The fluted fragment with transitive relations

DOI:
10.1016/j.apal.2021.103042

Citation for published version (APA):
Pratt-hartmann, I., & Tendera, L. (2022). The fluted fragment with transitive relations. *Annals of Pure and Applied Logic*, *173*(1), 103042. https://doi.org/10.1016/j.apal.2021.103042

Published in:
Annals of Pure and Applied Logic

The Fluted Fragment with Transitive Relations[⋆]

Ian Pratt-Hartmann[a], Lidia Tendera[b,*]

[a] *University of Opole, Poland/University of Manchester, UK*
[b] *University of Opole, Institute of Computer Science, Oleska 48, 45-052 Opole, Poland*

Abstract

The fluted fragment is a fragment of first-order logic (without equality) in which, roughly speaking, the order of quantification of variables coincides with the order in which those variables appear as arguments of predicates. It is known that this fragment has the finite model property. We consider extensions of the fluted fragment with various numbers of transitive relations, as well as the equality predicate. In the presence of one transitive relation (together with equality), the finite model property is lost; nevertheless, we show that the satisfiability and finite satisfiability problems for this extension remain decidable. We also show that the corresponding problems in the presence of two transitive relations (with equality) or three transitive relations (without equality) are undecidable, even for the two-variable sub-fragment.

Keywords: fluted logic, transitivity, satisfiability, decidability
2000 MSC: 03B25, 03B70

1. Introduction

The *fluted fragment*, here denoted $\mathcal{FL}$, is a fragment of first-order logic in which, roughly speaking, the order of quantification of variables coincides with the order in which those variables appear as arguments of predicates. The allusion is presumably architectural: we are invited to think of arguments of predicates as being 'lined up' in columns. The following formulas are sentences of $\mathcal{FL}$

$$\begin{array}{l}\text{No student admires every professor}\\ \forall x_1(\text{student}(x_1) \rightarrow \neg\forall x_2(\text{prof}(x_2) \rightarrow \text{admires}(x_1, x_2)))\end{array} \tag{1}$$

 This is a revised and substantially extended version of the MFCS 2019 paper [19].

[*] Corresponding author
Email addresses: ipratt@cs.man.ac.uk (Ian Pratt-Hartmann), tendera@uni.opole.pl (Lidia Tendera)

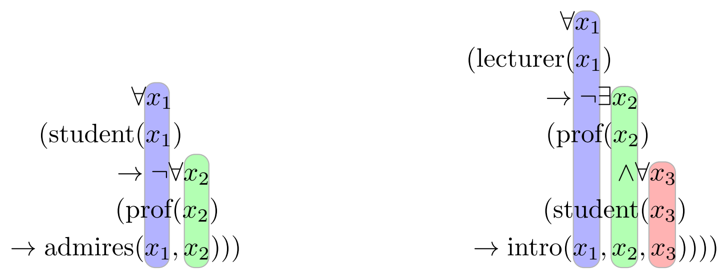

Figure 1: The 'lining up' of variables in the fluted formulas (1) and (2); all quantification is executed on the right-most available column.

$$\begin{aligned}&\text{No lecturer introduces any professor to every student}\\&\forall x_1(\text{lecturer}(x_1) \to\\&\qquad \neg\exists x_2(\text{prof}(x_2) \wedge \forall x_3(\text{student}(x_3) \to \text{intro}(x_1, x_2, x_3)))),\end{aligned} \tag{2}$$

with the 'lining up' of variables illustrated in Fig. 1. By contrast, none of the formulas

$$\forall x_1 . r(x_1, x_1) \tag{3}$$

$$\forall x_1 \forall x_2 (r(x_1, x_2) \to r(x_2, x_1)) \tag{4}$$

$$\forall x_1 \forall x_2 \forall x_3 (r(x_1, x_2) \wedge r(x_2, x_3) \to r(x_1, x_3)), \tag{5}$$

expressing, respectively, the reflexivity, symmetry and transitivity of the relation r, is fluted. If equality is present, then reflexivity can be expressed using a fluted sentence, since we may equivalently write (3) as, for example $\forall x_1 \forall x_2 (x_1 = x_2 \to r(x_1, x_2))$. On the other hand, no similar trick is available in the case of (4) or (5). For example, $\forall x_1 \forall x_2 (r(x_1, x_2) \to \exists x_3 (r(x_2, x_3) \wedge x_1 = x_3))$ is not fluted, as the equality atom violates the 'lining-up' of variables in the same way as does the atom $r(x_1, x_3)$ in (5).

The history of this fragment is somewhat tortuous. The basic idea of *fluted logic* can be traced to a paper given by W.V. Quine to the 1968 *International Congress of Philosophy* [22], in which the author defined the *homogeneous m-adic formulas*. Quine later relaxed this fragment, in the context of a discussion of predicate-functor logic, to what he called 'fluted' quantificational schemata [23], claiming that the satisfiability problem for the relaxed fragment is decidable. The viability of the proof strategy sketched by Quine was explicitly called into question by Noah [15], and the subject then taken up by W.C. Purdy [20], who gave his own definition of 'fluted formulas', proving decidability. It is questionable whether Purdy's reconstruction is faithful to Quine's intentions: the matter is clouded by differences between the definitions of predicate functors in Noah's and Quine's respective papers [15] and [23], both of which Purdy cites. In fact, Quine's original definition of 'fluted' quantificational schemata appears to coincide with a logic introduced—apparently independently—by A. Herzig [4]. Rightly or wrongly, however, the name 'fluted fragment' has now attached itself to Purdy's definition in [20]; and we shall continue to use it in that way in the present article. See Sec. 2 for a formal definition.

To complicate matters further, Purdy claimed in [21] that $\mathcal{FL}$ (i.e. the fluted fragment, in our sense, and his) has the exponential-sized model property: if a fluted formula φ is satisfiable, then it is satisfiable over a domain of size bounded by an exponential function of the number of symbols in φ. Purdy concluded that the satisfiability problem for $\mathcal{FL}$ is NExpTime-complete. These latter claims are false. It was shown in [17] that, although $\mathcal{FL}$ has the finite model property, there is no elementary bound on the sizes of the models required, and the satisfiability problem for $\mathcal{FL}$ is non-elementary. More precisely, define $\mathcal{FL}^m$ to be the subfragment of $\mathcal{FL}$ in which at most m variables (free or bound) appear. Then the satisfiability problem for $\mathcal{FL}^m$ is $\lfloor m/2 \rfloor$-NExpTime-hard for all $m \geq 2$ and in $(m-2)$-NExpTime for all $m \geq 3$ [18]. It follows that the satisfiability problem for $\mathcal{FL}$ is Tower-complete, in the framework of [24]. These results fix the exact complexity of satisfiability of $\mathcal{FL}^m$ for small values of m. Indeed, the satisfiability problem for FO^2, the two-variable fragment of first-order logic, is known to be NExpTime-complete [3], whence the corresponding problem for $\mathcal{FL}^2$ is certainly in NExpTime. Moreover, $\mathcal{FL}^1$ coincides with the 1-variable fragment of first-order logic, whence its satisfiability problem is NPTime-complete. Thus, taking 0-NExpTime to mean NPTime, we see that the satisfiability problem for $\mathcal{FL}^m$ is $\lfloor m/2 \rfloor$-NExpTime-complete, at least for $m \leq 4$.

The focus of the present paper is what happens when we add to the fluted fragment the ability to stipulate that certain designated binary relations are *transitive*, or are *equivalence relations*. The motivation comes from analogous results obtained for other decidable fragments of first-order logic. Consider basic propositional modal logic K. Under the standard translation into first-order logic (yielded by Kripke semantics), we can regard K as a fragment of first-order logic—indeed as a fragment of $\mathcal{FL}^2$. From basic modal logic K, we obtain the logic K4 under the supposition that the accessibility relation on possible worlds is transitive, and the logic S5 under the supposition that it is an equivalence relation: it is well-known that the satisfiability problems for K and K4 are PSpace-complete, whereas that for S5 is NPTime-complete [13]. (For analogous results on *graded* modal logic, see [6].) Closely related are also description logics (cf. [1]) with *role hierarchies* and *transitive roles*. In particular, the description logic $\mathcal{SH}$, which has the finite model property, is an ExpTime-complete fragment of $\mathcal{FL}^2$ with transitivity. Similar investigations have been carried out in respect of FO^2, which has the finite model property and whose satisfiability problem, as just mentioned, is NExpTime-complete. The finite model property is lost when one transitive relation or two equivalence relations are allowed. For equivalence, everything is known: the (finite) satisfiability problem for FO^2 in the presence of a single equivalence relation remains NExpTime-complete, but this increases to 2-NExpTime-complete in the presence of two equivalence relations [9, 10], and becomes undecidable with three. For transitivity, we have an incomplete picture: the *finite* satisfiability problem for FO^2 in the presence of a single transitive relation is decidable in 3-NExpTime [16], while the decidability of the satisfiability problem remains open (cf. [27]); the corresponding problems with two transitive relations are both undecidable [11]. The latter results can be easily adapted to show undecidability of the fluted fragment with unrestricted equality, where one can talk about transitivity of a binary relation.

Adding equivalence relations to the fluted fragment poses no new problems. Existing results on of FO^2 with two equivalence relations can be used to show that the satisfiability and finite satisfiability problems for $\mathcal{FL}$ (not just $\mathcal{FL}^2$) with *two* equivalence relations are decidable. Furthermore, the proof that the corresponding problems for FO^2 in the presence of *three* equivalence relations are undecidable can easily be seen to apply also to $\mathcal{FL}^2$. On the other hand, the situation with transitivity is less straightforward. We show in the sequel that the satisfiability and finite satisfiability problems for $\mathcal{FL}$ remain decidable in the presence of a single transitive relation and equality. (This logic lacks the finite model property.) On the other hand, the satisfiability and the finite satisfiability problems for $\mathcal{FL}$ in the presence of two transitive relations and equality, or indeed, in the presence of three transitive relations (but without equality) are all undecidable. For the fluted fragment with two transitive relations but *without* equality, the situation is not fully resolved. We show in the sequel that this fragment lacks the finite model property; this contrasts with the situation in description logics, where not only $\mathcal{SH}$ but also its extension $\mathcal{SHI}$ retain the finite model property, independently of the number of transitive relations [14]. However, the decidability of both satisfiability and finite satisfiability for this fragment remain open. Table 1 gives an overview of these results in comparison with known results on FO^2.

Some indication that flutedness interacts in interesting ways with transitivity is given by known complexity results on various extensions of guarded two-variable fragment with transitive relations. The *guarded fragment*, denoted GF, is that fragment of first-order logic in which all quantification is of either of the forms $\forall \bar{v}(\alpha \to \psi)$ or $\exists \bar{v}(\alpha \wedge \psi)$, where α is an atomic formula (a so-called *guard*) featuring all free variables of ψ. The *guarded two-variable fragment*, denoted GF^2, is the intersection of GF and FO^2. It is straightforward to show that the addition of two transitive relations to GF^2 yields a logic whose satisfiability problem is undecidable. However, as long as the distinguished transitive relations appear only in guards, we can extend the whole of GF with any number of transitive relations, yielding the so-called *guarded fragment with transitive guards*, whose satisfiability problem is in 2-ExpTime [26]. Intriguingly, in the two-variable case, we obtain a reduction in complexity if we require transitive relations in guards to point *forward*—i.e. allowing only $\forall v(t(u,v) \to \psi)$ rather than $\forall v(t(v,u) \to \psi)$, and similarly for existential quantification. These restrictions resemble flutedness, of course, except that they prescribe the order of variables only in *guards*, rather than in the whole formula. Thus, the extension of GF^2 with (any number of) transitive guards has a 2-ExpTime-complete satisfiability problem; however, the corresponding problem under the restriction to one-way transitive guards is ExpSpace-complete [8]. Since the above-mentioned extensions of GF^2 lack the finite model property, their satisfiability and the finite satisfiability problems do not coincide. Decidability and complexity bounds for the finite satisfiability problems are established in [11, 12].

Special symbols	**Decidability and Complexity**	
	$\mathcal{FL}^m$ ($m \geq 2$)	FO^2
no transitive r.	$\lfloor m/2 \rfloor$-NExpTime-hard in $(m-2)$-NExpTime$^{*)}$ [17, 18]	FMP NExpTime-compl. [3]
1 transitive r.	FMP [19] (*Fin*)*Sat*: in m-NExpTime	*Sat*: ? *FinSat*: in 3-NExpTime
1 transitive r. with =	*Sat*: in m-NExpTime **Theorem 21** *FinSat*: in $(m+1)$-NExpTime **Corollary 22**	*Sat*: ? *FinSat*: in 3-NExpTime [16]
2 transitive r.	*Sat*: ? *FinSat*: ?	undecidable [7, 5]
2 transitive r. with =	undecidable **Theorem 26**	undecidable
1 trans.&1 equiv. with =	undecidable **Corollary 27**	undecidable
3 transitive r.	undecidable *Sat*: **Theorem 38** *FinSat*: **Theorem 39**	undecidable
3 equivalence r.	undecidable **Corollary 40**	undecidable

Table 1: Overview of $\mathcal{FL}^m$ and FO^2 over restricted classes of structures. $^{*)}$ in case $m > 2$, and NExpTime-complete for $\mathcal{FL}^2$. Undecidability of extensions of FO^2 shown in grey were known earlier, but now can be inherited from remaining results of the Table.

2. Preliminaries

All signatures in this paper are purely relational, i.e., there are no individual constants or function symbols. We do, however, allow 0-ary relations (proposition letters). We use the notation $\varphi \dot{\vee} \psi$ to denote the exclusive disjunction of φ and ψ.

Let $\bar{x}_\omega = x_1, x_2, \ldots$ be a fixed sequence of variables. We define the sets of formulas $\mathcal{FL}^{[m]}$ (for $m \geq 0$) by structural induction as follows: (i) any non-equality atom $\alpha(x_\ell, \ldots, x_m)$, where $x_\ell, \ldots, x_m$ is a contiguous (possibly empty) subsequence of $\bar{x}_\omega$, is in $\mathcal{FL}^{[m]}$; (ii) $\mathcal{FL}^{[m]}$ is closed under boolean combinations; (iii) if φ is in $\mathcal{FL}^{[m+1]}$, then $\exists x_{m+1}\varphi$ and $\forall x_{m+1}\varphi$ are in $\mathcal{FL}^{[m]}$. The set of *fluted formulas* is defined as $\mathcal{FL} = \bigcup_{m \geq 0} \mathcal{FL}^{[m]}$. A *fluted sentence* is a fluted formula with no free variables. Thus, when forming Boolean combinations in the fluted fragment, there is some $m \geq 0$ such that each of the combined formulas has, as its free variables, some suffix of the sequence $x_1, \ldots, x_m$; and, when quantifying, only the last variable in this suffix may be bound. Note also that proposition letters (0-ary predicates) may, according to the above definitions, be combined freely with formulas: if φ is in $\mathcal{FL}^{[m]}$, then so, for example, is $\varphi \wedge P$, where P is a proposition letter. For $m \geq 0$, denote by $\mathcal{FL}^m$ the *m-variable sub-fragment* of $\mathcal{FL}$, i.e. the set of formulas of $\mathcal{FL}$ featuring at most m variables, free or bound. Do not confuse $\mathcal{FL}^m$ with $\mathcal{FL}^{[m]}$. For example, (1) is in $\mathcal{FL}^m$ just in case $m \geq 2$, and (2) is in $\mathcal{FL}^m$ just in case $m \geq 3$; but they are both in $\mathcal{FL}^{[0]}$. Note that $\mathcal{FL}^m$-formulas cannot, by force of syntax, feature predicates of arity greater than m. The fragments $\mathcal{FL}_=^{[m]}$, $\mathcal{FL}_=$ and $\mathcal{FL}_=^m$ are defined analogously, except that equality atoms $x_{m-1} = x_m$ are allowed in $\mathcal{FL}_=^{[m]}$ for $m \geq 2$. We call formulas of $\mathcal{FL}_=^{[0]}$ *sentences*, and those of $\mathcal{FL}_=^{[1]}$, *unary formulas*.

We denote by $\mathcal{FL}k\mathrm{T}$ the extension of $\mathcal{FL}$ with k distinguished binary predicates assumed to be interpreted as transitive relations; and we denote by $\mathcal{FL}_=k\mathrm{T}$ the corresponding extension of $\mathcal{FL}_=$. We denote their m-variable sub-fragments ($m \geq 2$) by $\mathcal{FL}^m k\mathrm{T}$, respectively $\mathcal{FL}_=^m k\mathrm{T}$. A predicate is called *ordinary* if it is neither the equality predicate nor one of the distinguished predicates.

If $\mathcal{L}$ is any logic, we denote its satisfiability problem by $Sat(\mathcal{L})$ and its finite satisfiability problem by $FinSat(\mathcal{L})$, understood in the usual way. If φ is a formula of any of the above fragments, $\|\varphi\|$ denotes the length of φ under any standard encoding, and similarly for finite sets of formulas.

2.1. Variable-free syntax for fluted formulas

Assuming, as we shall, that the arity of every predicate is fixed in advance, variables in fluted formulas carry no information, and therefore can be omitted. Thus, for example, sentences (1) and (2) can be written as follows

$$\begin{array}{l}\text{No student admires every professor}\\ \forall(\text{student} \rightarrow \neg\forall(\text{prof} \rightarrow \text{admires}))\end{array} \tag{6}$$

$$\begin{array}{l}\text{No lecturer introduces any professor to every student}\\ \forall(\text{lecturer} \rightarrow \neg\exists(\text{prof} \wedge \forall(\text{student} \rightarrow \text{intro}))),\end{array} \tag{7}$$

As an exercise, try converting (7) back into (2). The only ambiguity here comes from the choice of the highest-indexed variable; for example, the notation $\forall(\text{prof} \rightarrow \text{admires})$ can mean $\forall x_{m+1}(\text{prof}(x_{m+1}) \rightarrow \text{admires}(x_m, x_{m+1}))$ for any $m \geq 1$. However, such ambiguity is perfectly harmless, and in fact—as the present authors have found—rather convenient. Variable-free syntax for fluted formulas takes a little getting used to, but makes for a compact presentation; we shall standardly employ it in the sequel. We write $\forall^m$ to denote a block of m universal quantifiers; thus, if $\varphi \in \mathcal{FL}^{[m]}$, then $\forall^m \varphi \in \mathcal{FL}^{[0]}$. The elimination of variables seems to have been part of Quine's original motivation for introducing the fluted fragment (or at least one of its close relatives).

2.2. Loss of the finite model property

The logic $\mathcal{FL}1\mathrm{T}$ possesses the finite model property (see Table 1). However, this is no longer true if we add either equality or a second transitive relation, as shown by the examples below.

Example 1. *Consider the* $\mathcal{FL}^2_=1\mathrm{T}$*-sentence* $\varphi_1 = \forall\exists.T_1 \wedge \forall\forall(T_1 \rightarrow \neg =)$, *where* T_1 *is a distinguished binary predicate denoting a transitive relation. This sentence is satisfiable, but not finitely satisfiable.*

Proof. In standard first-order syntax, φ_1 reads as follows:

$$\varphi_1 = \forall x \exists y. T_1(x,y) \wedge \forall x \forall y (T_1(x,y) \rightarrow x \neq y).$$

It is obvious that φ_1 is satisfiable (for example by the structure $\mathbb{N}$ with T_1 interpreted as $<$), but not finitely satisfiable. $\square$

Example 2. *Consider the* $\mathcal{FL}^2 2\mathrm{T}$*-sentence*

$$\varphi_2 = \exists p_0 \wedge \forall(p_0 \dot{\vee} p_1 \dot{\vee} p_2) \wedge \forall\forall\neg(T_1 \wedge T_2) \wedge \bigwedge_{i=0,1,2} \forall\Big(p_i \rightarrow \big(\exists(p_{i+1} \wedge \neg(T_1 \vee T_2)) \;\wedge\; \forall(p_{i+2} \rightarrow T_1 \vee T_2)\big)\Big),$$

where the p_i $(0 \leq i \leq 2)$ *are unary predicates (addition in subscripts interpreted modulo* 3*), and* T_1*,* T_2 *are distinguished binary predicates denoting transitive relations. This sentence is satisfiable, but not finitely satisfiable.*

Proof. For readers still getting used to variable-free notation, we again restore the variables in φ_2:

$$\exists x_1. p_0(x_1) \wedge \forall x_1 \big(p_0(x_1) \dot{\vee} p_1(x_1) \dot{\vee} p_2(x_1)\big) \wedge \forall x_1 \forall x_2 \neg\big(T_1(x_1,x_2) \wedge T_2(x_1,x_2)\big) \wedge \bigwedge_{i=0,1,2} \forall x_1 \Big(p_i(x_1) \rightarrow \big(\exists x_2 (p_{i+1}(x_2) \wedge \neg(T_1(x_1,x_2) \vee T_2(x_1,x_2))) \wedge \forall x_2 (p_{i+2}(x_2) \rightarrow (T_1(x_1,x_2) \vee T_2(x_1,x_2)))\big)\Big).$$

One can easily check that the structure $\mathbb{N}$ with the following interpretation of the predicate letters

$$\begin{aligned} p_i[n] &\quad \text{iff} \quad n \mod 3 = i \\ T_1[n,m] &\quad \text{iff} \quad n+1 < m \\ T_2[n,m] &\quad \text{iff} \quad n > m \end{aligned}$$

is a model of φ_2.

To see that φ_2 is not finitely satisfiable, suppose it has a model, $\mathfrak{A}$. By the existential conjuncts of φ_2, there exist distinct elements $a_0, a_1, a_2 \in A$ such that a_i satisfies p_i $(0 \leq i < 3)$, and neither $\langle a_0, a_1 \rangle$ nor $\langle a_1, a_2 \rangle$ satisfy $T_1 \vee T_2$. The universal conjuncts of φ_2 imply that $\langle a_0, a_2 \rangle$, $\langle a_1, a_0 \rangle$ and $\langle a_2, a_1 \rangle$ satisfy $T_1 \vee T_2$ but not $T_1 \wedge T_2$. By the transitivity of $T_1^{\mathfrak{A}}$ and $T_2^{\mathfrak{A}}$, this allows for only two options: (i) $\mathfrak{A} \models T_1[a_1, a_0]$, $\mathfrak{A} \models T_1[a_2, a_1]$ and $\mathfrak{A} \models T_2[a_0, a_2]$; or (ii) $\mathfrak{A} \models T_2[a_1, a_0]$, $\mathfrak{A} \models T_2[a_2, a_1]$ and $\mathfrak{A} \models T_1[a_0, a_2]$ (shown in Figure 2). In both cases, applying transitivity of T_1 and T_2, we have $\mathfrak{A} \models (T_1 \vee T_2)[a_2, a_0]$. But then the existential conjuncts require a new witness, say a_3, for a_2 such that

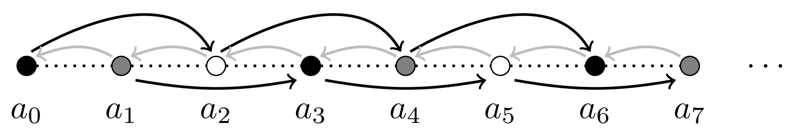

Figure 2: Infinite chain in models of φ_2 from Example 2. Pairs (a_i, a_{i+1}) are neither in T_1 nor in T_2; depicted by dotted lines. Black and grey arrows depict pairs belonging to the transitive relations T_1 and T_2.

$\mathfrak{A} \not\models (T_1 \vee T_2)[a_2, a_3]$. Again, taking the universal conjuncts into consideration, we see that a_3 must be related to a_0, a_1 and a_2 either by T_1 or by T_2 (in the case of Figure 2, it is T_2). So the situation repeats, and indeed $\mathfrak{A}$ embeds an infinite chain of elements such that, for each consecutive pair, $\mathfrak{A} \not\models (T_1 \vee T_2)[a_i, a_{i+1}]$. □

2.3. Fluted types and cliques

Suppose $\mathfrak{A}$ is a structure interpreting the distinguished binary predicate T as a transitive relation. A *clique* of $\mathfrak{A}$ is a maximal subset $B \subseteq A$ with the property that, for all distinct $a, b \in B$, $\mathfrak{A} \models T[a, b]$. Every element $a \in A$ is a member of exactly one clique, and if that clique has size greater than 1, then, necessarily $\mathfrak{A} \models T[a, a]$. Furthermore, if B_1 and B_2 are cliques, then either every element of B_1 is related to every element of B_2 by T, or no element of B_1 is related to any element of B_2 by T. In this way, $T^{\mathfrak{A}}$ induces a strict partial order on the set of cliques. If a singleton $\{a\}$ is a clique, then it may or may not be the case that $\mathfrak{A} \models T[a, a]$. If $\mathfrak{A} \models \neg T[a, a]$, then we call a (or sometimes $\{a\}$) a *soliton*.

In this paper, we adapt the familiar notions of atom, literal, and m-type to the fluted environment. Let Σ be a purely relational signature. A *fluted m-atom* (*over* Σ) is an atomic formula of $\mathcal{FL}^{[m]}_{=}$ featuring a predicate $p \in \Sigma \cup \{=\}$. Indeed, using variable-free syntax, we can simply say that a fluted m-atom *is* an element of $\Sigma \cup \{=\}$ having arity at most m. A *fluted m-literal* (*over* Σ) is a fluted m-atom (over Σ) or its negation; a *fluted m-type* (*over* Σ) is a maximal consistent conjunction of fluted m-literals (over Σ). (Observe that, if $m \geq 2$, then $=$ and $\neq$ are fluted literals.) If $\bar{a} = a_1, \ldots, a_m$ is a tuple of elements in some structure $\mathfrak{A}$ interpreting Σ, then $\bar{a}$ satisfies a unique fluted m-type over Σ, denoted $\mathrm{ftp}^{\mathfrak{A}}[\bar{a}]$. We silently identify fluted m-types with their conjunctions where appropriate; thus, any fluted m-type may be regarded as a (quantifier-free) $\mathcal{FL}^{[m]}$-formula. A fluted m-atom/literal is automatically a fluted m'-atom/literal for all $m' > m$.

3. Fluted logic with one transitive relation and equality

In this section, we study the logic $\mathcal{FL}_{=}1\mathrm{T}$, the fluted fragment with equality and a single transitive relation; we also consider its m-variable sub-fragment,

$\mathcal{FL}^m_= 1\mathrm{T} = \mathcal{FL}_= 1\mathrm{T} \cap \mathrm{FO}^m$, for all $m \geq 2$. As already mentioned, even the smallest of these fragments lacks the finite model property. Nevertheless, we show that the satisfiability problem for $\mathcal{FL}_= 1\mathrm{T}$ is decidable; indeed, $\mathit{Sat}(\mathcal{FL}^m_= 1\mathrm{T})$ is in m-NExpTime for $m \geq 2$. The proof is divided into three parts. In Sec. 3.1, we define a restricted class of $\mathcal{FL}^2_= 1\mathrm{T}$-formulas, called *basic formulas*, in which no binary predicates appear other than T and $=$. We give a procedure for determining satisfiability for finite sets of basic formulas, and give a parametrized analysis of its complexity. In Sec. 3.2, we show that $\mathit{Sat}(\mathcal{FL}^2_= 1\mathrm{T})$ is in 2-NExpTime, by reduction to the case considered in the previous section. In Sec. 3.3, we show that $\mathit{Sat}(\mathcal{FL}^m_= 1\mathrm{T})$ is in m-NExpTime, via a series of exponential-sized reductions to $\mathit{Sat}(\mathcal{FL}^2_= 1\mathrm{T})$.

It was shown in [18] that the satisfiability problem for $\mathcal{FL}^m$ (the m-variable fluted fragment without equality or transitive relations) is in $(m-2)$-NExpTime for $m \geq 3$. The strategy adopted there was to reduce this problem, via a series of exponential-sized reductions, to the satisfiability problem for $\mathcal{FL}^3$, at which point a direct procedure was described showing this last problem to be in NExpTime. In the presence of equality and transitivity, however, this direct procedure is unavailable, and we must carry out the further reduction to basic formulas as described in Sec. 3.2, at which point we can apply the procedure of Sec. 3.1. These final steps result in a weakening of the complexity bound by two exponentials.

