Quantifier rank for parity of embedded finite models

Hervé Fournier

▶ To cite this version:

Hervé Fournier. Quantifier rank for parity of embedded finite models. [Research Report] LIP RR-2001-09, Laboratoire de l'informatique du parallélisme. 2001, 2+15p. hal-02101892

HAL Id: hal-02101892

https://hal-lara.archives-ouvertes.fr/hal-02101892

Submitted on 17 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Laboratoire de l'Informatique du Parallélisme

École Normale Supérieure de Lyon
Unité Mixte de Recherche CNRS-INRIA-ENS LYON
n° 8512

Quantifier rank for parity of embedded finite models

Hervé Fournier

Février 2001

Research Report N° RR2001-09

École Normale Supérieure de Lyon
46 Allée d'Italie, 69364 Lyon Cedex 07, France
Téléphone : +33(0)4.72.72.80.37
Télécopieur : +33(0)4.72.72.80.80
Adresse électronique : `lip@ens-lyon.fr`

Quantifier rank for parity of embedded finite models

Hervé Fournier

Février 2001

Abstract

We prove some lower bounds for quantifier rank of formulas expressing parity of a finite set $\mathcal{I}$ of bounded cardinal embedded in an algebraically closed field or an ordered $\mathbb{Q}$-vector space. We show that these bounds are tight when elements of $\mathcal{I}$ are known to be linearly independent. In the second part, we prove that strongly minimal structures with quantifier elimination and zero characteristic differentially closed fields admit the active-natural collapse.

Keywords: Constraint databases, Active-natural collapse

Résumé

On prouve des bornes inférieures pour le rang de quantification de formules exprimant la parité d'un ensemble fini $\mathcal{I}$ de cardinal borné, plongé dans un corps algébriquement clos ou un $\mathbb{Q}$-espace vectoriel ordonné. De plus, ces bornes se trouvent être précises dans le cas où on impose aux éléments de $\mathcal{I}$ d'être linéairement indépendents. Dans la seconde partie, on montre que les structures fortememnt minimales éliminant les quantificateurs et les corps différentiellement clos de caractéristique nulle admettent le collapse actif-naturel.

Mots-clés: Bases de données avec contraintes, Collapse actif-naturel

Quantifier rank for parity of embedded finite models

Hervé Fournier

20th February 2001

Abstract

We prove some lower bounds for quantifier rank of formulas expressing parity of a finite set $\mathcal{I}$ of bounded cardinal embedded in an algebraically closed field or an ordered $\mathbb{Q}$-vector space. We show that these bounds are tight when elements of $\mathcal{I}$ are known to be linearly independent. In the second part, we prove that strongly minimal structures with quantifier elimination and zero characteristic differentially closed fields admit the active-natural collapse.

1 Introduction

There are numerous works about the expressiveness obtained by embedding a finite structure into an infinite one M. These studies have been carried out because of their fundamental role in the constraint database model. Among these results, the generic collapse results are of great importance. They state that embedding a finite model into some infinite structures does not help to express a large class of queries, called generic. These results hold for structures M having some good model-theoretic properties : the stronger result deals with structures without the independance property [1]. One of these generic queries is parity, which asks if the cardinal of a finite set $\mathcal{I}$ is even. As a special case of some general collapse theorems [7, 1, 4, 3], we obtain, for some structures M, that there is no first-order sentence defining parity. However, when restricting to the case where $|\mathcal{I}|$ is smaller than a given bound, such a formula exists and one can wonder which is the minimal quantifier rank possible. Can we do better than in the case where the finite set stands alone? In section 2, we establish some lower bounds on this quantifier rank when M is an algebraically closed field or an ordered $\mathbb{Q}$-vector space – for example $(\mathbb{R}, 0, +, -, <)$. Moreover these bounds happen to be tight when we restrict ourselves to the case where the elements of $\mathcal{I}$ are known to be linearly independent in M : indeed we build formulas whose quantifier ranks match the lower bounds. In section 3, we use these results to give a decision algorithm for a theory defined in [4] : this theory is the one of algebraically closed fields with an infinite discrete linear order with extremities, composed of algebraically independent elements. The last section is motivated by some works of Basu, where the notion of uniform quantifier elimination is defined, and where it is proved to hold over real closed fields. There the question of knowing if it holds over algebraically closed fields, and differentially closed fields, is raised [2]. We can give a positive anwer to these questions. We prove in fact a stronger result : we show that strongly minimal structures with quantifier elimination and zero characteristic differentially closed fields have the active-natural collapse – such a collapse is known to hold over o-minimal structures with quantifier elimination [3]. This means that any query that can be expressed with quantifiers running over the whole universe can be expressed with quantifiers which run over the active domain only (the coordinates of the points of the database).

2 Quantifier rank for parity

Here we are interested in the following problem. We embed a finite set $\mathcal{I}$ in either an algebraically closed field or an ordered $\mathbb{Q}$-vector space : we shall call M this structure, and $\mathcal{L}$ its signature. Thus we add a new predicate I that is interpreted as $\mathcal{I}$ to obtain the language $\mathcal{L}^*$. We note $\mathrm{QR}_M(\mathrm{Even}, n)$ the smallest possible quantifier rank of a first-order formula expressing that $|\mathcal{I}|$ is even, when it is known that $|\mathcal{I}| \leqslant n$. Our aim is to find some bounds on this number. We recall that the quantifier rank $\mathrm{qr}(\phi)$ of a formula ϕ is defined by induction on its structure. If ϕ is an atomic formula, $\mathrm{qr}(\phi) = 0$. Otherwise $\mathrm{qr}(\phi \vee \psi) = \mathrm{qr}(\phi \wedge \psi) = \max(\mathrm{qr}(\phi), \mathrm{qr}(\psi))$ and $\mathrm{qr}(\exists x \phi) = \mathrm{qr}(\forall x \phi) = 1 + \mathrm{qr}(\phi)$.

Our main tool will be back-and-forth games. We recall that if there exists a strategy to play a back-and-forth game of length n between two structures M and M', then the same formulas of quantifier rank at most n hold in M and M'. We shall note $\mathrm{qr}(\phi)$ the quantifier rank of a formula ϕ. First let us examine some bounds on the quantifier rank when the finite structure stands alone. When no order is available, we shall write $\mathrm{QR}_=(\mathrm{Even}, n)$ the minimal quantifier rank of a first-order formula expressing parity. In the same way, we write it $\mathrm{QR}_<(\mathrm{Even}, n)$ when the universe is ordered. By some usual back-and-forth games [5], we have $\mathrm{QR}_=(\mathrm{Even}, n) = n$ and $\mathrm{QR}_<(\mathrm{Even}, n) = \log n + O(1)$. These bounds are still the best ones we know in the embedded setting, so we are going to focus on lower bounds. However, we present better upper bounds in some restricted cases.

Now we need a simple remark. If two structures M and M' are elementary equivalent, then $\forall n_0,\ \mathrm{QR}_M(\mathrm{Even}, n_0) = \mathrm{QR}_{M'}(\mathrm{Even}, n_0)$. Indeed, let n_0 be fixed and suppose we have a fisrt order formula ϕ such that if $|\mathcal{I}| \leqslant n_0$, $(M, \mathcal{I}) \models \phi$ iff the property Even holds. Let $\tilde{\phi}(x_1, \ldots, x_n)$ be the formula ϕ where $I(x)$ is replaced with $\bigvee_i x = x_i$. Let $\psi_n = \forall x_1, \ldots, x_n \ \bigwedge_{i<j} x_i \neq x_j \to \tilde{\phi}(x_1, \ldots, x_n)$. Let $|\mathcal{I}| \leqslant n_0$. If Even holds then $(M, \mathcal{I}) \models \phi$. Hence $M \models \psi_n$, then $M' \models \psi_n$ and we obtain $(M', \mathcal{I}) \models \phi$. If Even does not hold, $(M, \mathcal{I}) \models \neg\phi$, and we obtain $(M', \mathcal{I}) \models \neg\phi$ as above. This shows that ϕ expresses Even for $|\mathcal{I}| \leqslant n_0$, so $\mathrm{QR}'_M(\mathrm{Even}, n_0) \leqslant \mathrm{QR}_M(\mathrm{Even}, n_0)$ and by symmetry $\mathrm{QR}'_M(\mathrm{Even}, n_0) = \mathrm{QR}_M(\mathrm{Even}, n_0)$. This justifies the notation $\mathrm{QR}_T(\mathrm{Even}, n)$ for a complete theory T.

2.1 In an algebraically closed field

We note ACF_p the theory of algebraically closed fields of characteristic p (p prime or $p = 0$). We note $\mathbb{F}_p$ the field with p elements and $\mathbb{F}_0 = \mathbb{Q}$. We shall note $\overline{A}$ the algebraic closure of the field generated by $A \subseteq K$, where K is an algebraically closed field. In this section, we shall prove the following lower bound.

Theorem 1 $\mathrm{QR}_{ACF_p}(\mathrm{Even}, n) \geqslant \lceil \log n \rceil + 1$.

