A game for counting logic formula size and an application to linear orders

Grégoire Fournier University of Illinois at Chicago Chicago, USA gfourn2@uic.edu György Turán University of Illinois at Chicago Chicago, USA ELRN Research Group on AI Szeged, Hungary gyt@uic.edu

Abstract—Ehrenfeucht - Fraïssé (EF) games are a basic tool in finite model theory for proving definability lower bounds, with many applications in complexity theory and related areas. They have been applied to study various logics, giving insights on quantifier rank and other logical complexity measures. In this paper, we present an EF game to capture formula size in counting logic with a bounded number of variables. The game combines games introduced previously for counting logic quantifier rank due to Immerman and Lander, and for first-order formula size due to Adler and Immerman, and Hella and Väänänen. The game is used to prove the main result of the paper, an extension of a formula size lower bound of Grohe and Schweikardt for distinguishing linear orders, from 3-variable first-order logic to 3-variable counting logic. As far as we know, this is the first formula size lower bound for counting logic.

Index Terms—Finite Model Theory, First-Order Logic, Counting Logic, Succinctness

I. INTRODUCTION

Ehrenfeucht - Fraïssé (EF) games [7], [11] are a basic tool of finite model theory for proving definability lower bounds [6], [26]. Combined with logical characterizations of complexity classes, they provide a logic-based approach to problems in complexity theory. The original form of EF games gives bounds for quantifier rank in first-order logic but the games have been extended and modified for many logics and formula complexity measures. An EF game for *formula size* in first-order logic (FO) is given by Hella and Väänänen [19], building on Adler and Immerman [1]. In what follows, we refer to this game as the HV-game.

Counting logic extends first-order logic by adding the counting quantifier $\exists^{\geq k}$, and is frequently used in complexity theory and combinatorics. Counting logic turns out to be relevant for understanding the computational power of graph neural networks (GNN) as well [12]. An EF game for distinguishing graphs in counting logic with a bounded number of variables is formulated by Immerman and Lander [22]. The game [22] extends the basic EF setup by an additional phase in each round involving the choice of subsets of the same cardinality in the two structures.

Understanding *formula size in counting logic with a bounded number of variables* would be useful, in particular, for a further analysis of the logical characterizations of

GNNS (Barceló et al. [2]). An EF game for this setup could be a useful tool in this endeavor.

While the standard EF game is played on two structures, the HV game is played on two *sets of structures*, referred to in this paper as *families*. Grohe and Schweikardt [14] prove a formula size lower bound for linear orders for the 3-variable fragment of first-order logic, using this technique implicitly. They show that every first-order 3-variable formula that distinguishes a linear order of size *n* from a larger one has size $\Omega(\sqrt{n})$. Their proof is based on the notion of a *separator* and an involved weighting technique, which allows for a refined analysis of the syntax tree of a formula.

The separator and the weighting scheme are useful tools to gauge the progress made in the subformulas of a distinguishing formula. The proof is a detailed case analysis, with numerous subcases required to deal with quantifiers.

In this paper, we formulate a game for capturing formula size complexity for counting logic (Theorem V.1). The game is a combination of the Immerman-Lander and Hella-Väänänen games mentioned above. Restricted versions characterize formula size for fragments of counting logic where the *number of variables* and the *counting rank* are bounded.

The main result of the paper is that every 3-variable counting logic formula with counting rank t distinguishing a linear order of size n from a larger one has size at least \sqrt{n}/t (Theorem VI.3). This result extends the result of [14] from first-order logic to counting logic. The theorem is proved using the game characterization. There is a simple distinguishing formula of size n/t (Proposition VI.2). In addition, for the case t = 1, i.e., for 3-variable FO, our result improves the formula size lower bound of [14] from $\sqrt{n}/2$ to \sqrt{n} .

The lower bound proof uses modifications of the separators and the weighting scheme. Some cases considered are identical to [14]. The overall argument is, however, different. The main difference is in the most technical part, the proof of the quantifier case. We introduce *gap sets* and *gap variables* (Definitions VII.8 and VII.9), which makes the argument more similar to standard EF arguments and generalizes the reasoning. The new proof of the FO case is somewhat simpler than the original. A tree summarizing the proof structure and the different cases is given in Fig. 4. The paper is structured as follows. After reviewing related work in Section II, we describe the counting and HV games in Section III. Section IV describes the game and Section V gives the correspondence between the game and counting logic formula size. Section VI contains the application on linear orders, with the proof of the lower bound in Section VII. Our main technical contribution, the proof of the main Lemma on counting quantifiers, is presented separately in Section VII-C. Finally, Section VIII contains remarks and directions for future research.

II. RELATED WORK

Counting logic has been discussed in several different forms. Grohe [12] defines counting logic C as first-order logic (FO) extended by counting quantifiers of the form $\exists^{\geq k}x$, and C_m as its fragment using at most m variables. This is the counting logic we consider in this paper. Previous work using this kind of counting logic includes Immerman and Lander [22] and Cai et al. [3]. Grohe [13], on the other hand, considers a more powerful counting logic, where formulae can include arithmetic operations on the number of elements satisfying a formula (see also Kuske and Schweikardt [25]).

EF games using sets of structures, capturing the number of quantifiers as opposed to quantifier rank, have been proposed by Immerman [20]. These *multi-structural (MS)* games receive increasing current attention (Fagin et al. [10], Carmosino et al. [4], Vinall-Smeeth [32]). HV games are essentially extensions of MS games, also modelling Boolean connectives in the formulae.

A graph neural network (GNN) is a variant of neural networks for machine learning problems involving graphs [31]. Such a network allows the use of deep learning techniques to classify graphs (graph classification), or to classify the nodes of a large graph (node classification). The computational power of GNN is closely related to the Weisfeiler - Leman (WL) graph isomorphism algorithm (Morris et al. [27], Xu et al. [33]). The connection of the WL algorithm to counting logic with a bounded number of variables ([3], [12]) brings these logics into the GNN picture as well. Barceló et al. [2] gave logical characterizations in terms of counting logic with a bounded number of variables using results established in modal logic (Otto [29]). The complexity aspects of the characterizations are not discussed in [2], and studying this aspect (also pointed out in Grohe [12]) has been a motivation for the topic of this paper (a brief further discussion is given at the end of the paper).

The GNN characterizations of Grohe [13] establish a connection of GNN to threshold circuits, a Boolean circuit model of neural networks. The computational power of such circuits corresponds to counting logic with an arbitrary built-in predicate. Proving superpolynomial lower bounds for threshold circuits is an open problem. Hajnal et al. [15], [16] prove an exponential lower bound for depth-2 circuits with polynomial weights, and so far this lower bound has not been extended to either depth-2 with unrestricted weights or to higher depths. The same papers prove a quantifier rank lower bound for counting logic with successor as the built-in relation. Similar results are also given in Etessami [8], [9]. Karchmer and Wigderson [24] formulate an approach, related to HV games, to proving monotone formula depth lower bounds. They also prove a depth version of the Krapchenko formula size lower bound. Krapchenko's Theorem is proved in [19] as an application of HV games.

General background for the topic of this paper is given in Immerman [21], Ebbinghaus and Flum [6], Libkin [26], Otto [28] and Hamilton [17].

III. BACKGROUND

In this section we introduce basic notation used in the paper and review EF and HV games.

A. Basic definitions

1) Logics: We consider relational structures over a fixed vocabulary.

Counting logic C is obtained by extending first-order logic with counting quantifiers $\exists^{\geq k} x \phi(x)$ and $\forall^{\geq k} x \phi(x)$. Here $\exists^{\geq k} x \phi(x)$ means that there are at least k distinct assignments to the variable x that satisfy ϕ . Thus $\exists^{\geq k} \phi(x)$ is logically equivalent to $\exists x_1 \dots \exists x_k (\bigwedge_i \phi(x_i) \land \bigwedge_{i,j} x_i \neq x_j)$. The quantifier $\forall^{\geq k} x \phi(x)$ stands for $\neg \exists^{\geq k} x \neg \phi(x)$. In $\exists^{\geq k} x \phi(x)$ and $\forall^{\geq k} x \phi(x)$, k is referred to as the *counting rank* of the quantifier. The counting rank of a formula is the maximum counting rank of its quantifiers.

As a counting quantifier can be replaced by standard quantifiers, adding counting quantifiers does not change the expressivity of first-order logic. It does, however, impact the succinctness, the minimum size of formulae expressing a property. In applications to finite model theory one usually considers a sequence $(\mathcal{A}_n, \mathcal{B}_n)$ of pairs of structures. Complexity bounds to be proven are also functions of n. Note that $\forall^{\geq k} x \varphi(x)$ is equivalent to $\exists^{\geq n-k+1} x \varphi(x)$ for an n-element structure. The transformation increases counting rank and thus it cannot be used in formula size bounds for the bounded counting rank case.

Parameters to be considered are the bound m on the number of variables, the bound t on the counting rank, and the bound w on the formula size. The fragment of counting logic of formulae containing at most m variables and counting rank at most t is denoted by C_m^t .

2) Structures and Families: The universe of a structure \mathcal{A} is denoted by $\mathcal{U}^{\mathcal{A}}$. We use $x_j, j \in \mathbb{N}$, to denote variables. A variable assignment for a structure \mathcal{A} is a finite partial mapping $\alpha : \mathbb{N} \to \mathcal{U}^{\mathcal{A}}$. The finite domain of α is denoted by $dom(\alpha)$.

An *interpretation* is a pair (\mathcal{A}, α) . For a formula ϕ , $(\mathcal{A}, \alpha) \models \phi$ means that the assignment α satisfies the formula ϕ in the structure \mathcal{A} , with $dom(\alpha)$ containing all the *j* for which the variable x_j is free in ϕ . A formula ϕ *distinguishes* interpretations (\mathcal{A}, α) and (\mathcal{B}, β) , denoted by $((\mathcal{A}, \alpha), (\mathcal{B}, \beta)) \models \phi$, if $(\mathcal{A}, \alpha) \models \phi$ and $(\mathcal{B}, \beta) \models \neg \phi$.

A family $A_{\mathcal{A},D}$ is a set of interpretations $\{(\mathcal{A}, \alpha_i) | i \in \Gamma\}$ where \mathcal{A} and $D = dom(\alpha_i)$ are fixed and Γ is some set. When the context is clear, we drop the subscript. We write $(A, B) \models \phi$ to express that for all $(\mathcal{A}, \alpha) \in A$, $(\mathcal{A}, \alpha) \models \phi$ and for all $(\mathcal{B}, \beta) \in B$, $(\mathcal{B}, \beta) \models \neg \phi$, and we say that ϕ distinguishes (A, B). For a structure \mathcal{A} , we denote by A_0 the family $\{(\mathcal{A}, \emptyset)\}$, containing a single interpretation with the empty assignment.

3) Operations: If α is an assignment on \mathcal{A} , $a \in \mathcal{U}^{\mathcal{A}}$ and $j \in \mathbb{N}$, then $\alpha(a/j)$ is the assignment that maps j to a and agrees with α otherwise.

Given a family A, a *choice function* is of the form $F : A \to U^A$. The set of all choice functions on the family A is denoted by F_A . We define two operations on families [19].

Change: Given a family A, a choice function F ∈ F_A and j ∈ N, the change operation on A with F for the variable x_j produces the family

$$A(F/j) := \{ (\mathcal{A}, \alpha(F(\mathcal{A}, \alpha)/j)) : (\mathcal{A}, \alpha) \in A \}.$$

In the new family, the assignment to x_j is changed based on the choice function F. If $j \notin D$ then x_j is a new variable, and D is updated to $D \cup \{j\}$. Thus a change operation may either leave the domain unchanged or add a new element to it.

