Math. Log. Quart. **64,** No. 3, 218–248 (2018) / **DOI** 10.1002/malq.201700055

Finite satisfiability for two-variable, first-order logic with one transitive relation is decidable

Ian Pratt-Hartmann[1,2]*

[1] School of Computer Science, University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom
[2] Instytut Matematyki i Informatyki, Uniwersytet Opolski, ul. Oleska 48, 45–040 Opole, Poland

Received 15 September 2017, revised 10 November 2017, accepted 20 December 2017
Published online 24 July 2018

We consider two-variable, first-order logic in which a single distinguished predicate is required to be interpreted as a transitive relation. We show that the finite satisfiability problem for this logic is decidable in triply exponential non-deterministic time. Complexity falls to doubly exponential non-deterministic time if the transitive relation is constrained to be a partial order.

1 Introduction

The *two-variable fragment*, henceforth denoted $\mathcal{L}^2$, is the fragment of first-order logic with equality but without function-symbols, in which only two logical variables may appear. It is well-known that $\mathcal{L}^2$ has the finite model property, and that its satisfiability (= finite satisfiability) problem is NEXPTIME-complete [3]. It follows that it is impossible, within $\mathcal{L}^2$, to express the condition that a given binary predicate r denotes a transitive relation, since in that case the $\mathcal{L}^2$-formula $\forall x \neg r(x, x) \wedge \forall x \exists y . r(x, y)$ becomes an axiom of infinity. This observation has prompted investigation of what happens when $\mathcal{L}^2$ is enriched by imposing various semantic restrictions on the interpretations of certain predicates. For $k > 0$, denote by $\mathcal{L}^2k$T the logic whose formulas are exactly those of $\mathcal{L}^2$, but where k distinguished predicates are required to be interpreted as *transitive relations*, and denote by $\mathcal{L}^2k$E the same set of formulas, but where k distinguished predicates are required to be interpreted as *equivalence relations*. For each of these logics, the question arises as to whether the satisfiability and finite satisfiability problems are decidable, and, if so, what their computational complexity is.

The following is known. (i) $\mathcal{L}^2$1E has the finite model property, and its satisfiability (= finite satisfiability) problem is NEXPTIME-complete [8]. (ii) $\mathcal{L}^2$2E lacks the finite model property, but its satisfiability and finite satisfiability problems are both 2-NEXPTIME-complete [7]. (iii) For $k \geq 3$, the satisfiability and finite satisfiability problems for $\mathcal{L}^2k$E are both undecidable [8]. (iv) $\mathcal{L}^2$1T lacks the finite model property and its satisfiability problem is 2-EXPTIME-hard [4] and in 2-NEXPTIME [13]. (v) For $k \geq 2$, the satisfiability and finite satisfiability problems for $\mathcal{L}^2k$T are both undecidable [5]. (In fact, the satisfiability and finite satisfiability problems for the two-variable fragment with one transitive relation and one equivalence relation are already undecidable [9].) This resolves the decidability and (within narrow limits) the complexity of the satisfiability and finite satisfiability problems for all of the logics $\mathcal{L}^2k$T and $\mathcal{L}^2k$E except for one case: the finite satisfiability problem for $\mathcal{L}^2$1T, where decidability is currently open. This article deals with that case by showing that the finite satisfiability problem for $\mathcal{L}^2$1T is in 3-NEXPTIME. The best currently known lower bound for this problem is 2-EXPTIME-hard [4]. We remark that the approach employed in [13] to establish the decidability of the satisfiability problem for $\mathcal{L}^2$1T breaks down if models are required to be finite: the algorithm presented here for determining finite satisfiability employs a quite different strategy.

Denote by $\mathcal{L}^2$1PO the logic defined in exactly the same way as $\mathcal{L}^2$1T, except that the distinguished binary relation is constrained to be interpreted as a (strict) *partial order*—i.e., as a transitive and irreflexive relation. Since the $\mathcal{L}^2$-formula $\forall x \neg r(x, x)$ asserts that r is irreflexive, it follows that $\mathcal{L}^2$1PO is no stronger, in terms of

* E-mail: ipratt@cs.man.ac.uk

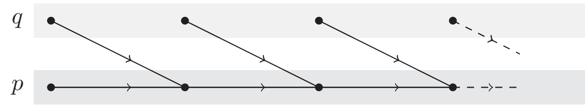


Fig. 1 A linear order on the elements satisfying p, and an anti-chain on the elements satisfying q.

expressive power, than $\mathcal{L}^2$1T. In addition, we take the logic $\mathcal{L}^2$1PO$^{\text{u}}$ to be the fragment of $\mathcal{L}^2$1PO in which—apart from equality and the distinguished (partial order) predicate—only *unary* predicates are allowed. Our strategy in the sequel is first to consider $\mathcal{L}^2$1PO$^{\text{u}}$. Structures interpreting this logic are, in effect, partial orders in which each element is assigned one of a finite number of *types*. We obtain a 2-NEXPTIME upper complexity-bound on the finite satisfiability problem for this logic, by introducing a method for 'factorizing' such typed partial orders into smaller partial orders on blocks of elements of the same type. We then extend this upper bound to $\mathcal{L}^2$1PO by exhibiting a method to eliminate all binary predicates in $\mathcal{L}^2$1PO-formulas (other than equality and the distinguished predicate). Finally, we obtain the 3-NEXPTIME upper complexity-bound on the finite satisfiability problem for $\mathcal{L}^2$1T by exhibiting a method to replace the distinguished transitive relation by a partial order. This latter reduction produces an exponential increase in the size of the formula in question.

Stronger complexity-theoretic upper bounds are available when the distinguished predicates are required to be interpreted as *linear* orders: the satisfiability and finite satisfiability problems for $\mathcal{L}^2$ together with one linear order are both NEXPTIME-complete [10]; the finite satisfiability problem for $\mathcal{L}^2$ together with two linear orders is in 2-NEXPTIME [14] (falling to EXPSPACE when all non-navigational predicates are unary [11]); with three linear orders, satisfiability and finite satisfiability are both undecidable [6, 10]. Also somewhat related to $\mathcal{L}^2$1PO$^{\text{u}}$ is the propositional modal logic known as *navigational XPATH*, which features a signature of proposition letters interpreted over vertices of some finite, ordered tree, together with modal operators giving access to vertices standing in the relations of *daughter* and *next-sister*, as well as their transitive closures. It is known that, over finite trees, navigational XPATH has the same expressive power as two-variable, first-order logic with a signature consisting of unary predicates (representing properties of vertices) together with binary 'navigational' predicates (representing the modal accessibility relations). The exact complexity of satisfiability for all natural variants of this logic is given in [1].

To convey a sense of the expressive power of the logics we are working with, we give an example showing that the logic $\mathcal{L}^2$1PO$^{\text{u}}$ can force the existence of an infinite anti-chain: that is, an infinite collection of elements none of which is related to any other in the partial ordering. The example is due to E. Kieroński (personal communication). In the following, we use $<$ as the distinguished binary predicate of $\mathcal{L}^2$1PO$^{\text{u}}$ (written using infix notation). First of all, the formulas

$$\exists x . p(x) \qquad\qquad \forall x \forall y (p(x) \wedge p(y) \to (x < y \vee x = y \vee y < x))$$

ensure that elements satisfying p form a non-empty linear order. Pick some such element a_1. Now the formulas

$$\forall x (p(x) \to \exists y (\neg x < y \wedge \neg y < x \wedge q(y))) \qquad\qquad \forall x (q(x) \to \exists y (x < y \wedge p(y)))$$

ensure that, for every element, say a_i, satisfying p, there is an incomparable element, say b_i, satisfying q, and, for every element b_i satisfying q, there is a greater element, say a_{i+1}, satisfying p. Thus, we generate sequences of elements $a_1, a_2, \ldots$, satisfying p, and $b_1, b_2, \ldots$, satisfying q. A moment's thought shows that for all i, $a_i < a_{i+1}$, so that, by a simple induction, $b_i < a_j$ for all $i < j$. This immediately implies that the b_j are all distinct, since a_j and b_j are, by construction, incomparable. The formula

$$\forall x \forall y (q(x) \wedge q(y) \to (\neg x < y \wedge \neg y < x))$$

then secures the sought-after infinite anti-chain. That the formulas are satisfiable is shown by the partially-ordered structure depicted in Fig. 1. We remark that, even under the assumption that structures are finite, $\mathcal{L}^2$1PO$^{\text{u}}$ can force doubly-exponential-sized models; this is demonstrated, for example, using the construction of [4].

2 Preliminaries

We employ standard model-theoretic notation: structures are indicated by (possibly decorated) *fraktur* letters $\mathfrak{A}$, $\mathfrak{B}, \ldots$, and their domains by the corresponding Roman letters $A, B, \ldots$. In this paper, we adopt the non-standard assumption that all structures have cardinality at least 2. Thus, the formula $\forall x \exists y (x \neq y)$ is for us a validity, and $\forall x \forall y (x = y)$ a contradiction. This assumption does not represent a significant restriction: over domains of size 1, first-order logic reduces to propositional logic.

A binary relation R on some carrier set A is *transitive* if aRb and bRc implies aRc, *reflexive* if aRa always holds, *irreflexive* if aRa never holds, and *anti-symmetric* if aRb and bRa implies $a = b$. Every transitive, irreflexive relation is trivially anti-symmetric. A *weak partial order* is a relation that is transitive, reflexive and anti-symmetric; a *strict partial order* is a relation that is transitive and irreflexive. If R is a weak partial order and I the identity (diagonal) relation on A, then $R \setminus I$ is a strict partial order; moreover, all strict partial orders on A arise in this way. Likewsie, if R is a strict partial order then $R \cup I$ is a weak partial order; moreover, all weak partial orders on A arise in this way. In the sequel, the unmodified phrase *partial order* will always mean *strict* partial order.

The *two-variable fragment*, here denoted $\mathcal{L}^2$, is the fragment of first-order logic with equality but without function-symbols, in which only two variables, x and y, may appear. There are no other syntactic restrictions. In particular, formulas such as $\forall x (p(x) \rightarrow \exists y (r(x, y) \wedge \exists x . s(y, x)))$, in which bound occurrences of a variable u may appear within the scope of a quantifier Qu, are allowed. It is routine to show that predicates having arity other than 1 or 2 add no effective expressive power in the context of $\mathcal{L}^2$. It is likewise routine to show that individual constants add no effective expressive power given the presence of the equality predicate. Henceforth, then, we shall take all signatures to consist only of unary and binary predicates.

We define $\mathcal{L}^2$1T to be the set of formulas of $\mathcal{L}^2$ over any signature of unary and binary predicates which features a distinguished binary predicate $\mathfrak{t}$. The semantics of $\mathcal{L}^2$1T is exactly as for $\mathcal{L}^2$, except that the interpretation of $\mathfrak{t}$ is required to be a *transitive relation*. Similarly, we define $\mathcal{L}^2$1PO to be the set of formulas of $\mathcal{L}^2$ over any signature of unary and binary predicates which features a distinguished binary predicate $<$ (written using infix notation). The semantics of $\mathcal{L}^2$1PO is exactly as for $\mathcal{L}^2$, except that the interpretation of $<$ is required to be a *partial order*. Finally, we define $\mathcal{L}^2$1PO$^{\mathrm{u}}$ to be the subset of $\mathcal{L}^2$1PO in which no binary predicates other than $=$ and $<$ appear.

A formula of $\mathcal{L}^2$ is said to be *unary* if it features just one free variable. A unary formula ζ is generally silently assumed to have x as its only free variable; if ζ is such a formula, we write $\zeta(y)$ for the result of replacing x in ζ by y. The unary formulas $\mu_1, \ldots, \mu_n$ are *mutually exclusive* if $\models \forall x (\mu_i \rightarrow \neg \mu_j)$ for all i $(1 \leq i < j \leq n)$. Any formula η of $\mathcal{L}^2$ with two free variables is assumed to have those variables taken in the order x, y. Thus, we write $\mathfrak{A} \models \eta[a, b]$, where a, b are elements of A, to indicate that η is satisfied in $\mathfrak{A}$ under the assignment $a \mapsto x$ and $b \mapsto y$. For the purposes of this paper, we may take the *size* of an $\mathcal{L}^2$-formula φ, denoted $\|\varphi\|$, to be the number of symbols it contains.

For any signature σ, a σ*-atom* is a formula of the form $p(\bar{x})$ where p is a predicate of $\sigma \cup \{=\}$ and $\bar{x}$ a tuple of variables of the appropriate arity. A σ*-literal* is a σ-atom or a negated σ-atom. The reference to σ is omitted if unimportant or clear from context. A *1-type* over σ is a maximal consistent set of σ-literals involving only the variable x; and a *2-type* over σ is a maximal consistent set of σ-literals involving the variables x and y. (Thus, 1- and 2-types are what are sometimes called *atomic* 1- and 2-types.) Consistency here is to be understood as taking into account the semantic constraints on distinguished predicates. Thus, if σ contains $<$, then the 1-type over σ contains the literals $x = x$ and $\neg x < x$, with similar restrictions applying to 2-types. Likewise, if σ contains $\mathfrak{t}$, then any 2-type containing the literals $\mathfrak{t}(x, y)$ and $\mathfrak{t}(y, x)$ also contains $\mathfrak{t}(x, x)$ and $\mathfrak{t}(y, y)$. We usually identify 1- and 2-types with the conjunction of their literals. If $\mathfrak{A}$ is a structure and $a, b \in A$, we write $\mathrm{tp}^{\mathfrak{A}}[a]$ for the unique 1-type satisfied in $\mathfrak{A}$ by a and $\mathrm{tp}^{\mathfrak{A}}[a, b]$ for the unique 2-type satisfied by $\langle a, b \rangle$.

A formula φ of $\mathcal{L}^2$ (or of $\mathcal{L}^2$1PO or $\mathcal{L}^2$1T) is said to be in *standard normal form* if it conforms to the pattern

$$\forall x \forall y (x = y \vee \eta) \wedge \bigwedge_{h=0}^{m-1} \forall x \exists y (x \neq y \wedge \vartheta_h), \tag{1}$$

where η, $\vartheta_0, \ldots, \vartheta_{m-1}$ are quantifier- and equality-free formulas, with $m \geq 1$. A formula is said to be in *weak normal form* if it conforms to the pattern

$$\bigwedge_{\zeta \in Z} \exists x.\zeta \wedge \forall x \forall y(x = y \vee \eta) \wedge \bigwedge_{h=0}^{m-1} \forall x \exists y(x \neq y \wedge \vartheta_h), \tag{2}$$

where Z is a finite set of unary quantifier- and equality-free formulas and the other components are as in (1). We refer to the parameter m in both (1) and (2) as the *multiplicity* of φ.

The following basic fact about $\mathcal{L}^2$ goes back, essentially, to [12], and is widely used in studies of $\mathcal{L}^2$ and its variants [2, Lemma 8.1.2]. Remembering our general assumption that all structures have cardinality at least 2, we have:

Lemma 2.1 *Let φ be an $\mathcal{L}^2$-formula. There exists a standard normal-form $\mathcal{L}^2$-formula φ' such that: (i) $\models \varphi' \to \varphi$; (ii) every model of φ can be expanded to a model of φ'; and (iii) $\|\varphi'\|$ is bounded by a polynomial function of $\|\varphi\|$.*

Obviously, Lemma 2.1 applies without change to $\mathcal{L}^2$1T, $\mathcal{L}^2$1PO and (following a simple check) $\mathcal{L}^2$1PO$^{\mathrm{u}}$. Under our general restriction to structures with at least 2 elements, $\exists x.\zeta$ is logically equivalent to $\forall x \exists y(x \neq y \wedge (\zeta \vee \zeta(y)))$. Hence any formula in weak normal form can be converted, in polynomial time, to a logically equivalent one in standard normal form. However, this process increases the multiplicity of the formula in question: in the sequel, we shall sometimes need (2) in full generality, in order to obtain finer control over this parameter.

3 Unary two-variable logic with one partial order

The purpose of this section is to show that the logic $\mathcal{L}^2$1PO$^{\mathrm{u}}$ has the doubly exponential-sized finite model property (Theorem 3.24): if φ is a finitely satisfiable $\mathcal{L}^2$1PO$^{\mathrm{u}}$-formula, then φ has a model of size bounded by some fixed doubly exponential function of $\|\varphi\|$. It follows that the finite satisfiability problem for $\mathcal{L}^2$1PO$^{\mathrm{u}}$ is in 2-NExpTime.

All structures in this section interpret a signature of unary predicates, together with the distinguished predicate $<$. To make reading easier, we typically write $x > y$ for $y < x$ and $x \sim y$ for $\neg(x = y \vee x < y \vee y < x)$. In practice, we shall simply treat the symbols $>$ and $\sim$ as if they were binary predicates (subject to the obvious constraints on their interpretations). With this concession to informality, we see that, in the logics $\mathcal{L}^2$1PO and $\mathcal{L}^2$1PO$^{\mathrm{u}}$, any pair of distinct elements of a structure satisfies exactly one of the atomic formulas $x < y$, $x > y$ or $x \sim y$. Where a structure $\mathfrak{A}$ is clear from context, we typically do not distinguish between the predicate $<$ and its interpretation in $\mathfrak{A}$, writing $a < b$ to mean $\langle a, b \rangle \in <^{\mathfrak{A}}$; similarly for $>$, $\sim$ and $=$. We sometimes refer to the distinguished predicates $\mathfrak{t}$, $<$, $>$, $\sim$ and $=$ as *navigational* predicates. (The allusion here is to the terminology employed in XPATH.) A predicate that is not navigational is called *ordinary*. A formula is navigation-free if it contains no navigational predicates. A formula is said to be *pure Boolean* if it is quantifier- and navigation-free—i.e., if it is a Boolean combination of literals featuring ordinary predicates. Notice that all 1-types contain the conjuncts $\neg x < x$ and $x = x$, and hence are not, technically speaking, pure Boolean formulas. However, they are of course logically equivalent to the pure Boolean formulas obtained by deleting all navigational conjuncts.

In this section, we use the (possibly decorated) variables α, β, γ, π to range over 1-types, μ to range over unary pure Boolean formulas, and ζ, η, ϑ, φ, χ, ψ to range over other $\mathcal{L}^2$1PO$^{\mathrm{u}}$-formulas.

3.1 Basic formulas

Structures interpreting $\mathcal{L}^2$1PO$^{\mathrm{u}}$-formulas have a very simple form, and it will be convenient to diverge slightly from standard model-theoretic terminology when discussing them. (Remember, all structures are taken to have cardinality at least 2 in this paper.) Let Π be a fixed set of 1-types over some unary signature σ. A *typed partial order* (*over*Π) is a triple $\mathfrak{A} = (X, <, \mathrm{tp})$, where X is a set of cardinality at least 2, $<$ a partial order on X, and $\mathrm{tp} : X \to \Pi$ a function. We can regard $\mathfrak{A}$ as a structure interpreting $\mathcal{L}^2$1PO$^{\mathrm{u}}$-formulas in the obvious way; and it is evident that all structures interpreting $\mathcal{L}^2$1PO$^{\mathrm{u}}$-formulas can be regarded as typed partial orders over some set of 1-types. This is what we shall do in the sequel, therefore. If $\mathfrak{A} = (X, <, \mathrm{tp})$ is a typed partial order and $a \in X$, we call *a maximal* if it is a largest element of its 1-type, i.e., there exists no a' such that $\mathrm{tp}(a') = \mathrm{tp}(a)$ and $a < a'$; similarly, *mutatis mutandis*, for *minimal*. We call *a extremal* if it is either maximal or minimal.

We begin by establishing a stronger normal form theorem for $\mathcal{L}^2 1\mathrm{PO}^{\mathrm{u}}$. Recall that we take unary formulas (including 1-types) to have free variable x unless otherwise indicated. We call a formula *basic* if it has one of the forms

$$\forall x(\alpha \to \forall y(\alpha(y) \to x = y)) \tag{B1a}$$

$$\forall x(\alpha \to \forall y(\beta(y) \to x = y)) \tag{B1b}$$

$$\forall x(\alpha \to \forall y(\alpha(y) \wedge x \neq y \to x \sim y)) \tag{B2a}$$

$$\forall x(\alpha \to \forall y(\beta(y) \to x \sim y)) \tag{B2b}$$

$$\forall x(\alpha \to \forall y(\beta(y) \to x < y)) \tag{B3}$$

$$\forall x(\alpha \to \forall y(\beta(y) \to (x < y \vee x \sim y))) \tag{B4}$$

$$\forall x(\alpha \to \forall y(\alpha(y) \wedge x \neq y \to (x < y \vee x > y))) \tag{B5a}$$

$$\forall x(\alpha \to \forall y(\beta(y) \to (x < y \vee x > y))) \tag{B5b}$$

$$\forall x(\alpha \to \exists y(\mu(y) \wedge \neg\alpha(y) \wedge x < y)) \tag{B6}$$

$$\forall x(\alpha \to \exists y(\mu(y) \wedge \neg\alpha(y) \wedge x > y)) \tag{B7}$$

$$\forall x(\alpha \to \exists y(\mu(y) \wedge x \sim y)) \tag{B8}$$

$$\forall x.\mu \tag{B9}$$

$$\exists x.\mu, \tag{B10}$$

where α and β are distinct 1-types and μ is a unary pure Boolean formula. We typically use the variable ψ to range over basic formulas and Ψ to range over finite sets of basic formulas. Formulas of the forms (B3) and (B5b) receive special treatment in the sequel, and will be referred to—for reasons that will become evident—as *factor-controllable* formulas. If Ψ is any finite set of basic formulas, we denote by $\mathrm{FC}(\Psi)$ the set of factor-controllable formulas in Ψ.

Lemma 3.1 *Let φ be a weak normal-form $\mathcal{L}^2 1\mathrm{PO}^{\mathrm{u}}$-formula with multiplicity m over signature σ. There exists an $\mathcal{L}^2 1\mathrm{PO}^{\mathrm{u}}$-sentence φ^* over a signature σ^*, such that: (i) φ and φ^* are satisfiable over the same finite domains; (ii) $|\sigma^*| = |\sigma| + 3m$; and (iii) φ^* is a conjunction of basic formulas.*

Proof. Let φ be as given in (2). The conjuncts $\exists x.\zeta$ are already of the form (B10), and so require no action. Consider next any conjunct $\chi_h = \forall x \exists y(x \neq y \wedge \vartheta_h)$, where $0 \leq h < m$. Letting $p_{h,<}$, $p_{h,>}$ and $p_{h,\sim}$ be fresh unary predicates, we may replace χ_h by the conjunction χ_h^* of the formulas

$$\forall x(p_{h,<}(x) \vee p_{h,>}(x) \vee p_{h,\sim}(x)) \tag{3}$$

$$\forall x(p_{h,<}(x) \to \exists y(\vartheta_h \wedge x < y)) \tag{4}$$

$$\forall x(p_{h,>}(x) \to \exists y(\vartheta_h \wedge x > y)) \tag{5}$$

$$\forall x(p_{h,\sim}(x) \to \exists y(\vartheta_h \wedge x \sim y)). \tag{6}$$

Obviously, $\models \chi_h^* \to \chi_h$; moreover, any model $\mathfrak{A}$ of χ_h can be expanded to a model $\mathfrak{A}'$ of χ_h by setting $p_{h,<}^{\mathfrak{A}'} = \{a \in A \mid \text{there exists } b \text{ s.t. } a < b \text{ and } \mathfrak{A} \models \vartheta_h[a, b]\}$, and similarly for $p_{h,>}$ and $p_{h,\sim}$. Carrying out this replacement for all h $(0 \leq h < m)$, let σ^* denote the enlarged signature. Evidently, $|\sigma^*| = |\sigma| + 3m$.

