MICHAEL PINSKER§, AND ANDRÁS PONGRÁCZ¶

Abstract. For $n \geq 3$, let (H_n, E) denote the nth Henson graph, i.e., the unique countable homogeneous graph with exactly those finite graphs as induced subgraphs that do not embed the complete graph on n vertices. We show that for all structures Γ with domain H_n whose relations are first-order definable in (H_n, E) the constraint satisfaction problem for Γ either is in P or is NP-complete. We moreover show a similar complexity dichotomy for all structures whose relations are first-order definable in a homogeneous graph whose reflexive closure is an equivalence relation. Together with earlier results, in particular for the random graph, this completes the complexity classification of constraint satisfaction problems of structures first-order definable in countably infinite homogeneous graphs: all such problems are either in P or NP-complete.

Key words. constraint satisfaction problems, homogeneous structures, first-order reducts, universal algebra, structural Ramsey theory, computational complexity

AMS subject classifications. 68W05, 68R05, 05C55

DOI. 10.1137/16M1082974

1. Introduction.

1.1. Constraint satisfaction problems. A *constraint satisfaction problem* (CSP) is a computational problem in which the input consists of a finite set of variables and a finite set of *constraints*, and where the question is whether there exists a mapping from the variables to some fixed domain such that all the constraints are satisfied. We can thus see the possible constraints as relations on that fixed domain, and in an instance of the CSP, we are asked to assign domain values to the variables such that certain specified tuples of variables become elements of certain specified relations.

When the domain is finite, and arbitrary constraints are permitted, the CSP is NP-complete. However, when only constraints from a restricted set of relations on the

*Received by the editors July 5, 2016; accepted for publication (in revised form) May 2, 2019; published electronically July 18, 2019. An extended abstract of this paper has appeared at the 43rd International Colloquium on Automata, Languages and Programming (ICALP) Track B, 2016.
https://doi.org/10.1137/16M1082974

Funding: The first and fourth authors have received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013 grant agreement 257039). The first author also received funding from the German Science Foundation (DFG, project 622397) and from the European Research Council (grant agreement 681988, CSP-Infinity). The third author has received funding from the Austrian Science Fund (FWF) through project P27600, and from the Czech Science Foundation (grant 13-01832S). The fourth author was supported by the Hungarian Scientific Research Fund (OTKA) grant K109185, by the National Research, Development and Innovation Fund of Hungary, financed under the FK 124814 and PD 125160 funding schemes, the János Bolyai Research Scholarship of the Hungarian Academy of Sciences, the New National Excellence Program Bolyai+ of the Ministry of Human Capacities, and by the European Social Fund (EFOP-3.6.2-16-2017-00015).

†Institut für Algebra, TU Dresden, 01062 Dresden, Germany (Manuel.Bodirsky@tu-dresden.de).
‡Department of Computer Science, Durham University, South Road, Durham, DH13LE, UK (barnabymartin@gmail.com).
§Institut für Diskrete Mathematik und Geometrie, FG Algebra, TU Wien, Wien, 1040, Austria, and Department of Algebra, Charles University, Czech Republic (marula@gmx.at).
¶Department of Algebra and Number Theory, University of Debrecen, 4032 Debrecen, Egyetem square 1, Hungary (pongracz.andras@science.unideb.hu).

domain are allowed in the input, there might be a polynomial-time algorithm for the CSP. The set of relations that is allowed to formulate the constraints in the input is often called the *constraint language*. The question of which constraint languages give rise to polynomial-time solvable CSPs has been the topic of intensive research over the years. It was conjectured by Feder and Vardi [FV98] that CSPs for constraint languages over finite domains have a complexity dichotomy: they are either in P or NP-complete. Over the years, the conjecture was proved for substantial classes (for example, when the domain has at most three elements [Sch78, Bul06] or when the constraint language contains a single binary relation without sources and sinks [HN90, BKN09]). Various methods, combinatorial (graph-theoretic), logical, and universal-algebraic, were brought to bear on this classification project, with many remarkable consequences. A conjectured delineation for the dichotomy was given in the algebraic language in [BKJ05], and finally the conjecture, and in particular this delineation, has recently been proven to be accurate [Bul17, Zhu17].

When the domain is infinite, the complexity of the CSP can be outside NP, and even undecidable [BN06]. But for natural classes of such CSPs there is often the potential for structured classifications, and this has proved to be the case for structures first-order definable over the order $(\mathbb{Q}, <)$ of the rationals [BK09] or over the integers with successor [BMM18]. Another classification of this type has been obtained for CSPs where the constraint language is first-order definable over the random (Rado) graph [BP15a], making use of structural Ramsey theory. This paper was titled "Schaefer's theorem for graphs" and it can be seen as lifting the famous classification of Schaefer [Sch78] from Boolean logic to logic over finite graphs, since the random graph is universal for the class of finite graphs.

1.2. Homogeneous graphs and their reducts. The notion of *homogeneity* from model theory plays an important role when applying techniques from finite-domain constraint satisfaction to constraint satisfaction over infinite domains. A relational structure is *homogeneous* if every isomorphism between finite induced substructures can be extended to an automorphism of the entire structure. Homogeneous structures are uniquely (up to isomorphism) given by the class of finite structures that embed into them. The structure $(\mathbb{Q}, <)$ and the random graph are among the most prominent examples of homogeneous structures. The class of structures that are first-order definable over a homogeneous structure with finite relational signature is a very large generalization of the class of all finite structures, and CSPs for those structures have been studied independently in many different areas of theoretical computer science, e.g., in temporal and spatial reasoning, phylogenetic analysis, computational linguistics, scheduling, graph homomorphisms, and many more; see [Bod12] for references.

While homogeneous relational structures are abundant, there are remarkably few countably infinite homogeneous (undirected, irreflexive) *graphs*; they have been classified by Lachlan and Woodrow [LW80]. Besides the random graph mentioned earlier, an example of such a graph is the countable homogeneous *universal triangle-free* graph, one of the fundamental structures that appears in most textbooks in model theory. This graph is the up to isomorphism unique countable triangle-free graph (H_3, E) with the property that for every finite independent set $X \subseteq H_3$ and for every finite set $Y \subseteq H_3$ there exists a vertex $x \in H_3 \setminus (X \cup Y)$ such that x is adjacent to every vertex in X and to no vertex in Y.

Further examples of homogeneous graphs are the graphs (H_4, E), (H_5, E), and so forth, which together with (H_3, E) are called the *Henson graphs*, and their complements. Here, (H_n, E) for $n > 3$ is the generalization of the graph (H_3, E) above from triangles to cliques of size n. Finally, the list of Lachlan and Woodrow contains only one more family of infinite graphs, namely the graphs (C_n^s, E) whose reflexive closure Eq is an equivalence relation with n classes of equal size s, where $1 \leq n, s \leq \omega$, and either n or s equals ω, as well as their complements. We remark that (C_n^s, Eq) is itself homogeneous and first-order interdefinable with (C_n^s, E), and so we shall sometimes refer to the *homogeneous equivalence relations*.

All countable homogeneous graphs, and even all structures which are first-order definable over homogeneous graphs, are *ω-categorical*; that is, all countable models of their first-order theory are isomorphic. Moreover, all countably infinite homogeneous graphs Γ are *finitely bounded* in the sense that the *age* of Γ, i.e., the class of finite structures that embed into Γ, can be described by finitely many forbidden substructures. Finitely bounded homogeneous structures also share with finite structures the property of having a finite description: up to isomorphism, they are uniquely given by the finite list of forbidden structures that describes their age. Recent work indicates the importance of finite boundedness for complexity classification [BPT13, BP11, BM16, BKO$^+$17], and it has been conjectured that all structures with a first-order definition in a finitely bounded homogeneous structure enjoy a complexity dichotomy, i.e., their CSP is either in P or NP-complete (cf. [BPP14, BP16a, BKO$^+$17]). The structures that are first-order definable in homogeneous graphs therefore provide the most natural class on which to further test the methods developed in [BP15a] specifically for the random graph.

In this article we obtain a complete classification of the computational complexity of CSPs where all constraints have a first-order definition in one of the Henson graphs. We moreover obtain such a classification for CSPs where all constraints have a first-order definition in a countably infinite homogeneous graph whose reflexive closure is an equivalence relation, expanding earlier results for the special cases of one single equivalence class (so-called *equality constraints* [BK08]) and infinitely many infinite classes [BW12]. Together with the above-mentioned result on the random graph, this completes the classification of CSPs for constraints with a first-order definition in any countably infinite homogeneous graph, by Lachlan and Woodrow's classification. Our result is in accordance with the delineations between tractability and hardness predicted in general for structures with a first-order definition in a finitely bounded homogeneous structure [BPP14, BP16a, BKO$^+$17].

Following an established convention (e.g., [Tho91, BP11] and many more) we call a relational structure Γ a *reduct* of a structure Δ if it has the same domain as Δ and all relations of Γ are first-order definable without parameters in Δ. That is, for us a reduct of Δ is similar to the classical definition of a reduct with the difference that we first allow a first-order expansion of Δ. With this terminology, the present article provides a complexity classification of the CSPs for all reducts of countably infinite homogeneous graphs. In other words, for every such reduct we determine the complexity of deciding its *primitive positive theory*, which consists of all sentences which are existentially quantified conjunctions of atomic formulas and which hold in the reduct. We remark that all reducts of such graphs can be defined by quantifier-free first-order formulas, by homogeneity and ω-categoricity.

For reducts of (H_n, E), the CSPs express computational problems where the task is to decide whether there exists a finite graph without any clique of size n that meets certain constraints. An example of a reduct whose CSP can be solved in polynomial time is $(H_n, E, \{(x, y, u, v) \mid E(x, y) \Rightarrow E(u, v)\})$, where $n \geq 3$ is arbitrary. As it turns out, for every CSP of a reduct of a Henson graph which is solvable in polynomial time, the corresponding reduct over the random graph, i.e., the reduct whose relations are defined by the same quantifier-free formulas, is also polynomial-time solvable. On the other hand, the CSP of the reduct $(H_n, \{(x,y,u,v) \mid E(x,y) \vee E(u,v)\})$ is NP-complete for all $n \geq 3$, but the corresponding reduct over the random graph can be decided in polynomial time.

Similarly, for reducts of the graph (C_n^s, E) whose reflexive closure is an equivalence relation with n classes of size s, where $1 \leq n, s \leq \omega$, the computational problem is to decide whether there exists an equivalence relation with n classes of size s that meets certain constraints. For example, consider the structure $(C_\omega^2; Eq, A)$ where

$$A := \big\{(x_1,y_1,x_2,y_2,x_3,y_3) \mid \text{if } Eq(x_1,y_1), Eq(x_2,y_2), \text{ and } Eq(x_3,y_3), \text{ then there is an odd number of } i \in \{1,2,3\} \text{ such that } x_i \neq y_i\big\}.$$

This structure is a reduct of $(C_\omega^2; E)$, and it follows from our results in section 7.2 that its CSP can be solved in polynomial time.

1.3. Results. Our first result is the complexity classification of the CSPs of all reducts of Henson graphs, showing in particular that a uniform approach to infinitely many "base structures" in the same language (namely, the nth Henson graph for each $n \geq 3$) is, in principle, possible.

THEOREM 1.1. *Let $n \geq 3$, and let Γ be a finite signature reduct of the nth Henson graph (H_n, E). Then* CSP(Γ) *is either in* P *or* NP*-complete.*

We then obtain a similar complexity dichotomy for reducts of homogeneous equivalence relations, expanding earlier results for special cases [BW12, BK08].

THEOREM 1.2. *Let (C_n^s, E) be a graph whose reflexive closure Eq is an equivalence relation with n classes of size s, where $1 \leq n, s \leq \omega$, and either s or n is ω. Then for any finite signature reduct Γ of (C_n^s, E), the problem* CSP(Γ) *is either in* P *or* NP*-complete.*

Together with the classification of countable homogeneous graphs, and the fact that the complexity of the CSPs of the reducts of the random graph have been classified [BP15a], this completes the CSP classification of reducts of all countably infinite homogeneous graphs, confirming further instances of the open conjecture that CSPs of reducts of finitely bounded homogeneous structures are either in P or NP-complete [BPP14, BP16a, BKO+17].

COROLLARY 1.3. *Let Γ be a finite signature reduct of a countably infinite homogeneous graph. Then* CSP(Γ) *is either in* P *or* NP*-complete.*

We are going to provide more detailed versions of Theorems 1.1 and 1.2, which describe in particular the delineation between the tractable and the NP-complete cases algebraically, in sections 5 and 8. We would like to emphasize that our proof does not assume or use the dichotomy for CSPs of finite structures, as opposed to some other dichotomy results for CSPs of infinite structures such as [BMM18].

1.4. The strategy. The method we employ follows broadly the method invented in [BP15a] for the corresponding classification problem where the "base structure" is the random graph. The key component of this method is the usage of Ramsey theory (in our case, a result of Nešetřil and Rödl [NR89]) and the concept of *canonical functions* introduced in [BP14]. There are, however, some interesting differences and novelties that appear in the present proof, as we now briefly outline.

1.4.1. Henson graphs. When studying the proofs in [BP15a], one might get the impression that the complexity of the method grows with the model-theoretic complexity of the base structure, and that for the random graph we have really reached the limits of bearableness for applying the Ramsey method.

However, quite surprisingly, when we step from the random graph to the graphs (H_n, E), which are in a sense more complicated structures from a model-theoretic point of view,[1] the classification and its proof become easier again. It is one of the contributions of the present article to explain the reasons behind this effect. Essentially, certain *behaviors* of canonical functions (cf. section 2) existing on the random graph cannot be realized in (H_n, E). For example, the behavior "max" (cf. section 2) plays no role for the present classification but accounts over the random graph for the tractability of, inter alia, the 4-ary relation defined by the formula $E(x, y) \vee E(u, v)$.

Remarkably, we are able to reuse results about canonical functions over the random graph, since the calculus for composing behaviors of canonical functions is the same for any other structure with a smaller type space, and in particular the Henson graphs. Via this meta-argument we can, on numerous occasions, make statements about canonical functions over the Henson graphs which were proven earlier for the random graph, ignoring completely the actual underlying structure; even more comfortably, we can a posteriori rule out some possibilities in those statements because of the K_n-freeness of the Henson graphs. Instances of this phenomenon appear in the analysis of canonical functions in section 3.11.

On the other hand, along with these simplifications, there are also new additional difficulties that appear when investigating reducts of (H_n, E) and that were not present in the classification of reducts of the random graph, which basically stem from the lower degree of symmetry of (H_n, E) compared to the random graph. For example, in expansions of Henson graphs by finitely many constants, not all orbits induce copies of Henson graphs; the fact that the analogous statement does hold for the random graph was used extensively in [BP15a], for example, in the rather technical proof of Proposition 7.18 of that paper.

1.4.2. Equivalence relations. Similarly to the situation for the equivalence relation with infinitely many infinite classes studied in [BW12], there are two interesting sources of NP-hardness for the reducts Γ of other homogeneous equivalence relations: namely, if the equivalence relation is invariant under the polymorphisms of Γ, then the structure obtained from Γ by factoring by the equivalence relation might have an NP-hard CSP, implying NP-hardness for the CSP of Γ itself; or, roughly, for a fixed equivalence class the restriction of Γ to that class might have an NP-hard CSP, again implying NP-hardness of the CSP of Γ (assuming that Γ is a *model-complete core*; see sections 3 and 6). But whereas for the equivalence relation with infinitely many infinite classes both the factor structure and the restriction to a class are again infinite structures, for the other homogeneous equivalence relations one of the two is a finite structure. This obliges us to combine results about CSPs of finite structures with those of infinite structures. As it turns out, the two-element case is, not surprisingly, different from the other finite cases and, quite surprisingly, significantly more involved than the other cases. One particularity of this case is that tractability is, for some reducts, implied by a ternary *noninjective* canonical function which we obtain by our Ramsey analysis. Among all the classification results for ω-categorical structures

[1]For example, the random graph has a *simple* theory [TZ12], whereas the Henson graphs are among the most basic examples of structures whose theory is *not* simple.

obtained so far, this ternary function is the first example of a noninjective canonical function leading to a maximal tractable class. The occurrence of this phenomenon is of technical interest in the quest for a proof of the CSP dichotomy conjecture for reducts of finitely bounded homogeneous structures via a reduction to the finite CSP dichotomy.

1.5. Overview. We organize the remainder of this article as follows. Basic notions and definitions, as well as the fundamental facts of the method we are going to use, are provided in section 2.

Sections 3 to 5 deal with the Henson graphs: Section 3 is complexity-free and investigates the structure of reducts of Henson graphs via polymorphisms and Ramsey theory. In section 4, we provide hardness and tractability proofs for different classes of reducts. Section 5 contains the proof of Theorem 1.1, and we discuss the complexity classification in more detail, formulating in particular a tractability criterion for CSPs of reducts of Henson graphs.

We then turn to homogeneous equivalence relations in sections 6 to 8. Similarly to the Henson graphs, the first section (section 6) is complexity-free and investigates the structure of reducts of homogeneous equivalence relations via polymorphisms and Ramsey theory. Section 7 contains the algorithms proving tractability where it applies. Finally, section 8 provides the proof of Theorem 1.2 and describes in detail the delineation between the tractable and the NP-complete cases.

We finish this work with further research directions in section 9.

2. Preliminaries.

2.1. General notational conventions. We use one single symbol, namely E, for the edge relation of all homogeneous graphs; since we never consider several such graphs at the same time, this should not cause confusion. Moreover, we use E for the symbol representing the relation E, for example in logical formulas. In general, we shall not distinguish between relation symbols and the relations which they denote. The binary relation $N(x,y)$ is defined by the formula $\neg E(x,y) \wedge x \neq y$.

When E is the edge relation of a homogeneous graph whose reflexive closure is an equivalence relation, we denote this equivalence relation by Eq; so $Eq(x,y)$ is defined by the formula $E(x,y) \vee x = y$.

When t is an n-tuple, we refer to its entries by $t_1, \ldots, t_n$. When $f\colon A \to B$ is a function and $C \subseteq A$, we write $f[C] := \{f(a) \mid a \in C\}$.

2.2. Henson graphs. For $n \geq 2$, denote the clique on n vertices by K_n. For $n \geq 3$, the graph (H_n, E) is the up to isomorphism unique countable graph which is

- *homogeneous*: any isomorphism between two finite induced subgraphs of (H_n, E) can be extended to an automorphism of (H_n, E), and
- *universal for the class of K_n-free graphs*: (H_n, E) contains all finite (in fact, all countable) K_n-free graphs as induced subgraphs.

The graph (H_n, E) has the *extension property*: for all disjoint finite $U, U' \subseteq H_n$ such that U is not inducing any isomorphic copy of K_{n-1} in (H_n, E), there exists $v \in H_n$ such that v is adjacent in (H_n, E) to all members of U and to none in U'. Up to isomorphism, there exists a unique countably infinite K_n-free graph with this extension property, and hence the property can be used as an alternative definition of (H_n, E).

2.3. Homogeneous equivalence relations. For $1 \leq n, s \leq \omega$ the graph (C_n^s, E) is the up to isomorphism unique countable graph whose reflexive closure is an equivalence relation Eq with n classes C_i, where $0 \leq i < n$, all of which have size s. Clearly, (C_n^s, E) is homogeneous and universal in a sense similar to that above.

2.4. Constraint satisfaction problems. For a relational signature τ, a first-order τ-formula is called *primitive positive* (pp) if it is of the form

$$\exists x_1, \ldots, x_n \, (\psi_1 \wedge \cdots \wedge \psi_m)$$

where the ψ_i are *atomic*, i.e., of the form $y_1 = y_2$ or $R(y_1, \ldots, y_k)$ for a k-ary relation symbol $R \in \tau$ and not necessarily distinct variables y_i.

Let Γ be a structure with a finite relational signature τ. The *constraint satisfaction problem for* Γ, denoted by $\mathrm{CSP}(\Gamma)$, is the computational problem of deciding for a given pp τ-sentence ϕ whether ϕ is true in Γ. The following lemma was first stated in [JCG97] for finite domain structures Γ only, but the proof there also works for arbitrary infinite structures.

LEMMA 2.1. *Let $\Gamma = (D, R_1, \ldots, R_\ell)$ be a relational structure, and let R be a relation that has a pp definition in Γ, i.e., a definition via a pp formula. Then* $\mathrm{CSP}(\Gamma)$ *and* $\mathrm{CSP}(D, R, R_1, \ldots, R_\ell)$ *are polynomial-time equivalent.*

When a relation R has a pp definition in a structure Γ, then we also say that Γ *pp-defines* R. Lemma 2.1 enables the so-called *universal-algebraic approach* to constraint satisfaction, as exposed in the following.

2.5. The universal-algebraic approach. We say that a k-ary function (also called *operation*) $f\colon D^k \to D$ *preserves* an m-ary relation $R \subseteq D^m$ if for all $t_1, \ldots, t_k \in R$ the tuple $f(t_1, \ldots, t_k)$, calculated componentwise, is also contained in R. If an operation f does not preserve a relation R, we say that f *violates* R. We say that a set of operations preserves a relation when all of its elements do.

If f preserves all relations of a structure Γ, we say that f is a *polymorphism* of Γ and that f *preserves* Γ. We write $\mathrm{Pol}(\Gamma)$ for the set of all polymorphisms of Γ. The unary polymorphisms of Γ are just the *endomorphisms* of Γ and are denoted by $\mathrm{End}(\Gamma)$.