• *Multiplication*: Given a family A and $j \in \mathbb{N}$, the multiplication operation A for the variable x_j produces the family

$$A(*/j) := \{ (\mathcal{A}, \alpha(a/j)) : (\mathcal{A}, \alpha) \in \mathcal{A}, a \in \mathcal{U}^{\mathcal{A}} \}.$$

The new family consists of interpretations with x_j assigned to all possible values in $\mathcal{U}^{\mathcal{A}}$. Here, again, the domain is either unchanged or a new element is added to it.

4) Formula complexity: The size of a formula is defined inductively: if ϕ is an atomic formula, $|\phi| = 1$; and for FO formulae $\phi, \psi, |\neg \phi| = |\phi| + 1$; $|\phi \lor \psi| = |\phi \land \psi| = |\phi| + |\psi|$; $|\exists x_j \phi| = |\forall x_j \phi| = 1 + |\phi|$. For the definition of the quantifier rank of a formula, we refer the reader to [6].

B. Review of games

In this section we review the counting logic game [22] and the first-order logic formula size game [19], referred to as the HV game.

Definition III.1 (*r*-round EF *m*-pebbling game). The game $EF(\mathcal{A}, \mathcal{B})$ is played on two relational structures \mathcal{A} and \mathcal{B} . There are two players, Spoiler and Duplicator, and *m* pairs of pebbles (a_i, b_i) for $i \in [m]$. It goes as follows:

- For r rounds:
 - Spoiler picks a set of elements S_1 of A or B and a number $i \in [m]$. Duplicator selects a set S_2 of elements of the same cardinality in the other structure.
 - Spoiler picks an element in S_2 and Duplicator selects an element in S_1 . The pebble a_i (resp., b_i) holds the value of the element picked in \mathcal{A} (resp., \mathcal{B}).
- The r-round game ends in the position $\vec{a} = (a_1, \dots, a_m)$, $\vec{b} = (b_1, \dots, b_m)$. Duplicator wins if the mapping from

 $((\vec{a}, \vec{c}^{\mathcal{A}}) \text{ to } (\vec{b}, \vec{c}^{\mathcal{B}}))$ is a partial isomorphism between \mathcal{A} and \mathcal{B} , where \vec{c} denotes the constants of the language.

Theorem III.2. The following are equivalent:

- A and B satisfy the same sentences of C_m of quantifier rank at most r.
- Duplicator has a winning strategy in the r-round mpebbling game.

A similar result holds for the bounded counting rank fragment C_m^t , by restricting the cardinality of the sets picked to be at most t.



Fig. 1. Two families, $A = \{(\mathcal{A}, \alpha)\}$ and $B = \{(\mathcal{B}, \beta)\}$, for the $\exists z$ move in the HV-game.

Definition III.3 (HV game for formula size on first-order logic). The game $HV_w(A, B)$ is played on two families, A and B, by Spoiler and Duplicator. The initial position is (w, A, B). There are five possibilities for the continuation of the game:

- \neg -move: the game continues from the position (w 1, B, A).
- V-move: Spoiler chooses 1 ≤ u, v < w such that u+v = w and partitions A to get C∪D. Duplicator chooses the next position as (u, C, B) or (v, D, B).
- \wedge -move: similar but played on B.
- \exists -move: Spoiler chooses $j \in \mathbb{N}$ and a choice function F from F_A . Then the game continues from (w - 1, A(F/j), B(*/j)).
- \forall -move: similar but Spoiler chooses on B.

Spoiler wins if the game reaches a position (w, A, B) for $w \ge 1$ and there is an atomic formula that distinguishes A and B. Duplicator wins if the game reaches a position (1, A, B) and Spoiler does not win.

Theorem III.4. Let (A, B) be a pair of families, and let w be a positive integer. Then the following are equivalent:

- 1) Spoiler has a winning strategy in the game $HV_w(A, B)$.
- 2) There is a formula ϕ of FO of size $|\phi| \leq w$ such that $(A, B) \models \phi$.

Note that for any $m \in \mathbb{N}$, one can define the variant of the game $HV_w^m(A, B)$, for which the moves \exists and \forall are restricted by $j \in \{1, \ldots, m\}$. This game then characterizes FO_m , the fragment of FO logic, with m variables. An illustration of the \exists -move is given in Fig. 1.

In Table I, we summarize the game characterizations discussed above and the game we are about to introduce.

TABLE I COMPLEXITY CHARACTERIZING GAMES.

	FO_m	\mathcal{C}_m
Quantifier Rank	[11], [7]	EF _m [22]
Size	HV_m [18]	CS_m (new)

IV. THE COUNTING LOGIC FORMULA SIZE GAME

In this section we define our game for counting logic formula size. We start by extending the operations of the HV game to this setting.

A. Extended operations

Definition IV.1. (k-choice function, selection) Given a family A, a k-choice function is of the form $F^k = (F_1^k, F_2^k, \ldots, F_k^k)$, where each F_i^k is a choice function, and for every $(\mathcal{A}, \alpha) \in A$ the elements $F_i^k(\mathcal{A}, \alpha)$ are pairwise distinct. The set of all the k-choice functions on the family A is denoted by F_A^k . A choice function F is a selection from F^k if $F(\mathcal{A}, \alpha) \in \{F_1^k(\mathcal{A}, \alpha), \ldots, F_k^k(\mathcal{A}, \alpha)\}$ for every $(\mathcal{A}, \alpha) \in A$. We say that F is selected from F^k .

Note that if F is a *selection* from F^k , then the choices on different interpretations may correspond to different F_i^k s. This concept will be used to represent the Spoiler's choices of one interpretation from a set of interpretations. The extended set of operations is the following:

k-Change: Given a family A and j, k ∈ N, the k-change operation associated to F^k ∈ F^k_A produces the family

$$A(F^k/j) := \{ (\mathcal{A}, \alpha(F_i^k(\mathcal{A}, \alpha)/j)) : (\mathcal{A}, \alpha) \in \mathcal{A}, 1 \le i \le k \}$$

Thus each interpretation gives rise to k interpretations composing the new family, where the assignments to x_j are changed based on the k-choice function F^k .

k-Multiplication: Given a family A and j ∈ N, the k-Multiplication operation on A does the following: for every k-choice function F^k on A, a choice function F is selected from F^k. The union of the A(F/j) over F^k ∈ F^k_A forms A(*^k/j).

In other words, for every $F^k \in F_A^k$ and $(\mathcal{A}, \alpha) \in A$, one interpretation is picked from $\{(\mathcal{A}, \alpha(F_i^k(\mathcal{A}, \alpha)/j)) : 1 \le i \le k\}$ to be part of $A(*^k/j)$. Thus each interpretation generates $card(F_A^k)$ interpretations that compose the family $A(*^k/j)$.

B. Formula size game for counting logic

We are now ready to define the game for counting logic formula size (referred to as the CS game for "counting size"). It is presented in the C_m version as this will be used in the rest of the paper. The new $\exists^{\geq k}$ -move is illustrated in Fig. 2.

Definition IV.2 (CS game for formula size on counting logic). The game $CS_w^m(A, B)$ has two players, Spoiler and Duplicator, m is the number of variables. A, B are two families with dom(A) = dom(B) of size at most m. Suppose after p moves we reach the position (w, A, B). Depending on Spoiler's choice, the game continues as follows:

- \neg -move: the game continues from the position (w 1, B, A).
- V-move: Spoiler first chooses numbers u and v such that I ≤ u, v < w and u + v = w. Then Spoiler partitions A into a pair of families C and D. The game continues either from the position (u, C, B) or from the position (v, D, B) according to Duplicator's choice.
- ∧-move: Spoiler first chooses numbers u and v such that 1 ≤ u, v < w and u + v = w. Then Spoiler turns B into a pair of families C and D. The game continues either from the position (u, A, C) or from the position (v, A, D) according to Duplicator's choice.
- ∃^{≥k}-move: Spoiler chooses j ∈ [m], k ∈ N and a k-choice function F^k on A. For every k-choice function G^k on B, Spoiler selects G from G^k. The union of the B(G/j) over G^k ∈ F^k_B forms B(*^k/j). Then the game continues from the position (w − 1, A(F^k/j), B(*^k/j)).
- ∀^{≥k}-move: Spoiler chooses j ∈ [m], k ∈ N and a kchoice function F^k on B. For every k-choice function G^k on A, Spoiler selects G from G^k. The union of the A(G/j) over A^k ∈ F^k_A forms A(*^k/j). Then the game continues from the position (w − 1, A(*^k/j), B(F^k/j)).

(Atomic) The game ends in a position (w, A, B) if either there is an atomic formula ϕ such that $(A, B) \models \phi$, in which case Spoiler wins, or if w = 1, in which case Duplicator wins if there is no such ϕ .



Fig. 2. The $\exists \geq k$ -move in the CS game. On the left the k-Change operation: k different elements are chosen by Spoiler. On the right the two steps of the k-Multiplication. The first step is to form all k-choice functions (each symbolized by a box). In the second step Spoiler picks one interpretation in each box to compose the new family.

C. An illustration of the CS game

Example IV.3. Consider two structures A and B of 3 and 4 elements respectively, onto which we define the unary relations blue and red, as illustrated in Fig. 3.

The sentence $\exists^{\geq 2}x$ (blue(x) \wedge red(x)) distinguishes \mathcal{A} and \mathcal{B} and is of size w = 3. It is easy to check by enumerating a finite number of formulae that no formula of size w = 2 in



Fig. 3. Structures A and B with relations red and blue, and a possible $\exists^{\geq 2}x$ move.

counting logic with counting rank 2 is able to distinguish A from B.

Consider the starting position $CS_3^{1,2}(A_0, B_0)$. In the following description we omit w for simplicity. Assume again that Spoiler plays $\exists^{\geq 2}x$ and picks $\{1,2\}$ in the left, corresponding to two copies of \mathcal{A} , with $\alpha(x) = 1$ and $\alpha(x) = 2$. The 2-choice functions on \mathcal{B} are $\{1,2\}, \{1,3\}, \{2,3\}$.

Spoiler now has several options to select a choice function: $\{\{1,1,3\}\} = \{1,3\}$ (as on the Figure), $\{1,2\}$, $\{2,3\}$ or $\{1,2,3\}$. In all these cases, neither red nor blue distinguishes the two families. So the game has to go on for Spoiler to be able to win. In fact, it is not too complicated to check that Spoiler does not win from the position $CS_2^{1,2}(A_0, B_0)$. Continuing from $(\{1,2\}, \{1,3\})$ as in the Figure, Spoiler can play the \wedge move to get to one of the two positions:

• $\{1,2\}$ on A and $\{1\}$ on B, which wins by playing red.

• $\{1,2\}$ on A and $\{3\}$ on B, which wins by playing blue. So, the " $\exists \geq^2 x (red(x) \land blue(x))$ " sequence of moves is winning for Spoiler and the corresponding formula separates A from B.

V. GAME CHARACTERIZATION THEOREM

In this section, we state the game characterization theorem and formulate some corollaries. The proof of the characterization theorem is given in Appendix A.

A. Game characterization and corollaries

Theorem V.1 (Characterization Theorem). Let A, B be families and w be a positive integer. Then the following are equivalent:

- 1) Spoiler has a winning strategy in the game $CS_w^m(A, B)$.
- 2) There is a C_m formula ϕ of size $|\phi| \leq w$ such that $(A, B) \models \phi$.

A corollary for structures is as follows:

Corollary V.2. Let (A and B) be structures and let w be a positive integer. Then the following conditions are equivalent:

- 1) Spoiler has a winning strategy in the game $CS_w^m(A_0, B_0)$.
- 2) There is a C_m -sentence ϕ of size $|\phi| \leq w$ such that $(\mathcal{A}, \mathcal{B}) \models \phi$.

Corollary V.3. Let (A and B) be structures and let w be a positive integer. Then the following conditions are equivalent:

- 1) Spoiler has a winning strategy in the game $CS_w^m(A_0, B_0)$.
- 2) There is a C_m -sentence ϕ of size $|\phi| \leq w$ such that $(\mathcal{A}, \mathcal{B}) \models \phi$.

We now describe the variation of the game for bounded counting rank.