We will be dealing here with logics featuring a single distinguished transitive relation, and we use the letter T for the corresponding binary predicate. Thus, if $\mathfrak{A}$ is a structure, we always assume that $T^{\mathfrak{A}}$ is a transitive relation on A. We additionally suppose that we have at our disposal a distinguished *unary* predicate $\hat{T}$, which we take to be satisfied, in any structure, by precisely those elements related to themselves by T. This constitutes no essential increase in the expressive power of any of the logics $\mathcal{FL}^m_= 1\mathrm{T}$ ($m \geq 2$), since we may fix the interpretation of $\hat{T}$ by writing the $\mathcal{FL}^2_= 1\mathrm{T}$-formula $\forall(\hat{T} \leftrightarrow \forall(= \rightarrow T))$. It follows from the constraints imposed on T and $\hat{T}$ that a T-clique containing any element not satisfying $\hat{T}$ in fact consists of just that element, and is a soliton.

3.1. Basic formulas

Throughout this section (3.1), Σ will always stand for a signature consisting of the distinguished predicates T and $\hat{T}$ together with any number of ordinary, *unary* predicates. We denote by Π_Σ the set of fluted 1-types over Σ. We always use the (possibly decorated) letters π to range over elements of Π_Σ, μ to range over arbitrary quantifier-free formulas of arity 1 in the signature Σ, and Π to range over subsets of Π_Σ. Call an $\mathcal{FL}^2_= 1\mathrm{T}$-formula over Σ *basic* if it is of one of the following forms:

(B1) $\forall(\pi \rightarrow \exists(\mu \wedge T \wedge \neq))$

(B2) $\forall(\pi \rightarrow \exists(\mu \wedge \neg T \wedge \neq))$

(B3) $\forall(\pi \rightarrow \forall(\pi' \rightarrow T))$ $\quad(\pi \neq \pi')$

(B4) $\forall(\pi \rightarrow \forall(\pi' \rightarrow \neg T))$ $\quad(\pi \neq \pi')$

(B5) $\forall(\pi \rightarrow \forall(\pi \rightarrow (= \vee T))$

(B6) $\forall(\pi \rightarrow \forall(\pi \rightarrow (= \vee \neg T))$

(B7) $\forall\mu$ (B8) $\exists\mu$.

We remark that the unary predicate $\hat{T}$ will always occur in the fluted 1-types π and π' in these forms. Our goal is to give a procedure for determining the satisfiability for finite sets of basic formulas. Our strategy is to characterize (possibly infinite) structures by means of finite data-structures, called *certificates*. Essentially, a certificate specifies which fluted 1-types are realized in the structure in question, and places just enough constraints on the arrangement of the cliques in that structure to determine whether it is a model of any basic formula. We remark that flutedness plays a crucial role here: none of the forms (B1)–(B8) corresponds to the first-order formula $\forall x(\pi(x) \to \forall y(\pi'(y) \to (T(x,y) \vee T(y,x))))$. In showing that the certificates guarantee the satisfiability of a set of basic formulas, we rely on the fact that such formulas do not occur.

To define certificates formally, we require the following notions. A *clique type* (over Σ) is a function $\xi : \Pi_\Sigma \to \{0,1,2\}$. Intuitively, a clique type is a multi-set of fluted 1-types, with multiplicities truncated at 2. We write $\pi \in \xi$ to mean that $\xi(\pi) \geq 1$, and treat ξ as the set of fluted 1-types $\{\pi \mid \pi \in \xi\}$ where convenient, thus writing, for example $\xi \cup \Pi$ for $\{\pi \mid \pi \in \xi \text{ or } \pi \in \Pi\}$, and so on. A *clique super-type* is a pair $\langle \xi, \Pi \rangle$, where ξ is a clique type and Π a set of fluted 1-types. We call ξ a *soliton clique type* if $\neg\hat{T} \in \bigcup \xi$, that is, if some fluted 1-type occurring in ξ with non-zero multiplicity contains the atom $\neg\hat{T}$. If $\mathfrak{A}$ is a structure interpreting Σ, B a clique of $\mathfrak{A}$, and $a \in B$, then the *clique type of* a is the function $\mathrm{ctp}^{\mathfrak{A}}[a] : \Pi_\Sigma \to \{0,1,2\}$ given by

$$\mathrm{ctp}^{\mathfrak{A}}[a](\pi) = \begin{cases} 2 & \text{if } \pi \text{ is realized in } \mathfrak{A} \text{ by at least two elements of } B \\ 1 & \text{if } \pi \text{ is realized in } \mathfrak{A} \text{ by exactly one element of } B \\ 0 & \text{otherwise,} \end{cases}$$

and the *clique super-type of* a is the pair $\mathrm{cstp}^{\mathfrak{A}}[a] = \langle \mathrm{ctp}^{\mathfrak{A}}[a], \Pi \rangle$, where

$$\Pi = \{\mathrm{ftp}^{\mathfrak{A}}[b] \mid \mathfrak{A} \models T[a,b] \text{ and } \mathfrak{A} \not\models T[b,a] \text{ for some } b \in A\}.$$

If B is a clique in $\mathfrak{A}$, then all elements of B obviously have the same clique type and the same clique super-type, which we denote by $\mathrm{ctp}^{\mathfrak{A}}[B]$ and $\mathrm{cstp}^{\mathfrak{A}}[B]$, respectively. It is easy to see that a clique B of $\mathfrak{A}$ is a soliton if and only if $\mathrm{ctp}^{\mathfrak{A}}[B]$ is a soliton clique type. Intuitively, the type of a clique is a specification of which fluted 1-types are realized exactly once in that clique, and which fluted 1-types are realized more than once; the super-type of a clique is its type paired with a specification of which fluted 1-types outside that clique can be reached from it via the predicate T.

A *certificate* (over Σ) is a triple $\mathcal{C} = \langle \Omega, \ll, V \rangle$, where Ω is a set of clique super-types over Σ, $\ll$ a strict partial order on Π_Σ, and $V \subseteq \Pi_\Sigma$, subject to the following conditions:

(C1) if $\langle \xi, \Pi \rangle \in \Omega$ and $\pi' \in \Pi$, then there exists $\langle \xi', \Pi' \rangle \in \Omega$ such that (i) $\pi' \in \xi'$, (ii) $\Pi' \cup \xi' \subseteq \Pi$, and (iii) $\xi \cap V \cap \Pi' = \emptyset$;

(C2) if $\langle \xi, \Pi \rangle, \langle \xi', \Pi' \rangle \in \Omega$ are distinct, $\pi \in \xi$, $\pi' \in \xi'$ and $\pi \ll \pi'$, then $\xi' \cup \Pi' \subseteq \Pi$;

(C3) if $\langle \xi, \Pi \rangle, \langle \xi', \Pi' \rangle \in \Omega$ and $\xi \cap \xi' \cap V \neq \emptyset$, then $\xi = \xi'$ and $\Pi = \Pi'$;

(C4) if $\langle \xi, \Pi \rangle \in \Omega$ and ξ is a soliton clique type, then there exists $\pi \in \Pi_\Sigma$ such that $\xi(\pi) = 1$ and $\xi(\pi') = 0$ for all $\pi' \in \Pi_\Sigma \setminus \{\pi\}$;

(C5) if $\langle \xi, \Pi \rangle \in \Omega$, $\pi' \in \xi$ and $\pi \ll \pi'$, then $\pi \notin \Pi$;

(C6) if $\langle \xi, \Pi \rangle \in \Omega$, $\pi, \pi' \in \xi$ and $\pi \ll \pi'$ then $\xi \cap V \neq \emptyset$.

To make sense of these conditions, imagine that Ω is the set of clique super-types realized in some structure $\mathfrak{A}$, let V be the set of fluted 1-types that are realized in exactly one clique of $\mathfrak{A}$, and define $\pi \ll \pi'$ (for fluted 1-types π and π' realized in $\mathfrak{A}$) to mean that $\mathfrak{A} \models \forall(\pi \to \forall(\pi' \to T))$ but $\mathfrak{A} \not\models \forall(\pi' \to \forall(\pi \to T))$. To understand (C1), suppose $a \in A$ has clique super-type $\langle \xi, \Pi \rangle$ and $\pi' \in \Pi$. Then $\mathfrak{A} \models T[a, a']$ for some element $a' \in A$ with fluted 1-type π' not lying in the same clique as a. Now a' must realize some clique super-type $\langle \xi', \Pi' \rangle$, with $\pi' \in \xi'$. But then a is related by T to every element in the clique of a' and, more generally, to every element to which a' is related by T, whence $\Pi' \cup \xi' \subseteq \Pi$. And furthermore, a' cannot be related by T to any element a'' whose fluted 1-type occurs only in the same clique as a (since then a, a' and a'' would all be in the same clique), whence $\xi \cap V \cap \Pi' = \emptyset$. Condition (C2) reflects the fact that, if $\pi \ll \pi'$, then any element a with fluted 1-type π is related by T to any element a' having fluted 1-type π', and indeed to all elements to which a' is related. The antecedents of (C3) state that there are cliques having super-types ξ and ξ' which both realize an element whose fluted 1-type is realized only in a single clique; hence these cliques are identical. Condition (C4) reflects the fact that soliton clique types are realized only by solitons. Condition (C5) follows from the fact that, if the clique super-type $\langle \xi, \Pi \rangle$ of some element a specifies that a is related by T to an element, say a' (of type π) in another clique, then a' cannot be required to be related by T to anything (of type π') in the clique of a. Finally, (C6) states that if some clique of type ξ contains elements whose fluted 1-types are related by $\ll$, then there must be a an element whose fluted 1-type is realized only in that clique. This condition is not satisfied by structures in general. However, it is satisfied by a particular class of structures which are easy to work with and to which—it transpires—we may confine attention. We return to this matter presently.

Having explained what certificates are, we turn our attention now to how they can be used to determine the truth of basic formulas. Let $\mathcal{C} = \langle \Omega, \ll, V \rangle$ be a certificate and ψ a basic formula, both over some signature Σ. We define the *satisfaction* relation $\mathcal{C} \models \psi$. In this definition, for any fluted 1-type π, we say that π *occurs* in $\mathcal{C}$ if, there exists $\langle \xi, \Pi \rangle \in \Omega$ such that $\pi \in \xi$. We proceed by cases.

1. ψ is $\forall(\pi \to \exists(\mu \wedge T \wedge \neq))$: $\mathcal{C} \models \psi$ if and only if, for all $\langle \xi, \Pi \rangle \in \Omega$, with $\pi \in \xi$, either (i) $\models \pi \to \mu$ and $\xi(\pi) = 2$; or (ii) there exists $\pi' \in \xi$ such that $\pi' \neq \pi$ and $\models \pi' \to \mu$; or (iii) there exists $\pi' \in \Pi$ such that $\models \pi' \to \mu$.

2. ψ is $\forall(\pi \to \exists(\mu \wedge \neg T \wedge \neq))$: $\mathcal{C} \models \psi$ if and only if, for all $\langle \xi, \Pi \rangle \in \Omega$ with $\pi \in \xi$, there exists $\langle \xi', \Pi' \rangle \in \Omega$ such that (i) $\models \pi' \to \mu$; (ii) there exist no $\pi'' \in \Pi$ and $\pi''' \in \xi'$ such that $\pi'' \ll \pi'''$; (iii) $\xi' \cap \Pi \cap V = \emptyset$; and (iv) $\langle \xi, \Pi \rangle = \langle \xi', \Pi' \rangle \Rightarrow \xi \cap V = \emptyset$.

3. ψ is $\forall(\pi \to \forall(\pi' \to T))$, where $\pi \neq \pi'$: $\mathcal{C} \models \psi$ if and only if one of the following obtains: (i) one of π or π' does not occur in $\mathcal{C}$; (ii) $\pi \ll \pi'$; or (iii) for all $\langle \xi, \Pi \rangle, \langle \xi', \Pi' \rangle \in \Omega$ such that $\pi \in \Pi$ and $\pi' \in \xi'$, we have $\xi = \xi'$, $\Pi = \Pi'$ and $\xi \cap V \neq \emptyset$.

4. ψ is $\forall(\pi \to \forall(\pi' \to \neg T))$, where $\pi \neq \pi'$: $\mathcal{C} \models \psi$ if and only if for all $\langle \xi, \Pi \rangle \in \Omega$ such that $\pi \in \xi$, $\pi' \notin \xi \cup \Pi$.

5. ψ is $\forall(\pi \to \forall(\pi \to (= \vee\, T)))$: $\mathcal{C} \models \psi$ if and only if there is at most one $\langle \xi, \Pi \rangle \in \Omega$ such that $\pi \in \xi$, and, if such a $\langle \xi, \Pi \rangle$ exists, then $\xi \cap V \neq \emptyset$.

6. ψ is $\forall(\pi \to \forall(\pi \to (= \vee \neg T)))$: $\mathcal{C} \models \psi$ if and only if for all $\langle \xi, \Pi \rangle \in \Omega$, $\pi \notin \xi \cap \Pi$, and $\xi(\pi) \leq 1$.

7. ψ is $\forall \mu$: $\mathcal{C} \models \psi$ if and only if, for all $\langle \xi, \Pi \rangle \in \Omega$ and $\pi \in \xi$, $\models \pi \to \mu$.

8. ψ is $\exists \mu$: $\mathcal{C} \models \psi$ if and only if there exist $\langle \xi, \Pi \rangle \in \Omega$ and $\pi \in \xi$ such that $\models \pi \to \mu$.

The reader may wish to postpone digesting the details of these conditions for the present. Roughly, the intention is that: (i) given a structure $\mathfrak{A}$, we can construct a certificate $\mathcal{C}$ satisfying all the basic formulas that $\mathfrak{A}$ satisfies (Lemma 11); and (ii) from a certificate $\mathcal{C}$, we can build a structure $\mathfrak{A}$, satisfying all the basic formulas that $\mathcal{C}$ satisfies (Lemma 12). This translates the search for a model of a set of basic formulas Ψ into a search for a certificate satisfying Ψ. We mention a minor complication at this point: claim (i) is not true in full generality, and we must restrict attention to a particular class of structures, which we call "quadratic structures"; however, we show that this class is as general as we need (Lemma 4).

With the technical apparatus of certificates at our disposal, we are in a position to transform the problem of determining the satisfiability of basic formulas in *structures* to that of determining their satisfiability by *certificates*. We begin by constructing certificates from structures. If $\mathfrak{A}$ is a structure, then we define $\mathcal{C}(\mathfrak{A})$ to be the triple $\langle \Omega, \ll, V \rangle$, where: $\Omega = \{\text{cstp}^{\mathfrak{A}}[a] \mid a \in A\}$ is the set of clique super-types realized in $\mathfrak{A}$; $\pi \ll \pi'$ if and only if π and π' are realized in $\mathfrak{A}$, $\mathfrak{A} \models \forall(\pi \to \forall(\pi' \to T))$ and $\mathfrak{A} \not\models \forall(\pi' \to \forall(\pi \to T))$; and V is the set of fluted 1-types realized in exactly one clique of $\mathfrak{A}$. We must ensure that the definition of $\ll$ has the requisite property:

Lemma 3. *The relation $\ll$ in the construction of $\mathcal{C}(\mathfrak{A})$ is a strict partial order.*

Proof. Antisymmetry is immediate. For transitivity, suppose, $\pi \ll \pi'$ and $\pi' \ll \pi''$. Trivially, $\mathfrak{A} \models \forall(\pi \to \forall(\pi'' \to T))$. On the other hand, if we also have $\mathfrak{A} \models \forall(\pi'' \to \forall(\pi \to T))$, then $\mathfrak{A} \models \forall(\pi'' \to \forall(\pi' \to T))$, contradicting $\pi' \ll \pi''$. Hence $\pi \ll \pi''$. □

Notice that we have not asserted that $\mathcal{C}(\mathfrak{A})$ is a certificate. Indeed this is true only for a specific class of structures, which we now define. Suppose $\mathfrak{A}$ is a structure, B a clique of $\mathfrak{A}$, and π, π' fluted 1-types. We say that B is *determined by* the pair $\{\pi, \pi'\}$ if it is the unique clique of $\mathfrak{A}$ in which π and π' are both realized. We call $\mathfrak{A}$ *quadratic* if, for any clique B determined by some pair of fluted 1-types $\{\pi, \pi'\}$, there exists a fluted 1-type π^* such that B is the unique clique of $\mathfrak{A}$ in which π^* is realized. That is, in a quadratic structure, any clique which can be uniquely identified as the only clique containing a given pair of fluted 1-types, π and π', can be uniquely identified as the only clique containing some (possibly different) fluted 1-type π^*. Our next task is to show that we may confine attention, without loss of generality, to quadratic structures.

Let Φ be a set of basic formulas over some signature Σ, and write $\ell = |\Sigma|$. Any $\varphi \in \Phi$ has one of the forms $\forall(\pi \to \exists(\mu \wedge \chi))$, $\forall(\pi \to \forall(\pi' \to \chi))$, $\forall\mu$ or $\exists\mu$, where π and π' are fluted 1-types over Σ, μ a unary, quantifier-free formula and χ a quantifier-free formula involving only the predicates T and $=$. Now let Σ^* be Σ together with the fresh unary predicates $p_0, \ldots, p_{2\ell-1}$, let $\bar{p}_0$ be the formula $\neg p_0 \wedge \cdots \wedge \neg p_{2\ell-1}$, and let $\Phi^* = \{\varphi^* \mid \varphi \in \Phi \cup \{\exists\top\}\}$, where

$$\varphi^* := \begin{cases} \forall(\pi \wedge \bar{p}_0 \to \exists(\mu \wedge \bar{p}_0 \wedge \chi) & \text{if } \varphi = \forall(\pi \to \exists(\mu \wedge \chi)) \\ \forall(\pi \wedge \bar{p}_0 \to \forall(\pi' \wedge \bar{p}_0 \to \chi) & \text{if } \varphi = \forall(\pi \to \forall(\pi' \to \chi)) \\ \forall(\bar{p}_0 \to \mu) & \text{if } \psi = \forall\mu \\ \exists(\mu \wedge \bar{p}_0) & \text{if } \psi = \exists\mu. \end{cases}$$

If π is a fluted 1-type over Σ, then $\pi \wedge \bar{p}_0$ is a fluted 1-type over Σ^*; hence Φ^* is a set of basic formulas over Σ^*. Moreover, Φ^* can be computed in time bounded by a polynomial function of $\|\Phi\|$. The following lemma tells us that Φ has a model if and only if Φ^* has a quadratic model.

Lemma 4. *Suppose Φ is a set of basic formulas. The following are equivalent:* (i) *Φ is satisfiable;* (ii) *$\Phi^* \cup \{\forall\bar{p}_0\}$ is satisfiable;* (iii) *Φ^* is satisfied in a quadratic structure;* (iv) *Φ^* is satisfiable.*

Proof. Let Σ be the signature of Φ, and Σ^*, the signature of Φ^*, as defined above. Call any fluted 1-type π over Σ^* such that $\models \pi \to \bar{p}_0$ *proper*. Clearly, the proper fluted 1-types over Σ^* are in natural 1–1 correspondence with the fluted 1-types over Σ.

(i) $\Rightarrow$ (ii): If $\mathfrak{A} \models \Phi$, let $\mathfrak{B}$ be the expansion of $\mathfrak{A}$ obtained by taking every element of A to satisfy $\bar{p}_0$. It is obvious that $\mathfrak{B} \models \Phi^* \cup \{\forall\bar{p}_0\}$. (ii) $\Rightarrow$ (iii): Suppose $\mathfrak{A} \models \Phi^* \cup \{\forall\bar{p}_0\}$. For each (unordered) pair, π, π' of distinct, proper fluted 1-types (over Σ^*) such that there is exactly one clique, u of $\mathfrak{A}$ in which both are realized, choose a fresh, *improper* fluted 1-type over Σ^*, and simply add a new element with that fluted 1-type to u. Because there are certainly $2^{2|\sigma|} - 1$ improper fluted 1-types, we never run out of fresh, improper fluted 1-types, so let $\mathfrak{B}$ be the resulting structure. Since the new elements do not satisfy $\bar{p}_0$, we have $\mathfrak{B} \models \Phi^*$. And since all the newly realized fluted 1-types occur only in single cliques, $\mathfrak{B}$ is quadratic. (iii) $\Rightarrow$ (iv) is trivial. (iv) $\Rightarrow$ (i):

Suppose $\mathfrak{A} \models \Phi^*$, and let $\mathfrak{B}$ be restriction of $\mathfrak{A}$ to the (necessarily non-empty) set of elements satisfying $\bar{p}_0$. It is obvious that $\mathfrak{B} \models \Phi$. $\square$

In the sequel, therefore, we may assume the conversion of Lemma 4 has been carried out, and concern ourselves with the problem of determining whether a given finite set of basic formulas Φ is satisfied in some *quadratic* structure. We show that this problem is equivalent to that of determining whether Φ is satisfied by some certificate.

Manufacturing a certificate from a quadratic structure is easy.

Lemma 5. *If $\mathfrak{A}$ is any quadratic structure interpreting Σ, then $\mathcal{C}(\mathfrak{A})$ is a certificate over Σ.*

Proof. Write $\mathcal{C}(\mathfrak{A}) = \langle \Omega, \ll, V \rangle$. By Lemma 3, $\ll$ is a strict partial order on Π_Σ. We must check conditions (C1)–(C6).

(C1): Suppose $\langle \xi, \Pi \rangle \in \Omega$ and $\pi' \in \Pi$. Let a be such that $\text{cstp}^{\mathfrak{A}}[a] = \langle \xi, \Pi \rangle$. Then there exists $b \in A$ such that $\text{ftp}^{\mathfrak{A}}[b] = \pi'$ and $\mathfrak{A} \models T[a, b]$, but with a and b lying in different cliques. Let $\text{cstp}^{\mathfrak{A}}[b] = \langle \xi', \Pi' \rangle$. Then: (i) $\langle \xi', \Pi' \rangle \in \Omega$ by construction of Ω; (ii) $\xi' \cup \Pi' \subseteq \Pi$ by transitivity of $T^{\mathfrak{A}}$; and (iii) if $\pi'' \in \xi \cap V$, then all elements with fluted 1-type π'' lie in the same clique as a. Since a and b are not in the same clique, b cannot be related by T to any of these elements, which is to say $\pi'' \notin \Pi'$.
(C2): Suppose $\langle \xi, \Pi \rangle, \langle \xi', \Pi' \rangle \in \Omega$ are distinct, $\pi \in \xi$, $\pi' \in \xi'$ and $\pi \ll \pi'$. Let $a, b \in A$ be such that $\text{cstp}^{\mathfrak{A}}[a] = \langle \xi, \Pi \rangle$ and $\text{cstp}^{\mathfrak{A}}[b] = \langle \xi', \Pi' \rangle$. If $\pi \ll \pi'$, then $\mathfrak{A} \models T[a, b]$. Moreover, if a and b belong to different cliques, then $\xi' \cup \Pi' \subseteq \Pi$, by the transitivity of T.
(C3): Suppose $\langle \xi, \Pi \rangle, \langle \xi', \Pi' \rangle \in \Omega$ and $\xi \cap \xi' \cap V \neq \emptyset$. Let $a, b \in A$ be such that $\text{cstp}^{\mathfrak{A}}[a] = \langle \xi, \Pi \rangle$ and $\text{cstp}^{\mathfrak{A}}[b] = \langle \xi', \Pi' \rangle$. If there exists a fluted 1-type π'' realized both in the clique of a and in the clique of b, and, moreover, in just one clique of $\mathfrak{A}$, then a and b are in the same clique.
(C4): Suppose $\langle \xi, \Pi \rangle \in \Omega$ and $\neg \hat{T} \in \bigcup \xi$. By construction, there exists $b \in A$ such that $\text{ctp}^{\mathfrak{A}}[b] = \xi$, and $\mathfrak{A} \not\models \hat{T}[b]$. But then b is the only element of its clique, and we may set $\pi = \text{ftp}^{\mathfrak{A}}[b]$.
(C5): Suppose $\langle \xi, \Pi \rangle \in \Omega$, $\pi' \in \xi$ and $\pi \ll \pi'$. Let $a, a' \in A$ be such that $\text{cstp}^{\mathfrak{A}}[a] = \langle \xi, \Pi \rangle$, $\text{ftp}^{\mathfrak{A}}[a'] = \pi'$, and a' is in the same clique as a. To show that $\pi \notin \Pi$, we must show that, for all $b \in A$ such that $\text{ftp}^{\mathfrak{A}}[b] = \pi$, either $\mathfrak{A} \not\models T[a, b]$ or b is in the same clique as a. But this follows immediately from $\pi \ll \pi'$.
(C6): Suppose $\langle \xi, \Pi \rangle \in \Omega$, $\pi, \pi' \in \xi$ and $\pi \ll \pi'$. It follows that there is exactly one clique of $\mathfrak{A}$, say u, in which π and π' are both realized, and that $\text{cstp}^{\mathfrak{A}}[u] = \langle \xi, \Pi \rangle$. Since $\mathfrak{A}$ is, by assumption, quadratic, there exists a fluted 1-type $\pi^* \in \xi$ realized only in u. Thus $\xi \cap V \neq \emptyset$. We mention in passing that this is the only point where we use the assumption that $\mathfrak{A}$ is quadratic. $\square$

Manufacturing a (quadratic) structure from a certificate is more difficult. As an aide to intuition, we give an informal sketch first. Suppose $\mathcal{C} = \langle \Omega, \ll, V \rangle$ is a certificate; we proceed to define a structure $\mathfrak{A}$. The domain A is the disjoint union of sets $A_{\xi,\Pi}$, where $\langle \xi, \Pi \rangle$ ranges over Ω; the elements of $A_{\xi,\Pi}$ will all be

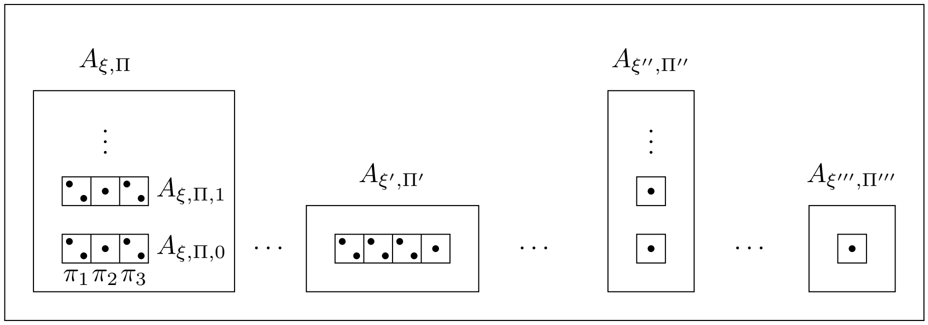

Figure 3: Construction of the domain A of $\mathfrak{A}(\mathcal{C})$ for $\mathcal{C}$ a certificate.

assigned the clique super-type $\langle \xi, \Pi \rangle$. If ξ contains no fluted 1-type π such that $\pi \in V$, then $A_{\xi,\Pi}$ will consist of infinitely many sets $A_{\xi,\Pi,i}$ ($i \geq 0$), referred to in the construction as 'cells'. If, on the other hand, ξ contains a fluted 1-type π such that $\pi \in V$, then $A_{\xi,\Pi}$ will consist of a single cell $A_{\xi,\Pi,0}$. It will later turn out that the cells are exactly the T-cliques of $\mathfrak{A}$; moreover, we identify certain cells (consisting of a single element) as 'soliton cells', which will turn out to be the solitons of $\mathfrak{A}$. Each cell $A_{\xi,\Pi,i}$ is in turn the disjoint union of sets $A_{\pi,\xi,\Pi,i}$, where π ranges over the fluted 1-types in ξ. The idea is that the elements of $A_{\pi,\xi,\Pi,i}$ will all be given fluted 1-type π; moreover, this set has cardinality equal to $\xi(\pi)$ (i.e. either 1 or 2). Fig. 3 gives a schematic representation of the domain A, featuring, among others, the clique super-types $\langle \xi, \Pi \rangle, \ldots, \langle \xi''', \Pi''' \rangle$. In this example, we are supposing that the clique types ξ and ξ'' are disjoint from V, so that the corresponding sets $A_{\xi,\Pi}$ and $A_{\xi'',\Pi''}$ comprise countably many cells; by contrast, ξ' and ξ''' intersect V, so that the corresponding sets $A_{\xi',\Pi'}$ and $A_{\xi''',\Pi'''}$ are unicellular. We are further supposing that the clique type ξ contains three fluted 1-types (with non-zero multiplicity), namely π_1, π_2 and π_3; ξ' contains four fluted 1-types; ξ'' and ξ''' each contain just one fluted 1-type. Within any cell, each fluted 1-type occurs either once or twice.