The set of all polymorphisms $\mathrm{Pol}(\Gamma)$ of a relational structure Γ forms an algebraic object called a *function clone* (see [Sze86], [GP08]), which is a set of finitary operations defined on a fixed domain that is closed under composition and that contains all projections. Moreover, $\mathrm{Pol}(\Gamma)$ is closed in the *topology of pointwise convergence*; i.e., an n-ary function f is contained in $\mathrm{Pol}(\Gamma)$ if and only if for all finite subsets A of Γ^n there exists an n-ary $g \in \mathrm{Pol}(\Gamma)$ which agrees with f on A. We will write $\overline{F}$ for the closure of a set F of functions on a fixed domain in this topology; so $\overline{\mathrm{Pol}(\Gamma)} = \mathrm{Pol}(\Gamma)$. This closure is sometimes referred to as *local closure*, and closed sets as locally closed, but we will use the terminology *topologically closed* throughout this work. For an arbitrary set F of functions on a fixed domain, when Γ is the structure whose relations are precisely those which are preserved by all functions in F, then $\mathrm{Pol}(\Gamma)$ is the smallest topologically closed function clone containing F (cf. [Sze86]).

When Γ is a countable and ω-categorical structure, we can characterize pp-definable relations via $\mathrm{Pol}(\Gamma)$, as follows.

THEOREM 2.2 (from [BN06]). *Let Γ be a countable ω-categorical structure. Then the relations preserved by* $\mathrm{Pol}(\Gamma)$ *are precisely those having a pp definition in* Γ.

Theorem 2.2 and Lemma 2.1 imply that if two countable ω-categorical structures Γ, Δ with finite relational signatures have the same clone of polymorphisms, then their

CSPs are polynomial-time equivalent. Moreover, if $\mathrm{Pol}(\Gamma)$ is contained in $\mathrm{Pol}(\Delta)$, then $\mathrm{CSP}(\Gamma)$ is, up to polynomial time, at least as hard as $\mathrm{CSP}(\Delta)$.

Note that the *automorphisms* of a structure Γ are just the bijective unary polymorphisms of Γ whose inverse function is also a polymorphism of Γ; the set of all automorphisms of Γ is denoted by $\mathrm{Aut}(\Gamma)$. For every reduct Γ of a structure Δ we have that $\mathrm{Pol}(\Gamma) \supseteq \mathrm{Aut}(\Gamma) \supseteq \mathrm{Aut}(\Delta)$. In particular, this is the case for reducts of the homogeneous graphs (H_n, E) and (C_n^s, E). Conversely, it follows from the ω-categoricity of homogeneous graphs (D, E) (in our case, $D = H_n$ or $D = C_n^s$) that every topologically closed function clone containing $\mathrm{Aut}(D, E)$ is the polymorphism clone of a reduct of (D, E).

When (D, E) is a homogeneous graph, and F is a set of functions, and g is a function on the domain D, we say that F *generates* g if g is contained in the smallest topologically closed function clone which contains $F \cup \mathrm{Aut}(D, E)$. This is the same as saying that for every finite $S \subseteq D$, there exists a term function over $F \cup \mathrm{Aut}(D, E)$ which agrees with g on S. By the discussion preceding Theorem 2.2, this is equivalent to g preserving all relations which are preserved by $F \cup \mathrm{Aut}(D, E)$.

We finish this section with a general lemma that we will refer to on numerous occasions; it allows us to restrict the arity of functions violating a relation. For a structure Γ and a tuple $t \in \Gamma^k$, the *orbit of* t in Γ is the set $\{\alpha(t) \mid \alpha \in \mathrm{Aut}(\Gamma)\}$. We also call this the orbit of t with respect to $\mathrm{Aut}(\Gamma)$.

LEMMA 2.3 (from [BK09]). *Let Γ be a relational structure. Suppose that $R \subseteq \Gamma^k$ intersects at most m orbits of k-tuples in Γ. If $\mathrm{Pol}(\Gamma)$ contains a function violating R, then $\mathrm{Pol}(\Gamma)$ also contains an m-ary operation violating R.*

2.6. Canonical functions. It will turn out that the polymorphisms relevant for the CSP classification show regular behavior with respect to the underlying homogeneous graph in a sense that we are now going to define.

DEFINITION 2.4. *Let Δ be a structure. The* type $\mathrm{tp}(a)$ *of an n-tuple $a = (a_1, \ldots, a_n)$ of elements in Δ is the set of first-order formulas with free variables $x_1, \ldots, x_n$ that hold for a in Δ. For structures $\Delta_1, \ldots, \Delta_k$ and k-tuples $a^1, \ldots, a^n \in \Delta_1 \times \cdots \times \Delta_k$, the type of $(a^1, \ldots, a^n)$ in $\Delta_1 \times \cdots \times \Delta_k$, denoted by $\mathrm{tp}(a^1, \ldots, a^n)$, is the k-tuple containing the types of $(a_i^1, \ldots, a_i^n)$ in Δ_i for each $1 \le i \le k$.*

We bring to the reader's attention the well-known fact that in ω-categorical structures, in particular in (H_n, E) and (C_n^k, E), two n-tuples have the same type if and only if their orbits coincide.

DEFINITION 2.5. *Let $\Delta_1, \ldots, \Delta_k$ and Λ be structures. A* behavior *B between $\Delta_1, \ldots, \Delta_k$ and Λ is a partial function from the types over $\Delta_1, \ldots, \Delta_k$ to the types over Λ. Pairs (s, t) with $B(s) = t$ are also called* type conditions. *We say that a function $f\colon \Delta_1 \times \cdots \times \Delta_k \to \Lambda$* satisfies the behavior *B if whenever $B(s) = t$ and $(a^1, \ldots, a^n)$ has type s in $\Delta_1 \times \cdots \times \Delta_k$ the n-tuple $(f(a_1^1, \ldots, a_k^1), \ldots, f(a_1^n, \ldots, a_k^n))$ has type t in Λ. A function $f\colon \Delta_1 \times \cdots \times \Delta_k \to \Lambda$ is* canonical *if it satisfies a behavior which is a total function from the types over $\Delta_1 \times \cdots \times \Delta_k$ to the types over Λ.*

We remark that since our structures are homogeneous and have only binary relations, the type of an n-tuple a is determined by its binary subtypes, i.e., the types of the pairs (a_i, a_j), where $1 \le i, j \le n$. In other words, the type of a is determined by which of its components are equal, and between which of its components there is an edge. Therefore, a function $f\colon (H_n, E)^k \to (H_n, E)$ or $f\colon (C_n^s, E)^k \to (C_n^s, E)$ is canonical if and only if it satisfies the condition of the definition for types of 2-tuples.

To provide immediate examples for these notions, we now define some behaviors that will appear in our proof as well as in the precise CSP classification. For m-ary relations $R_1, \dots, R_k$ over a set D, we will in the following write $R_1 \cdots R_k$ for the m-ary relation on D^k defined as follows: $R_1 \cdots R_k(x^1, \dots, x^m)$ holds for k-tuples $x^1, \dots, x^m \in D^k$ if and only if $R_i(x_i^1, \dots, x_i^m)$ holds for all $1 \leq i \leq k$. For example, when $p, q \in D^3$ are triples of elements in a homogeneous graph (D, E), then $EN{=}(p, q)$ holds if and only if $E(p_1, q_1)$, $N(p_2, q_2)$, and $p_3 = q_3$ hold in (D, E). We start with behaviors of binary injective functions f on homogeneous graphs.

Definition 2.6. *Let (D, E) be a homogeneous graph. We say that a binary injective operation $f \colon D^2 \to D$ is*

- balanced in the first argument *if for all $u, v \in D^2$ we have that $E{=}(u, v)$ implies $E(f(u), f(v))$ and $N{=}(u, v)$ implies $N(f(u), f(v))$;*
- balanced in the second argument *if $(x, y) \mapsto f(y, x)$ is balanced in the first argument;*
- balanced *if f is balanced in both arguments;*
- E-dominated (N-dominated) in the first argument *if for all $u, v \in D^2$ with $\neq{=}(u, v)$ we have that $E(f(u), f(v))$ $(N(f(u), f(v)))$;*
- E-dominated (N-dominated) in the second argument *if $(x, y) \mapsto f(y, x)$ is E-dominated (N-dominated) in the first argument;*
- E-dominated (N-dominated) *if it is E-dominated (N-dominated) in both arguments;*
- of behavior min *if for all $u, v \in D^2$ with $\neq\neq(u, v)$ we have $E(f(u), f(v))$ if and only if $EE(u, v)$;*
- of behavior max *if for all $u, v \in D^2$ with $\neq\neq(u, v)$ we have $N(f(u), f(v))$ if and only if $NN(u, v)$;*
- of behavior p_1 *if for all $u, v \in D^2$ with $\neq\neq(u, v)$ we have $E(f(u), f(v))$ if and only if $E(u_1, v_1)$;*
- of behavior p_2 *if $(x, y) \mapsto f(y, x)$ is of behavior p_1;*
- of behavior projection *if it is of behavior p_1 or p_2;*
- of behavior xnor if for all $u, v \in D^2$ with $\neq\neq(u, v)$ we have $E(f(u), f(v))$ if and only if $EE(u, v)$ or $NN(u, v)$.

Each of these properties describes the set of all functions of a certain behavior. We explain this for the first item defining functions which are balanced in the first argument, which can be expressed by the behavior consisting of the following two type conditions. Let (u, v) be any pair of elements $u, v \in D^2$ such that $E{=}(u, v)$, and let s be the type of the pair (u, v) in $(D, E) \times (D, E)$. Let $x, y \in D$ satisfy $E(x, y)$, and let t be the type of (x, y) in (D, E). Then the first type condition is (s, t). Now let s' be the type in $(D, E) \times (D, E)$ of any pair (u, v), where $u, v \in D^2$ satisfy $N{=}(u, v)$, and let t' be the type in (D, E) of any $x, y \in D$ with $N(x, y)$. The second type condition is (s', t').

To justify the less obvious names of some of the above behaviors, we would like to point out that a binary injection of behavior min is reminiscent of the Boolean minimum function on $\{0, 1\}$, where E takes the role of 1 and N the role of 0: for $u, v \in H_n^2$ with $\neq\neq(u, v)$, we have $E(f(u), f(v))$ if u, v are connected by an edge in both coordinates, and $N(f(u), f(v))$ otherwise. The names "max" and "projection" can be explained similarly.

Definition 2.7. *Let (D, E) be a homogeneous graph. We say that a ternary injective operation $f \colon D^3 \to D$ is of behavior*

- majority *if for all $u, v \in D^3$ with $\neq\neq\neq(u, v)$ we have that $E(f(u), f(v))$ if*

and only if $EEE(u,v)$, $EEN(u,v)$, $ENE(u,v)$, *or* $NEE(u,v)$*;*
- minority *if for all* $u,v \in D^3$ *with* $\neq\neq\neq(u,v)$ *we have* $E(f(u),f(v))$ *if and only if* $EEE(u,v)$, $NNE(u,v)$, $NEN(u,v)$, *or* $ENN(u,v)$.

In this article, contrary to min and minority, neither max nor majority will play a role, but we introduce them for the sake of completeness since they occur in [BP15a].

When we want to explain a type condition over a homogeneous graph (D,E), we are going to express it in the form $f(R_1,\dots,R_k) = S$ for binary relations $R_1,\dots,R_k$ and a binary relation S; the meaning is that whenever $p,q \in D^k$, then $R_1 \cdots R_k(p,q)$ implies $S(f(p),f(q))$. The relations we use in this notation range among $\{E,N,Eq,\neq,=\}$. Examples of type conditions expressed this way include $f(E,N) = N$ (meaning that $EN(p,q)$ implies $N(f(p),f(q))$ for all $p,q \in D^2$) and $f(E,=) = E$. In the latter, note that the second $=$ has different semantic content from the first. Similarly, the majority behavior in Definition 2.7 can be expressed by writing $f(E,E,E) = f(E,E,N) = f(E,N,E) = f(N,E,E) = E$ and $f(N,N,N) = f(E,N,N) = f(N,E,N) = f(N,N,E) = N$. As another example, note that E-dominated in the first argument can be expressed as $f(\neq,=) = E$, or equivalently, as the conjunction of $f(E,=) = E$ and $f(N,=) = E$. Our notation is justified by the fact that the type conditions satisfied by a function induce a partial function from types to types, and that in the case of homogeneous graphs, all that matters is the three types of pairs, given by the relations E, N, and $=$; the relation $\neq$ is the union of E and N and is used as a shortcut.

Definition 2.8. *Let* (D,E) *be a homogeneous graph. We say a ternary canonical injection* $f\colon D^3 \to D$ *is* hyperplanely of behavior projection *if the functions* $(u,v) \mapsto f(c,u,v)$, $(u,v) \mapsto f(u,c,v)$, *and* $(u,v) \mapsto f(u,v,c)$ *are of behavior projection for all* $c \in D$. *Other hyperplane behaviors, such as hyperplanely E-dominated, are defined similarly.*

Note that hyperplane behaviors are defined by conditions for the type functions $f(=,\cdot,\cdot)$, $f(\cdot,=,\cdot)$, and $f(\cdot,\cdot,=)$. For example, hyperplanely E-dominated precisely means that

$$f(=,=,\neq) = f(=,\neq,=) = f(\neq,=,=) = E\,.$$

2.7. Achieving canonicity in Ramsey structures. The next proposition, which is an instance of more general statements from [BP11, BPT13], provides us with the main combinatorial tool for analyzing functions on Henson graphs. Equip H_n with a total order $\prec$ in such a way that $(H_n,E,\prec)$ is homogeneous; up to isomorphism, there is only one such structure $(H_n,E,\prec)$, called the *random ordered K_n-free graph*. The order $(H_n,\prec)$ is then isomorphic to the order $(\mathbb{Q},<)$ of the rationals. By [NR89], $(H_n,E,\prec)$ is a *Ramsey structure*, which implies the following proposition—for more details, see the survey [BP11].

Proposition 2.9. *Let* $f\colon H_n^k \to H_n$, *let* $c_1,\dots,c_r \in H_n$, *and let* $(H_n,E,\prec,c_1,\dots,c_r)$ *be the expansion of* $(H_n,E,\prec)$ *by the constants* $c_1,\dots,c_r$. *Then*

$$\overline{\{\alpha \circ f \circ (\beta_1,\dots,\beta_r) \mid \alpha \in \mathrm{Aut}(H_n,E,\prec),\ \beta_1,\dots,\beta_r \in \mathrm{Aut}(H_n,E,\prec,c_1,\dots,c_r)\}}$$

contains a function g such that
- *g is canonical as a function from* $(H_n,E,\prec,c_1,\dots,c_r)$ *to* $(H_n,E,\prec)$*;*
- *g agrees with f on* $\{c_1,\dots,c_r\}^k$.

In particular, f generates a function g with these properties.

Similarly, Ramsey theory allows us to produce canonical functions on (C_n^s,E), expanded with a certain linear order. Equip C_n^s with a total order $\prec$ so that the equivalence classes of (C^s_n, Eq) are *convex* with respect to $\prec$; i.e., whenever $Eq(u, v)$ holds and $u \prec w \prec v$, then $Eq(u, w)$. Moreover, in the case where the size of the classes $s = \omega$, we require the order $\prec$ to be isomorphic to the order of the rational numbers on each equivalence class, and in the case where the number of classes $n = \omega$, we require the order to be isomorphic to the order of the rational numbers between the classes (note that we already required convexity, so that $\prec$ naturally induces a linear order between the classes).

If the number of classes n is finite and their size $s = \omega$ infinite, let $P_1, \ldots, P_n$ denote unary predicates such that P_i contains precisely the elements in the ith equivalence class of Eq with respect to the order on the classes induced by $\prec$. The structure $(C^\omega_n, E, \prec, P_1, \ldots, P_n)$ is homogeneous and a Ramsey structure, since its automorphism group is, as a topological group, isomorphic to $\mathrm{Aut}(\mathbb{Q}; <)^n$, and since being a Ramsey structure is a property of the automorphism group (as a topological group) [KPT05]. Thus, by [BP11, BPT13], we have the following analogous statement to Proposition 2.9 for this structure. In the statement, we may drop the mention of the auxiliary relations $P_1, \ldots, P_n$, since these are first-order definable in $(C^s_n, E, \prec)$ and since the types over first-order interdefinable structures coincide; in other words, the relations were only needed temporarily in order to achieve homogeneity, required in [BP11, BPT13], but not for the Ramsey property.

PROPOSITION 2.10. *Let $n \geq 1$ be finite. Let $f\colon (C^\omega_n)^k \to C^\omega_n$, and let $c_1, \ldots, c_r \in C^\omega_n$. Then*

$$\overline{\{\alpha \circ f \circ (\beta_1, \ldots, \beta_r) \mid \alpha \in \mathrm{Aut}(C^\omega_n, E, \prec),\ \beta_1, \ldots, \beta_r \in \mathrm{Aut}(C^\omega_n, E, \prec, c_1, \ldots, c_r)\}}$$

contains a function g such that

- *g is canonical as a function from $(C^\omega_n, E, \prec, c_1, \ldots, c_r)$ to $(C^\omega_n, E, \prec)$;*
- *g agrees with f on $\{c_1, \ldots, c_r\}^k$.*

In particular, f generates a function g with these properties.

If the class size s is finite and their number $n = \omega$, we add s unary predicates $Q_1, \ldots, Q_s$ where Q_i contains precisely the ith element for each equivalence class with respect to the order $\prec$. Then $(C^s_\omega, E, \prec, Q_1, \ldots, Q_s)$ is homogeneous and Ramsey, since its automorphism group is isomorphic as a topological group to $\mathrm{Aut}(\mathbb{Q}; <)$, so that we obtain an analogue of Propositions 2.9 and 2.10 also in this case. Again, we may drop the relations $Q_1, \ldots, Q_n$, which are first-order definable in $(C^n_\omega, E, \prec)$, in the statement.

PROPOSITION 2.11. *Let $s \geq 1$ be finite. Let $f\colon (C^s_\omega)^k \to C^s_\omega$, and let $c_1, \ldots, c_r \in C^s_\omega$. Then*

$$\overline{\{\alpha \circ f \circ (\beta_1, \ldots, \beta_r) \mid \alpha \in \mathrm{Aut}(C^s_\omega, E, \prec),\ \beta_1, \ldots, \beta_r \in \mathrm{Aut}(C^s_\omega, E, \prec, c_1, \ldots, c_r)\}}$$

contains a function g such that

- *g is canonical as a function from $(C^s_\omega, E, \prec, c_1, \ldots, c_r)$ to $(C^s_\omega, E, \prec)$;*
- *g agrees with f on $\{c_1, \ldots, c_r\}^k$.*

In particular, f generates a function g with these properties.

3. Polymorphisms over Henson graphs.

We investigate polymorphisms of reducts of (H_n, E). We start with unary polymorphisms in section 3.1, obtaining that we can assume that the relations E and N are pp-definable in our reducts, since otherwise their CSP can be modeled by a reduct of equality and hence has already been classified in [BK08].

We then turn to binary polymorphisms in section 3.2, obtaining Lemma 3.6 telling us that, excluding in addition just one degenerate case where all polymorphisms are essentially unary functions, we may further assume the existence of a binary injective polymorphism.

Building on the results of those sections, we show in section 3.3 via an analysis of ternary polymorphisms that for any reduct which pp-defines the relations E and N, either the polymorphisms preserve a certain relation H (and hence, H is pp-definable in the reduct by Theorem 2.2), or there is a polymorphism of behavior min (Proposition 3.9).

3.1. The unary case: Model-complete cores. A countable ω-categorical structure Δ is called a *model-complete core* if $\mathrm{Aut}(\Delta)$ is dense in $\mathrm{End}(\Delta)$, or equivalently, every endomorphism of Δ is an elementary self-embedding, i.e., preserves all first-order formulas over Δ. Every countable ω-categorical structure Γ is *homomorphically equivalent* to an up to isomorphism unique ω-categorical model-complete core Δ; that is, there exist homomorphisms from Γ into Δ and vice-versa [Bod07]. Since the CSPs of homomorphically equivalent structures are equal, it has proven fruitful in classification projects to always work with model-complete cores. The following proposition essentially calculates the model-complete cores of the reducts of Henson graphs.

PROPOSITION 3.1. *Let* Γ *be a reduct of* (H_n, E). *Then either* $\mathrm{End}(\Gamma)$ *contains a function whose image induces an independent set or* $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(\Gamma)} = \overline{\mathrm{Aut}(H_n, E)}$.

Proof. Assume that $\mathrm{End}(\Gamma) \neq \overline{\mathrm{Aut}(H_n, E)}$. Then, since Γ is ω-categorical and by Theorem 2.2 and Lemma 2.3, there exists an $f \in \mathrm{End}(\Gamma)$ which violates E or N. If f violated N but not E, then there would be a copy of K_n in the range of f, a contradiction.