Formula (3) is of the form (B9). Now replace any formula of the form (4) by the conjunction of all formulas of the forms $\forall x(\alpha(x) \to \exists y([\vartheta_h/\alpha] \wedge x < y))$, where α ranges over the set of 1-types over σ^*, and $[\vartheta_h/\alpha]$ denotes the result of replacing each ordinary literal $q(x)$ in ϑ_h by $\top$ or $\perp$ as determined by $\alpha(x)$. Doing the same for (5) and (6), and replacing any navigational literals in $[\vartheta_h/\alpha]$ by $\top$ or $\perp$ in the obvious way yields logically equivalent conjunctions of formulas of the respective forms

$$\forall x(\alpha \to \exists y(\mu(y) \wedge x < y)) \tag{B6$'$}$$

$$\forall x(\alpha \to \exists y(\mu(y) \wedge x > y)) \tag{B7$'$}$$

$$\forall x(\alpha \to \exists y(\mu(y) \wedge x \sim y)), \tag{B8$'$}$$

where μ is a pure Boolean (i.e., quantifier- and navigation-free) formula not involving the variable x. Notice that, over finite structures $\mathfrak{A}$, (B6$'$) entails $\forall x(\alpha \to \exists y(\mu(y) \wedge \neg\alpha(y) \wedge x < y))$. This is obvious since, if $\mathfrak{A} \models \alpha[a]$, let a' be a maximal element of 1-type α above a. That is: $a \leq a'$, $\mathfrak{A} \models \alpha[a']$ and there does not exist a'' such that $a' < a''$ and $\mathfrak{A} \models \alpha[a'']$. By (B6$'$), let b be such that $a' < b$ and $\mathfrak{A} \models \mu[b]$. But then $\mathfrak{A} \models \neg\alpha[b]$ and $a < b$, as required. Thus, (B6$'$) can be replaced by (B6). Likewise, (B7$'$) can be replaced by (B7). Notice that (B8$'$) is just (B8).

Consider finally the conjunct $\chi = \forall x \forall y(x = y \vee \eta)$ of φ. Clearly, we may replace this formula by the conjunction χ^* of all formulas of the forms

$$\forall x(\alpha \to \forall y(\beta(y) \wedge x \neq y \to [\eta/(\alpha, \beta)])),$$

where α and β range over the set of 1-types over σ^*, and $[\eta/(\alpha, \beta)]$ denotes the result of replacing each unary literal in η by its truth-value as determined by α and $\beta(y)$. Clearly, $\models \chi \leftrightarrow \chi^*$. Furthermore, any sub-formula $[\eta/(\alpha, \beta)]$ features only the navigational predicates $>$, $<$ and $\sim$, and thus is logically equivalent to one of the forms $\perp$, $x \sim y$, $x > y$, $x < y$, $(x > y \vee x \sim y)$, $(x < y \vee x \sim y)$, $(x < y \vee x > y)$ or $\top$. Ignoring the trivial case $\top$, and exchanging the variables x and y if necessary, we obtain the forms

$$\forall x(\alpha \to \forall y(\beta(y) \wedge x \neq y \to \perp)) \tag{B1$'$}$$

$$\forall x(\alpha \to \forall(\beta(y) \wedge x \neq y \to x \sim y)) \tag{B2$'$}$$

$$\forall x(\alpha \to \forall y(\beta(y) \wedge x \neq y \to x < y)) \tag{B3$'$}$$

$$\forall x(\alpha \to \forall y(\beta(y) \wedge x \neq y \to (x < y \vee x \sim y))) \tag{B4$'$}$$

$$\forall x(\alpha \to \forall y(\beta(y) \wedge x \neq y \to (x < y \vee x > y))). \tag{B5$'$}$$

We consider these forms in turn, according as α and β are identical or distinct. For (B1$'$), we have (B1a) and (B1b). For (B2$'$), we obtain (B2a) and (B2b). For (B3$'$), if $\alpha = \beta$, we have (B1a) again; if $\alpha \neq \beta$, we have (B3). For (B4$'$), if $\alpha = \beta$, we have (B2a) again; if $\alpha \neq \beta$, we have (B4). For (B5$'$), we obtain (B5a) and (B5b). □

3.2 Factorizations

The following notion will play a crucial role in the sequel. Let $\mathfrak{A} = (X, <, \mathrm{tp})$ be a typed partial order. A *factorization* of $\mathfrak{A}$ is a pair $\mathbb{B} = (\mathbf{B}, \ll)$, where $\mathbf{B}$ is a partition of X, and $\ll$ is a partial order on $\mathbf{B}$ satisfying:

(F1) for all $B \in \mathbf{B}$, there exists $\pi \in \Pi$, denoted $\mathrm{tp}(B)$, such that, for all $b \in B$, $\mathrm{tp}(b) = \pi$;
(F2) for all $\pi \in \Pi$, the set $\{B \in \mathbf{B} \mid \mathrm{tp}(B) = \pi\}$ is linearly ordered by $\ll$;
(F3) for all $A, B \in \mathbf{B}$, if $A \ll B$, then, for all $a \in A$ and all $b \in B$, $a < b$.

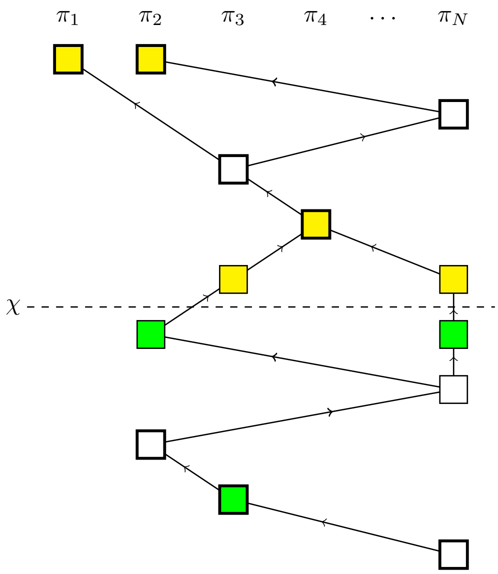


Fig. 2 Factorization of a finite typed partial order over 1-types $\pi_1, \ldots, \pi_N$, with cut χ, showing $\mathbf{B}^{\times}$ (thick lines), and $F^{+}(\chi)$ and $F^{-}(\chi)$ (shading).

We refer to the elements of $\mathbf{B}$ as *blocks*, and to the ordering $\ll$ as the *block ordering* (in contradistinction to the *element ordering* $<$). Notice that, if $|\mathbf{B}| \geq 2$, the triple $(\mathbf{B}, \ll, \text{tp})$ is itself a typed partial order. If $\text{tp}(B) = \alpha$, we call B an α-block. We say that a block B is *of type* $\alpha \vee \beta$ if it is either of type α or of type β, and we call B an $(\alpha \vee \beta)$-block.

Our strategy in the sequel will be as follows. Suppose Ψ is a finite set of basic formulas and $\mathfrak{A}$ a finite model of Ψ. We first obtain a factorization $\mathbf{B}$ of $\mathfrak{A}$ which guarantees (in a sense to be explained below) the truth of all formulas in $\text{FC}(\Psi)$. Then we show how to remove blocks from $\mathfrak{A}$, thus obtaining a finite model of Ψ with factorization $\mathbf{B}'$ such that the number of blocks $|\mathbf{B}'|$ is *small* (bounded by a doubly exponential function of the size of the signature of Ψ). Finally, we replace the blocks of $\mathbf{B}'$ with similarly *small* groups of elements (but retaining the block-structure of $\mathbf{B}'$) so as to obtain the sought-after model of Ψ.

In the context of a factorization $(\mathbf{B}, \ll)$, we use $A \gg B$ as an alternative to $B \ll A$. If A and B are blocks, we write $A \approx B$ to mean that A and B are distinct and neither $A \ll B$ nor $B \ll A$. Thus, $\approx$ stands in the same relation to $\ll$ as $\sim$ does to $<$. Note that, if $A \approx B$, it is possible for there to be $a, a' \in A$ and $b, b' \in B$ such that $a < b$ and $a' > b'$. We carry over the terms *maximal*, *minimal* and *extremal* from elements to blocks in the expected way: a block B is *maximal* if there exists no block B' such that $\text{tp}(B) = \text{tp}(B')$ and $B \ll B'$; similarly for *minimal*. A block is *extremal* if it is either maximal or minimal. We denote the set of extremal blocks of $\mathbf{B}$ by $\mathbf{B}^{\times}$.

Thus, a factorization of a typed partial order is an organization of its elements into blocks of uniform type, with a partial order on the blocks such that all blocks of a given type are linearly ordered, and such that, whenever one block is less than another in the block ordering, every element of the first block is less than every element of the second in the element ordering. Figure 2 shows a factorization of a finite typed partial order over 1-types $\pi_1, \ldots, \pi_N$, depicted as an acyclic directed graph: the block order $\ll$ is the transitive closure of the edges; extremal blocks are marked with thick boundaries. The shaded blocks and the line marked χ will be explained in Sec. 3.3.

It is important to realize that the factorization $(\mathbf{B}, \ll)$ does not determine the partial order $(X, <)$. Indeed, any typed partial order $\mathfrak{A}$ has a factorization, namely, the trivial factorization in which the blocks are simply the non-empty sets $\{a \in X \mid \text{tp}(a) = \pi\}$ for $\pi \in \Pi$, and the block-order is empty. The next two lemmas show that we can generally find more informative factorizations than this. Recall in this context that a factor-controllable basic formula is one of either of the forms (B3) or (B5b).

Lemma 3.2 *Let $\mathfrak{A}$ be a typed partial order over a set of types Π, with $\alpha, \beta \in \Pi$ distinct. Suppose $\mathbb{B}$ is a factorization of $\mathfrak{A}$ in which every block of type α lies below every block of type β in the block order. Then the formula* (B3)*, namely*

$$\forall x(\alpha(x) \to \forall y(\beta(y) \to x < y)),$$

is true in $\mathfrak{A}$. Conversely, if (B3) *is true in $\mathfrak{A}$, then there exists a factorization $\mathbb{B}$ of $\mathfrak{A}$ in which every block of type α lies below every block of type β in the block order.*

P r o o f . The first statement of the lemma is obvious. For the converse, let $\mathbf{B}$ consist of the non-empty sets $\{a \in X \mid \mathrm{tp}(a) = \pi\}$ for $\pi \in \Pi$. If there is no α-block or no β-block, let $\ll$ be the empty partial order. Otherwise, let $A \in \mathbf{B}$ be the α-block, let $B \in \mathbf{B}$ be the β-block, and let $\ll = \{\langle A, B\rangle\}$. □

If the typed partial order $\mathfrak{A}$ is clear from context, and $\mathbb{B}$ is a factorization of $\mathfrak{A}$, we write $\mathbb{B} \models \forall x(\alpha(x) \to \forall y(\beta(y) \to x < y))$ to mean that every block of type α is less than every block of type β in the block order. The motivation for this notation should be obvious from Lemma 3.2.

Lemma 3.3 *Let $\mathfrak{A}$ be a typed partial order over a set of types Π, with $\alpha, \beta \in \Pi$ distinct. Suppose $\mathbb{B}$ is a factorization of $\mathfrak{A}$ in which the set of blocks of type $\alpha \vee \beta$ is linearly ordered. Then the formula* (B5b)*, namely*

$$\forall x(\alpha(x) \to \forall y(\beta(y) \to (x < y \vee y < x))),$$

is true in $\mathfrak{A}$. Conversely, if (B5b) *is true in $\mathfrak{A}$, then there exists a factorization $\mathbb{B}$ of $\mathfrak{A}$ such that the set of blocks of type $\alpha \vee \beta$ is linearly ordered.*

P r o o f . The first statement of the lemma is obvious. For the converse, let $A_0 = \{a \in X \mid \mathrm{tp}(a) = \alpha\}$ and $B_0 = \{a \in X \mid \mathrm{tp}(a) = \beta\}$. We may assume that both these sets are non-empty, since otherwise the trivial factorization satisfies the conditions of the lemma. Define an equivalence relation $\equiv$ on A_0 by setting $a \equiv a'$ if, for all $b \in B_0$, $a < b \Leftrightarrow a' < b$. Similarly, define an equivalence relation $\equiv$ on B_0 by setting $b \equiv b'$ if, for all $a \in A_0$, $b' < a \Leftrightarrow b' < a$. Let $\mathbf{B}$ be the partition of X whose cells are: (i) the equivalence classes of $\equiv$ in A_0, (ii) the equivalence classes of $\equiv$ in B_0, and (iii) the non-empty sets $\{a \in X \mid \mathrm{tp}(a) = \pi\}$, where $\pi \in \Pi\backslash\{\alpha, \beta\}$. For any $C, D \in \mathbf{B}$ write $C \ll D$ just in case C and D are distinct $(\alpha \vee \beta)$-blocks such that there exist $c \in C$ and $d \in D$ with $c < d$.

To show that $\mathbb{B} = (\mathbf{B}, \ll)$ has the desired properties, we first observe that, if A is an α-block and B a β-block, then $A \ll B$ if and only if, for all $a \in A$ and $b \in B$, $a < b$. Now suppose that A and A' are distinct α-blocks, and pick $a_0 \in A$ and $a_1 \in A'$. From the definition of $\equiv$ on A_0 and the fact that $A_0 \cup B_0$ is linearly ordered by $<$, let $b_0 \in B_0$ be such that either $a_0 < b_0 < a_1$ or $a_1 < b_0 < a_0$. In the former case, again using the definition of $\equiv$ on A_0, we have $a < a'$ for all $a \in A$ and $a' \in A'$; and in the latter, $a > a'$ for all $a \in A$ and $a' \in A'$. Similar remarks apply to β-blocks. Thus, for any $(\alpha \vee \beta)$-blocks C and D, $C \ll D$ if and only if, for all $c \in C$ and $d \in D$, $c < d$. This relation is obviously transitive and irreflexive; that is, $\ll$ is a partial order. Moreover, the the collection of $(\alpha \vee \beta)$-blocks is linearly ordered by $\ll$, because $A_0 \cup B_0$ is linearly ordered by $<$. In particular, the collection of α-blocks and the collection of β-blocks are also both linearly ordered by $\ll$; and, for $\gamma \in \Pi\backslash\{\alpha, \beta\}$, there is at most one γ-block. Hence, $(\mathbf{B}, \ll)$ is a factorization of $\mathfrak{A}$. □

If the typed partial order $\mathfrak{A}$ is clear from context, and $\mathbb{B}$ is a factorization of $\mathfrak{A}$, we write $\mathbb{B} \models \forall x(\alpha(x) \to \forall y(\beta(y) \to (x < y \vee x > y))$ to mean that the set of blocks of type $\alpha \vee \beta$ is linearly ordered. The motivation for this notation should be obvious from Lemma 3.3.

Suppose $\mathfrak{A}$ is a typed partial order and $\mathbb{B}_1 = (\mathbf{B}_1, \ll_1)$, $\mathbb{B}_2 = (\mathbf{B}_2, \ll_2)$ are factorizations of $\mathfrak{A}$. We say that $\mathbb{B}_2$ is a *refinement* of $\mathbb{B}_1$ if, for all $A_2 \in \mathbf{B}_2$, there exists a (necessarily unique) $A_1 \in \mathbf{B}_1$ such that $A_2 \subseteq A_1$, and moreover, for all $A_2, B_2 \in \mathbf{B}_2$, and all $A_1, B_1 \in \mathbf{B}_1$ such that $A_2 \subseteq A_1$ and $B_2 \subseteq B_1$, $A_1 \ll_1 B_1$ implies $A_2 \ll_2 B_2$.

Lemma 3.4 *Any two factorizations of a typed partial order have a common refinement.*

P r o o f . Let $\mathbb{B}_1 = (\mathbf{B}_1, \ll_1)$ and $\mathbb{B}_2 = (\mathbf{B}_2, \ll_2)$ be factorizations of the typed partial order $\mathfrak{A}$. Define $\mathbb{B} = (\mathbf{B}, \ll)$ as follows: let

$$\mathbf{B} = \{B_1 \cap B_2 \mid B_1 \in \mathbf{B}_1, B_2 \in \mathbf{B}_2\}\backslash\{\varnothing\};$$

and let $\ll$ be the transitive closure of the relation

$$\{\langle A, B\rangle \mid A = A_1 \cap A_2, B = B_1 \cap B_2, A_1 \ll_1 B_1 \text{ or } A_2 \ll_2 B_2\}.$$

A simple induction shows that, if A is related to B by $\ll$ then, for all $a \in A$ and all $b \in B$, $a < b$. It follows that $\ll$ is irreflexive and hence is a partial order. It is thus immediate from the definition of $\ll$ that the partially ordered set $\mathbb{B} = (\mathbf{B}, \ll)$ is a factorization of $\mathfrak{A}$ and moreover that it is a refinement of both $\mathbb{B}_1$ and $\mathbb{B}_2$. □

Refinements of block orders are useful because they preserve the properties featured in Lemmas 3.2 and 3.3. The following Lemma is immediate.

Lemma 3.5 *Let $\mathfrak{A}$ be a typed partial order and $\mathbb{B}$, $\mathbb{B}'$ factorizations of $\mathfrak{A}$ with $\mathbb{B}'$ a refinement of $\mathbb{B}$. Let ψ be a factor-controllable basic formula. If $\mathbb{B} \models \psi$, then $\mathbb{B}' \models \psi$.*

A *unit block* of $\mathbb{B}$ is a block containing exactly one element of X. Trivially, every unit block is linearly ordered by $<$. We say that $\mathbb{B}$ is *unitary* if every block of $\mathbf{B}$ which is linearly ordered by $<$ is a unit block. Combining all of the above lemmas, we have:

Lemma 3.6 *Let $\mathfrak{A}$ be a typed partial order and Ψ a finite set of basic formulas such that $\mathfrak{A} \models \Psi$. Then there is a unitary factorization $\mathbb{B}$ of $\mathfrak{A}$ such that $\mathbb{B} \models \mathrm{FC}(\Psi)$.*

P r o o f . For each $\psi \in \mathrm{FC}(\Psi)$, we apply Lemmas 3.2 or 3.3 as appropriate, and take a common refinement of all the resulting factorizations by Lemma 3.4. Now further refine by replacing all linearly ordered blocks with unit blocks having the obvious block order. The result then follows by Lemma 3.5. □

Let $\mathfrak{A} = (X, <, \mathrm{tp})$ be a typed partial order and $\mathbb{B} = (\mathbf{B}, \ll)$ a factorization of $\mathfrak{A}$. We have already observed that $\mathbb{B}$ does not contain all the information required to reconstruct the element order $<$. However, it nearly does, in a sense that we can make precise. Let us first overload the block-order $\ll$ by writing, for all $a, b \in X$, $a \ll b$ if there exist $A, B \in \mathbf{B}$ such that $a \in A$, $b \in B$ and $A \ll B$. We might call $\ll$ the *inter-block order* on X. It is obvious that the inter-block order is a partial order, and, from (F3), that it is contained in the element order $<$. Now, for all $a, b \in X$, write $a <_0 b$ if $a < b$, and both a and b belong to the same block of $\mathbf{B}$. Again, this is clearly a partial order: we call it the *intra-block order*. Finally, define $a <_\times b$ if $a < b$, and both a and b are extremal elements of $(X, <)$; once again, $<_\times$ is clearly a partial order: we call it the *extremal order*. Now define the binary relation $\lessdot$ on X to be the transitive closure of $(\ll \cup <_0 \cup <_\times)$. It is obvious that $\lessdot$ is a partial order no stronger than (i.e., included in) $<$, but that, nevertheless, $\mathbb{B}$ is a factorization of the typed partial order $(X, \lessdot, \mathrm{tp})$. It is also obvious that, when restricted to elements of some fixed 1-type π, $<$ and $\lessdot$ coincide. We say that $\mathfrak{A}$ is *thin* over $\mathbb{B}$ if $<$ and $\lessdot$ coincide over the whole of X.

Lemma 3.7 *Suppose $\mathfrak{A} = (X, <, \mathrm{tp})$ is a finite typed partial order and Ψ a finite set of basic formulas such that $\mathfrak{A} \models \Psi$. Let $\mathbb{B}$ be a factorization of $\mathfrak{A}$ such that $\mathbb{B} \models \mathrm{FC}(\Psi)$. Then there exists a typed partial order $\dot{\mathfrak{A}}$ over the domain X such that $\mathbb{B}$ is a factorization of $\dot{\mathfrak{A}}$, $\dot{\mathfrak{A}} \models \Psi$, and $\dot{\mathfrak{A}}$ is thin over $\mathbb{B}$.*

P r o o f . Define $\lessdot$ to be the transitive closure of $(\ll \cup <_0 \cup <_\times)$, as just described, and let $\dot{\mathfrak{A}} = (X, \lessdot, \mathrm{tp})$. Thus, $\mathbb{B}$ is a factorization of $\dot{\mathfrak{A}}$, with $\dot{\mathfrak{A}}$ thin over $\mathbb{B}$. We show that $\dot{\mathfrak{A}} \models \psi$, where $\psi \in \Psi$ is of each of the possible forms (B1a)–(B10) in turn.

(B1a), (B1b), (B9), (B10): ψ does not involve the ordering.
(B2a), (B2b), (B4), (B8): $\lessdot$ is no stronger than $<$.
(B3), (B5b): $\mathbb{B} \models \psi$.
(B5a): When restricted to elements of some fixed type, $<$ and $\lessdot$ coincide.
(B6): Suppose that $a \in X$ is of type α. Since $<$ and $\lessdot$ coincide on elements of some fixed type, let a^* be a maximal element of type α such that either $a = a^*$ or $a < a^*$ (equivalently: $a = a^*$ or $a \lessdot a^*$). But $\mathfrak{A} \models \psi$, so there exists an element b satisfying μ—say of type $\beta \neq \alpha$—such that $a^* < b$, and hence a maximal element b^* of type β such that $a^* <_\times b^*$. But then $a \lessdot b^*$, whence $\dot{\mathfrak{A}} \models \psi$.
(B7): Similar to (B6). □

3.3 Reducing the number of blocks

Suppose $\mathfrak{A} = (X, <, \mathrm{tp})$ is a finite typed partial order with factorization $\mathbb{B} = (\mathbf{B}, \ll)$. The following notions will help us to reason about $\mathbb{B}$. Recall that $\mathbf{B}^\times$ denotes the set of extremal blocks of $\mathbf{B}$. If $B \in \mathbf{B}$, define the *depth* of B, denoted $d(B)$, to be the length m of the longest path $B = B_0 \ll \cdots \ll B_m$. The *depth* of $\mathbb{B}$, denoted $d(\mathbb{B})$, is the maximum value attained by $d(B)$ for $B \in \mathbf{B}$. A *cut* is a number $\chi = i + 0.5$ where $0 \leq i < d(\mathbb{B})$. If χ and χ' are cuts, we say χ' is *above* χ (and χ is *below* χ') if $\chi' < \chi$. (Depth increases as we go down.) Similarly, if $B \in \mathbf{B}$, we say that B is *above* χ if $d(B) < \chi$, and *below* χ if $d(B) > \chi$. If χ' is also a cut of $\mathbb{B}$ with χ below χ', we say that B is *between* χ and χ' if it is above χ and below χ'.

For any cut χ, and any 1-type π, a *minimal* π-block *above* χ is a block B such that $\mathrm{tp}(B) = \pi$, $d(B) < \chi$ and, for all $B' \in \mathbf{B}$ such that $\mathrm{tp}(B') = \pi$ and $d(B') < \chi$, $d(B) \geq d(B')$. A *minimal* block *above* χ is a minimal π-block above χ for some π. The notion of *maximal* (π)-block *below* χ is defined analogously. Denote by $F^+(\chi)$ the set of *minimal* blocks above χ, and by $F^-(\chi)$ the set of *maximal* blocks below χ. Note that $F^+(\chi)$ contains at most one block of each type, and similarly for $F^-(\chi)$. Let $F(\chi) = F^-(\chi) \cup F^+(\chi) \cup \mathbf{B}^\times$; we call $F(\chi)$ the *frontier* of χ. Figure 2 shows a cut $\chi = 4.5$ in a factorization of a typed partial order over $\pi_1, \ldots, \pi_N$. The sets of blocks $F^+(\chi)$ and $F^-(\chi)$ are shown by shading.