Thanks to the remark from the introduction, it is enough to show this lower bound in a given algebraically closed field of characteristic p. Let K be an algebraically closed field of characteristic p of transcendance degree $2^{n-1}+n+1$. Let $\mathcal{M}$ be a set of 2^{n-1} algebraically independent elements of K; the same for $\mathcal{N}$, with $|\mathcal{N}| = 2^{n-1}+1$. We are going to prove that it is possible to make a back-and-forth game of length n between $(K, \mathcal{M})$ and $(K, \mathcal{N})$.

When it remains j steps to do, we note E_j the space where φ is defined and $F_j = \varphi(E_j)$. Let $E_n = F_n = \overline{\mathbb{F}_p}$ and $\varphi = \mathrm{Id}_{\overline{\mathbb{F}_p}}$. At each step, φ_j is an isomorphism of algebraically closed field "with points" from E_j onto F_j. We shall also maintain the following property $\mathcal{P}_j$.

First $|\mathcal{M} \setminus E_j|, |\mathcal{N} \setminus E_j| \geqslant 2^{j-1}$. Moreover, if there exists $a \in \mathcal{M} \setminus E_j$ such that $a \in \overline{E_j \cup A}$ and $A \subset \mathcal{M} \setminus (E_j \cup \{a\})$, then $|A| \geqslant 2^{j-1}$. And the corresponding property in $(K, \mathcal{N})$.

First let us check that $\mathcal{P}_n$ is verified. We have $|\mathcal{M} \setminus E_n| = |\mathcal{M}| \geqslant 2^{n-1}$. Moreover, there is no $a \in \mathcal{M}$ with $a \in \overline{\mathbb{F}_p \cup A}$ such that $A \subseteq \mathcal{M}$ and $a \notin A$ because elements of $\mathcal{M}$ are algebraically independent over $\mathbb{F}_p$. And the same in $(K, \mathcal{N})$.

Let us suppose that $n-j-1$ steps have been done. The isomorphism φ is defined on E_{j+1} and it remains $j+1$ steps to do. Property $\mathcal{P}_{j+1}$ is verified by induction hypothesis. By symetry, we can assume that the point is chosen in $(K, \mathcal{M})$. Let us note v this point. We can also assume $v \notin E_{j+1}$. There are two cases. First case : $v \in \overline{E_{j+1} \cup \{a_1, \ldots, a_r\}}$ with $a_i \in \mathcal{M} \setminus E_{j+1}$ distinct and $r \leqslant 2^{j-1}$. Then we choose some distinct elements $b_1, \ldots, b_r$ in $\mathcal{N} \setminus F_{j+1}$ and we define $\varphi(a_i) = b_i$. Thus $E_j = \overline{E_{j+1} \cup \{a_1, \ldots, a_r\}}$. Let us show that $\mathcal{P}_j$ is verified. If there exists $d \in \mathcal{M} \setminus E_j$, with $d \in \overline{E_j \cup \{c_1, \ldots, c_l\}}$, $c_i \in \mathcal{M} \setminus (E_j \cup \{d\})$ and $l \leqslant 2^{j-1}-1$, then $d \in \overline{E_{j+1} \cup \{a_1, \ldots, a_r, c_1, \ldots, c_l\}}$. But $r + l \leqslant 2^{j-1} + 2^{j-1} - 1 = 2^j - 1$. Therefore we should have $d \in E_{j+1}$ by property $\mathcal{P}_{j+1}$, this is absurd. We have the same property in $(K, \mathcal{N})$. Moreover, $|\mathcal{M} \setminus E_j|, |\mathcal{N} \setminus E_j| \geqslant 2^j - 2^{j-1} = 2^{j-1}$ so $\mathcal{P}_j$ is verified. Exactly in the same way, we show that there are no other points from $\mathcal{M} \setminus E_{j+1}$ in E_j besides the a_i : if $d \in (\mathcal{M} \cap E_j) \setminus (E_{j+1} \cup \{a_1, \ldots, a_r\})$, then $d \in \overline{E_{j+1} \cup \{a_1, \ldots, a_r\}}$ and we conclude with $\mathcal{P}_{j+1}$. This also holds in $(K, \mathcal{N})$, and it shows that φ is an isomorphism. Second case : let $f \notin \overline{E_{j+1} \cup \mathcal{N}}$. Such a point exists because the transcendance degree of K is big enough. Let $\varphi(v) = f$. We set $E_j = \overline{E_{j+1} \cup \{v\}}$. Let us show that $\mathcal{P}_j$ is verified. Let $a \in \mathcal{M} \setminus E_j$ such that $a \in \overline{E_j \cup A}$ for $A \subseteq \mathcal{M} \setminus (E_j \cup \{a\})$ with $|A| < 2^{j-1}$. Thus $a \in \overline{E_{j+1} \cup \{v\} \cup A}$. This shows $v \in \overline{E_{j+1} \cup \{a\} \cup A}$, because $a \in \overline{E_{j+1} \cup A}$ is impossible by $\mathcal{P}_{j+1}$. But we should be in the first case since $|A \cup \{a\}| \leqslant 2^{j-1}$. This also holds in $(K, \mathcal{N})$ by the choice of f. Moreover, there is no point of $\mathcal{M}$ in $E_j \setminus E_{j+1}$ because if $a \in \mathcal{M} \cap E_j \setminus E_{j+1}$, then $a \in \overline{E_{j+1} \cup \{v\}}$ and as $a \notin E_{j+1}$ we would have $v \in \overline{E_{j+1} \cup \{a\}}$ which is absurd. This also holds in $(L, \mathcal{N})$ thanks to the choice of f, thus φ remains an isomorphism. Moreover, $|\mathcal{M} \setminus E_j| = |\mathcal{M} \setminus E_{j+1}| \geqslant 2^{j-1}$

which ends to show $\mathcal{P}_j$. This ends the back-and-forth game. Thus we have shown $\mathrm{QR}_{ACF_p}(\mathrm{Even}, 2^{n-1}+1) > n$. As $\mathrm{QR}_{ACF_p}(\mathrm{Even}, \cdot)$ is an increasing function, we obtain $\mathrm{QR}_{ACF_p}(\mathrm{Even}, n) \geqslant \lceil \log n \rceil + 1$. $\square$

We note $\mathcal{V}_p$ the theory of $\mathbb{F}_p$-vector spaces. We are now interested in a special case, the one where the elements of $\mathcal{I}$ are known to be linearly independent over $\mathbb{F}_p$. We shall use the following notations in this case : $\mathrm{QR}'_{\mathcal{V}_p}(\mathrm{Even}, n)$ and $\mathrm{QR}'_{ACF_p}(\mathrm{Even}, n)$.

Proposition 1 $\mathrm{QR}'_{\mathcal{V}_p}(\mathrm{Even}, n) \leqslant \lceil \log n \rceil + 1$.

We shall prove that it is possible to express that $|\mathcal{I}| \geqslant m$ with a formula of quantifier rank $\lceil \log m \rceil + 1$. First we remark that $|\mathcal{I}| \geqslant m$ iff

$$\exists x\ \exists y_1, \ldots, y_m \in \mathcal{I}\ x = \sum_i y_i \wedge \neg \exists z_1, \ldots, z_{m-1} \in \mathcal{I}\ x = 2z_1 + z_2 + \ldots + z_{m-1}.$$

Now we shall design some formulas $F_m(x)$ such that $F_m(x)$ is true iff $\exists y_1, \ldots, y_m \in \mathcal{I}\ x = \sum_i y_i$. We define $F_1(x) = I(x)$ and for $k \geqslant 2$ $F_k(x) = \exists y F_{\lfloor k/2 \rfloor}(y) \wedge F_{\lceil k/2 \rceil}(x-y)$. One can check that $\mathrm{qr}(F_m(x)) = \lceil \log m \rceil$. In the same way, we define $F'_m(x)$, meaning that $\exists z_1, \ldots, z_{m-1} \in \mathcal{I}\ x = 2z_1 + z_2 + \ldots + z_{m-1}$. For any $m \geqslant 2$, the formula $F'_m(x)$ is obtained by replacing in the previous construction of $F_m(x)$ one $F_2(x)$ by $F'_2(x) = \exists y\ F_1(y) \wedge x = y + y$. Clearly $\mathrm{qr}(F'_m(x)) = \mathrm{qr}(F_m(x))$. The formula $G_m = \exists x\ F_m(x) \wedge \neg F'_m(x)$ expresses that $|\mathcal{I}| \geqslant m$. Of course $G_m \wedge \neg G_{m+1}$ expresses that $|\mathcal{I}| = m$. Now if we know that $|\mathcal{I}| \leqslant n$, $|\mathcal{I}|$ is even if and only if $\bigvee_{2k \leqslant n} |\mathcal{I}| = 2k$. Remark that $|\mathcal{I}| \geqslant n$ is equivalent to $|\mathcal{I}| = n$ since we know that $|\mathcal{I}| \leqslant n$. Thus our formula will be $\bigvee_{2k \leqslant n} G_{2k} \wedge \neg G_{2k+1}$ when n is odd, and $\bigvee_{2k<n}(G_{2k} \wedge \neg G_{2k+1}) \vee G_n$ when n is even. That allows to obtain the desired bound. $\square$

Corollary 1 $\mathrm{QR}'_{ACF_p}(\mathrm{Even}, n) = \mathrm{QR}'_{\mathcal{V}_p}(\mathrm{Even}, n) = \lceil \log n \rceil + 1$.