Thus, we may assume that f violates E, i.e., there exists $(u, v) \in E$ such that $(f(u), f(v)) \in N$ or $f(u) = f(v)$. If for some such (u, v) we have $f(u) = f(v)$, then one can generate by topological closure from f a function whose image is an independent set. Since this is the first time we appeal to an argument with a flavor of topological closure, let us give it in longhand. First, fix $u, v \in H_n$ such that $E(u, v)$ with $f(u) = f(v)$. Given a subset A of vertices containing $m \geq 1$ edges, we argue that there is a g generated by f so that $g[A]$ contains fewer vertices than A. Indeed, take any $a, b \in A$ with $E(a, b)$, and an automorphism $\alpha \in \mathrm{Aut}(H_n, E)$ mapping (a, b) to (u, v), and use $g = f(\alpha(x), \alpha(y))$. Note that g maps the edge (a, b) to a single vertex, so that $g[A]$ is indeed smaller than A. By iterating this method, we can see that for every finite subset A of H_n, there is a function g generated by f so that $g[A]$ is an independent set. The conclusion that then f also generates a function which sends the entire domain H_n onto an independent set is achieved via a typical compactness argument which appears in one form or another in most works on polymorphism clones of ω-categorical structures; it uses topological closure together with ω-categoricity. The modern and perhaps most elegant way to present it is to consider an equivalence relation $\sim$ on the set F of all functions generated by f, defined by $g \sim g'$ if and only if $\overline{\{\alpha \circ g \mid \alpha \in \mathrm{Aut}(H_n, E)\}} = \overline{\{\alpha \circ g' \mid \alpha \in \mathrm{Aut}(H_n, E)\}}$. Then the factor space $F/_\sim$ is compact since (H_n, E) is ω-categorical. This has first been observed, in a slightly different form, in [BP15b]; we refer the reader to [BP16b] for a proof of the variant we are using here. Let $(A_i)_{i\in\omega}$ be an increasing sequence of finite sets so that $\bigcup_{i\in\omega} A_i = H_n$. Fix a function g_i generated by f which sends A_i onto an independent set. By compactness, a subsequence of $([g_i]_\sim \mid i \in \omega)$ converges in $F/_\sim$ to a class $[g]_\sim$. This means that there are $\alpha_i \in \mathrm{Aut}(H_n, E)$, for $i \in \omega$, such that a subsequence of $(\alpha_i \circ g_i \mid i \in \omega)$ converges to g. But then g maps H_n onto an independent set.

Thus, we may assume that there exists $(u, v) \in E$ such that $(f(u), f(v)) \in N$, and such that no edges are collapsed to $=$ by f. If there existed $(u', v') \in N$ such that $f(u') = f(v')$, then picking an automorphism $\alpha \in \mathrm{Aut}(H_n, E)$ such that $(\alpha(f(u)), \alpha(f(v))) = (u', v')$, we would have that $f \circ \alpha \circ f$ collapses an edge to $=$. Having considered this situation above, we may hence assume that f is injective.

By Proposition 2.9, the operation f generates an injective canonical function $g\colon (H_n, E, \prec, u, v) \to (H_n, E, \prec)$ such that $f(u) = g(u)$ and $f(v) = g(v)$; in fact, since f is unary, we can disregard the order $\prec$ and assume that g is canonical as a function from (H_n, E, u, v) to (H_n, E) [Pon17, Proposition 3.7].

Let

$$\begin{aligned} U_{uv} &:= \{x \in H_n \mid E(u, x) \wedge E(v, x)\}\ , \\ U_{u\overline{v}} &:= \{x \in H_n \mid E(u, x) \wedge N(v, x)\}\ , \\ U_{\overline{u}v} &:= \{x \in H_n \mid N(u, x) \wedge E(v, x)\}\ , \\ \text{and} \quad U_{\overline{uv}} &:= \{x \in H_n \mid N(u, x) \wedge N(v, x)\}. \end{aligned}$$

As all four of these sets contain an independent set of size n, we cannot have $g(N) = E$ on any of them, as this would introduce a copy of K_n. Since no nonedges are collapsed to $=$ by our assumption above, we infer that N is preserved by g on all four sets.

If g violates E on $U_{\overline{uv}}$, then, since $U_{\overline{uv}}$ induces an isomorphic copy of (H_n, E) therein, g generates a function whose image is an independent set. Thus, we may assume that g preserves E on $U_{\overline{uv}}$.

Then g preserves N between $U_{\overline{uv}}$ and any other orbit X of $\mathrm{Aut}(H_n, E, u, v)$, as otherwise it would send nonedges to edges between these orbits, and the image of the n-element induced subgraph of (H_n, E) induced by any point in X together with a copy of K_{n-1} in $U_{\overline{uv}}$ would be isomorphic to K_n.

Assume that g violates E between $U_{\overline{uv}}$ and another orbit X of $\mathrm{Aut}(H_n, E, u, v)$. Let $A \subseteq H_n$ be finite with an edge (x, y) in A. Then there exists an $\alpha \in \mathrm{Aut}(H_n, E)$ such that $\alpha(x) \in X$ and $\alpha[A \setminus \{x\}] \subseteq U_{\overline{uv}}$. The function $(g \circ \alpha) \restriction_A$ preserves N, and it maps (x, y) to a nonedge. By an iterative application of this step we can systematically delete all edges of A. Hence, by topological closure, g generates a function whose image is an independent set. Thus, we may assume that g preserves E between $U_{\overline{uv}}$ and any other orbit of $\mathrm{Aut}(H_n, E, u, v)$.

Let X and Y be infinite orbits of $\mathrm{Aut}(H_n, E, u, v)$, and assume that g violates N between X and Y. There exist vertices $x \in X$ and $y \in Y$ and a copy of K_{n-2} in $U_{\overline{uv}}$ such that (x, y) is the only nonedge in the graph induced by these n vertices. Then, by the above, the image of this n-element set under g induces a copy of K_n, a contradiction. Hence, we may assume that g preserves N on $H_n \setminus \{u, v\}$.

If g violates E on $H_n \setminus \{u, v\}$, then we can systematically delete the edges of any finite subgraph of (H_n, E) while preserving the nonedges and conclude that g generates a function whose image is an independent set. Thus, we may assume that g preserves E on $H_n \setminus \{u, v\}$.

Assume that g violates E between u and $U_{u\overline{v}}$. Given any finite $A \subseteq H_n$ with a vertex $x \in A$, there exists a $\beta \in \mathrm{Aut}(H_n, E)$ such that $\beta(x) = u$ and $\beta[A \setminus \{x\}] \subseteq U_{u\overline{v}} \cup U_{\overline{uv}}$. Since, as observed earlier, g preserves N between $U_{\overline{uv}}$ and any other orbit of $\mathrm{Aut}(H_n, E, u, v)$, including the orbits $U_{u\overline{v}}$ and $\{u\}$, we conclude that $(g \circ \beta) \restriction_A$ preserves N, and it maps edges from x to nonedges. Thus, we can systematically delete the edges of A, and consequently, g generates a function whose image is an independent set. Hence, we may assume that g preserves E between u and $U_{u\overline{v}}$.

There exist a vertex $x \in U_{\overline{u}v}$ and a copy of K_{n-2} in $U_{u\overline{v}}$ such that (x, u) is the only nonedge in the graph induced by these $n-1$ vertices together with u. Thus, if g violates N between $\{u\}$ and $U_{\overline{u}v}$, then the image of this n-element set under g induces a copy of K_n, a contradiction. Hence, g preserves N between u and $U_{\overline{u}v}$.

By symmetry, we may assume that g preserves N between v and $U_{u\overline{v}}$. Thus, g preserves N. As g deletes the edge between u and v, we can systematically delete the edges of any finite subgraph of (H_n, E). Hence, g generates a function whose image is an independent set. ∎

In the first case of Proposition 3.1, the model-complete core of the reduct is in fact a reduct of equality. Since the CSPs of reducts of equality have been classified [BK08], we do not have to consider any further reducts with an endomorphism whose image induces an independent set.

LEMMA 3.2. *Let Γ be a reduct of (H_n, E), and assume that* $\mathrm{End}(\Gamma)$ *contains a function whose image is an independent set. Then Γ is homomorphically equivalent to a reduct of $(H_n, =)$.*

Proof. The proof is trivial. ∎

In the second case of Proposition 3.1, it turns out that all polymorphisms preserve the relations E, N, and $\neq$, by the following lemma and Theorem 2.2.

LEMMA 3.3. *Let Γ be a reduct of (H_n, E). Then the following are equivalent:*
(1) $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(H_n, E)}$.
(2) *E and N have pp definitions in Γ.*
(3) *E, N, and $\neq$ have pp definitions in Γ.*

Proof. Since E and N are orbits of pairs with respect to $\mathrm{Aut}(H_n, E)$, the implication from (1) to (2) is an immediate consequence of Theorem 2.2 and Lemma 2.3. For the implication from (2) to (3), it is enough to observe that the pp formula $\exists z(E(x, z) \wedge N(y, z))$ defines $x \neq y$. Finally, the implication from (3) to (1) follows from the homogeneity of (H_n, E). ∎

Before moving on to binary polymorphisms, we observe the following corollary of Proposition 3.1, first mentioned in [Tho91].

COROLLARY 3.4. *For every $n \geq 3$, the permutation group* $\mathrm{Aut}(H_n, E)$ *is a maximal closed subgroup of the full symmetric group on H_n; i.e., every closed subgroup of the full symmetric group containing* $\mathrm{Aut}(H_n, E)$ *equals either* $\mathrm{Aut}(H_n, E)$ *or the full symmetric group.*

Proof. Let $G \supseteq \mathrm{Aut}(H_n, E)$ be a closed subgroup of the full symmetric group on H_n. Its closure $\overline{G}$ in the set of all unary functions on H_n is a closed transformation monoid, i.e., a topologically closed monoid of unary functions, and hence the monoid of endomorphisms of a reduct of (H_n, E) (cf., for example, [BP14]). By Proposition 3.1, $\overline{G}$ either contains a function e whose image induces an independent set or equals $\overline{\mathrm{Aut}(H_n, E)}$. In the latter case, $G = \mathrm{Aut}(H_n, E)$, and in the first case we prove that G equals the full symmetric group. Since G is closed in the full symmetric group, it suffices to prove that for every $k \geq 1$ and all $s, t \in H_n^k$ there exists an element of G which sends s to t. Since $e \in \overline{G}$, there exists a $\beta \in G$ such that e and β agree on the tuples s and t, and consequently β sends the two tuples into independent sets. By the homogeneity of (H_n, E), we have that $\beta(s)$ and $\beta(t)$ lie in the same orbit of G, and hence so do s and t. ∎

We remark that the automorphism group of the random graph has five closed supergroups [Tho91], which leads to more cases in the corresponding CSP classification in [BP15a].

3.2. Higher arities: Generating injective polymorphisms. We investigate at least binary functions preserving E and N (and hence, by Theorem 2.2, also $\neq$, since this relation is pp-definable from E and N by Lemma 3.3); our goal in this section is to show that they generate injections. Every unary function gives rise to a binary function by adding a dummy variable; the following definition rules out such "improper" higher-arity functions.

DEFINITION 3.5. *A finitary operation $f(x_1,\dots,x_n)$ on a set is* essential *if it depends on more than one of its variables x_i.*

LEMMA 3.6. *Let $f\colon H_n^2 \to H_n$ be a binary essential function that preserves E and N. Then f generates a binary injection.*

Proof. Let Δ be the structure with domain H_n and whose relations are those preserved by $\{f\} \cup \mathrm{Aut}(H_n,E)$; in particular, E, N, and $\neq$ are relations of Δ. It is sufficient to show that $\mathrm{Pol}(\Delta)$ contains a binary injection (see section 2.5).

We follow the strategy of the proof of [BP14, Theorem 38]. By [BP14, Lemma 42] it is enough to show that for all pp formulas ϕ over Δ we have that whenever $\phi \wedge x \neq y$ and $\phi \wedge s \neq t$ are satisfiable in Δ, the formula $\phi \wedge x \neq y \wedge s \neq t$ is also satisfiable in Δ. Still following the proof of [BP14, Proposition 38] it is enough to show the following claim.

Claim. Given two 4-tuples $a = (x,y,z,z)$ and $b = (p,p,q,r)$ in H_n^4 such that $x \neq y$ and $q \neq r$, there exist 4-tuples a' and b' such that $\mathrm{tp}(a) = \mathrm{tp}(a')$ and $\mathrm{tp}(b) = \mathrm{tp}(b')$ in (H_n,E) and such that $f(a',b')$ is a 4-tuple whose first two coordinates are different and whose last two coordinates are different.

Proof of claim. We may assume that $x \neq z$ and $p \neq q$. We may also assume that f itself is not a binary injection.

In the following, we say that a point $(u,v) \in H_n^2$ is *good* if $f(u,v) \neq f(u,w)$ for all $v \neq w$. Assume without loss of generality that there exist $u_1 \neq u_2$, $v \in H_n$ such that $f(u_1,v) = f(u_2,v)$. In particular, as f preserves $\neq$, the points (u_1,v) and (u_2,v) are good. First fix any values z',q' such that (z',q') is good. We may assume that for any $x',y',p' \in H_n$ with $\mathrm{tp}(x',y',z') = \mathrm{tp}(x,y,z)$ and $\mathrm{tp}(p',q') = \mathrm{tp}(p,q)$ we have $f(x',p') = f(y',p')$; otherwise the tuples $a' = (x',y',z',z')$ and $b' = (p',p',q',r')$ are appropriate with any $r' \in H_n$ with $\mathrm{tp}(p',q',r') = \mathrm{tp}(p,q,r)$. Hence, as f preserves $\neq$, all the points (x',p') with $\mathrm{tp}(x',z') = \mathrm{tp}(x,z)$ and $\mathrm{tp}(p',q') = \mathrm{tp}(p,q)$ are good. So we obtained that whenever the point (s,t) is good, and $s_0,t_0 \in H_n$ are such that $\mathrm{tp}(s,s_0) = \mathrm{tp}(x,z)$ and $\mathrm{tp}(t,t_0) = \mathrm{tp}(p,q)$, then (s_0,t_0) is also good, or otherwise we are done. We show that whatever the types $Q_1 = \mathrm{tp}(x,z)$ and $Q_2 = \mathrm{tp}(p,q)$ are, we can reach any point (s_4,t_4) in H_n^2 from a given good point (s_0,t_0) by at most four such steps. To see this, note that Q_1 and Q_2 are different from $=$ by assumption. Now let s_1,s_2,s_3,t_1,t_2,t_3 be such that

- s_0,s_1,s_2,s_3,s_4 are pairwise different except that $s_0 = s_4$ is possible, and
- t_0,t_1,t_2,t_3,t_4 are pairwise different except that $t_0 = t_4$ is possible, and
- $(s_0,s_1),(s_1,s_2),(s_2,s_3),(s_3,s_4) \in Q_1$ and all other pairs (s_i,s_j) are in N except that $s_0 = s_4$ is possible, and
- $(t_0,t_1),(t_1,t_2),(t_2,t_3),(t_3,t_4) \in Q_2$ and all other pairs (t_i,t_j) are in N except that $t_0 = t_4$ is possible.

These rules are not in contradiction with the extension property of (H_n,E); thus such vertices exist, and we can propagate the good property from (s_0, t_0) to (s_4, t_4). Hence, every point is good, or we are done. If $f(u_1, v) = f(u_2, v)$ for all $u_1, u_2, v \in H_n$ with $\mathrm{tp}(u_1, u_2) = \mathrm{tp}(x, y)$, then f would be essentially unary, since (H_n, E) and its complement have diameter 2. As f is a binary essential function, we can choose $x', y', p' \in H_n$ such that $\mathrm{tp}(x', y') = \mathrm{tp}(x, y)$ and $f(x', p') \neq f(y', p')$. By choosing points $z', q', r' \in H_n$ such that $\mathrm{tp}(x', y', z') = \mathrm{tp}(x, y, z)$ and $\mathrm{tp}(p', q', r') = \mathrm{tp}(p, q, r)$ the tuples $a' = (x', y', z', z')$ and $b' = (p', p', q', r')$ are appropriate. ∎

The following lemma allows us to drop the restriction to binary essential functions.

LEMMA 3.7. *Let $k \geq 2$. Every essential function $f\colon H_n^k \to H_n$ that preserves E and N generates a binary injection.*

Proof. By [BP14, Lemma 40], every essential operation generates a binary essential operation over the random graph; the very same proof works for the Henson graphs. Therefore, we may assume that f itself is binary. The assertion now follows from Lemma 3.6. ∎

3.3. The relation *H*. Let us investigate the case in which Γ, a reduct of (H_n, E), pp-defines E and N (and hence, $\neq$). The following relation characterizes the NP-complete cases in this situation.

DEFINITION 3.8. *We define a 6-ary relation $H(x_1, y_1, x_2, y_2, x_3, y_3)$ on H_n by*

$$\begin{aligned} &\bigwedge_{i,j \in \{1,2,3\}, i \neq j, u \in \{x_i, y_i\}, v \in \{x_j, y_j\}} N(u, v) \\ &\wedge \big((E(x_1, y_1) \wedge N(x_2, y_2) \wedge N(x_3, y_3)) \\ &\quad \vee \ (N(x_1, y_1) \wedge E(x_2, y_2) \wedge N(x_3, y_3)) \\ &\quad \vee \ (N(x_1, y_1) \wedge N(x_2, y_2) \wedge E(x_3, y_3))\big). \end{aligned}$$

Our goal for this section is to prove the following proposition, which states that if Γ is a reduct of (H_n, E) with E and N pp-definable in Γ, then either H has a pp definition in Γ, in which case $\mathrm{CSP}(\Gamma)$ is NP-complete, or $\mathrm{Pol}(\Gamma)$ has a certain canonical polymorphism which will imply tractability of the CSP. NP-completeness and tractability for those cases will be shown in section 4.

PROPOSITION 3.9. *Let Γ be a reduct of (H_n, E) with E and N pp-definable in Γ. Then at least one of the following holds:*

(a) *There is a pp definition of H in Γ.*

(b) $\mathrm{Pol}(\Gamma)$ *contains a canonical binary injection of behavior* min.

3.3.1. Arity reduction: Down to binary. With the ultimate goal of producing a binary canonical polymorphism of behavior min, we now show that under the assumption that Γ has a polymorphism preserving E and N yet violating H, it also has a binary polymorphism which is not of behavior projection. We begin by ruling out some ternary behaviors which do play a role on the random graph.

LEMMA 3.10. *On (H_n, E), there are no ternary functions of behavior majority or satisfying the type conditions $f(N, N, E) = f(E, N, N) = E$.*

Proof. These could introduce a K_n in the K_n-free graph (H_n, E), in the following fashions.

Suppose f has behavior majority, and choose $x_1, \ldots, x_{n-1} \in H_n$ inducing a copy of K_{n-1}, as well as a distinct $x_0 \in H_n$ adjacent to x_1 and no other x_i. Then $\{f(x_0, x_1, x_2), f(x_1, x_2, x_0), f(x_2, x_0, x_1)\}$ induces K_3 and is adjacent to any element in $\{f(x_i, x_i, x_i) \mid 2 < i \leq n-1\}$ since E is preserved, so that altogether we obtain a copy of K_n.

Suppose now that f satisfies the type conditions $f(N,N,E) = f(E,N,N) = E$, and choose elements $x_1, \ldots, x_{n-1} \in H_n^3$ such that $NNE(x_i, x_j)$ holds for distinct $1 \leq i, j \leq n-1$. Pick furthermore $x_0 \in H_n^3$ with $ENN(x_0, x_i)$ for all $1 \leq i \leq n-1$. Then $\{f(x_0), \ldots, f(x_{n-1})\}$ induces a K_n. ∎

PROPOSITION 3.11. *Let $f\colon H_n^k \to H_n$ be an operation that preserves E and N and violates H. Then f generates a binary injection which is not of behavior projection.*

Proof. Since H consists of three orbits of 6-tuples in (H_n, E), we may assume that f is ternary, by Lemma 2.3. Moreover, since f preserves E and N, it can only violate H if it is essential. Thus, by Lemma 3.7, f generates a binary injection g. If g is not of behavior projection, then we have proved the proposition. Otherwise, assume without loss of generality that it is of behavior p_1. Consider the ternary function $g(g(g(f(x,y,z),x),y),z)$. This function is injective since g is. Moreover, it violates H: if $x^1, x^2, x^3 \in H$ are so that $t := f(x^1, x^2, x^3) \notin H$, then t has pairwise distinct entries since f preserves $\neq$. Hence, because g is of behavior p_1, $t' := g(t, x^1)$ has the same type as t, and so do $t'' := g(t', x^2)$ and $t''' := g(t'', x^3)$, proving the claim. By substituting f by this function, we can therefore in the following assume that f is itself injective.

We now prove the proposition by showing that a function of the form $(x,y) \mapsto f(x,y,\alpha(x))$, or $(x,y) \mapsto f(x,\alpha(x),y)$, or $(x,y) \mapsto f(y,x,\alpha(x))$, where $\alpha \in \mathrm{Aut}(H_n, E)$, is not of type projection.

Fix $x^1, x^2, x^3 \in H$ such that $f(x^1, x^2, x^3) \notin H$. In the following, we will write $x_i := (x_i^1, x_i^2, x_i^3)$ for $1 \leq i \leq 6$. So $(f(x_1), \ldots, f(x_6)) \notin H$. If there exists $\alpha \in \mathrm{Aut}(H_n, E)$ such that $\alpha(x^i) = x^j$ for $1 \leq i \neq j \leq 3$, then our claim follows: for example, if $i = 1$ and $j = 3$, then the function $(x,y) \mapsto f(x,y,\alpha(x))$ violates H and hence cannot be of behavior projection.