If χ and χ' are cuts of $\mathbb{B}$, with χ below χ', we say that χ and χ' are *equivalent* if there exists a function $f : F(\chi) \to F(\chi')$ satisfying the following conditions:

(E1) f maps $F^-(\chi)$ to $F^-(\chi')$, f maps $F^+(\chi)$ to $F^+(\chi')$, and f is the identity on $\mathbf{B}^\times$;
(E2) $f : F(\chi) \to F(\chi')$ is a typed partial order isomorphism, i.e., f is 1–1 and onto, for all $B \in F(\chi)$, $\mathrm{tp}(B) = \mathrm{tp}(f(B))$, and for all $A, B \in F(\chi)$, $A \ll B \Leftrightarrow f(A) \ll f(B)$.

Suppose χ and χ' are equivalent cuts, with χ' above χ. Obviously, no extremal block can lie between χ and χ'. For example, if A is minimal, then $f(A) = A$, $A \in F^+(\chi)$ but $A \notin F^+(\chi')$, contradicting the requirements of (E1); a similar argument applies if A is maximal. Equally obviously, $F^+(\chi')$ and $F^-(\chi')$ each contain at most one block of any given type, by (F2), so that the function f, if it exists, is unique by the requirements of (E2); we denote this function by $f_{\chi,\chi'}$. Observe finally that, for any block B in $F^-(\chi) \cup F^+(\chi)$, the blocks B and $f_{\chi,\chi'}(B)$ stand in the same relations ($\ll$, $\gg$ or $=$) to all extremal blocks.

Fixing $\mathfrak{A} = (X, <, \mathrm{tp})$ and $\mathbb{B} = (\mathbf{B}, \ll)$, suppose χ, χ' are equivalent cuts in $\mathbb{B}$ with χ below χ'. Let $\mathbf{B}^- = \{B \in \mathbf{B} \mid d(B) > \chi\}$ be the set of blocks below χ and $\mathbf{B}^+ = \{B \in \mathbf{B} \mid d(B) < \chi'\}$ the set of blocks above χ'. Define $\mathbf{B}^* = \mathbf{B}^- \cup \mathbf{B}^+$, and define the relation $\ll^*$ on $\mathbf{B}^*$ to be the transitive closure of the union of the three relations

$$\{\langle A, B\rangle \in (\mathbf{B}^-)^2 \mid A \ll B\}$$

$$\{\langle A, B\rangle \in (\mathbf{B}^+)^2 \mid A \ll B\}$$

$$\{\langle B, f_{\chi,\chi'}(C)\rangle \mid B \in F^-(\chi), C \in F^+(\chi), B \ll C\}.$$

Denote by $\mathbb{B}^*$ the pair $(\mathbf{B}^*, \ll^*)$. The intuition behind the definition of $\ll^*$ is as follows. Below χ and above χ', $\ll^*$ coincides with $\ll$; for $\ll^*$-edges from blocks below χ to blocks above χ', we take those $\ll$-edges from $F^-(\chi)$ to $F^+(\chi)$, re-direct their targets to the corresponding blocks above χ', and then take the transitive closure. As an example, consider the configuration of blocks in Fig. 3, where $\ll$-edges are indicated by solid arrows and the function $f_{\chi,\chi'}$ by dotted arrows; here we obtain $A \ll^* C \ll^* D' \ll^* B'$.

Let $X^* = \bigcup \mathbf{B}^*$, and let tp^* be the restriction of the function tp to X^*. Noting that $\mathbf{B}^\times \subseteq \mathbf{B}^*$, we see that the extremal order $<_\times$ is defined on X^*. Let $<_0^*$ be the restriction of the intra-block order $<_0$ to X^*. As before, we overload the symbol $\ll^*$ so that it denotes the inter-block order on X^* under $\mathbb{B}^*$: $a \ll^* b$ if the blocks $A, B \in \mathbf{B}^*$ such that $a \in A$, $b \in B$ satisfy $A \ll^* B$. Finally, let $<^*$ be the transitive closure of $(\ll^* \cup <_0^* \cup <_\times)$. We denote by $\mathfrak{A}^*$ the triple $(X^*, <^*, \mathrm{tp}^*)$.

We now prove a sequence of lemmas concerning $\mathfrak{A}^*$ and $\mathbb{B}^*$. The most salient of these are Lemma 3.10, which states that $\mathfrak{A}^*$ is a typed partial order with $\mathbb{B}^*$ a factorization of $\mathfrak{A}^*$, and Lemma 3.13, which states that witnesses for basic formulas of the form (B8) can still be found in $\mathfrak{A}^*$. We then prove Lemma 3.14, which states that blocks between equivalent cuts can always be removed without compromising the truth of basic formulas. It is worth reflecting briefly on the relationship between $\mathfrak{A}$ and $\mathfrak{A}^*$ at this point. In particular, it is important to

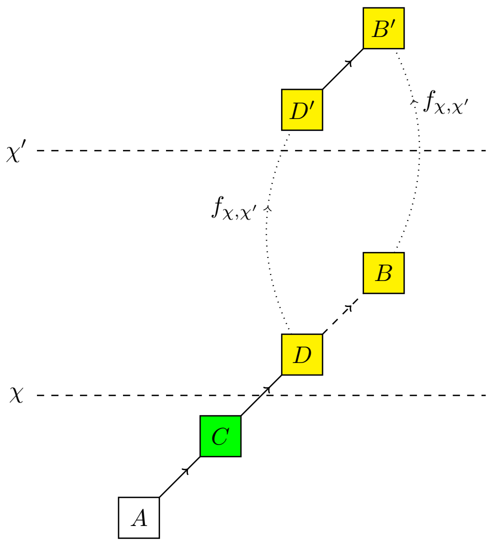


Fig. 3 Proof of Lemma 3.13 in the case where A is below χ: if $A \ll B'$, then $A \ll B$.

observe that the partial order $<^*$ is in general strictly weaker than the restriction of $<$ to the elements above χ' and below χ. This is obvious if we consider the case where $a < b$, a lies in a block A below χ, b lies in a block B above χ', and $A \not\ll B$. Unless a and b are extremal, we are not in general guaranteed that $a <^* b$. Slightly less obviously, corresponding remarks apply to the block orderings. Consider a non-extremal α-block A' below χ and a non-extremal β-block B above χ', where $A' \ll B'$. Let A^* be the maximal α-block below χ, and B^* be the minimal β-block above χ (so that $f_{\chi,\chi'}(B^*)$ is the minimal β-block above χ'). Unless $A^* \ll B^*$, we are not in general guaranteed that $A' \ll^* B'$. This thinning out of the element- and block-orderings turns out, however, not to compromise the truth of any basic formulas. Indeed, it is essential: without it, we should be in danger of losing witnesses for basic formulas of the form (B8), as guaranteed by Lemma 3.13.

In Lemmas 3.8 to 3.13, we keep $\mathfrak{A}$, $\mathbb{B}$, χ and χ' fixed, with $\mathbb{B}^*$ and $\mathfrak{A}^*$ as defined above.

Lemma 3.8 *For all $A, B \in \mathbf{B}^*$, $A \ll^* B$ implies $A \ll B$. In addition, for any 1-type α, the α-blocks of $\mathbf{B}^*$ are linearly ordered by $\ll^*$.*

P r o o f . For the first assertion, observe that, by (F2), the set of pairs $\{\langle B, f_{\chi,\chi'}(C)\rangle \mid B \in F^-(\chi), C \in F^+(\chi), B \ll C\}$ is included in $\ll$. Thus, $\ll^*$ is the transitive closure of a relation included in $\ll$, whence $\ll^*$ is included in $\ll$. For the second assertion, observe first that, if there are no α-blocks below χ or above χ', the result is immediate; hence we may assume otherwise. It suffices to show that, if $B \in F^-(\chi)$ is the maximal α-block in $\mathbb{B}$ below χ and A the minimal α-block above χ', then $B \ll^* A$. Let $C \in F^+(\chi)$ be the minimal α-block above χ, so that $B \ll C$. But then $B \ll^* f_{\chi,\chi'}(C) = A$, and we are done. □

Lemma 3.9 *For all $a, b \in X^*$, $a <^* b$ implies $a < b$. If, in addition, $\mathrm{tp}^*(a) = \mathrm{tp}^*(b)$, the converse implication holds.*

P r o o f . The first assertion is immediate from Lemma 3.8: $<^*$ is the transitive closure of a relation included in $<$. For the second assertion, suppose that $a < b$, and that a and b belong to the respective blocks A and B. If $A = B$, the result follows from the fact that $<^*$ extends the intra-block order $<^*_0$. Otherwise, we have $A \ll B$, and hence $B \not\ll A$. By Lemma 3.8, $A \ll^* B$ and so $a <^* b$ by the fact that $<^*$ extends the inter-block order $\ll^*$. □

Lemma 3.10 *$\mathfrak{A}^*$ is a typed partial order, and $\mathbb{B}^*$ is a factorization of $\mathfrak{A}^*$. Moreover, $\mathfrak{A}^*$ is thin over $\mathbb{B}^*$.*

P r o o f . By the first assertions of Lemmas 3.8 and 3.9, both $\ll^*$ and $<^*$ are partial orders. By construction, for all $B \in \mathbf{B}^*$, every element $b \in B$ satisfies $\mathrm{tp}^*(b) = \mathrm{tp}(b) = \mathrm{tp}(B)$. Thus, we can write $\mathrm{tp}^*(B) = \mathrm{tp}(B)$ to denote the 1-type of B in $\mathbb{B}^*$. The second assertion of Lemma 3.8 ensures that all blocks of any fixed 1-type in $\mathbb{B}^*$ are linearly ordered. Finally, the requirement that $A \ll^* B$ implies $a <^* b$ for all $a \in A$, $b \in B$ is secured by the fact that $<^*$ extends the inter-block ordering on X^*. The final statement of the lemma is immediate from the definition of $<^*$. □

Lemma 3.11 *If all blocks of type α are smaller than all blocks of type β in the block ordering $\mathbb{B}$, then the same is true of the block ordering $\mathbb{B}^*$.*

P r o o f . We may assume that there exist α-blocks and β-blocks, for otherwise the lemma is trivial. Let A be the largest α-block in $\mathbf{B}$ and B the smallest β-block, so that $A \ll B$. Since A and B are extremal, they do not lie between χ and χ', whence by Lemma 3.8, A is the largest α-block in $\mathbf{B}^*$ and B the smallest β-block. Hence it suffices to show that $A \ll^* B$. If A and B both lie below χ, this is immediate from the definition of $\ll^*$; similarly if both lie above χ'. If A lies below χ and B above χ' then the same conclusion follows from the fact that $A \in F^-(\chi)$, $B \in F^+(\chi)$ and $f_{\chi,\chi'}(B) = B$. □

Lemma 3.12 *If the blocks of type $(\alpha \vee \beta)$ are linearly ordered in $\mathbb{B}$, then the blocks of this type are linearly ordered in $\mathbb{B}^*$.*

P r o o f . It suffices to show that, if A is the maximal $(\alpha \vee \beta)$-block below χ, and B' the minimal $(\alpha \vee \beta)$ block above χ', then $A \ll^* B'$. Let B be the minimal $(\alpha \vee \beta)$-block above χ. Since χ and χ' are equivalent, B and B' must be of the same type, so $f_{\chi,\chi'}(B) = B'$. But $A \ll B$, whence $A \ll^* B'$, as required. □

Lemma 3.13 *Suppose $a \in X^*$ and $b \in X$ are such that $a \sim b$. Then there exists $b' \in X^*$ such that $a \sim^* b'$ and* $\mathrm{tp}(b') = \mathrm{tp}(b)$.

P r o o f . Let A be the block of $\mathbb{B}$ containing a and B_0 the block containing b. If $B_0 \in \mathbf{B}^*$, the result follows immediately from Lemma 3.9 by setting $b' = b$. So we may suppose B_0 lies between χ and χ'.

Assume, for definiteness, that A lies below χ. Let $\alpha = \mathrm{tp}(a) = \mathrm{tp}(A)$, let $\beta = \mathrm{tp}(b) = \mathrm{tp}(B_0)$, and let B be the minimal β-block above χ. We remark that $\beta \neq \alpha$: for if a and b are of the same 1-type, then they must be in the same block of $\mathbb{B}$, contradicting the supposition that B_0 is not in $\mathbf{B}^*$. We claim that $A \approx B$. (That is: $A \not\ll B$ and $B \not\ll A$.) For either $B = B_0$ or $B \ll B_0$, and by (F3), $A \approx B_0$, so that $A \not\ll B$; on the other hand no block B above χ satisfies $B \ll A$, which proves the claim. Now let B' be the minimal β-block above χ', so that $f_{\chi,\chi'}(B) = B'$. Since no extremal block can lie between the equivalent cuts χ and χ', B is not extremal; hence B' is not extremal, by (E1) and the fact that $f_{\chi,\chi'}$ is injective. We claim that $A \approx^* B'$. Certainly, $B' \not\ll^* A$, so it suffices to suppose $A \ll^* B'$, and derive a contradiction. We write $X \ll_= Y$ to mean either $X \ll Y$ or $X = Y$. By the construction of $\ll^*$, there exist C, D, $D' = f_{\chi,\chi'}(D)$ such that C is a maximal block below χ, D is a minimal block above χ, and such that $A \ll_= C$, $C \ll D$ and $D' \ll_= B'$. Since $B, D \in F^+(\chi)$ and $B', D' \in F^+(\chi')$, it follows from (E3) that $D \ll_= B$. But then $A \ll B$, which is the desired contradiction (cf. Fig. 3).

Recall now that B' is not extremal, and that $a \in A$, $b \in B_0$ with $a \sim b$. Pick any $b' \in B'$. From $a \sim b$ and $B \ll B'$, we know that $b' \not< a$, whence, by Lemma 3.9, $b' \not<^* a$. So suppose, for contradiction, that $a <^* b'$. By the definition of $<^*$, there exists a sequence $a = a_0, \ldots, a_m = b'$ such that, for all i $(0 \leq i < m)$, the pair $\langle a_i, a_{i+1}\rangle$ is in $(\ll^* \cup <_0^* \cup <_\times)$. But we already know that $A \approx^* B'$, so there must be some ℓ $(0 \leq \ell < m)$ such that $a_\ell <_\times a_{\ell+1}$. Take the largest such value of ℓ, and let C be the block containing $a_{\ell+1}$. Thus, C is extremal, and indeed, since B' is non-extremal, we have $\ell < m - 1$, and $C \ll^* B'$, whence $C \ll B'$ by Lemma 3.8. By the fact that $B' = f_{\chi,\chi}(B)$, and C is extremal, we have $C \ll B$ and hence $C \ll B_0$. On the other hand, $a < a_{\ell+1} \in C$, contradicting the fact that $a \sim b$. Thus, $a \sim^* b'$, as required. The case where A lies above χ' proceeds similarly. □

Let us summarize. We started by taking any partial order $\mathfrak{A} = (X, <\mathrm{tp})$ with factorization $\mathbb{B} = (\mathbf{B}, \ll)$. We supposed that there existed equivalent cuts χ, χ' of $\mathbb{B}$, with χ below χ'. We then constructed a new partial order $\mathfrak{A}^* = (X^*, <^* \mathrm{tp}^*)$ with factorization $\mathbb{B}^* = (\mathbf{B}^*, \ll^*)$, as established by Lemma 3.10. Let us write $\mathfrak{A}/(\chi, \chi')$ for $\mathfrak{A}^*$ and $\mathbb{B}/(\chi, \chi')$ for $\mathbb{B}^*$. Notice that the size of $\mathbb{B}/(\chi, \chi')$—i.e., the number of blocks it contains—is strictly smaller than that of $\mathbb{B}$.

Lemma 3.14 *Let Ψ be a finite set of basic formulas, and suppose $\mathfrak{A}$ is a typed partial order such that $\mathfrak{A} \models \Psi$. Let $\mathbb{B}$ be a factorization of $\mathfrak{A}$ such that $\mathbb{B} \models \mathrm{FC}(\Psi)$, and suppose χ, χ' are equivalent cuts in $\mathbb{B}$. Then $\mathfrak{A}/(\chi, \chi') \models \Psi$, $\mathbb{B}/(\chi, \chi') \models \mathrm{FC}(\Psi)$ and $\mathfrak{A}/(\chi, \chi')$ is thin over $\mathbb{B}/(\chi, \chi')$. Moreover, if $\mathbb{B}$ is unitary, then so is $\mathbb{B}/(\chi, \chi')$.*

P r o o f . Write $\mathfrak{A} = (X, <, \mathrm{tp})$ and $\mathfrak{A}/(\chi, \chi') = (X^*, <^*, \mathrm{tp}^*)$. We consider the various basic forms in turn.

(B1a), (B1b): $X^* \subseteq X$.
(B2a), (B2b), (B4): $X^* \subseteq X$ and, by Lemma 3.9, $<^*$ is no stronger than $<$.
(B3): By Lemma 3.11, $\mathbb{B}/(\chi, \chi') \models \psi$; now apply Lemma 3.2.
(B5a): By Lemma 3.9, $<^*$ coincides with $<$ on $(\mathrm{tp}^*)^{-1}(\alpha)$.
(B5b): By Lemma 3.12, $\mathbb{B}/(\chi, \chi') \models \psi$; now apply Lemma 3.3.
(B6): Pick any $a \in X^*$ such that $\mathrm{tp}^*(a) = \alpha$. Let a^* be a maximal α-element of X such that $a < a^*$. Since $\mathfrak{A} \models \psi$, let $b \in X$ be such that $\mathfrak{A} \models \mu[b]$, $\mathrm{tp}[b] \neq \alpha$ and $a^* < b$. Without loss of generality, we may assume that b is a maximal element of its 1-type in $\mathfrak{A}$. Since a^* and b are maximal elements, we have $a^*, b \in X^*$, and indeed $a^* <^* b$. Finally, by the second statement of Lemma 3.9, $a <^* a^*$, whence $a <^* b$, whence $\mathfrak{A}/(\chi, \chi') \models \psi$.
(B7): Proceed symmetrically to the case (B6).
(B8): By Lemma 3.13.
(B9): $X^* \subseteq X$.
(B10): All extremal blocks of $\mathbb{B}$ are blocks of $\mathbb{B}/(\chi, \chi')$.

Lemmas 3.11 and 3.12 ensure that $\mathbb{B}(\chi, \chi') \models \mathrm{FC}(\Psi)$. The remaining statements of the lemma are obvious. □

Lemma 3.15 *Suppose Ψ is a finitely satisfiable finite set of basic formulas. Then there is a finite model $\mathfrak{A} \models \Psi$ with unitary factorization $\mathbb{B}$ of size bounded by a doubly exponential function of the size of the signature of Ψ, such that $\mathfrak{A}$ is thin over $\mathbb{B}$.*

P r o o f . Suppose $\mathfrak{A}_0$ is a finite typed partial order such that $\mathfrak{A}_0 \models \Psi$. By Lemma 3.6, let $\mathbb{B}_0$ be a factorization of $\mathfrak{A}_0$ such that $\mathbb{B}_0 \models \mathrm{FC}(\Psi)$. By Lemma 3.7, we may assume that $\mathbb{B}_0$ is unitary and that $\mathfrak{A}_0$ is thin over $\mathbb{B}_0$. Assuming $\mathfrak{A}_i$ and $\mathbb{B}_i$ have been defined, if $\mathbb{B}_i$ contains a pair of equivalent cuts, χ and χ', let $\mathfrak{A}_{i+1} = \mathfrak{A}_i/(\chi, \chi')$ and $\mathbb{B}_{i+1} = \mathbb{B}_i/(\chi, \chi')$. By Lemma 3.14, $\mathfrak{A}_{i+1} \models \Psi$ and $\mathbb{B}_{i+1} \models \mathrm{FC}(\Psi)$; moreover, $\mathbb{B}_{i+1}$ is unitary, and $\mathfrak{A}_{i+1}$ is thin over $\mathbb{B}_{i+1}$. Since the number of blocks in $\mathbb{B}_i$ is strictly decreasing, we eventually reach a structure $\mathfrak{A}_m$ with factorization $\mathbb{B}_m$, in which no two cuts are equivalent. Since the frontier of any cut is at most exponential in size, there can be at most doubly exponentially many cuts in $\mathbb{B}_m$, and hence at most doubly exponentially many blocks in $\mathbb{B}_m$. This proves the lemma. □

3.4 Reducing the size of blocks

With Lemma 3.15, we have established that, if a collection Ψ of basic formulas has a finite model, then it has a finite model $\mathfrak{A}$ with a small factorization $\mathbb{B}$ (i.e., a factorization having a small number of blocks), such that $\mathbb{B}$ guarantees the truth of all factor-controllable members of Ψ, and $\mathfrak{A}$ is thin over $\mathbb{B}$. However, while the *number* of the blocks in $\mathbb{B}$ was bounded by a doubly exponential function of the size of the signature of Ψ, nothing was said about the *size* of these blocks. In this section we show that the blocks themselves can bounded in size.

Fix some finite typed partial order $\mathfrak{A} = (X, <, \mathrm{tp})$ with unitary factorization $\mathbb{B} = (\mathbf{B}, \ll)$, such that $\mathfrak{A}$ is thin over $\mathbb{B}$. Let us suppose that, for some finite set Ψ of basic formulas, $\mathfrak{A} \models \Psi$ and $\mathbb{B} \models \mathrm{FC}(\Psi)$. Our strategy in the sequel will be to divide up the blocks of $\mathbf{B}$ into *sub-blocks*, and then to replace each sub-block by a set of either one or two elements, imposing a partial order on these elements that secures satisfaction of Ψ. Unfortunately, reducing the size of each block is not a simple matter of removing elements: the elements in the reduced block do not clearly represent any particular elements in the original blocks. Thus, we have to work to ensure that, in providing witnesses for elements in these reduced blocks, we do not create connections between elements in previously unrelated blocks, and in particular do not create any cycles in the partial order we are trying to define. A sub-block will be replaced by a singleton if the block that includes it is itself is a unit block; otherwise, it will be replaced by a pair of incomparable elements. The assumption that $\mathfrak{A}$ is thin over $\mathbb{B}$ underpins an inductive argument in Lemma 3.16 crucial in showing that the order we eventually define contains no cycles. The assumption that $\mathbb{B}$ is unitary rules out the possibility that some sub-block is made to contain a pair of incomparable elements when the including block is required to be linearly ordered—in particular, if Ψ contains a basic formula of type (B5a).

For $a \in X$, and $B \in \mathbf{B}$, we say that B is *below* a if there exists $b \in B$ such that $b < a$; similarly, we say that B is *above* a if there exists $b \in B$ such that $a < b$. Notice that it is possible for B to be both above and below a. We define the *sub-type* of a to be the triple $\langle \mathbf{B}^-, B, \mathbf{B}^+ \rangle$, where B is the block containing a, $\mathbf{B}^-$ is the set

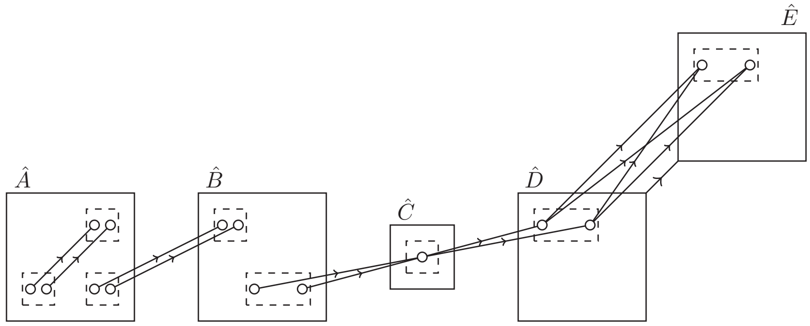


Fig. 4 Schematic depiction of the relation $\prec$ on $\hat{X}$. Elements $\hat{A}, \ldots \hat{E}$ of $\hat{\mathbf{B}}$—corresponding to blocks $A, \ldots, E$ of $\mathbf{B}$—are indicated by solid boxes; their decomposition into sets $\hat{s}$ (where s is a sub-block of one of $A, \ldots, E$) is indicated by dashed boxes. Block C is a unit block; D is smaller than E in the block ordering on $\mathbf{B}$, and other blocks are unrelated.

of blocks below a and $\mathbf{B}^+$ the set of blocks above a. A *sub-block* is a maximal set s of elements all having the same sub-type. We write $\mathrm{tp}(s) = \mathrm{tp}(B)$. Thus, each block is partitioned into a finite number of sub-blocks, and all elements of a sub-block s have 1-type $\mathrm{tp}(s)$. If s and t are sub-blocks contained in the respective blocks A and B, and $A \ll B$, we write $s \ll t$.