Definition V.4 (CS game for C_m^t -formula size). The $CS_w^{m,t}(A, B)$ game is the version of the $CS_w^m(A, B)$ game where $k \leq t$ for the $\exists^{\geq k}$ and $\forall^{\geq k}$ moves.

Corollary V.5. Suppose A and B be families and let w be a positive integer. Then the following are equivalent:

- 1) Spoiler has a winning strategy in the game $CS_w^{m,t}(A, B)$.
- 2) There is a C_m^t -formula ϕ of size $|\phi| \leq w$ such that $(A, B) \models \phi$.

The proof is similar to the CS game characterisation of C_m formula size presented in Section A, but with the aforementioned parameter k bounded, and is omitted.

We also derive a variation of the game for graded modal logic, where the modal binary relation is denoted by E.

Definition V.6 (CS game for graded modal logic formula size). The E- $CS_w^2(A, B)$ game is the two variables version of the $CS_w^2(A, B)$ game where additionally for the $\exists^{\geq k}$ and $\forall^{\geq k}$ moves: Spoiler chooses $i \in [m]$ that has already been chosen (i.e a variable x_j with an assignment), and the spaces of the k-choice functions are limited to $\{e \in A | E(\alpha(x_i), e)\}$ and $\{e \in B | E(\beta(x_i), e)\}$.

Following the equivalence between the expressivity of GNNS and the graded modal logic established in [2], we obtain the following corollary linking the distinguishing power of GNNS and EF games:

Corollary V.7. Suppose A and B be families and let w be a positive integer. Then the following are equivalent:

- 1) Spoiler has a winning strategy in the game E- $CS_w^2(A, B)$.
- 2) There is a GNN that expresses an FO formula of size w that is capable of distinguishing A from B.

VI. LINEAR ORDERS DISTINGUINSHABILITY THROUGH COUNTING LOGICS

In this section, we present the new results derived from the complexity games for distinguishability and indistinguishability of finite linear orders.

A linear order is defined over the signature $\{min, max, < , succ\}$, where \langle is a linear ordering and *succ* is the successor

relation. \mathcal{A}_n denotes the structure ($\{0, \ldots, n\}$, <), where < is the standard linear order of $\{0, \ldots, n\}$. In the rest of the paper, without loss of generality we will take linear orders on the integers, and we will denote by d(b, a) the quantity |b-a|for a, b integers.

A. Lower bound for the quantifier rank necessary to distinguish linear orders

Theorem VI.1. Let t, k > 0, and let L_1, L_2 be linear orders of length at least $(t+1)^k$.

Then $L_1 \equiv_{\mathcal{C}^t[k]} L_2$;

where $\mathcal{C}^{t}[k]$ denotes the fragment of the counting logic with counting rank at most t and quantifier rank at most k.

This result extends the theorem from [26] stating that two linear orders of size at least 2^k cannot be distinguished by a FO[k] formula.

The proof of this result is relatively short, and provides a good intuition about the framework developed for the more complex Theorem VI.3. The proof follows the pebble EF game of Definition III.1, with bounded counting rank [22] and is presented in Appendix B.

B. Upper bound for the minimal size necessary to distinguish linear orders

The following proposition gives an upper bound on the size of C_2^t - formulas distinguishing linear orders of different sizes.

Proposition VI.2. There is a C_2^t -formula ϕ of size O(n/t)such that for n < m it holds that $(\mathcal{A}_n, \mathcal{A}_m) \models \phi$.

Proof. Consider the formulae:

- 1) $\phi_0(x) = (x = x).$
- 2) $\phi_{t(l+1)}(x) = \exists^{\geq t} y((y < x) \land \phi_{tl}(y)).$
- 3) For n = tl + p with p < t: $\phi_{tl+p}(x) = \exists^{\geq p} y((y < t))$ $(x) \wedge \phi_{tl}(y)$

Let $\phi = \neg \exists x \ \phi_{n+1}(x)$ Then a linear order satisfies ϕ iff it has size at most n. The size of ϕ is O(n/t).

C. Lower bound for the size necessary of a C_3^t formula to distinguish linear orders

Using the $CS^{3,t}$ game introduced in Section V, we derive the following lower bound for the size of a \mathcal{C}_3^t formula distinguishing linear orders:

Theorem VI.3. If ϕ is a C_3^t -formula distinguishing A_n and \mathcal{A}_m , where n < m, then $|\phi| \geq \frac{\sqrt{n}}{t}$.

For t = 1, i.e. for FO₃, this theorem improves on the best known lower bound of [14] from $\sqrt{n}/2$ to \sqrt{n} . The rest of the paper is dedicated to proving this result.

VII. PROOF OF THEOREM VI.3

A. Description of the framework

We extend the framework for FO of [14] to counting logic. We start by describing the main concepts: the counting game derived extended syntax trees, the separators, and the weighting scheme.

1) Extended Syntax Tree: Given two families A, B, the extended syntax tree represents a winning strategy for Spoiler in the $CS^{3,t}$ game on (A, B). It assigns to each tree node v a pair of families il(v) ("interpretation label"), along with Spoiler's move sl(v) ("syntax label") in this position such that:

- 1) A node v_1 is a child of node v if position $il(v_1)$ can be obtained in one move from il(v),
- 2) A node v is a leaf if il(v) satisfies the atomic win condition.

The root is associated to the starting position of the game $CS^{3,t}$, and we consider the nodes associated to positions reached through a winning strategy.

Definition VII.1. Let ψ be an C_3^t -formula, let A and B be families of interpretations such that $(A, B) \models \psi$. By induction on the construction of ψ , we define an extended syntax tree $T_{\psi}^{\langle A,B\rangle}$ as follows:

- If ψ is an atomic formula, then $T_{\psi}^{\langle A,B\rangle}$ consists of a single node v that has a syntax label $sl(v) := \psi$ and an interpretation label $il(v) := \langle A, B \rangle$.
- If ψ is of the form $\neg \psi_1$, then $T_{\psi}^{\langle A,B \rangle}$ has a root node v
- If ψ is of the form ¬ψ₁, then T_ψ ' has a root node v with sl(v) := ¬ and il(v) := ⟨A, B⟩. The unique child of v is the root of T_{ψ1}^{⟨B,A⟩}. Note that (B, A) ⊨ ψ₁.
 If ψ is of the form ψ₁∨ψ₂, then T_ψ^{⟨A,B⟩} has a root node v with sl(v) := ∨ and il(v) := ⟨A, B⟩. The first child of v is the root of T_{ψ1}^{⟨A,B⟩} and the second child of v is the root of c_π^{⟨A,B⟩}. of $T_{\psi_2}^{\langle A_2, B \rangle}$, where $A_i = \{(\mathcal{A}, \alpha) \in A : (\mathcal{A}, \alpha) \models \psi_i\}$ for $i \in \{1, 2\}$. Note that $A = A_1 \cup A_2$ and $(A_i, B) \models \psi_i$. • If ψ is of the form $\psi_1 \land \psi_2$, then $T_{\psi}^{\langle A, B \rangle}$ has a root node v
- with $sl(v) := \wedge$ and $il(v) := \langle A, B \rangle$. The first child of v is the root of $T_{\psi_1}^{\langle A, B_1 \rangle}$ and the second child of v is the root $\langle A, B_2 \rangle$ of $T_{\psi_2}^{\langle A, B_2 \rangle}$, where $B_i = \{ (\mathcal{B}, \beta) \in B : (\mathcal{B}, \beta) \models \neg \psi_i \}$ for $i \in \{1,2\}$. Note that $B = B_1 \cup B_2$, $(A, B_i) \models \psi_i$.
- If ψ is of the form $\exists^{\geq k} u \psi_1$, for a variable $u \in \{x, y, z\}$ and $k \leq t$, then $T_{\psi}^{\langle A, B \rangle}$ has a root node v with $sl(v) := \exists^{\geq k} u$ and $il(v) := \langle A, B \rangle$. The unique child of v is the root of $T_{\psi_1}^{\langle A(F^k/u), B(*^k/u) \rangle}$, where $F^k \in F_A^k$ is chosen so that $A(F^k/u) \models \psi_1$, and for every $G^k \in F_B^k$, G is selected from G^k so that $B(G/u) \models \neg \psi_1$. Since $B(*^k/u) = \bigcup_{G^k \in F_D^k} B(G/u)$, note that $(A(F^k/u), B(*^k/u)) \models \psi_1.$
- If ψ is of the form $\forall^{\geq k} u \psi_1$, for a variable $u \in \{x, y, z\}$ and $k \leq t$, then $T_{\psi}^{\langle A,B \rangle}$ has a root node v with $sl(v) := \forall^{\geq k}u$ and $il(v) := \langle A,B \rangle$. The unique child of v is the root of $T_{\psi_1}^{\langle A(*^k/u),B(F^k/j) \rangle}$, where $F^k \in F_B^k$ is chosen so that $B(F^k/u) \models \neg \psi_1$, and for every $G^k \in F_A^k$, G is set lected from G^k so that $A(G/u) \models \psi_1$. Since $A(*^k/u) =$ $\bigcup_{G^k \in F_*^k} A(G/u)$, note that $(A(*^k/u), B(F^k/u)) \models \psi_1$.

We define $|T_{\psi}^{\langle A,B\rangle}|$ to be the number of nodes in the tree $T_{\psi}^{\langle A,B\rangle}$. It is trivial to check by induction that $|T_{\psi}^{\langle A,B\rangle}| = |\psi|$.

Next we specify separators, the key concept to study succinctness on linear orders. Separators express the distance from which, at a given position in the game (i.e. at a certain size in the formula), we are able to distinguish two elements of linear orders. This distance differs for each element of $\mathcal{P}_2(\{\min, \max, x, y, z\})$, where the \mathcal{P}_2 operator describes the subsets of size two.

2) Separators and Weighting Scheme:

Definition VII.2 (separator). Let A and Bbe sets of interpretations. A separator is a mapping δ : $\mathcal{P}_2(\{\min, \max, x, y, z\}) \rightarrow \mathbb{N}$. δ is called separator for $\langle A, B \rangle$, if the following is satisfied:

For every $I := (\mathcal{A}, \alpha) \in A$ and $J := (\mathcal{B}, \beta) \in B$, there are $u, u' \in \{min, max, x, y, z\}$ with $u \neq u'$, such that :

1)
$$<-type(\alpha(u), \alpha(u')) \neq <-type(\beta(u), \beta(u'))$$
 or

2) *both:*

• $MIN[d(\alpha(u), \alpha(u')), d(\beta(u), \beta(u'))] \leq \delta(\{u, u'\})$ • $d(\alpha(u), \alpha(u')) \neq d(\beta(u), \beta(u')).$

where, for $a, b \in \mathbb{N}$, d(a, b) := |a - b| and $\langle -type(a, b)$ is " = "," < " or " > " reflecting the order between a and b.

Definition VII.3 (weight of a separator and minimal separator). Let δ be a separator, we define its:

- 1) border-distance $b(\delta) := MAX \{\delta(\{min, max\}), \}$ $\delta(\{\min, u\}) + \delta(\{u', \max\}) : u, u' \in \{x, y, z\}\}$
- 2) centre-distance $c(\delta) := MAX \{\delta(p) + \delta(q) : p, q \in$ $\mathcal{P}_2(\{x, y, z\}), p \neq q\}$ 3) weight $w(\delta) := \sqrt{c(\delta)^2 + b(\delta)}$

The weight is a measure of the distinguishing power of separators. We call *minimal separator* of a pair of structures a separator of minimal weight.

B. Proof of Theorem VI.3

We introduce the main technical lemma of the proof, which links separators weights to the structure of the tree.

