Ramsey’s theorem for pairs, collection, and proof size

Leszek Aleksander Kołodziejczyk  
Institute of Mathematics  
University of Warsaw  
Banacha 2  
02-097 Warszawa, Poland

Tin Lok Wong  
Department of Mathematics  
National University of Singapore  
10 Lower Kent Ridge Road  
Singapore 119076

Keita Yokoyama  
School of Information Science  
Japan Advanced Institute of Science and Technology  
1-1 Asahidai  
Nomi, Ishikawa 923-1292, Japan

15 May, 2020

Abstract

We prove that any proof of a $\forall \Sigma^0_2$ sentence in the theory WKL$_0 + \text{RT}^2_2$ can be translated into a proof in RCA$_0$ at the cost of a polynomial increase in size. In fact, the proof in RCA$_0$ can be found by a polynomial-time algorithm. On the other hand, RT$_2^2$ has non-elementary speedup over the weaker base theory RCA$_0^*$ for proofs of $\Sigma^1_1$ sentences.

We also show that for $n \geq 0$, proofs of $\Pi^n_{n+2}$ sentences in $\text{B} \Sigma^{n+1} + \text{exp}$ can be translated into proofs in $\text{I} \Sigma^1_n + \text{exp}$ at polynomial cost. Moreover, the $\Pi_{n+2}$-conservativity of $\text{B} \Sigma^{n+1} + \text{exp}$ over $\Sigma^1_n + \text{exp}$ can be proved in PV, a fragment of bounded arithmetic corresponding to polynomial-time computation. For $n \geq 1$, this answers a question of Clote, Hájek, and Paris.

The logical strength of Ramsey’s theorem for pairs and two colours, RT$_2^2$, has been a major topic of interest in reverse mathematics and related areas of logic for over 25 years [26, 4, 5, 6]. Recently, Patey and the third author [24] showed that RT$_2^2$ is $\forall \Sigma^0_2$-conservative over the usual base theory considered in reverse mathematics, RCA$_0$, which is axiomatized by recursive comprehension and $\Sigma^0_1$ induction. This new insight into the strength of RT$_2^2$ naturally leads to some further questions. One of these is, of course, whether the $\forall \Sigma^0_2$-conservativity over RCA$_0$ extends to full $\Pi^1_1$-conservativity over RCA$_0$ strengthened by the
Σ₀² collection principle (Π¹�-conservativity over RCA₀ being out of the question, as witnessed by Σ₀² collection itself). Here, we consider a different problem, already raised as Question 9.5 in [24]: does RT₂² have significant proof speedup over RCA₀ with respect to ∀Σ₀² sentences?

The study of proof size in axiomatic theories and proof speedup phenomena, a topic going back to Gödel [15], was given an excellent (if no longer fully up-to-date) survey in [25]. An important phenomenon that has been observed empirically is that, from a quantitative perspective, prominent cases in which an arithmetic theory T+ is conservative over a theory T for sentences of some class Γ typically fit one of two patterns. Either T+ has iterated exponential (“tower function”) speedup over T on proofs of sentences from Γ, or each proof of such a sentence in T+ can be translated into a proof in T with at most polynomial blowup. The former behaviour is illustrated for instance by the conservativity of arithmetical comprehension ACA₀ over PA and the Π₂� conservativity of IΣ₁ over primitive recursive arithmetic. The latter, by the conservativity of RCA₀ over Σ₁ and the Π¹�-conservativity of Weak König’s Lemma over RCA₀. The question we are interested in is whether the conservativity result of [24] also fits one of these patterns and, if so, which one.

Our main result is that RT₂² has at most polynomial speedup over RCA₀ for ∀Σ₀² sentences. Moreover, this is witnessed by a polynomial simulation: there is a polytime procedure that takes a proof of a ∀Σ₀² sentence in WKL₀ + RT₂² (where WKL₀ is RCA₀ extended by Weak König’s Lemma) and outputs a proof of the same sentence in RCA₀. By [3, Proposition 4.6], it follows that proofs of (purely first-order) Π₃ sentences in WKL₀ + RT₂² can be translated in polynomial time into proofs in Σ₁. To obtain our result, we make use of a general-purpose technique due to Avigad [1]: in order to show polynomial simulation of a theory T+ by T, build what could be called a “forcing interpretation” of T+ in T—that is, formalize within T a forcing construction that leads to a model of T+ and verify via small proofs in T that sentences of the appropriate class are forced if they are true in the ground model. Interestingly, however, while most forcing arguments give rise to a generic extension of the ground model, the one we work with produces a generic cut, refining an initial segment construction used in [24]. Both the construction of [24] and ours rely on non-trivial finite combinatorics. In the case of [24], this took the form of upper bound on Ramsey’s theorem for finite sets expressed in terms of so-called α-largeness. Here, we need a considerable strengthening of that bound with a more constructive proof, which was provided in [21].