We assume henceforth that there is no such automorphism α. In this situation, by permuting arguments of f if necessary, we can assume without loss of generality that

$$ENN(x_1, x_2),\ NEN(x_3, x_4), \text{ and } NNE(x_5, x_6).$$

We set

$$S := \{y \in H_n^3 \mid NNN(x_i, y) \text{ for all } 1 \leq i \leq 6\}.$$

Consider the binary relations $Q_1Q_2Q_3$ on H_n^3, where $Q_i \in \{E, N\}$ for $1 \leq i \leq 3$. We show that either our claim above proving the proposition holds, or for each such relation $Q_1Q_2Q_3$, whether $E(f(u), f(v))$ or $N(f(u), f(v))$ holds for $u, v \in S$ with $Q_1Q_2Q_3(u,v)$ does not depend on u, v; that is, whenever $u, v, u', v' \in S$ satisfy $Q_1Q_2Q_3(u,v)$ and $Q_1Q_2Q_3(u', v')$, then $E(f(u), f(v))$ if and only if $E(f(u'), f(v'))$. Note that this is another way of saying that f satisfies some type conditions on S. We go through all possibilities of $Q_1Q_2Q_3$.

(1) $Q_1Q_2Q_3 = ENN$. Let $\alpha \in \mathrm{Aut}(H_n, E)$ be such that (x_1^2, x_2^2, u_2, v_2) is mapped to (x_1^3, x_2^3, u_3, v_3); such an automorphism exists since

$$NNN(x_1, u), NNN(x_1, v), NNN(x_2, u), NNN(x_2, v)$$

hold, and since (x_1^2, x_2^2) has the same type as (x_1^3, x_2^3), and (u_2, v_2) has the same type as (u_3, v_3). We are done if the operation g defined by $g(x,y) := f(x,y,\alpha(y))$ is not of type projection. Otherwise, $E(g(u_1, u_2), g(v_1, v_2))$

if and only if $E(g(x_1^1, x_1^2), g(x_2^1, x_2^2))$. Combining this with the equations $(f(u), f(v)) = (g(u_1, u_2), g(v_1, v_2))$ and $(g(x_1^1, x_1^2), g(x_2^1, x_2^2)) = (f(x_1), f(x_2))$, we get that $E(f(u), f(v))$ if and only if $E(f(x_1), f(x_2))$, and so our claim holds for this case.

(2) $Q_1Q_2Q_3 = NEN$ or $Q_1Q_2Q_3 = NNE$. These cases are analogous to the previous case.

(3) $Q_1Q_2Q_3 = NEE$. Let α be defined as in the first case. Reasoning as above, if the operation defined by $f(x, y, \alpha(y))$ is of type projection, then one gets that $E(f(u), f(v))$ if and only if $N(f(x_1), f(x_2))$.

(4) $Q_1Q_2Q_3 = ENE$ or $Q_1Q_2Q_3 = EEN$. These cases are analogous to the previous case.

(5) $Q_1Q_2Q_3 = EEE$ or $Q_1Q_2Q_3 = NNN$. This is trivial since f preserves E and N.

Now we show that f actually cannot satisfy the type conditions above on S. First note that setting $h(x, y, z) := f(e^1(x), e^2(y), e^3(z))$ for self-embeddings e_1, e_2, e_3 of (H_n, E) such that $(e_1, e_2, e_3)(u) \in S$ for all $u \in H_n^3$, we obtain a function h which satisfies the same type conditions everywhere; such embeddings exist since by its definition, the projection of S onto any coordinate has an induced copy of (H_n, E). Then, if $\{(f(x_1), f(x_2)), (f(x_3), f(x_4)), (f(x_5), f(x_6))\}$ has E twice or more, by (1) and (2) we get that h satisfies two type conditions from the minority behavior, say $h(N, N, E) = E$ and $h(E, N, N) = E$, contradicting Lemma 3.10. If $\{(f(x_1), f(x_2)), (f(x_3), f(x_4)), (f(x_5), f(x_6))\}$ has E no times, then by (3) and (4) h is of behavior majority, again contradicting Lemma 3.10. Thus, the set must have precisely one E, contradicting $f(x^1, x^2, x^3) \notin H$. ∎

3.3.2. Producing min. By Proposition 3.11, it remains to show the following to obtain a proof of Proposition 3.9.

PROPOSITION 3.12. *Let $f\colon H_n^2 \to H_n$ be a binary injection preserving E and N that is not of behavior projection. Then f generates a binary canonical injection of behavior* min.

In the remainder of this section we will prove this proposition by a Ramsey theoretic analysis of f, which requires the following definitions and facts from [BP14] concerning behaviors with respect to the homogeneous expansion of the graphs (H_n, E) by the total order $\prec$ from section 2.7. At this point, it might be appropriate to remark that canonicity of functions on H_n, and even the notion of behavior, does depend on which underlying structure we have in mind, in particular, whether or not we consider the order $\prec$ (which we have almost managed to ignore so far). Let us define the following behaviors for functions from $(H_n, E, \prec)^2$ to (H_n, E); we write $\succ$ for the relation $\{(a, b) \mid b \prec a\}$.

DEFINITION 3.13. *Let $f\colon H_n^2 \to H_n$ be injective. If for all $u, v \in H_n^2$ with $u_1 \prec v_1$ and $u_2 \prec v_2$*

- *$E(f(u), f(v))$ if and only if $EE(u, v)$, then f* behaves like min on input $(\prec, \prec)$*;*
- *$E(f(u), f(v))$ if and only if $E(u_1, v_1)$, then f* behaves like p_1 on input $(\prec, \prec)$*;*
- *$E(f(u), f(v))$ if and only if $E(u_2, v_2)$, then f* behaves like p_2 on input $(\prec, \prec)$.

Analogously, we define behaviors on input $(\prec, \succ)$ using pairs $u, v \in H_n^2$ with $u_1 \prec v_1$ and $u_2 \succ v_2$.

PROPOSITION 3.14. *Let $f\colon H_n^2 \to H_n$ be an injection which is canonical as a function from $(H_n, E, \prec)^2$ to $(H_n, E, \prec)$ and suppose f preserves E and N. Then it behaves like* min, *p_1 or p_2 on input $(\prec, \prec)$ (and similarly on input $(\prec, \succ)$).*

Proof. By definition of the term canonical; one only needs to enumerate all possible types of pairs (u,v), where $u,v \in H_n^2$, and recall that (H_n,E) does not contain any clique of size n, which makes some behaviors impossible to be realized by f. ∎

DEFINITION 3.15. *If an injection* $f\colon H_n^2 \to H_n$ *behaves like* X *on input* $(\prec,\prec)$ *and like* Y *on input* $(\prec,\succ)$*, where* $X,Y \in \{\min, p_1, p_2\}$*, then we say that* f *is of* behavior X/Y.

In the following lemmas, we show that every injective canonical binary function which behaves differently on input $(\prec,\prec)$ and on input $(\prec,\succ)$ generates a function which behaves the same way on both inputs, allowing us to ignore the order again.

LEMMA 3.16. *Suppose that* $f\colon H_n^2 \to H_n$ *is injective and canonical from* $(H_n,E,\prec)^2$ *to* $(H_n,E,\prec)$*, and suppose that it is of type* $\min/p_i$ *or of type* $p_i/\min$*, where* $i \in \{1,2\}$*. Then* f *generates a binary injection of type* $\min$.

Proof. Since the calculus for behaviors on the Henson graphs is the same as that on the random graph, the same proof as in [BP15a] works. ∎

LEMMA 3.17. *No binary injection* $f\colon H_n^2 \to H_n$ *can have behavior* p_1/p_2.

Proof. Such a behavior would introduce a K_n in a K_n-free graph. ∎

Having ruled out some behaviors without constants, we finally introduce constants to the language to prove Proposition 3.12.

Proof of Proposition 3.12. Fix a finite set $C := \{c_1,\dots,c_m\} \subseteq H_n$ on which the fact that f is not of behavior projection is witnessed. Invoking Proposition 2.9, we may henceforth assume that f is canonical as a function from $(H_n,E,\prec,c_1,\dots,c_m)^2$ to $(H_n,E,\prec)$. We are going to show that f generates a binary injection g of behavior min. Then another application of Proposition 2.9 to g yields a canonical function g'; this function is still of behavior min because any function of the form $\alpha(g(\beta(x),\gamma(y)))$ is of type min, for automorphisms α,β,γ of (H_n,E), and g' is generated from operations of this type by topological closure.

To obtain g, consider in the structure $(H_n,E,\prec,c_1,\dots,c_m)$ the orbit

$$O := \{a \in H_n \mid N(a,c_i) \text{ and } a \prec c_i \text{ for all } 1 \le i \le m\}.$$

Then O induces a structure isomorphic to $(H_n,E,\prec)$, as it satisfies the extension property for totally ordered K_n-free graphs: the same extensions can be realized in O as in $(H_n,E,\prec)$. Therefore, by Lemma 3.14, f has one of the three mentioned behaviors on input $(\prec,\prec)$ and on input $(\prec,\succ)$. By Lemmas 3.16 and 3.17, we may assume that f behaves like a projection on O, for any other combination of behaviors implies that it generates a binary injection of behavior min.

Assume without loss of generality that f behaves like p_1 on O. Let $u \in O^2$ and $v \in (H_n \setminus \{c_1,\dots,c_m\})^2$ satisfy $\neq\neq(u,v)$; we claim that f behaves like p_1 or like min on $\{u,v\}$. Otherwise we must have $NE(u,v)$ and $E(f(u),f(v))$. Pick $q_1,\dots,q_{n-1} \in O^2$ forming a clique in the first coordinate, an independent set in the second coordinate, and such that the type of (q_i,v) equals the type of (u,v) in $(H_n,E,\prec,c_1,\dots,c_n)$. Then by canonicity, the image of $\{q_1,\dots,q_{n-1},v\}$ under f forms a clique of size n, a contradiction.

Suppose next that there exist $u \in O^2$ and $v \in (H_n \setminus C)^2$ with $\neq\neq(u,v)$ such that f does not behave like p_1 (and hence, by the above, behaves like min) on $\{u,v\}$. This means that $EN(u,v)$ and $N(f(u),f(v))$. We use topological closure to show that f

generates a binary injection which behaves like min. To this end, set

$$S := \{p \in H_n^2 \mid \operatorname{tp}(p,v) = \operatorname{tp}(u,v) \text{ in } (H_n, E, \prec, c_1, \ldots, c_n)\} \subseteq O^2.$$

Now let $q_0 \in H_n^2$ be arbitrary. Pick a self-embedding e of (H_n, E) whose range is contained in O. Then the function $r\colon H_n^2 \to H_n^2$ defined by $(x,y) \mapsto (f(e(x), e(y)), f(e(y), e(x)))$ has the property that $EN(p,q)$ implies $EN(r(p), r(q))$ and $NE(p,q)$ implies $NE(r(p), r(q))$ for all $p, q \in H_n^2$, since f behaves like p_1 on O. Moreover, since f is injective, we have that $p \neq q$ implies $\neq\neq(r(p), r(q))$. By the latter property, there exist self-embeddings e_1, e_2 of (H_n, E) such that for the function $r'\colon H_n^2 \to H_n^2$ defined by $r' := (e_1, e_2) \circ r$ we have that $r'(q_0) = v$, that $r'(p) \in O^2$ for all $p \in H_n^2 \setminus \{q_0\}$, and that $r'(p) \in S$ for all $p \in H_n^2$ with $EN(p, q_0)$. Then the function $h\colon H_n^2 \to H_n^2$ defined by $h(x,y) := (f(r'(x,y)), y)$ has the property that $NN(h(p), h(q_0))$ holds for all $p \in H_n^2$ with $EN(p, q_0)$, since f behaves like min between S and v. Moreover, $NE(h(p), h(q_0))$ holds for all $p \in H_n^2$ with $NE(p, q_0)$, since f behaves like p_1 or like min between O^2 and v. Finally, for any $p, p' \in H_n^2$ distinct from q_0 we have that $EN(p,p')$ implies $EN(h(p), h(p'))$ and $NE(p,p')$ implies $NE(h(p), h(p'))$, since f behaves like p_1 on O. Similarly, one can construct a function h' on H_n^2 which preserves EN and NE between any $p, p' \in H_n^2$ distinct from q_0, and such that $NE(p, q_0)$ implies $NN(h'(p), h'(q_0))$. Iterating such functions for different choices of q_0, we obtain functions $r_A\colon H_n^2 \to H_n^2$ for every finite subset $A \subseteq H_n^2$ such that $EN(p,p')$ or $NE(p,p')$ implies $NN(r_A(p), r_A(p'))$ for all $p, p' \in A$. By topological closure (cf. Proposition 3.1), one then gets a function $r\colon H_n^2 \to H_n^2$ which has this property everywhere, and then $f(r)$ is the desired binary injection of behavior min.

So we assume henceforth that f behaves like p_1 on $\{u,v\}$ for all $u \in O^2$ and all $v \in (H_n \setminus C)^2$ with $\neq\neq(u,v)$. We then claim that f must behave like p_1 or like min on $\{u,v\}$ for all $u, v \in (H_n \setminus C)^2$ with $\neq\neq(u,v)$. Otherwise, we must have $NE(u,v)$ and $E(f(u), f(v))$. Pick $q_1, \ldots, q_{n-2} \in O^2$ forming a clique in the first coordinate, forming an independent set in the second coordinate, and adjacent to u and v in the first coordinate. Applying f, we get a clique of size n, a contradiction.

If there exist $u, v \in (H_n \setminus C)^2$ with $EN(u,v)$ and $N(f(u), f(v))$, then by precomposing f with a self-embedding e of (H_n, E) whose range equals $H_n \setminus C$, we may moreover assume that f behaves like p_1 or like min on $\{u', v'\}$ for all $u', v' \in H_n^2$. A standard iterating argument, similar to the one above (or the one given in detail in the proof of Proposition 3.1), then shows that f generates a binary injection g of type min.

We thus henceforth assume that f behaves like p_1 on $\{u,v\}$ for all $u, v \in (H_n \setminus C)^2$. We next claim that f must behave like p_1 or like min on $\{u,v\}$ for all $u \in H_n^2$ and all $v \in (H_n \setminus C)^2$ with $\neq\neq(u,v)$. Otherwise, we must have $NE(u,v)$ and $E(f(u), f(v))$. Pick $q_1, \ldots, q_{n-2} \in H_n^2$ forming a clique in the first coordinate, forming an independent set in the second coordinate, adjacent to u in both coordinates, and adjacent to v in precisely the first coordinate. Applying f, we get a clique of size n, a contradiction.

An argument similar to that two paragraphs above now shows that if there exist $u, v \in H_n^2$ with $EN(u,v)$ and $N(f(u), f(v))$, then f generates a binary injection g of type min.

It thus remains to consider the case where f behaves like p_1 on $\{u,v\}$ whenever $u \in H_n^2$ and $v \in (H_n \setminus C)^2$, and where there exist $u, v \in H_n^2$ such that $NE(u,v)$ and $E(f(u), f(v))$. Pick $q_1, \ldots, q_{n-2} \in (H_n \setminus C)^2$ which are adjacent to u and v in the first coordinate and not the second, which form a clique in the first coordinate, and which form an independent set in the second coordinate. The image of the set $\{u, v, q_1, \ldots, q_{n-2}\}$ under f then is a clique of size n so that this case cannot occur. ☐

4. CSPs over Henson graphs.

4.1. Hardness of $\boldsymbol{H}$. We now show that any reduct of (H_n, E) which has H among its relations, and hence by Lemma 2.1 every reduct which pp-defines H, has an NP-hard CSP. We first show hardness directly by reduction from positive 1-in-3-SAT; then, we provide another proof via *h*1 *clone homomorphisms*, which gives further insight into the mathematical structure of such reducts and draws connections to the general dichotomy conjecture for reducts of finitely bounded homogeneous structures.

4.1.1. Reduction from positive 1-in-3-SAT. We start by showing hardness directly, which however does not tell us anything about the structure of the polymorphism clones of reducts which pp-define H.

PROPOSITION 4.1. $\mathrm{CSP}(H_n, H)$ *is NP-hard.*

Proof. We reduce positive 1-in-3-SAT to $\mathrm{CSP}(H_n, H)$. Each variable v in an instance ϕ of the former becomes two variables v, v' in the corresponding instance ψ of the latter. Each clause (u, v, w) from ϕ becomes a tuple $H(u, u', v, v', w, w')$ in ψ. It is easy to see that ϕ is a yes-instance of 1-in-3-SAT if and only if ψ is a yes-instance of $\mathrm{CSP}(H_n, H)$, and the result follows. ☐

4.1.2. Clone homomorphisms. We will now show another way to prove NP-hardness of $\mathrm{CSP}(H_n, H)$ via a structural property of $\mathrm{Pol}(H_n, H)$, using general results from [BP16a] (a strengthening of the structural hardness proof in [BP15b]). This will allow us to show that the dichotomy for the Henson graphs is in line with the dichotomy conjecture for CSPs of reducts of finitely bounded homogeneous structures, from [BP16a] (and the earlier dichotomy conjecture for the same class, due to Bodirsky and Pinsker (cf. [BPP14]), which has recently been proved equivalent [BKO+17]).

DEFINITION 4.2. *Let* Γ *be a structure. A* projective clone homomorphism *of* Γ *(or* $\mathrm{Pol}(\Gamma)$*) is a mapping from* $\mathrm{Pol}(\Gamma)$ *onto its projections which*
- *preserves arities;*
- *fixes each projection;*
- *preserves composition.*

A projective strong h1 clone homomorphism *of* Γ *is a mapping such as that above, where the third condition is weakened to preservation of composition of any function in* $\mathrm{Pol}(\Gamma)$ *with projections only.*

Recall that $\mathrm{Pol}(\Gamma)$ is equipped with the topology of pointwise convergence for any structure Γ.

THEOREM 4.3 (from [BP16a]). *Let* Γ *be a countable* ω*-categorical structure in a finite relational language which has a uniformly continuous projective strong* $h1$ *clone homomorphism. Then* $\mathrm{CSP}(\Gamma)$ *is NP-hard.*

PROPOSITION 4.4. *The structure* (H_n, H) *has a uniformly continuous projective strong* $h1$ *clone homomorphism. Consequently,* $\mathrm{CSP}(H_n, H)$ *is NP-hard.*

Proof. Note that H consists of three orbits of 6-tuples with respect to $\mathrm{Aut}(H_n, E)$. Let $a^1, a^2, a^3 \in H$ be representatives of those three orbits. By reshuffling the a^i we may assume that $\mathit{ENN}(a_1, a_2)$, $\mathit{NEN}(a_3, a_4)$, $\mathit{NNE}(a_5, a_6)$ (where a_i denotes the ith row of the matrix (a^1, a^2, a^3) for $1 \leq i \leq 6$).

We claim that whenever $f \in \mathrm{Pol}(H_n, H)$ is ternary, and $b^1, b^2, b^3 \in H$ are so that $\mathrm{tp}(b^1, b^2, b^3) = \mathrm{tp}(a^1, a^2, a^3)$, then $\mathrm{tp}(f(b^1, b^2, b^3)) = \mathrm{tp}(f(a^1, a^2, a^3))$ in (H_n, E). To see this, let $c^1, c^2, c^3 \in H$ be so that $\mathrm{tp}(c^1, c^2, c^3) = \mathrm{tp}(b^1, b^2, b^3)$, and such that no entry of any c^i is adjacent to any component of any b^j or a^j. Suppose that $f(b^1, b^2, b^3)$ and $f(a^1, a^2, a^3)$ do not have the same type; then one of them, say $f(a^1, a^2, a^3)$, does not have the same type as $f(c^1, c^2, c^3)$. Without loss of generality, this is witnessed on the first two components of the 6-tuples $f(c^1, c^2, c^3)$ and $f(a^1, a^2, a^3)$. For $1 \leq i \leq 3$, consider the 6-tuple $d^i := (c^i_1, c^i_2, a^i_3, \ldots, a^i_6)$; i.e., in a^i we replace the first two components by the components from c^i. Then $d^i \in H$, but $f(d^1, d^2, d^3) \notin H$, a contradiction.

Let $f \in \mathrm{Pol}(H_n, H)$. Then precisely one out of $(f(a_1), f(a_2))$, $(f(a_3), f(a_4))$, and $(f(a_5), f(a_6))$ is contained in E. If this is the case for the first pair, then it follows from the claim above that f satisfies the three type conditions $f(E, N, N) = E$ and $f(N, E, N) = f(N, N, E) = N$; in the other two cases we obtain similar type conditions.

Let ξ be the mapping which sends every ternary $f \in \mathrm{Pol}(H_n, H)$ to the ternary projection which is consistent with the type conditions satisfied by f (in the case considered above, the projection onto the first coordinate). Then ξ clearly preserves arities and projections. Moreover, let $f \in \mathrm{Pol}(H_n, H)$ and say without loss of generality that $f(E, N, N) = E$, so that $\xi(f)$ is the first projection. Then whenever $g_1, g_2, g_3 \in \mathrm{Pol}(H_n, H)$ are ternary projections, we have $\xi(f(g_1, g_2, g_3)) = g_1$; this is easy to verify by checking the behavior of $f(g_1, g_2, g_3)$ on a suitable triple of the form (E, N, N), (N, E, N), or (N, N, E). Hence, ξ satisfies the definition of a strong projective h1 clone homomorphism for ternary functions. It has been observed (see, e.g., [BP18]) that ξ then uniquely extends to a strong projective h1 clone homomorphism of the entire clone $\mathrm{Pol}(H_n, H)$. Since the value of every f under ξ can be seen on any test matrix (a^1, a^2, a^3) as above, we have that ξ is uniformly continuous, and so is its extension (the latter follows from the proof in [BP18]). ∎

4.2. Tractability of min. We now show that if a reduct Γ of (H_n, E) with finite relational signature has a polymorphism which is of behavior min, then $\mathrm{CSP}(\Gamma)$ is in P. We are going to apply Theorem 4.5 below for the structure $\Delta := (H_n, E)$. In the theorem, $\hat{\Delta}$ denotes the expansion of Δ by the inequality relation $\neq$ and by the complement $\hat{R}$ of each relation R in Δ.