It is just necessary to remark that the previous lower bound holds in the case where the elements of $\mathcal{I}$ are known to be linearly independent. $\square$

Corollary 2 *For ϕ a first-order formula of* ACF_p *with an extra unary predicate I, let $a(\phi)$ be the minimum quantifier rank of an equivalent active semantics formula. Let $\alpha(r) = \max\{a(\phi),\ \mathrm{qr}(\phi) = r\}$. We have $\alpha(p) \geqslant 2^p$.*

Consider the formula ϕ_n expressing that $\mathcal{I}$ has at least n elements, assuming they are linearly independant – see proposition 1. Let ϕ_n^a be an equivalent active semantics formula. When restricting to the case where $\mathcal{I} \subset D$ with $D = \{d_i,\ i \in \mathbb{N}\}$ a set of indiscernibles, we obtain $\tilde{\phi}_n^a$ a pure equality formula expressing that $\mathcal{I}$ has at least n elements. As $\mathrm{qr}(\phi_n^a) = \mathrm{qr}(\tilde{\phi}_n^a) \geqslant n$ and $\mathrm{qr}(\phi_n) = \lceil \log n \rceil$, taking $n = 2^p$ gives the result. $\square$

2.2 In an ordered $\mathbb{Q}$-vector space

We first show a lower bound. We note $\mathcal{O}vs$ the theory of $\mathbb{Q}$-ordered vector spaces. We define N_p the following way:

$$\begin{cases} N_0 &= 1 \\ N_{p+1} &= (2^p+1)N_p \end{cases}$$

We define an algebraic measure d_∞ as follows. For $x \leqslant y$, we define $d_\infty(x,y) = |\{z,\ x < z \leqslant y\}|$. Then, for $j \in \mathbb{N}$, we define $d_j(x,y) = d_\infty(x,y)$ if $d_\infty(x,y) < N_j$, $d_j(x,y) = \infty$ otherwise. At last, we take $d_\infty(y,x) = -d_\infty(x,y)$ and $d_j(y,x) = -d_j(x,y)$. First we need a simple remark.

Lemma 1 *We consider a back-and-forth game between two finite ordered sets A and B where it is possible to choose 2^j elements (on the same side) when it remains j moves to play. If $|A|, |B| \geqslant N_{n+1}$, then it is possible to play a game of length n.*

Let $|A|, |B| \geqslant N_{n+1}$. We show how to play a game of length n. Before the game begins, we define our partial isomorphism α to send the extremities of A onto the extremities of B. We can assume that it remains j moves to play. Let us call $D \subseteq A$ the set where α is defined. By induction hypothesis, we assume that $d_{j+1}(a,a') = d_{j+1}(\alpha(a),\alpha(a'))$ for $a, a' \in D$. We proceed as in the case of back-and-forth games between two finite linear orderings – see [5] – except that we can take 2^j elements. We shall handle all at once the elements $a_1 < a_2 < \ldots < a_k$ lying in an interval $]c,d[$ with $c,d \in D$, $]c,d[\cap D = \emptyset$. First case : $d_{j+1}(c,d) < \infty$. By induction, $d_{j+1}(c,d) = d_{j+1}(\alpha(c),\alpha(d))$ and we chose the $\alpha(a_i)$ in the obvious way. Second case : $d_{j+1}(c,d) = \infty$. Let $a_0 = c$ and $a_{k+1} = d$. We successively choose $\alpha(a_l)$ for $l = 1, 2, \ldots, s$ such that $d_j(a_l, a_{l+1}) = d_j(\alpha(a_l), \alpha(a_{l+1}))$, where s is the smaller subscript such that $d_j(a_s, a_{s+1}) = \infty$. We proceed in the same way for $l = k, k-1, \ldots, t$ where t is the larger subscript such that $d_j(a_{t-1}, a_t) = \infty$. If the images of all the a_i for $1 \leqslant i \leqslant k$ have not yet been determined, then we successively choose the images of a_l for $l = s+1, \ldots, t-1$: we choose $\alpha(a_l)$ such that $d_\infty(\alpha(a_{l-1}), \alpha(a_l)) = \min\{N_j, d_j(a_{l-1}, a_l)\}$. Let us show we have enough points from B in $]\alpha(a_0), \alpha(a_{k+1})[$. As $d_{j+1}(\alpha(a_0), \alpha(a_{k+1})) = \infty$ by induction, we have $d_\infty(\alpha(a_0), \alpha(a_{k+1})) \geqslant N_{j+1}$. Taking into account that $k \leqslant 2^j$ and $N_{j+1} = (2^j + 1)N_j$, there are indeed enough points to proceed this way. □

We now establish the following lower bound.

Theorem 2 $\mathrm{QR}_{\mathcal{O}vs}(\mathrm{Even}, n) = \Omega(\sqrt{\log n})$.

In order to prove this, we shall make a back-and-forth game of length n between two ordered $\mathbb{Q}$-vector spaces with points $(V, \mathcal{M})$ and $(W, \mathcal{N})$. We shall chose $\mathcal{M}$ (resp. $\mathcal{N}$) such that it is a basis of V (resp. W). We note $a \ll b$ or $a = o(b)$ if $\forall n \in \mathbb{N},\ n|a| \leqslant |b|$. By decomposing a point $v \in V$ in the basis $\mathcal{M}$, it can be written $v = \sum_{i=1}^{r} \alpha_i a_i$ with $\alpha_i \in \mathbb{Q}^*$, $a_i \in \mathcal{M}$ and $a_1 \gg \ldots \gg a_r$. We use the following notations : $\mathrm{supp}(v) = \{a_1, \ldots, a_r\}$, $\mathrm{supp}(v,l) = \{a_i,\ i \leqslant \min(l,r)\}$, $z(v,j) = a_{\min(2^j,r)}$ and $\mathrm{T}_j(v) = \sum_{i=1}^{\min(2^j,r)} \alpha_i a_i$. Thus $z(v,j) = z(\mathrm{T}_j(v), j)$. Let us remark that if $|\mathrm{supp}(\mathrm{T}_j(x))| < 2^j$ then $x = \mathrm{T}_j(x)$. We note π the canonical projection from $\mathcal{M} \times \mathcal{N}$ onto $\mathcal{M}$. Given $R \subseteq \mathcal{M} \times \mathcal{N}$ a one to one function from a part of $\mathcal{M}$ in $\mathcal{N}$, we note $\mathcal{L}_R$ the linear application defined on $\mathrm{Vect}(\pi(R))$ and extending R. We shall also use the corresponding notations in $(W, \mathcal{N})$.

We it remains j steps to do, we shall have an isomorphism φ_{j+1} defined from E_{j+1} onto F_{j+1}. We shall also need the distances d_∞ and d_j defined previously, but relativized to the set $\mathcal{I}$.

Let V be the ordered $\mathbb{Q}$-vector space spanned by $\mathcal{M} = \{\varepsilon_1, \ldots, \varepsilon_{n_v}\}$ with $0 < \varepsilon_1 \ll \ldots \ll \varepsilon_{n_v}$ and $n_v = N_{n+1}$ – this is the same as considering $\mathbb{Q}^{n_v}$ with the lexicographic order. In the same way, $W = \mathrm{Vect}(\mathcal{N})$ with $\mathcal{N} = \{\eta_1, \dots, \eta_{n_w}\}$ such that $0 < \eta_1 \ll \dots \ll \eta_{n_w}$ and $n_w = N_{n+1} + 1$. We set $E_{n+1} = \mathrm{Vect}(\{\varepsilon_1, \varepsilon_{n_v}\})$, $F_{n+1} = \mathrm{Vect}(\{\eta_1, \eta_{n_w}\})$ and φ_{n+1} is defined by $\varphi(\varepsilon_1) = \eta_1$ et $\varphi(\varepsilon_{n_v}) = \eta_{n_w}$. We also set $R_{n+1} = \{(\varepsilon_1, \eta_1), (\varepsilon_{n_v}, \eta_{n_w})\} \subseteq \mathcal{M} \times \mathcal{N}$. At each step we shall maintain the following property $\mathcal{P}_j$.

a) For all $x, y \in \pi(R_j)$, $d_j(x, y) = d_j(R_j(x), R_j(y))$.
b) For all $v \in E_j$, we have $\mathrm{T}_j(\varphi_j(v)) = \mathcal{L}_{R_j}(\mathrm{T}_j(v))$. Similarily, for all $w \in F_j$, $\mathrm{T}_j(\varphi_j^{-1}(w)) = \mathcal{L}_{R_j^{-1}}(\mathrm{T}_j(w))$.
c) The application φ_j is an isomorphism of ordered $\mathbb{Q}$-vector space with points from E_j onto F_j.