We will take the relation $T^{\mathfrak{A}}$ to be the transitive closure of the union of three relations, R_{cell}, R_{ex} and R_{all}, each of which plays a specific role. The relation R_{cell} specifies T within each cell, $A_{\xi,\Pi,i}$. It will ensure that, excepting the case of soliton cells, all pairs of elements within a single cell are related by T. The relation R_{cell}, in essence, secures the existential commitments required by the clique super-types. Specifically, if $a \in A_{\xi,\Pi,i}$ and $\pi' \in \Pi$, we select some $\langle \xi', \Pi' \rangle \in \Omega$ such that $\xi' \cup \Pi' \subseteq \Pi$ (possible by (C1)), and choose cells included in $A_{\xi',\Pi'}$ whose elements will act as 'witnesses' for the fact that a has to be related by T to something of type π'. Finally, the relation R_{all} deals with the T-relations mandated by $\ll$: if $A_{\xi,\Pi,i}$ and $A_{\xi',\Pi',j}$ are distinct cells with $\pi \in \xi$ and $\pi' \in \xi'$, then, in order to secure that $\mathfrak{A} \models \forall(\pi \to \forall(\pi' \to T))$, we take all elements of the former cell to be related by R_{all} to all elements of the latter. We explain how R_{cell}, R_{ex} and R_{all} work in more detail after we have defined them formally; at this stage however, it is helpful to keep the following objective in mind. Since we wish the cells of A to become cliques of the model $\mathfrak{A}$, we must be careful not to create cycles in the resulting graph of R_{ex}- and R_{all}-links. (Any such cycle would cause several cells to be merged into a single clique.) Most of the technical complications in the definitions of R_{ex} and R_{all} derive from this objective.

We are now ready to give the formal definition of $\mathfrak{A}$, beginning with the construction of the domain, A. For all $\langle \xi, \Pi \rangle \in \Omega$, all $\pi \in \xi$ and all $i \in \mathbb{N}$, let $a^+_{\pi,\xi,\Pi,i}$ and $a^-_{\pi,\xi,\Pi,i}$ be fresh objects. Set

$$A_{\pi,\xi,\Pi,i} = \begin{cases} \{a^+_{\pi,\xi,\Pi,i}, a^-_{\pi,\xi,\Pi,i}\} & \text{if } \xi(\pi) = 2 \\ \{a^+_{\pi,\xi,\Pi,i}\} & \text{otherwise (i.e. if } \xi(\pi) = 1) \end{cases}$$

$$A_{\xi,\Pi,i} = \bigcup_{\pi \in \xi} A_{\pi,\xi,\Pi,i}$$

$$A_{\xi,\Pi} = \begin{cases} \bigcup_{i \in \mathbb{N}} A_{\xi,\Pi,i} & \text{if } \xi \cap V = \emptyset \\ A_{\xi,\Pi,0} & \text{otherwise} \end{cases}$$

$$A = \bigcup_{\langle \xi, \Pi \rangle \in \Omega} A_{\xi,\Pi}.$$

We call the sets $A_{\xi,\Pi,i}$ *cells*. Thus, any cell $A_{\xi,\Pi,i}$ is the union of non-empty sets $A_{\pi,\xi,\Pi,i}$, each with cardinality $\xi(\pi)$, where π ranges over ξ (considered as a set). The sets $A_{\xi,\Pi}$ are unions of cells defined in one of two ways, depending on ξ. If $\xi \cap V = \emptyset$, then $A_{\xi,\Pi}$ includes $A_{\xi,\Pi,i}$ for all $i \geq 0$; otherwise, $A_{\xi,\Pi}$ is simply the cell $A_{\xi,\Pi,0}$. We fix the extensions of unary predicates in $\mathfrak{A}$ (including $\hat{T}$) by setting $\text{ftp}^{\mathfrak{A}}[a] = \pi$ for all $a = a^p_{\pi,\xi,\Pi,i} \in A$, where $p \in \{+, -\}$. As already announced, the intention is that each cell $A_{\xi,\Pi,i}$ will become a clique of $\mathfrak{A}$ with super-type $\langle \xi, \Pi \rangle$. This explains why, if $\xi \cap V \neq \emptyset$, the set $A_{\xi,\Pi}$ consists of just one cell: V is meant to represent the set of 1-types occurring in exactly one clique; hence we certainly do not want infinitely many different cells all realizing some fluted 1-type in V. We remark in this connection that, in this case, there will only ever be a single value of Π such that $\langle \xi, \Pi \rangle \in \Omega$, by (C3). If ξ is a soliton clique type, we call the cell $A_{\xi,\Pi,i}$ a *soliton-cell*. It follows from (C4) that, in this case, $A_{\xi,\Pi,i} = \{a^+_{\pi,\xi,\Pi,i}\}$ for some fluted 1-type π. Note that the converse does not hold: it is perfectly feasible for the cell $A_{\xi,\Pi,i}$ to consist of the single element $a^+_{\pi,\xi,\Pi,i}$ even though $\hat{T} \in \pi$.

It remains only to set the extension of the distinguished predicate T. As mentioned earlier, we define three binary relations, R_{cell}, R_{ex} and R_{all}, taking T to be the transitive closure of $R_{\text{cell}} \cup R_{\text{ex}} \cup R_{\text{all}}$. The relations R_{cell} and R_{all} are relatively straightforward, and we consider them first. Let $a = a^p_{\pi,\xi,\Pi,i}$ and $a' = a^{p'}_{\pi',\xi',\Pi',j}$; and let $u = A_{\xi,\Pi,i}$ and $v = A_{\xi',\Pi',j}$ be the respective cells of a and a'. We declare $a R_{\text{cell}} a'$ if and only if $u = v$ (i.e. $\xi = \xi'$, $\Pi = \Pi'$, and $i = j$), and ξ is not a soliton clique type. That is: R_{cell} holds between all pairs of elements in the same non-soliton cell. (As just mentioned, any soliton cell is composed of a single element.) Turning now to the relation R_{all}, we declare

$aR_{\rm all}a'$ if $u \neq v$ and, for some fluted 1-types $\pi \in \xi$ and $\pi' \in \xi'$, $\pi \ll \pi'$. That is: $R_{\rm all}$ holds between a and a' if u contains an element with fluted 1-type π and $v \neq u$ contains an element with fluted 1-type π', such that all elements with fluted 1-type π are required to be related by T to all those having fluted 1-type π'. Note that the relation $R_{\rm all}$ depends only on the *cells* of its relata: that is to say, if $b \in u$ and $b' \in v$, then $aR_{\rm all}a'$ implies $bR_{\rm all}b'$. There being no ambiguity, we shall write, in this case, $uR_{\rm all}v$.

Turning finally to the relation $R_{\rm ex}$, let $u = A_{\xi,\Pi,i}$ and $v = A_{\xi',\Pi',j}$ be the respective cells of a and a', as in the previous paragraph. Declare $aR_{\rm ex}a'$ if (a) $\xi' \cup \Pi' \subseteq \Pi$; (b) $\xi' \cap V = \emptyset \Rightarrow j \geq i+2$; and (c) $\xi \cap V \cap \Pi' = \emptyset$. As explained above, the idea is to pick witnesses to realize the fact that a is supposed to have clique super-type $\langle \xi, \Pi \rangle$, and in particular is supposed to be related via the predicate T to some element with fluted 1-type π' for each $\pi' \in \Pi$. Condition (a) simply ensures that the chosen witnesses have clique super-type $\langle \xi', \Pi' \rangle$ compatible with that of a. To understand (b), recall that, if $\xi' \cap V = \emptyset$ then $A_{\xi',\Pi',j}$ exists for all $j \geq 0$, whereas if $\xi' \cap V \neq \emptyset$, then $A_{\xi',\Pi',j}$ exists only for $j = 0$ (reflecting the fact that the cell contains an element whose fluted 1-type should be realized in only one clique). Condition (b) states that, in the former case, we are to pick elements of $A_{\xi',\Pi',j}$ as witnesses when $j \geq i+2$: the fact that $j > i$ avoids cycles of $R_{\rm ex}$-links being created when this process is repeated. In the latter case, (b) does not prevent us from selecting elements of $A_{\xi',\Pi',0}$ as witnesses: here, the various conditions on certificates will ensure that doing so still will not give rise to any $R_{\rm ex}$-cycles. Condition (c) is a technical condition used in Lemma 8, but essentially, it states that, when seeking witnesses for a, we may not select any $a' \in A_{\xi',\Pi',j}$ if Π' requires a' to be related by T to some elements whose fluted 1-type is realized only in the same clique as a. Again, the relation $R_{\rm ex}$ depends only on the cells of its relata, and we write $uR_{\rm ex}v$ if some (equivalently, all) elements of u are related by $R_{\rm ex}$ to some (equivalently, all) elements of v. Having defined the relations $R_{\rm cell}$, $R_{\rm ex}$ and $R_{\rm all}$, we let $T^{\mathfrak{A}}$ be the transitive closure of $R_{\rm cell} \cup R_{\rm ex} \cup R_{\rm all}$. We denote the structure $\mathfrak{A}$, constructed from the certificate $\mathcal{C}$ as just described, by $\mathfrak{A}(\mathcal{C})$. Notice that $\mathfrak{A}(\mathcal{C})$ will in general be infinite.

We must check that $\mathfrak{A}(\mathcal{C})$ interprets the predicates T and $\hat{T}$ in a proper fashion. Lemmas 6–9 do precisely this.

Lemma 6. *In the construction of* $\mathfrak{A}(\mathcal{C})$*, if* $aR_{\rm ex}a'$*, then* a *and* a' *occupy different cells of* A*.*

Proof. Suppose for contradiction that $aR_{\rm ex}a'$ with $a = a^p_{\pi,\xi,\Pi,i}$ and $a' = a^{p'}_{\pi',\xi,\Pi,i}$. By condition (a) in the definition of $R_{\rm ex}$, we have $\xi \subseteq \Pi$, and, by (b), we have, $\xi \cap V \neq \emptyset$, whence $\xi \cap V \cap \Pi \neq \emptyset$, contradicting (c). $\square$

In this paper, we take all (directed) graphs to be *simple*—i.e., not to have multiple edges or self-loops. Now, another way of expressing Lemma 6 is to say that no cell of A is related to itself by $R_{\rm ex}$; and by definition, no cell of A is related to itself by $R_{\rm all}$. Hence, we may consider the directed graph on the set of cells of A defined by the relation $R_{\rm ex} \cup R_{\rm all}$. We show that this graph is acyclic. It follows that the cells (both soliton and non-soliton) are the cliques of the relation $T^{\mathfrak{A}}$, and hence that $T^{\mathfrak{A}}$ induces a strict partial order on these cells.

Lemma 7. *In the construction of* $\mathfrak{A}(\mathcal{C})$*, suppose* $u_0, \ldots, u_k$ $(k \geq 1)$ *is a sequence of cells such that, for all* h $(0 \leq h < k)$ *either* $u_h R_{\text{ex}} u_{h+1}$ *or* $u_h R_{\text{all}} u_{h+1}$*. Writing* $u_h = A_{\xi_h, \Pi_h, i_h}$ *for all* h $(0 \leq h \leq k)$*, we have* $\xi_k \cup \Pi_k \subseteq \Pi_0$*.*

Proof. We proceed by induction on k. For the base case $(k = 1)$, if $u_0 R_{\text{ex}} u_1$, then the result is immediate by (a) in the definition of R_{ex}. If $u_0 R_{\text{all}} u_1$, then there exist $\pi_0 \in \xi_0$ and $\pi_1 \in \xi_1$ such that $\pi_0 \ll \pi_1$. The result then follows from (C2). For the inductive case $(k > 1)$, we have by inductive hypothesis, $\xi_{k-1} \cup \Pi_{k-1} \subseteq \Pi_0$; and from the base case applied to the sequence u_{k-1}, u_k, we have $\xi_k \cup \Pi_k \subseteq \Pi_{k-1}$. □

Lemma 8. *In the construction of* $\mathfrak{A}(\mathcal{C})$*, there exists no sequence of cells* $u_0, \ldots, u_k = u_0$ $(k \geq 2)$ *such that, for all* h $(0 \leq h < k)$ *either* $u_h R_{\text{ex}} u_{h+1}$ *or* $u_h R_{\text{all}} u_{h+1}$*.*

Proof. Suppose for contradiction that such a sequence exists, again writing $u_h = A_{\xi_h, \Pi_h, i_h}$ for all h $(0 \leq h \leq k)$. By Lemma 7, $\Pi_0 = \cdots = \Pi_k = \Pi$, say, and $\xi_h \in \Pi$ for all h $(0 \leq h \leq k)$. It follows that we cannot have $u_h R_{\text{all}} u_{h+1}$ for any h $(0 \leq h < k)$, since, if there exist $\pi_h \in \xi_h$ and $\pi_{h+1} \in \xi_{h+1}$ with $\pi_h \ll \pi_{h+1}$, then, by (C5), $\pi_{h+1} \not\in \Pi_h = \Pi$, contradicting $\xi_{h+1} \subseteq \Pi$. Thus, we may assume that $u_h R_{\text{ex}} u_{h+1}$ for all h $(0 \leq h < k)$. Necessarily, $i_{h+1} \leq i_h$ for some h in the same range; indeed, by rotating the original sequence if necessary, we may assume without loss of generality that $h < k - 1$. By (b) in the definition of R_{ex}, $\xi_{h+1} \cap V \neq \emptyset$, and by (c), $\xi_{h+1} \cap V \cap \Pi_{h+2} = \emptyset$. But we have just argued that $\xi_{h+1} \subseteq \Pi$ and $\Pi_{h+2} = \Pi$. This is a contradiction. □

Lemma 9. *In the structure* $\mathfrak{A} = \mathfrak{A}(\mathcal{C})$*, we have* $\hat{T}^{\mathfrak{A}} = \{a \in A \mid \mathfrak{A} \models T[a, a]\}$*.*

Proof. Fix $a \in A_{\pi, \xi, \Pi, i}$. If $\mathfrak{A} \models \hat{T}[a]$, then $\hat{T} \in \pi$, whence, by (C4), ξ is not a soliton clique type. Hence $a R_{\text{cell}} a$, and $\mathfrak{A} \models T[a, a]$. Conversely, if $\mathfrak{A} \not\models \hat{T}[a]$, then $\neg\hat{T} \in \pi$, so that ξ is certainly a soliton type, and a is not related to itself by R_{cell}. On the other hand, by Lemma 8, there is no sequence of cells $u_0, \ldots, u_k$ $(k \geq 2)$ with $a \in u_0 = u_k$, such that, for all h $(0 \leq h < k)$, either $u_h R_{\text{ex}} u_{h+1}$ or $u_h R_{\text{all}} u_{h+1}$. Since $T^{\mathfrak{A}}$ is the transitive closure of $R_{\text{cell}} \cup R_{\text{ex}} \cup R_{\text{all}}$, we see that $\mathfrak{A} \not\models T[a, a]$, as required. □

Summing up the above discussion, we have

Lemma 10. *If* $\mathcal{C}$ *is a certificate over* Σ*, then* $\mathfrak{A}(\mathcal{C})$ *is a properly defined structure interpreting* Σ*.*

Thus, from a quadratic structure $\mathfrak{A}$, we can define a certificate $\mathcal{C}(\mathfrak{A})$, and from a certificate $\mathcal{C}$, we can define a structure $\mathfrak{A}(\mathcal{C})$. (It is easy to see that $\mathfrak{A}$ will in fact be quadratic, though this is inessential.) It remains only show that satisfaction of formulas by certificates corresponds to satisfaction of formulas by structures in the sense captured by the following two lemmas.

Lemma 11. *Let ψ be a basic formula, and suppose $\mathfrak{A} \models \psi$ for some quadratic structure $\mathfrak{A}$. Then $\mathcal{C}(\mathfrak{A}) \models \psi$.*

Proof. Write $\mathcal{C}(\mathfrak{A}) = \langle \Omega, \ll, V \rangle$. We consider the forms of ψ in turn.

1. ψ is $\forall(\pi \to \exists(\mu \wedge T \wedge \neq))$: Suppose $\mathfrak{A} \models \psi$ and $\langle \xi, \Pi \rangle \in \Omega$ with $\pi \in \xi$. Let $a \in A$ be such that $\text{cstp}^{\mathfrak{A}}[a] = \langle \xi, \Pi \rangle$ and $\text{ftp}^{\mathfrak{A}}[a] = \pi$. Pick $b \in A \setminus \{a\}$ such that $\mathfrak{A} \models \mu[b]$ and $\mathfrak{A} \models T[a, b]$, and let $\text{ftp}^{\mathfrak{A}}[b] = \pi'$. Thus, $\models \pi' \to \mu$. (i) If a and b are in the same clique of $\mathfrak{A}$ and $\pi = \pi'$, then $\models \pi \to \mu$, and $\xi(\pi) = 2$. (ii) If a and b are in the same clique, but $\pi' \neq \pi$, then $\pi' \in \xi$. (iii) If a and b are not in the same clique, then $\pi \in \Pi$.

2. ψ is $\forall(\pi \to \exists(\mu \wedge \neg T \wedge \neq))$: Suppose $\mathfrak{A} \models \psi$ and $\langle \xi, \Pi \rangle \in \Omega$ with $\pi \in \xi$. Let $a \in A$ be such that $\text{cstp}^{\mathfrak{A}}[a] = \langle \xi, \Pi \rangle$ and $\text{ftp}^{\mathfrak{A}}[a] = \pi$. Pick $b \in A \setminus \{a\}$ such that $\mathfrak{A} \models \mu[b]$ and $\mathfrak{A} \not\models T[a, b]$, and let $\text{cstp}^{\mathfrak{A}}[b] = \langle \xi', \Pi' \rangle$, and $\text{ftp}^{\mathfrak{A}}[b] = \pi'$. (i) Thus, $\models \pi' \to \mu$. (ii) Suppose, for contradiction, that there exist $\pi'' \in \Pi$ and $\pi''' \in \xi'$ such that $\pi'' \ll \pi'''$. Then there exist $b'', b''' \in A$ such that $\mathfrak{A} \models T[a, b'']$, $\mathfrak{A} \models T[b'', b''']$, with b''' in the same clique as b, contradicting the assumption that $\mathfrak{A} \not\models T[a, b]$. (iii) Suppose, for contradiction, that $\pi'' \in \xi' \cap \Pi \cap V$. Then there exists $b'' \in A$ with $\text{ftp}^{\mathfrak{A}}[b''] = \pi''$, realized in just one clique (namely, the clique of b) and an element b''' with $\text{ftp}^{\mathfrak{A}}[b'''] = \pi''$ and $\mathfrak{A} \models T[a, b''']$. This contradicts the supposition that $\mathfrak{A} \not\models T[a, b]$. (iv) Suppose, for contradiction, that $\langle \xi, \Pi \rangle = \langle \xi', \Pi' \rangle$ and $\pi'' \in \xi \cap V$. Then the cliques of both a and b contain elements of fluted 1-type π'', with such elements realized in just one clique. Thus a and b are in the same clique, which contradicts the supposition that $\mathfrak{A} \not\models T[a, b]$.

3. ψ is $\forall(\pi \to \forall(\pi' \to T))$, where $\pi \neq \pi'$: Suppose $\mathfrak{A} \models \psi$. (i) If π and π' are not both realized in $\mathfrak{A}$, then they do not both occur in $\mathcal{C}$. If π and π' are both realized in $\mathfrak{A}$, and $\mathfrak{A} \not\models \forall(\pi' \to \forall(\pi \to T))$, then $\pi \ll \pi'$. (iii) Otherwise, π and π' are realized in $\mathfrak{A}$, but there is a clique, say u, containing all these realizing elements. Hence, if $\langle \xi, \Pi \rangle, \langle \xi', \Pi' \rangle \in \Omega$ with $\pi \in \xi$ and $\pi' \in \xi'$, then $\langle \xi, \Pi \rangle = \langle \xi', \Pi' \rangle$, and $\pi \in V$, whence $\xi \cap V \neq \emptyset$.

4. ψ is $\forall(\pi \to \forall(\pi' \to \neg T))$, where $\pi \neq \pi'$: Suppose $\mathfrak{A} \models \psi$ and $\langle \xi, \Pi \rangle \in \Omega$ with $\pi \in \xi$. Then there exist $a \in A$ such that $\text{cstp}^{\mathfrak{A}}[a] = \langle \xi, \Pi \rangle$. By the definition of $\text{cstp}^{\mathfrak{A}}[a]$, $\pi' \notin \xi \cup \Pi$.

5. ψ is $\forall(\pi \to \forall(\pi \to (= \vee\, T)))$: Suppose $\mathfrak{A} \models \psi$. Then all elements $a \in A$ such that $\text{ftp}^{\mathfrak{A}}[a] = \pi$ lie in a single clique, so let their common clique super-type be $\langle \xi, \Pi \rangle$. Thus, $\langle \xi, \Pi \rangle$ is the only element of Ω such that $\pi \in \xi$; moreover, if this element exists, we have $\pi \in V$, and hence $\xi \cap V \neq \emptyset$.

6. ψ is $\forall(\pi \to \forall(\pi \to (= \vee \neg T)))$: Suppose $\mathfrak{A} \models \psi$ and $\langle \xi, \Pi \rangle \in \Omega$ with $\pi \in \xi$. Let $a \in A$ be such that $\text{cstp}^{\mathfrak{A}}[a] = \langle \xi, \Pi \rangle$ and $\text{ftp}^{\mathfrak{A}}[a] = \pi$, and let u be the clique of a in $\mathfrak{A}$. Since $\mathfrak{A} \models \psi$, there is certainly no element $b \in A \setminus u$ such that $\text{ftp}^{\mathfrak{A}}[b] = \pi$ and $\mathfrak{A} \models T[b, a]$, whence $\pi \notin \Pi$. One the other hand, there is no element $b \in u \setminus \{a\}$ such that $\text{ftp}^{\mathfrak{A}}[b] = \pi$, whence $\xi(\pi) = 1$.

The cases $\forall\mu$ and $\exists\mu$ are routine. $\square$

Lemma 12. *Let ψ be a basic formula, and suppose $\mathcal{C} \models \psi$ for some certificate $\mathcal{C}$. Then $\mathfrak{A}(\mathcal{C}) \models \psi$.*

Proof. Write $\mathcal{C} = \langle \Omega, \ll, V \rangle$ and $\mathfrak{A} = \mathfrak{A}(\mathcal{C})$. We consider the forms of ψ in turn.

1. ψ is $\forall(\pi \to \exists(\mu \wedge T \wedge \neq))$: Suppose $\mathcal{C} \models \psi$ and $a \in A$ with $\text{ftp}^{\mathfrak{A}}[a] = \pi$. We may write $a = a^p_{\pi,\xi,\Pi,i}$, for $\langle \xi, \Pi \rangle \in \Omega$ with $\pi \in \xi$. We must show that there exists $b \in A \setminus \{a\}$ such that $\mathfrak{A} \models \mu[b]$ and $\mathfrak{A} \models T[a,b]$. (i) If $\models \pi \to \mu$ and $\xi(\pi) = 2$, then, by construction of $\mathfrak{A}$, there exists $b = a^{p'}_{\pi,\xi,\Pi,i}$ with $p' \neq p$. Thus, $\text{ftp}^{\mathfrak{A}}[b] = \pi$ and $aR_{\text{cell}}b$, whence $\mathfrak{A} \models T[a,b]$. (ii) If there exists $\pi' \in \xi$ such that $\pi' \neq \pi$ and $\models \pi' \to \mu$, there exists $b = a^p_{\pi',\xi,\Pi,i}$. Thus, $\text{ftp}^{\mathfrak{A}}[b] = \pi'$ and $aR_{\text{cell}}b$, whence $\mathfrak{A} \models T[a,b]$. (iii) If there exists $\pi' \in \Pi$ such that $\models \pi' \to \mu$, then, by (C1), choose $\langle \xi', \Pi' \rangle \in \Omega$ with $\pi' \in \xi'$, $\xi' \cup \Pi' \subseteq \Pi$ and $\xi \cap \Pi' \cap V = \emptyset$. Suppose on the one hand that $\xi' \cap V = \emptyset$. Then we may let $b = a^+_{\pi',\xi,\Pi,i+2}$. Certainly, $\text{ftp}^{\mathfrak{A}}[b] = \pi'$. It suffices to prove that $aR_{\text{ex}}b$, whence $\mathfrak{A} \models T[a,b]$. We consider conditions (a)–(c) in the definition of R_{ex}. (a) We have already established that $\xi' \cup \Pi' \subseteq \Pi$. (b) Trivially, $i+2 \geq i+2$. (c) *A fortiori*, $\xi' \cap V \cap \Pi = \emptyset$. Suppose on the other hand that $\xi' \cap V \neq \emptyset$. Then we may let $b = a^+_{\pi',\xi,\Pi,0}$. Since $\xi \cap \Pi' \cap V = \emptyset$, we have $\xi \neq \xi'$, so that $b \neq a$. Again, consider conditions (b) and (c) in the definition of R_{ex}. For (b), we are supposing anyway that $\xi' \cap V \neq \emptyset$, and for (c), we have already established that $\xi \cap \Pi' \cap V = \emptyset$. Thus, in all cases, we have $\mathfrak{A} \models \mu[b]$ and $\mathfrak{A} \models T[a,b]$, as required.

2. ψ is $\forall(\pi \to \exists(\mu \wedge \neg T \wedge \neq))$: Suppose $\mathcal{C} \models \psi$ and $a \in A$ with $\text{ftp}^{\mathfrak{A}}[a] = \pi$. We may write $a = a^p_{\pi,\xi,\Pi,i}$, for $\langle \xi, \Pi \rangle \in \Omega$ with $\pi \in \xi$. Then we may select $\langle \xi', \Pi' \rangle \in \Omega$ with $\pi' \in \xi'$ such that: (i) $\models \pi' \to \mu$; (ii) there exists no $\pi'' \in \Pi$ and $\pi''' \in \xi'$ such that $\pi'' \ll \pi'''$; (iii) $\xi' \cap \Pi \cap V = \emptyset$; and (iv) $\langle \xi, \Pi \rangle = \langle \xi', \Pi' \rangle \Rightarrow \xi \cap V = \emptyset$. We consider first the case where $\langle \xi', \Pi' \rangle \neq \langle \xi, \Pi \rangle$. Let $b = a^+_{\pi',\xi',\Pi',0}$, so that, by construction of $\mathfrak{A}$, $\text{ftp}^{\mathfrak{A}}[b] = \pi'$. We must show that $a \neq b$ and $\mathfrak{A} \not\models T[a,b]$. Let u be the cell containing a and u' the cell containing b. Since $\langle \xi', \Pi' \rangle \neq \langle \xi, \Pi \rangle$, we have $u \neq u'$, whence, certainly $a \neq b$. So suppose for contradiction that there is a sequence of $(R_{\text{ex}} \cup R_{\text{all}})$-links from u to u'. Let $u'' \in A_{\xi'',\Pi''}$, say, be the penultimate element of this sequence. Certainly, there is no R_{all}-link from u'' to u', since this would require $\pi'' \in \xi''$ and $\pi''' \in \xi'$ with $\pi'' \ll \pi'''$. But by Lemma 7, we would then have $\pi'' \in \Pi$, which is ruled out by (ii). On the other hand, if there were a R_{ex}-link from u'' to u', then we would have $\xi' \cap V \neq \emptyset$, and again by Lemma 7, $\xi' \subseteq \Pi$, whence $\xi' \cap V \cap \Pi \neq \emptyset$, which is ruled out by (iii). We consider next the case where $\langle \xi', \Pi' \rangle = \langle \xi, \Pi \rangle$. But then (iv) implies $\xi \cap V = \emptyset$, so that we may select $b = a^+_{\pi',\xi,\Pi,j}$, where $j = 1$ if $i = 0$ and $j = 0$ otherwise. Again, let u be the cell containing a and u' the cell containing b. Thus $u \neq u'$, whence certainly $a \neq b$. Moreover, $\mathfrak{A} \models \mu[b]$. Again, it remains to show that $\mathfrak{A} \not\models T[a,b]$. Suppose there is a chain $u = u_0, \ldots, u_k = u'$ of $(R_{\text{ex}}, \cup R_{\text{all}})$-links. By construction of R_{ex}, we must have $u_{k-1} R_{\text{all}} u_k$, since

$j \leq 1$. Then there exists $\pi'' \in \xi_{k-1}$ and $\pi''' \in \xi_k$ such that $\pi'' \ll \pi'''$. Then, certainly, $k > 1$ since, otherwise, $\xi_0 = \xi_k = \xi$ contains both π'' and π''' with $\pi'' \ll \pi'''$ and $\xi \cap V = \emptyset$, which contravenes (C6). But if $k > 1$, then $\pi \in \Pi$ by Lemma 7, which contravenes (C5). Thus, we have shown that $\mathfrak{A} \not\models T[a, b]$ as required.