THEOREM 4.5 (Proposition 14 in [BCKvO09]). *Let Δ be an ω-categorical structure, and let Γ be a reduct of Δ. If Γ has a polymorphism e which is an embedding of Δ^2 into Δ, and if $\mathrm{CSP}(\hat{\Delta})$ is in P, then $\mathrm{CSP}(\Gamma)$ is in P as well.*

PROPOSITION 4.6. *Let Γ be a reduct of (H_n, E) which has a polymorphism of behavior* min. *Then $\mathrm{CSP}(\Gamma)$ is in P.*

Proof. To apply Theorem 4.5 to $\Delta = (H_n, E)$, we first show that the CSP for $\hat{\Delta} = (H_n, E, \hat{E}, \neq)$ can be solved in polynomial time. Given an input primitive positive formula, we identify all variables x, y such that $x = y$ is a constraint in the input. Then the formula is satisfiable if and only if for all variables $x_1, \ldots, x_n$ we have the following conditions:

- $E(x_i, x_j)$ is not in the input for some distinct $i, j \in \{1, \ldots, n\}$ (in particular, the statement for $x_1 = \cdots = x_n$ implies that the input does not contain constraints of the form $E(x, x)$).
- There are no constraints of the form $x_1 \neq x_1$.
- There are no constraints of the form $E(x_1, x_2)$ and $\hat{E}(x_1, x_2)$.

Since n is fixed, it is clear that these conditions can be checked in polynomial time.

Now let $f \in \mathrm{Pol}(\Gamma)$ be a canonical binary injection of behavior min. Each of the type conditions $f(N,=) = E$ and $f(=,N) = E$ is impossible, because they would introduce a K_n. Further, $f(E,=) = N$ or $f(=,E) = N$ for the same reason. But then $g(x,y) := f(f(x,y), f(y,x))$ is of behavior min and N-dominated and therefore an embedding from $(H_n,E)^2$ into (H_n,E). Hence, CSP(Γ) is in P by Theorem 4.5. ∎

5. Summary for the Henson graphs.

5.1. Proof of the complexity dichotomy. We are ready to assemble our results to prove the dichotomy for the CSPs of reducts of Henson graphs.

Proof of Theorem 1.1. Let Γ be a reduct of (H_n,E). If End(Γ) contains a function whose image is an independent set, then CSP(Γ) equals the CSP for a reduct of $(H_n,=)$ by Lemma 3.2, and such CSPs are either in P or NP-complete [BK08]. Otherwise, End(Γ) $= \overline{\mathrm{Aut}(H_n,E)}$ by Proposition 3.1. Lemma 3.3 shows that E, N, and $\neq$ are pp-definable in Γ.

If the relation H is also pp-definable in Γ, then CSP(Γ) is NP-hard by Proposition 4.4 (or Proposition 4.1); it is in NP since Γ is a reduct of (H_n,E), which is a finitely bounded homogeneous structure.

So let us assume that H is not pp-definable in Γ; then Proposition 3.9 shows that Pol(Γ) contains a canonical binary injection f of behavior min. Hence, CSP(Γ) is in P by Proposition 4.6. ∎

5.2. Discussion. We can restate Theorem 1.1 in a more detailed fashion as follows.

THEOREM 5.1. *Let Γ be a reduct of a Henson graph (H_n,E). Then one of the following holds.*

(1) Γ *has an endomorphism inducing an independent set and is homomorphically equivalent to a reduct of* $(H_n,=)$.
(2) Pol(Γ) *has a uniformly continuous projective clone homomorphism.*
(3) Pol(Γ) *contains a binary canonical injection which is of behavior* min *and* N*-dominated.*

Items (2) *and* (3) *cannot simultaneously hold, and when Γ has a finite relational signature,* (2) *implies NP-completeness and* (3) *implies tractability of its CSP.*

The first statement of Theorem 5.1 follows directly from the proof of Theorem 1.1, with the additional observation that the strong h1 clone homomorphism defined in Proposition 4.4 is in fact a clone homomorphism. When (3) holds for a reduct, (2) cannot hold, because (3) implies the existence of $f(x,y) \in \mathrm{Pol}(\Gamma)$ and $\alpha \in \overline{\mathrm{Aut}(\Gamma)}$ satisfying the equation $f(x,y) = \alpha f(y,x)$, an equation impossible to satisfy by projections. In fact, by further analyzing case (1), using what is known about reducts of equality, one can easily show that it also implies either (2) or (3), so that we have the following.

COROLLARY 5.2. *For every reduct Γ of a Henson graph (H_n,E), precisely one of the following holds.*

- Pol(Γ) *has a uniformly continuous projective clone homomorphism.*
- Pol(Γ) *contains* $f(x,y) \in \mathrm{Pol}(\Gamma)$ *and* $\alpha \in \overline{\mathrm{Aut}(\Gamma)}$ *such that* $f(x,y) = \alpha f(y,x)$.

When Γ has a finite relational signature, the first case implies NP-completeness and the second case implies tractability of its CSP.

6. Polymorphisms over homogeneous equivalence relations. We now investigate polymorphisms of reducts of the graphs (C_n^s,E) for $2 \le n, s \le \omega$, with precisely one of n, s equal to ω. Recall from section 2 that we write Eq for the reflexive closure of E, that Eq is an equivalence relation with n classes of size s, and that we denote its equivalence classes by C_i for $0 \leq i < n$.

Similarly to the case of the Henson graphs, we start with unary polymorphisms in section 6.1, reducing the problem to model-complete cores.

We then turn to higher-arity polymorphisms; here, the organization somewhat differs from the case of the Henson graphs. The role of the NP-hard relation H from the Henson graphs is now taken by the two sources of NP-hardness mentioned in the introduction: the first source is that factoring by the equivalence relation Eq yields a structure with an NP-hard problem, and the second source is that restriction to some equivalence class yields a structure with an NP-hard problem. In section 6.2, we show that, in fact, one of the two sources always applies for model-complete cores when $2 < n < \omega$ or $2 < s < \omega$. Consequently, only the higher-arity polymorphisms of the reducts of (C_2^ω, E) and (C_ω^2, E) require deeper investigation using Ramsey theory; this will be dealt with in sections 6.3 and 6.4, respectively.

6.1. The unary case: Model-complete cores.

PROPOSITION 6.1. *Let Γ be a reduct of (C_n^s, E), where $1 \leq n, s \leq \omega$, and at least one of n, s equals ω. Then either* $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(\Gamma)} = \overline{\mathrm{Aut}(C_n^s, E)}$ *or* $\mathrm{End}(\Gamma)$ *contains an endomorphism onto a clique or an independent set.*

Proof. Assume that $\mathrm{End}(\Gamma) \neq \overline{\mathrm{Aut}(C_n^s, E)}$, so there is an endomorphism f of Γ violating either E or N.

Case 0. If $n = 1$ or $s = 1$, then the statement is trivial.

Case 1. If $n = s = \omega$, and so Eq has infinitely many infinite classes, then we can refer to [BW12].

Case 2. Assume that $1 < n < \omega$ and $s = \omega$.

Suppose that f violates Eq and preserves N; then clearly, iterating applications of automorphisms of (C_n^ω, E) and f, we could send any finite subset of C_n^ω to an independent set in (C_n^ω, E), contradicting that the number of equivalence classes is the fixed finite number n.

If f preserves both Eq and N, then there exist a, b with $E(a, b)$ and $f(a) = f(b)$. Via a standard iterative argument using topological closure, one then sees that f generates a function whose range is an independent set.

Therefore, it remains to consider the case where f violates N. Fix $u, v \in C_n^\omega$ with $N(u, v)$ and $Eq(f(u), f(v))$. Without loss of generality we may assume $u \in C_0$ and $v \in C_1$. By Proposition 2.10, we may assume that f is canonical as a function from $(C_n^\omega, E, \prec, u, v)$ to $(C_n^\omega, E, \prec)$. Clearly, f must preserve Eq on each class C_i with $i > 1$, as otherwise canonicity would imply the existence of an infinite independent set in (C_n^ω, E). For the same reason, f preserves Eq on each of the four sets

$$\begin{aligned} C_0^- &:= \{a \in C_0 \mid a \prec u\}\ , \\ C_0^+ &:= \{a \in C_0 \mid u \prec a\}\ , \\ C_1^- &:= \{a \in C_1 \mid a \prec v\}\ , \\ \text{and}\ \ C_1^+ &:= \{a \in C_1 \mid v \prec a\}. \end{aligned}$$

If N is not preserved between two sets among $S := \{C_0^-, C_0^+, C_1^-, C_1^+, C_2, C_3, \ldots\}$, then we pick these two sets along with $n-2$ further sets from S belonging to distinct equivalence classes. The union of this collection induces a copy of (C_n^ω, E) on which

f preserves Eq but not N, and a standard iterative argument shows that f generates a function whose range is contained in a single equivalence class. Hence, we may assume that N is preserved between any two sets in S. Since n is finite, this is only possible if Eq is preserved on $C_0^- \cup C_0^+$ and on $C_1^- \cup C_1^+$. By composing f with an automorphism of (C_n^ω, E), we may thus assume that $f[C_i^- \cup C_i^+] \subseteq C_i$ for $i \in \{0,1\}$ and that f preserves the classes C_i for $i > 1$. Either $f(u) \notin C_0$ or $f(v) \notin C_1$. Assume without loss of generality that $f(u) \in C_i$ where $i > 0$. Let e be a self-embedding of (C_n^ω, E) with range $C_n^\omega \setminus \{v\}$. Then $f \circ e$ preserves all equivalence classes except for the element u, which it moves from C_0 to C_i. Iterating applications of $f \circ e$ and automorphisms, and using topological closure, we obtain a function which joins C_0 and C_i. By further iteration, we obtain a function which joins all classes.

Case 3. Assume that $s < \omega$ and $n = \omega$.

Suppose that f violates N and preserves Eq; then, by topological closure, f generates a mapping onto a clique. If it preserves both Eq and N, then, as above, f generates a function whose range is an independent set.

Therefore, we may assume that f violates Eq. Fix $u, v \in C_\omega^s$ with $E(u,v)$ such that $N(f(u), f(v))$. By Proposition 2.11, we may assume that f is canonical as a function from $(C_\omega^s, E, \prec, u, v)$ to $(C_\omega^s, E, \prec)$. If f preserves N, then by topological closure f generates a function whose range induces an independent set. Otherwise, there exist $a, b \in C_\omega^s$ with $N(a,b)$ and $Eq(f(a), f(b))$. Without loss of generality, a is not contained in the class of u and v. Then $\{a' \in C_\omega^s \mid \operatorname{tp}(a', b) = \operatorname{tp}(a,b) \text{ in } (C_\omega^s, E, \prec, u, v)\}$ contains an infinite independent set S. By canonicity, we have $Eq(f(a'), f(b))$ for all $a' \in S$, so that S is mapped into a single class. Since this class is finite, there exist $a', a'' \in S$ with $f(a') = f(a'')$, and so by topological closure, we can generate a function from f whose range is contained in a single equivalence class. ∎

If the second case of Proposition 6.1 applies to a reduct Γ of (C_n^s, E), then Γ is homomorphically equivalent to a reduct of equality, and its CSP is understood. In the following sections, we investigate essential polymorphisms of reducts Γ of (C_n^s, E) satisfying $\operatorname{End}(\Gamma) = \overline{\operatorname{Aut}(\Gamma)} = \overline{\operatorname{Aut}(C_n^s, E)}$. In particular, such reducts are model-complete cores. The following proposition implies that in the situation where $2 < s$ the equivalence relation Eq is invariant under $\operatorname{Pol}(\Gamma)$.

PROPOSITION 6.2. *Let* Γ *be a reduct of* (C_n^s, E)*, where* $1 \leq n \leq \omega$ *and* $2 < s \leq \omega$. *If* $\operatorname{End}(\Gamma) = \overline{\operatorname{Aut}(C_n^s, E)}$*, then* E*,* N*, and* Eq *are preserved by* $\operatorname{Pol}(\Gamma)$.

Proof. By Lemma 2.3, the condition $\operatorname{End}(\Gamma) = \overline{\operatorname{Aut}(C_n^s, E)}$ implies that all polymorphisms of Γ preserve E and N, and hence also Eq since $Eq(x,y)$ has the pp definition $\exists z\, (E(x,z) \wedge E(z,y))$. Note that we need the classes to contain at least three elements for this definition to work. ∎

If $s = 1$, then Eq is pp-definable as equality, but if $s = 2$, then Eq is not in general pp-definable; this will account for an additional nontrivial (tractable) case in our analysis.

Since, in the situation of Proposition 6.2, Eq is an equivalence relation which is invariant under $\operatorname{Pol}(\Gamma)$, it follows that $\operatorname{Pol}(\Gamma)$ acts naturally on the equivalence classes of Eq: for $f(x_1, \ldots, x_n) \in \operatorname{Pol}(\Gamma)$ and classes $C_{i_1}, \ldots, C_{i_n}$ of Eq, the class $f(C_{i_1}, \ldots, C_{i_n})$ is then defined as the equivalence class of $f(c_{i_1}, \ldots, c_{i_n})$, where $c_{i_1} \in C_{i_1}, \ldots, c_{i_n} \in C_{i_n}$ are arbitrary.

Moreover, if we fix any class C of Eq and expand the structure Γ by the predicate C to a structure (Γ, C), then $\operatorname{Pol}(\Gamma, C)$ acts naturally on C via restriction of its functions. Since $\operatorname{Aut}(C_n^s, E)$ can flip any two equivalence classes, all such actions are isomorphic; i.e., for any two classes C, C' there exists a bijection $i\colon C \to C'$ such that

$$\mathrm{Pol}(\Gamma, C') = \{(x_1, \ldots, x_n) \mapsto i(f(i^{-1}(x_1), \ldots, i^{-1}(x_n))) \mid f \in \mathrm{Pol}(\Gamma, C)\},$$
$$\mathrm{Pol}(\Gamma, C) = \{(x_1, \ldots, x_n) \mapsto i^{-1}(f(i(x_1), \ldots, i(x_n))) \mid f \in \mathrm{Pol}(\Gamma, C')\}$$

(in fact any bijection i works, since any permutation on C extends to an automorphism of (C_n^s, E) which fixes the elements of C' pointwise). It is for this reason that in the following, it will not matter if we make statements about all such actions or a single action.

In the following sections, we analyze these two types of actions.

6.2. The case $2 < n < \omega$ or $2 < s < \omega$. It turns out that in these cases, one of the two types of actions always yields hardness of the CSP. We are going to use the following fact about function clones on a finite domain.

PROPOSITION 6.3 (from [HR94]). *Every function clone on a finite domain of at least three elements which contains all permutations as well as an essential function contains a unary constant function.*

We can immediately apply this fact to the action of $\mathrm{Pol}(\Gamma)$ on the equivalence classes, when there are more than two but finitely many classes.

PROPOSITION 6.4. *Let Γ be a reduct of (C_n^ω, E), where $2 < n < \omega$, such that* $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C_n^\omega, E)}$. *Then the action of* $\mathrm{Pol}(\Gamma)$ *on the equivalence classes of Eq has no essential and no constant operation.*

Proof. The action has no constant operation because N is preserved. Therefore, it cannot have an essential operation either, by Proposition 6.3. ∎

Similarly, we can apply the same fact to the action of $\mathrm{Pol}(\Gamma, C)$ on any equivalence class C on C_ω^s if this class is finite and has more than two elements.

PROPOSITION 6.5. *Let Γ be a reduct of (C_ω^s, E), where $2 < s < \omega$, such that* $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C_\omega^s, E)}$. *Then for any equivalence class C of Eq, the action of* $\mathrm{Pol}(\Gamma, C)$ *on C has no essential and no constant operation.*

Proof. The action has no constant operation because E is preserved. Therefore, it cannot have an essential operation either, by Proposition 6.3. ∎

6.3. The case of two infinite classes: $n = 2$ and $s = \omega$. The following proposition states that either one of the two sources of hardness applies, or $\mathrm{Pol}(\Gamma)$ contains a ternary canonical function with a certain behavior.

PROPOSITION 6.6. *Let Γ be a reduct of (C_2^ω, E) such that* $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C_2^\omega, E)}$. *Then one of the following holds:*

- *the action of* $\mathrm{Pol}(\Gamma)$ *on the equivalence classes of Eq has no essential function;*
- *the action of* $\mathrm{Pol}(\Gamma, C)$ *on some (or any) class C has no essential function;*
- $\mathrm{Pol}(\Gamma)$ *contains a canonical ternary injection of behavior minority which is hyperplanely of behavior balanced xnor.*

To prove the proposition, we need to recall a special case of Post's classical result about function clones acting on a two-element set. Comparing this statement with Proposition 6.3 sheds light on why the case of this section is more involved than the cases of the preceding section.

PROPOSITION 6.7 (Post [Pos41]). *Every function clone with domain $\{0, 1\}$ containing both permutations of $\{0, 1\}$ as well as an essential function contains a unary constant operation or the ternary addition modulo 2.*

We moreover require the following result on polymorphism clones on a countable set.

PROPOSITION 6.8 (from [BK08]). *Every polymorphism clone on a countably infinite set which contains all permutations as well as an essential operation contains a binary injection.*

We now combine these two results into a proof of Proposition 6.6.

Proof of Proposition 6.6. Recall that the equivalence classes of Eq are denoted by C_0 and C_1, and that E, N, and Eq are preserved by the functions of $\mathrm{Pol}(\Gamma)$, by Proposition 6.2. Suppose that the first statement of the proposition does not hold. Then by Proposition 6.7, the action of $\mathrm{Pol}(\Gamma)$ on $\{C_0, C_1\}$ contains a unary constant operation, or a function which behaves like ternary addition modulo 2. The first case is impossible since the unary functions in $\mathrm{Pol}(\Gamma)$ preserve N, so the latter case holds and $\mathrm{Pol}(\Gamma)$ contains a ternary function g which acts like $x + y + z$ modulo 2 on the classes.

Suppose now in addition that the second statement of the proposition does not hold either, and fix some equivalence class C. Since the action of $\mathrm{Pol}(\Gamma, C)$ on C contains all permutations of C, by Proposition 6.8 it also contains a binary injection. Therefore, $\mathrm{Pol}(\Gamma)$ contains for each $i \in \{0, 1\}$ a binary function f_i whose restriction to C_i is an injection on this set.

We claim that there is a single function $f \in \mathrm{Pol}(\Gamma)$ which has this property for both C_0 and C_1. Note that since N is preserved by f_0, it maps C_1 into itself. If f_0 is essential on C_1, then Proposition 6.8 implies that together with all permutations which fix the classes, it generates a function which is injective on C_1; this function is then injective on both classes C_0, C_1. So assume that f_0 is not essential on C_1; say without loss of generality that it depends only on the first coordinate (and injectively so, since it preserves E). Then $f_0(f_1(x, y), f_0(x, y))$ preserves both classes and is injective on each of them.

By Proposition 2.9, we may assume that f is canonical as a function from $(C_2^\omega, E, \prec) \times (C_2^\omega, E, \prec)$ to $(C_2^\omega, E, \prec)$. We claim that f is also canonical as a function from $(C_2^\omega, E) \times (C_2^\omega, E)$ to (C_2^ω, E). To prove this, it suffices to show that if $u, v, u', v' \in C_2^\omega \times C_2^\omega$ are so that (u, v) and (u', v') have the same type in $(C_2^\omega, E) \times (C_2^\omega, E)$, then $(f(u), f(v))$ and $(f(u'), f(v'))$ have the same type in (C_2^ω, E). There exist $u'', v'' \in C_2^\omega \times C_2^\omega$ such that (u', v') and (u'', v'') have the same type in $(C_2^\omega, E, \prec) \times (C_2^\omega, E, \prec)$ and such that $EqEq(u, u'')$ and $EqEq(v, v'')$; by the canonicity of f as a function from $(C_2^\omega, E, \prec) \times (C_2^\omega, E, \prec)$ to $(C_2^\omega, E, \prec)$, it suffices to show that $(f(u), f(v))$ and $(f(u''), f(v''))$ have the same type in (C_2^ω, E). Since Eq is preserved, we have $Eq(f(u), f(u''))$ and $Eq(f(v), f(v''))$, and so $Eq(f(u), f(v))$ implies $Eq(f(u''), f(v''))$ and vice-versa, by the transitivity of Eq. Failure of canonicity can therefore only happen if $Eq(f(u), f(v))$ and $Eq(f(u''), f(v''))$, and precisely one of $f(u) = f(v)$ and $f(u'') = f(v'')$ holds, say without loss of generality the former. But then picking any $v''' \in C_2^\omega \times C_2^\omega$ distinct from v such that $EqEq(v, v''')$ and such that the type of (u, v) equals the type of (u, v''') in $(C_2^\omega, E, \prec) \times (C_2^\omega, E, \prec)$ shows that $f(v) = f(u) = f(v''')$ by canonicity, contradicting the fact that f is injective on each equivalence class.