Let us remark that point b) means that $\mathcal{L}_{R_j}(\mathrm{T}_j(v))$ makes sense, so it implies $\mathrm{supp}(\mathrm{T}_j(v)) \subseteq \pi(R_j)$. Let us also remark that, as a consequence of a), R_j is a strictly increasing application from $\pi(R_j) \subseteq \mathcal{M}$ to $\mathcal{N}$. Let us show that $\mathcal{P}_{n+1}$ holds : a) comes from $|\mathcal{M}|, |\mathcal{N}| \geqslant N_{n+1}$, the other points are clear. Let us assume that $n - j$ steps of then back-and-forth game have been carried out. It remains $j \geqslant 1$ steps to do. The isomorphism φ_{j+1} is defined from E_{j+1} onto F_{j+1}. By symmetry, we can assume that point v is chosen in $(V, \mathcal{M})$. Without loss of generality, we assume that $v \notin E_{j+1}$. Let $u \in \mathrm{Vect}(E_{j+1} \cup \{v\}) \setminus E_{j+1}$ *such that $z(u, j)$ is minimal* with respect to the order on V. Let $S = \mathrm{supp}(u, 2^j) \setminus \pi(R_{j+1})$. Thanks to lemma 1, we now define the relation R_j extending R_{j+1} such that $\pi(R_j) = \pi(R_{j+1}) \cup S$.

Let φ_j be the linear application extending φ_{j+1} and such that $\varphi_j(u) = \mathcal{L}_{R_j}(\mathrm{T}_j(u))$. Let $E_j = \mathrm{Vect}(E_{j+1} \cup \{u\})$ et $F_j = \varphi_j(E_j)$. In what follows, φ_j will be noted φ, R_j will be noted R and $\mathcal{L}_{R_j}$ sometimes noted $\mathcal{L}_j$. Let us show we have $\mathcal{P}_j$. Let us first remark that φ is a linear application from E_j onto F_j. Let us show that φ is one to one. Let $w \in E_j$, $\varphi(w) = 0$. We wrote $w = \alpha u + e$ with $e \in E_{j+1}$ and $\alpha \in \mathbb{Q}$. If $\alpha = 0$, then $e = 0$ because φ_{j+1} is one to one. Let us suppose $\alpha \neq 0$. Thus $\varphi(e) = \varphi_{j+1}(e) = -\alpha\varphi(u) = -\alpha\mathcal{L}_j\mathrm{T}_j(u)$. Thanks to $\mathcal{P}_{j+1}$ *b)* for φ_{j+1}^{-1} we obtain $\mathrm{T}_{j+1}(e) = -\alpha\mathcal{L}_{R_{j+1}^{-1}}\mathrm{T}_{j+1}\mathcal{L}_j\mathrm{T}_j u$. Therefore $\mathrm{T}_{j+1}(e) = -\alpha\mathcal{L}_{R_{j+1}^{-1}}\mathcal{L}_j\mathrm{T}_j u = -\alpha\mathrm{T}_j(u)$ because this expression makes sense and R_j extends R_{j+1}. But $2^j < 2^{j+1}$, so $e = -\alpha\mathrm{T}_j(u)$. Now if $u \neq \mathrm{T}_j(u)$, this gives $w = e + \alpha u \notin E_{j+1}$ with $w = o(z(u, j))$ which is impossible by the choice of u. As $u = \mathrm{T}_j(u)$, we have $e = -\alpha u$ that is to say $w = 0$.

Point a) stems from construction. Let us show point b) for φ. Let $v \in E_j$. If $v \in E_{j+1}$, it is clear by $\mathcal{P}_{j+1}$ since $2^{j+1} \geqslant 2^j$, $R_{j+1} \subseteq R_j$ and φ extends φ_{j+1}. Hence we suppose $v \notin E_{j+1}$ Thus $v = \alpha u + e$ where u is the vector chosen above, $\alpha \in \mathbb{Q}^*$ and $e \in E_{j+1}$. The following holds.

$$\mathrm{T}_j(v) = \mathrm{T}_j(\alpha\mathrm{T}_j(u) + \mathrm{T}_{j+1}(e)) \,. \tag{1}$$

Proof:
i) Let us suppose $z(\mathrm{T}_{j+1}(e)) \leqslant z(\mathrm{T}_j(u))$. Thus $\mathrm{T}_{j+1}(e) = e + o(z(\mathrm{T}_j(u)))$ and $v = \alpha u + e = \alpha\mathrm{T}_j(u) + \mathrm{T}_{j+1}(e) + o(z(\mathrm{T}_j(u)))$. As $z(\mathrm{T}_j(v)) \geqslant z(\mathrm{T}_j(u))$ by the choice of u, we obtain relation (1) by truncating the previous equality at the order 2^j.
ii) Now let us suppose $z(\mathrm{T}_{j+1}(e)) \geqslant z(\mathrm{T}_j(u))$.
ii-a) If $\mathrm{T}_{j+1}(e) = e$, in particular we have $\mathrm{T}_{j+1}(e) = e + o(z(\mathrm{T}_j(u)))$ and we finish as previously.

ii-b) Otherwise, $|\mathrm{supp}(e, 2^{j+1})| = 2^{j+1}$. Moreover $v = e + \alpha u = \mathrm{T}_{j+1}(e) + \alpha \mathrm{T}_j(u) + o(z(\mathrm{T}_{j+1}(e)))$. As the sum $\mathrm{T}_{j+1}(e) + \alpha\mathrm{T}_j(u)$ has at least 2^j terms from $\mathrm{T}_{j+1}(e)$, we obtain (1) by truncating the previous equality at the order 2^j.

The following also holds.

$$\mathrm{T}_j(\varphi(v)) = \mathrm{T}_j(\alpha\mathrm{T}_j(\varphi(u)) + \mathrm{T}_{j+1}(\varphi(e))) \; . \tag{2}$$

Proof:

i) Let us suppose $|\mathrm{supp}(e, 2^{j+1})| < 2^{j+1}$. Thus $e = \mathrm{T}_{j+1}(e)$. By $\mathcal{P}_{j+1}$, we obtain $\mathrm{T}_{j+1}(\varphi(e)) = \mathcal{L}_{j+1}(\mathrm{T}_{j+1}(e)) = \mathcal{L}_{j+1}(e)$. But $\mathcal{L}_{j+1}(e)$ has strictly less than 2^{j+1} terms so $\mathrm{T}_{j+1}(\varphi(e)) = \varphi(e)$. Let us recall that $\varphi(u) = \mathrm{T}_j(\varphi(u))$ by the choice of $\varphi(u)$. By subsituting these terms in $\mathrm{T}_j(\varphi(v)) = \mathrm{T}_j(\alpha\varphi(u) + \varphi(e))$ we obtain (2).

ii) Otherwise $|\mathrm{supp}(e, 2^{j+1})| = 2^{j+1}$. Thus $\mathrm{T}_{j+1}(\varphi(e)) = \mathcal{L}_{j+1}(\mathrm{T}_{j+1}(e))$ has 2^{j+1} terms. But $\varphi(v) = \alpha\varphi(u) + \varphi(e) = \alpha\mathrm{T}_j(\varphi(u)) + \mathrm{T}_{j+1}(\varphi(e)) + o(z(\mathrm{T}_{j+1}(\varphi(e))))$. As $\alpha\mathrm{T}_j(\varphi(u)) + \mathrm{T}_{j+1}(\varphi(e))$ has at least 2^j terms from $\mathrm{T}_{j+1}(\varphi(e))$, we obtain (2) by truncating the previous equality at the order 2^j.

Now let us prove P_j *b)* for φ. Let $v \in E_j$. We write $v = \alpha u + e$ with $e \in E_{j+1}$ and $\alpha \in \mathbb{Q}$. By (2), $\mathrm{T}_j(\varphi(v)) = \mathrm{T}_j(\alpha\mathrm{T}_j(\varphi(u)) + \mathrm{T}_{j+1}(\varphi(e)))$. But $\mathrm{T}_j(\varphi(u)) = \varphi(u) = \mathcal{L}_j\mathrm{T}_j(u)$ by the choice of $\varphi(u)$. Moreover, by $\mathcal{P}_{j+1}$, $\mathrm{T}_{j+1}(\varphi(e)) = \mathcal{L}_{j+1}\mathrm{T}_{j+1}(e)$. And $\mathcal{L}_{j+1}\mathrm{T}_{j+1}(e) = \mathcal{L}_j\mathrm{T}_{j+1}(e)$ since $\mathcal{L}_j$ extends $\mathcal{L}_{j+1}$. By the linearity of $\mathcal{L}_j$, this gives $\mathrm{T}_j(\varphi(v)) = \mathrm{T}_j(\mathcal{L}_j(\alpha\mathrm{T}_j(u) + \mathrm{T}_{j+1}(e)))$. Clearly, if $\mathrm{T}_j\mathcal{L}_j(x)$ makes sense for $x \in E_j$, then $\mathcal{L}_j\mathrm{T}_j(x) = \mathrm{T}_j\mathcal{L}_j(x)$. Thus we have $\mathrm{T}_j(\varphi(v)) = \mathcal{L}_j\mathrm{T}_j(\alpha\mathrm{T}_j(u) + \mathrm{T}_{j+1}(e))$. With relation (1) we obtain $\mathrm{T}_j(\varphi(v)) = \mathcal{L}_j(\mathrm{T}_j(v))$.