3. ψ is $\forall(\pi \to \forall(\pi' \to T))$, where $\pi \neq \pi'$: Suppose $\mathcal{C} \models \psi$, and that $a, a' \in A$ with $\mathrm{ftp}^{\mathfrak{A}}[a] = \pi$ and $\mathrm{ftp}^{\mathfrak{A}}[a'] = \pi'$. Write $a = a^p_{\pi,\xi,\Pi,i}$ and $a' = a^{p'}_{\pi',\xi',\Pi',j}$. We must show that $\mathfrak{A} \not\models T[a, a']$. We consider the three possibilities in the definition of $\mathcal{C} \models \psi$. (i) By construction of $\mathfrak{A}$, π and π' both occur in $\mathcal{C}$, so the first possibility does not arise. (ii) Suppose that $\pi \ll \pi'$. If a and a' are in different cells, then then we immediately have $aR_{\mathrm{all}}a'$. If, on the other hand, a and a' are in the same cell, then since $\pi \neq \pi'$, by (C4), $\neg\hat{T} \notin \bigcup \xi$, whence $aR_{\mathrm{cell}}a'$. (iii) Suppose that there is a single clique super-type $\langle \xi, \Pi \rangle \in \Omega$ such that ξ contains either π or π' and that $\xi \cap V \neq \emptyset$. By the construction of $\mathfrak{A}$, a and a' belong to the same cell $A_{\xi,\Pi,0}$, and again by (C4), $\neg\hat{T} \notin \bigcup \xi$, whence $aR_{\mathrm{cell}}a'$. In all cases, then, $\mathfrak{A} \models T[a, a']$, as required.

4. ψ is $\forall(\pi \to \forall(\pi' \to \neg T))$, where $\pi \neq \pi'$: Suppose $\mathcal{C} \models \psi$, and that $a, a' \in A$ with $\mathrm{ftp}^{\mathfrak{A}}[a] = \pi$ and $\mathrm{ftp}^{\mathfrak{A}}[a'] = \pi'$. Write $a = a^p_{\pi,\xi,\Pi,i}$ and $a' = a^{p'}_{\pi',\xi',\Pi',j}$. From the definition of $\mathcal{C} \models \psi$, we have $\pi' \notin \xi \cup \Pi$, whence $\xi \neq \xi'$. Thus, a and a' occupy different cells, say, u and u', respectively. By Lemma 7, there is no chain $u = u_0, \ldots, u_k = u'$ of $(R_{\mathrm{ex}} \cup R_{\mathrm{all}})$-links. Therefore, $\mathfrak{A} \not\models T[a, a']$, as required.

5. ψ is $\forall(\pi \to \forall(\pi \to (= \vee\ T)))$: Suppose $\mathcal{C} \models \psi$, and that $a, a' \in A$ with $\mathrm{ftp}^{\mathfrak{A}}[a] = \mathrm{ftp}^{\mathfrak{A}}[a'] = \pi$ and $a \neq a'$. From the definition of $\mathcal{C} \models \psi$ and the construction of $\mathfrak{A}$, a and a' belong to the same set $A_{\xi,\Pi}$ and, moreover, $\xi \cap V \neq \emptyset$. It follows that a and a' belong to the cell $A_{\pi,\xi,\Pi,0}$. Since $a \neq a'$, by the construction of A, $\xi(\pi) = 2$, whence by (C4), ξ is not a soliton clique type, whence $aR_{\mathrm{cell}}a'$. Thus $\mathfrak{A} \models T[a, a']$, as required.

6. ψ is $\forall(\pi \to \forall(\pi \to (= \vee \neg T)))$: Suppose $\mathcal{C} \models \psi$. It follows immediately by construction of $\mathfrak{A}$ that no set $A_{\pi,\xi,\Pi,i}$ can have cardinality greater than 1. Now suppose $a, a' \in A$ with $\mathrm{ftp}^{\mathfrak{A}}[a] = \mathrm{ftp}^{\mathfrak{A}}[a'] = \pi$ and $a \neq a'$. Thus, a and a' are not in the same cell, and hence by Lemma 7, $\mathfrak{A} \models T[a, a']$ implies $\xi' \subseteq \Pi$, whence $\pi \in \Pi$, contradicting the definition of $\mathcal{C} \models \psi$. Thus, $\mathfrak{A} \not\models T[a, a']$, as required.

The cases $\forall\mu$ and $\exists\mu$ are routine. □

We can now achieve the goal of this section.

Lemma 13. *There exists a non-deterministic procedure which, when given a set Φ of basic formulas over a signature Σ, will terminate in time bounded by $g(2^{2^{g(|\Sigma|)}} + \|\Phi\|)$, for some fixed polynomial g not depending on Φ or Σ, and which has an accepting run if and only if Φ is satisfiable.*

Proof. Let Φ be given. By Lemma 4, the following are equivalent: Φ is satisfiable; Φ^* is satisfied in a quadratic structure; Φ^* is satisfiable. Observe that Φ^* (a set of basic formulas over some signature $\Sigma^* \supseteq \Sigma$) can be computed in time bounded by a polynomial function of Φ. By Lemmas 5 and 11, if Φ^* is satisfiable over a quadratic structure, then there exists a certificate $\mathcal{C}$, interpreting Σ^*, such that $\mathcal{C} \models \Phi^*$. By Lemmas 10 and 12, if there exists a certificate $\mathcal{C}$ over Σ^*, such that $\mathcal{C} \models \Phi^*$, then Φ^* is satisfiable. Evidently $\|\mathcal{C}\|$ is bounded by a doubly exponential function of $|\Sigma^*|$, and the condition $\mathcal{C} \models \Phi^*$ may be checked in time bounded by a polynomial function of $\|\Phi^*\| + \|\mathcal{C}\|$. □

It is worth remarking that the construction of certificates for satisfiability described in this section above depends heavily on flutedness, which is built into the notion of a basic formula and the associated apparatus of super-types. The most obvious difficulty in generalizing this result to the non-fluted case is posed by formulas of the form $\forall x(\pi(x) \to \forall y(\pi'(y) \to T(x,y) \vee T(y,x)))$, stating that every element satisfying some formula π must be related by T *in one direction or the other* to every element satisfying the formula π'. Such formulas were dealt with in [16] in the context of the *finite satisfiability* problem for $\mathcal{FL}^2$; however, the corresponding *satisfiability* problem remains open. In particular, no obvious way of augmenting the information contained in super-types appears to lead to a notion of certificates which would ensure the truth of such formulas.

3.2. The logic $\mathcal{FL}^2_=1\mathrm{T}$

In this section, we tackle the satisfiability problem for $\mathcal{FL}^2_=1\mathrm{T}$, the 2-variable fluted fragment with equality and a single, distinguished, transitive relation T. Our strategy is to reduce this problem to the corresponding problem for basic formulas, which we solved in Lemma 13. Unlike basic formulas, $\mathcal{FL}^2_=1\mathrm{T}$-formulas in general contain ordinary binary predicates; our principal task, consequently, is to get rid of them. A crucial role in this endeavour will be played by *normal-form* formulas in the style of [25]. In fact, we shall have recourse to several variant normal forms in the sequel. However, we begin with the most fundamental, which will also recur in Sec. 3.3. A formula of $\mathcal{FL}^m_=1\mathrm{T}$ is in *normal form* if it has the shape

$$\bigwedge_{i \in I} \forall^{m-1}(\mu_i \to \exists(\kappa_i \wedge \gamma_i)) \wedge \bigwedge_{j \in J} \forall^{m-1}(\nu_j \to \forall \delta_j), \tag{8}$$

where I and J are finite sets of indices, such that, for $i \in I$ and $j \in J$, μ_i and ν_j are quantifier-free $\mathcal{FL}^{m-1}_=1\mathrm{T}$-formulas, κ_i is a formula of any of the four forms $(T\wedge =)$, $(T\wedge \neq)$, $(\neg T\wedge =)$, $(\neg T\wedge \neq)$, and γ_i and δ_j are quantifier-free $\mathcal{FL}^m_=1\mathrm{T}$ formulas.

A few remarks on the forms (8) at this point may help to guide the reader's intuition. We call a conjunct of the form $\forall^{m-1}(\mu_i \to \exists(\kappa_i \wedge \gamma_i))$ a *witness requirement*. It ensures that any $(m-1)$-tuple of elements $\bar{a}$ satisfying μ_i can be extended to an m-tuple $\bar{a}b_i$ satisfying $\kappa_i \wedge \gamma_i$. We call the formula κ_i a *control formula*. It ensures that the relationship between the final element of $\bar{a}$

and the *witness* b_i is specified completely *vis-à-vis* the binary predicates T and $=$. We call a conjunct of the form $\forall^{m-1}(\nu_j \to \forall \delta_j)$ a *universal requirement*. It constrains the possible ways in which any $(m-1)$-tuple $\bar{a}$ satisfying ν_i may be extended to an m-tuple $\bar{a}b$. In the case where ν_j is the trivial formula, $\top$, a universal requirement constrains the fluted m-types that can be realized.

The following lemma is slightly modified from [18, Lemma 4.1], where it was proved for the fluted fragment without transitivity or equality. The proof, however, is virtually identical, and we may simply state:

Lemma 14. *Let φ be an $\mathcal{FL}^m_=1\mathrm{T}$-sentence. We can compute, in time bounded by a polynomial function of $\|\varphi\|$, a normal-form $\mathcal{FL}^m_=1\mathrm{T}$-formula ψ such that:* (i) $\models \psi \to \varphi$*; and* (ii) *any model of φ can be expanded to a model of ψ.*

In the context of a structure interpreting a relational signature, a *king* is an element whose fluted 1-type is not realized by any other element in that structure. The fluted 1-types of kings are called *royal*. (This terminology is taken from [3].) We make use of the well-known fact that, in two-variable logic, parts of structures may be duplicated as long as they contain no king. We use the formulation appearing in [16, Lemma 4.1]. The proof given there concerns two-variable first-order logic with a single distinguished predicate interpreted as a *partial order*; however, the proof for the (present) case in which it is interpreted as a *transitive relation* is identical, and we need not repeat it here.

Lemma 15. *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 for any pair of elements $a, b \in A_i$, there exist $a', b' \in A_1$ such that $\mathrm{ftp}^{\mathfrak{A}_i}[a,a] = \mathrm{ftp}^{\mathfrak{A}_1}[a',a']$ and $\mathrm{ftp}^{\mathfrak{A}_i}[a,b] = \mathrm{ftp}^{\mathfrak{A}_1}[a',b']$;*

(ii) *for all $a \in B_i$, $\mathrm{ftp}^{\mathfrak{A}_i}[a,a] = \mathrm{ftp}^{\mathfrak{A}_1}[f_i(a), f_i(a)]$ and, for all $b \in A_1 \setminus \{f_i(a)\}$, $\mathrm{ftp}^{\mathfrak{A}_i}[a,b] = \mathrm{ftp}^{\mathfrak{A}_1}[f_i(a), b]$ and $\mathrm{ftp}^{\mathfrak{A}_i}[b,a] = \mathrm{ftp}^{\mathfrak{A}_1}[b, f_i(a)]$;*

(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 $\mathrm{ftp}^{\mathfrak{A}_i}[a,b] = \mathrm{ftp}^{\mathfrak{A}_1}[f_i(a), f_j(b)]$;*

(iv) *$T^{\mathfrak{A}_i}$ is a transitive relation.*

Intuitively, the sets $B_2, \ldots, B_i$ are indistinguishable copies of the set of non-royal elements, B_1. This copying process may be continued indefinitely (or even infinitely); we require only finitely many iterations in this paper. This construction is useful in the case where φ is a normal-form formula of $\mathcal{FL}^2_=1\mathrm{T}$. For suppose φ has some model $\mathfrak{A}$, and let $\mathfrak{A}_i$ $(i \geq 1)$ be as constructed in Lemma 15. A simple check shows that $\mathfrak{A}_i \models \varphi$. Indeed, if $a, b \in A_i$ are distinct elements of A_i with b non-royal, there exists a b' in each of the sets

B_j $(1 \leq j \leq i)$ such that $\text{ftp}^{\mathfrak{A}_i}[a,b] = \text{ftp}^{\mathfrak{A}_i}[a,b']$. In other words: non-royal witnesses in $\mathfrak{A}_i$ occur at least i times.

We next introduce a variant normal form for $\mathcal{FL}^2_=1\text{T}$, motivated by the observation that the witnesses corresponding to the different witness requirements of (8) need not be distinct. Say that a formula of $\mathcal{FL}^2_=1\text{T}$ is in *spread normal form* if it has the shape

$$\bigwedge_{h\in H} \exists\lambda_h \wedge \bigwedge_{i\in I} \forall(\mu_i \to \exists(o_i \wedge \kappa_i \wedge \gamma_i))\wedge \bigwedge_{j\in J} \forall(\nu_j \to \forall\delta_j) \wedge \bigwedge_{i,i'\in I}^{i\neq i'} \forall(o_i \to \neg o_{i'}), \tag{9}$$

where H is an index set, the λ_h $(h \in H)$ are quantifier-free, unary formulas, the o_i $(i \in I)$ are unary predicates, and I, J, μ_i, ν_j, κ_i, γ_i, δ_j are as before. The essential change here is the insertion of the atoms o_i into the witness requirements, together with the addition of the conjuncts $\forall(o_i \to \neg o_{i'})$ for distinct indices i and i'. Thus, if an object satisfies μ_i for several indices i, the corresponding witnesses of the formula $\exists(o_i \wedge \kappa_i \wedge \gamma_i)$ for that element are all distinct, a feature which we shall rely on when eliminating ordinary binary predicates. The other change—more technical in character—is the addition of the conjuncts $\exists\lambda_h$. They will enable us to convert normal-form $\mathcal{FL}^2_=1\text{T}$-sentences into spread normal form without an unacceptable inflation in the size of the signature, as promised by the next lemma.

Lemma 16. *Let φ be a normal-form $\mathcal{FL}^2_=1\text{T}$-formula and $\Pi = \{\pi_1, \ldots, \pi_L\}$ a set of fluted 1-types over the signature of φ. We can compute, in time bounded by an exponential function of $\|\varphi\|$, a formula ψ in spread normal form, such that: (i) the signature of ψ is bounded in size by a polynomial function of $\|\varphi\|$; (ii) $\models \psi \to \varphi$; and (iii) if φ has a (finite) model in which Π is the set of royal fluted 1-types, then ψ has a (finite) model.*

Proof. Putting $m = 2$ in (8), let φ have the shape

$$\bigwedge_{i\in I} \forall(\mu_i \to \exists(\kappa_i \wedge \gamma_i)) \wedge \bigwedge_{j\in J} \forall(\nu_j \to \forall\delta_j).$$

Observe that, in any structure in which Π is the set of royal fluted 1-types, there are exactly L kings; denote the king with fluted 1-type π_k by c_k $(1 \leq k \leq L)$. Setting $\ell = \lceil\log(L+1)\rceil$, let $o_i, w_{i,0}, \ldots, w_{i,\ell-1}$ be a collection of fresh unary predicates, for each $i \in I$. Observe that the number of these new predicates is polynomially bounded as a function of $\|\varphi\|$. As a guide to intuition, think of the predicates $w_{i,0}, \ldots, w_{i,\ell-1}$ as representing the binary digits of some number between 0 and L (inclusive). Then a formula of the form $\pm w_{i,0} \wedge \cdots \wedge \pm w_{i,\ell}$ defines a number k $(0 \leq k \leq L)$ by taking the jth binary digit of k to be given by the polarity of the literal $\pm w_{i,j}$. We are invited to read the (unary) formula $\bar{w}\langle i, 0\rangle$ as stating that there is a non-royal *ith* witness for a, and $\bar{w}\langle i, k\rangle$ (for

$1 \leq k \leq L$) as stating that the kth king is an ith witness for a. We mention at this point that the role of the predicates o_i is to pick out pairwise disjoint collections of non-royal elements; we will say presently how these sets are chosen.

We define ψ to be the conjunction of the following formulas.

$$\bigwedge_{k=1}^{L} \exists \pi_k \tag{10}$$

$$\bigwedge_{i \in I} \forall(\bar{w}\langle i, 0\rangle \to \exists(o_i \wedge \kappa_i \wedge \gamma_i)) \tag{11}$$

$$\bigwedge_{i \in I} \bigwedge_{k=1}^{L} \forall(\bar{w}\langle i, k\rangle \to \forall(\pi_k \to (\kappa_i \wedge \gamma_i))) \tag{12}$$

$$\bigwedge_{i \in I} \forall\big(\mu_i \to \bigvee_{k=0}^{L} \bar{w}\langle i, k\rangle\big) \tag{13}$$

$$\bigwedge_{j \in J} \forall\,(\nu_j \to \forall \delta_j) \tag{14}$$

$$\bigwedge_{i,i' \in I}^{i \neq i'} \forall \neg(o_i \wedge o'_i). \tag{15}$$

Modulo trivial re-arrangement, ψ is in spread normal form. The conjuncts (10)–(13) clearly entail the witness requirements of φ. Moreover, the conjuncts (14) are the universal requirements. Thus, $\models \psi \to \varphi$. Suppose, on the other hand, $\mathfrak{A}_1 \models \varphi$, with the set of royal types in $\mathfrak{A}_1$ equal to Π. We may assume without loss of generality that $I = \{1, \ldots, s\}$. Let the set of kings in $\mathfrak{A}_1$ be $A_0 = \{c_1, \ldots, c_L\}$, where $\mathfrak{A}_1 \models \pi_k[c_k]$, and apply the construction of Lemma 15 to obtain the structures $\mathfrak{A}_2, \ldots, \mathfrak{A}_s$. Let $\mathfrak{B} = \mathfrak{A}_s$, a model of φ with domain $B = A_0 \cup B_1 \cup \cdots \cup B_s$. For all $a \in B$ and $i \in I$, if there exists a non-royal element $b \neq a$ such that $\mathfrak{B} \models \gamma_i[a, b]$, then there exists such a b in each of the sets $B_1, \ldots, B_s$. Now expand $\mathfrak{B}$ to a model $\mathfrak{B}^+ \models \psi$ by setting $o_i^{\mathfrak{B}^+} = B_i$, and interpreting the predicates $w_{i,0}, \ldots, w_{i,k-1}$ as follows. We set $\mathfrak{B}^+ \models \bar{w}\langle i, 0\rangle[a]$ if $\mathfrak{B} \models \mu_i[a]$ and there exists a *non-royal* b such that $\mathfrak{B} \models \gamma_i[a, b]$, and we set $\mathfrak{B}^+ \models \bar{w}\langle i, k\rangle[a]$ (for $1 \leq k \leq L$) if $\mathfrak{B} \models \mu_i[a]$ and there exists no *non-royal* b such that $\mathfrak{B} \models \gamma_i[a, b]$, but k is the smallest number such that $\mathfrak{B} \models \gamma_i[a, c_k]$. Otherwise, set the $w_{i,k}$ arbitrarily. It is then simple to check that $\mathfrak{B}^+ \models \psi$. □

Spread normal form formulas are useful because they can be easily converted into equisatisfiable sets of basic formulas using a simple device which we now introduce (and which will be re-used in Sec. 3.3). Let χ be any quantifier-free fluted formula in some signature Σ, and let $\Sigma^- \subseteq \Sigma$ be a smaller signature containing the distinguished predicates T and $\hat{T}$. Then there is a strongest quantifier-free fluted formula χ° in Σ^- such that $\models \chi \to \chi^\circ$. To obtain χ°, write χ in disjunctive normal form and then delete all literals involving predicates not in Σ^-. This can be done in time bounded by an exponential function of $\|\chi\|$.

Lemma 17. *Let χ be a quantifier-free $\mathcal{FL}_=$-formula in some signature Σ and let Σ^- be a signature such that $\{T, \hat{T}\} \subseteq \Sigma^- \subseteq \Sigma$. Now let χ° be the strongest quantifier-free $\mathcal{FL}_=$-formula in Σ^- such that $\models \chi \rightarrow \chi^\circ$, and let m be the maximum arity of any predicate in Σ and m^- the maximum arity of any predicate in Σ^-. If τ^- is a fluted m^--type over Σ^- such that $\models \tau^- \rightarrow \chi^\circ$, then τ^- can be extended to a fluted m-type $\tau \supseteq \tau^-$ over Σ such that $\models \tau \rightarrow \chi$.*

Proof. The disjuncts in the disjunctive normal form of χ are the fluted m-types τ over Σ such that $\models \tau \rightarrow \chi$. The lemma is then immediate from the construction of χ° just described. □

Lemma 18. *Let φ be a spread normal-form $\mathcal{FL}^2_=$1T-formula. We can compute, in time bounded by an exponential function of $\|\varphi\|$, a set Φ of basic formulas, such that: (i) the signature of Φ consists of the unary predicates occurring in φ together with the distinguished binary predicate T; (ii) $\models \varphi \rightarrow \bigwedge \Phi$; and (iii) any model of Φ can be expanded to a model of φ.*

Proof. Let φ in spread normal form be given. From (9), φ has the shape

$$\bigwedge_{h \in H} \exists \lambda_h \wedge \bigwedge_{i \in I} \forall(\mu_i \rightarrow \exists(o_i \wedge \kappa_i \wedge \gamma_i)) \wedge \bigwedge_{j \in J} \forall(\nu_j \rightarrow \forall \delta_j) \wedge \bigwedge_{i,i' \in I}^{i \neq i'} \forall(o_i \rightarrow \neg o_{i'}).$$

Let Σ be the signature of φ, and Σ^- the signature obtained by removing all ordinary binary predicates from Σ. If χ is a any quantifier-free $\mathcal{FL}$1T-formula in Σ, denote by χ° be the strongest quantifier-free $\mathcal{FL}$1T-formula in Σ^- such that $\models \chi \rightarrow \chi^\circ$.

To motivate the construction of Φ, suppose φ has a model $\mathfrak{A}$, select any element $a \in A$, and define $J' = \{j \in J : \mathfrak{A} \models \nu_j[a]\}$. If $i \in I$ is such that $\mathfrak{A} \models \mu_i[a]$, then there exists a witness b_i such that $\mathfrak{A} \models \chi[a, b_i]$, where χ is the (quantifier-free) formula

$$o_i \wedge \kappa_i \wedge \gamma_i \wedge \bigwedge_{j \in J'} \delta_j. \tag{16}$$

It follows that $\mathfrak{A} \models \chi^\circ[a, b_i]$, and furthermore, since $o_i \wedge \kappa_i$ is in the signature Σ^-, that $\models \chi^\circ \rightarrow (o_i \wedge \kappa_i)$.

Now let ψ be the conjunction of all the following formulas:

$$\bigwedge_{h \in H} \exists \lambda_h \tag{17}$$

$$\bigwedge_{i \in I} \bigwedge_{J' \subseteq J} \forall\big(\big(\mu_i \wedge \bigwedge_{j \in J'} \nu_j\big) \rightarrow \exists\big(o_i \wedge \kappa_i \wedge \gamma_i \wedge \bigwedge_{j \in J} \delta_j\big)^\circ\big) \tag{18}$$

$$\bigwedge_{J' \subseteq J} \forall\big(\bigwedge_{j \in J'} \nu_j \rightarrow \forall\big(\bigwedge_{j \in J'} \delta_j\big)^\circ\big) \tag{19}$$

$$\bigwedge_{i,i' \in I}^{i \neq i'} \forall \neg(o_i \wedge o'_i). \tag{20}$$

It is simple to check that $\models \varphi \to \psi$. Moreover, the formula ψ is in the signature Σ^-, and in fact, following some trivial re-arrangement, can be equivalently written as a conjunction of basic formulas, of the forms (B1)–(B8). The only non-obvious case in this regard is (18). We have already noted, however, that if χ is the formula (16), then $\models \chi^\circ \to \kappa_i$. Hence (18) can be equivalently written

$$\bigwedge_{i \in I} \bigwedge_{J' \subseteq J} \forall\big((\mu_i \wedge \bigwedge_{j \in J'} \nu_j) \to \exists\big[\kappa_i \wedge \big(o_i \wedge \kappa_i \wedge \gamma_i \wedge \bigwedge_{j \in J} \delta_j\big)^\circ\big]\big).$$

And re-writing this as a conjunction of formulas of the forms (B1), (B2) and (B8) in time bounded by an exponential function of $\|\varphi\|$ is completely straightforward.

It remains only to show that any model of ψ may be expanded to a model of φ. Suppose then that $\mathfrak{B}^- \models \psi$; we expand to a structure $\mathfrak{B}$ interpreting Σ over the same domain B as follows. Fix any $a \in B$, and let $J' = \{j \in J \mid \mathfrak{B}^- \models \nu_j[a]\}$. For each $i \in I$, if $\mathfrak{B}^- \models \mu_i[a]$, let χ be as in (16), so that, by (18), there exists $b_i \in B$ be such that $\mathfrak{B}^- \models \chi^\circ[a, b_i]$ (and hence $\mathfrak{B}^- \models o_i[b_i]$). Letting $\tau^- = \mathrm{ftp}^{\mathfrak{B}^-}[a, b_i]$, Lemma 17 guarantees that there exists a fluted 2-type τ over Σ consistent with τ^- such that $\models \tau \to \chi$. Therefore we may set the extensions of the ordinary binary predicates in Σ so that $\mathrm{ftp}^{\mathfrak{B}}[a, b] = \tau$. If a satisfies μ_i for several indices $i \in I$, then by (20), the corresponding elements b_i will be distinct; therefore, these assignments may be carried out for all the required values of i without interference. (This is the point of using spread normal form.) Indeed, we may perform these assignments for each $a \in A$, again without pairs of elements being considered twice, since only *fluted* literals occur in fluted 2-types. At the end of this process, all elements $a \in B$ have the witnesses required by the conjuncts $\forall(\mu_i \to \exists(o_i \wedge \kappa_i \wedge \gamma_i))$ of φ; moreover, no pair $\langle a, b\rangle$ for which $\mathrm{ftp}^{\mathfrak{B}}[a, b]$ has been defined can violate any of the conjuncts $\forall(\nu_j \to \forall\delta_j)$.

We now complete the definition of $\mathfrak{B}$ for all remaining pairs $\langle a, b\rangle$ in such a way that no conjunct $\forall(\nu_j \to \forall\delta_j)$ is violated. Suppose then that $\mathrm{ftp}^{\mathfrak{B}}[a, b]$ has not been defined, and let $J' = \{j \in J \mid \mathfrak{B}^- \models \nu_j[a]\}$. Setting χ now to be the formula $\bigwedge_{j \in J'} \delta_j$, it follows from (19) that $\mathfrak{B}^- \models \chi^\circ[a, b]$, whence, taking τ^- in Lemma 17 to be $\mathrm{ftp}^{\mathfrak{B}^-}[a, b]$, we may consistently set $\mathrm{ftp}^{\mathfrak{B}}[a, b] = \tau$ such that $\models \tau \to \chi$. At the end of this process, $\mathfrak{B}$ is fully defined, and $\mathfrak{B} \models \varphi$. □

Thus, we have the promised upper bound for the problem $Sat\mathcal{FL}^2_{=}1\mathrm{T}$.

Lemma 19. *The satisfiability problem for $\mathcal{FL}^2_{=}1\mathrm{T}$ is in 2-*NExpTime.