We analyze the behavior of the canonical function $f : (C_2^\omega, E) \times (C_2^\omega, E) \to (C_2^\omega, E)$. Because E and N are preserved, we have $f(E, E) = E$ and $f(N, N) = N$. Moreover, because f is injective on the classes, and because Eq is preserved, we have $f(=, E) = f(E, =) = E$.

We next claim that either $f(\cdot, N) = N$ or $f(N, \cdot) = N$. Otherwise, there exist

$Q, P \in \{E, =\}$ such that $f(Q, N) \neq N$ and $f(N, P) \neq N$. Pick $u, v, w \in (C_2^\omega)^2$ such that $QN(u, v), NP(v, w)$, and $NN(u, w)$. Then $Eq(f(u), f(w))$ and $N(f(u), f(w))$, a contradiction.

Assume henceforth without loss of generality that $f(N, \cdot) = N$. Then $f(P, N) \neq N$ for $P \neq N$, because there are only two equivalence classes. Moreover, $f(E, N) = \,=$ or $f(=, N) = \,=$ would imply that f is not injective on the classes, so we have $f(E, N) = f(=, N) = E$.

Summarizing, f is a binary injection of behavior p_1, balanced in the first argument, and E-dominated in the second argument.

Let $q \in \mathrm{Pol}(\Gamma)$ be any ternary injection (for example, $(x, y, z) \mapsto f(x, f(y, z))$), and set $h(x, y, z) := f(g(x, y, z), q(x, y, z))$. We now show that h is canonical by establishing all type conditions satisfied by it. To this end, we use the behavior of f and the fact that g acts like $x+y+z$ modulo 2 on the classes. The latter fact implies that g satisfies certain type conditions as well, as is easily verified: $g(Eq, Eq, N) = g(Eq, N, Eq) = g(N, Eq, Eq) = N$, $g(Eq, Eq, Eq) = Eq$, and moreover $g(Eq, N, N) = Eq$, $g(N, Eq, N) = Eq$, and $g(N, N, Eq) = Eq$. In the following table, $u, v, w \in (C_2^\omega)^2$ are three pairs for which $===(u, v, w)$ does not hold, and according to the type of (u, v, w) in $(C_2^\omega, E) \times (C_2^\omega, E)$ the type of $h(u, v, w)$ in (C_2^ω, E) is computed. By the symmetry of the type conditions of g listed above, and since of q we only use injectivity so that $\neq(q(u, v, w))$ holds, the value of a triple of types does not change if its components are permuted. Therefore, we only list all possibilities of types for (u, v, w) up to permutations.

$\mathrm{tp}(u, v, w)$	$\mathrm{tp}(g(u, v, w), q(u, v, w))$	$\mathrm{tp}(h(u, v, w))$
EEE	$(E, \neq)$	E
NNN	$(N, \neq)$	N
EEN	$(N, \neq)$	N
ENN	$(Eq, \neq)$	E
$=EE$	$(Eq, \neq)$	E
$=NN$	$(Eq, \neq)$	E
$=EN$	$(N, \neq)$	N
$==E$	$(Eq, \neq)$	E
$==N$	$(N, \neq)$	N

So h acts like a minority which is hyperplanely of behavior balanced xnor. □

6.4. The case of infinitely many classes of size two: $n = \omega$ and $s = 2$. Recall that in this situation, Proposition 6.2 does not apply, and Eq might not be pp-definable in a reduct Γ of (C_ω^2, E), even if Γ is a model-complete core. We first show that if this happens, then $\mathrm{Pol}(\Gamma)$ contains a certain binary canonical function (Proposition 6.9). We then show, in Proposition 6.10, that if Eq does have a primitive positive definition in Γ, then either one of the two sources of hardness applies, or $\mathrm{Pol}(\Gamma)$ contains a ternary function of a certain behavior.

PROPOSITION 6.9. *Let Γ be a reduct of (C_ω^2, E) such that $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C_\omega^2, E)}$, and such that Eq is not pp-definable. Then Γ enjoys a binary canonical polymorphism of behavior* min *which is N-dominated.*

Proof. By Theorem 2.2, Γ has a polymorphism f which violates Eq. By the assumption, all endomorphisms preserve E and N, and hence, by Lemma 2.3, so does f. By the same lemma, because Eq consists of two orbits with respect to the action of the automorphism group of (C_ω^2, E) on pairs, we may assume that f is binary.

We refer to sets of the form $C \times D$, where C, D are equivalence classes of Eq, as *squares*. Note that each square is the disjoint union of precisely two edges in the product graph $(C^2_\omega, E)^2$, and that each of these edges is mapped by f to an edge in (C^2_ω, E), since f preserves E. We say that f *splits* a square if it does not map this square into a single class; in this case, it necessarily maps it into two classes, by the previous observation.

By composing f with automorphisms from the inside, we may assume that f violates Eq on a square of the form $C \times C$. Writing $C = \{u, v\}$, we may invoke Proposition 2.11 and assume that f is canonical when viewed as a function from $(C^2_\omega, E, \prec, u, v) \times (C^2_\omega, E, \prec, u, v)$ to $(C^2_\omega, E, \prec)$. We set $S := C^2_\omega \setminus C$ and $S' := \{x \in S \mid x \prec u \wedge x \prec v\}$.

We now distinguish two cases to show the following.

Claim. f generates a binary function f' which still splits $C \times C$ and satisfies either $f'(N, \cdot) = N$ or $f'(\cdot, N) = N$.

Case 1. We first assume that f splits a square within $(S')^2$. Then, by canonicity, it splits all squares within $(S')^2$. In that case, the function $f(e(x), e(y))$, where e is a self-embedding from $(C^2_\omega, E, \prec)$ onto the structure induced therein by S', is canonical while splitting all squares. Replacing f by this function, we henceforth assume f to split all squares. The constants u, v which were introduced to witness the occurrence of a splitting will no longer be of importance to us in the further discussion of this case.

The function g on $(C^2_\omega)^2$ sending every pair (x, y) to the pair $(f(x, y), f(y, x))$ is canonical when viewed as a function

$$(C^2_\omega, E, \prec) \times (C^2_\omega, E, \prec) \to ((C^2_\omega)^2, EE, EN, NE, NN, E{=}, {=}E, N{=}, {=}N, \prec\prec)\,,$$

by the canonicity of f. In the following, we analyze the behavior of g.

We start by observing that every square consists of an *upward edge* and a *downward edge* in $(C^2_\omega, E, \prec)^2$, the orientation being induced by the order $\prec$: by the upward edge $(p, q) \in EE$ we refer to the one on which the order $\prec$ agrees in both coordinates between p and q, and by the downward one we refer to the other edge in the square (on which $\prec$ disagrees between the coordinates). Let U be the set of points contained in an upward edge, and V the set of points contained in a downward edge, so that $(C^2_\omega)^2$ is the disjoint union of U and V. We are going to verify the following properties of g:

(i) $g[U] \subseteq U$ and $g[V] \subseteq V$.
(ii) $E{=}(p, q)$, ${=}E(p, q)$, and $NN(p, q)$ all imply $NN(g(p), g(q))$ for all $p, q \in (C^2_\omega)^2$.
(iii) $NN(g(p), g(q))$ for all $p \in U$ and all $q \in V$.
(iv) On U as well as on V, either $f(N, \cdot) = N$ or $f(\cdot, N) = N$ holds.

Of property (i), we give the argument that $g[U] \subseteq U$; proving $g[V] \subseteq V$ is similar. Let $p = (p_1, p_2)$, $q = (q_1, q_2) \in (C^2_\omega)^2$ be so that (p, q) forms an upward edge, and say that $p_1 \prec q_1$ and $p_2 \prec q_2$. If $f(p_1, p_2) \prec f(q_1, q_2)$, then by canonicity also $f(p_2, p_1) \prec f(q_2, q_1)$, and so $(g(p), g(q))$ is related by $\prec$ in both coordinates. Since f preserves E, $(g(p), g(q))$ is also related by E in both coordinates, and hence $(g(p), g(q))$ forms an upward edge. If $f(q_1, q_2) \prec f(p_1, p_2)$, then a similar argument shows that $(g(p), g(q))$ forms an upward edge.

Property (ii) follows since f preserves N and because f splits all squares.

For (iii), suppose that $NN(g(p), g(q))$ does not hold for some $p \in U$ and $q \in V$. We cannot have $EE(p, q)$ since p is contained in an upward edge and q is contained in a downward edge, and so by (ii), p and q must be related by N in one coordinate.

Say we have $N{=}(p,q)$; the other situations are handled similarly. Pick $q' \in V$ distinct from q such that the types of (p,q) and (p,q') in $(C^2_\omega, E, \prec) \times (C^2_\omega, E, \prec)$ coincide. Then, by canonicity, we have that $g(p), g(q)$ are equivalent with respect to Eq in the same coordinate as $g(p), g(q')$; hence, so are $g(q), g(q')$, by the transitivity of Eq. By canonicity, we then know that for the unique $q'' \in V$ with $E{=}(p,q'')$, we have that $g(q)$ and $g(q'')$ are equivalent in that very same coordinate, since the types of (q,q') and either (q,q'') or (q'',q) agree. Again by transitivity, $g(p), g(q'')$ are then equivalent in that coordinate, contradicting (ii).

To see property (iv), suppose that neither $f(N,\cdot) = N$ nor $f(\cdot,N) = N$ hold on U. Then there exist $p,q,p',q' \in U$ such that p,q are related by N in the first coordinate, p',q' are related by N in the second coordinate, and $Eq(f(p),f(q))$ and $Eq(f(p'),f(q'))$ hold. But then we could pick $q'' \in U$ such that $\mathrm{tp}(p,q'') = \mathrm{tp}(p',q')$ in $(C^2_\omega, E, \prec) \times (C^2_\omega, E, \prec)$; any such q'' necessarily satisfies $NN(q,q'')$. By canonicity we would have $Eq(f(p),f(q''))$, and hence, by transitivity, this would imply $Eq(f(q),f(q''))$, a contradiction since f preserves N.

Now suppose that $f(N,\cdot) = N$ on both U and V. Then the function $f'(x,y) := f(g(x,y)) = f(f(x,y),f(y,x))$ has the same property by (iii), and moreover it splits all squares, and so we are done. If $f(\cdot,N) = N$ on both U and V, then by symmetry $f'(x,y) := f(f(y,x),f(x,y))$ has the same property everywhere and splits all squares. It remains to consider the case where, say, $f(N,\cdot) = N$ on U and $f(\cdot,N) = N$ on V. The function $f'(x,y) := f \circ g$ satisfies $f'(N,\cdot) = N$. To see this, let $p,q \in (C^2_\omega)^2$ be related by N in the first coordinate. If $p,q \in U$, then $g(p), g(q)$ are related by N in the first coordinate, and because $g[U] \subseteq U$, we have $N(f(g(p)), f(g(q)))$. When $p \in U$ and $q \in V$, then $NN(g(p),g(q))$ by (iii), and so $N(f'(p),f'(q))$ since f preserves N. Finally, if $p,q \in V$, then $g(p), g(q)$ are related by N in the second coordinate, and using $g[V] \subseteq V$, we see that $N(f'(p),f'(q))$. Since f' moreover splits all squares by (ii), we are done.

Case 2. Assume now that f does not split any square within $(S')^2$. We claim that $f(N,\cdot) = N$ or $f(\cdot,N) = N$ on $(S')^2$; otherwise, there would exist $p,q,p',q' \in (S')^2$ such that p,q are related by N in the first coordinate, p',q' are related by N in the second coordinate, and $Eq(f(p),f(q))$ and $Eq(f(p'),f(q'))$ hold. But then we could pick $q'' \in (S')^2$ such that $\mathrm{tp}(p,q'') = \mathrm{tp}(p',q')$ in $(C^2_\omega, E, \prec, u, v) \times (C^2_\omega, E, \prec, u, v)$; any such q'' necessarily satisfies $NN(q,q'')$. By canonicity we would have $Eq(f(p),f(q''))$, and hence, by transitivity, this would imply $Eq(f(q),f(q''))$, a contradiction since f preserves N. We assume without loss of generality that $f(N,\cdot) = N$ on $(S')^2$.

The function $f(e(x),e(y))$, where e is a self-embedding from (C^2_ω, E, u, v) onto the structure induced therein by $S' \cup \{u,v\}$, still splits $C \times C$, splits no square within $S \times S$, and satisfies $f(N,\cdot) = N$ on S^2. Invoking Proposition 2.11 again, we may assume that that function is moreover canonical as a function from $(C^2_\omega, E, \prec, u, v) \times (C^2_\omega, E, \prec, u, v)$ to $(C^2_\omega, E, \prec)$. Replacing f by this function, we may therefore henceforth assume that f itself enjoys the listed properties.

Note that the function $f(e(x),e(y))$ as in the preceding paragraph does not distinguish between elements of S' and those of $S \setminus S'$, since e sends the entire set S into S' before f is applied. In particular, it has the property that for any $p \in C \times C$ and any $q \in (C^2_\omega)^2$, the type of $(f(p),f(q))$ in (C^2_ω, E) only depends on the type of (p,q) in $(C^2_\omega, E, u, v) \times (C^2_\omega, E, u, v)$ and not on the more precise type of (p,q) in $(C^2_\omega, E, u, v, \prec) \times (C^2_\omega, E, u, v, \prec)$ (which does distinguish between S' and $S \setminus S'$).

We now distinguish two subcases to show that f generates a binary function f' which splits $C \times C$ and such that $f'(N,\cdot) = N$ everywhere, thus proving the claim.

Case 2.1. If $f(N,\cdot) = N$ on $S \times C$, then by canonicity and the remark above, one easily concludes $N(f(p), f(q))$ for all $p \in C \times C$ and all $q \in S \times C$, so that altogether $f(N,\cdot) = N$ everywhere. Hence, setting $f' := f$, we have achieved our goal.

Case 2.2. If $f(N,\cdot) = N$ does not hold on $S \times C$, then there exists $c \in S \times C$ such that $N(f(c), f(q))$ for any $q \in S^2$. To see this, we can pick any $c \in S \times C$ so that there exists $q' \in S \times C$ related to c by N in the first coordinate and such that $Eq(f(c), f(q'))$. Then, if there existed $q \in S^2$ with $Eq(f(c), f(q))$, we would have $Eq(f(q), f(q'))$; replacing q' by $q'' \in S \times C$ such that $\mathrm{tp}(c, q') = \mathrm{tp}(c, q'')$ in $(C^2_\omega, E, \prec, u, v)$ and such that q', q'' are related by N in both coordinates would yield a contradiction to the preservation of N.

We are going to check the following properties of the function g on $(C^2_\omega)^2$ defined by $(x, y) \mapsto (x, f(x, y))$.

(i) Whenever $p, q \in (C^2_\omega)^2$ are related by N in the first coordinate, then so are $g(p), g(q)$.

(ii) If $p \in (C^2_\omega)^2$, and $q \in S^2$ is related to p by N in the first coordinate, then $NN(g(p), g(q))$.

(iii) Writing $a := (u, u)$ and $b := (v, u)$, we have $E{=}(a, b)$ and $EN(g(a), g(b))$.

Property (i) is obvious from the definition of g. Property (ii) is clear if $p \in C^2_\omega \times C$, since in that case $NN(p, q)$ and since f preserves N. If $p \in S^2$, then it follows from the fact that $f(N,\cdot) = N$ on S. Finally, consider the case where $p \in C \times S$. If we had $Eq(f(p), f(q))$, then picking $q' \in S^2$ such that $N{=}(q, q')$ and such that $\mathrm{tp}(p, q) = \mathrm{tp}(p, q')$ in $(C^2_\omega, E, \prec) \times (C^2_\omega, E, \prec)$, we would get $Eq(f(p), f(q'))$ by canonicity, and so $Eq(f(q), f(q'))$, contradicting that $f(N,\cdot) = N$ on S. Property (iii) just restates that f splits $C \times C$.

Let e_1, e_2 be self-embeddings of (C^2_ω, E) such that the range of $(e_1, e_2) \circ g$ is contained in $S \times C^2_\omega$ and such that $(e_1, e_2) \circ g(a) = c$. Using that assumption, $g' := g \circ (e_1, e_2) \circ g$ clearly also satisfies (i) and (ii). Moreover, since $(e_1, e_2) \circ g(a) = c$, and since $EN((e_1, e_2) \circ g(a), (e_1, e_2) \circ g(b))$, we have $(e_1, e_2) \circ g(b) \in S^2$; this implies $EN(g'(a), g'(b))$, since $N(f(c), f(q))$ for all $q \in S^2$. Hence, g' still satisfies (iii).

We then pick a pair (e'_1, e'_2) of self-embeddings of (C^2_ω, E) with $(e'_1, e'_2) \circ g'(b) = c$, and consequently $(e'_1, e'_2) \circ g'(a) \in S^2$. Then $g'' := g \circ (e'_1, e'_2) \circ g' = g \circ (e'_1, e'_2) \circ g \circ (e_1, e_2) \circ g$ has the property that whenever $p, q \in (C^2_\omega)^2$ are related by N in the first coordinate, then $NN(g''(p), g''(q))$; this is because every point went through S^2 in one of the applications of g, and because of (ii). Moreover, we have $EN(g''(a), g''(b))$.

Setting f' to be the projection of g'' onto the second coordinate then completes the proof.

Wrap-up. Replacing f by f' from the claim, we thus henceforth assume that $f(N,\cdot) = N$. For the function h on $(C^2_\omega)^2$ defined by $(x, y) \mapsto (f(x, y), f(y, x))$, we are going to prove the following properties.

(i) If $p, q \in (C^2_\omega)^2$ are related by N in some coordinate, then $h(p), h(q)$ are related by N in the same coordinate.

(ii) There are $p', q' \in (C^2_\omega)^2$ with $E{=}(p', q')$ such that $h(p'), h(q')$ are related by N in the first coordinate.

(iii) There are $p'', q'' \in (C^2_\omega)^2$ with $EN(p'', q'')$ such that $NN(h(p''), h(q''))$.

(iv) There are $p''', q''' \in (C^2_\omega)^2$ with ${=}N(p'', q'')$ such that $NN(h(p'''), h(q'''))$.

Property (i) is obvious because $f(N,\cdot) = N$, and (ii) follows because f splits a square. To see (iii), we first observe that there exist $p, q \in (C^2_\omega)^2$ with equal first coordinate and such that $h(p), h(q)$ are related by N in the first coordinate: simply pick p, p' with ${=}E(p, p')$ within the square that is split; then $NN(h(p), h(p'))$, and so for any $q \in (C^2_\omega)^2$ with ${=}N(p, q)$ and ${=}N(p', q)$ we have that $h(q)$ must be related by N in the first coordinate to either $h(p)$ or $h(p')$, showing the observation. Now fix p, q with this property, and pick $v \in (C_\omega^2)^2$ with $EN(p,v)$ and $EN(q,v)$. Then $h(v)$ is related to $h(p)$ and $h(q)$ by N in the second coordinate by (i), but also necessarily to one of them in the first coordinate, showing (iii). The proof of (iv) is similar.

Using these properties, we first construct, by composition and topological closure, a function h' on $(C_\omega^2)^2$ which yields $NN(h'(p), h'(q))$ for all $p, q \in (C_\omega^2)^2$ which are related by N in at least one coordinate. To do this, let $\{(p_i, q_i) \mid i > 0\}$ be an enumeration of all pairs in $(C_\omega^2)^2$ which are related by N in at least one coordinate. We proceed inductively, constructing functions $h_0, h_1, \ldots$ with the property that $NN(h_n(p_j), h_n(q_j))$ for all $0 < j \leq n$, and $h_n(p_j)$ and $h_n(q_j)$ are related by N in at least one coordinate, for all $j > n$. For the base case, we set $h_0 := h$ (note that the first conjunct of the inductive hypothesis acts here on an empty set of pairs). Suppose we have already constructed h_n. Then $h_n(p_{n+1})$ and $h_n(q_{n+1})$ are related by N in at least one coordinate. If $NN(h_n(p_{n+1}), h_n(q_{n+1}))$, then we set $h_{n+1} := h_n$. If $EN(h_n(p_{n+1}), h_n(q_{n+1}))$, then let (α, β) be a pair of automorphisms of (C_ω^2, E) such that $(\alpha, \beta)(h_n(p_{n+1})) = p''$ (from (iii)), and $(\alpha, \beta)(h_n(q_{n+1})) = q''$. Setting $h_{n+1} := h \circ (\alpha, \beta) \circ h_n$ then yields the desired property for (p_{n+1}, q_{n+1}). If $NE(h_n(p_{n+1}), h_n(q_{n+1}))$, then $EN(\pi \circ h_n(p_{n+1}), \pi \circ h_n(q_{n+1}))$, where $\pi \colon (C_\omega^2)^2 \to (C_\omega^2)^2$ is defined by $(x, y) \mapsto (y, x)$ (i.e., π is a pair of projections); we can thus proceed as before. The cases $=N(h_n(p_{n+1}), h_n(q_{n+1}))$ and $N{=}(h_n(p_{n+1}), h_n(q_{n+1}))$ are treated similarly using (iv) instead of (iii). By topological closure, we obtain the function h'.