Lemma 3.16 *Let $a, b \in X$ with $a < b$. Let the sub-type of a be $s = \langle \mathbf{A}^-, A, \mathbf{A}^+ \rangle$ and the sub-type of b be $t = \langle \mathbf{B}^-, B, \mathbf{B}^+ \rangle$, and suppose $s \neq t$. Then (i) $\mathbf{A}^- \subseteq \mathbf{B}^-$; (ii) $\mathbf{A}^+ \supseteq \mathbf{B}^+$; and (iii) $\{A\} \cup \mathbf{A}^+ \supseteq \{B\} \cup \mathbf{B}^+$. Furthermore, at least one of these inclusions is strict.*

Proof. The inclusions themselves are immediate. To show strictness, therefore, since $\mathfrak{A}$ is thin over $\mathbb{B}$, it evidently suffices to prove the result in the cases where one of $a <_0 b$, $a \ll b$ or $a <_\times b$ holds. If $a <_0 b$, then $A = B$, and strictness of either (i) or (ii) follows from the assumption $s \neq t$. If $a \ll b$, then $A \ll B$, and hence $A \notin \mathbf{B}^+$, since there certainly cannot exist $a' \in A$ with $b < a'$. Hence, inclusion (iii) is strict. Suppose, then that $a <_\times b$, so that $A \neq B$, and a is either a maximal α-element or a minimal α-element, where $\alpha = \mathrm{tp}(a)$. If the former, then, since $a < b$, we again have $A \notin \mathbf{B}^+$, so that inclusion (iii) is strict. If the latter, then $A \notin \mathbf{A}^-$, so inclusion (i) is strict. $\square$

For every sub-block s, let $\hat{s} = \{\hat{s}(0), \hat{s}(1)\}$, where $\hat{s}(0)$ and $\hat{s}(1)$ are some objects. If s is contained in (and hence is equal to) a unit block, we set $\hat{s}(0) = \hat{s}(1)$; otherwise, we set $\hat{s}(0) \neq \hat{s}(1)$. Thus, each $\hat{s}$ has cardinality either 1 or 2. We call objects of the form $\hat{s}(0)$ *left-objects*, and those of the form $\hat{s}(1)$ *right-objects*. For $s \neq t$, we insist that $\hat{s} \cap \hat{t} = \varnothing$.

For each $B \in \mathbf{B}$, let $\hat{B} = \bigcup\{\hat{s} \mid s \text{ a sub-block of } B\}$. Now let $\hat{\mathbf{B}} = \{\hat{B} \mid B \in \mathbf{B}\}$, and $\hat{X} = \bigcup \hat{\mathbf{B}}$. Define the relation $\prec$ on $\hat{X}$ to be the transitive closure of $r_\exists \cup r_\forall$, where

$$r_\exists = \{\langle \hat{s}(i), \hat{t}(i) \rangle \mid i \in \{0, 1\}, s \neq t \text{ and there exist } a \in s, b \in t \text{ such that } a < b\}$$

$$r_\forall = \{\langle \hat{s}(i), \hat{t}(j) \rangle \mid s \ll t,\ i, j \in \{0, 1\}\}.$$

The idea behind this definition is as follows. When sub-blocks s and t contain elements related by $<$, we relate the corresponding left-objects of $\hat{s}$ and $\hat{t}$ by $\prec$, and likewise for the corresponding right-objects. However, we do not relate left-objects to right-objects or vice versa unless our hand is forced by one of two conditions: (i) either of s or t has cardinality 1; (ii) s and t lie in blocks related the block order $\mathbb{B}$ (Fig. 4).

Lemma 3.17 *If $c \in \hat{s}$, $d \in \hat{t}$ and $c \prec d$, then $\hat{s} \neq \hat{t}$. Hence, $\prec$ is a partial order on $\hat{X}$.*

Proof. Suppose $c = c_0, \ldots, c_m = d$ is a sequence of elements of $\hat{X}$, where $\langle c_i, c_{i+1} \rangle \in r_\exists \cup r_\forall$ for all i $(0 \leq i < m)$. For all i $(0 \leq i \leq m)$, let $c_i \subseteq \hat{s}_i$, and let $\langle \mathbf{A}_i^-, A_i, \mathbf{A}_i^+ \rangle$ be the sub-type defining the sub-block s_i. It is immediate from Lemma 3.16 that this sequence of sub-types cannot contain repeated elements. Hence, $\prec$ is irreflexive. $\square$

We observe that $|\hat{X}| \geq 2$. Indeed, if $|\mathbf{B}| \geq 2$, this is immediate. If, on the other hand, $\mathbf{B} = \{B\}$, then $|B| = |X| \geq 2$, whence $|\hat{X}| = |\hat{B}| \geq 2$. Define the function $\widehat{\mathrm{tp}}$ on $\hat{X}$ by setting $\widehat{\mathrm{tp}}(\hat{s}(i)) = \mathrm{tp}(s)$ for every sub-block s

and every $i \in \{0, 1\}$. Thus, $(\hat{X}, \prec, \widehat{\text{tp}})$ is a typed partial order, which we shall denote $\hat{\mathfrak{A}}$. Note that $\hat{\mathfrak{A}}$ does not violate our general restriction to structures of cardinality at least 2. Note also that, if $\hat{B} \in \hat{\mathbf{B}}$, the 1-type $\widehat{\text{tp}}(\hat{s}(i))$ is constant for all $\hat{s}(i) \in \hat{B}$; we denote this value by $\widehat{\text{tp}}(\hat{B})$. Finally, we define a partial order $\preccurlyeq$ on $\hat{\mathbf{B}}$ by setting $\hat{A} \preccurlyeq \hat{B}$ just in case $A \ll B$, and define $\hat{\mathbb{B}} = (\hat{\mathbf{B}}, \preccurlyeq)$.

The number of sub-types is bounded by $|\mathbf{B}|^{2N+1}$, where N is the number of 1-types. To see this, observe that, since blocks of any given type are linearly ordered, the sets $\mathbf{B}^-$ and $\mathbf{B}^+$ are each specified by a sequence of at most N blocks. We now prove a sequence of lemmas culminating in Lemma 3.23, which states that $\hat{\mathfrak{A}} \models \Psi$.

Lemma 3.18 *$\hat{\mathbb{B}}$ is a factorization of $\hat{\mathfrak{A}}$. Moreover, the mapping $B \mapsto \hat{B}$ is an isomorphism of typed partial-orders $(\mathbb{B}, \ll, \text{tp}) \to (\hat{\mathbb{B}}, \preccurlyeq, \widehat{\text{tp}})$.*

Proof. Immediate from the above construction. □

Lemma 3.19 *Let $c \in \hat{s}$ and $d \in \hat{t}$, where s, t are sub-blocks of $\mathbb{B}$, and let $s \subseteq A$, $t \subseteq B$, where A, B are blocks of $\mathbb{B}$. If $c \prec d$, then: (i) for all $a \in s$, there exists $b \in B$ such that $a < b$; and (ii) for all $b \in t$, there exists $a \in A$ such that $a < b$.*

Proof. We may suppose $c = c_0, \ldots, c_m = d$ are elements of $\hat{X}$ such that $(c_i, c_{i+1}) \in r_\exists \cup r_\forall$ for all i $(0 \le i < m)$. We establish (i) by induction on m. If $m = 1$, from the definition of $r_\exists$ and $r_\forall$, there exist $a \in s$ and $b \in t \subseteq B$ such that $a < b$. Since s is a sub-block, for all $a \in s$, there exists $b \in B$ such that $a < b$. If $m > 1$, suppose the result holds for c, d joined by shorter sequences. Let $c_1 \in \hat{s}_1$. From the definition of $r_\exists$ and $r_\forall$, there exist $a \in s$ and $a' \in s_1$ such that $a < a'$. By inductive hypothesis, there exists $b \in B$ such that $a' < b$, whence $a < b$. Since s is a sub-block, for all $a \in s$, there exists $b \in B$ such that $a < b$. The proof of (ii) is similar. □

Lemma 3.20 *Let $c_1 \in \hat{s}_1$, $c_2 \in \hat{s}_2$ and $c_3 \in \hat{s}_3$, where s_1, s_2, s_3 are sub-blocks of $\mathbb{B}$, with s_2 included in (and hence equal to) a unit block B. If $c_1 \prec c_2 \prec c_3$, then, for all $a \in s_1$ and all $b \in s_3$, $a < b$.*

Proof. By Lemma 3.19, for all $a \in s_1$ there exists $b' \in B$ such that $a < b$, and, for all $b \in s_3$ there exists $b' \in B$ such that $b' < b$. But B is a singleton, whence $a < b$. □

Lemma 3.21 *Let $c_1 \in \hat{s}_1$, $c_2 \in \hat{s}_2$, $c_3 \in \hat{s}_3$ and $c_4 \in \hat{s}_4$ where s_1, s_2, s_3 s_4 are sub-blocks of $\mathbb{B}$. If $c_1 \prec c_2$, $\langle c_2, c_3 \rangle \in r_\forall$ and $c_3 \prec c_4$, then, for all $a \in s_1$ and all $b \in s_4$, $a < b$.*

Proof. Similar reasoning to Lemma 3.20. □

Now for the crucial lemma guaranteeing the existence of incomparable witnesses in the typed partial order $\hat{\mathfrak{A}}$. For $c, d \in \hat{X}$, we write $c \asymp d$ to mean that $c \neq d$, $c \not\prec d$ and $d \not\prec c$. That is, $\asymp$ stands in the same relation to $\prec$ as $\sim$ does to $<$.

Lemma 3.22 *Suppose $a, b \in X$ with $a \sim b$. Let s be the sub-block containing a and t the sub-block containing b. Then, for every $c \in \hat{s}$, there exists $d \in \hat{t}$ such that $c \asymp d$.*

Proof. Assume without loss of generality that $c = \hat{s}(0)$ is a left-element. We claim that $d = \hat{t}(1)$ is incomparable to c. For suppose $c \prec d$. Then there is a sequence $c = c_0, \ldots, c_m = d$ of elements of $\hat{X}$ such that $(c_i, c_{i+1}) \in r_\exists \cup r_\forall$ for all i $(0 \le i < m)$. Let $c_i \in \hat{s}_i$ and $s_i \subseteq A_i$ for all i. Since c is a left-element and d is a right element, either A_i is a unit block for some i $(0 \le i \le m)$ or $(c_i, c_{i+1}) \in r_\forall$ for some i $(0 \le i < m)$ (cf. Fig. 4). It then follows from Lemmas 3.20 or 3.21 that, for all $a' \in s$, and all $b' \in t$, $a' < b'$, contradicting the supposition that $a \sim b$. Hence $c \not\prec d$. By a similar argument, $d \not\prec c$. □

Lemma 3.23 $\hat{\mathfrak{A}} \models \Psi$.

Proof. We consider the possible forms of $\psi \in \Psi$ in turn.

(B1a): $\mathfrak{A} \models \psi$ implies that there is just one block A of $\mathbb{B}$ having 1-type α, and $A = s$ is a unit block. But then there is only one element of $\hat{X}$ having 1-type α, namely $\hat{s}(0) = \hat{s}(1)$. Thus $\hat{\mathfrak{A}} \models \psi$.
(B1b): This formula is equivalent to $\forall x \neg\alpha \vee \forall x \neg\beta$; but the realized 1-types in $\mathfrak{A}$ and $\hat{\mathfrak{A}}$ are the same.
(B2a): Suppose $c \in \hat{s}$, $d \in \hat{t}$ be such that $\widehat{\text{tp}}(c) = \alpha$ and $\widehat{\text{tp}}(d) = \alpha$. Let s, t be sub-blocks of the respective blocks A and B. If $c \prec d$, then, by Lemma 3.19, there certainly exist $a \in A$ and $b \in B$ such that $a < b$, contradicting $\mathfrak{A} \models \psi$. Similarly if $d \prec c$. Thus, $\hat{\mathfrak{A}} \models \psi$.
(B2b): Similar to (B2a).

(B3), (B5b): By Lemma 3.18, $\hat{\mathbb{B}}$ is isomorphic to $\mathbb{B}$ (as a typed partial order), so that $\hat{\mathbb{B}} \models \psi$. But since $\hat{\mathbb{B}}$ is a factorization of $\hat{\mathfrak{A}}$, we have, by the first statements of Lemmas 3.2 and 3.3, $\hat{\mathfrak{A}} \models \psi$.

(B4): Let $c = \hat{s}(i)$ be of 1-type β and $d = \hat{t}(j)$ be of 1-type α. Let t lie in the block B of $\mathbb{B}$. Suppose, for contradiction, $c \prec d$. Pick any $a \in s$. By Lemma 3.19, there exists $b \in B$ such that $a < b$. But a is of type β and b of type α, contradicting $\mathfrak{A} \models \psi$. Hence $\hat{\mathfrak{A}} \models \psi$.

(B5a): $\mathfrak{A} \models \psi$ implies that every block of $\mathbb{B}$ having 1-type α is linearly ordered, and hence, by assumption, a unit-block. But then every block of $\hat{\mathbb{B}}$ having 1-type α is a unit-block, whence $\hat{\mathfrak{A}} \models \psi$.

(B6): Let $c = \hat{s}(i)$ be of 1-type α, and pick any $a \in s$. Since $\mathfrak{A} \models \psi$, we have $b > a$ such that $\mathrm{tp}(b) \neq \alpha$ and $\mathfrak{A} \models \mu[b]$. Let b be in the sub-block t, and let $d = \hat{t}(i)$. By construction, $c \prec d$ and $\hat{\mathfrak{A}} \models \mu[d]$, whence $\hat{\mathfrak{A}} \models \psi$.

(B7): Similar to (B6).

(B8): Let $c \in \hat{s}$ be of 1-type α, and pick any $a \in s$. Since $\mathfrak{A} \models \psi$, we have $b \sim a$ such that $\mathfrak{A} \models \mu[b]$. Let b be in the sub-block t. By Lemma 3.22 there exists $d \in \hat{t}$ such that $c \asymp d$. Hence $\hat{\mathfrak{A}} \models \psi$.

(B9), (B10): The realized 1-types in $\mathfrak{A}$ and $\hat{\mathfrak{A}}$ are the same. □

Theorem 3.24 *Let φ be an $\mathcal{L}^2$1PO$^{\mathrm{u}}$-formula in weak normal form with multiplicity m over a signature σ. If φ has a finite model, then it has a model of size bounded by a doubly exponential function of $|\sigma| + m$. Hence, any finitely satisfiable $\mathcal{L}^2$1PO$^{\mathrm{u}}$-formula φ has a model of size bounded by a doubly exponential function of $\|\varphi\|$, and so FinSat($\mathcal{L}^2$1PO$^{\mathrm{u}}$) is in* 2-NEXPTIME.

P r o o f . For the first statement, by Lemma 3.1, we may replace φ by a set Ψ of basic formulas over a signature σ^* of size at most $|\sigma| + 3m$. By Lemma 3.15, let $\mathfrak{A}$ be a typed partial order with unitary factorization $\mathbb{B} = (\mathbf{B}, \ll)$. such that $\mathfrak{A} \models \Psi$, $\mathbb{B} \models \mathrm{FC}(\Psi)$, $\mathbb{B}$ is of size doubly exponential in $|\sigma^*|$, and $\mathfrak{A}$ is thin over $\mathbb{B}$. Now let $\hat{\mathfrak{A}}$ be as defined before Lemma 3.17. By Lemma 3.23, $\hat{\mathfrak{A}} \models \Psi$. But $\hat{\mathfrak{A}}$ is of size at most $2(|\mathbf{B}|^{2N+1})$, where N is the number of 1-types over σ^*. Thus, Ψ is satisfiable over a domain doubly exponential in $|\sigma| + m$. The remainder of the theorem follows by Lemma 2.1. □

4 Two-variable logic with one partial order

The purpose of this section is to show that the logic $\mathcal{L}^2$1PO has the doubly exponential-sized finite model property (Theorem 4.5): if φ is a finitely satisfiable $\mathcal{L}^2$1PO-formula, then φ has a model of size bounded by some fixed doubly exponential function of $\|\varphi\|$. It follows that the finite satisfiability problem for $\mathcal{L}^2$1PO is in 2-NEXPTIME. We proceed by reduction to the corresponding problem for weak normal-form $\mathcal{L}^2$1PO$^{\mathrm{u}}$-formulas, paying particular attention to the size of the relevant signature, and the multiplicities of the formulas in question. In this section, we continue to assume that all signatures contain the navigational predicates $<$, $>$ and $\sim$, subject to the usual semantic constraints. We use the (possibly decorated) variable τ to range over 2-types, λ, μ, ν to range over unary pure Boolean formulas and $\zeta, \eta, \vartheta, \varphi, \chi, \psi, \omega$ to range over arbitrary formulas. Henceforth, for any integer n, we denote by $\lfloor n \rfloor$ the value n modulo 3.

Our strategy is to use resolution theorem-proving to get rid of ordinary binary predicates in some given $\mathcal{L}^2$1PO-formula φ (Lemmas 4.3 and 4.4). To do this, however, we must transform φ into a specialized normal form (Lemmas 4.1 and 4.2). Say that an $\mathcal{L}^2$1PO-formula is in *spread normal form* if it conforms to the pattern

$$\bigwedge_{\zeta \in Z} \exists x.\zeta \wedge \forall x \forall y (x = y \vee \eta) \wedge \bigwedge_{k=0}^{2} \bigwedge_{h=0}^{m-1} \forall x \exists y (\lambda_k \to (\lambda_{\lfloor k+1 \rfloor}(y) \wedge \mu_h(y) \wedge \vartheta_h)), \tag{7}$$

where: (i) Z is a set of unary pure Boolean formulas; (ii) $\eta, \vartheta_0, \ldots, \vartheta_{m-1}$ are quantifier- and equality-free formulas, with $m \geq 1$; (iii) $\lambda_0, \lambda_1, \lambda_2$ are mutually exclusive unary pure Boolean formulas; and (iv) $\mu_0, \ldots, \mu_{m-1}$ are mutually exclusive unary pure Boolean formulas. Spread normal form is—modulo insertion of harmless conjuncts $x \neq y$—a special case of weak normal form (2). We take the *multiplicity* of the spread normal form formula (7) to be the quantity $3m$. (Thus, the definitions of multiplicity for spread normal form and weak normal form agree.) The essential feature of spread normal form is that witnesses are required to be 'spread' over disjoint sets of elements. Thus, suppose $\mathfrak{A}$ is a model of the formula (7), and $\mathfrak{A} \models \lambda_k[a]$ for some $a \in A$ and some k $(0 \leq k < 3)$. Then there exist $b_0, \ldots, b_{m-1} \in A$ such that, for each h $(0 \leq h < m)$, $\mathfrak{A} \models \vartheta_h[a, b_h]$ and $\models \mu_h[b_h]$. It follows that the $b_0, \ldots, b_{m-1}$ are distinct; moreover, all of these elements satisfy $\lambda_{\lfloor k+1 \rfloor}(y)$, so that *their* witnesses, which satisfy $\lambda_{\lfloor k+2 \rfloor}(y)$, cannot include a. Thus, the witnesses for an element of $\mathfrak{A}$ are distinct, and nothing is a witness of a witness of itself.

In order to transform $\mathcal{L}^2$1PO-formulas into spread normal form, we must first establish a lemma allowing us to create copies of certain parts of structures without compromising the truth of those formulas. If $\mathfrak{A}$ is any structure interpreting a signature σ, we call any element of a a *king* if it is the unique element of A realizing its 1-type (over σ): $\mathrm{tp}^{\mathfrak{A}}[b] = \mathrm{tp}^{\mathfrak{A}}[a]$ implies $b = a$ for all $b \in A$. Elements which are kings are said to be *royal*. The following lemma says that we may duplicate the non-royal elements of any structure any (finite) number of times. The technique by no means originates here: it is implicit in the decision procedure for $\mathcal{L}^2$ given in [3] (cf. also [3, Ch. 8]), and had been used by various authors since.

Lemma 4.1 *Let $\mathfrak{A}_1$ be a structure over domain A_1, A_0 the set of kings of $\mathfrak{A}_1$, $\mathfrak{A}_0$ the restriction of $\mathfrak{A}$ to A_0, and $B_1 = A_1 \setminus A_0$. There exists a family of sets $\{B_i\}_{i\geq 2}$, pairwise disjoint and disjoint from A_1, a family of bijections $\{f_i\}_{i\geq 1}$, where $f_i : B_i \to B_1$, and a sequence of structures $\{\mathfrak{A}_i\}_{i\geq 2}$, where $\mathfrak{A}_i$ has domain $A_i = A_0 \cup B_1 \cup B_2 \cup \cdots \cup B_i$, such that, for all $i \geq 1$:*

(i) $\mathfrak{A}_{i-1} \subseteq \mathfrak{A}_i$, and all 2-types realized in $\mathfrak{A}_i$ are realized in $\mathfrak{A}_1$;
(ii) for all $a \in B_i$ and all $b \in A_1$, if $f_i(a) \neq b$, then $\mathrm{tp}^{\mathfrak{A}_i}[a, b] = \mathrm{tp}^{\mathfrak{A}_1}[f_i(a), b]$;
(iii) for all $a \in B_i$, all j $(1 \leq j \leq i)$ and all $b \in B_j$, if $f_i(a) \neq f_j(b)$, then $\mathrm{tp}^{\mathfrak{A}_i}[a, b] = \mathrm{tp}^{\mathfrak{A}_1}[f_i(a), f_j(b)]$;
(iv) $<^{\mathfrak{A}_i}$ is a partial order.

Proof. Enumerate B_1 as $\{a^1, a^2, a^3, \ldots\}$. Let the set of indices of this enumeration (which may be finite or infinite) be K. For each $k \in K$, let $b^k \in B_1$ be such that $\mathrm{tp}^{\mathfrak{A}}[a^k] = \mathrm{tp}^{\mathfrak{A}}[b^k]$ but $a^k \neq b^k$. This is possible because $A_0 = A_1 \setminus B_1$ is the set of kings of $\mathfrak{A}_1$.

We prove the lemma by induction on i. For the base case, observe that the domain of $\mathfrak{A}_1$ is the disjoint union of A_0 and B_1, and define the bijection $f_1 : B_1 \to B_1$ to be the identity map. It is trivial to check that Statements (i)–(iv) then hold for $i = 1$. Notice, incidentally, that the set $A_0 \subseteq A_1$ has been defined. For the inductive case, let $i \geq i$, and suppose the sets A_{i-1}, B_i, the structure $\mathfrak{A}_i$, and the bijection $f_i : B_i \to B_1$ have been defined, such that the domain A_i of $\mathfrak{A}_i$ is the disjoint union of A_{i-1} and B_i, and Statements (i)–(iv) hold. We proceed to define A_{i+1}, B_{i+1}, $\mathfrak{A}_{i+1}$ and f_{i+1}, and establish the corresponding properties for these objects.