Lemma VII.4. Suppose $(A, B) \models \psi$ and let T be the extended syntax tree $T_{\psi}^{\langle A,B \rangle}$. For every node v of T the following is true:

- 1) If v is a leaf, then $w(\delta) < 1$.
- 2) If v has 2 children v_1 and v_2 , and δ_i is a minimal separator for $il(v_i)$, for $i \in \{1, 2\}$, then $w(\delta) \le w(\delta_1) + w(\delta_2)$.
- 3) If v has exactly one child v_1 , and δ_1 is a minimal separator for $il(v_1)$, then $w(\delta) \leq w(\delta_1) + t$.

where δ is a minimal separator for il(v).

The proofs of parts 1. and 2., related only to the syntax of first-order logic, are almost the same in our counting logic extended syntax tree framework as in [14], and are given in Appendix C. However, the proof of Lemma VII.4 part 3., which accounts for the counting quantifiers, requires different arguments, to take into account the "multiple choices" needed in counting logic, and is given in Section VII-C.

The bound of part 3. extends the corresponding bound of [14] to counting logic. In fact, for counting rank 1, i.e., for firstorder logic without counting, it improves the bound from $w(\delta) \leq w(\delta_1) + 2$ to $w(\delta) \leq w(\delta_1) + 1$. This leads to the slight improvement of the lower bound of [14] mentioned after Theorem VI.3.

From Lemma VII.4's inductive relation on the minimal weight of extended syntax tree nodes, we derive a lower bound on the size of the tree in terms of the minimal weight of the root.

Lemma VII.5. Let T be a finite binary tree where each node v is equipped with a weight w(v) > 0 such that the following is true:

- If v is a leaf, then $w(v) \leq 1$.
- If v has 2 children v_1 and v_2 , then $w(v) \le w(v_1) + w(v_2)$.
- If v has exactly one child v_1 , then $w(v) < w(v_1) + t$.

Then, $|T| \geq \frac{w(r)}{t}$, where r is the root of T and |T| is the number of nodes of T.

Proof. The proof is by induction:

- If T is a leaf, |T| = 1 ≥ w(r)/t.
 If T's root starts with two children v₁, v₂ associated to the trees T_1, T_2 . By inductive hypothesis $|T_i| \geq \frac{w(v_i)}{t}$, so $|T| = |T_1| + |T_2| + 1 \ge \frac{w(v_1)}{t} + \frac{w(v_2)}{t} + 1 \ge \frac{w(r)}{t}.$ • Finally if T's root starts with one child v_1 associated to
- the tree T_1 . By inductive hypothesis $|T_1| \geq \frac{w(v_1)}{t}$, so $|T| = |T_1| + 1 \ge \frac{w(v_1)}{t} + \frac{t}{t} \ge \frac{w(r)}{t}.$

Finally we can derive a lower bound for the size of the extended syntax trees on the counting logic with the main result:

Theorem VII.6. Suppose $(A, B) \models \psi$ and δ is a minimal separator for $\langle A, B \rangle$, then $|\psi| \ge \frac{w(\delta)}{t}$.

Proof. Consider $T_{\psi}^{\langle A,B\rangle}$ the syntax tree for the pair $\langle A,B\rangle$, ψ .

We associate to each node v of $T_{\psi}^{\langle A,B\rangle}$ a weight w(v):= $w(\delta_v)$, where δ_v is a minimal separator for il(v). From Lemmas VII.4 and VII.5, we get that $|\psi| = |T_{\psi}^{\langle A,B \rangle}| \ge \frac{w(\delta)}{t}$, where δ is a minimal separator of $il(r) = \langle A, B \rangle$.

Finally, as a consequence, we get Theorem VI.3.

Proof. Suppose ψ is an \mathcal{C}_3^t -sentence such that $\mathcal{A}_n \models \psi$ and $\mathcal{A}_m \models \neg \psi$. Let α be the assignment that assigns x, y and zto 0. Consider the separator $\delta : \mathcal{P}_2(\{\min, \max, x, y, z\}) \to \mathbb{N}$ defined as:

- $\delta(\{\min, \max\}) = n$,
- $\delta(\lbrace u, u' \rbrace) = 0$ for $(u, u') \in \mathcal{P}_2(\lbrace \min, \max, x, y, z \rbrace) \setminus$ $\{(min, max)\}.$

It is easy to see that $w(\delta) = \sqrt{n}$ and that δ is a minimal separator for $\langle (\mathcal{A}_n, \alpha), (\mathcal{A}_m, \alpha) \rangle$. Then Theorem VI.3 follows from Theorem VII.6 and this observation.



Fig. 4. The structure of the proof.

C. Proof of Lemma VII.4

The proof of Lemma VII.4 part 3. relies on the following key lemma:

Lemma VII.7. Let v be a node of T that has syntax-label $sl(v) = Q^{\geq k}u$ for $Q \in \{\exists, \forall\}, k \leq t \text{ and } u \in \{x, y, z\}$. Let δ_1 be a separator for $il(v_1)$, where v_1 is the unique child of v in T. Let δ be the separator defined via:

- $\delta(\{u, u'\}) := 0$, for all $u' \in \{\min, \max, x, y, z\} \setminus \{u\}$,
- $\delta(\{\min, \max\}) := MAX\{\delta_1(\{\min, \max\}), \delta_1(\{\min, u\}) + \delta_1(\{u, \max\}) + k 1\}.$ And for all u', u'' such that $\{x, y, z\} = \{u, u', u''\}$ and all $m \in \{\min, \max\}$:
- $\delta(\{u', u''\}) := MAX\{\delta_1(\{u', u''\}), \delta_1(\{u', u\}) + \delta_1(\{u, u''\}) + k 1\},$
- $\delta(\{m, u'\}) := MAX\{\delta_1(\{m, u'\}), \delta_1(\{m, u\}) + \delta_1(\{u, u'\}) + k 1\}.$

Then, δ is a separator for il(v).

Lemma VII.7 precisely captures how counting quantifiers

increase the distinguishing power of separators. From Lemma VII.7, we obtain Lemma VII.4 part 3. by building inequalities on the weight, which is done in Appendix D.

1) Overview of the proof of Lemma VII.7:

Proof. Due to symmetry, we only consider the case $\exists^{\geq k}z$. It has to be shown that δ is a separator for $\langle A, B \rangle = il(v)$. By definition, $il(v_1) = \langle A(F^k/z), B(*^k/z) \rangle$. Consider $I = (\mathcal{A}, \alpha) \in A$ and $J = (\mathcal{B}, \beta) \in B$. Let:

- $I_{F^k} := \{(\mathcal{A}, \alpha(F_i^k(\mathcal{A}, \alpha)/z)\}$ be the set of interpretations generated from I in v_1 , and let $a_i = F_i^k(\mathcal{A}, \alpha)$ and $I_{F^k}^i = (\mathcal{A}, \alpha(a_i/z)), 1 \le i \le k$.
- For any k-choice function G^k on B, let G be the choice function selected from G^k, and let J_G be the interpretation selected from J_{G^k} = {(B, β(G_i^k(B, β)/z))}. We define b_i := G_i^k(B, β); note that J_G corresponds to some b_j, 1 ≤ j ≤ k.

For every G^k on B, δ_1 is a separator for $\langle I_{F^k}, J_G \rangle$. Thus for every a_i there is a pair u_G^i, v_G^i in $\{min, max, x, y, z\}$ such that $\langle I_{F^k}^i, J_G \rangle$ is separated by $\delta_1(\{u_G^i, v_G^i\})$. The proof of the Lemma follows by considering several cases, as illustrated in Fig. 4. In each case, it is shown that δ separates $\langle I, J \rangle$. **Cases** 1-5 are *overlapping*, and after considering these cases we show that they cover all the possibilities, i.e. **Case 6** leads to a contradiction.

The first case, when z is not needed in δ_1 , is the same as the first case in [14]. **Cases 2-4** show that if there is an a_i in some *nearness relationship* with the interpretation (\mathcal{A}, α) according to δ_1 , then δ is a separator. In **Case 5**, we introduce the concept of *gap*, which is the set of elements not covered by the nearness conditions imposed by δ_1 . The condition of **Case 5** depends on the *number* of elements a_i in the gaps. This is the point where the additive term k enters the bound. Finally, in **Case 6** is shown that these cases cover all possibilities, as otherwise there is a set $\{b_1, \ldots, b_k\}$ duplicating $\{a_1, \ldots, a_k\}$ in the sense that no b_j allows δ_1 to separate. **Case 6** in particular explains the improvement of our proof over the result of [14] for FO, as a "+1" is introduced in [14] to deal with what would encompass that case.

x, y, and for all $i \in [k]$, $\alpha(F_i^k(\mathcal{A}, \alpha)/z)(u) = \alpha(u)$. So we will omit subscripts and just write $\alpha(x)$ for the assignment of x and min_A for the minimum in the k copies of (\mathcal{A}, α) (and similarly with y, max). Similarly, for all $u \in \{min, max, x, y\}$, for all $i \in [k]$, and for all $G^k \in F_B^k$,

 $\beta(G_i^k(\mathcal{B},\beta)/z)(u) = \beta(u)$, so we will also write $\beta(x)$ for the assignment of x and min_B for the minimum in the k copies of (\mathcal{B},β) (and similarly with y, max). We will say that $\delta(\{u, u'\})$ separates $\langle I, J \rangle$ to signify that the pair of variable $\{u, u'\}$ witnesses the separation property of δ on $\langle I, J \rangle$.

Case 1: There is G^k for which one of the $I^i_{F^k}$ is separated from J_G by $\delta_1(\{u, u'\})$ for $\{u, u'\} \in \{min, x, y, max\}$.

If there exists $\{u, u'\} \in \{min, x, y, max\}$ such that $\delta_1(\{u, u'\})$ separates $\langle I_{F^k}, J_G \rangle$ then by definition $\delta(\{u, u'\})$ separates $\langle I, J \rangle$.

Assumption 1 In the rest of the proof we will suppose that none of the $I_{F^k}^i$ is separated from J_G by $\delta_1(\{u, u'\})$ for $\{u, u'\} \in \{min, x, y, max\}.$

Without loss of generality, we may assume that $\alpha(x) \leq \alpha(y)$. If $\beta(x) > \beta(y)$, $\delta_1(\{x, y\})$ separates $\langle I_{F^k}, J_G \rangle$ for every G^k , which is a contradiction, so $\beta(x) \leq \beta(y)$.

To be able to consider the k choices simultaneously, we are looking to find for each choice a_i , a $b_i \in \mathcal{U}^{\mathcal{B}}$ that δ_1 will not be able to distinguish, in an adversary mindset. An element a_i has a particular location in (\mathcal{A}, α) with regards to $\delta_1(\{z, u\})$ for $u \in \{x, y, min, max\}$, namely that either $d(a_i, \alpha(u)) \leq \delta_1(\{z, u\})$ or $d(a_i, \alpha(u)) > \delta_1(\{z, u\})$. An a_i location is also based on its type (" = ", " < ", " > ") relative to $min_A, \alpha(x), \alpha(y), max_A$, which we will refer to as $\langle -type(a_i) (\langle -type(a_i, \alpha(x)), \langle -type(a_i, \alpha(y)), \langle -type(a_i, max_A)) \rangle$.

Case 2: There is an a_i and a $u \in \{x, y, min, max\}$ such that $d(a_i, \alpha(u)) \leq \delta_1(\{z, u\})$ and there is no $b \in \mathcal{U}^{\mathcal{B}}$ that satisfies $d(a_i, \alpha(u)) = d(b, \beta(u))$ and $\langle -\text{type}(a_i) = \langle -\text{type}(b) \rangle$.

We consider two sub-cases to better handle several variables at the same time. **Case 2.2** is represented in Fig. 5.

Case 2.1: There is an a_i and a $u \in \{x, y, min, max\}$ such that $d(a_i, \alpha(u)) \leq \delta_1(\{z, u\})$ and there is no $b \in \mathcal{U}^{\mathcal{B}}$ such that $d(a_i, \alpha(u)) = d(b, \beta(u))$ and $\langle -type(a_i, \alpha(u)) = \langle -type(b, \beta(u)) \rangle$.