Setting $h'' := h' \circ h$, we retain the defining property of h' by (i) but moreover have $NN(h''(p'), h''(q'))$ for the pair (p', q') from (ii).

The function $g_0 := f \circ h''$ then satisfies $g_0(N, \cdot) = g_0(\cdot, N) = N$ and moreover satisfies $N(g(p'), g(q'))$, since $NN(h''(p'), h''(q'))$ and since f preserves N.

Let $\{(p_i, q_i) \mid i \geq 0\}$ be an enumeration of all pairs in $(C_\omega^2)^2$ related by $E{=}$, where $(p_0, q_0) = (p', q')$. As above, we obtain, by composition and topological closure, for every $i \geq 0$ a function g_i which satisfies $g_i(N, \cdot) = g_i(\cdot, N) = N$ and such that $N(g_i(p_i), g_i(q_i))$. Setting $t_0 := g_0$, and $t_{n+1} := f(t_n(x, y), g_{n+1}(x, y))$ for all $n \geq 0$, we obtain binary functions $t_0, t_1, \ldots$ satisfying $t_i(N, \cdot) = t_i(\cdot, N) = N$ and with the property that $N(t_i(p_j), t_i(q_j))$ for all $j \leq i$. By topological closure, we obtain a binary function t satisfying $t(N, \cdot) = t(\cdot, N) = N$ and $N(t(p), t(q))$ for all $p, q \in (C_\omega^2)^2$ with $E{=}(p, q)$. This function clearly has behavior min and is N-dominated in the first argument; since it preserves E, these properties also imply that it is N-dominated in the second argument. ∎

We now turn to the case where Eq is pp-definable in a reduct Γ so that $\mathrm{Pol}(\Gamma)$ acts on its equivalence classes.

PROPOSITION 6.10. *Let Γ be a reduct of (C_ω^2, E) such that $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C_\omega^2, E)}$ and such that Eq is pp-definable. Then one of the following holds:*

- *the action of $\mathrm{Pol}(\Gamma)$ on the equivalence classes of Eq has no essential function;*
- *the action of $\mathrm{Pol}(\Gamma, C)$ on some (or any) equivalence class of C has no essential function;*
- *$\mathrm{Pol}(\Gamma)$ contains a ternary canonical function h such that $h(N, \cdot, \cdot) = h(\cdot, N, \cdot) = h(\cdot, \cdot, N) = N$ which behaves like a minority on $\{E, =\}$ (so $h(E, =, =) = E$, etc.).*

To prove the proposition, we are again going to make use of Propositions 6.7 and 6.8, and the following lemma. We are going to say that a ternary function f on C_ω^2 behaves like $x + y + z$ modulo 2 on an equivalence class $C = \{0, 1\}$ of Eq if the restriction of f to C is of the form $\alpha \circ g_C$, where $\alpha \in \mathrm{Aut}(C^2_\omega, E)$ and g_C is the ternary function on C defined by $g_C(x, y, z) = x + y + z$ modulo 2. Note that this property can be expressed in terms of type conditions satisfied on C: namely, f behaves like $x + y + z$ modulo 2 on C if and only if it satisfies $f(E, E, E) = E$, $f(E, E, =) = f(E, =, E) = f(=, E, E) = =$, and $f(E, =, =) = f(=, =, E) = f(=, E, =) = E$ on C. In other words, f behaves like a minority on the types $\{E, =\}$.

LEMMA 6.11. *Let* Γ *be a reduct of* (C^2_ω, E) *such that* $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C^2_\omega, E)}$*,* Eq *is pp-definable, and* $\mathrm{Pol}(\Gamma)$ *contains a ternary function which behaves like* $x + y + z$ *modulo* 2 *on some equivalence class. Then* $\mathrm{Pol}(\Gamma)$ *contains a ternary function which behaves like* $x + y + z$ *modulo* 2 *on all equivalence classes.*

Proof. Let $C_0, C_1, \ldots$ be the equivalence classes of Eq. We show, by induction over n, that for all $n \in \omega$, $\mathrm{Pol}(\Gamma)$ contains a function g_n which equals $x+y+z$ modulo 2 on each class $C_0, \ldots, C_n$. The lemma then follows by a standard compactness argument: by ω-categoricity, there exist $\alpha_n \in \mathrm{Aut}(C^2_\omega, E)$ for $n \in \omega$, such that $(\alpha_n \circ g_n)_{n\in\omega}$ converges to a function $g \in \mathrm{Pol}(\Gamma)$ (cf., for example, the proof of Proposition 3.1). That function then has the desired property: for every $i \in \omega$, there exists $n > i$ such that g agrees with $\alpha_n \circ g_n$ on C_i, and hence it behaves like $x + y + z$ modulo 2 on C_i.

For the base case $n = 0$, the statement follows from the assumption of the lemma. Now suppose it holds for n. By the assumption that $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(D, E)}$, we may assume that $g_n(x, x, x) = x$ for all $x \in C_0 \cup \cdots \cup C_{n+1}$, and in particular g_n preserves each of the classes $C_0, \ldots, C_{n+1}$. In particular, the restriction of g_n to any C_i with $0 \leq i \leq n$ actually equals the function $x + y + z$ modulo 2 on that class.

Assume first that g_n is not essential on C_{n+1}; by composing it with an automorphism of (C^2_ω, E), we may assume it is a projection, without loss of generality to the first coordinate, on C_{n+1}. Let $g'_n \in \mathrm{Pol}(\Gamma)$ be a ternary function which has the properties of g_n, but with the roles of C_n and C_{n+1} switched. Such a function g'_n can be obtained by composing g_n in all arguments with the same automorphism that switches C_n and C_{n+1}. Then

$$g_{n+1}(x, y, z) := g_n(g'_n(x, y, z), g'_n(y, z, x), g'_n(z, x, y))$$

has the desired property.

Next assume that g_n is essential on C_{n+1}, and write g'_n for its restriction to C_{n+1}. Let $\alpha \in \mathrm{Aut}(C^2_\omega, E)$ flip the two elements of C_{n+1}, and fix all other elements of C^2_ω; then the restriction α' of α to C_{n+1} is the only nontrivial permutation of C_{n+1}. By Proposition 6.7, there exists a term $h'(x, y, z)$ over $\{g'_n, \alpha'\}$ which induces either a constant function or the function $x + y + z$ modulo 2 on C_{n+1}. The term $h(x, y, z)$ obtained from h' by replacing all occurrences of α' by α and all occurrences of g'_n by g_n induces a ternary function on C^2_ω whose restriction to C_{n+1} equals h'. Since h preserves E, it cannot be constant on C_{n+1}, and hence it is equal to $x+y+z$ modulo 2 on C_{n+1}. For each $0 \leq i \leq n$, since g_n equals $x + y + z$ modulo 2 on C_i, and since α is the identity on C_i, it is easy to see that the term function h, restricted to C_i, is of the form $\beta' \circ g$, where β' is a permutation on C_i and g equals either $x+y+z$ modulo 2 or a projection on C_i. Hence, iterating the preceding case, we obtain the desired function. ∎

Proof of Proposition 6.10. Suppose that neither of the first two items hold. Then by Proposition 6.8, $\mathrm{Pol}(\Gamma)$ contains a binary function f acting injectively on the classes of Eq; moreover, using Proposition 6.7 and since E is preserved, we see that $\mathrm{Pol}(\Gamma)$ contains a ternary function which equals $x+y+z$ modulo 2 on some equivalence class.

Hence, by Lemma 6.11 it contains a ternary function g which behaves like $x+y+z$ modulo 2 on all equivalence classes.

Observe first that since f acts injectively on the classes of Eq, we have that whenever $p,q \in (C^2_\omega)^2$ are not equivalent with respect to Eq in at least one coordinate, then $Eq(f(p),f(q))$ cannot hold. In other words, we have the type conditions $f(N,Eq) = f(Eq,N) = f(N,N) = N$.

We next argue that on each class C the operation f is essentially unary. Write $C = \{0,1\}$. Since E is preserved, we have $E(f(0,0),f(1,1))$; similarly, we know that $E(f(0,1),f(1,0))$. Since f moreover preserves Eq, the four values are contained in a single class. Hence either $f(0,1) = f(0,0)$ and $f(1,0) = f(1,1)$, or $f(1,0) = f(0,0)$ and $f(0,1) = f(1,1)$. In the first case, the restriction of f to C only depends on its first argument, and in the second case on its second argument. Assume without loss of generality that the former, i.e., $f(E,=) = E$ and $f(=,E) = =$, holds on infinitely many equivalence classes C. By precomposing f with self-embeddings of (C^2_ω, E) we may assume that f satisfies these type conditions everywhere. In particular, we then have that f is also canonical as a function from $(C^2_\omega,E)^2$ to (C^2_ω,E).

The function $q(x,y,z) := f(x,f(y,z))$ satisfies $q(N,\cdot,\cdot) = q(\cdot,N,\cdot) = q(\cdot,\cdot,N) = N$, and $q(P,Q,R) = P$ if $P,Q,R \in \{E,=\}$.

Consider the function t on $(C^2_\omega)^3$ which sends every triple (x,y,z) to the triple $(q(x,y,z),q(y,z,x),q(z,x,y))$. Then, whenever $P,Q,R \in \{E,=\}$ and $p,q \in (C^2_\omega)^3$ satisfy $PQR(p,q)$, then also $PQR(t(p),t(q))$, by the properties of q. Moreover, whenever $p,q \in (C^2_\omega)^3$ are related by N in at least one coordinate, then $NNN(t(p),t(q))$. By the latter property of t, there exist $\alpha,\beta,\gamma \in \overline{\mathrm{Aut}(C^2_\omega,E)}$ such that the function

$$(\alpha,\beta,\gamma)\circ t(x,y,z) := (\alpha(q(x,y,z)),\beta(q(y,z,x)),\gamma(q(z,x,y)))$$

sends any product $C_i \times C_j \times C_k$ of three equivalence classes into the cube C^3 of a single equivalence class; moreover, this function still has the properties of t mentioned above. Set $h(x,y,z) := g\circ(\alpha,\beta,\gamma)\circ t(x,y,z) = g(\alpha(q(x,y,z)),\beta(q(y,z,x)),\gamma(q(z,x,y)))$. Then $h(N,\cdot,\cdot) = h(\cdot,N,\cdot) = h(\cdot,\cdot,N) = g(N,N,N) = N$. We claim that h behaves like a minority on $\{E,=\}$. If $P,Q,R \in \{E,=\}$, then $h(P,Q,R) = g(P,Q,R)$. Since $(\alpha,\beta,\gamma)\circ t(x,y,z)$ maps the product of three equivalence classes into the cube of a single equivalence class, and since g behaves like $x+y+z$ modulo 2 on each equivalence class, the claim follows. ∎

7. Polynomial-time tractable CSPs over homogeneous equivalence relations.

We provide two polynomial-time algorithms: the first one is designed for the CSPs of reducts of (C^ω_2,E) with a ternary injective canonical polymorphism of behavior minority which is hyperplanely of behavior balanced xnor (section 7.1), and the second one for reducts of (C^2_ω,E) with a ternary canonical polymorphism h such that

$$h(N,\cdot,\cdot) = h(\cdot,N,\cdot) = h(\cdot,\cdot,N) = N$$

and which behaves like a minority on $\{=,E\}$ (section 7.2).

7.1. Two infinite classes.

We consider the case where Γ is a reduct of (C^ω_2,E) which is preserved by a canonical injection h of behavior minority which is hyperplanely of behavior balanced xnor (cf. Proposition 6.6). Our algorithm for CSP(Γ) is an adaptation of an algorithm for reducts of the random graph [BP15a].

We first reduce CSP(Γ) to the CSP of a structure that we call the *injectivization* of Γ, which can then be reduced to a tractable CSP over a Boolean domain.

Definition 7.1. *A tuple is called* injective *if all its entries are pairwise distinct. A relation is called* injective *if all its tuples are injective. A structure is called* injective *if all its relations are injective.*

Definition 7.2. *We define* injectivizations *for relations, atomic formulas, and structures.*

- *Let R be any relation. Then the* injectivization of R*, denoted by* $\mathrm{inj}(R)$*, is the (injective) relation consisting of all injective tuples of R.*
- *Let $\phi(x_1,\ldots,x_n)$ be an atomic formula in the language of Γ, where $x_1,\ldots,x_n$ is a list of the variables that appear in ϕ. Then the* injectivization of $\phi(x_1,\ldots,x_n)$ *is the formula $R^{\mathrm{inj}}_\phi(x_1,\ldots,x_n)$, where R^{inj}_ϕ is a relation symbol which stands for the injectivization of the relation defined by ϕ.*
- *The* injectivization *of a relational structure Γ, denoted by* $\mathrm{inj}(\Gamma)$*, is the relational structure with the same domain as Γ whose relations are the injectivizations of the atomic formulas over Γ, i.e., the relations R^{inj}_ϕ.*

To state the reduction to the CSP of an injectivization, we also need the following operations on instances of $\mathrm{CSP}(\Gamma)$. Here, it will be convenient to view instances of $\mathrm{CSP}(\Gamma)$ as primitive positive τ-sentences.

Definition 7.3. *Let Φ be an instance of* $\mathrm{CSP}(\Gamma)$*. Then the* injectivization of Φ*, denoted by* $\mathrm{inj}(\Phi)$*, is the instance ψ of* $\mathrm{CSP}(\mathrm{inj}(\Gamma))$ *obtained from ϕ by replacing each conjunct $\phi(x_1,\ldots,x_n)$ of Φ by $R^{\mathrm{inj}}_\phi(x_1,\ldots,x_n)$.*

We say that a constraint in an instance of $\mathrm{CSP}(\Gamma)$ is *false* if it defines an empty relation in Γ. Note that a constraint $R(x_1,\ldots,x_k)$ might be false even if the relation R is nonempty (simply because some of the variables from $x_1,\ldots,x_k$ might be equal). The proof of the following statement is identical to the proof for the random graph instead of (C_2^ω, Eq) in [BP15a].

Proposition 7.4 (Lemma 71 in [BP15a]). *Let Γ be preserved by a binary injection f of behavior E-dominated projection. Then* $\mathrm{CSP}(\Gamma)$ *can be reduced to* $\mathrm{CSP}(\mathrm{inj}(\Gamma))$ *in polynomial time.*

We are now in a position to give our reduction.

Proposition 7.5. *Let Γ be a reduct of (C^2_ω, E) such that $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C^2_\omega, E)}$ and Γ has a ternary injection f which behaves like minority. Further, let Δ be $(\{0,1\}; 0, 1, \{(x,y,z) : z+y+z = 1 \bmod 2\})$. There is a polynomial time reduction from* $\mathrm{CSP}(\mathrm{inj}(\Gamma))$ *to* $\mathrm{CSP}(\Delta)$*.*

Proof. Note that f preserves $\mathrm{inj}(\Gamma)$ since f is injective. From f one can derive a polymorphism f' on the two-element structure obtained from Γ by factoring by the equivalence classes, which behaves like the ternary minimum function on domain $\{0,1\}$.

Take an instance ϕ for $\mathrm{CSP}(\mathrm{inj}(\Gamma))$ and build an instance ϕ' for $\mathrm{CSP}(\Delta)$ in the following manner. The variable set remains the same, and every constraint $(b_1,\ldots,b_k) \in R$ from ϕ becomes $(a_1,\ldots,a_k) \in R'$ in ϕ' where $b_i \in C_{a_i}$. From Proposition 6.7, through the presence of f' and the lack of a polymorphism of Γ identifying one equivalence class alone, we can assume that the relations of ϕ' are preserved by $x+y+z \bmod 2$ and can thus be taken to be pp-definable in the relation $(x+y+z = 1 \bmod 2)$ (see, e.g., [CKS01]).

Suppose ϕ is a yes-instance of $\mathrm{CSP}(\mathrm{inj}(\Gamma))$; then ϕ' is a yes-instance of $\mathrm{CSP}(\Delta)$, by application of the polymorphism f'.

Suppose ϕ' is a yes-instance of $\mathrm{CSP}(\Delta)$, with solution $f \colon V \to \{0,1\}$. Then we can build a satisfying assignment for ϕ by choosing any injective function from V to (C^2_ω, E) sending $x \to C_{f(x)}$. $\square$

COROLLARY 7.6. *Let* Γ *be a reduct of* (C^ω_2, E) *which is preserved by a ternary injection* h *of behavior minority which is hyperplanely of behavior balanced xnor. Then* $\mathrm{CSP}(\Gamma)$ *can be solved in polynomial time.*

Proof. Note that the binary function $h(x, y, y)$ is of type p_1 and is E-dominated in the second argument. So the statement is a consequence of Propositions 7.4 and 7.5. $\square$

7.2. Infinitely many classes of size two. We now prove tractability of $\mathrm{CSP}(\Gamma)$ for reducts Γ of (C^2_ω, Eq) in a finite language such that $\mathrm{Pol}(\Gamma)$ contains a ternary canonical function h such that

$$h(N, \cdot, \cdot) = h(\cdot, N, \cdot) = h(\cdot, \cdot, N) = N$$

which behaves like a minority on $\{=, E\}$.

PROPOSITION 7.7. *A relation* R *with a first-order definition in* (C^2_ω, Eq) *is preserved by* h *if and only if it can be defined by a conjunction of formulas of the form*

$$N(x_1, y_1) \vee \cdots \vee N(x_k, y_k) \vee Eq(z_1, z_2) \tag{7.1}$$

for $k \geq 0$*, or of the form*

$$N(x_1, y_1) \vee \cdots \vee N(x_k, y_k) \vee (|\{i \in S : x_i \neq y_i\}| \equiv_2 p), \tag{7.2}$$

where $p \in \{0, 1\}$ *and* $S \subseteq \{1, \ldots, k\}$.

The proof is inspired from a proof for tractable phylogeny constraints [BJP17].

Proof. For the backwards implication, it suffices to verify that formulas of the form in the statement are preserved by h. Let $o, p, q \in R$, and let $r := h(o, p, q)$. Assume that R has a definition by a formula ϕ of the form as described in the statement. Suppose for contradiction that r does not satisfy ϕ. For any conjunct of ϕ violated by r, of the form $N(x_1, y_1) \vee \cdots \vee N(x_k, y_k) \vee \theta$, the tuple r must therefore satisfy $Eq(x_1, y_1) \wedge \cdots \wedge Eq(x_k, y_k)$. Since h has the property that $h(N, \cdot, \cdot) = h(\cdot, N, \cdot) = h(\cdot, \cdot, N) = N$, this means that each of o, p, and q also satisfies this formula. This in turn implies that o, p, and q must satisfy the formula θ. It suffices to prove that r satisfies θ, too, since this contradicts the assumption that r does not satisfy ϕ. Suppose first that θ is of the form $Eq(z_1, z_2)$. In this case, r must also satisfy $Eq(z_1, z_2)$ since h preserves Eq. So assume that θ is of the form $|\{i \in S : x_i \neq y_i\}| \equiv_2 p$ for $S \subseteq \{1, \ldots, k\}$ and $p \in \{0, 1\}$. Since each of o, p, q satisfies this formula and h behaves like a minority on $\{E, =\}$, we have that r satisfies this formula, too.

For the forwards implication, let R be an n-ary relation with a first-order definition in (C^2_ω, Eq) that is preserved by h. Define $\sim$ to be the equivalence relation on $(C^2_\omega)^n$ where $a \sim b$ if and only if $Eq(a_i, a_j) \Leftrightarrow Eq(b_i, b_j)$ for all $i, j \leq n$. Note that h preserves $\sim$. For $a \in (C^2_\omega)^n$, let R_a be the relation that contains all $t \in R$ with $t \sim a$. Let ψ_a be the formula

$$\bigwedge_{i<j\leq n, Eq(a_i, a_j)} Eq(x_i, x_j)$$

and ψ'_a be the formula

$$\bigwedge_{i<j\leq n, N(a_i, a_j)} N(x_i, x_j)\,.$$

Note that $t \in (C^2_\omega)^n$ satisfies $\psi_a \wedge \psi'_a$ if and only if $t \sim a$, and hence a tuple from R is in R_a if and only if it satisfies $\psi_a \wedge \psi'_a$.

Pick representatives $a_1, \ldots, a_m$ for all orbits of n-tuples in R.

Claim 1. $\bigvee_{i \leq m}(\psi_{a_i} \wedge \psi'_{a_i})$ is equivalent to a conjunction of formulas of the form (7.1) from the statement.

Rewrite the formula into an equivalent formula ψ_0 in conjunctive normal form of minimal size where every literal is either of the form $Eq(x, y)$ or of the form $N(x, y)$. Suppose that ψ_0 contains a conjunct with literals $Eq(a, b)$ and $Eq(c, d)$. Since ψ_0 is of minimal size, there exists $r \in (C^2_\omega)^n$ that satisfies $Eq(a, b)$ and none of the other literals in the conjunct, and similarly there exists $s \in (C^2_\omega)^n$ that satisfies $Eq(c, d)$ and none of the other literals. By assumption, $r \sim r' \in R$ and $s \sim s' \in R$. Since R is preserved by h, we have $t' := h(r', s', s') \in R$. Then $t \sim t'$ since h preserves $\sim$, and hence t satisfies ψ_0. But t satisfies none of the literals in the conjunct, a contradiction. Hence, all conjuncts of ψ_0 have form (7.1) from the statement.