We now show point b) for φ^{-1}. Let $w \in F_j$ and $v \in E_j$ such that $w = \varphi(v)$. We have $\mathrm{T}_j(w) = \mathrm{T}_j(\varphi(v)) = \mathcal{L}_j(\mathrm{T}_j(v))$ by $\mathcal{P}_j$ *b)* for φ. Moreover, $\mathcal{L}_{R_j}^{-1} = \mathcal{L}_{R_j^{-1}}$; therefore $\mathrm{T}_j(\varphi^{-1}(w)) = \mathcal{L}_{R_j^{-1}}(\mathrm{T}_j(w))$. This proves point b) of $\mathcal{P}_j$ for φ^{-1}.

It remains to prove c). If $a \in E_j \cap \mathcal{M}$, then by $\mathcal{P}_j$ b) we have $\mathrm{T}_j(\varphi(a)) = \mathcal{L}_j\mathrm{T}_j(a) = \mathcal{L}_j(a) = R(a)$. But $|\mathrm{supp}(R(a))| = 1 < 2^j$, so $\varphi(a) = R(a) \in \mathcal{N}$. In the same way, if $x \in E_j$ is positive, then $x = \alpha a + o(a)$ with $a \in \mathcal{M}$ and $\alpha > 0$. By point b) of $\mathcal{P}_j$, we have $\varphi(a) = \alpha R_j(a) + o(R_j(a))$. But $R_j(a) \in \mathcal{N}$; thus $R_j(a) > 0$, which proves $\varphi(x) > 0$. The same works for φ^{-1}, so it ends the proof of point c).

This ends the back-and-forth game. We have shown $\mathrm{QR}_{\mathcal{O}vs}(\mathrm{Even}, N_n + 1) > n$. As $N_p = \prod_{i=0}^{p}(2^i + 1) \leqslant 2^{(p+1)(p+2)/2}$, this leads to $\mathrm{QR}_{\mathcal{O}vs}(\mathrm{Even}, n) = \Omega(\sqrt{\log n})$. $\square$

Does a similar result hold in real-closed fields ? We have a weaker bound in o-minimal structures having quantifier elimination.

Proposition 2 *Let M be an o-minimal structure that admits quantifier elimination. Then* $\mathrm{QR}_M(\mathrm{Even}, n) \geqslant \log\log n + O(1)$.

Let ψ be a formula expressing parity of $|\mathcal{I}|$ for $|\mathcal{I}| \leqslant n$. We consider a structure M', elementary equivalent to M, that contains a sequence of indiscernibles $D = \{d_i,\ i \in \mathbb{N}\}$. The same formula ψ still works in M'. Let ψ' be the formula obtained by replacing $I(t)$ by $\forall z\ (z = t \to I(z))$ in ψ, where z is a new variable. Remark that $\mathrm{qr}(\psi') \leqslant \mathrm{qr}(\psi) + 1$. Here we can apply to ψ' the algorithm of [3] to obtain an active semantics equivalent formula ψ_{act} (as mentionned there, there is no need for ψ to be in prenex form to apply this algorithm). One can check that $\mathrm{qr}(\psi_{act}) \leqslant 2^{qr(\psi')+O(1)}$. Now when we restrict ourselves to the case where $\mathcal{I} \subset D$, the formula ψ_{act} is equivalent to a pure order formula ψ_o with $\mathrm{qr}(\psi_o) = \mathrm{qr}(\psi_{act})$. Applying the bound about the pure ordered case recalled in the introduction, we obtain $\mathrm{qr}(\psi_o) \geqslant \log n + O(1)$. Thus $\mathrm{qr}(\psi) \geqslant \log\log n + O(1)$. □

We can obtain result similar to theorem 2 for a finite graph embedded in an ordered $\mathbb{Q}$-vector space. The query Connected asks if the graph is connected.

Corollary 3 $\mathrm{QR}_{\mathcal{O}vs}(\mathrm{Connected}, n) = \Omega(\sqrt{\log n})$.

We use a usual first-order reduction from parity to connectivity. We consider the graph $G_n = (V, E)$ over $V = \{v_1, \ldots, v_n\}$ where $E(v_i, v_j)$ holds iff $|i-j| = 2$ or $\{i, j\} = \{1, n\}$. For $n \geqslant 2$, G_n is connected iff n is even. As we can express E with a formula of quantifier rank 2 in any ordered structure M, we obtain $\mathrm{QR}_M(\mathrm{Even}, n) \leqslant \mathrm{QR}_M(\mathrm{Connected}, n)+2$. It remains to apply this to the theory $\mathcal{O}vs$. □

Once again we are interested in the special case where the elements of $\mathcal{I}$ are known to be linearly independent. We use the special notation $\mathrm{QR}'_{\mathcal{O}vs}$ in this case. The following holds.

Proposition 3 $\mathrm{QR}'_{\mathcal{O}vs}(\mathrm{Even}, n) = O(\sqrt{\log n})$.

The rough idea is this one. To express that $|\mathcal{I}\cap]a, b[| \geqslant 2^{p^2}$, it is enough to have a set S of 2^{2p} elements of $\mathcal{I}$ such that between two consecutive elements of S, there are at least $2^{(p-1)^2}$ elements of $\mathcal{I}$. And the set S will be represented by the sum of its elements, from which it is possible to extract the elements with a formula of quantifier depth p thanks to the assumption of linear independence. Remark that this does not work anymore if we remove this assumption : there would be no canonical elements to extract from the sum, and intervals considered in the recursion step could overlap.

Now let us explain this more precisely. We shall define several families of formulas. First $S_{2^p}(a, b, x)$ means that x is the sum of 2^p distinct elements of $\mathcal{I}$ which lie in $]a, b[$. We define $S_{2^0}(a, b, x) = I(x) \wedge x \in]a, b[$ and

$$S_{2^p}(a, b, x) = \exists c, y\ S_{2^{p-1}}(a, c, y) \wedge S_{2^{p-1}}(c, b, x-y).$$

We define also $E_{i,2^p}(a, b, x, z)$ for $0 \leqslant i < 2^p$. If $S_{2^p}(a, b, x)$ holds, then $x = \sum_i z_i$ with $z_0 < \ldots < z_{2^p-1}$ and $E_{i,2^p}(a, b, x, z)$ means that $z = z_i$. We define

$$E_{i,2^p}(a, b, x, z) = \exists c, y\ S_{2^{p-1}}(a, c, y) \wedge S_{2^{p-1}}(c, b, x-y) \wedge E'$$

where $E' = E_{i,2^{p-1}}(a, c, y, z)$ if $i < 2^{p-1}$ and $E' = E_{i-2^{p-1},2^{p-1}}(c, b, x-y, z)$ if $i \geqslant 2^{p-1}$. Of course $S_1(a, b, x) = I(x) \wedge a < x < b$. Let us define $m_0 = 1$ and $m_p = 2^p + (2^p+1)m_{p-1}$. For a given n, we take p such that $m_{p-1} < n \leqslant m_p$. Now we can define $F_n(a, b) =$

$$\exists x\ S_{2^p}(a, b, x) \wedge \bigwedge_{i=0}^{2^p-2} \{\exists y_1, y_2\ E_{i,2^p}(a, b, x, y_1) \wedge E_{i+1,2^p}(a, b, x, y_2) \wedge F_{m_{p-1}}(y_1, y_2)\} \wedge F'_n$$

where F'_n handles the first and last intervals :

$$F'_n = \exists y\ E_{0,2^p}(a,b,x,y) \wedge F_{m_{p-1}}(a,y) \wedge \exists y\ E_{2^p-1,2^p}(a,b,x,y_1) \wedge F_{n-(m_{p-1}+1)2^p}(y,b).$$

Thus we can express that $|\mathcal{I}| \geqslant n$ for any $n \leqslant m_p$ with a formula of quantifier rank at most p. The conclusion is the same as in proposition 1. □

Corollary 4 $\mathrm{QR}'_{\mathcal{O}vs}(\mathrm{Even}, n) = \Theta(\sqrt{\log n})$. □

3 Application

We now give a procedure to decide a sentence of the following theory T defined in [4], where it is shown to be complete. Let p be fixed. The theory T is defined over $\mathcal{L} = \mathcal{L}_{rings} \cup \{I, <\}$, where I is a unary predicate, and consists of the following axioms :

- the axioms of ACF_p
- $<$ is defined exactly on the elements of I
- $<$ is a discrete linear order with extremities
- I is infinite
- the elements of I are algebraically independent over $\mathbb{F}_p$.