We have four cases:

- *u* is min: then then $\delta(\{\min, \max\}) \ge \delta_1(\{\min, z\}) \ge d(a_i, \min_A) > d(\max_B, \min_B)$. Therefore:
 - $d(max_B, min_B) \neq d(max_A, min_A);$
 - $d(max_B, min_B) \leq \delta(\{min, max\}).$

So $\delta(\{\min, \max\})$ separates $\langle I, J \rangle$.

- u is max: the above reasoning applies by symmetry.
- u is x: If $\alpha(x) < a_i$, $\delta(\{x, max\}) \ge \delta_1(\{x, z\}) \ge d(a_i, \alpha(x)) > d(max_B, \beta(x))$. Then:
 - $d(max_A, \alpha(x)) \neq d(max_B, \beta(x));$
 - $d(max_B, \beta(x)) \leq \delta(\{x, max\}).$

So $\delta(\{x, max\})$ separates $\langle I, J \rangle$. If $\alpha(x) > a_i$, we reach a similar conclusion. Note that $\alpha(x) = a_i$ contradicts the assumption.

- u is y: the above reasoning applies by symmetry. \Box

Case 2.2: There is an a_i and a $u \in \{x, y, min, max\}$ such that $d(a_i, \alpha(u)) \leq \delta_1(\{z, u\})$ and there is no $b \in \mathcal{U}^{\mathcal{B}}$ such that $d(a_i, \alpha(u)) = d(b, \beta(u))$ and $\langle -type(a_i) = \langle -type(b) \rangle$.

We assume we are not in **Case 2.1**, so there exists a $b \in U^{\mathcal{B}}$ verifying $d(a_i, \alpha(u)) = d(b, \beta(u))$ and $\langle -type(a_i, \alpha(u)) = \langle -type(b, \beta(u)) \rangle$, note that such a b is unique, as we are working on linear orders. We fix b and call u' the variable in $\{x, y, min, max\}$ such that $\langle -type(a_i, \alpha(u')) \neq \langle -type(b, \beta(u')) \rangle$ by hypothesis of **Case 2.2**. The type is different, so it must be that both:

• $\mathbf{d}(\alpha(u'), \alpha(u)) \neq \mathbf{d}(\beta(u'), \beta(u));$

• MIN[d($\alpha(u'), \alpha(u)$), d($\beta(u'), \beta(u)$)] \leq d($a_i, \alpha(u)$). Since d($a_i, \alpha(u)$) $\leq \delta_1(\{z, u\}) \leq \delta(\{u', u\})$ we get:

• $\mathbf{d}(\alpha(u'), \alpha(u)) \neq \mathbf{d}(\beta(u'), \beta(u));$ • $\mathbf{MIN}[\mathbf{d}(\alpha(u'), \alpha(u)), \mathbf{d}(\beta(u'), \beta(u))] \leq \delta(\{u', u\})$

• MIN[d(
$$\alpha(u'), \alpha(u)$$
), d($\beta(u'), \beta(u)$)] $\leq \delta(\{u', u\})$.
So $\delta(\{u', u\})$ separates $\langle I, J \rangle$.

Assumption 2 In the rest of the proof, we will suppose that if there is an a_i and a $u \in \{x, y, min, max\}$ such that $d(a_i, \alpha(u)) \leq \delta_1(\{z, u\})$, then there is a $b \in \mathcal{U}^{\mathcal{B}}$ such that $d(a_i, \alpha(u)) = d(b, \beta(u))$ and $\langle -\text{type}(a_i) = \langle -\text{type}(b). \rangle$

We define $S_i := \{u \in \{x, y, min, max\} : d(a_i, \alpha(u)) \le \delta_1(\{z, u\})\}$, the set of variables to which a_i is "near".

Case 3: There is an a_i such that $|S_i| \ge 2$ and there is no $b \in \mathcal{U}^{\mathcal{B}}$ verifying $d(a_i, \alpha(u)) = d(b, \beta(u))$ for every $u \in S_i$.

We give a visual representation of this case in Fig. 7 in Appendix E.

Suppose there is no such b, we call u, u' two elements of S_i . Then Assumption 2 implies that $d(\alpha(u), \alpha(u')) \neq d(\beta(u), \beta(u'))$. Notice that $d(\alpha(u), \alpha(u')) \leq d(a_i, \alpha(u)) + d(a_i, \alpha(u')) \leq \delta_1(\{z, u\}) + \delta_1(\{z, u'\}) \leq \delta(\{u, u'\})$. So: • $d(\alpha(u), \alpha(u')) \neq d(\beta(u), \beta(u'))$;



Fig. 5. Case 2 of the proof of Lemma VII.7.

We associate a ball with center $\alpha(u)$ of radius $\delta_1(u, v)$ for every $u \in \{x, y, min, max\}$ on \mathcal{A} (resp. $\beta(u)$ on \mathcal{B}). Each ball associated to variable u accounts for the ability of the pair $\{u, z\}$ to separate. The crossed dot on \mathcal{A} represents a choice on the linear order from F^k , the crossed dot on \mathcal{B} represents "a best attempt" at a response for indistinguishability. In **Case 2**, the choice on \mathcal{A} is in a nearness relationship (for $\delta_1(\{x, z\})$ on the figure), and there is no element of \mathcal{B} that matches the separating conditions for indistinguishability (in the figure, $d(a_1, \alpha(x)) = d(b_1, \beta(x))$ but $\langle -type(a_1) \neq \langle -type(b_1) \rangle$).

• $d(\alpha(u), \alpha(u')) \le \delta(\{u, u'\}).$ Therefore $\delta(\{u, u'\})$ separates $\langle I, J \rangle$.

Assumption 3 In the rest of the proof, we will suppose that if there is an a_i such that $|S_i| \ge 2$, then there is $b \in \mathcal{U}^{\mathcal{B}}$ such that $d(a_i, \alpha(u)) = d(b, \beta(u))$ for all $u \in S_i$.

Case 4: There is an a_i and there are $u, u' \in \{x, y, min, max\}$ such that $d(a_i, \alpha(u)) > \delta_1(\{z, u\})$ and $d(a_i, \alpha(u')) \leq \delta_1(\{z, u'\})$ and there is no $b \in \mathcal{U}^{\mathcal{B}}$ that satisfies $d(a_i, \alpha(u')) = d(b, \beta(u')), <-type(a_i) = <-type(b)$ and $d(b, \beta(u)) > \delta_1(\{z, u\}).$

We give a visual representation of this case in Fig. 8 in Appendix E.

Assume such an a_i exists, we set $u, u' \in \{x, y, min, max\}$ and $b \in \mathcal{U}^{\mathcal{B}}$ such that $d(a_i, \alpha(u')) = d(b, \beta(u'))$, <type $(a_i) = \langle -type(b), note that such a b exists by Assumption$ **2**.

By hypothesis $d(b,\beta(u)) < d(a_i,\alpha(u))$, so we must have that $d(\beta(u'),\beta(u)) \neq d(\alpha(u'),\alpha(u))$. Notice that $d(\beta(u'),\beta(u)) \leq d(b,\beta(u))+d(b,\beta(u')) \leq \delta_1(\{z,u\}) + \delta_1(\{z,u'\}) \leq \delta(\{u,u'\})$. So:

•
$$\mathbf{d}(\alpha(u'), \alpha(u)) \neq \mathbf{d}(\beta(u'), \beta(u));$$

•
$$d(\beta(u'), \beta(u)) \le \delta(\{u, u'\}).$$

So $\delta(\{u, u'\})$ separates $\langle I, J \rangle$.

Assumption 4 In the rest of the proof, we will suppose that if there is an a_i and $u, u' \in \{x, y, min, max\}$ such that $d(a_i, \alpha(u)) > \delta_1(\{z, u\})$ and $d(a_i, \alpha(u')) \le \delta_1(\{z, u'\})$ then there is $b \in \mathcal{U}^{\mathcal{B}}$ such that $d(a_i, \alpha(u')) = d(b, \beta(u'))$ and $d(b, \beta(u)) > \delta_1(\{z, u\})$.

Before proceeding to the next case, we define quantities, called gaps, to characterize the elements of a linear order not in any nearness relationship to any of the assignments.

Definition VII.8 (Gap). For $(\mathcal{U}, \gamma) \in \{(\mathcal{A}, \alpha), (\mathcal{B}, \beta)\}$, we derive three ¹ particular quantities called gaps:

$$Gap_{(\mathcal{U},\gamma)}(\{min,x\}) = \{a \in [min_U,\gamma(x)] | \\ \forall u \in \{x, y, min, max\}, d(a,\gamma(u)) > \delta_1(\{u,z\})\}$$

$$Gap_{(\mathcal{U},\gamma)}(\{y, max\}) = \{a \in [\gamma(y), max_U] | \\ \forall u \in \{x, y, min, max\}, d(a, \gamma(u)) > \delta_1(\{u, z\})\}$$

¹The three cases are analogous, but a unified formal definition appears to be less readable.

$$Gap_{(\mathcal{U},\gamma)}(\{x,y\}) = \{a \in [\gamma(x),\gamma(y)] | \\ \forall u \in \{x,y,min,max\}, d(a,\gamma(u)) > \delta_1(\{u,z\})\}$$

The motivation for gaps is that a separator cannot distinguish between two elements located within the same gap. We will refer to as "gap variables" any two variables among $\{x, y, min, max\}$ that bound a given gap². More formally:

Definition VII.9 (Gap variables). For an interpretation (\mathcal{U}, γ) and a pair of variables $(u, u') \in \{(min, x), (x, y), (y, max)\}$, the gap variables (v, v') of $Gap_{(\mathcal{U}, \gamma)}(\{u, u'\})$ are defined as:

• (u, u') = (min, x): (v, v') is a pair of variables in $\{(min, x), (min, y), (min, max)\}$ such that

$$|Gap_{(\mathcal{U},\gamma)}(\{min,x\})| + \delta_1(\{z,v\}) + \delta_1(\{z,x\}) \ge d(\gamma(v'),\gamma(v))$$

• (u, u') = (y, max): (v, v') is a pair of variables in $\{(y, max), (x, max), (min, max)\}$ such that

$$|Gap_{(\mathcal{U},\gamma)}(\{y,max\})| + \delta_1(\{z,v\}) + \delta_1(\{z,x\}) \ge d(\gamma(v'),\gamma(v))$$

 (u, u') = (x, y): (v, v') is a pair of variables in {(x, y), (min, y), (x, max), (min, max)} such that

$$|Gap_{(\mathcal{U},\gamma)}(\{x,y\})| + \delta_1(\{z,v\}) + \delta_1(\{z,x\}) \ge d(\gamma(v'),\gamma(v))$$

In the definition, the (v, v') are "bounding" the gap (u, u')on U. It is clear that such pairs (v, v') exist by the definition of gaps, and if multiple pairs verify one of the inequalities, we pick the smallest pair in lexicographic order to be the gap variables. We give a visual representation of the definitions in Fig. 10 in Appendix F.

A gap $\{u, u'\}$ needs not have gap variables $\{u, u'\}$. For example, the gap $\{min, x\}$ might have gap variables $\{min, y\}$ when the $\delta_1(y, v)$ radius circle contains the $\delta_1(x, v)$ radius circle, hence covering a bigger chunk of the $\{min, x\}$ interval. If the gap is nonempty, the inequalities in Definition VII.9 become equalities, as the two circles corresponding to the gap variables are non-overlapping, and so the two radii added to the length of the gap add up to the length of the interval.

We derive a couple results for gap variables before proceeding to the next case:

Proposition VII.10. If there are $u, u' \in \{x, y, min, max\}$ such that MIN $[d(\alpha(u), \alpha(u')), d(\beta(u), \beta(u'))] \leq \delta_1(\{z, u\})$ then $d(\alpha(u), \alpha(u')) = d(\beta(u), \beta(u')).$

Proof. Suppose not, since $\delta_1(\{z, u\}) \leq \delta(\{u, u'\})$ we have:

• $\mathbf{d}(\alpha(u), \alpha(u')) \neq \mathbf{d}(\beta(u), \beta(u')),$

• MIN
$$[d(\alpha(u), \alpha(u')), d(\beta(u), \beta(u'))] \le \delta(\{u, u'\})$$

So $\delta(\{u, u'\})$ separates $\langle I, J \rangle$.