Proof. Let an $\mathcal{FL}^2_{=}1\mathrm{T}$-sentence φ be given. By Lemma 14, we may assume without loss of generality that φ is in normal form. Guess a set Π of fluted 1-types over the signature of φ and apply the procedure guaranteed by Lemma 16 to obtain, in time bounded by an exponential function of $\|\varphi\|$, a spread normal-form formula ψ_Π, over a signature bounded by a polynomial function of $\|\varphi\|$, such that $\models \psi_\Pi \to \varphi$, and, if φ has a (finite) model in which the set of royal fluted 1-types is Π, then ψ_Π has a such a model too. By Lemma 18 we may then obtain, in time bounded by an exponential function of $\|\psi_\Pi\|$, a set Φ_Π of basic formulas, over the signature consisting of the unary predicates of ψ_Π

together with the distinguished predicate T, such that Φ_Π is satisfiable over the same domains as ψ_Π. Thus, if φ is satisfiable over some domain, taking Π to be the set of royal 1-types realized, we find that ψ_Π, and hence also Φ_Π, is satisfiable over that domain. Conversely, is Φ_Π is satisfiable over some domain, so is ψ_Π, and hence φ. Thus, it suffices to check the satisfiability of each such Φ_Π, non-deterministically, in time bounded by a doubly exponential function of $\|\varphi\|$. But this we can do by Lemma 13, since the signature of Φ_Π is bounded by a polynomial function of $\|\varphi\|$. □

3.3. The logic $\mathcal{FL}^m_=1\mathrm{T}$

Finally, we show how the satisfiability problem for $\mathcal{FL}^m_=1\mathrm{T}$ ($m \geq 3$) can be reduced to the corresponding problem for $\mathcal{FL}^{m-1}_=1\mathrm{T}$, but with exponential blow-up. The following notion will be useful. Let X be a finite set. A *cover* of X is a set $M = \{C_1, \ldots, C_\ell\}$ of subsets of X such that $C_1 \cup \cdots \cup C_\ell = X$; the elements of M will be referred to as *patches*. A *minimal cover* of X is a cover M of X such that no proper subset of M is a cover of X. Denote by $MC(X)$ the set of minimal covers of X. Since no *minimal* cover of X can have more than $|X|$ patches and each patch is a subset of X, we have $|MC(X)| \leq 2^{|X|^2}$. If X is a set of integers, and M is a minimal cover of X, we may assume the patches of M to be enumerated in some standard way as $C_1, \ldots, C_\ell$.

Lemma 20. *Let φ be an $\mathcal{FL}^m_=1\mathrm{T}$-formula ($m \geq 3$). We can compute, in time bounded by an exponential function of $\|\varphi\|$, an $\mathcal{FL}^{m-1}_=1\mathrm{T}$-formula ψ such that φ and ψ are satisfiable over the same domains.*

Proof. By Lemma 14, we may take φ to be in normal form, as given by (8):

$$\bigwedge_{i \in I} \forall^{m-1}(\mu_i \to \exists(\kappa_i \wedge \gamma_i)) \wedge \bigwedge_{j \in J} \forall^{m-1}(\nu_j \to \forall \delta_j).$$

We may further assume without loss of generality that the indices in I are integers. Let Σ be the signature of φ and Σ^- the result of removing from Σ all predicates of arity m. If χ is any quantifier-free $\mathcal{FL}$-formula over Σ, let χ° be the strongest quantifier-free $\mathcal{FL}$-formula over Σ^- such that $\models \chi \to \chi^\circ$, as in Lemma 17. To motivate the ensuing construction, we suppose that φ has a model $\mathfrak{A}$, and describe how to expand $\mathfrak{A}$ to a structure $\mathfrak{A}^+$ interpreting certain additional predicates of arities $m-1$ and $m-2$. We then define the sought-after $\mathcal{FL}^{m-1}_=1\mathrm{T}$-formula ψ in such a way that $\mathfrak{A}^+ \models \psi$. To complete the proof, we show that any model $\mathfrak{B} \models \psi$ can be expanded to a model $\mathfrak{B}^+ \models \varphi$.

For each $I' \subseteq I$ and each $J' \subseteq J$, let $p_{I',J'}$ and $q_{J'}$ be fresh $(m-2)$-ary predicates. Further, for each minimal cover $M = \{C_1, \ldots, C_\ell\} \in MC(I')$ (enumerated in the standard way), let $p_{I',J',M}$ be a fresh $(m-2)$-ary predicate, and for each h ($1 \leq h \leq \ell$), let $p_{I',J',M,h}$ be a fresh $(m-1)$-ary predicate. The structure $\mathfrak{A}^+$ interprets these new predicates over A as follows. For any $(m-2)$-tuple $\bar{a}$ and any $I' \subseteq I$ and $J' \subseteq J$, we set $\mathfrak{A}^+ \models p_{I',J'}[\bar{a}]$ just in case there exists an $a \in A$ such that $\mathfrak{A} \models \mu_i[a, \bar{a}]$ for all $i \in I'$ and $\mathfrak{A} \models \nu_j[a, \bar{a}]$ for all $j \in J'$. Similarly, we set $\mathfrak{A}^+ \models q_{J'}[\bar{a}]$ just in case there exists an $a \in A$ such that

$\mathfrak{A} \models \nu_j[a, \bar{a}]$ for all $j \in J'$. Thus, $p_{I',J'}$ tells that a given $(m-2)$ tuple $\bar{a}$ is the 'tail' of an $(m-1)$-tuple $a\bar{a}$ satisfying the μ_i and ν_j specified in the respective sets I' and J'; and similarly for $q_{J'}$. Now suppose that $\mathfrak{A}^+ \models p_{I',J'}[\bar{a}]$, and pick some $a \in A$ such that $a\bar{a}$ satisfies the μ_i and ν_j as specified by I' and J'. (If there are several possibilities for a, choose one arbitrarily.) Since $\mathfrak{A} \models \varphi$, there exists, for each $i \in I'$, some $b'_i \in A$ such that $\mathfrak{A} \models \mu_i[a\bar{a}b'_i]$. Hence there exists a minimal cover $M = \{C_1, \ldots, C_\ell\}$ of I', and *distinct* elements $b_1, \ldots, b_\ell$ of A such that, for all h $(1 \leq h \leq \ell)$, the m-tuple $\langle a, \bar{a}, b_h \rangle$, satisfies the quantifier-free formula $\bigwedge_{i \in C_h}(\kappa_i \wedge \gamma_i) \wedge \bigwedge_{j \in J'} \delta_j$. Actually, we can say a little more. The various κ_i mentioned in this formula are all control formulas—i.e. of the forms $\pm T \wedge \pm =$. Since these four possibilities are mutually exclusive, the κ_i must be identical. Call a minimal cover $M = \{C_1, \ldots, C_\ell\}$ of I' *consistent* if, for each patch C_h, the formulas κ_i are all the same for all $i \in C_h$, and thus may be denoted κ_h. Write $MCC(I')$ for the set of minimal covers of I' that are consistent in this sense. Thus, there exist $M = \{C_1, \ldots, C_\ell\}$ in $MCC(I')$, and *distinct* elements $b_1, \ldots, b_\ell$ of A such that, for all h $(1 \leq h \leq \ell)$, $\mathfrak{A} \models \chi[a, \bar{a}, b_h]$, where χ is the quantifier-free formula

$$\kappa_h \wedge \bigwedge_{i \in C_h} \gamma_i \wedge \bigwedge_{j \in J'} \delta_j. \tag{21}$$

It follows that $\mathfrak{A} \models \chi^\circ[a, \bar{a}, b_h]$, and thence, since χ° features no predicates of arity m, that $\mathfrak{A} \models \chi^\circ[\bar{a}, b_h]$. Notice that we are using the fact that χ° is fluted: no predicate of arity $m-1$ or lower can 'see' the element a. Now set $\mathfrak{A}^+ \models p_{I',J',M}[\bar{a}]$, and for all h $(1 \leq h \leq \ell)$, set $\mathfrak{A}^+ \models p_{I',J',M,h}[\bar{a}, b_h]$. Thus, $p_{I',J',M}$ tells us that the required witnesses for $a\bar{a}$ may be selected in accordance with the minimal cover M of I', while the predicates $p_{I',J',M,h}$ actually identify those witnesses. Defining ψ to be the conjunction of the formulas

$$\bigwedge_{I' \subseteq I} \bigwedge_{J' \subseteq J} \forall^{m-1}\big(\bigwedge_{i \in I'} \mu_i \wedge \bigwedge_{j \in J'} \nu_j \rightarrow p_{I',J'}\big) \tag{22}$$

$$\bigwedge_{J' \subseteq J} \forall^{m-1}\big(\bigwedge_{j \in J'} \nu_j \rightarrow q_{J'}\big) \tag{23}$$

$$\bigwedge_{I' \subseteq I} \bigwedge_{J' \subseteq J} \forall^{m-2}\big(p_{I',J'} \rightarrow \bigvee_{M \in MCC(I')} p_{I',J',M}\big) \tag{24}$$

$$\begin{aligned}&\bigwedge_{I' \subseteq I} \bigwedge_{J' \subseteq J} \bigwedge_{M \in MCC(I')} \forall^{m-2}\big(p_{I',J',M} \rightarrow \\ &\qquad \bigwedge_{h=1}^{|M|} \exists\big(p_{I',J',M,h} \wedge \big(\kappa_h \wedge \bigwedge_{i \in C_h} \gamma_i \wedge \bigwedge_{j \in J'} \delta_j\big)^\circ\big)\big)\end{aligned} \tag{25}$$

$$\bigwedge_{J' \subseteq J} \forall^{m-2}\big(q_{J'} \rightarrow \forall\big(\bigwedge_{j \in J'} \delta_j\big)^\circ\big) \tag{26}$$

$$\bigwedge_{I' \subseteq I} \bigwedge_{J' \subseteq J} \bigwedge_{M \in MC(I')} \bigwedge_{1 \leq h < h' \leq |M|} \forall^{m-1}\neg(p_{I',J',M,h} \wedge p_{I',J',M,h'}). \tag{27}$$

we have $\mathfrak{A}^+ \models \psi$. Indeed, the truth of conjuncts (22)–(25) is immediate by construction of $\mathfrak{A}^+$; the conjuncts (26) follow straightforwardly from the universal requirements of φ; and the conjuncts (27) reflect the fact that, for each possible choice of I', J' and $M \in MCC(I')$, the various witnesses $b_1, \dots, b_\ell$ chosen in the construction of the predicates $p_{I',J',M,h}$ are distinct. Of course, ψ depends only on φ, and not on the structure $\mathfrak{A}$; nevertheless, we have shown that if φ is satisfiable over some domain A, then so is ψ. Observe also that ψ is in $\mathcal{FL}^{m-1}_{=}1\mathrm{T}$; in particular, there are no predicates of arity m.

We claim that, if ψ is satisfiable over some domain B, then so is φ. For suppose $\mathfrak{B} \models \psi$. We expand to a structure $\mathfrak{B}^+$ interpreting the m-ary predicates of φ. Fix for the moment some element a and $(m-2)$-tuple of elements $\bar{a}$, and define $I' = \{i \in I \mid \mathfrak{B} \models \mu_i[a, \bar{a}]\}$ and $J' = \{j \in J \mid \mathfrak{B} \models \nu_j[a, \bar{a}]\}$. It follows from (22) that $\mathfrak{B} \models p_{I',J'}[\bar{a}]$. Indeed, from (24), there exists a consistent minimal cover $M = \{C_1, \dots, C_\ell\}$ of I' such that $\mathfrak{B} \models p_{I',J',M}[\bar{a}]$, whence from (25), we can find elements $b_1, \dots, b_\ell$ such that, for all h $(1 \leq h \leq \ell)$, $\mathfrak{B} \models p_{I',J',M,h}[\bar{a}, b_h]$ and $\mathfrak{B} \models \chi^\circ[\bar{a}, b_h]$, where χ is as in (21). Letting $\tau^- = \mathrm{ftp}^{\mathfrak{B}}[\bar{a}, b_h]$, it follows from Lemma 17 that there exists a fluted m-type $\tau \supseteq \tau^-$ such that $\models \tau \rightarrow \chi$. Thus, we may interpret the m-ary predicates of φ in $\mathfrak{B}^+$ in such a way that $\mathrm{ftp}^{\mathfrak{B}^+}[a, \bar{a}, b_h] = \tau$. From (27), the b_h are all distinct, so that this may be done for all h $(1 \leq h \leq \ell)$ without clashes. Indeed, we may carry out this process for all m-tuples $(a, \bar{a})$, again without fear of clashes, since it is only *fluted* m-types that are being assigned. We have thus ensured that, however $\mathfrak{B}^+$ is completed, all of the witness requirements in φ are satisfied, and, furthermore, none of the fluted m-types so far assigned violates any of the universal requirements of φ. To complete the definition of $\mathfrak{B}^+$, let $\langle a, \bar{a}, b \rangle$ be an m-tuple for which the extensions of the m-ary predicates have not been fixed. Again, let $J' = \{j \in J \mid \mathfrak{B} \models \nu_j[a, \bar{a}]\}$. It follows from (23) that $\mathfrak{B} \models q_{J'}[\bar{a}]$, and thence from (26) that $\mathfrak{B} \models \left(\bigwedge_{j \in J'} \delta_j\right)^\circ [\bar{a}, b]$. Now let $\tau^- = \mathrm{ftp}^{\mathfrak{B}}[\bar{a}, b]$, so that, by Lemma 17, there exists a fluted m-type $\tau \supseteq \tau^-$ such that $\models \tau \rightarrow \left(\bigwedge_{j \in J'} \delta_j\right)$. Hence we may interpret the m-ary predicates of φ in $\mathfrak{B}^+$ so that $\mathrm{ftp}^{\mathfrak{B}^+}[a, \bar{a}, b] = \tau$. At the end of this process, $\mathfrak{B}^+ \models \bigwedge_{j \in J} \forall^m (\nu_j \rightarrow \forall \delta_j)$. Thus, $\mathfrak{B}^+ \models \varphi$. □

We remark that, in the proof of Lemma 20, no unary or binary predicates encountered in the structures considered were disturbed. Thus, the lemma does not hinge on any properties of the distinguished predicates $\hat{T}$ or T except their arity: it would hold whatever constraints are imposed on their interpretations. The only feature of the logic we are using here is flutedness.

We have finally reached the goal of this section.

Theorem 21. *The satisfiability problem for* $\mathcal{FL}^{m}_{=}1\mathrm{T}$ *is in* m-NExpTime *for all* $m \geq 1$.

Proof. The case $m = 1$ is trivial, so suppose $m \geq 2$. Let an $\mathcal{FL}^{m}_{=}1\mathrm{T}$-sentence φ be given. By Lemma 14, we may assume without loss of generality that φ is in normal form. We proceed by induction, starting with $m = 2$. The base case is Lemma 19. For the recursive case, Lemma 20 reduces the original problem to the corresponding problem for $m - 1$, but with an exponential blow-up. □

Before moving to the next section we obtain a corollary concerning the finite satisfiability problem.

Corollary 22. *The finite satisfiability problem for* $\mathcal{FL}^m_{=}1\mathrm{T}$ *is in* $(m+1)$-NExpTime.

Proof. The proof differs from the proof of Theorem 21 only in the base case, where we apply the fact that the finite satisfiability problem for FO^2 with one transitive relation and equality is decidable in 3-NExpTime [16]; this complexity bound obviously applies to $\mathcal{FL}_{=}1\mathrm{T}$. □

4. Fluted Logic with more Transitive Relations

In the previous section, we considered $\mathcal{FL}^m$ extended with a single transitive relation and equality. In this section we consider $\mathcal{FL}^2$ extended with more transitive relations. Specifically, we show that the satisfiability and finite satisfiability problems for $\mathcal{FL}^2_{=}2T$ (two-variable fluted logic with two transitive relations and equality) or for $\mathcal{FL}^2 3T$ (two-variable fluted logic with three transitive relations but without equality), are all undecidable.

A *tiling system* is a tuple $\boldsymbol{\mathcal{C}} = (\mathcal{C}, H, V)$, where $\mathcal{C}$ is a finite set of *tiles*, and H, $V \subseteq \mathcal{C} \times \mathcal{C}$ are the *horizontal* and *vertical* constraints. A *tiling* of $\mathbb{N}^2$ for $\boldsymbol{\mathcal{C}}$ is a function $f : \mathbb{N}^2 \to \mathcal{C}$, such that for all $X, Y \in \mathbb{N}$, $(f(X,Y), f(X+1,Y)) \in H$ and $(f(X,Y), f(X,Y+1)) \in V$. Intuitively, we think of f as assigning (a copy of) some tile in $\mathcal{C}$ to each point with integer coordinates in the upper-right quadrant of the plane: this assignment must respect the horizontal and vertical constraints, understood as a list of which tiles may be placed immediately to the right of—respectively, immediately above—which others. A tiling is *periodic* if there exist m, n such that, for all X and Y, $f(X+m,Y) = f(X,Y+n) = f(X,Y)$; we refer to m and n as *periods* of the tiling. Alternatively, such a periodic tiling of $\mathbb{N}^2$ can be seen as a tiling of the toroidal grid obtained by identifying all pairs of points (X,Y) and $(X+m,Y)$, and all pairs of points (X,Y) and $(X,Y+n)$. We also consider tilings of finite initial segments of $\mathbb{N}^2$. If m and n are natural numbers, denote by $\mathbb{N}^2_{m,n}$ the subset $[0, m-1] \times [0, n-1]$ of $\mathbb{N}^2$. A *tiling* of $\mathbb{N}^2_{m,n}$ is a function $f : \mathbb{N}^2_{m,n} \to \mathcal{C}$, such that for all X, Y $(0 \le X < m-1,\ 0 \le Y \le n-1)$, $(f(X,Y), f(X+1,Y)) \in H$ and for all X, Y $(0 \le X \le m-1,\ 0 \le Y < n-1)$, $(f(X,Y), f(X,Y+1)) \in V$. If f is a tiling (of either $\mathbb{N}^2_{m,n}$ or $\mathbb{N}^2$), we call the value $f(0,0)$ the *initial condition*, and, if f is a tiling of $\mathbb{N}^2_{m,n}$, we call the value $f(m-1,n-1)$ the *final condition*.

We shall require several undecidability results concerning tiling systems. The *infinite tiling problem with initial condition* is the following: given a tiling system $\boldsymbol{\mathcal{C}} = (\mathcal{C}, H, V)$ and a tile $C_0 \in \mathcal{C}$, does there exist a tiling of $\mathbb{N}^2$ for $\boldsymbol{\mathcal{C}}$ with initial condition C_0? The *finite tiling problem with initial and final conditions* is the following: given a tiling system $\boldsymbol{\mathcal{C}}$ and tiles $C_0, C_1 \in \mathcal{C}$, do there exist positive

m, n and a tiling of $\mathbb{N}^2_{m,n}$ for $\mathcal{C}$ with initial condition C_0 and final condition C_1? The *unconstrained infinite tiling problem* is the following: given a tiling system $\mathcal{C}$, does there exist a tiling of $\mathbb{N}^2$ for $\mathcal{C}$? The *periodic tiling problem* is the following: given a tiling system $\mathcal{C}$, does there exist a periodic tiling of $\mathbb{N}^2$ for $\mathcal{C}$?

The use of constrained tiling systems to prove undecidability results was pioneered in [28]; by encoding infinite runs of Turing machines as tilings of $\mathbb{N}^2$ (and finite runs as tilings of $\mathbb{N}^2_{m,n}$) it is relatively straightforward to show:

Proposition 23. *The infinite tiling problem with initial condition and the finite tiling problem with initial and final conditions are both undecidable.*

Equating a problem with its set of positive instances, we recall that problems A and B are *recursively inseparable* if there exists no decidable problem S such that $A \subseteq S$ and $B \cap S = \emptyset$. Obviously, if A and B are recursively inseparable, then neither problem is decidable. The following result requires an elaborate analysis (see e.g. [2, p. 90 and Appendix A] for a comprehensive treatment).

Proposition 24. *The periodic tiling problem and the complement of the unconstrained infinite tiling problem are recursively inseparable.*

Typically, fragments of first-order logic whose satisfiability problems are undecidable have undecidable finite satisfiability problems and vice versa, and Prop. 24 is sometimes used in this context to give simultaneous proofs that both problems for some fragment are undecidable. This, indeed, is the strategy we employ in Sec. 4.1. In Sec. 4.2, by contrast, we are forced to adopt a more straightforward, if rather less neat, approach based on Prop. 23.

4.1. The case of two transitive relations

In this section we show that both the satisfiability and the finite satisfiability problems for $\mathcal{FL}^2_=2\mathrm{T}$ are undecidable. (Recall from Example 2 that $\mathcal{FL}^2 2\mathrm{T}$ admits infinity axioms.) We do this by showing how to map a tiling system $\mathcal{C}$ effectively to an $\mathcal{FL}^2_=2\mathrm{T}$-formula $\eta_{\mathcal{C}}$ in such a way that: (i) if $\mathbb{N}^2$ has a periodic tiling for $\mathcal{C}$, then $\eta_{\mathcal{C}}$ is finitely satisfiable; and (ii) if $\eta_{\mathcal{C}}$ is satisfiable, then $\mathbb{N}^2$ has a tiling for $\mathcal{C}$. The result then follows from Prop. 24.

The formula $\eta_{\mathcal{C}}$ features a conjunct φ_{grid} whose canonical model, shown in Fig. 4, has as its domain the unbounded integer plane $\mathbb{Z}^2$. The signature of φ_{grid} consists of the two distinguished binary predicates T_1 and T_2, together with the unary predicates $c_{i,j}$ $(0 \leq i, j \leq 3)$, which we call *local address predicates*. The element with coordinates (X, Y) satisfies the local address predicate $c_{i,j}$, where $i = X \mod 4$ and $j = Y \mod 4$; and the T_1- and T_2-links connect nearby elements in the regular pattern shown. We proceed to construct φ_{grid} as the conjunction of (28)-(36).

There is an initial element:

$$\exists c_{0,0}. \tag{28}$$

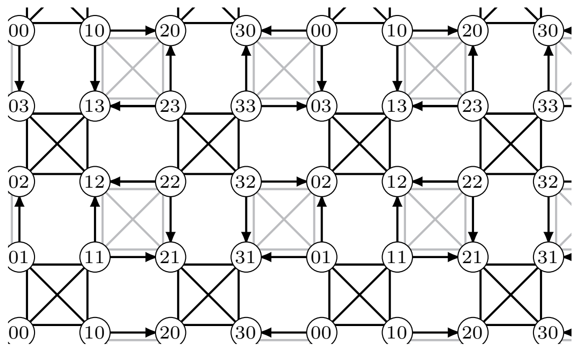

Figure 4: Canonical model of φ_{grid} over $\mathbb{N}^2$, showing the two transitive relations T_1 and T_2. Edges without arrows represent connections in both direction. Elements are labelled with the indices of the local address predicate c_{ij} they satisfy.

The predicates $c_{i,j}$ enforce a partition of the universe:

$$\forall\big(\dot{\bigvee}_{0\leq i\leq 3}\dot{\bigvee}_{0\leq j\leq 3} c_{i,j}\big). \tag{29}$$

Transitive paths do not connect distinct elements with the same local address:

$$\bigwedge_{0\leq i,j\leq 3} \forall(c_{i,j} \to \forall((T_1 \vee T_2) \wedge c_{i,j} \to =)). \tag{30}$$

Each element belongs to a 4-element T_1-clique:

$$\begin{aligned}\bigwedge_{i,j\in\{0,2\}} \forall \quad &\big((c_{i,j} \to \exists(T_1 \wedge c_{i+1,j})) \wedge (c_{i+1,j} \to \exists(T_1 \wedge c_{i+1,j+1}))\wedge \\ &(c_{i+1,j+1} \to \exists(T_1 \wedge c_{i,j+1})) \wedge (c_{i,j+1} \to \exists(T_1 \wedge c_{i,j}))\big).\end{aligned} \tag{31}$$

Each element belongs to a 4-element T_2-clique:

$$\begin{aligned}\bigwedge_{i,j\in\{1,3\}} \forall \quad &\big((c_{i,j} \to \exists(T_2 \wedge c_{i+1,j})) \wedge (c_{i+1,j} \to \exists(T_2 \wedge c_{i+1,j+1}))\wedge \\ &(c_{i+1,j+1} \to \exists(T_2 \wedge c_{i,j+1})) \wedge (c_{i,j+1} \to \exists(T_2 \wedge c_{ij}))\big).\end{aligned} \tag{32}$$

Certain pairs of elements connected by one transitive relation are also connected by the other one, specifically:

$$\bigwedge_{i=0,2} \forall(c_{i,i} \to \forall((T_1 \vee T_2) \wedge (c_{i,i-1} \vee c_{i-1,i}) \to (T_1 \wedge T_2)) \tag{33}$$

$$\bigwedge_{i=1,3} \forall(c_{i,i} \to \forall((T_1 \vee T_2) \wedge (c_{i,i+1} \vee c_{i+1,i}) \to (T_1 \wedge T_2)) \tag{34}$$

$$\bigwedge_{i=0,2} \forall(c_{i,i+1} \to \forall((T_1 \vee T_2) \wedge (c_{i,i+2} \vee c_{i-1,i+1}) \to (T_1 \wedge T_2)) \tag{35}$$

$$\bigwedge_{i=1,3} \forall (c_{i,i-1} \to \forall ((T_1 \vee T_2) \wedge (c_{i,i} \vee c_{i,i-2}) \to (T_1 \wedge T_2)). \quad (36)$$

A quick check shows that the structure depicted in Fig. 4 is indeed a model of φ_{grid}. Actually, a little more is true. As this structure repeats in a regular fashion, we easily obtain a *finite* model of φ_{grid} over the toroidal grid $\mathbb{Z}_{4m} \times \mathbb{Z}_{4n}$ (for all $m, n > 0$) by identifying elements from columns 0 and $4m$ and from rows 0 and $4n$. Observe that the formulas (31) and (32) work in tandem with (30). Specifically, both (31) and (32) generate, for a given element a with local address $c_{i,j}$ four new elements of certain local addresses such that the fourth element, say a', has the same local address as the element a. Formula (30) then implies $a = a'$, hence the element a is a member of a 4-element T_1-clique and a member of a (distinct) 4-element T_2-clique; members of these cliques can be uniquely identified by their local addresses (cf. Fig. 4).

We now show that, conversely, if $\mathfrak{A} \models \varphi_{grid}$, then the standard grid $\mathbb{N}^2$ can be embedded in $\mathfrak{A}$. The core of the proof is the (rather technical) Lemma 25, concerning the formulas $h_{i,j}$ and $v_{i,j}$, defined, for all i, j in the range $0 \leq i, j < 4$, as follows:

$$h_{i,j} := \begin{cases} T_1 \wedge c_{i+1,j} & \text{if } i \text{ is even} \\ T_2 \wedge c_{i+1,j} & \text{otherwise} \end{cases} \qquad v_{i,j} := \begin{cases} T_1 \wedge c_{i,j+1} & \text{if } j \text{ is even} \\ T_2 \wedge c_{i,j+1} & \text{otherwise.} \end{cases}$$

Intuitively, the lemma states that every element a satisfying $c_{i,j}$ is related by $h_{i,j}$ to some 'horizontal neighbour', b, and by $v_{i,j}$ to some 'vertical neighbour', a' (both satisfying the expected local address predicates); moreover, the resulting system of horizontal and vertical neighbours exhibits the usual grid confluence pattern.

Lemma 25. *In any model $\mathfrak{A}$, of φ_{grid}, the following hold for any i, j in the range $0 \leq i, j < 4$:*

$$\mathfrak{A} \models c_{i,j}[a] \;\Rightarrow\; \text{there exists } b \text{ s.t. } \mathfrak{A} \models h_{i,j}[a, b] \text{ and } a' \text{ s.t. } \mathfrak{A} \models v_{i,j}[a, a'] \quad (37)$$

$$\mathfrak{A} \models c_{i,j}[a] \wedge h_{i,j}[a, b] \wedge v_{i,j}[a, a'] \wedge v_{i+1,j}[b, b'] \quad \Rightarrow \quad \mathfrak{A} \models h_{i,j+1}[a', b']. \quad (38)$$

Proof. Let $a \in A$ and $\mathfrak{A} \models c_{i,j}[a]$. The existence of b in (37) is immediate from (31) for i, j even, and from (32) for i, j odd. Suppose i is even and j is odd. By the last conjunct of (31), there is $a_1 \in A$ such that $\mathfrak{A} \models T_1[a, a_1] \wedge c_{i,j-1}[a_1]$. By (31) again, there are $a_2, a_3, a_4 \in A$ such that $\mathfrak{A} \models T_1[a_1, a_2] \wedge c_{i+1,j-1}[a_2] \wedge T_1[a_2, a_3] \wedge c_{i+1,j}[a_3] \wedge T_1[a_3, a_4] \wedge c_{i,j}[a_4]$. By transitivity of T_1, $\mathfrak{A} \models T_1[a, a_4]$ and by (30), $a = a_4$, so the elements a, a_1, a_2, a_3 form a T_1-clique in $\mathfrak{A}$, hence $T_1[a, a_3]$ holds and, indeed, $\mathfrak{A} \models h_{i,j}[a, a_3]$. In the same way we show the existence of b when i is odd and j even, and, also, the existence of a'. We should regard the witnesses for the formulas $\exists h_{i,j}$ and $\exists v_{i,j}$ with respect to any element a are the horizontal and vertical neighbours, respectively, of a.