Here is how the decision algorithm proceeds. Let ϕ be a first-order sentence over $\mathcal{L}$, and r be the quantifier rank of ϕ. We can play a back-and-forth game as explained in the section about algebraically closed fields, except that elements of $\mathcal{I}$ (in the first case of the proof of theorem 1) are chosen as described in lemma 1 since $\mathcal{I}$ is ordered. Therefore, $T \vdash \phi$ is and only if $(K, \mathcal{I}) \models \phi$ where K is an algebraically closed field of characteristic p, $\mathcal{I} \subset K$, $\mathcal{I}$ is finite with $|\mathcal{I}| = N$, where $N = N_r$ – defined in section 2.2, $N_r = O(2^{r^2})$ – and the elements of $\mathcal{I}$ are algebraically independent. Let $\psi(x_1, \ldots, x_N)$ be the formula ϕ where $I(x)$ has been replaced with $\bigvee_{i=1}^{N} x = x_i$ and $x < y$ with $\bigvee_{1 \leqslant i < j \leqslant N} x = x_i \wedge y = x_j$. Now either $\psi(x_1, \ldots, x_N)$ is true whenever $(x_1, \ldots, x_N)$ are algebraically independent (in the case where $T \vdash \phi$), or it is false. Thus it just remains to check if $\dim \psi(K^N) = N$. Here we can eliminate quantifiers and check that set obtained is of full dimension. Another method is this one : $\dim \psi(K^N) = N$ if and only if K^N is covered by $N+1$ translated of $\psi(K^N)$ – see [6]. Thus we just have to decide the following formula of ACF_p :

$$\exists \bar{t}_1, \ldots, \bar{t}_{N+1} \forall \bar{x} \exists \bar{y}\ \psi(\bar{y}) \wedge \bigvee_{i=1}^{N+1} \bar{x} = \bar{y} + \bar{t}_i\ .$$

4 Active-natural collapse results

We consider a relational database of signature $\mathrm{SC} = \{R_1, R_2, \ldots, R_f\}$, with R_i of arity r_i, embedded in an infinite L-structure M. Moreover, we deal here with the finite case : each R_i interprets a finite set of M^{r_i}. Let us note D the active domain, that is to say the set of the coordinates of all points in the database.

Here D will be a finite subset of M. We recall that in an active semantics formula, quantifiers are of the type $\exists x \in D$ ($\forall x \in D$), that we shall write $\exists^a x$ ($\forall^a x$). We shall prove several active-natural collapses, based on the algorithm of Benedikt et Libkin for o-minimal structures with quantifier elimination [3]. This part is motivated by a question asked by Basu [2] – see section 4.3.

To prove active-natural collapse, all we need is to suppress a natural existential quantifier in front of an active formula with parameters. Let us consider the formula $\varphi(\bar{x}) := \exists z\ \alpha(\bar{x}, z)$ where $\alpha(\bar{x}, z)$ is an active formula. We can assume without loss of generality that $\alpha(\bar{x}, z)$ is under prenex form. Thus

$$\alpha(\bar{x}, z) = Q_1^a y_1 \ldots Q_m^a y_m \beta(\bar{x}, \bar{y}, z)$$

where $Q_i \in \{\exists, \forall\}$. Moreover we can also assume that

- every atomic subformula of $\alpha(\bar{x}, z)$ is either from M (more precisely L) or from SC,
- $m > 0$ and $\alpha(\bar{x}, z)$ has at least one atomic subformula from M,
- z does not appear in any subformula of SC.

Thus any atomic subformula of $\alpha(\bar{x}, z)$ from M is of the form $\tau(\bar{x}, \bar{y}, z)$ with $\bar{y} = (y_1, \ldots, y_m)$.

Definition 1 *The active domain $D \subset M$, the parameters $\bar{x}$ and a set $\mathcal{T} = \{\tau_1(\bar{x}, \bar{y}, z), \ldots, \tau_k(\bar{x}, \bar{y}, z)\}$ of atomic formulas being fixed, we call sign vector an application from $\mathcal{T} \times D^m$ to $\{\text{true}, \text{false}\}$. We call sign vector of a point $u \in M$ the application $(\tau, \bar{a}) \mapsto \tau(\bar{x}, \bar{a}, u)$.*

Proposition 4 (Sufficent condition for active-natural collapse) *Let M be a structure with the following property. Given any finite family $\mathcal{T} = \{\tau_1(\bar{x}, \bar{y}, z), \ldots\}$ of atomic formulas of M, there exists $\mathcal{B} \in \mathbb{N}$, a finite set Γ, a family of active formulas $\mathcal{S}_\gamma[\tau]$ of M for $(\gamma, \tau) \in \Gamma \times \mathcal{T}$ and a set of active formulas $\mathcal{F}_\gamma$ of M for $\gamma \in \Gamma$ such that for all $\bar{x} \in M^n$, for any finite active domain $D \subset M$ and for any sign vector $\vec{v}$, there exists $z \in M$ with sign vector $\vec{v}$ iff there exists $(\gamma, \bar{t}) \in \Gamma \times D^{\mathcal{B}}$ such that*
i) $\vec{v} = (\tau, \bar{a}) \mapsto \mathcal{S}_\gamma[\tau](\bar{x}, \bar{a}, \bar{t})$
ii) (M,D) $\models \mathcal{F}_\gamma(\bar{x}, \bar{t})$.
Then M has the active-natural collapse.

We take the notations from the beginning of the section. Let $\mathcal{T}$ be the set of atomic formulas of M appearing in $\alpha(\bar{x}, z)$. Let $\alpha_\gamma(\bar{x}, \bar{t})$ be the formula $\alpha(\bar{x}, z)$ where $\tau(\bar{x}, \bar{y}, z)$ is replaced with $\mathcal{S}_\gamma[\tau](\bar{x}, \bar{y}, \bar{t})$. Let us set

$$\varphi_{act}(\bar{x}) := \exists^a \bar{t} \in D^{\mathcal{B}} \bigvee_{\gamma \in \Gamma} \mathcal{F}_\gamma(\bar{x}, \bar{t}) \wedge \alpha_\gamma(\bar{x}, \bar{t}).$$

It is immediate that this formula is equivalent to $\varphi(\bar{x})$ when the database is not empty. We can handle this case too, just as in [3]. Let $\varphi_{nat}(\bar{x})$ be a natural formula equivalent to $\varphi(\bar{x})$. Let $\varphi_\emptyset(\bar{x})$ be the formula $\varphi_{nat}(\bar{x})$ where each subformula of the type $R(\ldots)$ for R in the database has been replaced with *false*. Let $\varphi'_\emptyset(\bar{x})$ be a quantifier free formula equivalent to $\varphi_\emptyset(\bar{x})$. For the active formula we take $(\exists^a x\ x = x \wedge \varphi_{act}(\bar{x})) \vee (\neg\exists^a x\ x = x \wedge \varphi'_\emptyset(\bar{x}))$. $\square$

4.1 On strongly minimal structures

For a formula $\phi(\bar{a},x)$, we note $\phi(\bar{a},M)$ the set $\{x \in M,\ M \models \phi(\bar{a},x)\}$. We recall that a structure M is said to be strongly minimal if for every formula $\phi(\bar{y},x)$, there exists $d \in \mathbb{N}$ such that for all $\bar{a} \in M$, one of the sets $\phi(\bar{a},M)$ and $\neg\phi(\bar{a},M)$ contains at most d points.

Proposition 5 *Strongly minimal structures with quantifier elimination have the active-natural collapse.*

Let M be a strongly minimal structure with quantifier elimination. We shall use proposition 4. Let Θ be a finite set of atomic formulas of M of the form $\theta(\bar{x},\bar{y},z)$. Let Ψ be the set made of the elements of Θ and their negations. Because M is strongly minimal and Ψ is finite, there is an integer d such that for all $\psi \in \Psi$ and all $\bar{a},\bar{b}$ we have

$$\min\{|\psi(\bar{a},\bar{b},M)|, |M \setminus \psi(\bar{a},\bar{b},M)|\} \leqslant d.$$

In the following, $\bar{x}$ is fixed. Let $z \in M$. For each $(\psi,\bar{y}) \in \Psi \times D^m$, the formula $\psi(\bar{x},\bar{y},\cdot)$ partitions M into a finite set of at most d points and its complement. Here there are two cases : either z is in the finite set for one element of $\Psi \times D^m$, or it always lies in the infinite set. We shall build some sign vector patterns for both cases.