Proposition VII.11. Let v, v' are the gap variables associated with $Gap_{(\mathcal{A},\alpha)}(\{u, u'\})$ if and only if v, v' are the gap variables associated with $Gap_{(\mathcal{B},\beta)}(\{u, u'\})$.

Proof. Let v, v' be the gap variables associated with $Gap_{(\mathcal{A},\alpha)}(\{u, u'\})$, and suppose w, w' are the gap variables associated with $Gap_{(\mathcal{B},\beta)}(\{u, u'\})$.

We will first reason from (\mathcal{A}, α) to (\mathcal{B}, β) , and suppose that $\alpha(v) < \alpha(w) \le \alpha(u)$. By the definition of the gap on (\mathcal{A}, α) , it must be that $\delta_1(\{z, v\}) \ge d(\alpha(w), \alpha(v)) + \delta_1(\{z, w\})$. Then proposition VII.10 implies that $d(\alpha(w), \alpha(u)) = d(\beta(w), \beta(u))$. So $\delta_1(\{z, v\}) \ge d(\beta(w), \beta(v)) + \delta_1(\{z, w\})$, therefore w cannot be a gap variable for $Gap_{(\mathcal{B},\beta)}(\{u, u'\})$. The same reasoning applies from the gap on (\mathcal{B}, β) to the gap on (\mathcal{A}, α) , yielding the result.

Case 5: There is $(u, u') \in \{(\min, x), (x, y), (y, \max)\}$ such that there are p choices of F^k in $Gap_{(\mathcal{A},\alpha)}$ and $|Gap_{(\mathcal{B},\beta)}(\{u, u'\})| < p$.

We give a visual representation of this case in Fig. 9 in Appendix E.

By proposition VII.10, there is a unique pair v, v' of gap variables for both $Gap_{(\mathcal{A},\alpha)}(\{u, u'\})$ and $Gap_{(\mathcal{B},\beta)}(\{u, u'\})$. On one hand by the gap variables definition on (\mathcal{B},β) , we have $d(\beta(v'),\beta(v)) \leq \delta_1(\{z,v\}) + \delta_1(\{z,v'\}) + |Gap_{(\mathcal{B},\beta)}(\{u,u'\})| \leq \delta_1(\{z,v\}) + \delta_1(\{z,v'\}) + (k-1) \leq \delta(\{v,v'\})$.

On the other hand by the gap variables definition on (\mathcal{A}, α) , $d(\alpha(v'), \alpha(v)) = \delta_1(\{z, v\}) + \delta_1(\{z, v'\}) + |Gap_{(\mathcal{A}, \alpha)}(\{u, u'\})| > d(\beta(v'), \beta(v))$. Finally $\delta(\{v, v'\})$ separates $\langle I, J \rangle$.

Assumption 5 In the rest of the proof, we assume that if there are p choices of F^k in $Gap_{(\mathcal{A},\alpha)}(\{u, u'\})$ for $(u, u') \in \{(min, x), (x, y), (y, max)\}$ then $|Gap_{(\mathcal{B},\beta)}(\{u, u'\})| \ge p$.

Case 6: Remaining cases. Note that $S_i = \emptyset$ if and only if a_i is in a gap. Recall the assumptions so far in the proof:

- From Assumption 1: for every G^k, ⟨I_{F^k}, J_G⟩ can only be distinguished by δ₁(z, u) for u ∈ {x, y, min, max};
- From Assumptions 2 to 4: for every a_i such that $S_i \neq \emptyset$, we have that there exists a unique $b_i \in \mathcal{U}^{\mathcal{B}}$ such that $\langle -\text{type}(a_i) = \langle -\text{type}(b_i), d(b_i, \beta(u)) = d(a_i, \alpha(u)) \text{ for}$ all $u \in S_i$ and $d(b_i, \beta(u)) > \delta_1(\{z, u\})$ for all $u \in \{x, y, \min, \max\} \setminus S_i$.
- From Assumption 5: for every a_i in Gap_(A,α)({u, u'}) for (u, u') ∈ {(min, x), (x, y), (y, max)}, there exists a distinct b_i ∈ Gap_(B,β)({u, u'}).

Consider the choice function G^k consisting of the k b_i 's matching the k a_i 's as described above. Suppose index j from the choice function G^k is selected to create J_G . Then by construction $\delta_1(\{z, u\})$ does not distinguish $\langle I_{F_j^k}, J_G \rangle$ for any $u \in \{x, y, min, max\}$, which contradicts **Assumption 1**. The situation is represented in Fig. 6. In conclusion, the δ defined separates $\langle I, J \rangle$.

 $^{^{2}}$ A note on terminology: here *min* and *max* are seen as arguments of a separator function.



Fig. 6. $\langle I_{F^k}, J_G \rangle$ cannot be separated by δ_1 in **Case 6** of the proof of Lemma VII.7. We associate a ball with center $\alpha(u)$ of radius $\delta_1(u, v)$ for every $u \in \{x, y, min, max\}$ on \mathcal{A} (resp. $\beta(u)$ on \mathcal{B}). Each ball associated to variable u accounts for the ability of the pair $\{u, z\}$ to separate. The regions of the linear orders that are not covered by any of the balls are the gaps, in purple on the Figure (in the example, the $\{min, x\}$ gap is empty. The crossed dots on \mathcal{A} represents the choices on the linear order from F^k , the crossed dots on \mathcal{B} represent the best attempts at a response for indistinguishability. In **Case 6**, every choice on \mathcal{A} can be matched by a different δ_1 -indistinguishable element of \mathcal{B} .

VIII. CONCLUSION AND PERSPECTIVES

In this paper an EF game is formulated for formula size for counting logic. It is used to prove a \sqrt{n}/t lower bound for the size of 3-variable counting logic formulae with counting rank t, distinguishing a linear order of size n from a larger one. The lower bound extends a $\Omega(\sqrt{n})$ lower bound of [14]. The proof is based on the approach of [14], with a different argument for the central case of handling counting quantifiers. Closing the gap between the lower bound and the upper bound of size O(n/t) is an open problem. This is open even in the FO case where, as far as we know, no improvement is known of the linear upper bound.

The lower bound has some implications for comparing formula sizes of various fragments of counting logic. Comparing the succinctness of various knowledge representation formalisms is studied in detail in knowledge compilation [5], Boolean complexity theory [23] and other areas, but, as noted in [14], perhaps less so for predicate logic. Bounds comparing m-variable counting logic formula sizes for m = 2, 3, 4 will be included in the final version of this paper. Separating the expressivity of the 2- and 3-variable fragments is an open problem, and it seems to be open for the FO case as well.

We conclude with a brief description of the connection between GNN and counting logic formula size, which is one of the motivations of this work and the subject of ongoing work. A GNN works on a graph with feature vectors assigned to the nodes. In each round these are updated by applying a *combination* function to the previous vector and an *aggregate* of the feature vectors of the neighbors [17]. A *logical classifier* computes a unary query on graphs (e.g., assigning to every graph the set of red vertices with all blue neighbors). Barceló et al. [2] showed that an FO logical classifier is computable by a GNN iff it is definable in 2-variable guarded counting logic. See also [12], [13], also noting that this is a "uniform" model. Every such formula has a GNN simulation with complexity (number of features and rounds) depending on the complexity of the formula.

In ongoing work with coauthors we study GNN learning a query corresponding to a formula. One question we consider is: to what extent can the model learned on one class of graphs be transferred to some other class of graphs? How does performance depend on the formula and the graph classes? The answer seems to be related to the connection between counting logic and the Weisfeiler-Leman algorithm mentioned in the introduction, and formula complexity may play a role here. Another related question, recently considered in [30], is whether the underlying formula be extracted from a learned GNN using explainability techniques?

References

- ADLER, M., AND IMMERMAN, N. An n! lower bound on formula size. ACM Trans. Comput. Log. 4, 3 (2003), 296–314.
- [2] BARCELÓ, P., KOSTYLEV, E. V., MONET, M., PÉREZ, J., REUTTER, J. L., AND SILVA, J. P. The logical expressiveness of graph neural networks. In 8th International Conference on Learning Representations, ICLR 2020 (2020).
- [3] CAI, J., FÜRER, M., AND IMMERMAN, N. An optimal lower bound on the number of variables for graph identification. *Comb.* 12, 4 (1992), 389–410.
- [4] CARMOSINO, M., FAGIN, R., IMMERMAN, N., KOLAITIS, P. G., LENCHNER, J., AND SENGUPTA, R. A finer analysis of multi-structural games and beyond. *CoRR abs/2301.13329* (2023).
- [5] DARWICHE, A., AND MARQUIS, P. A knowledge compilation map. J. Artif. Intell. Res. 17 (2002), 229–264.
- [6] EBBINGHAUS, H., AND FLUM, J. *Finite model theory*. Perspectives in Mathematical Logic. Springer, 1995.
- [7] EHRENFEUCHT, A. An application of games to the completeness problem for formalized theories. *Fundamenta Mathematicae* (1961).
- [8] ETESSAMI, K. Counting quantifiers, successor relations, and logarithmic space. In Proceedings of the Tenth Annual Structure in Complexity Theory Conference, Minneapolis, Minnesota, USA, June 19-22, 1995 (1995), IEEE Computer Society, pp. 2–11.
- [9] ETESSAMI, K. Counting quantifiers, successor relations, and logarithmic space. J. Comput. Syst. Sci. 54, 3 (1997), 400–411.
- [10] FAGIN, R., LENCHNER, J., VYAS, N., AND WILLIAMS, R. R. On the number of quantifiers as a complexity measure. In 47th International Symposium on Mathematical Foundations of Computer Science, MFCS 2022, August 22-26, 2022, Vienna, Austria (2022), S. Szeider, R. Ganian, and A. Silva, Eds., vol. 241 of LIPIcs, Schloss Dagstuhl - Leibniz-Zentrum für Informatik, pp. 48:1–48:14.
- [11] FRAISSE, R. Sur quelques classifications des systemes de relations.
- [12] GROHE, M. The logic of graph neural networks. In 36th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS (2021), IEEE, pp. 1–17.
- [13] GROHE, M. The descriptive complexity of graph neural networks. In LICS (2023), pp. 1–14.
- [14] GROHE, M., AND SCHWEIKARDT, N. The succinctness of first-order logic on linear orders. *Log. Methods Comput. Sci.* 1, 1 (2005).
- [15] HAJNAL, A., MAASS, W., PUDLÁK, P., SZEGEDY, M., AND TURÁN, G. Threshold circuits of bounded depth. In 28th Annual Symposium on Foundations of Computer Science (1987), IEEE Computer Society, pp. 99–110.
- [16] HAJNAL, A., MAASS, W., PUDLÁK, P., SZEGEDY, M., AND TURÁN, G. Threshold circuits of bounded depth. J. Comput. Syst. Sci. 46, 2 (1993), 129–154.
- [17] HAMILTON, W. L. Graph Representation Learning. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers, 2020.
- [18] HELLA, L. Logical hierarchies in PTIME. Inf. Comput. 129, 1 (1996), 1–19.
- [19] HELLA, L., AND VÄÄNÄNEN, J. The size of a formula as a measure of complexity. In Logic Without Borders - Essays on Set Theory, Model Theory, Philosophical Logic and Philosophy of Mathematics, Å. Hirvonen, J. Kontinen, R. Kossak, and A. Villaveces, Eds., vol. 5 of Ontos Mathematical Logic. De Gruyter, 2015, pp. 193–214.
- [20] IMMERMAN, N. Number of quantifiers is better than number of tape cells. J. Comput. Syst. Sci. 22, 3 (1981), 384–406.
- [21] IMMERMAN, N. *Descriptive complexity*. Graduate texts in computer science. Springer, 1999.
- [22] IMMERMAN, N., AND LANDER, E. Describing graphs: A first-order approach to graph canonization. 1990.
- [23] JUKNA, S. Boolean Function Complexity Advances and Frontiers, vol. 27 of Algorithms and combinatorics. Springer, 2012.
- [24] KARCHMER, M., AND WIGDERSON, A. Monotone circuits for connectivity require super-logarithmic depth. SIAM J. Discret. Math. 3, 2 (1990), 255–265.