We now establish (38) proceeding separately for the possible indices i and j. Consider first the case $i = j = 0$, and suppose a, a', b and b' are elements of $\mathfrak{A}$ such that $\mathfrak{A} \models c_{0,0}[a] \wedge T_1[a,b] \wedge c_{1,0}[b] \wedge T_1[a,a'] \wedge c_{0,1}[a'] \wedge T_1[b,b'] \wedge c_{1,1}[b']$. By (31) b' is a member of a T_1-clique consisting of elements of local addresses $c_{1,1}, c_{0,1}, c_{0,0}, c_{1,0}$. Since by (30) the relation T_1 does not connect distinct elements of the same local address, a' belongs to the T_1-clique of b', so $\mathfrak{A} \models T_1[a',b']$, and the claim follows.

Consider now the case $i = 3$, $j = 0$, and suppose a, a', b and b' are elements such that $\mathfrak{A} \models c_{3,0}[a] \wedge T_1[a,a'] \wedge c_{3,1}[a'] \wedge T_2[a,b] \wedge c_{0,0}[b] \wedge T_1[b,b'] \wedge c_{0,1}[b']$. Applying (32) together with (30) to b, we see that b is a member of a 4-element T_2-clique consisting of elements of local addresses $c_{0,0}, c_{3,0}, c_{3,3}, c_{0,3}$. By (30), a is a member of this clique, whence $\mathfrak{A} \models T_2[b,a]$. By (33), $\mathfrak{A} \models T_1[b,a]$. Moreover, b' is in a T_1-clique of b, and so $\mathfrak{A} \models T_1[b',b]$. By transitivity of T_1, $\mathfrak{A} \models T_1[b',a']$. Now, by (35), $\mathfrak{A} \models T_2[b',a']$. By (32), a' is a member of a T_2-clique that, by (30), must contain b'. Hence $h_{3,0}[a',b']$ holds and the claim follows.

The remaining cases are dealt with similarly. □

Using Lemma 25, we show how the infinite grid $\mathbb{N}^2$ can be embedded into any model $\mathfrak{A} \models \varphi_{grid}$. Specifically, we define a function $\iota : \mathbb{N}^2 \to A$ as follows. Set $\iota(0,0)$ to be some witness for (1). By (37), we may choose $\iota(1,0), \iota(2,0), \ldots$ such that, for all $X \geq 0$, setting $i = X \mod 4$, we have $\mathfrak{A} \models h_{i,0}[\iota(X,0), \iota(X+1,0)]$; and then, for every $X \geq 0$, we may choose $\iota(X,1), \iota(X,2), \ldots$ such that for every $Y \geq 0$, setting $j = Y \mod 4$, we have $\mathfrak{A} \models v_{i,j}[\iota(X,Y), \iota(X,Y+1)]$. A simple induction on Y using (38) then shows that, for all X and Y, $\mathfrak{A} \models h_{i,j}[\iota(X,Y), \iota(X+1,Y)]$.

We can now effectively map any tiling system $\boldsymbol{\mathcal{C}}$ to an $\mathcal{FL}^2_=$2T-formula $\eta_{\boldsymbol{\mathcal{C}}}$, formed by the conjunction of φ_{grid} with the following formulas.

Each node encodes precisely one tile:

$$\forall \Big(\bigvee_{C \in \mathcal{C}} C \wedge \bigwedge_{C \neq D} (\neg C \vee \neg D) \Big). \tag{39}$$

Adjacent tiles respect H and V:

$$\bigwedge_{C \in \mathcal{C}} \bigwedge_{0 \leq i,j < 4} \forall \Big(C \wedge c_{i,j} \to \forall \big((h_{i,j} \to \bigvee_{C' : (C,C') \in H} C') \wedge (v_{i,j} \to \bigvee_{C' : (C,C') \in V} C') \big) \Big). \tag{40}$$

We make two observations. (i) If f is a periodic tiling of $\mathbb{N}^2$ for $\boldsymbol{\mathcal{C}}$, with periods m and n, then we can take the $4m \times 4n$ toroidal model of φ_{grid}, and expand to a model of $\eta_{\boldsymbol{\mathcal{C}}}$ by setting any predicate $C \in \mathcal{C}$ to be satisfied by $(X,Y) \in \mathbb{N}^2$ just in case $f(X,Y) = C$. It is a simple matter to check that $\eta_{\boldsymbol{\mathcal{C}}}$ is true in the resulting structure. (ii) If $\mathfrak{A} \models \eta_{\boldsymbol{\mathcal{C}}}$, then $\mathfrak{A} \models \varphi_{grid}$, and so there exists a grid embedding $\iota : \mathbb{N}^2 \to A$. We then define a function $f : \mathbb{N}^2 \to \mathcal{C}$ by setting $f(X,Y)$ to be the unique tile $C \in \mathcal{C}$ such that $\mathfrak{A} \models C[\iota(X,Y)]$, which is well-defined by (39). By (40), f is a tiling for $\boldsymbol{\mathcal{C}}$.

Theorem 26. *The satisfiability problem and finite satisfiability problems for* $\mathcal{FL}^2_=$2T *are both undecidable.*

Proof. It suffices to show that $Sat(\mathcal{FL}^2_{=}2\mathrm{T})$ and the complement of the problem $FinSat(\mathcal{FL}^2_{=}2\mathrm{T})$ are recursively inseparable. Now the mapping $\boldsymbol{\mathcal{C}} \mapsto \eta_{\boldsymbol{\mathcal{C}}}$ constructed above is effective. But we have just shown that: (i) if $\mathbb{N}^2$ has a periodic tiling for $\boldsymbol{\mathcal{C}}$, then $\eta_{\boldsymbol{\mathcal{C}}}$ is finitely satisfiable; and (ii) if $\eta_{\boldsymbol{\mathcal{C}}}$ is satisfiable, then $\mathbb{N}^2$ has a tiling for $\boldsymbol{\mathcal{C}}$. The theorem then follows by Prop. 24. □

A quick check reveals that the formula $\eta_{\boldsymbol{\mathcal{C}}}$ lies in the guarded fragment of first-order logic. Moreover, the proof of Lemma 25 remains valid even if T_2 is required to be an equivalence relation. Thus we have:

Corollary 27. *The satisfiability problem and the finite satisfiability problems for the intersection of* $\mathcal{FL}^2_{=}2\mathrm{T}$ *with the guarded fragment are both undecidable. This result continues to hold if, in place of* $\mathcal{FL}^2_{=}2\mathrm{T}$*, we have* $\mathcal{FL}^2_{=}1\mathrm{T}1\mathrm{E}$*, the two-variable fluted fragment together with identity, one transitive relation and one equivalence relation.*

We conclude Sec. 4 by remarking that decidability of the satisfiability and the finite satisfiability problems for $\mathcal{FL}^m2\mathrm{T}$ remains open for every $m \geq 2$. We showed in Example 2 that these two problems are distinct.

4.2. The case of three transitive relations

In this section we show that the satisfiability problem and the finite satisfiability problem for $\mathcal{FL}^2 3\mathrm{T}$ are both undecidable. (Note that equality is not available in this logic.) We start by reducing the infinite tiling problem to the satisfiability problem.

We write a formula φ_{grid} whose canonical model has as its domain the upper-right quadrant of the integer plane, $\mathbb{N}^2$. Our construction employs a *boustrophedon*,[1] that is, a bijection $\varsigma : \mathbb{N} \mapsto \mathbb{N}^2$, such that for all n, $\varsigma(n+1)$ is a grid neighbour of $\varsigma(n)$. Many such functions exist; we shall avail ourselves of the one depicted by the thick grey arrow in Fig. 5. Denoting $\varsigma(t)$ by (X_t, Y_t), it is defined inductively by setting $\varsigma(0) = (0,0)$ and, for all $t \geq 0$:

$$\varsigma(t+1) = \begin{cases} (X_t, Y_t+1) & \text{if } X_t = 0 \text{ and } Y_t \text{ even,} \\ (X_t+1, Y_t) & \text{if } Y_t > X_t \text{ and } Y_t \text{ odd,} \\ (X_t, Y_t-1) & \text{if } X_t \geq Y_t > 0 \text{ and } X_t \text{ odd,} \\ (X_t+1, Y_t) & \text{if } Y_t = 0 \text{ and } X_t \text{ odd,} \\ (X_t, Y_t+1) & \text{if } X_t > Y_t \geq 0 \text{ and } X_t \text{ even,} \\ (X_t-1, Y_t) & \text{if } Y_t \geq X_t > 0 \text{ and } Y_t \text{ even.} \end{cases} \tag{41}$$

Beginning with $\varsigma(0) = (0,0)$, the first clause of (41) yields $\varsigma(1) = (0,1)$. If now $u \in \mathbb{N}$ is such that $\sigma(u) = (0,k)$ for some odd k, the remaining clauses yield as

[1] In linguistics: written from right to left and from left to right in alternate lines. Etymology: Greek, literally 'as an ox turns in ploughing', from bous 'ox' + -strophos 'turning'.

subsequent values:

$$
\begin{array}{l}
\text{a rightward row } (0,k),\ldots,(k-1,k) \text{ of } k \text{ points,}\\
\text{a downward column } (k,k),\ldots,(k,0) \text{ of } k+1 \text{ points,}\\
\text{an upward column } (k+1,0),\ldots,(k+1,k+1) \text{ of } k+2 \text{ points, and}\\
\text{a leftward row } (k,k+1),\ldots,(0,k+1) \text{ of } k+1 \text{ points.}
\end{array}
\tag{42}
$$

At that point, the first clause again yields the point $(0,k+2)$, and the process repeats.

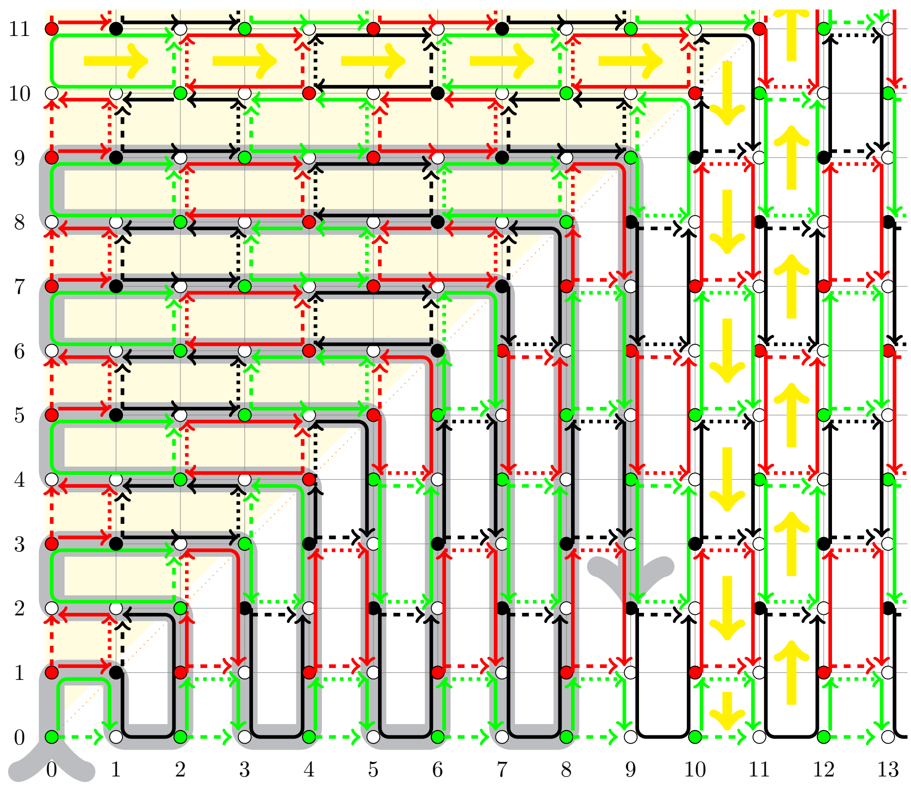

Figure 5: Intended expansion of the $\mathbb{N}^2$ grid with three transitive relations T_0, T_1 and T_2. Solid arrows represent *generated* edges; direction of these arrows follows the boustrophedon order indicated by the thick grey path starting at the origin. Dashed arrows represent edges induced by transitivity. Dotted arrows represent edges induced by *transfer* formulas. The meaning of the yellow arrows is explained later (cf. page 50).

The formula φ_{grid} comprises a large number of conjuncts. To help give an overview of the construction, we have organized these conjuncts into four groups called *labelling constraints*, *generation rules*, *transfer formulas*, and *control formulas*; each of them secures a particular property exhibited by its models. We use the following notational conventions. If i is an integer, $i/2$ indicates integer division without remainder (e.g., $5/2 = 2$); moreover, $\lfloor i \rfloor_k$ denotes the remainder of i on division by k, and $\lfloor i \rfloor$ (i.e., without the subscript) denotes $\lfloor i \rfloor_6$.

The signature of φ_{grid} features, in addition to the (distinguished) binary predicates T_0, T_1 and T_2, the unary predicates $c_{i,j}$ and $d_{i,j}$ ($0 \leq i, j \leq 5$) and bt, lf, dg and dg$^+$. We call the $c_{i,j}$ and $d_{i,j}$ *local address predicates*, and the predicates bt, lf, dg and dg$^+$ *control predicates*. To aid visualization, we also refer to the distinguished predicates T_0, T_1 and T_2 as *colours*, often replacing them by the respective synonyms black, green and red.

To motivate the definition of φ_{grid}, we first describe an intended model $\mathfrak{A}_0 \models \varphi_{grid}$, illustrated in Fig. 5. As we shall see in the sequel (Lemma 36), this model is *canonical*: it embeds into any other model of φ_{grid}. The domain of $\mathfrak{A}_0$ is the set of grid points $\mathbb{N}^2$. We interpret the *local address predicates* as follows: if $a = (X, Y)$, then $\mathfrak{A}_0 \models c_{i,j}[a]$ just in case $Y > X$, $i = \lfloor X \rfloor$ and $j = \lfloor Y \rfloor$, and $\mathfrak{A}_0 \models d_{i,j}[a]$ just in case $Y \leq X$, $i = \lfloor X \rfloor$ and $j = \lfloor Y \rfloor$. (Remember that $\lfloor n \rfloor$ denotes n modulo 6.) Thus, $c_{i,j}$ indicates an element strictly above the diagonal, and $d_{i,j}$ an element on or strictly below the diagonal, with the indices giving the coordinates of that element modulo 6. The control predicates are interpreted as follows: if $a = (X, Y)$, then $\mathfrak{A}_0 \models \text{bt}[a]$ just in case $Y = 0$, $\mathfrak{A}_0 \models \text{lf}[a]$ just in case $X = 0$ and $Y > 0$, $\mathfrak{A}_0 \models \text{dg}[a]$ just in case $Y = X$, and $\mathfrak{A}_0 \models \text{dg}^+[a]$ just in case $Y = X + 1$. Thus, we gloss bt as "is on the bottom row", lf as "is on the left-most column, but not the origin", dg as "is on the diagonal" and dg$^+$ as "is on the super-diagonal". The colours T_0 (black), T_1 (green) and T_2 (red) are interpreted as the transitive closures of the solid or dotted arrows having the respective colours in Fig. 5. (The dashed arrows are induced by transitivity; we refer to them in the argument below.) For example: (3,2) is related by black to (3,1), (3,0), (4,0), (4,1), (4,2), by green to (4,2), (4,3), (4,4), (3,4) and by red to (3,1). Note that pairs of elements can be joined by up to two colours, and that very distant elements are always unrelated. The general pattern is that the arrows along *downward pointing* sequences cycle in pairs through the colours black, red, and green, while *upward pointing* sequences cycle through the colours black, green and red. Similarly for *leftward pointing* and *rightward pointing* sequences. Links between adjacent columns below the diagonal and adjacent rows above the diagonal follow a similar repeating pattern. Always remember also that these colours are *transitive* relations. We shall show in the course of the ensuing argument that any model $\mathfrak{A} \models \varphi_{grid}$ homomorphically embeds a copy of $\mathfrak{A}_0$, in a sense that we shall make precise.

We begin with two very simple conjuncts of φ_{grid}, which we together refer to as the *labelling constraints*. The first states that the local address predicates partition the universe:

$$\forall \Big(\dot{\bigvee_{0 \leq i,j \leq 5}} c_{i,j} \ \dot{\vee} \ \dot{\bigvee_{0 \leq i,j \leq 5}} d_{i,j} \Big). \tag{43}$$

The second relates local address predicates and the control predicates as suggested by the glosses above:

$$\forall(\mathrm{bt} \to \bigvee_{i=0}^{5} d_{i,0}) \wedge \forall(\mathrm{lf} \to \bigvee_{j=0}^{5} c_{0,j}) \wedge \forall(\mathrm{dg} \to \bigvee_{i=0}^{5} d_{i,i}) \wedge \forall(\mathrm{dg}^{+} \to \bigvee_{j=0}^{5} c_{j,\lfloor j+1 \rfloor}). \quad (44)$$

A quick check shows that these formulas are both true in $\mathfrak{A}_0$.

The next group of conjuncts of φ_{grid} all involve existential quantifiers, and we call them the *generation rules*. The first two are simple. There is an 'initial' element satisfying the predicates $d_{0,0}$, dg and bt.

$$\exists(d_{0,0} \wedge \mathrm{dg} \wedge \mathrm{bt}). \quad (45)$$

Any model $\mathfrak{A} \models \varphi_{grid}$ contains an element satisfying $d_{0,0}$, dg and bt. To aid readability in the sequel, we suppress reference to the model $\mathfrak{A}$ where it is clear from context; thus we simply write $d_{0,0}[a]$ instead of $\mathfrak{A} \models d_{0,0}[a]$, and so on. Hence, (45) gives us an element a_0 of $\mathfrak{A}$ such that $d_{0,0}[a_0]$, $\mathrm{bt}[a_0]$, $\mathrm{dg}[a_0]$. The next generation rule ensures that this element has a T_1- (i.e. green-) successor, say a_1, satisfying $c_{0,1}$, dg^+ and lf:

$$\forall(\mathrm{bt} \wedge \mathrm{dg} \to \exists(c_{0,1} \wedge \mathrm{dg}^{+} \wedge \mathrm{lf} \wedge T_1). \quad (46)$$

Applying this universally quantified formula to our element a_0, we obtain a witness a_1 such that $c_{0,1}[a_1]$, $\mathrm{lf}[a_1]$, $\mathrm{dg}^+[a_1]$ and $T_1[a_0, a_1]$. We remark that these formulas are satisfied in $\mathfrak{A}_0$, by putting $a_0 = (0,0)$ and $a_1 = (0,1)$.

The remaining generation rules are more complicated. Elements satisfying $d_{i,j}$ (but not both bt and dg) have successors given by the following conjuncts:

$$\bigwedge_{i=0,2,4} \bigwedge_{j=0}^{5} \forall(d_{i,j} \wedge \neg\mathrm{dg} \to \exists(d_{i,\lfloor j+1 \rfloor} \wedge \neg\mathrm{bt} \wedge T_{\lfloor j/2 \rfloor_3} \wedge T_{\lfloor (j+1)/2+1 \rfloor_3})) \quad (47)$$

$$\bigwedge_{i=1,3,5} \bigwedge_{j=0}^{5} \forall(d_{i,j} \wedge \neg\mathrm{bt} \to \exists(d_{i,\lfloor j-1 \rfloor} \wedge \neg\mathrm{dg} \wedge T_{\lfloor j/2+1 \rfloor_3} \wedge T_{\lfloor (j+1)/2-1 \rfloor_3})) \quad (48)$$

$$\bigwedge_{i=1,3,5} \forall(d_{i,0} \wedge \mathrm{bt} \wedge \neg\mathrm{dg} \to \exists(d_{\lfloor i+1 \rfloor,0} \wedge \mathrm{bt} \wedge \neg\mathrm{dg} \wedge T_0)) \quad (49)$$

$$\bigwedge_{i=0,2,4} \forall(d_{i,i} \wedge \neg\mathrm{bt} \wedge \mathrm{dg} \to \exists(c_{\lfloor i-1 \rfloor,i} \wedge \mathrm{dg}^{+} \wedge \neg\mathrm{lf} \wedge T_{\lfloor i/2-1 \rfloor_3} \wedge T_{\lfloor i/2 \rfloor_3})). \quad (50)$$

Likewise, elements satisfying $c_{i,j}$ have successors given by the following conjuncts:

$$\bigwedge_{j=0,2,4} \bigwedge_{i=0}^{5} \forall(c_{i,j} \wedge \neg\mathrm{lf} \to \exists(c_{\lfloor i-1 \rfloor,j} \wedge \neg\mathrm{dg}^{+} \wedge T_{\lfloor i/2-1 \rfloor_3} \wedge T_{\lfloor (i+1)/2 \rfloor_3})) \quad (51)$$

$$\bigwedge_{j=1,3,5} \bigwedge_{i=0}^{5} \forall(c_{i,j} \wedge \neg\mathrm{dg}^{+} \to \exists(c_{\lfloor i+1 \rfloor,j} \wedge \neg\mathrm{lf} \wedge T_{\lfloor i/2+1 \rfloor_3} \wedge T_{\lfloor (i+1)/2-1 \rfloor_3})) \quad (52)$$

$$\bigwedge_{j=0,2,4} \forall(c_{0,j} \wedge \mathrm{lf} \to \exists(c_{0,j+1} \wedge \mathrm{lf} \wedge \neg \mathrm{dg}^+ \wedge T_1)) \tag{53}$$

$$\bigwedge_{j=1,3,5} \forall(c_{j-1,j} \wedge \mathrm{dg}^+ \to \exists(d_{j,j} \wedge \mathrm{dg} \wedge \neg \mathrm{bt} \wedge T_{\lfloor (j+1)/2 \rfloor_3} \wedge T_{\lfloor (j+3)/2 \rfloor_3})). \tag{54}$$

To understand these formulas, consider our arbitrary model $\mathfrak{A} \models \varphi_{grid}$, and suppose that we have generated a sequence $a_0, \ldots, a_u$ of elements such that $\mathrm{lf}[a_u]$ and $c_{0,j}[a_u]$, with j odd. (Note that this is actually the case: the sequence a_0, a_1 satisfies these conditions.) Then either of the generation rules (52) and (54) applies to a_u, depending on whether $\mathrm{dg}^+[a_u]$. If $\neg\mathrm{dg}^+[a_u]$, then rule (52) applies, and a_u is related, this time by T_1 and T_2, to some element, say a_{u+1} such that $c_{1,j}[a_{u+1}]$. If also $\neg\mathrm{dg}^+[a_{u+1}]$, then rule (52) again applies, and a_{u+1} is related by T_1 and T_0 to some element, say a_{u+2}, such that $c_{2,j}[a_{u+2}]$. In this way we obtain a sequence of elements successively satisfying the predicates $c_{i,j}$, with i cycling through the numbers $0, \ldots, 5$ and j constant. We call any sequence $a_u, \ldots, a_s$ in which $\mathrm{lf}[a_u]$, $s \geq u$, and successive elements are generated by rule (52), a *rightward sequence*, and we refer to j as the *latitude* of that sequence. Necessarily, $j \in \{1, 3, 5\}$. If, in addition, $\mathrm{dg}^+[a_s]$, then this process stops, and we call $a_u, \ldots, a_s$ a *rightward row*. We remark that the existence of at least one rightward row in any model $\mathfrak{A} \models \varphi_{grid}$ is guaranteed: since $\mathrm{dg}^+[a_1]$ and $c_{0,1}[a_1]$, it follows that a_1 is a rightward row of length 1 and latitude 1. Notice in particular the sequence of colours with which each element in a rightward row is related to its successor, as specified by rule (52). Let us call these—in the order they appear in the formula—the *primary* colour and the *secondary* colour, respectively. Remembering our mnemonics *black*, *green* and *red* for T_0, T_1 and T_2, respectively, the sequences of primary and secondary colours on any rightward sequence are

green, green, red, red, black, black, green, ...
red, black, black, green, green, red, red, ...

repeating (as long as the sequence continues) with a period of six. In the structure $\mathfrak{A}_0$ (Fig. 5), the elements of the odd-numbered rows from the left margin to the super-diagonal (inclusive) form rightward rows; the primary colours are drawn below, and the secondaries above. (For example, look at row 9.) Summarizing these observations:

Lemma 28. *Let $a_u, \ldots, a_s$ be a rightward row of latitude $j \in \{1, 3, 5\}$ in a model of φ_{grid}. Then $c_{\lfloor h \rfloor, j}[a_{u+h}]$ for all h ($0 \leq h \leq s - u$), $\neg\mathrm{lf}[a_{u+h}]$ for all h ($0 < h \leq s - u$) and $\neg\mathrm{dg}^+[a_{u+h}]$ for all h ($0 \leq h < s - u$). Moreover, $T_{\lfloor h/2+1 \rfloor_3}[a_{u+h}, a_{u+h+1}]$ and $T_{\lfloor (h+1)/2-1 \rfloor_3}[a_{u+h}, a_{u+h+1}]$ for all h ($0 \leq h < s - u$).*

Now consider any rightward row $a_u, \ldots, a_s$ of latitude j. In particular, $\mathrm{dg}^+[a_s]$, and for some i, $c_{i,j}[a_s]$. In fact, by the third conjunct of (44), there is only one possibility for i, namely, $c_{\lfloor j-1 \rfloor, j}[a_s]$. In this case, therefore, rule (54) applies, and generates a successor, a_{s+1}, satisfying the predicates $d_{j,j}$, dg, and

$\neg$bt. To make the subsequent argument easier to follow, we write i for the current (odd) value of j. That is, we have $d_{i,i}[a_{s+1}]$, dg$[a_{s+1}]$, and $\neg$bt$[a_{s+1}]$. At this point, rule (48) applies, yielding a_{s+2} such that $d_{i,\lfloor i-1 \rfloor}[a_{s+2}]$. If $\neg$bt$[a_{s+2}]$, then rule (48) again applies, and we have $d_{i,\lfloor i-2 \rfloor}[a_{s+3}]$. In this way we obtain a sequence of elements satisfying the respective predicates $d_{i,j}$, with j cycling backwards (starting at $j = i$) through the numbers $0, \ldots, 5$, and i constant. We call any sequence $a_{s+1}, \ldots, a_t$ in which dg$[a_{s+1}]$, $t \geq s+1$, and successive elements are generated by rule (48), a *downward sequence*, and we refer to i as the *longitude* of that sequence. Necessarily, $i \in \{1, 3, 5\}$. If, in addition, bt$[a_t]$, then this process stops, and we call $a_{s+1}, \ldots, a_t$ a *downward column*. In the structure $\mathfrak{A}_0$, the elements of the odd-numbered columns from the diagonal down to the bottom edge (inclusive) form downward columns. Rule (48) again assigns two sequences of colours, repeating with a period of 6, linking successive elements of downward columns, and which we may again call the *primary* and *secondary* sequences. Since a downward column necessarily ends in an element satisfying a local address predicate $d_{i,0}$, we see from rule (48) that these sequences *end* as follows,

$$\ldots, \text{black red, red, green, green, black, black}$$
$$\ldots, \text{green, green, black, black, red, red, green}$$

having repeated in a period of 6 as long as there have been elements in the sequence. In the structure $\mathfrak{A}_0$ (Fig. 5), the elements of the odd-numbered columns from the diagonal down to the bottom form downward columns; the primary colours are drawn to the right, and the secondaries to the left.