First case : there exists $(\psi_0,\bar{y}_0) \in \Psi \times D^m$ such that $M \models \psi_0(\bar{x},\bar{y}_0,M)$ is finite and $\psi_0(\bar{x},\bar{y}_0,z)$. In this case the sign vector of z is completely determined by a formula of the type

$$\mathrm{loc}_{\bar{\psi}}(\bar{x},\bar{t},u) := \bigwedge_{i=1}^{d} \psi_i(\bar{x},\bar{t}_i,u)$$

for some $(\psi_i,\bar{t}_i) \in \Psi \times D^m$. Indeed the set of points of M satisfying $\psi_0(\bar{x},\bar{y}_0,\cdot)$ is finite, non empty, and contains at most d points. We set $E(u) := \psi_0(\bar{x},\bar{y}_0,u)$. Then we test in turn for each $(\psi,\bar{y}) \in \Psi \times D^m$ whether $\psi(\bar{x},\bar{y},z)$. For each pair, this gives a new formula $E'(u) := E(u) \wedge \psi(\bar{x},\bar{y},u)$. If $E'(M) \subsetneq E(M)$, then we go on with $E = E'$. Otherwise we leave E unchanged. Each time we add an atomic formula to E, this decreases $|E(M)|$ by at least one. This shows that E is made of at most d atomic formulas at the end of the process. We are now ready to define the first set of sign vector formulas. Let $\mathcal{B} = md$. For $\bar{t} \in D^{\mathcal{B}}$, we shall note $\bar{t} = \bar{t}_1.\bar{t}_2 \ldots \bar{t}_d$ with $|\bar{t}_i| = m$. For $\bar{\psi} \in \Psi^d$ and $\bar{t} \in D^{\mathcal{B}}$, let $\mathcal{S}_{\bar{\psi}}[\tau](\bar{x},\bar{y},\bar{t})$ be a quantifier free formula equivalent to

$$\forall u \in M \ \mathrm{loc}_{\bar{\psi}}(\bar{x},\bar{t},u) \rightarrow \tau(\bar{x},\bar{y},u).$$

Moreover, the formula $\mathcal{F}_{\bar{\psi}}$ will express that the set of points satisfying $\mathrm{loc}_{\bar{\psi}}(\bar{x},\bar{t},\cdot)$ is not empty and that all these points have the same sign vector. Let $\mathrm{NotEmpty}_{\bar{\psi}}(\bar{x},\bar{t})$ be a quantifier free formula equivalent to $\exists u \ \mathrm{loc}_{\bar{\psi}}(\bar{x},\bar{t},u)$. Let $\mathrm{same}(\bar{x},\bar{r},u,v)$ be the formula $\bigwedge_{\psi\in\Psi} \psi(\bar{x},\bar{r},u) \leftrightarrow \psi(\bar{x},\bar{r},v)$ and $\theta_{\bar{\psi}}(\bar{x},\bar{t},\bar{r})$ a quantifier free formula equivalent to

$$\forall u,v \ (\mathrm{loc}_{\bar{\psi}}(\bar{x},\bar{t},u) \wedge \mathrm{loc}_{\bar{\psi}}(\bar{x},\bar{t},v)) \rightarrow \mathrm{same}(\bar{x},\bar{r},u,v)\ .$$

At last we define $\mathrm{PreciseLoc}_{\bar{\psi}}(\bar{x},\bar{t}) := \forall^a \bar{r}\ \theta_{\bar{\psi}}(\bar{x},\bar{t},\bar{r})$. Then we can set

$$\mathcal{F}_{\bar{\psi}}(\bar{x},\bar{t}) := \mathrm{NotEmpty}_{\bar{\psi}}(\bar{x},\bar{t}) \wedge \mathrm{PreciseLoc}_{\bar{\psi}}(\bar{x},\bar{t}).$$

Second case : for all $(\psi,\bar{y}) \in \Psi \times D^m$, z is in the infinite set among $\psi(\bar{x},\bar{y},K)$ and its complement. Let $\mathrm{Infini}[\tau](\bar{x},\bar{y})$ be a quantifier free formula equivalent to

$$\exists u_1,\ldots,u_{d+1} \bigwedge_{1\leqslant i<j\leqslant d+1} u_i \neq u_j \wedge \bigwedge_{i=1}^{d+1} \tau(\bar{x},\bar{y},u_i).$$

Hence the sign vector of z is given by $\mathcal{S}_G[\tau](\bar{x},\bar{y},\bar{t}) := \mathrm{Infini}[\tau](\bar{x},\bar{y})$. And we set $\mathcal{F}_G(\bar{x},\bar{t}) :=$ true.

For $\Gamma = \Psi^d \cup \{G\}$, $\mathcal{B} = md$, and the associated formulas defined above, we shall check that the hypothesis of proposition 4 hold. If $\vec{v}$ is a feasible sign vector, there is a point $z \in M$ that proves it. If the point z is in the finite part of $\psi_0(\bar{x},\bar{y}_0,\cdot)$ for a pair $(\psi_0,\bar{y}_0)$, there exists $(\bar{\psi},\bar{t})$ locating precisely z. Thus the formula $\mathcal{F}_{\bar{\psi}}(\bar{x},\bar{t})$ is satisfied and $\vec{v} = \mathcal{S}_{\bar{\psi}}[\cdot](\bar{x},\cdot,\bar{t})$. Otherwise $\vec{v} = \mathcal{S}_G[\cdot](\bar{x},\cdot,\bar{t})$ (for any $\bar{t}$). Conversely suppose there are $\gamma \in \Gamma$ and $\bar{t} \in D^{\mathcal{B}}$ such that i) and ii) hold true. If $\gamma = \bar{\psi}$, then $\mathcal{F}_\gamma(\bar{x},\bar{t})$ proves that there is a point satisfying $\mathrm{loc}_{\bar{\psi}}(\bar{x},\bar{t},\cdot)$ (by NotEmpty), that the points satisfying $\mathrm{loc}_{\bar{\psi}}(\bar{x},\bar{t},\cdot)$ have all the same sign vector (by PreciseLoc), and that the formulas $\mathcal{S}_{\bar{\psi}}[\cdot](\bar{x},\cdot,\bar{t})$ define this sign vector. Otherwise $\gamma = G$; let

$$A_{gen} := \{x \in M, \bigwedge_{\psi\in\Psi} \bigwedge_{\bar{y}\in D^m} |\psi(\bar{x},\bar{y},M)| < \infty \to x \in M \setminus \psi(\bar{x},\bar{y},M)\}.$$

Of course $A_{gen} \neq \emptyset$ because it is the intersection of a finite number of cofinite sets. Moreover it is clear that points from A_{gen} have the sign vector given by $\mathcal{S}_G[\cdot](\bar{x},\cdot,\bar{t})$ (for any $\bar{t}$). $\square$

4.2 On differentially closed fields

We recall that a derivation on a field K is an application $d : K \to K$ such that for all x,y in K, $d(x+y) = d(x)+d(y)$ and $d(xy) = xd(y)+yd(x)$. A differential field K is a field equipped with a derivation. A differential polynomial in the variables $x_1,\ldots,x_k$ is a polynomial in the $d^j(x_i)$ for $1 \leqslant i \leqslant k$ and $j \in \mathbb{N}$. The order of a differential polynomial $p(x)$ is the greatest n such that $x^{(n)}$ appears in p. We say that K is a differentially closed field if for any non-constant polynomials f and g where the order of g is strictly less than the order of f there is an x such that $f(x) = 0 \wedge g(x) \neq 0$ [8]. The structure M is now a zero characteristic differentially closed field K. We begin with two remarks.

Lemma 2 *Let $d,n \in \mathbb{N}$. Then there exists $B' \in \mathbb{N}$ such that for all $N \in \mathbb{N}$ and all differential polynomials $p_1(x),\ldots,p_N(x)$ of order and degree (in $x,x',x'',\ldots$) bounded by n et d respectively, there exists $i_1,\ldots,i_{B'}$ such that*

$$\forall x \in K \bigwedge_{j=1}^{B'} p_{i_j}(x) = 0 \to \bigwedge_{j=1}^{N} p_j(x) = 0.$$

It is enough to take for B' the dimension of the K-vector space of polynomials of order and degree bounded by n et d : indeed we take a family $\{p_{i_1},\ldots\}$ that generates $\{p_1,\ldots,p_N\}$. Thus we can take $B' = (d+1)^{n+1}$. $\square$

Lemma 3 *Let $d,n \in \mathbb{N}$. Then there exists $B'' \in \mathbb{N}$ such that for all polynomials $p_1,\ldots,p_s$ of order and degree bounded by n et d, the following holds. Let $V =$*

$\{x \in K,\ p_1(x) = \ldots = p_s(x) = 0\}$. *Then for all $N \in \mathbb{N}$ and all $q_1(x), \ldots, q_N(x)$ of degree and order less that n et d, there exists $i_1, \ldots, i_{B''}$ such that*

$$\{x \in V, \bigwedge_{i=1}^{N} q_i(x) \neq 0\} = \emptyset \to \{x \in V, \bigwedge_{j=1}^{B''} q_{i_j}(x) \neq 0\} = \emptyset.$$

By lemma 2, we can assume $0 \leqslant s \leqslant B'$. Now s is fixed. For $k \in \{1, \ldots, N\}$, let $\bar{a}_k$ be the tuple containing the coefficients of $p_1, \ldots, p_s$ followed by the coefficients of q_k. Let $\theta(\bar{a}_k, x) := \bigwedge_{i=1}^{s} p_i(x) = 0 \wedge q_k(x) \neq 0$. The NFCP (Not Finite Cover Property), which is verified in differentially closed fields – see [8] section 2, gives us a bound $B''(s)$. Thus $B'' = \max\{B''(s),\ 0 \leqslant s \leqslant B'\}$ works. $\square$