Intuitively, $\mathfrak{A}_{i+1}$ is just like $\mathfrak{A}_i$ except that we have added an extra copy of the set B_1, relating the new elements to each other and to $\mathfrak{A}_i$ as specified by $\mathfrak{A}_1$. Formally, we proceed via a subsidiary induction. Let $A_i^1 = A_i$ and $\mathfrak{A}_i^1 = \mathfrak{A}_i$. We shall construct a sequence of structures $\{\mathfrak{A}_i^k\}_{k\in K}$ over the corresponding sequence of domains $\{A_i^k\}_{k\in K}$. Assume that $\mathfrak{A}_i^k$ has been defined over domain A_i^k and $k+1 \in K$. Let a_{i+1}^k be a new element (not in A_i^k), and let $A_i^{k+1} = A_i^k \cup \{a_{i+1}^k\}$. (The indexing reflects the intuition that a_{i+1}^k will form the kth new element in the structure $\mathfrak{A}_{i+1}$ when this is completed.) We extend $\mathfrak{A}_i^k$ to a structure $\mathfrak{A}_i^{k+1}$ over A_i^{k+1} by setting:

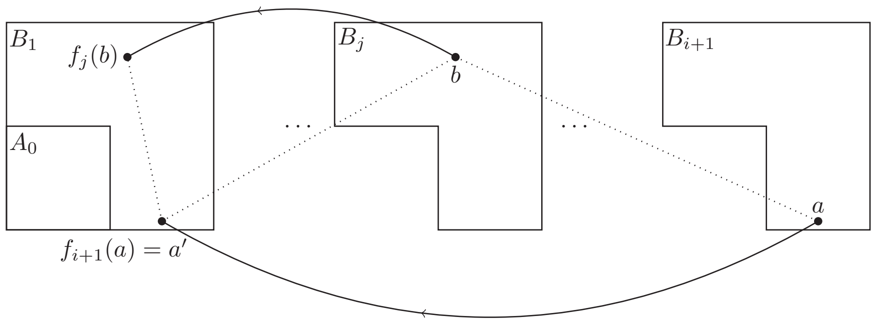
$$\mathrm{tp}^{\mathfrak{A}_i^{k+1}}[a_{i+1}^k] = \mathrm{tp}^{\mathfrak{A}_i^k}[a^k] \tag{8}$$

$$\mathrm{tp}^{\mathfrak{A}_i^{k+1}}[a_{i+1}^k, a^k] = \mathrm{tp}^{\mathfrak{A}_i^k}[b^k, a^k] \tag{9}$$

$$\mathrm{tp}^{\mathfrak{A}_i^{k+1}}[a_{i+1}^k, b] = \mathrm{tp}^{\mathfrak{A}_i^k}[a^k, b] \quad \text{for all } b \in A_i^k \setminus \{a^k\}. \tag{10}$$

From the fact that $\mathrm{tp}^{\mathfrak{A}_i^k}[a^k] = \mathrm{tp}^{\mathfrak{A}_i^k}[b^k]$, these type-assignments involve no clashes. Moreover, since $\mathfrak{A}_i^1 \subsetneq \mathfrak{A}_i^2 \subsetneq \cdots$, we may define $\mathfrak{A}_{i+1} = \bigcup_{k\in K} \mathfrak{A}_i^k$, taking $\mathfrak{A}_{i+1}$ to have domain A_{i+1}. Letting $B_{i+1} = \{a_{i+1}^1, a_{i+1}^2, \ldots\}$, we see that $A_{i+1} = A_i \cup B_{i+1} = A_0 \cup B_0 \cup \cdots \cup B_{i+1}$. Define the bijection $f_{i+1} : B_{i+1} \to B_1$ by setting $f_{i+1}(a_{i+1}^k) = a^k$ for all $k \in K$.

We need to secure Statements (i)–(iv) of the lemma, but with i replaced by $i+1$. For Statement (i), it is immediate by construction that $\mathfrak{A}_i \subseteq \mathfrak{A}_{i+1}$ and from (9) and (10), via a subsidiary induction on k, we see that $\mathfrak{A}_{i+1}$ realizes only those 2-types realized in $\mathfrak{A}_i$, and hence, by inductive hypothesis, in $\mathfrak{A}_1$. For Statement (ii), it follows from (10), again via a subsidiary induction on k, that, for all $a \in B_{i+1}$ and all $b \in A_1$, if $f_{i+1}(a) \neq b$, then $\mathrm{tp}^{\mathfrak{A}_{i+1}}[a, b] = \mathrm{tp}^{\mathfrak{A}_1}[f_{i+1}(a), b]$. For Statement (iii), we consider separately the cases $j = i+1$ and $j \leq i$. The former is the simpler: observe that, for $k, \ell \in K$ with $k < \ell$, $\mathrm{tp}^{\mathfrak{A}_i^\ell}[a_{i+1}^k, a_{i+1}^\ell] = \mathrm{tp}^{\mathfrak{A}_1}[a^k, a^\ell]$. Indeed,



$$\mathrm{tp}^{\mathfrak{A}_{i+1}}[a, b] = \mathrm{tp}^{\mathfrak{A}_j}[f_{i+1}(a), b] = \mathrm{tp}^{\mathfrak{A}_1}[f_{i+1}(a), f_j(b)]$$

Fig. 5 Construction of the structure $\mathfrak{A}_{i+1}$ (Lemma 4.1), where $a \in B_{i+1}$, $b \in B_j$, and $j \le i$.

$\mathrm{tp}^{\mathfrak{A}_i^\ell}[a^k_{i+1}, a^\ell_{i+1}] = \mathrm{tp}^{\mathfrak{A}_i^{\ell-1}}[a^k_{i+1}, a^\ell] = \mathrm{tp}^{\mathfrak{A}_i^k}[a^k_{i+1}, a^\ell] = \mathrm{tp}^{\mathfrak{A}_i^{k-1}}[a^k, a^\ell] = \mathrm{tp}^{\mathfrak{A}_1}[a^k, a^\ell]$. Thus, for distinct $a, b \in B_{i+1}$, we have

$$\mathrm{tp}^{\mathfrak{A}_{i+1}}[a, b] = \mathrm{tp}^{\mathfrak{A}_1}[f_{i+1}(a), f_{i+1}(b)].$$

This secures Statement (iii) for the case $j = i + 1$. The case $1 \le j \le i$ is illustrated in Fig. 5. Writing $a' = f_{i+1}(a)$, by statement (ii) of the inductive hypothesis, if $f_j(b) \ne a'$, then $\mathrm{tp}^{\mathfrak{A}_i}[a', b] = \mathrm{tp}^{\mathfrak{A}_1}[a', f_j(b)]$, and by construction of $\mathfrak{A}_{i+1}$, $\mathrm{tp}^{\mathfrak{A}_{i+1}}[a, b] = \mathrm{tp}^{\mathfrak{A}_i}[a', b]$. Thus, if $f_{i+1}(a) \ne f_j(b)$, then $\mathrm{tp}^{\mathfrak{A}_{i+1}}[a, b] = \mathrm{tp}^{\mathfrak{A}_1}[f_{i+1}(a), f_j(b)]$.

Turning to Statement (iv), it follows from (8) that $<^{\mathfrak{A}^k_{i+1}}$ is not reflexive. We claim that, in addition, this relation is transitive. It evidently suffices to show that if $<^{\mathfrak{A}^k_i}$ is a transitive relation, then so is $<^{\mathfrak{A}^{k+1}_i}$. As a guide to intuition here, it helps to observe that the new element, a^k_{i+1} has been inserted 'close to' the element a^k in the partial order of $\mathfrak{A}^k_i$. Suppose then that $<^{\mathfrak{A}^k_i}$ is transitive, and let a, b, c be elements of A^{k+1}_i such that $\mathfrak{A}^{k+1}_i \models a < b$ and $\mathfrak{A}^{k+1}_i \models b < c$. We must show $\mathfrak{A}^{k+1}_i \models a < c$. If $a, b, c \in A^k_i$, this is immediate. Moreover, if $a = b$ or $b = c$ there is nothing to show. On the other hand, if $a = c = a^k_{i+1}$ and $b \in A^k_i$, then either $\mathrm{tp}^{\mathfrak{A}^{k+1}_i}[a, b] = \mathrm{tp}^{\mathfrak{A}^{k+1}_i}[c, b] = \mathrm{tp}^{\mathfrak{A}^k_i}[a^k, b]$ or $\mathrm{tp}^{\mathfrak{A}^{k+1}_i}[a, b] = \mathrm{tp}^{\mathfrak{A}^{k+1}_i}[c, b] = \mathrm{tp}^{\mathfrak{A}^k_i}[b^k, b]$, in either case contradicting the supposition that $\mathfrak{A}^{k+1}_i \models a < b$ and $\mathfrak{A}^{k+1}_i \models b < c$. Moreover, an exactly similar argument applies if $a = c \in A^k_i$ and $b = a^k_{i+1}$. Hence, we may assume that the elements a, b and c are distinct, and that exactly one of them is equal to a^k_{i+1}. We therefore have three cases to consider.

Case 1. $a^k_{i+1} = a$. We claim first of all that $c \ne a^k$. For suppose $c = a^k \ne b$. Then $\mathfrak{A}^{k+1}_i \models b < c$ implies $\mathfrak{A}^k_i \models b < a^k$, whence $\mathfrak{A}^k_i \not\models a^k < b$, and therefore $\mathfrak{A}^{k+1}_i \not\models a^k_{i+1} < b$, contradicting the supposition that $\mathfrak{A}^{k+1}_i \models a < b$. If, on the other hand, $b = a^k \ne c$, then $\mathfrak{A}^k_i \models b < c$ is the statement $\mathfrak{A}^k_i \models a^k < c$, which implies $\mathfrak{A}^{k+1}_i \models a < c$. Thus we may suppose that neither b nor c is equal to a^k. But then $\mathfrak{A}^k_i \models a^k < b$ and $\mathfrak{A}^k_i \models b < c$, whence $\mathfrak{A}^k_i \models a^k < c$, whence $\mathfrak{A}^{k+1}_i \models a < c$.

Case 2. $a^k_{i+1} = b$. Suppose first that $a = a^k$. Then $\mathfrak{A}^{k+1}_i \models b < c$ implies $\mathfrak{A}^k_i \models a^k < c$, and hence $\mathfrak{A}^{k+1}_i \models a^k < c$, which is the required statement $\mathfrak{A}^{k+1}_i \models a < c$. A similar argument applies if $c = a^k$. Thus we may suppose that neither a nor c is equal to a^k. But then $\mathfrak{A}^k_i \models a < a^k$ and $\mathfrak{A}^k_i \models a^k < c$, whence $\mathfrak{A}^k_i \models a < c$, whence $\mathfrak{A}^{k+1}_i \models a < c$.

Case 3. $a^k_{i+1} = c$. The same as Case 1, but with the order reversed.

This completes the induction. □

We now come to the lemma allowing us to transform any $\mathcal{L}^2$1PO-formula in standard normal form into one in spread normal form. We require some additional notation. Let $\bar{p} = p_1, \dots, p_n$ be a sequence of unary predicates. For all i $(0 \le i < 2^n)$, we abbreviate by $\bar{p}\langle i\rangle$ the unary, pure Boolean formula $\varrho_1 \wedge \dots \wedge \varrho_n$, where, for all j $(1 \le j \le n)$, ϱ_j is $p_j(x)$ if the jth bit in the n-digit binary representation of i is 1, and $\neg p_j(x)$ otherwise. We call $\bar{p}\langle i\rangle(x)$ the ith *labelling formula* (*over* $p_1, \dots, p_n$). Evidently, if $A = \{a_0, \dots, a_{M-1}\}$ is a set of cardinality $M \le 2^n$, then we can interpret the predicates in p_j $(1 \le j \le n)$ over A so as to ensure that, for all i $(0 \le i < M)$, a_i satisfies $\bar{p}\langle i\rangle$.

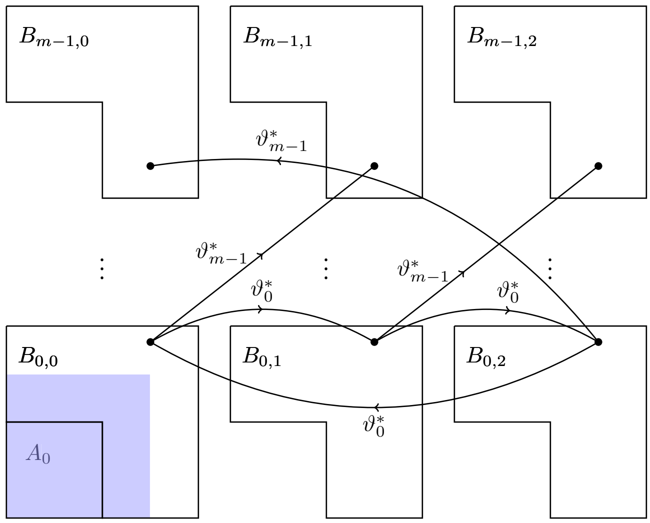


Fig. 6 Construction of the structure $\mathfrak{A}$ (Lemma 4.2), illustrating 'circular witnessing', in the case where B_1 has been re-indexed as $B_{0,0}$; the shaded region depicts the court, $C_0 \subseteq A_1 = A_0 \cup B_1$.

Lemma 4.2 *Let φ be an $\mathcal{L}^2$1PO-formula in standard normal form over a signature σ, having multiplicity m. There exists a formula φ^* in spread normal form over a signature σ^* with the following properties: (i) $\models \varphi^* \to \varphi$; (ii) if φ has a (finite) model then so has φ^*; and (iii) $|\sigma^*|$ is polynomially bounded as a function of $|\sigma| + m$, and φ^* has multiplicity $3m$.*

P r o o f . The strategy is to begin with any (finite) model, say $\mathfrak{A}_1$, of φ and apply Lemma 4.1 so as to obtain a structure containing $3m$ copies of the set of non-royal elements of $\mathfrak{A}_1$—that is, 3 copies for each $\forall\exists$-conjunct of φ. By interpreting additional unary predicates to label the royal elements and to encode certain relations between royal and non-royal elements, we obtain a (finite) structure $\mathfrak{A}'''$ satisfying a certain spread normal form formula φ^*. To complete the proof, we show that φ^* entails φ.

Write φ as

$$\forall x \forall y (x = y \vee \eta) \wedge \bigwedge_{h=0}^{m-1} \forall x \exists y (x \neq y \wedge \vartheta_h).$$

Suppose $\mathfrak{A}_1 \models \varphi$, and let A_0 be the set of kings of $\mathfrak{A}_1$. We may assume $|A_0| \geq 2$. (If not, expand $\mathfrak{A}_1$ with two new unary predicates, and interpret them so that they are uniquely instantiated by distinct elements.) Let $\mathfrak{A}_0$ be the restriction of $\mathfrak{A}$ to A_0, and let $B_1 = A_1 \setminus A_0$. Now take $\{B_i\}_{i \geq 2}$, $\{f_i\}_{i \geq 1}$ and $\{\mathfrak{A}_i\}_{i \geq 2}$ to be the series of sets, bijections and structures guaranteed by Lemma 4.1. Let $\mathfrak{A} = \mathfrak{A}_{3m}$; thus, $\mathfrak{A}$ is finite if $\mathfrak{A}_1$ is. Finally, re-index the sets $B_1, \ldots, B_{3m}$ (in any order whatever) as $B_{h,k}$, where $0 \leq h < m$ and $0 \leq k < 3$; and re-index the $f_1, \ldots, f_{3m}$ correspondingly as $f_{h,k}$.

For each $a \in A_0$ and each h ($0 \leq h < m$) select some $b \in A_1 \setminus \{a\}$ such that $\mathfrak{A}_1 \models \vartheta_h[a, b]$, and let C_0 consist of the elements of A_0 together with all of the (at most $m \cdot |A_0|$) elements thus selected. We refer to C_0 as the *court* of $\mathfrak{A}$. The situation is illustrated in Fig. 6, in the case where B_1 has been re-indexed as $B_{0,0}$. Let us enumerate A_0 as $c_0, \ldots, c_{S-1}$ and the rest of C_0 as $c_S, \ldots, c_{T-1}$. Thus, $0 \leq S \leq T \leq (m+1)2^{|\sigma|}$. Let $t = \lceil \log(T+1) \rceil$, and let $q_1, \ldots, q_t$ be new unary predicates. Writing $\bar{q}\langle i \rangle$ for the ith labelling formula over $q_1, \ldots, q_t$, let $\mathfrak{A}$ be expanded to a structure $\mathfrak{A}'$ such that, for all i ($0 \leq i < T$), $\mathfrak{A}' \models \bar{q}\langle i \rangle[c_i]$, and $\mathfrak{A}' \models \bar{q}\langle T \rangle[a]$ for all $a \in A \setminus C_0$.

Thus, under the interpretation $\mathfrak{A}'$, for $0 \leq i < S$, we may read $\bar{q}\langle i \rangle(x)$ as "x is the ith king;" and for $0 \leq i < T$, we may read $\bar{q}\langle i \rangle$ as "x is the ith member of the court." (Hence, the kings come before the non-royal courtiers in the numbering.) Now let χ be the formula

$$\bigwedge_{i=0}^{T-1} \exists x . \bar{q}\langle i \rangle(x)$$

and ψ_1 the formula

$$\bigwedge_{i=0}^{T-2} \bigwedge_{j=i+1}^{T-1} \forall x \forall y (x = y \vee (\bar{q}\langle i \rangle(x) \wedge \bar{q}\langle j \rangle(y)) \to \mathrm{tp}^{\mathfrak{A}}[c_i, c_j]),$$

recording the diagram of $\mathfrak{A}$ over C_0. (Technically, we also need to remove the literals $x \neq y$ from the formulas $\mathrm{tp}^{\mathfrak{A}}[c_i, c_j]$; but this is a minor detail.) Obviously, $\mathfrak{A}' \models \chi \wedge \psi_1$. Conversely, in any model of $\chi \wedge \psi_1$, we see that for all h $(0 \leq h < m)$ and for any element a satisfying $\bar{q}\langle i \rangle$ for some i $(0 \leq i < S)$, there exists $b \neq a$ such that the pair $\langle a, b \rangle$ satisfies ϑ_h.

Let $s = \lceil \log(S + 1) \rceil$. For each h $(0 \leq h < m)$, let $q_1^h, \ldots, q_s^h$ be new unary predicates, and write $\bar{q}^h\langle i \rangle(x)$ for the ith labelling formula over these predicates. Expand $\mathfrak{A}'$ to a model $\mathfrak{A}''$ as follows. For each $a \in A \backslash A_0$, and each h $(0 \leq h < m)$, if there exists any king $b \in A_0$ such that $\mathfrak{A} \models \vartheta_h[a, b]$, choose some such element, say, c_i (with i depending on a and h), and interpret the predicates $q_1^h, \ldots, q_s^h$ so that $\mathfrak{A}'' \models \bar{q}^h\langle i \rangle[a]$; otherwise, interpret the predicates $q_1^h, \ldots, q_s^h$ so that $\mathfrak{A}'' \models \bar{q}^h\langle S \rangle[a]$. Thus, under the interpretation $\mathfrak{A}''$, for $0 \leq i < S$, we may read $\bar{q}^h\langle i \rangle(x)$ as "x is a non-royal element such that the ith king provides a ϑ_h-witness for x." Now let ψ_2 be the formula

$$\bigwedge_{i=0}^{S-1} \bigwedge_{h=0}^{m-1} \forall x \forall y (x = y \vee (\bar{q}^h\langle i \rangle(x) \wedge \bar{q}\langle i \rangle(y) \to \vartheta_h)),$$

recording this fact. Obviously, $\mathfrak{A}'' \models \psi_2$. Conversely, in any model of $\chi \wedge \psi_2$, we see that for all h $(0 \leq h < m)$, and all elements a satisfying $\bar{q}^h\langle i \rangle(x)$ for some i $(0 \leq i < S)$, there exists $b \neq a$ such that the pair $\langle a, b \rangle$ satisfies ϑ_h.

Finally, let o_0, o_1, o_2 and $p_0, \ldots, p_{m-1}$ be new unary predicates, and expand $\mathfrak{A}''$ to a structure $\mathfrak{A}'''$ by setting

$$(o_k)^{\mathfrak{A}} = \bigcup_{h=0}^{m-1} B_{h,k} \qquad \text{for all } k \ (0 \leq k < 3)$$

$$(p_h)^{\mathfrak{A}} = \bigcup_{k=0}^{2} B_{h,k} \qquad \text{for all } h \ (0 \leq h < m).$$

Thus, we may read $o_k(x)$ as "x is in $B_{h,k}$ for some h", and $p_h(x)$ as "x is in $B_{h,k}$ for some k". Let $\lambda_0 = o_0(x)$, $\lambda_1 = o_1(x) \wedge \neg o_0(x)$, $\lambda_2 = o_2(x) \wedge \neg o_0(x) \wedge \neg o_1(x)$. Thus, $\lambda_0, \lambda_1, \lambda_2$ are mutually exclusive pure unary formulas. Similarly, let $\mu_h(x) = p_h(x) \wedge \bigwedge_{h'=0}^{h-1} \neg p_{h'}(x)$ for all h $(0 \leq h < m)$. Thus, $\mu_0, \ldots, \mu_{m-1}$ are also mutually exclusive unary pure Boolean formulas.

Now let ψ_3 be the formula

$$\forall x \forall y \left(x = y \vee \bigvee_{i=0}^{S-1} \bar{q}\langle i \rangle(x) \vee \bigvee_{k=0}^{2} \lambda_k \right),$$

which, we note, is equivalent (over structures with cardinality at least 2) to

$$\forall x \left(\bigvee_{i=0}^{S-1} \bar{q}\langle i \rangle(x) \vee \bigvee_{k=0}^{2} \lambda_k \right).$$

It is immediate by construction that $\mathfrak{A}''' \models \psi_3$, since every $a \in A_0$ satisfies $\bar{q}\langle i \rangle(x)$ for some i $(0 \leq i < S)$, and every $a \in A \backslash A_0$ lies in one of the sets $B_{h,k}$. In addition, let $\vartheta_h^*(x, y)$ be the formula

$$\left(\bigwedge_{i=0}^{S-1} \neg \bar{q}^h\langle i \rangle(x) \right) \to \vartheta_h,$$

for all h $(0 \leq h < m)$, and let ω be the formula

$$\bigwedge_{h=0}^{m-1} \bigwedge_{k=0}^{2} \forall x (\lambda_k \to \exists y (\lambda_{\lfloor k+1 \rfloor}(y) \wedge \mu_h(y) \wedge \vartheta_h^*)).$$

We claim that $\mathfrak{A}''' \models \omega$. To see this, fix $0 \leq h < m$ and $0 \leq k < 3$, and suppose $a \in A$ is such that $\mathfrak{A}''' \models \lambda_k[a]$. If $\mathfrak{A}''' \models \bar{q}^h\langle i \rangle[a]$ for some i $(0 \leq i < S)$, then we may pick any element $b \in B_{h,\lfloor k+1 \rfloor}$ as a witness, since $\mathfrak{A}''' \models \vartheta_h^*[a, b]$ holds by failure of the antecedent. Otherwise, by the construction of $\mathfrak{A}'''$, $a \in B_{h',k}$ for some h' $(1 \leq h' \leq m)$ and, moreover, there is no $b \in A_0$ for which $\mathfrak{A} \models \vartheta_h[a, b]$. Now let $a' = f_{h',k}(a)$. Since $\mathrm{tp}^{\mathfrak{A}}[a, b] = \mathrm{tp}^{\mathfrak{A}}[a', b]$ for all $b \in A_0$, it follows that there is no $b \in A_0$ for which $\mathfrak{A} \models \vartheta_h[a', b]$. Since $\mathfrak{A}_1 \models \varphi$, therefore, let $b' \in B_1$ be such that $\mathfrak{A} \models \vartheta_h[a', b']$ and let $b \in B_{h,\lfloor k+1 \rfloor}$ be such that $f_{h,\lfloor k+1 \rfloor}(b) = b'$. Since $\mathrm{tp}^{\mathfrak{A}}[a, b] = \mathrm{tp}^{\mathfrak{A}}[a', b']$, we have $\mathfrak{A} \models \vartheta_h[a, b]$. That is: we obtain the 'circular witnessing' pattern illustrated in Fig. 6. Moreover, by the construction of $\mathfrak{A}'''$, $\mathfrak{A}''' \models \lambda_{\lfloor k+1 \rfloor}[b]$ and $\mathfrak{A}''' \models \mu_h[b]$. Therefore, $\mathfrak{A}''' \models \omega$ as claimed. Conversely, in any model of ω, we see that for all h $(0 \leq h < m)$ and all elements a satisfying $\lambda_k(x)$ but not satisfying $\bar{q}^h\langle i \rangle(x)$ for any i $(0 \leq i < S)$, there exists some $b \neq a$ such that the pair $\langle a, b \rangle$ satisfies ϑ_h.