- [25] KUSKE, D., AND SCHWEIKARDT, N. First-order logic with counting: At least, weak hanf normal forms always exist and can be computed! *CoRR abs/1703.01122* (2017).
- [26] LIBKIN, L. Elements of Finite Model Theory. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2004.
- [27] MORRIS, C., RITZERT, M., FEY, M., HAMILTON, W. L., LENSSEN, J. E., RATTAN, G., AND GROHE, M. Weisfeiler and leman go neural: Higher-order graph neural networks. In *The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019* (2019), AAAI Press, pp. 4602–4609.
- [28] OTTO, M. Bounded Variable Logics and Counting: A Study in Finite Models, vol. 9 of Lecture Notes in Logic. Cambridge University Press, 2017.
- [29] OTTO, M. Graded modal logic and counting bisimulation. CoRR abs/1910.00039 (2019).
- [30] PLUSKA, A., WELKE, P., GÄRTNER, T., AND MALHOTRA, S. Logical distillation of graph neural networks. In Proceedings of the 21st International Conference on Principles of Knowledge Representation and Reasoning, KR 2024 (2024).
- [31] SCARSELLI, F., GORI, M., TSOI, A. C., HAGENBUCHNER, M., AND MONFARDINI, G. The graph neural network model. *IEEE Trans. Neural Networks* 20, 1 (2009), 61–80.
- [32] VINALL-SMEETH, H. From quantifier depth to quantifier number: Separating structures with k variables, 2024.
- [33] XU, K., HU, W., LESKOVEC, J., AND JEGELKA, S. How powerful are graph neural networks? In 7th International Conference on Learning Representations, ICLR 2019 (2019).

APPENDIX A PROOF OF THE COUNTING GAME FOR SIZE CHARACTERISATION THEOREM

Proof. We use induction on w. If w = 1, Spoiler wins in $CS_1^m(A, B)$ only if there is an atomic formula ϕ , so verifying $|\phi| = 1$, such that $(A, B) \models \phi$. Reciprocally, if $|\phi| = 1$ and $(A, B) \models \phi$, then ϕ is an atomic formula and spoiler wins in $CS_1^m(A, B)$.

We now assume the equivalence for v < w.

To prove the forward direction of the equivalence, suppose Spoiler has a winning strategy in the game $CS_w^m(A, B)$ starting with:

- ¬-move. Spoiler has a winning strategy in $CS_{w-1}^{m}(B, A)$, so by induction hypothesis there is a formula ψ verifying $|\psi| \le w 1$ and $(B, A) \models \psi$. Finally $(A, B) \models \neg \psi$ and $|\neg \psi| = 1 + |\psi| \le w$ so we get (2).
- V-move, choosing u, v, C and D such that 1 ≤ u, v < w with u + v = w, splitting A into C and D. Spoiler has a winning strategy in CS^m_u(C, B) and CS^m_u(D, B), so by the induction hypothesis, there are formulae ψ and θ such that |ψ| ≤ u, |θ| ≤ v, (C, B) ⊨ ψ and (D, B) ⊨ θ.

We have that $C \models \psi$ and $D \models \theta$, so $A \models \psi \lor \theta$. Conversely, $B \models \neg \psi$ and $B \models \neg \theta$, and consequently $B \models \neg (\psi \lor \theta)$.

It follows that $(A, B) \models \psi \lor \theta$, and since $|\psi \lor \theta| = |\psi| + |\theta| \le u + v = w$, we get (2).

- ∧-move; and by a symmetrical argument to the ∨-move we get (2).
- $\exists^{\geq k}$ -move; choosing $j \in [m], k \in \mathbb{N}$, a k-choice function F^{k} on A and for every $G^{k} \in F_{B}^{k}$, selecting G from G^{k} to form $B(*^{k}/u) = \bigcup_{G^{k} \in F_{B}^{k}} B(G/u)$. Spoiler has a winning strategy in $CS_{w-1}^{m}(A(F^{k}/j), B(*^{k}/j))$, so by the induction hypothesis, there is a formula ψ such that $|\psi| \leq w 1$ and $(A(F^{k}/j), B(*^{k}/j)) \models \psi$.

Define $\phi := \exists^{\geq k} x_j \psi$. We have that $A(F^k/j) \models \psi$, and so for all $1 \leq i \leq k$, $(\mathcal{A}, \alpha(F_i^k/j)) \models \psi$ and therefore $A \models \phi$.

Now suppose that there exists a k-choice function G^k on B and a $(\mathcal{B},\beta) \in B$ such that $\{(\mathcal{B},\beta(G_i^k(\mathcal{B},\beta)/j)) : 1 \leq i \leq k\} \models \psi$. Then there is an $i \in [k]$ such that $(\mathcal{B},\beta(G_i^k(\mathcal{B},\beta)/j))$ is part of the family $B(*^k/j)$ through the selected G, but $B(*^k/j) \models \neg \psi$, which is a contradiction. Therefore $B \models \neg \phi$.

Finally $(A, B) \models \phi$ and $|\phi| = |\psi| + 1 \le w$, so we get (2).

∀^{≥k}-move; a symmetrical argument to the V-move yields
 (2).

To prove the backward direction, assume there is a formula ϕ of size w > 1 such that $(A, B) \models \phi$. We show that Spoiler has a winning strategy in $CS_w^m(A, B)$. Let us consider the form of the formula ϕ :

• $\phi = \neg \psi$: Spoiler plays the \neg -move and gets in the position $CS_{w-1}^m(B, A)$. Since $|\psi| < w$ and $(B, A) \models \phi$, Spoiler has a winning strategy in $CS_{w-1}^m(B, A)$ by induc-

tion hypothesis, and therefore Spoiler has also a winning strategy in $CS_w^m(A, B)$.

- $\phi = \psi \lor \theta$: Define $C = \{(\mathcal{A}, \alpha) \in A | (\mathcal{A}, \alpha) \models \psi\},$ $D = \{(\mathcal{A}, \alpha) \in A | (\mathcal{A}, \alpha) \models \theta\}, u \text{ and } v \text{ such that}$ $w = u + v, |\psi| \le u \text{ and } |\theta| \le v.$ Spoiler plays the \bigvee -move associated to u, v, C, D and gets to $\operatorname{CS}_u^m(C, B)$ or $\operatorname{CS}_v^m(D, B)$ according to Duplicator's choice. We have that $(C, B) \models \psi, (D, B) \models \theta$ and therefore by induction hypothesis Spoiler has a winning strategy in $\operatorname{CS}_u^m(C, B)$ or $\operatorname{CS}_v^m(D, B)$, and therefore Spoiler has also a winning strategy in $\operatorname{CS}_w^m(A, B)$.
- $\phi = \psi \wedge \theta$: a symmetrical argument shows that Spoiler has a winning strategy in $CS_w^m(A, B)$.
- φ = ∃^{≥k}x_jψ: Since A ⊨ φ, there is a k-choice function F^k on A such that (A, α(F^k_i((A, α)/j)) ⊨ ψ, for every 1 ≤ i ≤ k and for all (A, α) ∈ A. Thus, A(F^k/j) ⊨ ψ. On the other hand, B ⊨ ¬φ, thus from every G^k on B we can select G such that B(G/j) ⊨ ¬ψ and therefore, B(*^k/j) ⊨ ¬ψ.

Finally we have that $(A(F^k/j), B(*^k/j)) \models \psi$ and $|\psi| = |\phi| - 1 < w$, so by induction hypothesis, Spoiler has a winning strategy in the game

 $CS_{w-1}^m(A(F^k/j), B(*^k/j))$ and therefore Spoiler has also a winning strategy in $CS_w^m(A, B)$.

• $\phi = \forall^{\geq k} x_j \psi$: a symmetrical argument shows that Spoiler has a winning strategy in $CS_w^m(A, B)$.

APPENDIX B

PROOF OF THEOREM VI.1

We want to show that Duplicator can win a k-round EF game on L_1 and L_2 taken to be of length at least $(t + 1)^k$. The key idea behind the proof is that, from a given assignment, Spoiler can "only look up to a certain distance" in the two directions on the line.

Proof. Suppose L_1 and L_2 are linear orders of length at least $(t+1)^k$, on which the EF game will be played.

After *i* moves, we denote by *a* the "position": *a* is a tuple consisting of min_{L_1}, max_{L_1} concatenated to the *i* moves played on L_1 : $a = (a_{-1}, a_0, a_1, \ldots, a_i), a_{-1} = min_{L_1}$, $a_0 = max_{L_1}$). Similarly, we define the tuple *b* of moves played on L_2 .

For $-1 \leq j, l \leq i$, we prove that regardless of Spoiler's choices, Duplicator can maintain the following inequalities:

- 1) if $d(a_j, a_l) \le (t+1)^{k-i}$, then $d(b_j, b_l) = d(a_j, a_l)$. 2) if $d(a_j, a_l) > (t+1)^{k-i}$, then $d(b_j, b_l) > (t+1)^{k-i}$.
- 3) $a_j \leq a_l$ iff $b_j \leq b_l$.

Using those inequalities for i = k moves, property 3. yields that Duplicator win the k-round EF game on L_1 and L_2 . Since this happens no matter what Spoiler plays, the game characterisation theorem for bounded counting rank implies that $L_1 \equiv_{\mathcal{C}^t[k]} L_2$.

We now prove by induction on the moves (on i) that the inequalities can be maintained. The base case of i = 0 is immediate.

For the induction step, assume the inequalities stand for i moves, and suppose without loss of generality that Spoiler makes his $(i + 1)^{\text{th}}$ move on L_1 . Spoiler plays the set M of cardinality k, in L_1 . We describe Duplicator's response, a set N on L_2 . In the second part of the move, if Spoiler picks N_q on L_2 for $q \in [t]$ (i.e. $b_{i+1} = N_q$), Duplicator will respond with M_q on L_1 (i.e. $a_{i+1} = M_q$).

If $a_j = M_q$ for $j \le i$ and $q \in [t]$, Duplicator sets N_q to be b_j . Suppose there is an element $M_q \in M$ for which this is not the case. We define $j, l \le i$ such that $a_j < M_q < a_l$ and that there is no other previously played moves on L_1 inside this interval. By property 3., the interval between b_j and b_l contains no other elements of b. Then we have two cases regarding the length of the interval:

- If $d(a_j, a_l) \leq (t+1)^{k-i}$. Then by property 1. and inductive hypothesis, $d(b_j, b_l) = d(a_j, a_l)$, and the intervals $[a_j, a_l]$ and $[b_j, b_l]$ are isomorphic. Duplicator picks N_q so that $d(a_j, M_q) = d(b_j, N_q)$ and $d(M_q, a_l) = d(M_q, b_l)$, which ensures that the three properties hold for i + 1 moves.
- $d(a_j, a_l) > (t+1)^{k-i}$. In this case by property 2., $d(b_j, b_l) > (t+1)^{k-i}$. We have three possibilities:
 - If $d(a_j, M_q) \leq (t+1)^{k-(i+1)}$. Then $d(M_q, a_l) > (t+1)^{k-(i+1)}$, and Duplicator picks N_q on $[b_j, b_l]$ so that $d(b_j, N_q) = d(a_j, M_q)$ maintaining properties 1 and 3. Since $d(M_q, b_l) > (t+1)^{k-(i+1)}$ this maintains $d(N_q, b_l) > (t+1)^{k-(i+1)}$ by property 2.(i), hence property 2.(i+1).
 - $d(M_q, a_l) \leq (t+1)^{k-(i+1)}$, in which case a similar reasoning applies.
 - Otherwise, Spoiler has picked M_q such that both: $d(a_j, M_q) > (t + 1)^{k-(i+1)}, \ d(M_q, a_l) > (t + 1)^{k-(i+1)}.$ There can be at most t such M_q as |M| = t.