We next add a conjunct to φ_{grid} ensuring that every model contains at least one downward column. Returning to our arbitrary model $\mathfrak{A} \models \varphi_{grid}$, recall that a_1 is a trivial (i.e. length 1) rightward row, and consider the next element, a_2, generated by (54). We know that dg$[a_2]$, $d_{1,1}[a_2]$ and $\neg$bt$[a_2]$. Thus rule (50) applies, and generates a successor, a_3, such that $d_{1,0}[a_3]$, $T_1[a_2, a_3]$ and $T_0[a_2, a_3]$ (for the chromatically inclined: T_1 is green and T_0 black). But we have already established that $T_1[a_0, a_1]$ and $T_1[a_1, a_2]$, whence, by transitivity, $T_1[a_0, a_3]$. (For the model $\mathfrak{A}_0$, this is indicated in by the dashed arrow from $(0, 0)$ to $(1, 0)$.) At this point, we add a conjunct to φ_{grid}:

$$\forall(d_{0,0} \wedge \text{bt} \rightarrow \forall(T_1 \wedge d_{1,0} \rightarrow \text{bt})).$$

In fact, this conjunct is subsumed by a collection conjuncts to be added presently, specifically by formula (62), below. Anticipating this addition, and bearing in mind that, as we have established, bt$[a_0]$ and $T_1[a_0, a_3]$, it follows from the assumption $\mathfrak{A} \models \varphi_{grid}$ that bt$[a_3]$. Thus a_2, a_3 is a downward column of length 2 and longitude 1. Summarizing these observations:

Lemma 29. *Any model of φ_{grid} contains elements $a_0, \ldots, a_3$ successively related by T_1, such that a_1 is a rightward row with latitude 1, and a_2, a_3 a downward column with longitude 1.*

Suppose now that $a_{s+1}, \ldots, a_t$ is any downward column of longitude $i \in \{1, 3, 5\}$. Since $\text{bt}[a_t]$ and $d_{i,0}[a_t]$ (i odd), generation rule (49) applies, yielding a_{t+1} such that $\text{bt}[a_{t+1}]$, $d_{\lfloor i+1 \rfloor, 0}[a_{t+1}]$ and $\neg\text{dg}[a_{t+1}]$. Now generation rule (47) therefore applies, yielding a_{t+2} such that $d_{\lfloor i+1 \rfloor, 1}[a_{t+2}]$. At this point, one of two possible generation rules apply, namely (47) and (50), depending on whether $\text{dg}[a_{t+2}]$. If $\neg\text{dg}[a_{t+2}]$, then rule (47) applies, and we have $d_{\lfloor i+1 \rfloor, 2}[a_{t+3}]$. In this way we obtain a sequence $a_{t+1}, a_{t+2}, \ldots$ satisfying the respective predicates $d_{\lfloor i+1 \rfloor, j}$, with j cycling forwards through the numbers $0, \ldots, 5$, and i constant. This process continues until we encounter an element satisfying dg (if we ever do). We call any sequence $a_{t+1}, \ldots, a_v$ in which $\text{bt}[a_{t+1}]$, $v \geq t+1$, and successive elements are generated by rule (47), an *upward sequence*; if, in addition, $\text{dg}[a_v]$, we call it an *upward column*. Note that the only element of this sequence satisfying bt is a_{t+1}. We refer to the constant index i of the local address predicates $d_{i,j}$ as the *longitude* of the upward sequence. Necessarily, $i \in \{0, 2, 4\}$. Furthermore, we observe that $a_{t+1}, a_{t+2}, \ldots$ defines primary and secondary colour sequences

black, black, green, green, red, red, ...
green, red, red, black, black, green, ...

in the by now familiar way, repeating (as long as the sequence continues) with a period of six. In the structure $\mathfrak{A}_0$ (Fig. 5), the elements of the even-numbered column from the bottom to the diagonal (inclusive) form upward columns; the primary colours are to the left, and the secondaries to the right. Observe that the sequences of colours obtained for an upward column is thus the reverse of that for a downward column. (Look, for example at columns 7 and 8 in Fig. 5.) Summarizing these observations:

Lemma 30. *Let $a_{s+1}, \ldots, a_t$ be a downward column of longitude $i \in \{1, 3, 5\}$. Then $d_{i,\lfloor h \rfloor}[a_{t-h}]$ for all h $(0 \leq h \leq t-s-1)$ and $\neg\text{dg}[a_{t-h}]$ for all h $(0 \leq h < t-s-1)$. Moreover, $T_{\lfloor (h+3)/2 \rfloor_3}[a_{t-h-1}, a_{t-h}]$ and $T_{\lfloor h/2 \rfloor_3}[a_{t-h-1}, a_{t-h}]$ for all h $(0 \leq h < t-s-1)$. Suppose $a_{t+1}, \ldots, a_v$ is an upward sequence following a_t. Then $d_{\lfloor i+1 \rfloor, \lfloor h \rfloor}[a_{t+1+h}]$ for all h $(0 \leq h \leq v-(t+1))$ and $\neg\text{bt}[a_{t+1+h}]$ for all h $(1 \leq h \leq v-(t+1))$. Moreover, $T_{\lfloor h/2 \rfloor_3}[a_{t+1+h}, a_{t+1+(h+1)}]$ and $T_{\lfloor (h+3)/2 \rfloor_3}[a_{t+1+h}, a_{t+1+(h+1)}]$ for all h $(0 \leq h < v-(t+1))$.*

As yet, we have no guarantee that the upward sequence mentioned in Lemma 30 cannot be extended indefinitely. We now add to φ_{grid} two final groups of conjuncts which provide such a guarantee. We refer to the first of these as the *transfer* formulas:

$$\bigwedge_{i=1,3,5} \bigwedge_{j=0,2,4} \forall(d_{i,j} \rightarrow \forall(d_{\lfloor i+1 \rfloor, j} \wedge T_{\lfloor j/2-1 \rfloor_3} \rightarrow T_{\lfloor j/2 \rfloor_3})) \tag{55}$$

$$\bigwedge_{i=0,2,4} \bigwedge_{j=1,3,5} \forall(d_{i,j} \rightarrow \forall(d_{\lfloor i+1 \rfloor, j} \wedge T_{\lfloor j/2-1 \rfloor_3} \rightarrow T_{\lfloor j/2+1 \rfloor_3})) \tag{56}$$

$$\bigwedge_{i=0,2,4} \forall(d_{i,i} \wedge \text{dg} \rightarrow \forall(c_{i,\lfloor i+1 \rfloor} \wedge T_{\lfloor i/2 \rfloor_3} \rightarrow T_{\lfloor i/2+1 \rfloor_3})) \tag{57}$$

$$\bigwedge_{i=1,3,5} \forall(d_{i,i} \wedge \mathrm{dg} \to \forall(c_{i,\lfloor i+1 \rfloor} \wedge T_{\lfloor i/2 \rfloor_3} \to T_{\lfloor i/2-1 \rfloor_3})) \tag{58}$$

$$\bigwedge_{i=0,2,4} \bigwedge_{j=0,2,4} \forall(c_{i,j} \to \forall(c_{i,\lfloor j+1 \rfloor} \wedge T_{\lfloor i/2 \rfloor_3} \to T_{\lfloor i/2+1 \rfloor_3})) \tag{59}$$

$$\bigwedge_{i=1,3,5} \bigwedge_{j=1,3,5} \forall(c_{i,j} \to \forall(c_{i,\lfloor j+1 \rfloor} \wedge T_{\lfloor i/2 \rfloor_3} \wedge \to T_{\lfloor i/2-1 \rfloor_3})). \tag{60}$$

We illustrate how these formulas work with reference to the model $\mathfrak{A}_0$ of Fig. 5. Consider the elements $(3,2)$ and $(4,2)$. Since the former element satisfies $d_{3,2}$ and the latter $d_{4,2}$, the transfer formula (55) tells us that if $(3,2)$ is related to $(4,2)$ by T_0 (i.e. black, shown by the dashed arrow in Fig. 5), then $(3,2)$ is related to $(4,2)$ by T_1 (i.e. green, shown by the dotted arrow). And similarly for all the other cases in Fig. 5: they allow us in all cases to infer a dotted arrow (horizontal or vertical) from the adjacent dashed arrow.

For the last group of conjuncts of φ_{grid}, we write $T_\diamond$ to abbreviate $T_0 \vee T_1 \vee T_2$; thus $T_\diamond[a,b]$ means that a is related to b by at least one of the colours. The *control formulas* are the following conjuncts:

$$\bigwedge_{i=0}^{5} \forall(d_{i,i} \wedge \pm\mathrm{dg} \to \forall(T_\diamond \wedge d_{\lfloor i+1 \rfloor, \lfloor i+1 \rfloor} \to \pm\mathrm{dg})) \tag{61}$$

$$\bigwedge_{i=0}^{5} \forall(d_{i,0} \wedge \pm\mathrm{bt} \to \forall(T_\diamond \wedge d_{\lfloor i+1 \rfloor, 0} \to \pm\mathrm{bt})) \tag{62}$$

$$\bigwedge_{j=0}^{5} \forall(c_{\lfloor j-1 \rfloor, j} \wedge \pm\mathrm{dg}^+ \to \forall(T_\diamond \wedge c_{j,\lfloor j+1 \rfloor} \to \pm\mathrm{dg}^+)) \tag{63}$$

$$\bigwedge_{j=0}^{5} \forall(c_{0,j} \wedge \pm\mathrm{lf} \to \forall(T_\diamond \wedge c_{0,\lfloor j+1 \rfloor} \to \pm\mathrm{lf})). \tag{64}$$

Here, the occurences of $\pm$ are assumed to be resolved in the same way within a numbered display, thus each of (61)-(64) is actually a *pair* of formulas. These formulas are again best illustrated with reference to $\mathfrak{A}_0$. Fix a value of i ($0 \leq i \leq 5$) and consider elements satisfying $d_{i,i}$. Note that there are infinitely many of these, dotted throughout the model. Formula (61) states that if any such element a, is related by one of the colour predicates to an element b satisfying $d_{\lfloor i+1 \rfloor, \lfloor i+1 \rfloor}$, then a lies on the diagonal if and only if b does. This is true in $\mathfrak{A}_0$, since elements are related by colours only to nearby elements. Similarly, (62) states that if any element a satisfying $d_{i,0}$ is related by one of the colour predicates to an element b satisfying $d_{\lfloor i+1 \rfloor, 0}$ then a satisfies the predicate bt if and only if b does. The formulas (63) and (64) function similarly.

We now show how the transfer and control formulas work together to link elements of a downward column $a_s, \ldots, a_t$ to those of any following upward sequence $a_{t+1}, \ldots, a_v$. It transpires that the latter sequence eventually stops at an element a_v satisfying dg (so that we have an upward column), and that this upward column is longer than the preceding downward column by exactly 1;

moreover, the elements of these columns are connected 'horizontally' by colours as illustrated in Fig. 5. More precisely:

Lemma 31. *Let $a_{s+1}, \ldots, a_t$ be a downward column of longitude $i \in \{1,3,5\}$ and length $\ell = t - s$ in some model of φ_{grid}. Then the generation rules yield an upward column $a_{t+1}, \ldots, a_v$ of longitude $\lfloor i+1 \rfloor$ and length exactly $\ell + 1$. Moreover, for all h ($0 \leq h \leq \ell - 1$), $T_{\lfloor h/2 \rfloor_3}[a_{t-h}, a_{t+1+h}]$.*

Proof. To show that there is an upward sequence $a_{t+1}, \ldots, a_v$ of length at least $\ell + 1$, it suffices to show that, for all h ($0 \leq h \leq \ell - 1$), $\neg\mathrm{dg}[a_{t+1+h}]$, since that is the condition for rule (47)—which generates the upward sequence—to keep on applying. We proceed by induction, showing the stronger claim, namely, that for all h ($0 \leq h \leq \ell - 1$), both $\neg\mathrm{dg}[a_{t+1+h}]$ and $T_{\lfloor h/2 \rfloor_3}[a_{t-h}, a_{t+1+h}]$. The case $h = 0$ is immediate, since rule (49) yields an element a_{t+1} such that $T_0[a_t, a_{t+1}]$ and $\neg\mathrm{dg}[a_{t+1}]$. Suppose, then, the claim holds for some non-negative value $h < \ell - 1$; we show that it holds for $h + 1$. By inductive hypothesis, then, $\neg\mathrm{dg}[a_{t+1+h}]$, whence rule (47) applies to a_{t+1+h}, yielding an element $a_{t+1+(h+1)}$. Also by inductive hypothesis, $T_{\lfloor h/2 \rfloor_3}[a_{t-h}, a_{t+1+h}]$, and by Lemma 30, $T_{\lfloor h/2 \rfloor_3}[a_{t-(h+1)}, a_{t-h}]$ and $T_{\lfloor h/2 \rfloor_3}[a_{t+1+h}, a_{t+1+(h+1)}]$. Hence, by transitivity, $T_{\lfloor h/2 \rfloor_3}[a_{t-(h+1)}, a_{t+1+(h+1)}]$. If h is even, then $(h+1)/2 = h/2$ and so we have $T_{\lfloor (h+1)/2 \rfloor_3}[a_{t-(h+1)}, a_{t+1+(h+1)}]$ as required. On the other hand, if h is odd, then $(h+1)/2 - 1 = h/2$, so that $T_{\lfloor (h+1)/2-1 \rfloor_3}[a_{t-(h+1)}, a_{t+1+(h+1)}]$. But by Lemma 30, we have $d_{i,\lfloor h+1 \rfloor}[a_{t-(h+1)}]$ and $d_{\lfloor i+1 \rfloor, \lfloor h+1 \rfloor}[a_{t+1+(h+1)}]$, and so, by the transfer formula (55) (setting $j = \lfloor h+1 \rfloor$), we have $T_{\lfloor (h+1)/2 \rfloor_3}[a_{t-(h+1)}, a_{t+1+(h+1)}]$, which completes the inductive step.

Finally, putting $h = t - s - 1 = \ell - 1$, we have $\neg\mathrm{dg}[a_{t+1+(\ell-1)}]$, whence rule (47) applies once again to yield an upward sequence $a_{t+1}, \ldots, a_{t+1+\ell}$ of length $\ell + 1$, with $T_{\lfloor (h-(\ell-1))/2 \rfloor_3}[a_{t+1+(\ell-1)}, a_{t+1+\ell}]$. To prove the lemma, it suffices to show that $\mathrm{dg}[a_{t+1+\ell}]$, since an element satisfying dg by definition terminates an upward column. From the claim of the previous paragraph, with $h = \ell - 1$, we have $T_{\lfloor (h-(\ell-1))/2 \rfloor_3}[a_{t-(\ell-1)}, a_{t+1+(\ell-1)}]$. By transitivity of this relation, it follows that $T_{\lfloor (h-(\ell-1))/2 \rfloor_3}[a_{t-(\ell-1)}, a_{t+1+\ell}]$, and since $a_{t-(\ell-1)} = a_{s+1}$ is the first element of our downward column, we have $\mathrm{dg}[a_{t-(\ell-1)}]$. Now apply (61) to infer that $\mathrm{dg}[a_{t+1+\ell}]$. Thus, setting $v = t + \ell + 2$, we obtain an upward column $a_{t+1}, \ldots, a_v$ of length exactly $\ell + 1$, as required. □

Having obtained a downward column $a_{s+1}, \ldots, a_t$ with longitude $i \in \{1,3,5\}$ together with a subsequent upward column $a_{t+1}, \ldots, a_v$ with longitude $\lfloor i+1 \rfloor$ in this way, we see that generation rule (50) applies to a_v, yielding an element a_{v+1} satisfying $c_{i,\lfloor i+1 \rfloor}$ and dg^+, and to which a_v is related by $T_{\lfloor (i+1)/2 \rfloor_3}$. We observe in passing that, by Lemma 30, $T_{\lfloor (i+1)/2 \rfloor_3}[a_{v-1}, a_v]$, and by Lemma 31, $T_{\lfloor (i+1)/2 \rfloor_3}[a_{s+1}, a_{v-1}]$, whence by transitivity, $T_{\lfloor (i+1)/2 \rfloor_3}[a_{s+1}, a_{v+1}]$, and hence, by the transfer formula (57),

$$T_{\lfloor (i+1)/2+1 \rfloor_3}[a_{s+1}, a_{v+1}]. \qquad (\star)$$

We return to this observation presently. One of the generation rules (51) or (53) now applies to a_{v+1}, depending on whether $\mathrm{lf}[a_{v+1}]$.

We call any sequence $a_{v+1}, \ldots, a_w$ in which $\mathrm{dg}^+[a_{v+1}]$, $w \geq v+1$, and successive elements are generated by rule (51), a *leftward sequence*. If, in addition, $\mathrm{lf}[a_w]$, then this process stops, and we call $a_{v+1}, \ldots, a_w$ a *leftward row*. Writing j for the value $\lfloor i+1 \rfloor$, we see that successive elements satisfy the local address predicates $c_{i',j}$ with j constant (the *latitude* of the sequence) and i' cycling through $5, \ldots, 0$ (starting at $\lfloor j-1 \rfloor$). In the case of a leftward row we have $\mathrm{lf}[a_w]$, and hence, by the second conjunct of (44), $c_{0,j}[a_w]$. We see therefore from rule (51) that the primary and secondary colour sequences must end in the pattern:

$$\ldots, \text{black, black, red, red, green, green}$$
$$\ldots, \text{red, green, green, black, black, red}$$

having repeated in a period of 6 as long as there have been elements in the sequence. In the structure $\mathfrak{A}_0$ (Fig. 5), the elements of the even-numbered rows from the super-diagonal to the left-hand edge (inclusive) form leftward rows; the primary colours are drawn above, and the secondaries below. Thus, the sequences of colours obtained for a leftward row is the reverse of that for a rightward row. (Look, for example at rows 7 and 8 in Fig. 5.) Summarizing these observations:

Lemma 32. *Let $a_{v+1}, \ldots, a_w$ be a leftward sequence of latitude j and length $\ell = w - v$. Then $c_{\lfloor j-1-h \rfloor, j}[a_{v+1+h}]$ for all h $(0 \leq h < \ell)$, and $\neg\mathrm{dg}^+[a_{v+1+h}]$ for all h $(0 < h < \ell)$. Moreover, $T_{\lfloor (j-h+1)/2 \rfloor_3}[a_{v+1+h}, a_{v+1+h+1}]$ and $T_{\lfloor (j-h+2)/2 \rfloor_3}[a_{v+1+h}, a_{v+1+h+1}]$ for all h $(0 \leq h < \ell)$.*

Now we establish the connection between a rightward row and its subsequent leftward sequence.

Lemma 33. *Let $a_u, \ldots, a_s$ be a rightward row of with latitude j and length $\ell = s - u + 1$ in some model of φ_{grid}. Let this be followed by the downward column $a_{s+1}, \ldots, a_t$ and upward column $a_{t+1}, \ldots, a_v$. Then the generation rules yield a leftward row $a_{v+1}, \ldots, a_w$ of latitude $\lfloor j+1 \rfloor$ and length exactly $\ell + 1$. Moreover, $T_{\lfloor (j-h)/2+2 \rfloor}[a_{s+1-h}, a_{v+1+h}]$ for all h $(0 \leq h \leq \ell - 1)$.*

The proof proceeds as for Lemma 31 by induction on h, but using Lemmas 28 and 32 instead of Lemma 30. The most significant difference is the base case ($h = 0$). Note that the latitude j of the rightward row $a_u, \ldots, a_s$ equals the longitude i of the subsequent downward column $a_{s+1}, \ldots, a_t$. Referring to the observation ($\star$), and remembering that $i = j$ is odd, we have $T_{\lfloor j/2+2 \rfloor_3}[a_{s+1}, a_{v+1}]$ as required. For the inductive case, we employ transitivity and the transfer formula (59) to move backwards through $a_u, \ldots, a_s$ and forwards through $a_{v+1}, \ldots, a_w$, analogously to Lemma 31.

At this point, the pattern repeats, but with axes transposed. By reasoning almost identical to that of Lemma 31, we have:

Lemma 34. *Let $a_{s+1}, \ldots, a_t$ be a leftward row with latitude j and length ℓ in some model of φ_{grid}. Then the generation rules yield a rightward row $a_{t+1}, \ldots, a_v$*

of latitude $\lfloor j+1 \rfloor$ *and length exactly* $\ell+1$*. Moreover, for all* h $(0 \leq h \leq \ell-1)$*,* $T_{\lfloor h/2+1 \rfloor_3}[a_{t-h}, a_{t+1+h}]$*.*

And by reasoning almost identical to that of Lemma 33, again with a slight change to the base case:

Lemma 35. *Let* $a_u, \ldots, a_s$ *be a upward column of with longitude* $i \in \{0, 2, 4\}$ *and length* $\ell = s - u + 1$ *in some model of* φ_{grid}*. Let this be followed by the leftward row* $a_{s+1}, \ldots, a_t$ *and rightward row* $a_{t+1}, \ldots, a_v$*. Then the generation rules yield a downward column* $a_{v+1}, \ldots, a_w$ *of longitude* $\lfloor i+1 \rfloor$ *and length exactly* $\ell+1$*. Moreover,* $T_{\lfloor (i-h+1)/2-1 \rfloor}[a_{s-h}, a_{v+2+h}]$ *for all* h $(1 \leq h \leq \ell-1)$*.*

Summarizing, suppose $\mathfrak{A} \models \varphi_{grid}$. The preceding lemmas yield a sequence of alternating rows and columns exactly matching the boustrophedon ς defined above, and illustrated in Fig. 5. Thus, we have a natural mapping $(X_t, Y_t) \mapsto a_t$ which is a homomorphism in the sense captured by the following lemma:

Lemma 36. *Let* $\mathfrak{A}_0$ *be the model of* φ_{grid} *defined above and illustrated in Fig. 5, and let* $\varsigma(t) = (X_t, Y_t)$ *be the boustrophedon defined in* (41)*. If* $\mathfrak{A}$ *is any model of* φ_{grid}*, then there exists a sequence of elements* $a_0, a_1, \ldots$ *from* A *such that, for all* $s, t \geq 0$*, all local address and control predicates* p*, and all colours* T_h*:* (i) *if* $\mathfrak{A}_0 \models p[\varsigma(s)]$ *then* $\mathfrak{A} \models p[a_s]$*; and* (ii) *if* $\mathfrak{A}_0 \models T_h[\varsigma(s), \varsigma(t)]$*, then* $\mathfrak{A} \models T_h[a_s, a_t]$*.*

Proof. Consider the characterization of ς in terms of alternating rows and columns in (42). Starting with Lemma 29, we obtain, after the element a_0, the first rightward row, a_1, of length 1 and latitude 1, and the first downward column, a_2, a_3 of length 2 and longitude 1. Lemmas 30 and 31 then give us the next upward column a_4, a_5, a_6 of length 3 and longitude 2, and Lemmas 32 and 33 a leftward row a_7, a_8 of length 2 and latitude 2. Lemma 34 then generates the next rightward row a_9, a_{10}, a_{11} of length 3 and latitude 3. And so the process repeats. As it does so, Lemmas 28, 30 and 32 ensure that the indices of the predicates $c_{i,j}$ and $d_{i,j}$ satisfied by the elements of the sequence $a_0, a_1, a_2 \ldots$ are given by the latitude and longitude of the corresponding points $\varsigma(0), \varsigma(1), \varsigma(2), \ldots$; moreover, extremal elements of either row or column satisfy the control predicates as required. The repeating patterns of colour predicates shown in Fig. 5 are likewise secured by the preceding lemmas. Specifically, the colours on the edges following the course of the boustrophedon are given by Lemmas 28, 30 and 32; the colours on the edges between neighbouring rows and columns are given by Lemmas 31, 33, 34 and 35. □

Lemma 36 yields an embedding ι of $\mathbb{N}^2$ into any model $\mathfrak{A}$ of φ_{grid} defined by setting $\iota(X_t, Y_t) = a_t$, where $\varsigma(t) = (X_t, Y_t)$. This justifies us in picturing the sequence $a_0, a_1, \ldots$ as laid out in Fig. 5. By inspection of Fig. 5, we see that ι has the following properties:

- If $X \geq Y$ then $T_\diamond[\iota(X,Y), \iota(X+1,Y)]$. Moreover, if X is even then $T_\diamond[\iota(X,Y), \iota(X,Y+1)]$, and if X is odd then $T_\diamond[\iota(X,Y+1), \iota(X,Y)]$.

- If $X < Y$ then $T_\diamond[\iota(X,Y), \iota(X,Y+1)]$. Moreover, if Y is even then $T_\diamond[\iota(X+1,Y), \iota(X,Y)]$, and if Y is odd then $T_\diamond[\iota(X,Y), \iota(X+1,Y)]$.

The above observation allows us to write formulas that properly assign tiles from a given tiling system $\boldsymbol{\mathcal{C}} = (\mathcal{C}, H, V)$ to elements of the model of φ_{grid}. We do this with a formula φ_{tile}, which again features several conjuncts. The first conjunct is straightforward. We require that each node encodes precisely one tile $C \in \mathcal{C}$ and the initial element satisfies the initial tiling condition by adding to φ_{tile} the formula:

$$\forall\big(\bigvee_{C\in\mathcal{C}} C \wedge \bigwedge_{C\neq D} (\neg C \vee \neg D) \wedge (\mathrm{lf} \wedge \mathrm{dg} \rightarrow C_0)\big). \tag{65}$$

The next formulas ensure that adjacent tiles respect the constraints H and V. To ensure that the horizontal constraints are satisfied we add to φ_{tile} the following conjuncts for every $C \in \mathcal{C}$:

$$\bigwedge_{0\leq i,j\leq 5} \forall(C \wedge d_{ij} \rightarrow \forall(T_\diamond \wedge d_{\lfloor i+1\rfloor,j} \rightarrow \bigvee_{C':(C,C')\in H} C')) \tag{66}$$

$$\bigwedge_{0\leq i\leq 5}\bigwedge_{j=1,3,5} \forall(C \wedge c_{ij} \rightarrow \forall(T_\diamond \wedge (c_{\lfloor i+1\rfloor,j} \vee d_{\lfloor i+1\rfloor,j}) \rightarrow \bigvee_{C':(C,C')\in H} C')) \tag{67}$$

$$\bigwedge_{0\leq i\leq 5}\bigwedge_{j=0,2,4} \forall(C \wedge (c_{i,j} \vee d_{i,j}) \rightarrow \forall(T_\diamond \wedge c_{\lfloor i-1\rfloor,j} \rightarrow \bigvee_{C':(C',C)\in H} C')). \tag{68}$$

A similar group of conjuncts is added to handle the vertical constraints. Again, we add to φ_{tile} the following conjuncts for every $C \in \mathcal{C}$:

$$\bigwedge_{0\leq i,j\leq 5} \forall(C \wedge (c_{i,j} \vee d_{i,j}) \rightarrow \forall(T_\diamond \wedge c_{i,\lfloor j+1\rfloor} \rightarrow \bigvee_{C':(C,C')\in V} C')) \tag{69}$$

$$\bigwedge_{i=0,2,4}\bigwedge_{0\leq j\leq 5} \forall(C \wedge d_{i,j} \rightarrow \forall(T_\diamond \wedge d_{i,\lfloor j+1\rfloor} \rightarrow \bigvee_{C':(C,C')\in V} C')) \tag{70}$$

$$\bigwedge_{i=1,3,5}\bigwedge_{0\leq j\leq 5} \forall(C \wedge d_{i,j} \rightarrow \forall(T_\diamond \wedge d_{i,\lfloor j-1\rfloor} \rightarrow \bigvee_{C':(C',C)\in V} C')). \tag{71}$$

This completes the definition of the formula φ_{tile}. Finally, let $\eta_{\mathcal{C}}$ be the conjunction of φ_{grid} and φ_{tile}. We show that

Lemma 37. *$\eta_{\mathcal{C}}$ is satisfiable iff $\boldsymbol{\mathcal{C}}$ tiles $\mathbb{N}^2$ with initial condition C_0.*

Proof. If $\boldsymbol{\mathcal{C}}$ tiles $\mathbb{N}^2$ with the initial condition then to show that $\eta_{\mathcal{C}}$ is satisfiable we expand our intended model $\mathfrak{G}$ for φ_{grid} assigning to every element of the grid a unique $C \in \mathcal{C}$ given by the tiling. It is clear that (65) holds in the model. Moreover, the transitive paths of any of the transitive relations connect at most three adjacent columns and at most three adjacent rows. So, the distribution of the local address predicates ensures that also the conjuncts (66)–(71) are satisfied.

Now, let $\mathfrak{A} \models \eta_{\mathcal{C}}$. Since $\mathfrak{A} \models \varphi_{grid}$ consider the embedding ι of the standard $\mathbb{N}^2$ grid into $\mathfrak{A}$ defined above. We define a tiling of the $\mathbb{N}^2$ grid assigning to every node $(X, Y) \in \mathbb{N}^2$ the unique tile C such that $\mathfrak{A} \models C(\iota(X, Y))$. Formula (65) ensures that this is well defined and satisfies the initial condition. Formulas (66)-(68) ensure that the horizontal constraints are satisfied, taking care to ensure the local address predicates change in the correct way for rows above the diagonal (and in particular near the diagonal); and formulas (69)-(71) ensure that the vertical constraints are satisfied. □

Hence, we have the following

Theorem 38. *The satisfiability problem for $\mathcal{FL}^2$3T, the two-variable fluted fragment with three transitive relations, is undecidable.*

We now turn towards the finite satisfiability problem. First, anticipating Theorem 41, we remark that the formula φ_{grid} is an axiom of infinity. Hence, there is no prospect of using it simultaneously to obtain undecidability of the satisfiability and finite satisfiability problems via Proposition 24. To prove undecidability of finite satisfiability for $\mathcal{FL}^2$3T, we instead reduce from the finite tiling problem with initial and final conditions, invoking Proposition 23. We proceed as follows. First, we modify the formula φ_{grid} so that it no longer constructs an infinite sequence of witnesses but the process is allowed to stop when the boustrophedon meets an element on the bottom row. In other words, the coordinates assigned to the sequence of witnesses correspond to a square domain $\mathbb{N}^2_{2n,2n}$, for some $n \geq 1$.