Finally, let φ^* be the formula

$$\chi \wedge \forall x \forall y(x = y \vee \eta) \wedge \psi_1 \wedge \psi_2 \wedge \psi_3 \wedge \omega,$$

and let σ^* be the signature of ψ^*. Thus, φ^* is in spread form, with multiplicity $3m$. Moreover, the only new predicates in σ^* are $o_0, o_1, o_2, p_0, \ldots, p_{m-1}, q_1, \ldots q_t$, and the $q_1^h, \ldots q_s^h$ $(0 \leq h < m)$, so that $|\sigma^*|$ is bounded by a polynomial function of $|\sigma| + m$. Moreover, we have shown that, if $\mathfrak{A} \models \varphi$, then $\mathfrak{A}''' \models \varphi^*$, and, furthermore, $\mathfrak{A}'''$ is finite if $\mathfrak{A}$ is. It remains to show that $\models \varphi^* \rightarrow \varphi$. So suppose $\mathfrak{B} \models \varphi^*$, $a \in B$ and $0 \leq h < m$. As we have observed, if $\mathfrak{B} \models \bar{q}\langle i \rangle[a]$ for some i $(0 \leq i < S)$, then $\chi \wedge \psi_1$ guarantees the existence of some $b \in B \backslash \{a\}$ such that $\mathfrak{B} \models \vartheta_h[a, b]$. Otherwise, by ψ_3, $\mathfrak{B} \models \lambda_k[a]$ for some k $(0 \leq k < 3)$. If, now $\mathfrak{B} \models \bar{q}^h\langle i \rangle[a]$ for some i $(0 \leq i < S)$, $\chi \wedge \psi_2$ guarantees the existence of some $b \in B \backslash \{a\}$ such that $\mathfrak{B} \models \vartheta_h[a, b]$. If, on the other hand, $\mathfrak{B} \not\models \bar{q}^h\langle i \rangle[a]$ for any i $(0 \leq i < S)$, ω guarantees the existence of some $b \in B \backslash \{a\}$ such that $\mathfrak{A}_1 \models \vartheta_h[a, b]$. Thus, $\mathfrak{B} \models \varphi$. □

In the sequel, we employ terminology and techniques familiar from the area of automated theorem proving. In particular, a *clause* over signature σ is a disjunction (possibly empty) of σ-literals. The empty disjunction is written as $\perp$, and is taken to denote the falsum. We assume that the only variables appearing in clauses are x and y, and we silently delete repeated literals from clauses. Given these restrictions, the set of clauses over a fixed, finite, relational signature is finite. We use (possibly decorated) lower-case Greek letters γ, δ, ε to range over clauses, and upper-case Greek letters Γ, Δ to range over finite sets of clauses. If Γ is a finite set of clauses, then we denote by Γ^{-1} the result of transposing the variables x and y in Γ. To avoid notational clutter, we frequently identify a finite set of clauses with its conjunction, writing, for example, Γ when we actually mean $\bigwedge \Gamma$. It is a familiar fact that, for any quantifier-free formula φ over a signature σ, there exists a collection of clauses Γ over σ such that $\models \varphi \leftrightarrow \Gamma$ (so-called conjunctive normal form). In general $|\Gamma|$ will be exponential in $\|\varphi\|$; however, the signature is unaffected.

Let ϱ be an atomic formula featuring two distinct variables–i.e., a formula of either of the forms $r(x, y)$ or $r(y, x)$, where r is a binary predicate, and let γ', δ' be clauses. Then, $\gamma = \varrho \vee \gamma'$ and $\delta = \neg\varrho \vee \delta'$ are also clauses, as indeed is $\gamma' \vee \delta'$. In that case, we call $\gamma' \vee \delta'$ a *binary resolvent* of γ and δ, and we say that $\gamma' \vee \delta'$ is *obtained by binary resolution* from γ and δ *on* ϱ, or simply: γ and δ *resolve to form* $\gamma' \vee \delta'$. Note that no unification of variables occurs in binary resolution, and no new variables are created: in fact, binary resolution is just the familiar rule of propositional resolution restricted to the case where the resolved-on atom is of the form $r(x, y)$ or $r(y, x)$, with r a binary predicate. Observe that: (i) if γ and δ resolve to form ε, then $\models \gamma \wedge \delta \rightarrow \varepsilon$; (ii) the binary resolvent of two clauses may or may not involve atoms of the form $r(x, y)$ or $r(y, x)$; (iii) if the clause γ involves no atoms of these forms, then it cannot undergo binary resolution at all.

If Γ is a set of clauses, denote by $[\Gamma]^*$ the smallest set of clauses including Γ and closed under binary resolution, in the sense that, if $\gamma, \delta \in [\Gamma]^*$, and ε is a binary resolvent of γ and δ, then $\varepsilon \in [\Gamma]^*$. Let σ' be any subset of σ such that $\sigma \backslash \sigma'$ contains only binary predicates. We denote by $[\Gamma]^*_{/\sigma'}$ the result of deleting from $[\Gamma]^*$ any clause involving an atom $r(x, y)$ or $r(y, x)$, where $r \notin \sigma'$. Notice, incidentally, that $[\Gamma]^*_{/\sigma'}$ may feature predicates in $\sigma \backslash \sigma'$: however, all occurrences of these must be in atoms of the forms $r(x, x)$ or $r(y, y)$.

This last observation prompts the introduction of some additional notation and terminology that will be used in the next lemma. Call a literal *diagonal* if it is of the form $\pm r(u, u)$, where r is a binary predicate and u a variable. Let σ be a relational signature and $\sigma' \subseteq \sigma$ such that $\sigma \backslash \sigma'$ consists only of binary predicates. A *semi-diagonal 2-type over* (σ, σ') is a maximal consistent set of literals over σ each one of which is either a literal over σ' or a diagonal literal. If $\mathfrak{A}$ is a structure interpreting σ, and a, b are elements of the domain $\mathfrak{A}$, we denote by $\mathrm{tp}^{\mathfrak{A}}_{/\sigma'}[a, b]$

the unique semi-diagonal 2-type over (σ, σ') satisfied by the pair $\langle a, b\rangle$. Thus, if a and b are distinct, $\mathrm{tp}^{\mathfrak{A}}_{/\sigma'}[a, b]$ is just like $\mathrm{tp}^{\mathfrak{A}}[a, b]$, except that it is silent on the question of which binary relations in $\sigma\backslash\sigma'$ are satisfied by the pairs $\langle a, b\rangle$ and $\langle b, a\rangle$.

The following lemma, which will form the core of our reduction of *FinSat*($\mathcal{L}^2$1PO) to *FinSat*($\mathcal{L}^2$1PO$^{\mathrm{u}}$), is, in effect, nothing more than the familiar completeness theorem for (ordered) propositional resolution.

Lemma 4.3 *Let Γ be a set of clauses over a signature σ, let $\sigma' \subseteq \sigma$ be such that $\sigma\backslash\sigma'$ consists only of ordinary binary predicates, and let τ^- be a semi-diagonal 2-type over (σ, σ'). If $\models \tau^- \to [\Gamma]^*_{/\sigma'}$, then there exists a 2-type τ over the signature σ such that $\models \tau \to \tau^-$ and $\models \tau \to \Gamma$.*

Proof. Enumerate the formulas of the forms $r(x, y)$ and $r(y, x)$, where r is an (ordinary binary) predicate in $\sigma\backslash\sigma'$, as $\varrho_1, \dots, \varrho_n$. Define a *level-$i$ extension* of τ^- inductively as follows: (i) τ^- is a level-0 extension of τ^-; (ii) if τ' is a level-i extension of τ^- $(0 \le i < n)$, then $\tau' \wedge \varrho_{i+1}$ and $\tau' \wedge \neg\varrho_{i+1}$ are level-$(i+1)$ extensions of τ^-. Thus, the level-n extensions of τ are exactly the 2-types over σ entailing τ^-. (Here we rely on the assumption that navigational predicates do not occur in $\sigma\backslash\sigma'$: we are free to add ϱ_{i+1} or $\neg\varrho_{i+1}$ to τ' without generating inconsistencies.) If τ' is a level-i extension of τ^- $(0 \le i < n)$, we say that τ' *violates* a clause δ if, for every literal in δ, the opposite literal is in τ'; we say that τ' *violates* a set of clauses Δ if τ' violates some $\delta \in \Delta$. We construct a sequence of level-i extensions of τ^- $(i = 0, 1, \dots)$ none of which violates $[\Gamma]^*$.

By definition, τ^- is a level-0 extension of itself. Suppose now that τ' is a level-i extension of τ^- $(0 \le i < n)$. We claim that, if both $\tau' \wedge \varrho_{i+1}$ and $\tau' \wedge \neg\varrho_{i+1}$ violate $[\Gamma]^*$, then so does τ^-. For otherwise, there must be a clause $\neg\varrho_{i+1} \vee \gamma' \in [\Gamma]^*$ violated by $\tau' \wedge \varrho_{i+1}$ and a clause $\varrho_{i+1} \vee \delta' \in [\Gamma]^*$ violated by $\tau' \wedge \neg\varrho_{i+1}$. But in that case τ' violates the binary resolvent $\gamma' \vee \delta'$, contradicting the supposition that τ' does not violate $[\Gamma]^*$. (Here we rely on the assumption that all predicates in $\sigma\backslash\sigma'$ are binary.) This proves the claim. Now, since τ^- by hypothesis entails $[\Gamma]^*_{/\sigma'}$, it certainly does not violate $[\Gamma]^*_{/\sigma'}$. Therefore, since τ^- involves no atoms of the form $r(x, y)$ or $r(y, x)$ for $r \in \sigma\backslash\sigma'$, it does not violate $[\Gamma]^*$ either. By the above claim, then, there must be at least one level-n extension τ of τ^- which does not violate $[\Gamma]^* \supseteq \Gamma$. Since τ is a 2-type, this proves the lemma. $\square$

The next lemma will allow us to eliminate atoms of the forms $r(x, y)$ and $r(y, x)$, where r is an ordinary binary predicate, from spread-form $\mathcal{L}^2$1PO-formulas. (We state it in a slightly more general form than is required for this purpose.) Recall that, if Γ is a finite set of clauses, Γ^{-1} denotes the result of transposing the variables x and y in Γ.

Lemma 4.4 *Let φ be the spread-form $\mathcal{L}^2$1PO-formula*

$$\bigwedge_{\zeta \in Z} \exists x.\zeta \wedge \forall x \forall y (x = y \vee \Gamma) \wedge \bigwedge_{k=0}^{2} \bigwedge_{h=0}^{m-1} \forall x \exists y (\lambda_k \to (\lambda_{\lfloor k+1 \rfloor}(y) \wedge \mu_h(y) \wedge \Delta_h)),$$

over some signature σ, where: Z is a set of pure unary formulas; $\lambda_0, \lambda_1, \lambda_2$ are mutually exclusive pure unary formulas; $\mu_0, \dots, \mu_{m-1}$ are mutually exclusive pure unary formulas (with $m \ge 1$); and $\Gamma, \Delta_0, \dots, \Delta_{m-1}$ are sets of clauses. Let $\sigma' \subseteq \sigma$ be such that $\sigma\backslash\sigma'$ consists only of ordinary binary predicates, and let φ° be the corresponding formula

$$\begin{aligned}&\bigwedge_{\zeta \in Z} \exists x.\zeta \wedge \forall x \forall y (x = y \vee [\Gamma \cup \Gamma^{-1}]^*_{/\sigma'}) \\ &\quad \wedge \bigwedge_{k=0}^{2} \bigwedge_{h=0}^{m-1} \forall x \exists y (\lambda_k \to (\lambda_{\lfloor k+1 \rfloor}(y) \wedge \mu_h(y) \wedge [\Delta_h \cup \Gamma \cup \Gamma^{-1}]^*_{/\sigma'})).\end{aligned}$$

Then $\models \varphi \to \varphi^\circ$, and, moreover, if φ° has a model over some domain A, then so has φ.

Proof. It is immediate that $\models \varphi \to \varphi^\circ$, by the validity of resolution. Now suppose $\mathfrak{A}$ is a structure such that $\mathfrak{A} \models \varphi^\circ$; we define a structure $\mathfrak{A}'$ over the same domain as $\mathfrak{A}$, such that $\mathfrak{A}' \models \varphi$. Fix $a \in A$ and h $(0 \le h < m)$. If a satisfies one (hence: exactly one) of the formulas $\lambda_0, \lambda_1, \lambda_2$, there exists b such that $\mathfrak{A} \models \lambda_{\lfloor k+1 \rfloor}[b]$, $\mathfrak{A} \models \mu_h[b]$ and $\mathfrak{A} \models [\Delta_h \cup \Gamma \cup \Gamma^{-1}]^*_{/\sigma'}[a, b]$. Let $\tau^- = \mathrm{tp}^{\mathfrak{A}}_{/\sigma'}[a, b]$. Since $[\Delta_h \cup \Gamma \cup \Gamma^{-1}]^*_{/\sigma'}$ involves no atoms of the forms $r(x, y)$ or $r(y, x)$, where $r \notin \sigma'$, we have $\models \tau^- \to [\Delta_h \cup \Gamma \cup \Gamma^{-1}]^*_{/\sigma'}$, and therefore, by Lemma 4.3, there is a 2-type τ such that $\models \tau \to \tau^-$ and $\models \tau \to (\Delta_h \cup \Gamma \cup \Gamma^{-1})$. So set the interpretations of the predicates in $\sigma\backslash\sigma'$ such that $\mathfrak{A}' \models \tau[a, b]$. Keeping a fixed, carry out the above procedure for all values of h, thus choosing m witnesses for a. Since, in each case, the chosen element b satisfies μ_h, these witnesses are all distinct, and so no clashes arise when setting 2-types in $\mathfrak{A}'$. Now carry out the above procedure for all values of a. If a satisfies λ_k, then any b chosen as a witness for a satisfies $\lambda_{\lfloor k+1 \rfloor}$, so that a could not previously have been chosen as a witness for b. Again, therefore, no clashes arise when setting 2-types in $\mathfrak{A}'$. At this stage, although $\mathfrak{A}'$ is not completely defined, we know that, however the construction of $\mathfrak{A}'$ is completed, for all $a \in A$ and k $(0 \leq k < 3)$ such that $\mathfrak{A}' \models \lambda_k[a]$, and all h $(0 \leq h < m)$, there will exist $b \in A \setminus \{a\}$ such that $\mathfrak{A}' \models (\Delta_h \wedge \Gamma \wedge \Gamma^{-1})[a, b]$. Finally, suppose a, b are distinct elements of A for which $\mathrm{tp}^{\mathfrak{A}'}[a, b]$ has not yet been defined, and let $\tau = \mathrm{tp}^{\mathfrak{A}}_{/\sigma'}[a, b]$. Since $\mathfrak{A} \models \varphi^\circ$, it follows that $\models \tau^- \rightarrow [\Gamma \cup \Gamma^{-1}]^*_{/\sigma'}$, and hence by Lemma 4.3 that there exists a 2-type τ such that $\models \tau \rightarrow \tau^-$ and $\models \tau \rightarrow (\Gamma \cup \Gamma^{-1})$. Again, set the interpretations of the predicates in $\sigma \setminus \sigma'$ so that $\mathfrak{A}' \models \tau[a, b]$; and repeat the process until $\mathfrak{A}'$ is completely defined. At the end of this process, for any distinct a, b of A, $\tau \models \Gamma[a, b]$. Thus, $\mathfrak{A}' \models \varphi$, as required. □

Theorem 4.5 *Let φ be an $\mathcal{L}^2$1PO-formula in standard normal form with multiplicity m over a signature σ. If φ has a finite model, then it has a model of size bounded by a doubly exponential function of $|\sigma| + m$. Hence, any finitely satisfiable $\mathcal{L}^2$1PO-formula φ has a model of size bounded by a doubly exponential function of $\|\varphi\|$, and so FinSat($\mathcal{L}^2$1PO) is in* 2-NEXPTIME.

Proof. We prove the first statement of the theorem. The remainder then follows by Lemma 2.1. By Lemma 4.2, let φ^* be an $\mathcal{L}^2$1PO-formula in spread normal form (7) with multiplicity $3m$ over a signature σ^* having the following properties: (i) $\models \varphi^* \rightarrow \varphi$; (ii) if φ has a (finite) model then so has φ^*; and (iii) $|\sigma^*|$ is polynomially bounded as a function of $|\sigma| + m$. By rewriting the sub-formulas $\eta, \vartheta_0, \ldots, \vartheta_{m-1}$ of φ^* in conjunctive normal form, we may assume that φ^* has the form required for Lemma 4.4. This re-writing will not affect the signature or multiplicity of φ^*. Let σ' be the signature obtained by removing all ordinary binary predicates from σ. Thus, σ' consists entirely of unary and navigational predicates. By Lemma 4.4, there is an $\mathcal{L}^2$1PO-formula φ° in weak normal form over the same signature as φ^*, having the same multiplicity, and satisfiable over the same domains, in which all occurrences of ordinary binary predicates are in atoms of the forms $r(x, x)$ or $r(y, y)$. Let φ' be the result of replacing any such atoms in φ° with the respective atoms $\hat{r}(x)$, $\hat{r}(y)$, where $\hat{r}$ is a fresh unary predicate for each ordinary binary predicate r. It is obvious that φ° and φ' are satisfiable over the same domains. Moreover, given that the formulas λ_0, λ_1 and λ_2 are mutually exclusive, we may insert the condition $x \neq y$ in all $\forall\exists$-conjuncts of φ'. Thus, φ' is an $\mathcal{L}^2$1PO$^\mathrm{u}$-formula in weak normal form over some signature σ' with multiplicity $m' = 3m$ such that $|\sigma'|$ is polynomially bounded as function of $|\sigma| + m$. By Theorem 3.24, if φ' has a finite model, then it has a model of size bounded by a doubly exponential function of $|\sigma'| + m'$. Therefore, φ has a model of size bounded by a doubly exponential function of $|\sigma| + m$. □

5 Two-variable logic with one transitive relation

The purpose of this section is to show that the logic $\mathcal{L}^2$1T has the triply exponential-sized finite model property (Theorem 5.7): if φ is a finitely satisfiable $\mathcal{L}^2$1T-formula, then φ has a model of size bounded by some fixed triply exponential function of $\|\varphi\|$. It follows that the finite satisfiability problem for $\mathcal{L}^2$1T is in 3-NEXPTIME. We proceed by reduction to the corresponding problem for standard normal-form $\mathcal{L}^2$1PO-formulas, but over signatures of exponential size, and with exponentially large multiplicities. Recall that, in $\mathcal{L}^2$1T, we have a distinguished binary predicate, $\mathfrak{t}$, which must be interpreted as a transitive relation. When speaking about 2-types, we take the assumed transitivity of $\mathfrak{t}$ into account: specifically, if a 2-type contains the literals $\mathfrak{t}(x, y)$ and $\mathfrak{t}(y, x)$, then it must also contain $\mathfrak{t}(x, x)$ and $\mathfrak{t}(y, y)$.

Let A be a set and T a transitive relation on A. A subset $B \subseteq A$ is *strongly connected* if, for all distinct $a, b \in B$, aTb. It is obvious that the maximal strongly-connected subsets of A form a partition: we refer to the cells of this partition as the T-*cliques* of A. If C is a T-clique of A and $|C| > 1$, then $T \supseteq C \times C$; if, however, $C = \{a\}$, then a may or may not be related to itself by T. If C and D are distinct T-cliques of A, then we write: (i) $C <_T D$ if, for all $a \in C$ and $b \in D$, aTb; (ii) $C >_T D$ if, for all $a \in C$ and $b \in D$, bTa; and (iii) $C \sim_T D$ if, for all $a \in C$ and $b \in D$, neither aTb nor bTa. It is routine to show:

Lemma 5.1 *Let A be a set and T a transitive relation on A. Then the relation $<_T$ is a partial order on the set of T-cliques of A. Moreover, if C and D are distinct T-cliques, then $C <_T D$ if and only if $D >_T C$ and, furthermore, exactly one of $C <_T D$, $C >_T D$ and $C \sim_T D$ obtains.*

If $\mathfrak{A}$ is a structure interpreting a distinguished binary predicate $\mathfrak{t}$ as a transitive relation over a domain A, we refer to the $\mathfrak{t}^{\mathfrak{A}}$-cliques, simply, as the *cliques* of $\mathfrak{A}$. We employ the following abbreviations:

$$\begin{aligned} &\mathfrak{t}_{\equiv}(x, y) \equiv \mathfrak{t}(x, y) \wedge \mathfrak{t}(y, x) \wedge x \neq y && \mathfrak{t}_{<}(x, y) \equiv \mathfrak{t}(x, y) \wedge \neg\mathfrak{t}(y, x) \\ &\mathfrak{t}_{\sim}(x, y) \equiv \neg\mathfrak{t}(x, y) \wedge \neg\mathfrak{t}(y, x) \wedge x \neq y && \mathfrak{t}_{>}(x, y) \equiv \neg\mathfrak{t}(x, y) \wedge \mathfrak{t}(y, x). \end{aligned} \tag{11}$$

It is then easy to see that the following validity holds:

$$\models \forall x \forall y (x = y \vee \mathfrak{t}_{\equiv}(x, y) \vee \mathfrak{t}_{<}(x, y) \vee \mathfrak{t}_{>}(x, y) \vee \mathfrak{t}_{\sim}(x, y)). \tag{12}$$

A formula of $\mathcal{L}^2 1\mathrm{T}$ is said to be in *transitive normal form* if it conforms to the pattern

$$\bigwedge_{\mathfrak{s} \in \{\equiv, <, >, \sim\}} \forall x \forall y (\mathfrak{t}_{\mathfrak{s}}(x, y) \to \eta_{\mathfrak{s}}) \ \wedge \bigwedge_{h=0}^{m-1} \bigwedge_{\mathfrak{s} \in \{\equiv, <, >, \sim\}} \forall x \exists y (p_{h,\mathfrak{s}}(x) \to (\mathfrak{t}_{\mathfrak{s}}(x, y) \wedge \vartheta_{h,\mathfrak{s}})). \tag{13}$$

where $m \geq 1$, the $p_{h,\mathfrak{s}}$ are unary predicates, and the $\eta_{\mathfrak{s}}$ and $\vartheta_{h,\mathfrak{s}}$ quantifier- and equality-free formulas not featuring either of the atoms $\mathfrak{t}(x, y)$ or $\mathfrak{t}(y, x)$. Transitive normal form is—modulo trivial logical manipulation—a special case of standard normal form (1). Note that, in the above definition, the sub-formulas $\eta_{\mathfrak{s}}$ and $\vartheta_{h,\mathfrak{s}}$ may contain the atoms $\mathfrak{t}(x, x)$ or $\mathfrak{t}(y, y)$.

We have the following normal-form theorem for $\mathcal{L}^2 1\mathrm{T}$.