Let $t \in (C^2_\omega)^n$, set $l := \binom{n}{2}$, and let $i_1 j_1, \ldots, i_l j_l$ be an enumeration of $\binom{\{1,\ldots,n\}}{2}$. The tuple $b \in \{0,1\}^{\binom{n}{2}}$ with $b_s = 1$ if $t_{i_s} \neq t_{j_s}$ and $b_s = 0$ otherwise is called the *split vector* of t. We associate to R_a the Boolean relation B_a consisting of all split vectors of tuples in R_a. Since R and R_a are preserved by h, the relation B_a is preserved by the Boolean minority operation and hence has a definition by a Boolean system of equations. Therefore, there exists a conjunction θ_a of equations of the form $|\{s \in S : x_{i_s} = y_{j_s}\}| \equiv_2 p$, $p \in \{0, 1\}$ such that $\theta_a \wedge \psi_a \wedge \psi'_a$ defines R_a.

Claim 2. The following formula ϕ defines R:

$$\phi := \psi_0 \wedge \bigwedge_{a \in \{a_1,\ldots,a_m\}} (\neg\psi_a \vee \theta_a).$$

It is straightforward to see that this formula can be rewritten into a formula of the form required in the statement.

To prove the claim, we first show that every $t \in R$ satisfies ϕ. Clearly, t satisfies ψ_0. Let $a \in \{a_1, \ldots, a_m\}$ be arbitrary; we have to verify that t satisfies $\neg\psi_a \vee \theta_a$. If there are indices $i, j \in \{1, \ldots, n\}$ such that $N(t_i, t_j)$ and $Eq(a_i, a_j)$, then t satisfies $\neg\psi_a$. We are left with the case that for all $i, j \in \{1, \ldots, n\}$ if $Eq(a_i, a_j)$, then $Eq(t_i, t_j)$. In order to show that t satisfies θ_a, it suffices to show that there exists a $t' \in R_a$ such that for all $i, j \leq n$ with $Eq(a_i, a_j)$ we have $t_i = t_j$ if and only if $t'_i = t'_j$. Note that $t' := h(a, a, t) \sim a$ since $h(N, \cdot, \cdot) = h(\cdot, N, \cdot) = h(\cdot, \cdot, N) = h$. Moreover, $t' \in R$ and thus $t' \in R_a$. Finally, for all $i, j \leq n$ with $Eq(a_i, a_j)$ we have $t_i = t_j$ if and only if $t'_i = t'_j$ because h behaves as a minority on $\{E, =\}$. Hence, t satisfies ϕ.

Next, we show that every tuple t that satisfies ϕ is in R. Since t satisfies ψ_0, we have that $t \sim a$ for some $a \in \{a_1, \ldots, a_m\}$. Thus, $t \models \psi_a \wedge \psi'_a$. By assumption, t satisfies $\neg\psi_a \vee \theta_a$ and hence $t \models \theta_a$. Therefore, $t \in R_a$ and in particular $t \in R$. ◻

PROPOSITION 7.8. *There is a polynomial-time algorithm that decides whether a given set Φ of formulas as in the statement of Proposition* 7.7 *is satisfiable.*

Proof. Let X be the set of variables which appear in Φ. Create a graph G with vertex set X that contains an edge between z_1 and z_2 if Φ contains a formula of the form $Eq(z_1, z_2)$. Eliminate all literals of the form $N(x_i, y_i)$ in formulas from Φ when x_i and y_i lie in the same connected component of G. Repeat this procedure until no more literals get removed.

We then create a Boolean system of equations Ψ with variable set $\binom{X}{2}$ as follows: if $x, y \in X$ are distinct, for better readability we write xy for the respective Boolean variable instead of $\{x, y\}$. For each formula $|\{i \in S \mid x_i \neq y_i\}| \equiv_2 p$ we add the Boolean equation $\sum_{i \in S} x_i y_i = p$. We additionally add for all $xy, yz, xz \in \binom{X}{2}$ the equation $xy + yz = xz$. If the resulting system of equations Ψ does not have a solution over $\{0, 1\}$, reject the instance. Otherwise accept.

To see that this algorithm is correct, observe that the literals that have been removed in the first part of the algorithm are false in all solutions, so removing them from the disjunctions does not change the set of solutions.

If the algorithm rejects, then there is indeed no solution to Φ. To see this, suppose that $s\colon C^2_\omega \to C^2_\omega$ is a solution to Φ. Define $b\colon \binom{X}{2} \to \{0, 1\}$ as follows. Note that for every variable $x_i y_i$ that appears in some Boolean equation in Ψ, a literal $N(x_i, y_i)$ has been deleted in the first phase of the algorithm (recall the syntactic form in (7.2); we only add Boolean equations to Ψ if all the literals involving N have been deleted), and hence we have $Eq(s(x_i), s(y_i))$. Define $s'(x_i y_i) := 1$ if $s(x_i) \neq s(y_i)$ and $s'(x_i y_i) := 0$ otherwise. Then s' is a satisfying assignment for Ψ.

We still have to show that there exists a solution to Φ if the algorithm accepts. Let $s'\colon \binom{X}{2} \to \{0, 1\}$ be a solution to Ψ. For each connected component C in the graph G at the final stage of the algorithm we pick two values $a_C, b_C \in C^2_\omega$ such that $Eq(a_C, b_C)$, and such that $N(a_C, d)$ and $N(b_C, d)$ for all previously picked values $d \in C^2_\omega$. Moreover, for each connected component C of G we pick a representative r_C. Define $s(r_C) := a_C$, and for $x \in C$ define $s(x) := a_C$ if $s'(xr_C) = 0$, and $s(x) := b_C$ otherwise.

Then s satisfies all formulas in Ψ that still contain disjuncts of the form $N(x_i, y_i)$, since these disjuncts are satisfied by s. Formulas of the form $|\{i \in S : x_i \neq y_i\}| \equiv_2 p$ are satisfied, too, since x_i and y_i lie in the same connected component C, and hence $s(x_i) \neq s(y_i)$ if and only if $s'(xr_C) \neq s'(y_i r_C)$, which is the case if and only if $s'(xr_C) + s'(y_i r_C) = s'(x_i y_i) = 1$ because of the additional equations we have added to Ψ. Therefore, $|\{i \in S : x_i = y_i\}| \equiv_2 p$ if and only if $\sum_{i \in S} s'(x_i y_i) = p$. ∎

COROLLARY 7.9. *Let* Γ *be a reduct of* (C^2_ω, Eq) *with finite signature and such that* $\mathrm{Pol}(\Gamma)$ *contains a ternary canonical injection* h *as described in the beginning of section* 7.2. *Then* $\mathrm{CSP}(\Gamma)$ *is in P.*

Proof. The proof is a direct consequence of Propositions 7.7 and 7.8. ∎

8. Summary for the homogeneous equivalence relations.

THEOREM 8.1. *Let* Γ *be a finite signature reduct of* (C^s_n, E)*, where either* $2 < n < \omega$ *or* $2 < s < \omega$*, and either* n *or* s *equals* ω*. Then one of the following holds.*

(1) Γ *is homomorphically equivalent to a reduct of* $(C^s_n, =)$*, and* $\mathrm{CSP}(\Gamma)$ *is in P or NP-complete by* [BK08].

(2) $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C^s_n, E)}$, $\mathrm{Pol}(\Gamma)$ *has a uniformly continuous h1 clone homomorphism, and* $\mathrm{CSP}(\Gamma)$ *is NP-complete.*

Proof. If Γ has an endomorphism whose image is a clique or an independent set, then Γ is homomorphically equivalent to a reduct of $(C^s_n, =)$ and the complexity classification is known from [BK08]. Otherwise, courtesy of Propositions 6.1 and 6.2, we may assume that $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C^s_n, E)}$ and that there is a pp definition of E, N, and Eq in Γ.

In the first case, that Eq has a finite number $n \geq 3$ of classes, we use Proposition 6.4 to see that the action of $\mathrm{Pol}(\Gamma)$ on the classes of Eq has no essential and no constant operation. It follows that this action has a uniformly continuous projective clone homomorphism as in Definition 4.2. The mapping which sends every function in $\mathrm{Pol}(\Gamma)$ to the function it becomes in the action on the classes of Eq is a uniformly continuous clone homomorphism [BP15b], and hence the original action of $\mathrm{Pol}(\Gamma)$ has a uniformly continuous projective clone homomorphism as well. This implies NP-completeness of $\mathrm{CSP}(\Gamma)$ (Theorem 4.3).

In the second case, that Eq has classes of finite size $s \geq 3$, we use Proposition 6.5 to see that the action of $\mathrm{Pol}(\Gamma, C)$ on some equivalence class C has no essential and no constant operation and hence has a uniformly continuous projective clone homomorphism. Picking any $c \in C$, we have that $\mathrm{Pol}(\Gamma, c) \subseteq \mathrm{Pol}(\Gamma, C)$ since C is pp-definable from c and Eq. Consequently, $\mathrm{Pol}(\Gamma, c)$ has a uniformly continuous projective clone homomorphism as well. Because Γ is a model-complete core, this implies that $\mathrm{Pol}(\Gamma)$ has a uniformly continuous projective h1 clone homomorphism [BP16a], and hence $\mathrm{CSP}(\Gamma)$ is NP-complete by Theorem 4.3. ◻

THEOREM 8.2. *Suppose* Γ *is a finite signature reduct of* (C_2^ω, E)*. Then one of the following holds.*

(1) Γ *is homomorphically equivalent to a reduct of* $(C_2^\omega, =)$*, and* $\mathrm{CSP}(\Gamma)$ *is in P or NP-complete by* [BK08].
(2) $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C_2^\omega, E)}$, $\mathrm{Pol}(\Gamma)$ *contains a canonical ternary injection of behavior minority which is hyperplanely of behavior balanced xnor, and* $\mathrm{CSP}(\Gamma)$ *is in P.*
(3) $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C_2^\omega, E)}$, $\mathrm{Pol}(\Gamma)$ *has a uniformly continuous h*1 *clone homomorphism, and* $\mathrm{CSP}(\Gamma)$ *is NP-complete.*

Proof. As in the proof of Theorem 8.1 we may assume that $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C_2^\omega, E)}$ and that E, N, and Eq are pp-definable. We apply Proposition 6.6. The first two cases from that proposition imply a uniformly continuous projective h1 clone homomorphism, and hence NP-completeness of the CSP, as in the proof of Theorem 8.1. The third case in Proposition 6.6 yields case (2) here, and tractability as detailed in section 7.1. ◻

THEOREM 8.3. *Suppose* Γ *is a finite signature reduct of* (C_ω^2, E)*. Then one of the following holds.*

(1) Γ *is homomorphically equivalent to a reduct of* $(C_\omega^2, =)$*, and* $\mathrm{CSP}(\Gamma)$ *is in P or NP-complete by* [BK08].
(2) $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C_\omega^2, E)}$*, Eq is not pp-definable,* $\mathrm{Pol}(\Gamma)$ *contains a canonical binary injective polymorphism of behavior* min *that is N-dominated, and* $\mathrm{CSP}(\Gamma)$ *is in P.*
(3) $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C_\omega^2, E)}$*, Eq is pp-definable,* $\mathrm{Pol}(\Gamma)$ *contains a ternary canonical function h with $h(N,\cdot,\cdot) = h(\cdot,N,\cdot) = h(\cdot,\cdot,N) = N$ and which behaves like a minority on $\{E,=\}$, and* $\mathrm{CSP}(\Gamma)$ *is in P.*
(4) $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C_\omega^2, E)}$, $\mathrm{Pol}(\Gamma)$ *has a uniformly continuous h*1 *clone homomorphism, and* $\mathrm{CSP}(\Gamma)$ *is NP-complete.*

Proof. As in Theorem 8.1, we may assume that $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C_\omega^2, E)}$, and that therefore E and N are pp-definable. If Eq is not pp-definable, then by Proposition 6.9, we have a binary injective polymorphism of behavior min that is N-dominated, and we have a polynomial algorithm from Theorem 4.5, similarly as in Proposition 4.6 for reducts of (H_n, E). Suppose now that Eq is pp-definable. We apply Proposition 6.10. As before, the first two cases imply NP-completeness of $\mathrm{CSP}(\Gamma)$. The third case from Proposition 6.10 yields tractability as detailed in section 7.2. ◻

Summarizing, we obtain a proof of Theorem 1.2.

Proof of Theorem 1.2. The statement follows from the preceding three theorems, together with [BW12] (for C^{ω}_{ω}) and [BK08] (for C^{1}_{ω} and C^{ω}_{1}). ☐

9. Outlook. We have classified the computational complexity of CSPs for reducts of the infinite homogeneous graphs. Our proof shows that the scope of the classification method from [BP15a] is much larger than one might expect at first sight. The general research goal here is to identify larger and larger classes of infinite-domain CSPs where systematic complexity classification is possible; two dichotomy conjectures are given for CSPs of reducts of finitely bounded homogeneous structures in [BPP14] and [BP16a], where these have now been proved equivalent in [BKO+17]. We have given additional evidence for these conjectures by proving that they hold for all reducts of homogeneous graphs. The next step in this direction might be to show a general complexity dichotomy for reducts of homogeneous structures whose age is finitely bounded and has the *free amalgamation property* (the Henson graphs provide natural examples for such structures). The present paper indicates that this problem might be within reach.

Acknowledgment. We are most grateful to our reviewers, whose patience and diligence has uncovered errors as well as encouraged us to make the paper significantly more readable.

REFERENCES

[BCKvO09] M. Bodirsky, H. Chen, J. Kára, and T. von Oertzen, *Maximal infinite-valued constraint languages*, Theoret. Comput. Sci., 410 (2009), pp. 1684–1693.

[BJP17] M. Bodirsky, P. Jonsson, and T. V. Pham, *The Complexity of Phylogeny Constraint Satisfaction Problems*, ACM Trans. Comput. Log., 18 (2017), 23.

[BK08] M. Bodirsky and J. Kára, *The complexity of equality constraint languages*, Theory Comput. Syst., 3 (2008), pp. 136–158.

[BK09] M. Bodirsky and J. Kára, *The complexity of temporal constraint satisfaction problems*, J. ACM, 57 (2009), pp. 1–41.

[BKJ05] A. Bulatov, P. Jeavons, and A. Krokhin, *Classifying the complexity of constraints using finite algebras*, SIAM J. Comput., 34 (2005), pp. 720–742, https://doi.org/10.1137/S0097539700376676.

[BKN09] L. Barto, M. Kozik, and T. Niven, *The CSP dichotomy holds for digraphs with no sources and no sinks (a positive answer to a conjecture of Bang-Jensen and Hell)*, SIAM J. Comput., 38 (2009), pp. 1782–1802, https://doi.org/10.1137/070708093.

[BKO+17] L. Barto, M. Kompatscher, M. Olšák, M. Pinsker, and T. V. Pham, *The equivalence of two dichotomy conjectures for infinite domain constraint satisfaction problems*, in Proceedings of the 32nd Annual ACM/IEEE Symposium on Logic in Computer Science, LICS'17, IEEE, Washington, DC, 2017, https://doi.org/10.1109/LICS.2017.8005128.

[BM16] M. Bodirsky and A. Mottet, *Reducts of finitely bounded homogeneous structures, and lifting tractability from finite-domain constraint satisfaction*, in Proceedings of the 31th Annual IEEE Symposium on Logic in Computer Science, LICS'16, IEEE, Washington, DC, 2016, pp. 623–632.

[BMM18] M. Bodirsky, B. Martin, and A. Mottet, *Discrete temporal constraint satisfaction problems*, J. ACM, 65 (2018), 9.

[BN06] M. Bodirsky and J. Nešetřil, *Constraint satisfaction with countable homogeneous templates*, J. Logic Comput., 16 (2006), pp. 359–373.

[Bod07] M. Bodirsky, *Cores of countably categorical structures*, Logical Methods Comput. Sci., 3 (2007), pp. 1–16.

[Bod12] M. Bodirsky, *Complexity Classification in Infinite-Domain Constraint Satisfaction*, Mémoire d'habilitation à diriger des recherches, Université Diderot, Paris 7; preprint, https://arxiv.org/abs/1201.0856, 2012.

[BP11] M. Bodirsky and M. Pinsker, *Reducts of Ramsey structures*, in Model Theoretic Methods in Finite Combinatorics, AMS Contemp. Math. 558, Providence, RI, 2011, pp. 489–519.

[BP14] M. Bodirsky and M. Pinsker, *Minimal functions on the random graph*, Israel J. Math., 200 (2014), pp. 251–296.

[BP15a] M. Bodirsky and M. Pinsker, *Schaefer's theorem for graphs*, J. ACM, 62 (2015), 19.

[BP15b] M. Bodirsky and M. Pinsker, *Topological Birkhoff*, Trans. Amer. Math. Soc., 367 (2015), pp. 2527–2549.

[BP16a] L. Barto and M. Pinsker, *The algebraic dichotomy conjecture for infinite domain constraint satisfaction problems*, in Proceedings of the 31st Annual IEEE Symposium on Logic in Computer Science, LICS'16, IEEE, Washington, DC, 2016, pp. 615–622.

[BP16b] M. Bodirsky and M. Pinsker, *Canonical Functions: A Proof via Topological Dynamics*, preprint, https://arxiv.org/abs/1610.09660, 2016.

[BP18] L. Barto and M. Pinsker, *Topology is Irrelevant*, preprint, available from http://www.karlin.mff.cuni.cz/~barto/Articles/pinsker_topologyirrelevant.pdf, 2018.

[BPP14] M. Bodirsky, M. Pinsker, and A. Pongrácz, *Projective clone homomorphisms*, J. Symbolic Logic, to appear, https://doi.org/10.1017/jsl.2019.23; preprint, https://arxiv.org/abs/1409.4601, 2014.

[BPT13] M. Bodirsky, M. Pinsker, and T. Tsankov, *Decidability of definability*, J. Symbolic Logic, 78 (2013), pp. 1036–1054.

[Bul06] A. A. Bulatov, *A dichotomy theorem for constraint satisfaction problems on a* 3*-element set*, J. ACM, 53 (2006), pp. 66–120.

[Bul17] A. A. Bulatov, *A dichotomy theorem for nonuniform CSPs*, in Proceedings of the 58th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, 2017, pp. 319–330.

[BW12] M. Bodirsky and M. Wrona, *Equivalence constraint satisfaction problems*, in Proceedings of Computer Science Logic, LIPICS 16, Dagstuhl Publishing, Dagstuhl, Germany, 2012, pp. 122–136.

[CKS01] N. Creignou, S. Khanna, and M. Sudan, *Complexity Classifications of Boolean Constraint Satisfaction Problems*, SIAM Monogr. Discrete Math. Appl. 7, SIAM, Philadelphia, 2001, https://doi.org/10.1137/1.9780898718546.

[FV98] T. Feder and M. Y. Vardi, *The computational structure of monotone monadic SNP and constraint satisfaction: A study through Datalog and group theory*, SIAM J. Comput., 28 (1998), pp. 57–104, https://doi.org/10.1137/S0097539794266766.

[GP08] M. Goldstern and M. Pinsker, *A survey of clones on infinite sets*, Algebra Universalis, 59 (2008), pp. 365–403.

[HN90] P. Hell and J. Nešetřil, *On the complexity of H-coloring*, J. Combin. Theory Ser. B, 48 (1990), pp. 92–110.

[HR94] L. Haddad and I. G. Rosenberg, *Finite clones containing all permutations*, Canad. J. Math., 46 (1994), pp. 951–970.

[JCG97] P. Jeavons, D. Cohen, and M. Gyssens, *Closure properties of constraints*, J. ACM, 44 (1997), pp. 527–548.

[KPT05] A. Kechris, V. Pestov, and S. Todorcevic, *Fraissé limits, Ramsey theory, and topological dynamics of automorphism groups*, Geom. Funct. Anal., 15 (2005), pp. 106–189.

[LW80] A. H. Lachlan and R. E. Woodrow, *Countable ultrahomogeneous undirected graphs*, Trans. Amer. Math. Soc., 262 (1980), pp. 51–94.

[NR89] J. Nešetřil and V. Rödl, *The partite construction and Ramsey set systems*, Discrete Math., 75 (1989), pp. 327–334.

[Pon17] A. Pongrácz, *Reducts of the Henson graphs with a constant*, Ann. Pure Appl. Logic, 168 (2017), pp. 1472–1489.

[Pos41] E. L. Post, *The Two-Valued Iterative Systems of Mathematical Logic*, Ann. Math. Stud. 5, Princeton University Press, Princeton, NJ, 1941.

[Sch78] T. J. Schaefer, *The complexity of satisfiability problems*, in Proceedings of the Symposium on Theory of Computing (STOC), ACM, New York, 1978, pp. 216–226.

[Sze86] Á. Szendrei, *Clones in Universal Algebra*, Séminaire de Mathématiques Supérieures, Les Presses de l'Université de Montréal, Montreal, Canada, 1986.

[Tho91] S. Thomas, *Reducts of the random graph*, J. Symbolic Logic, 56 (1991), pp. 176–181.

[TZ12] K. Tent and M. Ziegler, *A Course in Model Theory*, Lecture Notes in Logic 40, Cambridge University Press, Cambridge, UK, 2012.

[Zhu17] D. Zhuk, *A proof of CSP dichotomy conjecture*, in Proceedings of the 58th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, 2017, pp. 331–342.