Denote the modified formula φ_{sgrid}. It contains some conjuncts taken directly from φ_{grid}, some that are modified versions of conjuncts in φ_{grid}, and some that are new. First of all, we employ an additional control predicate rt intended to mark the rightmost column of the square domain. This is secured by adding the following new conjunct to φ_{sgrid} (complementing the formula (44)):

$$\forall(\mathrm{rt} \to \bigvee_{i=0,2,4} \bigvee_{j=0}^{5} d_{i,j}) \tag{72}$$

and the following new control formula:

$$\bigwedge_{i=0}^{5} \bigwedge_{j=0}^{5} \forall(d_{i,j} \wedge \pm\mathrm{rt} \to \forall(T_\diamond \wedge d_{i,\lfloor j-1 \rfloor} \to \pm\mathrm{rt})). \tag{73}$$

In φ_{sgrid} we modify the formula (45) by ensuring that the initial element does not satisfy rt as follows:

$$\exists(d_{0,0} \wedge \mathrm{dg} \wedge \mathrm{bt} \wedge \neg\mathrm{rt}). \tag{74}$$

Finally, we modify the generation rule (49); now we require a new witness only for bottom elements that are not on the rightmost column, writing:

$$\bigwedge_{i=1,3,5} \forall(d_{i,0} \wedge \mathrm{bt} \wedge \neg\mathrm{dg} \wedge \neg\mathrm{rt} \to \exists(d_{\lfloor i+1 \rfloor,0} \wedge \mathrm{bt} \wedge \neg\mathrm{dg} \wedge T_0)). \tag{75}$$

The remaining conjuncts of φ_{grid} constitute conjuncts of φ_{sgrid} without modification.

Observe that φ_{sgrid} has finite models: if a witness a_t of the conjunct (54) happens to satisfy rt then the following witnesses $a_{t'}$ with $t' > t$, corresponding to a downward column in the model, also satisfy rt due to the new control formula (73). As argued earlier, the sequence of witnesses eventually reaches an element $a_{t''}$ satisfying bt, and this is where no new witnesses are required due to the modified generation rule (75). Moreover in every finite model of φ_{sgrid} one can embed a square grid $\mathbb{N}^2_{2n,2n}$ similarly as we did before embedding the $\mathbb{N}^2$ grid in models of φ_{grid}.

In order to complete the reduction of the finite tiling problem with initial and final conditions, we need one more conjunct ensuring the final condition:

$$\forall(\mathrm{dg} \wedge \mathrm{rt} \to C_1). \tag{76}$$

It should be now straightforward to check that the conjunction of (76) with $\varphi_{sgrid} \wedge \varphi_{tile}$ is *finitely* satisfiable iff $\mathcal{C}$ tiles $\mathbb{N}^2_{2n,2n}$ with initial condition C_0 and final condition C_1, for some $n \geq 1$. Hence, by Proposition 23:

Theorem 39. *The finite satisfiability problem for* $\mathcal{FL}^2$3T*, the two-variable fluted fragment with three transitive relations, is undecidable.*

We observe additionally that all formulas used in the proofs of Theorems 38 and 39 are either guarded or can be rewritten as guarded. Furthermore, in the proof it would suffice to assume that T_0, T_1 and T_2 are interpreted as equivalence relations. Hence, we can strengthen the above theorem as follows.

Corollary 40. *The* (*finite*) *satisfiability problem for the intersection of the fluted fragment with the two-variable guarded fragment is undecidable in the presence of three transitive relations* (*or three equivalence relations*).

Since the satisfiability problem for $\mathcal{FL}^2$3T is undecidable, it follows that $\mathcal{FL}^2$3T lacks the finite model property; however, as yet, we do not have an actual axiom of infinity. Recalling the model $\mathfrak{A}_0 \models \varphi_{grid}$ and the boustrophedon $\varsigma : \mathbb{N} \to \mathbb{N}^2$, suppose $\mathfrak{A} \models \varphi_{grid}$ and consider the sequence $a_0, a_1, \ldots$ constructed by the generation rules. Lemma 36 states that if $\mathfrak{A}_0 \models p[\varsigma(t)]$, where p is in the signature of φ_{grid}, then $\mathfrak{A} \models p[a_t]$. Similarly, if $\mathfrak{A}_0 \models p[\varsigma(s), \varsigma(t)]$, then $\mathfrak{A} \models p[a_s, a_t]$. This result invites us, informally, to picture the sequence $a_0, a_1, \ldots$ as being laid out as shown in Fig. 5. It is important to realize, however, that we have not shown that these elements are *distinct*. We do so now.

Theorem 41. *The formula* φ_{grid} *is an axiom of infinity.*

Proof. Suppose $\mathfrak{A} \models \varphi_{grid}$. Let ι be the embedding of $\mathbb{N}^2$ into $\mathfrak{A}$ defined above, i.e. $\iota(X_t, Y_t) = a_t$, where $\varsigma(t) = (X_t, Y_t)$ is the tth point on the boustrophedon (41). We show that ι is injective, which proves the theorem.

As a preliminary, consider the rectangles into which the upper-right quadrant of the plane is divided by the black, green and red lines in Fig. 5. We refer to these rectangles as *bricks*. Each brick consists of four or six points in the plane, with the former kind confined to the left-hand and bottom edges; moreover, the bricks form a natural sequence following the boustrophedon. Since every point $\varsigma(t) = (X_t, Y_t)$ is associated with an element a_t in some model of φ_{grid}, we can think of bricks as the set of associated elements in a given model. And by condition (ii) of Lemma 36, we see that for any brick B, there exists k $(0 \leq k < 3)$ such that, for all elements $a_s, a_t \in B$ with $s < t$, we have $T_k[a_s, a_t]$. In other words, each brick has a *colour*, and, furthermore, an *orientation* induced by the ordering of points on the boustrophedon. We call the bricks below the diagonal having their left-hand margins in even columns *downward-pointing*, while those below the diagonal having their left-hand margins in odd columns are *upward-pointing*; similarly for *leftward-* and *rightward-pointing* bricks above the diagonal, depicted by yellow arrows in Figure 5. Of course, while the elements of B lie in order as the periphery of B is traversed, they are not in general consecutive in the sequence $\{a_t\}$. From Fig. 5, the following are evident.

(E1) Every element satisfying $d_{i,j}$ except for a_0 lies on at least one upward-pointing brick and at least one downward pointing brick.

(E2) Any two elements satisfying the same local address and control predicates lie on bricks with the same set of colours/orientations. (Example: in Fig, 5, (5,2) and (11,8) both lie on a red downward-pointing brick, a black upward-pointing brick and a green upward-pointing brick.)

(E3) If B is a 6-element upward-pointing brick and its first element is a non-diagonal element, then that element has local address $d_{i,j}$ (i odd, j even), while the last element has local address $d_{\lfloor i+1 \rfloor, j}$, and the colour of B is $T_{\lfloor j/2-1 \rfloor_3}$; similarly for downward-pointing bricks.

(E4) The first element of each brick B is related to all the others by the colour of B, and all the elements but the last are related to the last element by the colour of B.

We are now ready to embark on the proof. Assume for contradiction that $a_s = a_t$ with $t < s$. We consider the case where $a_s = a_t$ satisfies some $d_{i,j}$; the case for elements satisfying some $c_{i,j}$ is handled similarly.

Assume first that $Y_s = Y_t$. Since $t < s$, and, a_s has the same local address as a_t (since they are identical), we must have $X_t < X_s$ and therefore, by condition (i) of Lemma 36, $X_t < X_s - 5$. As a preliminary, we claim that, if a_s lies on a brick B and a_t on a brick D, then no element of either B or D can satisfy dg. For if B has an element $a_{s'}$ such that $\text{dg}[a_{s'}]$, then $X_t < X_s - 5 \leq X_{s'} - 4 = Y_{s'} - 4 \leq Y_s - 2 = Y_t - 2 < Y_t$ contradicting condition (i) of Lemma 36 and the fact that $a_s = a_t$ satisfies some predicate $d_{i,j}$. In particular, $a_s = a_t$ itself does not satisfy dg. If, on the other hand, D has an element $a_{t'}$ satisfying dg, then, again by Lemma 36, there is such a t' satisfying $t' > t$. Letting $s' = s + (t' - t)$, we see that the sequences $a_s, \ldots a_{s'}$ and $a_t, \ldots a_{t'}$ are the same (and thus have the same local addresses), whence $X_s, \ldots X_{s'}$ and

$X_t, \ldots X_{t'}$ move in the same way (i.e. $X_{s+h+1} - X_{s+h} = X_{t+h+1} - X_{t+h}$, for all h with $0 \leq h \leq s' - s$) so that $X_t < X_s - 5$ implies $X_{t'} < X_{s'} - 5$. Thus, recalling that $\mathrm{dg}[a_{s'}]$ implies $X_{t'} = Y_{t'}$, and that $Y_s = Y_t$ by assumption, we have $X_{s'} > X_{t'} + 5 = Y_{t'} + 5 \geq Y_t + 3 = Y_s + 3 \geq Y_{s'} + 1 > Y_{s'}$, contradicting the supposition that $a_{t'} = a_{s'}$ satisfies dg. This proves the claim that neither a_s nor a_t lie on any brick containing a diagonal element.

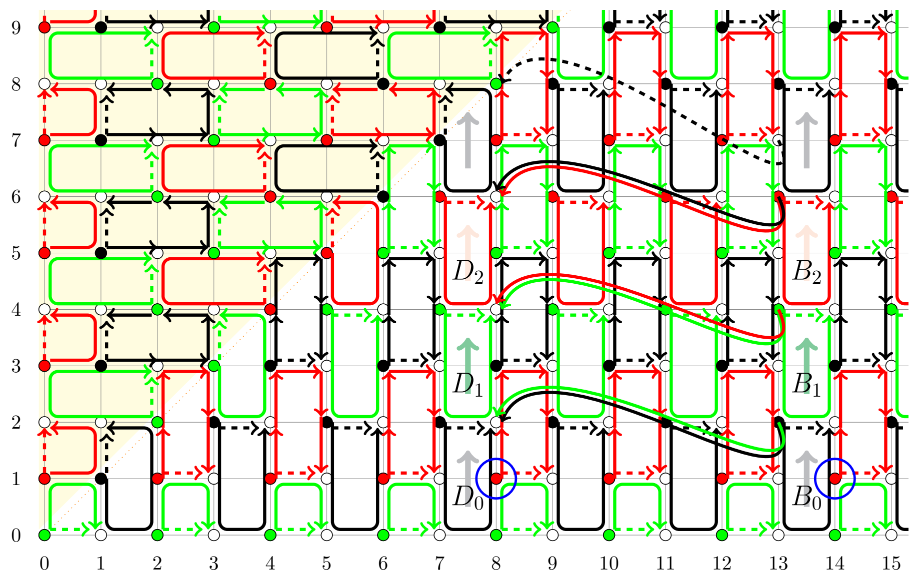

Figure 6: Proof of Lemma 41, illustrating the example in which we suppose $\varsigma(s) = (14, 1)$ and $\varsigma(t) = (8, 1)$ (circled) are mapped by ι to the same element $a_s = a_t$. We have $\varsigma(s_0) = (13, 2)$, $\varsigma(t_0) = (8, 2)$. Now $T_0[a_{s_0}, a_{t_0}]$ implies $T_1[a_{s_0}, a_{t_0}]$ implies $T_1[a_{s_1}, a_{t_1}]$ implies $T_2[a_{s_1}, a_{t_1}]$ implies $T_2[a_{s_2}, a_{t_2}]$ implies $T_0[a_{s_2}, a_{t_2}]$. The black edge from $(13, 7)$ to $(8, 8)$ yields the desired contradiction.

This claim having been established, we proceed to derive the promised contradiction. To make the proof easier, we suggest the reader follows with reference to the example $\varsigma(s) = (14, 1)$ and $\varsigma(t) = (8, 1)$ (see Fig. 6). By (E1) and (E2), let B_0 and D_0 be the upward-pointing bricks containing, respectively, a_s and a_t, and having the same colour, say T_{k_0}. Let a_{s_0} be the first element on the brick B_0, and a_{t_0}—the last element on the brick D_0, in our example, $\varsigma(s_0) = (13, 2)$ and $\varsigma(t_0) = (8, 2)$. By (E3), if s_0 satisfies—say—$c_{i',j'}$, then t_0 satisfies $c_{i',\lfloor j'+1 \rfloor}$; and by (E4), $T_{k_0}[a_{s_0}, a_{t_0}]$, i.e. a_{s_0} is connected to a_{t_0} by an edge of some colour, T_{k_0}—in our example, black. The transfer formula (55) implies that $T_{\lfloor k_0+1 \rfloor_3}[a_{s_0}, a_{t_0}]$, in our example green. Now, write $k_1 = \lfloor k_0 + 1 \rfloor_3$, and let B_1 and D_1 be the upward-pointing bricks of colour T_{k_1} (in our example, green), containing, respectively, a_{s_0} and a_{t_0}. Let a_{s_1} be the first element on the brick B_1, and a_{t_1}—the last element on the brick D_1, i.e. $\varsigma(s_1) = (13, 4)$ and

$\varsigma(t_1) = (8, 4)$. By the same reasoning as for the previous pair of bricks, a_{s_1} is connected to a_{t_1} by a T_{k_1}-edge (in our example, green); hence by (55), a_{s_1} is connected to a_{t_1} also by an edge of colour T_{k_2}, where $k_2 = \lfloor k_1 + 1 \rfloor_3$ (in our example, red).

Now the reasoning simply repeats, generating sequences of bricks $B_0, B_1, \ldots$ and $D_0, D_1, \ldots$ until we reach $\ell \geq 0$ where either the brick above B_ℓ or the brick above D_ℓ contains an element satisfying dg. In particular, in our example, we consider B_2 and D_2—the red upward-pointing bricks containing, respectively, a_{s_1} and a_{t_1}, and we let a_{s_2} be the first element on the brick B_2, and a_{t_2}—the last element on the brick D_2. So, $\varsigma(s_2) = (13, 6)$ and $\varsigma(s_2) = (8, 6)$. Again, by (E4), a_{s_2} is connected to a_{t_2} by a red edge, hence by (55), also by a black one. Now the black brick above D_2 contains diagonal elements (i.e. $\ell = 2$); in particular, $\text{dg}[a_{t_2+2}]$, where $\varsigma(t_2 + 2) = (8, 8)$.

Recall that we are assuming that $Y_s = Y_t$. By (E2), we have $Y_{s_0} = Y_{t_0}$, and, since we have been following the two columns of the boustrophedon upward, $Y_{s_\ell} = Y_{t_\ell}$. Moreover, since $t < s$, we have $X_{t_\ell} < X_{s_\ell}$, and indeed, $X_{t_\ell} < X_{s_\ell} - 5$. So, indeed, the process stops when the brick above D_ℓ contains an element satisfying dg and, then, we necessarily have $\text{dg}[a_{t_\ell+2}]$. We have already established that $d_{i_0,\lfloor j_0+2\ell \rfloor}[a_{s_\ell}]$, $d_{\lfloor i_0+1 \rfloor,\lfloor j_0+2\ell \rfloor}[a_{t_\ell}]$ and $T_{k_0+l+1}[a_{s_\ell}, a_{t_\ell}]$. By condition (ii) of Lemma 36, we see that $T_{k_0+l+1}[a_{s_\ell-1}, a_{s_\ell}]$, and, indeed, $T_{k_0+l+1}[a_{t_\ell}, a_{t_\ell+2}]$. By transitivity, therefore $T_{k_0+l+1}[a_{s_\ell-1}, a_{t_\ell+2}]$. On the other hand, since $X_{s_\ell-1} > Y_{s_\ell-1}$, condition (i) of Lemma 36 implies that $a_{s_\ell-1}$ does not satisfy dg . But then we have $d_{i_0,\lfloor j_0+2\ell+1 \rfloor}[a_{s_\ell-1}]$, $d_{\lfloor i_0+1 \rfloor,\lfloor j_0+2\ell+2 \rfloor}[a_{t_\ell+2}]$ and $T_{k+l+1}[a_{s_\ell-1}, a_{t_\ell+2}]$, which, in the presence of (44), violates the control formula (61). In our example, $\varsigma(s_2 - 1) = (13, 7)$ and we have $d_{1,1}[a_{s_2-1}]$, $d_{2,2}[a_{t_2+2}]$, $T_0[a_{s_2-1}, a_{t_2+2}]$, $\neg\text{dg}[a_{s_2-1}]$ and $\text{dg}[a_{t_2+2}]$. Thus, if $a_s = a_t$ with $s \neq t$ but $Y_s = Y_t$, we obtain two distinct sequences of bricks, marching upwards until one (and only one) reaches the diagonal; this contradicts the control formulas.

This deals with the case $Y_s = Y_t$. If $Y_s \neq Y_t$, then we let B_0 be any *downward*-pointing brick containing a_s, T_k be the colour of B_0, and D_0 the downward-pointing brick containing a_t and having the same colour as D_0. Again, we let s_0 be the first element on B_0 and t_0 be the last element on D_0, following the preceding bricks $B_1, B_2, \ldots$ and $D_1, D_2, \ldots$. This time, however, we will be marching *down* the columns until we reach B_ℓ and D_ℓ such that one of the elements $a_{s_\ell-1}$ or $a_{t_\ell+1}$ satisfies bt. Now, the assumption that $Y_s \neq Y_t$ implies that at most one of $a_{s_\ell-1}$ and $a_{t_\ell+1}$ satisfies bt, which yields a violation of the control formula (62) using parallel reasoning to the upward case. The process is depicted in Figure 7 for the case where $\varsigma(s) = (11, 10)$ and $\varsigma(t) = (5, 4)$. □

5. Conclusions

In this paper, we considered the logics $\mathcal{FL}^m k\mathrm{T}$ and $\mathcal{FL}^m_= k\mathrm{T}$, the m-variable fluted fragment in the presence of (equality and) k transitive relations. We showed that the satisfiability problem for $\mathcal{FL}^m_= 1\mathrm{T}$ is in m-NExpTime, and indeed that the corresponding finite satisfiability problem is in $(m+1)$-NExpTime.

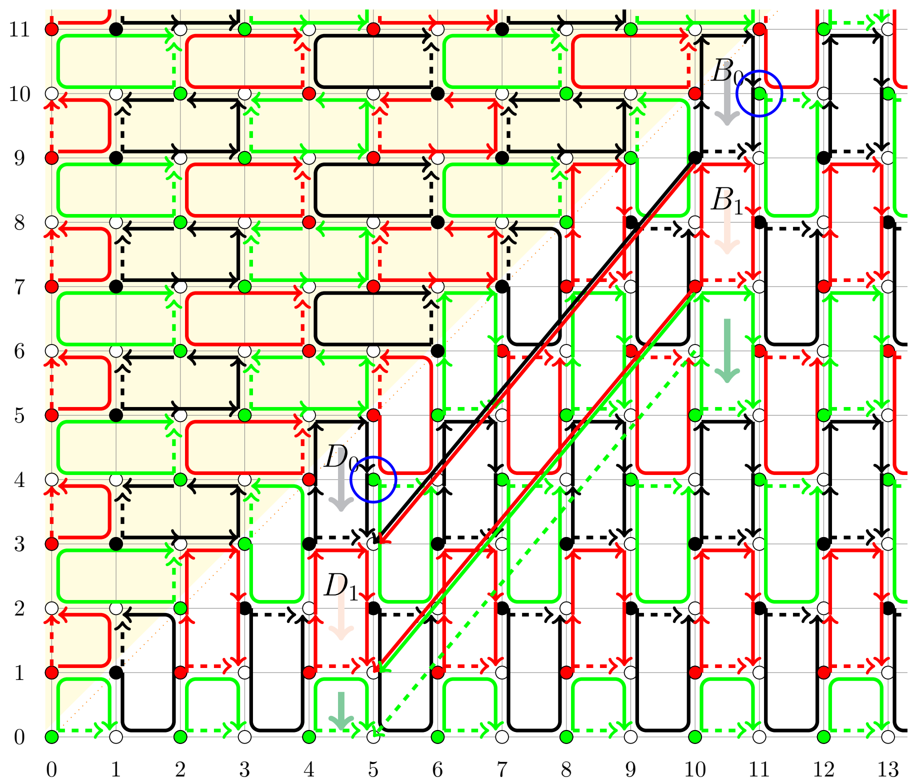

Figure 7: Proof of Lemma 41, illustrating the example in which we suppose $\varsigma(s) = (11, 10)$ and $\varsigma(t) = (5, 4)$ (circled) are mapped by ι to the same element $a_s = a_t$. We have $\varsigma(s_0) = (10, 9)$, $\varsigma(t_0) = (5, 3)$. Now $T_0[a_{s_0}, a_{t_0}]$ implies $T_2[a_{s_0}, a_{t_0}]$ implies $T_2[a_{s_1}, a_{t_1}]$ implies $T_1[a_{s_1}, a_{t_1}]$. Now $\varsigma(s_1 - 1) = (10, 6)$, $\varsigma(t_1 + 1) = (5, 0)$ and the green edge from $(10, 6)$ to $(5, 0)$, which follows by transitivity, yields the desired contradiction with (62).

(It seems probable that this latter bound, at least, can be improved.) Together with known lower bounds on the m-variable fluted fragment, it follows that the satisfaibility and finite satisfiability problems for $\mathcal{FL}_=1\mathrm{T}$, the fluted fragment with equality and a single transitive relation, are both TOWER-complete. (This extends the result of [18], which establishes the same complexity for the fluted fragment without equality or any transitive relations.) We also showed, however, that decidability is easily lost when additional transitive relations are added: even the two-variable fluted fragments $\mathcal{FL}^2_=2\mathrm{T}$ (two transitive relations plus equality) and $\mathcal{FL}^2_=3\mathrm{T}$ (three transitive relations, but without equality) have undecidable satisfiability and finite satisfiability problems.

It is open whether the satisfiability or finite satisfiability problems for $\mathcal{FL}2\mathrm{T}$ (*two* transitive relations, but without equality) are decidable. We point out that Lemma 20 in Section 3 could be generalized to normal-form formulas of

$\mathcal{FL}^{m+1}$2T (defined in the natural way). Hence, the (finite) satisfiability problem for $\mathcal{FL}^{m}$2T ($m > 2$) is decidable if and only if the corresponding problem $\mathcal{FL}^{2}$2T is. Unfortunately neither the method of Sec. 3 (to show decidability) nor that of Sec. 4 (to show undecidability) appears to apply here. The barrier in the former case is that pairs of elements can be related by both of the transitive relations, T_1 and T_2, via *distinct* T_1- and T_2-chains, so that simple certificates of the kind employed for $\mathcal{FL}^{2}_{=}1\mathrm{T}^{u}$ do not guarantee the existence of models. The barrier in the latter case is that the grid construction has to build models featuring transitive paths of *bounded* length, and this seems not to be achievable with just two transitive relations.

Acknowledgements. This work is supported by the Polish National Science Centre grant 2018/31/B/ST6/03662. The authors highly appreciate the reviewer's insightful comments that substantially helped to improve the original manuscript.

References

[1] F. Baader, D. Calvanese, D. L. McGuinness, D. Nardi, and P. F. Patel-Schneider, editors. *The Description Logic Handbook: Theory, Implementation, and Applications.* Cambridge University Press, 2003.

[2] E. Börger, E. Grädel, and Y. Gurevich. *The Classical Decision Problem.* Springer, 1997.

[3] E. Grädel, P. Kolaitis, and M. Vardi. On the decision problem for two-variable first-order logic. *Bulletin of Symbolic Logic*, 3(1):53–69, 1997.

[4] A. Herzig. A new decidable fragment of first order logic. In *Abstracts of the 3rd Logical Biennial Summer School and Conference in Honour of S. C.Kleene*, June 1990.

[5] Y. Kazakov. *Saturation-based decision procedures for extensions of the guarded fragment.* PhD thesis, Universität des Saarlandes, Saarbrücken, Germany, 2006.

[6] Y. Kazakov and I. Pratt-Hartmann. A note on the complexity of the satisfiability problem for graded modal logic. In *Logic in Computer Science*, pages 407–416. IEEE, 2009.

[7] E. Kieroński. Results on the guarded fragment with equivalence or transitive relations. In *Computer Science Logic*, volume 3634, pages 309–324. Springer Verlag, 2005.

[8] E. Kieroński. On the complexity of the two-variable guarded fragment with transitive guards. *Information and Computation*, 204:1663–1703, 2006.

[9] E. Kieroński, J. Michaliszyn, I. Pratt-Hartmann, and L. Tendera. Two-variable first-order logic with equivalence closure. *SIAM Journal on Computing*, 43(3):1012–1063, 2014.

[10] E. Kieroński and M. Otto. Small substructures and decidability issues for first-order logic with two variables. *Journal of Symbolic Logic*, 77:729–765, 2012.

[11] E. Kieroński and L. Tendera. On finite satisfiability of two-variable first-order logic with equivalence relations. In *Logic in Computer Science*. IEEE, 2009.

[12] E. Kieroński and L. Tendera. Finite satisfiability of the two-variable guarded fragment with transitive guards and related variants. *ACM Transactions on Computational Logic*, 19(2):8:1–8:34, 2018.

[13] R. Ladner. The computational complexity of provability in systems of modal propositional logic. *SIAM Journal on Computing*, 6:467–480, 1980.

[14] C. Le Duc and M. Lamolle. Decidability of description logics with transitive closure of roles in concept and role inclusion axioms. In *Proceedings of the 23rd International Workshop on Description Logics (DL 2010), Waterloo, Ontario, Canada, May 4-7, 2010*, 2010.

[15] A. Noah. Predicate-functors and the limits of decidability in logic. *Notre Dame Journal of Formal Logic*, 21(4):701–707, 1980.

[16] I. Pratt-Hartmann. The finite satisfiability problem for two-variable, first-order logic with one transitive relation is decidable. *Mathematical Logic Quarterly*, 64(3):218–248, 2018.

[17] I. Pratt-Hartmann, W. Szwast, and L. Tendera. Quine's fluted fragment is non-elementary. In *25th EACSL Annual Conference on Computer Science Logic, CSL*, volume 62 of *LIPIcs*, pages 39:1–39:21. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2016.

[18] I. Pratt-Hartmann, W. Szwast, and L. Tendera. The fluted fragment revisited. *Journal of Symbolic Logic*, 84(3):1020–1048, 2019.

[19] I. Pratt-Hartmann and L. Tendera. The fluted fragment with transitivity. In Peter Rossmanith, Pinar Heggernes, and Joost-Pieter Katoen, editors, *44th International Symposium on Mathematical Foundations of Computer Science, MFCS 2019, August 26-30, 2019, Aachen, Germany*, volume 138 of *LIPIcs*, pages 18:1–18:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.

[20] W. C. Purdy. Fluted formulas and the limits of decidability. *Journal of Symbolic Logic*, 61(2):608–620, 1996.

[21] W. C. Purdy. Complexity and nicety of fluted logic. *Studia Logica*, 71:177–198, 2002.

[22] W. V. Quine. On the limits of decision. In *Proceedings of the 14th International Congress of Philosophy*, volume III, pages 57–62. University of Vienna, 1969.

[23] W. V. Quine. The variable. In *The Ways of Paradox*, pages 272–282. Harvard University Press, revised and enlarged edition, 1976.

[24] S. Schmitz. Complexity hierarchies beyond Elementary. *ACM Transactions on Computation Theory*, 8(1:3):1–36, 2016.

[25] D. Scott. A decision method for validity of sentences in two variables. *Journal of Symbolic Logic*, 27:477, 1962.

[26] W. Szwast and L. Tendera. The guarded fragment with transitive guards. *Annals of Pure and Applied Logic*, 128:227–276, 2004.

[27] W. Szwast and L. Tendera. On the satisfiability problem for fragments of the two-variable logic with one transitive relation. *Journal of Logic and Computation*, 29(6):881–911, 2019.

[28] Hao Wang. Dominoes and the ∀∃∀-case of the decision problem. In *Proceedings of Symposium on the Mathematical Theory of Automata*, pages 23–55. Brooklyn Polytechnic Institute, 1962.