Lemma 5.2 *Let φ be an $\mathcal{L}^2 1T$-formula. There exists an $\mathcal{L}^2 1T$-formula φ^* in transitive normal form such that: (i) $\models \varphi^* \to \varphi$; (ii) every model of φ can be expanded to a model of φ^*; and (iii) $\|\varphi^*\|$ is bounded by a polynomial function of $\|\varphi\|$.*

Proof. By Lemma 2.1, we may without loss of generality assume φ to be in standard normal form:

$$\forall x \forall y (x = y \vee \eta) \wedge \bigwedge_{h=0}^{m-1} \forall x \exists y (x \neq y \wedge \vartheta_h).$$

where $\eta, \vartheta_0, \ldots, \vartheta_{m-1}$ are equality- and quantifier-free. Suppose $\mathfrak{A} \models \varphi$. For all h ($0 \leq h < m$) and all $\mathfrak{s} \in \{\equiv, <, >, \sim\}$, let $p_{h,\mathfrak{s}}$ be a fresh unary predicate, and expand $\mathfrak{A}$ to an interpretation $\mathfrak{A}'$ by setting $\mathfrak{A}' \models p_{h,\mathfrak{s}}[a]$ if there exists $b \in A \setminus \{a\}$ such that $\mathfrak{A} \models \mathfrak{t}_{\mathfrak{s}}[a, b]$ and $\mathfrak{A} \models \vartheta_h[a, b]$. Further, set $\vartheta_{h,\mathfrak{s}}$ to be the result of replacing all atoms of the forms $\mathfrak{t}(x, y)$ or $\mathfrak{t}(y, x)$ in ϑ_h by either $\top$ or $\bot$ as specified by $\mathfrak{t}_{\mathfrak{s}}(x, y)$. Thus, setting ω to be the formula

$$\bigwedge_{h=0}^{m-1} \bigwedge_{\mathfrak{s} \in \{\equiv, <, >, \sim\}} \forall x (p_{h,\mathfrak{s}}(x) \to \exists y (\mathfrak{t}_{\mathfrak{s}}(x, y) \wedge \vartheta_{h,\mathfrak{s}})),$$

we see by construction of $\mathfrak{A}'$ that $\mathfrak{A}' \models \omega$. Observe that none of the $\vartheta_{h,\mathfrak{s}}$ contains either of the atoms $\mathfrak{t}(x, y)$ or $\mathfrak{t}(y, x)$. Let ψ_1' be the formula

$$\forall x \bigwedge_{h=0}^{m-1} \bigvee_{\mathfrak{s} \in \{\equiv, <, >, \sim\}} p_{h,\mathfrak{s}}(x).$$

Since $\mathfrak{A} \models \bigwedge_{h=0}^{m-1} \forall x \exists y (x \neq y \wedge \vartheta_h)$, and bearing in mind the validity (12), it follows that $\mathfrak{A}' \models \psi_1'$. Moreover, under our general assumption that all domains have cardinality at least 2, ψ_1' is logically equivalent to the formula ψ_1 given by:

$$\forall x \forall y \left(x = y \vee \left(\bigwedge_{h=0}^{m-1} \bigvee_{\mathfrak{s} \in \{\equiv, <, >, \sim\}} p_{h,\mathfrak{s}}(x) \right) \right).$$

For all $\mathfrak{s} \in \{\equiv, <, >, \sim\}$, let $\eta_\mathfrak{s}$ be the result of replacing all atoms of the forms $\mathfrak{t}(x, y)$ or $\mathfrak{t}(y, x)$ in η by either $\top$ or $\bot$ as specified by $\mathfrak{t}_\mathfrak{s}(x, y)$; and let ψ_2 be the formula

$$\mathfrak{A} \models \bigwedge_{\mathfrak{s} \in \{\equiv, <, >, \sim\}} \forall x \forall y (\mathfrak{t}_\mathfrak{s}(x, y) \to (x = y \vee \eta_\mathfrak{s})).$$

Since $\mathfrak{A} \models \forall x \forall y (x = y \vee \eta)$, we have $\mathfrak{A}' \models \psi_2$. Observe that none of the $\eta_\mathfrak{s}$ contains either of the atoms $\mathfrak{t}(x, y)$ or $\mathfrak{t}(y, x)$.

Let $\varphi^* = \psi_1 \wedge \psi_2 \wedge \omega$. Thus, $\|\varphi^*\|$ is bounded by a polynomial function of $\|\varphi\|$. We have shown that any model of φ can be expanded to a model of φ^*. Moreover, it follows easily from (12) that $\models \varphi^* \to \varphi$. □

The following lemma, taken from [8], gives us a simple way to replace a collection B of elements in some structure $\mathfrak{A}$ interpreting a purely relational signature σ with a 'small' set of elements B' in such a way that formulas of $\mathcal{L}^2$ do not notice the difference. context, We employ the following notation where $\mathfrak{A}$ is a structure and $B, B' \subseteq A$. We denote the set of 1-types realized over B by $\mathrm{tp}^{\mathfrak{A}}[B]$. Likewise, we denote the set of 2-types realized by pairs of elements $\langle b, b' \rangle$, where $b \in B$ and $b' \in B'$, by $\mathrm{tp}^{\mathfrak{A}}[B, B']$. (There is no requirement that B and B' be disjoint.) When B is a singleton, we write $\mathrm{tp}^{\mathfrak{A}}[b, B']$ in place of $\mathrm{tp}^{\mathfrak{A}}[\{b\}, B']$.

Lemma 5.3 ([8, Prop. 4]) *Let $\mathfrak{A}$ be a σ-structure not containing the distinguished predicate $\mathfrak{t}$, $B \subseteq A$, and $C := A \setminus B$. Then there is a σ-structure $\mathfrak{A}'$ with domain $A' = B' \cup C$ for some set B' of size exponential in $|\sigma|$, such that*

(i) $\mathfrak{A}'_{|C} = \mathfrak{A}_{|C}$.
(ii) $\mathrm{tp}^{\mathfrak{A}'}[B'] = \mathrm{tp}^{\mathfrak{A}}[B]$, *whence* $\mathrm{tp}^{\mathfrak{A}'}[A'] = \mathrm{tp}^{\mathfrak{A}}[A]$;
(iii) $\mathrm{tp}^{\mathfrak{A}'}[B', B'] = \mathrm{tp}^{\mathfrak{A}}[B, B]$ *and* $\mathrm{tp}^{\mathfrak{A}'}[B', C] = \mathrm{tp}^{\mathfrak{A}}[B, C]$, *whence* $\mathrm{tp}^{\mathfrak{A}'}[A', A'] = \mathrm{tp}^{\mathfrak{A}}[A, A]$;
(iv) for each $b' \in B'$ there is some $b \in B$ with $\mathrm{tp}^{\mathfrak{A}'}[b', A'] \supseteq \mathrm{tp}^{\mathfrak{A}}[b, A]$;
(v) for each $a \in C$: $\mathrm{tp}^{\mathfrak{A}'}[a, B'] \supseteq \mathrm{tp}^{\mathfrak{A}}[a, B]$.

The above Lemma applies to arbitrary structures (without any distinguished predicates). If, now, $\mathfrak{t}$ is a distinguished predicate required to be interpreted as a transitive relation, let us write $\mathrm{tp}^{\mathfrak{A}}_{<}[B, B']$ to denote the subset of 2-types $\beta \in \mathrm{tp}^{\mathfrak{A}}[B, B']$ such that $\models \beta \to \mathfrak{t}_<(x, y)$, and similarly for $\mathrm{tp}^{\mathfrak{A}}_{>}[B, B']$ and $\mathrm{tp}^{\mathfrak{A}}_{\sim}[B, B']$. The following two Lemmas, due to [13], allow us to replace any clique in a structure $\mathfrak{A}$ interpreting $\mathfrak{t}$ by an equivalent one of bounded size. The proofs were kindly supplied to the present author by the authors of [13] (private communication).

Lemma 5.4 *Let σ be a signature containing the distinguished transitive predicate $\mathfrak{t}$, $\mathfrak{A}$ a σ-structure, and $\mathbb{A}$ the set of cliques of $\mathfrak{A}$. Let $B \in \mathbb{A}$ and $C = A \setminus B$. Then there is a σ-structure $\mathfrak{A}'$ with domain $A' = B' \cup C$ for some set B', with $|B'|$ bounded exponentially in $|\sigma|$, such that (i)–(v) are as in Lemma* 5.3*, and the set of cliques of $\mathfrak{A}'$ is* $(\mathbb{A} \setminus \{B\}) \cup \{B'\}$.

Proof. If $|B| = 1$, then we simply put $B' = B$ and we are done. Otherwise, let u, $u_<$, $u_>$ and $u_\sim$ be fresh unary predicates. Let $\bar{\mathfrak{A}}$ be the expansion of $\mathfrak{A}$ obtained by setting $u^{\bar{\mathfrak{A}}} = B$ and

$$u_\mathfrak{s}^{\bar{\mathfrak{A}}} = \{a \in C : \quad \mathfrak{A} \models \mathfrak{t}_\mathfrak{s}[a, b] \text{ for some } (= \text{all}) \ b \in B\},$$

for $\mathfrak{s} \in \{<, >, \sim\}$; and now rename the distinguished predicate $\mathfrak{t}$ in $\bar{\mathfrak{A}}$ with an ordinary binary predicate—say—q_0. (Of course, even though q_0 is not a distinguished predicate, $q_0^{\bar{\mathfrak{A}}}$ is still a transitive relation.) Let the result of applying Lemma 5.3 to $\bar{\mathfrak{A}}$ and B be a structure $\bar{\mathfrak{A}}'$, in which B' is the replacement for B; and write A' for the domain of $\bar{\mathfrak{A}}'$. Notice that, if τ is a 2-type realized in $\bar{\mathfrak{A}}$ containing the literals $u_<(x)$ and $u(y)$, then τ also contains the literals $q_0(x, y)$ and $\neg q_0(y, x)$, and similarly, *mutatis mutandis*, with $u_<$ replaced by $u_>$ and $u_\sim$. But from property (iii) of Lemma 5.3, we have $\mathrm{tp}^{\bar{\mathfrak{A}}'}[A', A'] = \mathrm{tp}^{\bar{\mathfrak{A}}}[A, A]$. Hence, if $a \in C$ is such that $\mathfrak{A} \models \mathfrak{t}_<[a, b]$ for some (and hence all) $b \in B$, then $\bar{\mathfrak{A}}' \models q_0[a, b']$ and $\bar{\mathfrak{A}}' \not\models q_0[b', a]$ for all (and hence some) $b' \in B$, and similarly for $\mathfrak{t}_>$ and $\mathfrak{t}_\sim$. It is then obvious that $q_0^{\bar{\mathfrak{A}}}$ is a transitive relation with set of cliques $(\mathbb{A} \setminus \{B\}) \cup \{B'\}$, and indeed that the clique ordering induced by $\mathfrak{t}$ on $\mathfrak{A}$ and clique ordering induced by q_0 on $\bar{\mathfrak{A}}'$ are isomorphic under replacement of B by B'. Now let $\mathfrak{A}'$ be the structure obtained from $\bar{\mathfrak{A}}'$ by dropping the interpretations of u, $u_<$, $u_>$ and $u_\sim$ and renaming q_0 back to $\mathfrak{t}$. □

Now for the promised lemma allowing us to confine attention to models with small cliques.

Lemma 5.5 *Let φ be a (finitely) satisfiable $\mathcal{L}^2 1T$-sentence in transitive normal form over a signature σ. Then there exists a (finite) model of φ in which the size of each clique is bounded exponentially in $|\sigma|$.*

Proof. Let φ be as given in (13), and suppose $\mathfrak{A} \models \varphi$. Let $B \subseteq A$ be a clique of $\mathfrak{A}$, let $C = A \setminus B$, and let $\mathfrak{A}'$, with domain $A' = B' \cup C$, be the result of applying of Lemma 5.4 to $\mathfrak{A}$. We claim that $\mathfrak{A}' \models \varphi$. The universally quantified conjuncts of φ are true in $\mathfrak{A}'$ thanks to property (iii) of Lemma 5.4. As for the existential conjuncts, for any $c \in C$, properties (i) and (v) guarantee that c has all required witnesses. For any $b \in B'$, the same thing is guaranteed by property (iv). This establishes the claim.

Now let $\mathfrak{A}$ be a countable σ-structure. Let $I_1, I_2, \dots$ be a (possibly infinite) sequence of all cliques in a $\mathfrak{A}$, $\mathfrak{A}_0 = \mathfrak{A}$ and $\mathfrak{A}_{j+1}$ be the structure $\mathfrak{A}_j$ modified by replacing clique I_{j+1} by its small replacement I'_{j+1} as described above. We define the limit structure $\mathfrak{A}_\infty$ with the domain $I'_1 \cup I'_2, \dots$ such that for all k, ℓ the connections between I'_k and I'_l are defined in the same way as in $\mathfrak{A}_{max(k,\ell)}$. It is easy to see that $\mathfrak{A}_\infty \models \varphi$ and all cliques in $\mathfrak{A}_\infty$ are bounded exponentially in $|\sigma|$. □

We are now ready to prove the main result of this section: an exponential reduction of the (finite) satisfiability problem for $\mathcal{L}^2$1T to the (finite) satisfiability problem for standard normal form $\mathcal{L}^2$1PO-formulas.

Lemma 5.6 *Let φ be an $\mathcal{L}^2 1T$-formula. There exists an $\mathcal{L}^2$1PO-formula $\hat{\varphi}$ in standard normal form over a signature $\hat{\sigma}$ with multiplicity $\hat{m}$, such that: (i) if φ has a (finite) model with at least 2 cliques, then $\hat{\varphi}$ has a (finite) model; (ii) if $\hat{\varphi}$ has a model of size L, then φ has a model of size at most $n \cdot L$, where n is bounded by an exponential function of $\|\varphi\|$; and (iii) both $|\hat{\sigma}|$ and $\hat{m}$ are bounded by an exponential function of $\|\varphi\|$.*

Proof. Let σ be the signature of φ. The strategy of the proof is as follows. We use Lemma 5.5 to show that if φ has any (finite) model, then it has a (finite) model $\mathfrak{A}$ in which all cliques have size at most n, where n is bounded by an exponential function of $\|\varphi\|$. We therefore consider domains whose elements are themselves σ-structures consisting of single cliques of size at most n. We interpret over those domains a signature of: (i) unary predicates encoding these cliques, and (ii) binary predicates encoding structures consisting of pairs of such cliques. In this way, the structure $\mathfrak{A}$ has an encoding in which the cliques are represented by single elements (and pairs of cliques by pairs of elements). We construct a forumula $\hat{\varphi}$ that is satisfied by the encoding of $\mathfrak{A}$. Then we show, conversely, that, from any (finite) model $\mathfrak{B}$ of $\hat{\varphi}$, a (finite) model of φ can be constructed by replacing the elements of B by σ-structures consisting of single cliques of size at most n, and interpreting the predicates in the signature of φ according to the encoding given by $\mathfrak{B}$. Since, in the encoding process, cliques collapse to single elements, the transitive relation $\mathfrak{t}$ appearing in φ collapses to a partial order, so that $\hat{\varphi}$ is an $\mathcal{L}^2$1PO-formula. We now proceed to the details of the argument.

By Lemma 5.2, we may without loss of generality assume φ to be in transitive normal form:

$$\bigwedge_{\mathfrak{s}\in\{\equiv,<,>,\sim\}} \forall x \forall y (\mathfrak{t}_\mathfrak{s}(x, y) \to \eta_\mathfrak{s}) \wedge \bigwedge_{h=0}^{m-1} \bigwedge_{\mathfrak{s}\in\{\equiv,<,>,\sim\}} \forall x \exists y (p_{h,\mathfrak{s}}(x) \to (\mathfrak{t}_\mathfrak{s}(x, y) \wedge \vartheta_{h,\mathfrak{s}})).$$

From Lemma 5.5, we know that, if φ has a (finite) model, then it has one in which each clique is of size at most n, where n (≥ 1) is bounded by an exponential function of $|\sigma|$. Let $\mathbf{C} = \{c_1, \dots, c_n\}$ be some set of n objects. We call any set $C = \{c_1, \dots, c_m\}$ for some m ($1 \leq m \leq n$) an *initial segment* of $\mathbf{C}$. Say that a *cell* is a σ-structure $\mathfrak{C}$ whose domain C is an initial segment of $\mathbf{C}$ such that $\mathfrak{C}$ has exactly one clique, namely C itself. Here (and here only) we lift our usual assumption that all structures have cardinality at least 2, thus allowing cells with the singleton domain $\{c_1\}$. Enumerate the cells as $\mathfrak{C}_0, \dots, \mathfrak{C}_{M-1}$. Thus, M is bounded by a doubly exponential function of $|\sigma|$. Notice that, if $\mathfrak{C}$ is a cell containing more than 1 element, then by the transitivity of $\mathfrak{t}^\mathfrak{C}$, we have $\mathfrak{C} \models \forall x.\mathfrak{t}(x, x)$, and hence $\mathfrak{t}^\mathfrak{C} = C \times C$. On the other hand, if $C = \{c_1\}$, then we may have either $\mathfrak{C} \models \mathfrak{t}[c_1, c_1]$ or $\mathfrak{C} \not\models \mathfrak{t}[c_1, c_1]$. (It follows, incidentally, that $M \geq 2$.)

Now let $\mathbf{E} = \{e_1, \dots, e_n\}$ and $\mathbf{E}' = \{e'_1, \dots, e'_n\}$ be disjoint sets of cardinality n, and define the notion of an *initial segment* of these sets in the same way as for $\mathbf{C}$. Say that a *diatom* is a σ-structure $\mathfrak{D}$ with domain $D = E \cup E'$, where E is an initial segment of $\mathbf{E}$ and E' an initial segment of $\mathbf{E}'$ (not necessarily of the same cardinality), such that the set of cliques in $\mathfrak{D}$ is exactly $\{E, E'\}$. Note that, if there is a $\mathfrak{t}$-edge between any element in E and any element in E', then there is a $\mathfrak{t}$-edge between every element in E and every element in E'. Enumerate the diatoms as $\mathfrak{D}_0, \dots, \mathfrak{D}_{N-1}$. Thus, N is bounded by a doubly exponential function of $|\sigma|$. (On the other hand, $N \geq M \geq 2$.)

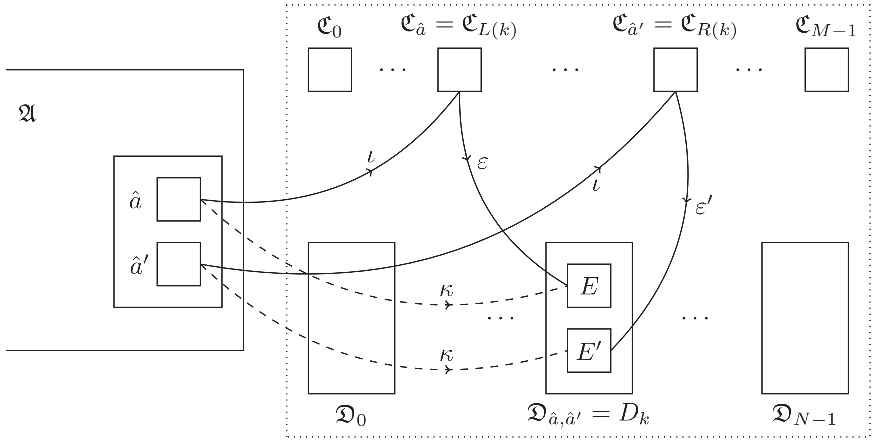


Fig. 7 The function κ mapping $\hat{a} \cup \hat{a}'$ to the reference diatom $\mathfrak{D} = \mathfrak{D}_{\hat{a},\hat{a}'}$. The construction of κ composes the function ι mapping $\hat{a}$ and $\hat{a}'$ to their respective reference cells $\mathfrak{C} = \mathfrak{C}_{\hat{a}}$ and $\mathfrak{C}' = \mathfrak{C}_{\hat{a}'}$ with the functions $\varepsilon : \mathbf{C} \to \mathbf{E}$ and $\varepsilon' : \mathbf{C} \to \mathbf{E}'$.

If C is an initial segment of $\mathbf{C}$ we define the mappings $\varepsilon : C \to \mathbf{E}$ and $\varepsilon' : C \to \mathbf{E}'$ by $\varepsilon(c_i) = e_i$ and $\varepsilon'(c_i) = e'_i$ for all i $(1 \leq i \leq |C|)$. Thus, if $\mathfrak{D} = \mathfrak{D}_k$ is a diatom with cliques $E \subseteq \mathbf{E}$ and $E' \subseteq \mathbf{E}'$, there exist unique cells $\mathfrak{C} = \mathfrak{C}_j$ and $\mathfrak{C}' = \mathfrak{C}_{j'}$ such that $\varepsilon : \mathfrak{C} \simeq \mathfrak{D}_{|E}$ and $\varepsilon' : \mathfrak{C}' \simeq \mathfrak{D}_{|E'}$. We refer to $\mathfrak{C}$ and $\mathfrak{C}'$ as the *left-* and *right-cells* of $\mathfrak{D}$, respectively, and, working with the corresponding indices, we define, for all k $(0 \leq k < N)$, $L(k) = j$ and $R(k) = j'$. Suppose now that we replaced the elements $e_1, e_2, \ldots$ of $E \subseteq \mathbf{E}$ with the corresponding elements $e'_1, e'_2, \ldots$ of $\mathbf{E}'$, and we replaced the elements $e'_1, e'_2, \ldots$ of $E' \subseteq \mathbf{E}'$ with the corresponding elements $e_1, e_2, \ldots$ of $\mathbf{E}$. The result would be another diatom, say, $\mathfrak{D}^{-1}$, obtained (in essence) by reversing the choice of which clique of $\mathfrak{D}$ defines the left-cell, and which the right-cell. We refer to $\mathfrak{D}^{-1} = \mathfrak{D}_{k'}$ as the *inverse* of $\mathfrak{D} = \mathfrak{D}_k$, and, working with the corresponding indices, we define, for all k $(0 \leq k < N)$, $I(k) = k'$. We introduce one final piece of terminology regarding diatoms. Recalling the abbreviations (11), consider any diatom $\mathfrak{D}$ with cliques $E \subseteq \mathbf{E}$ and $E' \subseteq \mathbf{E}'$. Evidently, for some $\mathfrak{s} \in \{<, >, \sim\}$, we have $\mathfrak{D} \models \mathfrak{t}_{\mathfrak{s}}[e, e']$ for all $e \in E$ and $e' \in E$. We call $\mathfrak{s}$ the *order-type* of $\mathfrak{D}$. Thus, the order type of $\mathfrak{D}$ is $>$ if and only if the order type of $\mathfrak{D}^{-1}$ is $<$, and the order type of $\mathfrak{D}$ is $\sim$ if and only if the order type of $\mathfrak{D}^{-1}$ is $\sim$. Working with the corresponding indices, we denote the order-type of $\mathfrak{D}_k$ by $\mathfrak{s}\langle k\rangle$.

Suppose φ has a model $\mathfrak{A}$ with at least two cliques, where no clique of $\mathfrak{A}$ has more than n elements. We proceed to construct an $\mathcal{L}^2$1PO-formula $\hat{\varphi}$ (depending only on φ, the signature σ, and the number n, but not on $\mathfrak{A}$ itself), together with a model $\hat{\mathfrak{A}}$ of $\hat{\varphi}$. To avoid confusion, we use the variables u and v in $\hat{\varphi}$ in place of x and y: it helps to think of u and v as ranging over the set of cliques of $\mathfrak{A}$. Let $\bar{p} = p_1, \ldots, p_s$ be a list of fresh unary predicates and $\bar{q} = q_1, \ldots, q_t$ a list of fresh binary predicates, where $\lceil s = \log M\rceil$ and $\lceil t = \log N\rceil$. Applying the same technique as employed in the proof of Lemma 4.2, we may form the labelling formulas $\bar{p}\langle j\rangle(u)$, for $0 \leq j < M$, and $\bar{q}\langle k\rangle(u, v)$ for $0 \leq k < N$. Now let $\hat{A}$ be the set of cliques of $\mathfrak{A}$, and for each $\hat{a} \in \hat{A}$, fix some (arbitrary) 1–1 function $\hat{a} \to C$, where C is the initial segment of $\mathbf{C}$ of cardinality $|\hat{a}|$. Denote by $\iota : A \to \mathbf{C}$ the union of all these functions. (In effect, ι orders the elements in each cell.) For any $\hat{a} \in \hat{A}$, the substructure $\mathfrak{A}_{|\hat{a}}$ is isomorphic, under ι, to some cell or other, say, $\mathfrak{C}_{\hat{a}}$, which we call the *reference cell* of $\hat{a}$. Now suppose that $\hat{a}, \hat{a}' \in \hat{A}$ are distinct, and let $\mathfrak{C}$ and $\mathfrak{C}'$ be their respective reference cells. (There is no requirement that $\mathfrak{C}$ and $\mathfrak{C}'$ be distinct.) Recalling the functions ε and ε' defined above, and setting $E = \varepsilon(C)$, $E' = \varepsilon'(C)$, define the function $\kappa : (\hat{a} \cup \hat{a}') \to E \cup E'$ (cf. Fig. 7) by

$$\kappa(a) = \begin{cases} \varepsilon(\iota(a)) & \text{if } a \in \hat{a}; \\ \varepsilon'(\iota(a)) & \text{otherwise (i.e., if } a \in \hat{a}'). \end{cases}$$

Evidently, κ defines an isomorphism from $\mathfrak{A}_{|(\hat{a}\cup\hat{a}')}$ to some diatom or other, say $\mathfrak{D}_{\hat{a},\hat{a}'}$, with cells E and E', which we call the *reference diatom* of the pair $\langle\hat{a}, \hat{a}'\rangle$. Observe that $\mathfrak{C}_{\hat{a}}$ is always the left-cell of $\mathfrak{D}_{\hat{a},\hat{a}'}$, and $\mathfrak{C}_{\hat{a}'}$ the right-cell. That is, if $\mathfrak{D}_{\hat{a},\hat{a}'} = \mathfrak{D}_k$, then $\mathfrak{C}_{\hat{a}} = \mathfrak{C}_{L(k)}$ and $\mathfrak{C}_{\hat{a}'} = \mathfrak{C}_{R(k)}$. Observe also that, if $\mathfrak{D} = \mathfrak{D}_{\hat{a},\hat{a}'}$, then $\mathfrak{D}^{-1} = \mathfrak{D}_{\hat{a}',\hat{a}}$. That is, if $\mathfrak{D}_{\hat{a},\hat{a}'} = \mathfrak{D}_k$, then $\mathfrak{D}_{\hat{a}',\hat{a}} = \mathfrak{D}_{I(k)}$.