Since $d(b_j, b_l) > (t+1)^{k-i}$, there are at least t distinct elements in L_2 between $b_j + (t+1)^{k-(i+1)}$ and $b_l - (t+1)^{k-(i+1)}$. Duplicator picks the first N_q on that interval that is not already in N. This ensures that, $d(b_j, N_q) > (t+1)^{k-(i+1)}$ and $d(N_q, b_l) > (t+1)^{k-(i+1)}$, therefore satisfying all three properties.

Thus, in all the cases, the induction is preserved. \Box

APPENDIX C

PROOF OF LEMMA VII.4 PARTS 1 AND 2

A. Part 1

Let v be a leaf of T, by definition VII.1, v if of the form $T_{\psi}^{\langle C,D\rangle}$ where ψ is an atomic formula, say R(u,u') where $u, u' \in \{x, y, min, max\}$ and $R \in \{<, =, succ\}$. Since $C, D \models R(u, u')$, it is easy to verify that for all $(\mathcal{A}, \alpha) \in C$ and $(\mathcal{B}, \beta) \in D$:

- 1) $;-type(\alpha(u), \alpha(u')) \neq ;-type(\beta(u), \beta(u')) \text{ or }$
- 2) both:
 - $MIN[d(\alpha(u), \alpha(u')), d(\beta(u), \beta(u'))] \le 1$ and,
 - $\mathbf{d}(\alpha(u), \alpha(u')) \neq \mathbf{d}(\beta(u), \beta(u')).$

First if $u \neq u'$, consider the separator of (C, D), δ_1 , defined as:

$$\delta_1 = \begin{cases} 1 & \text{if } p = \{u, u'\} \\ 0 & \text{otherwise.} \end{cases}$$

Since $w(\delta_1) = 1$, any minimal separator of C < D has weight at most 1.

Now, if u = u', then $C = \emptyset$ or $D = \emptyset$, and then $\delta_1 = 0$ is a separator for C, D.

Finally, $w(\delta) \leq 1$ for δ a minimal separator of v.

B. Part 2

Let v be a node of T of the form $T_{\psi}^{\langle C,D\rangle}$ with two children v_1 and v_2 . Let δ_1 and δ_2 be minimum separators of v_1 and v_2 respectively. We first prove the straightforward proposition:

Proposition C.1. $\delta' = \delta_1 + \delta_2$ is a separator of v.

Proof. Let us fix $(\mathcal{A}, \alpha) \in C$ and $(\mathcal{B}, \beta) \in D$. Suppose ψ is of the form $\psi_1 \vee \psi_2$. The case ψ of the form $\psi_1 \wedge \psi_2$ is symmetrical.

By definition VII.1, there are C_1, C_2 such that δ_1 is a separator of $T_{\psi_1}^{\langle C_1, D \rangle}$ and δ_2 is a separator of $T_{\psi_2}^{\langle C_2, D \rangle}$ and with $C_1 \cup C_2 = C$.

Whether $(\mathcal{A}, \alpha) \in C_1$ or $(\mathcal{A}, \alpha) \in C_2$, since $\delta' \ge \delta_1, \delta_2$ it is clear that δ' defined as above satisfies the separation condition:

- 1) $i\text{-type}(\alpha(u), \alpha(u')) \neq i\text{-type}(\beta(u), \beta(u'))$ or 2) both:
 - MIN[d(α(u), α(u')), d(β(u), β(u'))] ≤ δ'({u, u'}) and,
 d(α(u), α(u')) ≠ d(β(u), β(u')).

Therefore δ' is a separator of v.

We conclude with the upper bounding of the weight by showing that $w(\delta') \leq w(\delta_1) + w(\delta_2)$. By definition VII.3, it is easy to check that both $b(\delta') \leq b(\delta_1) + b(\delta_2)$ and $c(\delta') \leq c(\delta_1) + c(\delta_2)$ and therefore:

$$w(\delta')^{2} = c(\delta')^{2} + b(\delta')$$

$$\leq (c(\delta_{1}) + c(\delta_{2})^{2} + b(\delta_{1}) + b(\delta_{2})$$

$$= c(\delta_{1})^{2} + c(\delta_{2})^{2} + b(\delta_{1}) + b(\delta_{2}) + 2c(\delta_{1})c(\delta_{2})$$

$$\leq w(\delta_{1})^{2} + w(\delta_{2})^{2} + 2w(\delta_{1})w(\delta_{2})$$

$$= (w(\delta_{1}) + w(\delta_{2}))^{2}$$

APPENDIX D

END OF THE PROOF OF LEMMA VII.4 PART 3.

We derive Lemma *VII.*4 part 3. from the consequence on the separator weight of the key Lemma VII.7:

Proof. Suppose v be a node of T that has exactly one child v_1 , we consider δ a minimal separator of il(v) and δ_1 be a minimal separator of $\langle A_1, B_1 \rangle := il(v_1)$.

We want to show that $w(\delta) \le w(\delta_1) + t$. From the definition of a syntax tree we know that either $sl(v) = \neg$ or $sl(v) = Q^{\ge k}u$, for some $Q \in \{\exists, \forall\}, k \leq t$, and $u \in \{x, y, z\}$. Case 1: $sl(v) = \neg$. In this case $il(v) = \langle B, A \rangle$ and δ_1 is a separator for il(v), therefore $w(\delta) \leq w(\delta_1) \leq w(\delta_1) + t$. Case 2: $sl(v) = Q^{\geq k}u$.

We call δ' the separator for il(v) defined in Lemma VII.7. Since δ is a minimal separator for il(v), $w(\delta) \leq w(\delta')$. We will show that $w(\delta') \leq w(\delta_1) + k$ to finish the proof. Let us choose u', u'' so that $\{x, y, z\} = \{u, u', u''\}$. From the

definition of δ' we get the following:

• $c(\delta') = \delta'(\{u', u''\}) \le c(\delta_1) + k - 1$ • $\delta'(\{\min, \max\}) \le b(\delta_1) + k - 1$

Proposition D.1. For any $f, g \in \{x, y, z\}$, $\delta'(\{\min, f\}) + \delta'(\{g, \max\}) \le b(\delta_1) + 2c(\delta_1) + 2(k-1).$

- - Otherwise we get $f, g \in \{x, y, z\} \setminus \{u\}$ and: - $\delta'(\{\min, f\}) = MAX(\delta_1(\{\min, f\})), \delta_1(\{\min, u\})) + \delta_1(\{u, f\}) + k - 1$ - $\delta'(\{\max, g\}) = MAX(\delta_1(\{\max, g\})), \delta_1(\{\max, u\})) + \delta_1(\{u, g\}) + k - 1$

So:

$$\begin{split} \delta'(\{\min, f\}) + \delta'(\{\max, g\}) &\leq \\ MAX[\delta_1(\{\min, f\}) + \delta_1(\{\max, u\}), \\ \delta_1(\{\min, f\}) + \delta_1(\{\max, u\})) + \delta_1(\{u, g\}) + k - 1, \\ \delta_1(\{\min, u\}) + \delta_1(\{u, f\}) + k - 1 + \delta_1(\{\max, g\}), \\ \delta_1(\{\min, u\}) + \delta_1(\{u, f\}) + \delta_1(\{\max, u\}) + \\ \delta_1(\{u, g\}) + 2(k - 1)] \end{split}$$

Therefore:

.

$$\delta'(\{\min, f\}) + \delta'(\{g, \max\}) \le b(\delta_1) + 2c(\delta_1) + 2(k-1)$$

From Proposition D.1 we can upper bound the border distance:

$$b(\delta') = MAX[\delta'(\{\min, f\}) + \delta'(\{g, \max\}), \delta'(\{\min, \max\})]$$
$$\leq b(\delta_1) + 2c(\delta_1) + 2(k-1)$$

Finally we can conclude:

$$w(\delta')^{2} = c(\delta')^{2} + b(\delta')$$

$$\leq (c(\delta_{1}) + k - 1)^{2} + b(\delta_{1}) + 2c(\delta_{1}) + 2(k - 1)$$

$$= c(\delta_{1})^{2} + b(\delta_{1}) + 2kc(\delta_{1}) + k^{2} - 1$$

$$\leq w(\delta_{1})^{2} + 2kw(\delta_{1}) + k^{2} = (w(\delta_{1}) + k)^{2}.$$

As a final consequence:

$$w(\delta) \le w(\delta') \le w(\delta_1) + k \le w(\delta_1) + t.$$

Appendix E Figures for the proof of Lemma VII.7 Appendix F Figure for gap variables illustration Definition



Fig. 7. Case 3 of the proof of Lemma VII.7. We associate a ball with center $\alpha(u)$ of radius $\delta_1(u, v)$ for every $u \in \{x, y, min, max\}$ on \mathcal{A} (resp. $\beta(u)$ on \mathcal{B}). Each ball associated to variable u accounts for the ability of the pair $\{u, z\}$ to separate. The crossed dot on \mathcal{A} represents the choices on the linear order from F^k , the crossed dot on \mathcal{B} represent the best attempt at a response for indistinguishability. In Case 3, a choice on \mathcal{A} is in two or more balls, which cannot be matched by an element of \mathcal{B} .



Fig. 8. Case 4 of the proof of Lemma VII.7.

We associate a ball with center $\alpha(u)$ of radius $\delta_1(u, v)$ for every $u \in \{x, y, min, max\}$ on \mathcal{A} (resp. $\beta(u)$ on \mathcal{B}). Each ball associated to variable u accounts for the ability of the pair $\{u, z\}$ to separate. The crossed dot on \mathcal{A} represents the choices on the linear order from F^k , the crossed dot on \mathcal{B} represent the best attempt at a response for indistinguishability. In **Case 4**, a choice on \mathcal{A} is in one ball and not in another, which cannot be matched by an element of \mathcal{B} .



Fig. 9. Case 5 of the proof of Lemma VII.7.

We associate a ball with center $\alpha(u)$ of radius $\delta_1(u, v)$ for every $u \in \{x, y, min, max\}$ on \mathcal{A} (resp. $\beta(u)$ on \mathcal{B}). Each ball associated to variable u accounts for the ability of the pair $\{u, z\}$ to separate. The regions of the linear orders that are not covered by any of the balls are the gaps, in purple on the Figure (in the example, the $\{min, x\}$ gap is empty. The crossed dots on \mathcal{A} represents the choices on the linear order from F^k , the crossed dots on \mathcal{B} represent the best attempts at a response for indistinguishability. In **Case 5**, p choices on \mathcal{A} are in a gap, which cannot be matched by p elements of \mathcal{B} from the corresponding gap.



Fig. 10. We associate a ball with center $\alpha(u)$ of radius $\delta_1(u, v)$ for every $u \in \{x, y, min, max\}$ on \mathcal{A} (resp. $\beta(u)$ on \mathcal{B}). Each ball associated to variable u accounts for the ability of the pair $\{u, z\}$ to separate. The regions of the linear orders that are not covered by any of the balls are the gaps, in purple on the Figure.

In the example on \mathcal{A} , the $\{min, x\}$ gap is empty, the gap variables for the $\{x, y\}$ gap are (x, y) and the gap variables for the $\{y, max\}$ gap are (y, max). On \mathcal{B} , the $\{min, x\}$ gap is empty. The gap variables for the $\{x, y\}$ gap on \mathcal{B} are (min, y), as the $\delta_1(\{min, z\})$ ball covers more of the gap on the left than the $\delta_1(\{min, z\})$ ball. Finally on \mathcal{B} the gap variables for the $\{y, max\}$ gap are (y, max).