Proposition 6 *Zero characteristic differentially closed fields have the active-natural collapse.*

We shall use proposition 4. Let $\mathcal{T}$ be a finite set of atomic formulas of M : we can assume that they are of the form $p(\bar{x}, \bar{y}, z) = 0$. Let $\mathcal{P}$ be the finite set of differential polynomials $p(\bar{x}, \bar{y}, z)$ appearing in Θ. Let n and d bounding their order and degree (in each of the variables $x, x', x'', \ldots$). Let B' et B'' be the integers given by lemma 2 and 3. Let $\Gamma = \emptyset \cup \mathcal{P} \cup \ldots \cup \mathcal{P}^{B'}$ and $\mathcal{B} = mB'$. We shall note $\bar{t} \in D^{\mathcal{B}}$ as $\bar{t}_1.\bar{t}_2 \ldots \bar{t}_{B'}$ with $|\bar{t}_i| = m$. For $\bar{p} \in \Gamma$ and $\bar{t} \in D^{\mathcal{B}}$, let us define $\mathrm{loc}_{\bar{p}}(\bar{x}, \bar{t}, u) := \bigwedge_{i=1}^{|\bar{p}|} p_i(\bar{x}, \bar{t}_i, u) = 0$ (which is true if $|\bar{p}| = 0$). For $q \in \mathcal{P}$, let $\mathcal{S}_{\bar{p}}[q = 0](\bar{x}, \bar{y}, \bar{t})$ be a quantifier free formula equivalent to $\forall u\ \mathrm{loc}_{\bar{p}}(\bar{x}, \bar{t}, u) \to q(\bar{x}, \bar{y}, u) = 0$. Let $\mathrm{PartialNotEmpty}_{\bar{p},\bar{q}}(\bar{x}, \bar{t}, \bar{r})$ – for $|r| = mB''$ and $|q| = B''$ – be a quantifier free formula equivalent to

$$\exists u \bigwedge_{i=1}^{|\bar{p}|} p_i(\bar{x}, \bar{t}_i, u) = 0 \wedge \bigwedge_{j=1}^{B''} (\mathcal{S}_{\bar{p}}[q_j = 0](\bar{x}, \bar{r}_j, \bar{t}) \vee q_j(\bar{x}, \bar{r}_j, u) \neq 0)$$

Let us define

$$\mathcal{F}_{\bar{p}}(\bar{x}, \bar{t}) := \forall^a \bar{r} \bigwedge_{\bar{q} \in \mathcal{P}^{B''}} \mathrm{PartialNotEmpty}_{\bar{p},\bar{q}}(\bar{x}, \bar{t}, \bar{r}).$$

Let us show that Γ and the associated formulas $\mathcal{F}_\gamma$ and $\mathcal{S}_\gamma$ defined above verify proposition 4. Indeed, let $\vec{v}$ be a sign vector realized by a point $z \in M$. Lemma 2 applied to the set of polynomials $q(\bar{x}, \bar{y}_0, \cdot)$ for all $(q, \bar{y}_0) \in \mathcal{P} \times D^m$ such that $q(\bar{x}, \bar{y}_0, z) = 0$ gives us s pairs $(p_i, \bar{t}_i) \in \mathcal{P} \times D^m$ with $0 \leqslant s \leqslant B'$. Let us show that $\gamma = \bar{p}$ et $\bar{t} = \bar{t}_1 \ldots . \bar{t}_s$ satisfy i) and ii). On one hand $\mathcal{S}_{\bar{p}}[\cdot](\bar{x}, \cdot, \bar{t}) = \vec{v}$: indeed $\models \mathcal{S}_{\bar{p}}[q = 0](\bar{x}, \bar{r}, \bar{t})$ iff $q(\bar{x}, \bar{r}, K) = 0$ contains $\bigcap_i p_i(\bar{x}, \bar{t}_i, K) = 0$, which is true iff $q(\bar{x}, \bar{r}, z) = 0$. On the other hand $\models \mathcal{F}_{\bar{p}}(\bar{x}, \bar{t})$ since z exists. Conversely, let $\bar{p}, \bar{t}$ be such that $\models \mathcal{F}_{\bar{p}}(\bar{x}, \bar{t})$. Let $\vec{v} = \mathcal{S}_{\bar{p}}[\cdot](\bar{x}, \cdot, \bar{t})$. As we have $\mathcal{F}_{\bar{p}}(\bar{x}, \bar{t})$, lemma 3 tells us that the set of points of K having this sign vector is not empty. $\square$

4.3 Uniform quantifier elimination

In this section we are interested in a notion introduced by Basu : uniform quantifier elimination [2]. In this article the question whether the theories ACF_p (p prime or zero) and DCF have uniform quantifier elimination was raised. We can bring a positive answer to these questions, since we remark that uniform quantifier elimination is equivalent to active-natural collapse in the case of a single unary relation.

Proposition 7 *A structure has uniform quantifier elimination iff it has the active natural collapse for a single unary relation.*

First of all, let us remark that in the case where there is a single unary relation I, any active formula can be written without using I : just replace $I(t(\bar{w}))$ with $\exists^a v \; v = t(\bar{w})$. That is what we are going to do in the following.

Now to each active formula we associate a uniform family of formulas, and conversely. We shall do it for prenex ones – but this is not imperative. Let $\Phi(\bar{x}, \bar{y})$ be the uniform family

$$\phi_n(\bar{x}, \bar{y}) = Q^1_{1 \leqslant i_1 \leqslant n} \dots Q^m_{1 \leqslant i_m \leqslant n} \phi(\bar{x}, y_{i_1}, \dots, y_{i_m})$$

where $Q_i \in \{\bigvee, \bigwedge\}$. To this family will correspond the following active formula

$$\psi(\bar{x}) := \mathcal{Q}^1 t_1 \dots \mathcal{Q}^m t_m \; \phi(\bar{x}, t_1, \dots, t_m)$$

with $\mathcal{Q}^i = \exists^a$ (resp. $\forall^a$) if $Q^i = \bigvee$ (resp. $\bigwedge$). Of course one can make the correspondance the other way. Moreover $\Phi(\bar{x}, \bar{y})$ and $\psi(\bar{x})$ are related this way : $(M, \mathcal{I}) \models \psi(\bar{x})$ if and only if $M \models \phi_n(\bar{x}, \bar{y}_{\mathcal{I}})$ where $n = |\mathcal{I}|$ and $\bar{y}_{\mathcal{I}}$ lists the elements of $\mathcal{I}$. A structure has uniform quantifier elimination iff for any uniform family $\Phi(\bar{x}, \bar{y}, z)$, there exists a uniform family $\tilde{\Phi}(\bar{x}, \bar{y})$ such that $\forall \bar{x} \in M \; \forall n \in \mathbb{N} \; (\exists z \in M \; \phi_n(\bar{x}, \bar{y}, z) \leftrightarrow \tilde{\phi}_n(\bar{x}, \bar{y}))$. In the same way, a structure has the active-natural collapse for a unary relation iff for any active formula $\psi(\bar{x}, z)$, there exists an active formula $\tilde{\psi}(\bar{x})$ such that $\forall \bar{x} \in M (\exists z \in M \; \psi(\bar{x}, z) \leftrightarrow \tilde{\psi}(\bar{x}))$. Thus it is clear that these two notions are equivalent. □

Corollary 5 *Strongly minimal structures with quantifier elimination and zero characteristic differentially closed fields have uniform quantifier elimination.* □

We have the following : active-natural collapse ⇒ uniform QE ⇒ QE. Moreover, there exists a structure that eliminates quantifiers but not uniformly : the ternary random structure [3]. Is there a structure that admits uniform quantifier elimination but not the (full) active-natural collapse ? One can also wonder about all intermediate questions.

References

[1] J. T. Baldwin and M. Benedikt. Embedded finite models, stability theory, and the impact of order. In *Proc. 13th IEEE Symposium on Logic in Computer Science*, pages 490–500, 1998.

[2] S. Basu. New results on quantifier elimination over real closed fields and applications to constraint databases. *Journal of the ACM*, 46(4):537–555, july 1999.

[3] M. Benedikt and L. Libkin. Relational queries over interpreted structures. *Journal of the ACM*, 47(4):644–680, 2000.

[4] O. Chapuis and P. Koiran. Definability of geometric properties in algebraically closed fields. *Mathematical Logic Quarterly*, 45(4):533–550, 1999.

[5] H.-D. Ebbinghaus and J. Flum. *Finite Model Theory*. Springer-Verlag, 1995.

[6] P. Koiran. Elimination of parameters in the polynomial hierarchy. *Theoretical Computer Science*, 215(1-2):289–304, 1999.

[7] G. Kuper, L. Libkin, and J. Paredaens, editors. *Constraint Databases*. Springer-Verlag, 2000.

[8] D Marker. Model theory of differential fields. In *Model theory of fields*, number 5 in Lecture notes in logic. Springer, 1996.