Now let $\hat{\mathfrak{A}}$ be the structure over $\hat{A}$ with signature $\hat{\sigma} = \bar{p} \cup \bar{q} \cup \{<, >, \sim\}$, defined as follows.

1. For all $\hat{a} \in \hat{A}$, and all j $(0 \leq j < M)$, $\hat{\mathfrak{A}} \models \bar{p}\langle j\rangle[\hat{a}]$ if and only if $\mathfrak{C}_{\hat{a}} = \mathfrak{C}_j$;
2. for all distinct $\hat{a}, \hat{a}' \in \hat{A}$ and all k $(0 \leq k < N)$, $\hat{\mathfrak{A}} \models \bar{q}\langle k\rangle[\hat{a}, \hat{a}']$ if and only if $\mathfrak{D}_{\hat{a},\hat{a}'} = \mathfrak{D}_k$;
3. for all distinct $\hat{a}, \hat{a}' \in \hat{A}$, and all $\mathfrak{s} \in \{<, >, \sim\}$, $\hat{\mathfrak{A}} \models \mathfrak{s}(\hat{a}, \hat{a}')$ if and only if $\mathfrak{D}_{\hat{a},\hat{a}'}$ is of order-type $\mathfrak{s}$.

Under this interpretation, and taking the variables u and v to range over the cliques of $\mathfrak{A}$, the formula $\bar{p}\langle j\rangle(u)$ says "the reference cell $\mathfrak{C}_u$ of u is $\mathfrak{C}_j$," while the formula $\bar{q}\langle k\rangle(u, v)$ says "the reference diatom $\mathfrak{D}_{u,v}$ of $\langle u, v\rangle$ is $\mathfrak{D}_k$." Furthermore, by Lemma 5.1, $<^{\hat{\mathfrak{A}}}$ is a partial order, with $>^{\hat{\mathfrak{A}}}$ and $\sim^{\hat{\mathfrak{A}}}$ standing in the expected relations to $<^{\hat{\mathfrak{A}}}$. Since $\mathfrak{A}$ by assumption has at least two cliques, $\hat{\mathfrak{A}}$ does not violate our general assumption that all structures have cardinality at least 2.

We now proceed to define the sought-after formula $\hat{\varphi}$, building it up conjunct-by-conjunct, verifying, as we do so, that all these conjuncts are true in $\hat{\mathfrak{A}}$. We begin with the conjuncts depending only on σ and n. Let ψ_1' be the formula

$$\forall u \bigvee_{j=0}^{M-1} \bar{p}\langle j\rangle(u) \wedge \forall u \forall v \left(u = v \vee \bigvee_{k=0}^{N-1} \bar{q}\langle k\rangle(u, v) \right).$$

Under the interpretation $\hat{\mathfrak{A}}$, we may read ψ_1' as saying: "Every clique of $\mathfrak{A}$ has some reference cell, and every pair of distinct cliques has some reference diatom." This is obviously true by construction. Under the general assumption that all domains have cardinality at least 2, ψ_1' is equivalent to the formula ψ_1 given by

$$\forall u \forall v \left(u = v \vee \bigvee_{j=0}^{M-1} \bar{p}\langle j\rangle(u) \right) \wedge \forall u \forall v \left(u = v \vee \bigvee_{k=0}^{N-1} \bar{q}\langle k\rangle(u, v) \right),$$

so that $\hat{\mathfrak{A}} \models \psi_1$. Now let ψ_2 be the formula

$$\bigwedge_{k=0}^{N-1} \forall u \forall v \left(u = v \vee (\bar{q}\langle k\rangle(u, v) \rightarrow \bar{p}\langle L(k)\rangle(u) \wedge \bar{p}\langle R(k)\rangle(v)) \right).$$

Under the interpretation $\hat{\mathfrak{A}}$, we may read ψ_2 as stating that, if u and v are distinct cliques of $\mathfrak{A}$ such that $\mathfrak{D}_{u,v} = \mathfrak{D}_k$, then $\mathfrak{C}_u = \mathfrak{C}_{L(k)}$ and $\mathfrak{C}_v = \mathfrak{C}_{R(k)}$. Now let ψ_3 be the formula

$$\bigwedge_{k=0}^{N-1} \forall u \forall v \left(u = v \vee (\bar{q}\langle k\rangle(u, v) \rightarrow \bar{q}\langle I(k)\rangle(v, u)) \right).$$

Under the interpretation $\hat{\mathfrak{A}}$, we may read ψ_3 as stating that, if u and v are distinct cliques of $\mathfrak{A}$ such that $\mathfrak{D}_{u,v} = \mathfrak{D}_k$, then $\mathfrak{D}_{v,u} = \mathfrak{D}_{I(k)}$. Further, let ψ_4 be the formula

$$\bigwedge_{k=0}^{N-1} \forall u \forall v \left(u = v \vee (\bar{q}\langle k\rangle(u, v) \rightarrow \mathfrak{s}\langle k\rangle(u, v)) \right).$$

Under the interpretation $\hat{\mathfrak{A}}$, we may read ψ_4 as stating that, if u and v are distinct cliques of $\mathfrak{A}$ such that $\mathfrak{D}_{u,v} = \mathfrak{D}_k$, then the order-type of $\mathfrak{D}_{u,v}$ is $\mathfrak{s}\langle k\rangle$. Again, we have already observed that all these statements are true. Thus, $\hat{\mathfrak{A}} \models \psi_2 \wedge \psi_3 \wedge \psi_4$.

This completes the list of conjuncts of $\hat{\varphi}$ depending only on σ and n. We now turn our attention to those conjuncts of $\hat{\varphi}$ corresponding to the various conjuncts of φ. We consider first the $\forall\forall$-conjuncts of φ. Let $\lambda(u)$ abbreviate the formula

$$\bigvee \{ \bar{p}\langle j\rangle(u) \mid 0 \leq j < M,\ \mathfrak{C}_j \models \forall x \forall y (x = y \vee \eta_{\equiv}(x, y)) \}.$$

Under the interpretation $\hat{\mathfrak{A}}$, we may read $\lambda(u)$ as "u is a clique of $\mathfrak{A}$ in which the formula $\eta_{\equiv}(x, y)$ is satisfied by all pairs of distinct elements." For each $\mathfrak{s} \in \{<, >, \sim\}$, let $\hat{\eta}_{\mathfrak{s}}(u, v)$ abbreviate the formula

$$\bigvee\{\bar{q}\langle k\rangle(u, v) \mid 0 \leq k < N,\ \mathfrak{D}_k \models \forall x \forall y(\mathfrak{t}_{\mathfrak{s}}(x, y) \rightarrow \eta_{\mathfrak{s}}(x, y))\}.$$

We may read $\hat{\eta}_{\mathfrak{s}}(u, v)$ as "u and v are a pair of cliques in which the formula $\eta_{\mathfrak{s}}(x, y)$ is satisfied by all pairs of elements related by $\mathfrak{t}_{\mathfrak{s}}$." Now let ψ_5' be the formula

$$\forall u.\lambda(u) \wedge \bigwedge_{\mathfrak{s} \in \{<,>,\sim\}} \forall u \forall v(x = y \vee \hat{\eta}_{\mathfrak{s}}(u, v)).$$

Under the interpretation $\hat{\mathfrak{A}}$, we may read the first conjunct of ψ_5' as stating: "if $\mathfrak{C}$ is a cell realized in $\mathfrak{A}$, then any pair of distinct elements in $\mathfrak{C}$ satisfies $\eta_{\equiv}(x, y)$." The truth of this statement follows from the fact that $\mathfrak{A} \models \forall x \forall x(\mathfrak{t}_{\equiv}(x, y) \rightarrow \eta_{\equiv}(x, y))$. Similarly, the remaining conjuncts state: "if $\mathfrak{D}$ is a diatom realized in $\mathfrak{A}$ having order-type $\mathfrak{s}$ then any pair of elements ordered by $\mathfrak{t}_s$ satisfies $\eta_{\mathfrak{s}}$." The truth of this statement follows from the fact that $\mathfrak{A} \models \forall x \forall x(\mathfrak{t}_{\mathfrak{s}}(x, y) \rightarrow \eta_{\mathfrak{s}}(x, y))$. Again, replacing ψ_5' with the equivalent formula ψ_5 given by

$$\forall u \forall v(u = v \vee \lambda(u)) \wedge \bigwedge_{\mathfrak{s} \in \{<,>,\sim\}} \forall u \forall v(x = y \vee \hat{\eta}_{\mathfrak{s}}(u, v)),$$

we see that, $\hat{\mathfrak{A}} \models \psi_5$.

Now we consider the $\forall\exists$-conjuncts of φ. For each h $(0 \leq h < m)$, let $\mu_h(u)$ abbreviate the formula

$$\bigvee\{\bar{p}\langle j\rangle(u) \mid 0 \leq j < M,\ \mathfrak{C}_j \models \forall x(p_{h,\equiv}(x) \rightarrow \exists y(x \neq y \wedge \vartheta_{h,\equiv}(x, y)))\}.$$

Under the interpretation $\hat{\mathfrak{A}}$, we may read $\mu_h(u)$ as "u is a clique isomorphic to some cell $\mathfrak{C}$ such that $\mathfrak{C} \models \forall x(p_{h,\equiv}(x) \rightarrow \exists y(x \neq y \wedge \vartheta_{h,\equiv}(x, y)))$." Now let ψ_6' be the formula

$$\forall u \bigwedge_{h=0}^{m-1} \mu_h(u).$$

Under the interpretation $\hat{\mathfrak{A}}$, we may read ψ_6' as stating: "if $\mathfrak{C}$ is a cell realized in $\mathfrak{A}$, then any element in $\mathfrak{C}$ satisfying $p_{h,\equiv}(x)$ has a witness for $\exists y(x \neq y \wedge \vartheta_{h,\equiv}(x, y))$ in $\mathfrak{C}$. The truth of this statement follows from the fact that $\mathfrak{A} \models \forall x(p_{h,\equiv}(x) \rightarrow \exists y(\mathfrak{t}_{\equiv}(x, y) \wedge \vartheta_{h,\equiv}(x, y)))$. Replacing ψ_6' with ψ_6, given by

$$\forall u \forall v \left(u = v \vee \bigwedge_{h=0}^{m-1} \mu_h(u)\right),$$

we thus have $\hat{\mathfrak{A}} \models \psi_6$. Further, for each h $(0 \leq h < m)$, each $\mathfrak{s} \in \{<, >, \sim\}$ and each i $(1 \leq i \leq n)$, let $\nu_{h,\mathfrak{s},i}(u)$ abbreviate the formula

$$\bigvee\{\bar{p}\langle j\rangle(u) \mid 0 \leq j < M,\ \mathfrak{C}_j \models p_{h,\mathfrak{s}}[c_i]\}.$$

We may read $\nu_{h,\mathfrak{s},i}(u)$ as "u is a clique whose reference cell $\mathfrak{C}_u$ is such that $\mathfrak{C}_u \models p_{h,\mathfrak{s}}[c_i]$." (We take the statement "$\mathfrak{C}_u \models p_{h,\mathfrak{s}}[c_i]$" to be false if c_i is not in the domain of $\mathfrak{C}_u$.) Finally, for each h $(0 \leq h < m)$, each $\mathfrak{s} \in \{<, >, \sim\}$, each i $(1 \leq i \leq n)$ and each i' $(1 \leq i' \leq n)$, let $\xi_{h,\mathfrak{s},i,i'}(u, v)$ be the formula

$$\bigvee\{\bar{q}\langle k\rangle(u, v) \mid 0 \leq k < N,\ \mathfrak{s}\langle k\rangle = \mathfrak{s} \text{ and } \mathfrak{D}_k \models \vartheta_{h,\mathfrak{s}}[c_i, c'_{i'}]\}.$$

We may read $\xi_{h,\mathfrak{s},i,i'}(u, v)$ as "u and v are cliques whose reference diatom $\mathfrak{D}_{u,v}$ has order-type $\mathfrak{s}$ and is such that $\mathfrak{D}_{u,v} \models \vartheta_{h,\mathfrak{s}}[c_i, c'_{i'}]$." (We take the statement "$\mathfrak{D}_{u,v} \models \vartheta_{h,\mathfrak{s}}[c_i, c'_{i'}]$" to be false if c_i or c'_i are not in the domain of $\mathfrak{D}_{u,v}$.) Now let ω be the conjunction

$$\bigwedge_{\mathfrak{s} \in \{<,>,\sim\}} \bigwedge_{i=1}^{n} \bigwedge_{h=0}^{m-1} \forall u \exists v \left(u \neq v \wedge \left(\nu_{h,\mathfrak{s},i}(u) \rightarrow \bigvee_{i'=1}^{n} \xi_{h,\mathfrak{s},i,i'}(u, v)\right)\right).$$

Under the interpretation $\hat{\mathfrak{A}}$, ω states: "for all $\mathfrak{s}$ and h, if u is an $\mathfrak{A}$-clique with reference cell $\mathfrak{C}$ such that some element a of C satisfies $p_{h,\mathfrak{s}}(x)$, then there is some other $\mathfrak{A}$-clique v of such that a has a witness for $\exists y(\mathfrak{t}_\mathfrak{s}(x, y) \wedge \vartheta_{h,\mathfrak{s}}(x, y))$ in $\mathfrak{A}_{|(u \cup v)}$." The truth of this statement follows from the fact that $\mathfrak{A} \models \forall x(p_{h,\mathfrak{s}}(x) \rightarrow \exists y(\mathfrak{t}_\mathfrak{s}(x, y) \wedge \vartheta_{h,\mathfrak{s}}(x, y)))$.

Now let $\hat{\varphi} = \psi_1 \wedge \cdots \wedge \psi_6 \wedge \omega$. Thus, $\hat{\varphi}$ is an $\mathcal{L}^2$1PO-formula in standard normal form over a signature $\hat{\sigma}$ consisting of the unary predicates $p_1, \ldots, p_s$, the ordinary binary predicates $q_1, \ldots, q_t$ and the navigational predicates $<$, $>$ and $\sim$, with multiplicity $\hat{m} = 3mn$. We see that both $|\hat{\sigma}|$ and $\hat{m}$ are bounded by an exponential function of $\|\varphi\|$. Moreover, we have shown that $\hat{\varphi}$ has the model $\hat{\mathfrak{A}}$, where $<$ is interpreted as the partial order $<_T$ on the cliques of $\mathfrak{A}$, and $>$ and $\sim$ stand in the usual relations to $<$. It is obvious that $\hat{\mathfrak{A}}$ is finite if $\mathfrak{A}$ is. This establishes conditions (i) and (iii) of the lemma.

To establish condition (ii), we show that, if $\hat{\varphi}$ has a model of size $L \geq 2$, then φ has a model of size at most $n \cdot L$. Suppose then that $\mathfrak{B} \models \hat{\varphi}$, with $|B| = L$. Consider any element $b \in B$. From ψ_1 there exists j $(0 \leq j < M)$ such that $\mathfrak{B} \models \bar{p}\langle j\rangle[b]$, so let $\mathfrak{C}_b$ be a fresh copy of the cell $\mathfrak{C}_j$, having domain, say, $\check{B}_b$. Let $\check{B} = \bigcup_{b \in B} \check{B}_b$, and define a structure $\check{\mathfrak{B}}$ over $\check{B}$ as follows. For all $b \in B$, let $\check{\mathfrak{B}}_{|\check{B}_b} = \mathfrak{C}_b$, so that it remains only to define the 2-types involving elements from different sets $\check{B}_b$. Suppose $b, c \in B$ are distinct. From ψ_1 again, there exists k $(0 \leq k < N)$ such that $\mathfrak{B} \models \bar{q}\langle k\rangle[b, c]$. Now set $\check{B}_{|(\check{B}_b \cup \check{B}_c)} = \mathfrak{D}_k$. That these assignments do not clash with the structures $\mathfrak{C}_b$ already established is immediate from ψ_2. That these assignments do not clash with each other is immediate from ψ_3. This completes the construction of $\check{\mathfrak{B}}$. Obviously, $|\check{B}| \leq n \cdot L$. From ψ_5, we have

$$\check{\mathfrak{B}} \models \bigwedge_{\mathfrak{s} \in \{\equiv, <, >, \sim\}} \forall x \forall y(\mathfrak{t}_\mathfrak{s}(x, y) \rightarrow (x = y \vee \eta_\mathfrak{s})).$$

Likewise, from $\psi_6 \wedge \omega$, we have

$$\check{\mathfrak{B}} \models \bigwedge_{h=0}^{m-1} \bigwedge_{\mathfrak{s} \in \{\equiv, <, >, \sim\}} \forall x(p_{h,\mathfrak{s}}(x) \rightarrow \exists y(\mathfrak{t}_\mathfrak{s}(x, y) \wedge \vartheta_{h,\mathfrak{s}})).$$

That is, $\check{\mathfrak{B}} \models \varphi$, as required. It remains only to check that $\mathfrak{t}^{\check{\mathfrak{B}}}$ is transitive. By assumption, $<^{\mathfrak{B}}$ is a partial order. By construction, if a and a' are distinct elements of $\check{B}_b$, for some $b \in B$, then $\check{\mathfrak{B}} \models \mathfrak{t}[a, a']$. From ψ_4, if $a \in \check{B}_b$ and $a' \in \check{B}_{b'}$, where b and b' are distinct elements of B, then $\check{\mathfrak{B}} \models \mathfrak{t}[a, a']$ if and only if $\mathfrak{B} \models b < b'$. It is then obvious that $\mathfrak{t}^{\check{\mathfrak{B}}}$ is transitive. □

Theorem 5.7 *Any finitely satisfiable $\mathcal{L}^2 1T$-formula φ has a model of size bounded by a triply exponential function of $\|\varphi\|$, and so FinSat($\mathcal{L}^2 1T$) is in* 3-NEXPTIME.

Proof. Let φ be a formula of $\mathcal{L}^2$1T. Recalling our general assumption that all structures have cardinality at least 2, any model of φ consisting of a single clique is one in which $\mathfrak{t}$ is total. Thus, we may test satisfiability of φ in single-clique structures by replacing all $\mathfrak{t}$-atoms by $\top$, and considering the resulting $\mathcal{L}^2$-formula. Since any satisfiable $\mathcal{L}^2$-formula φ' has a model of cardinality bounded by an exponential function of $\|\varphi'\|$, the result is established. Thus, we may confine our attention to determining whether φ has a finite model with at least 2 cliques.

By Lemma 5.6, let $\hat{\varphi}$ be an $\mathcal{L}^2$1PO-formula in standard normal form with multiplicity $\hat{m}$ over a signature $\hat{\sigma}$, such that: (i) if φ has a finite model with at least 2 cliques, then $\hat{\varphi}$ is finitely satisfiable, (ii) if $\hat{\varphi}$ has a model of size L, then φ has a model of size $n \cdot L$, where n is bounded by an exponential function of $\|\varphi\|$; and (iii) both $|\hat{\sigma}|$ and $\hat{m}$ are bounded by an exponential function of $\|\varphi\|$. By Theorem 4.5, if φ, and therefore $\hat{\varphi}$, is finitely satisfiable, then $\hat{\varphi}$ has a model of size L bounded by a doubly exponential function of $|\hat{\sigma}| + \hat{m}$, and hence by a triply exponential function of $\|\varphi\|$, whence φ also has a model of size bounded by a triply exponential function of $\|\varphi\|$. □

Acknowledgements The author would like to thank Wiesław Szwast, Lidia Tendera, and the anonymous referee for valuable comments and suggestions.

References

[1] S. Benaim, M. Benedikt, W. Charatonik, E. Kieronski, R. Lenhardt, F. Mazowiecki, and J. Worrell, Complexity of two-variable logic on finite trees, in: Automata, Languages, and Programming, 40th International Colloquium, ICALP

2013, Riga, Latvia, July 8–12, 2013, Proceedings, Part II, edited by F. Fomin, R. Freivalds, M. Kwiatkowska, and D. Peleg, Lecture Notes in Computer Science Vol. 7966 (Springer, 2013), pp. 74–88.
[2] E. Börger, E. Grädel, and Y. Gurevich, The Classical Decision Problem, Universitext (Springer, 1997).
[3] E. Grädel, P. Kolaitis, and M. Vardi, On the decision problem for two-variable first-order logic, Bull. Symb. Log. **3**(1), 53–69 (1997).
[4] E. Kieroński, The two-variable guarded fragment with transitive guards is 2EXPTIME-hard, in: Foundations of Software Science and Computational Structures, 6th International Conference, FOSSACS 2003 Held as Part of the Joint European Conference on Theory and Practice of Software, ETAPS 2003, Warsaw, Poland, April 7–11, 2003, Proceedings, edited by A. D. Gordon, Lecture Notes in Computer Science Vol. 2620 (Springer, 2003), pp. 299–312.
[5] E. Kieroński, Results on the guarded fragment with equivalence or transitive relations, in: Computer Science Logic, 19th International Workshop, CSL 2005, 14th Annual Conference of the EACSL, Oxford, UK, August 22–25, 2005, Proceedings, edited by L. Ong, Lecture Notes in Computer Science Vol. 3634 (Springer, 2005), pp. 309–324.
[6] E. Kieroński, Decidability issues for two-variable logics with several linear orders, in: Computer Science Logic, 25th International Workshop / 20th Annual Conference of the EACSL, CSL 2011, September 12–15, 2011, Bergen, Norway, Proceedings, edited by M. Bezem, Leibniz International Proceedings in Informatics Vol. 12 (Schloß Dagstuhl, 2011), pp. 337–351.
[7] E. Kieroński, J. Michalyszyn, I. Pratt-Hartmann, and L. Tendera, Two-variable first-order logic with equivalence closure, SIAM J. Comput. **43**(3), 1012–1063 (2014).
[8] E. Kieroński and M. Otto, Small substructures and decidability issues for first-order logic with two variables, in: 20th IEEE Symposium on Logic in Computer Science (LICS 2005), 26-29 June 2005, Chicago, IL, USA, Proceedings, (IEEE Computer Society, 2005), pp. 448–457.
[9] E. Kieroński and L. Tendera, On finite satisfiability of two-variable first-order logic with equivalence relations, in: Proceedings of the 24th Annual IEEE Symposium on Logic in Computer Science, LICS 2009, 11-14 August 2009, Los Angeles, CA, USA, (IEEE Computer Society, 2009), pp. 123–132.
[10] M. Otto, Two-variable first-order logic over ordered domains, J. Symb. Log. **66**, 685–702 (2001).
[11] T. Schwentick and T. Zeume, Two-variable logic with two order relations, Log. Methods Comput. Sci. **8**(1:15), 1–27 (2012).
[12] D. Scott, A decision method for validity of sentences in two variables, J. Symb. Log. **27**, 477 (1962).
[13] W. Szwast and L. Tendera, FO^2 with one transitive relation is decidable, in: 30th International Symposium on Theoretical Aspects of Computer Science, STACS 2013, February 27 – March 2, 2013, Kiel, Germany, edited by N. Portier and T. Wilke, Leibniz International Proceedings in Informatics Vol. 20 (Schloß Dagstuhl, 2013), pp. 317–328.
[14] T. Zeume and F. Harwath, Order-invariance of two-variable logic is decidable, in: Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science, LICS '16, New York, NY, USA, July 5–8, 2016, (ACM, 2016), pp. 807–816.