We also consider the question whether an analogous polynomial simulation result still holds if we weaken the base theory to RCA₀, which differs from RCA₀ in that the Σ₀¹ induction scheme is replaced by Δ₀¹ induction (plus a separate axiom exp guaranteeing totality of the exponential function). It turns out that, perhaps surprisingly, changing the base theory makes a major difference: even though RT₂² is ∀Σ₂⁰-conservative over RCA₀ (as can be shown by the method used to derive Π₀�-conservativity in [30]), it has iterated exponential speedup over RCA₀ already for relatively simple sentences. We give a detailed proof of the speedup for Σ₁ sentences; in fact, it can also be witnessed by finite consistency statements, which can be written as bounded formulas with exponential terms.

Since RCA₀ + RT₂² proves Σ₀² collection, our main theorem immediately implies that the Paris-Friedman theorem, i.e. Πⁿ�₂-conservativity of BΣₙ₊₁ over Πⁿₙ, can be strengthened to a polynomial simulation for n = 1. It makes sense
to ask whether this extends to other values of \( n \), especially because a related question – whether the conservativity can be proved in bounded arithmetic – was asked by Clote et al. \[8\]. We prove that the answer to both questions is positive for all \( n \geq 1 \), and that it is positive for \( n = 0 \) if both \( B\Sigma_1 \) and \( I\Delta_0 \) are extended by \( \exp \). Our argument is based on turning one of the standard model-theoretic proofs of Paris-Friedman into a forcing interpretation. As discussed in slightly greater detail in Section 4, Fedor Pakhomov [private communication] has independently devised a completely different approach that proves a more general result; in particular, his argument obviates the need for \( \exp \) in the \( n = 0 \) case.

The remainder of this paper has a relatively straightforward structure. After introducing some basic definitions and notational conventions below, we discuss the general concept of forcing interpretations in Section 1. We prove our main theorem in Section 2 and the contrasting speedup result over \( \text{RCA}_0^\ast \) in Section 3. The polynomial simulation of \( B\Sigma_{n+1} + \exp \) by \( I\Sigma_n + \exp \) is presented in Section 4.

* 

We assume that the reader has some basic knowledge of fragments of first- and second-arithmetic as described in \[27\] or \[18\]. We will write \( \mathcal{L}_1 \) and \( \mathcal{L}_2 \) for the languages of first- and second-order arithmetic, respectively. Note that these languages do not have a symbol for exponentiation. We use lowercase letters for objects of the numerical (“first-order”) sort, and uppercase letters for objects of the set (“second-order”) sort. Notation like \( \Sigma^0_n, \Pi^0_n \) represents the usual formula classes defined in terms of first-order quantifier alternations, but allowing second-order free variables. On the other hand, notation without the superscript \( 0 \), like \( \Sigma_n, \Pi_n \), represents analogously defined classes of \( \mathcal{L}_1 \) formulas – that is, without any second-order variables at all. If we want to specify the second-order parameters appearing in a \( \Sigma^0_n \) formula, we use notation like \( \Sigma_n(X) \).

We extend these conventions to naming theories: thus, for example, \( B\Sigma_2 \) is the \( \mathcal{L}_2 \)-theory axiomatized by \( \Delta^0_0 \) induction and \( \Sigma^0_n \) collection, whereas \( B\Sigma_2 \) is the \( \mathcal{L}_1 \)-theory axiomatized by \( \Delta^0_0 \) induction and \( \Sigma^0_2 \) collection.

Recall that \( \text{RCA}_0^\ast \) is the theory defined in \[23\] which differs from \( \text{RCA}_0 \) in that \( \Sigma^0_1 \) induction is replaced by \( \Delta^0_0 \) induction plus an axiom known as \( \exp \) which states the totality of exponentiation. The first-order consequences of \( \text{RCA}_0^\ast \) coincide with the theory \( B\Sigma_1 + \exp \), which in particular means that \( \text{RCA}_0^\ast \) is \( \Pi_2 \)-conservative over \( I\Delta_0 + \exp \).

The symbol \( \mathbb{N} \) stands for the set of natural numbers – both as understood in the metatheory (the standard natural numbers) or as formalized in an \( \mathcal{L}_2 \) or \( \mathcal{L}_1 \) theory (which, if consistent, will of course have nonstandard models). It should be clear from the context which is meant. A set \( X \subseteq \mathbb{N} \) is finite if it is bounded, i.e. there is \( k \in \mathbb{N} \) such that \( \ell \leq k \) for all \( \ell \in X \); otherwise, \( X \) is infinite. Each finite set \( X \) is coded in a standard way by the binary representation of some natural number \( x \), and may be identified with \( x \) in many contexts. The symbol \( \omega \) stands for the smallest infinite ordinal – again, we use the same symbol in the metatheory and in formal theories.

We will be interested in comparing the sizes of proofs of statements in various theories, which requires us to fix some terminology and conventions regarding syntax. We use the word \textit{theory} to mean a set of sentences. Every theory \( T \)
comes with a language \( \mathcal{L}(T) \) with the property that each non-logical symbol appearing in \( T \) is in \( \mathcal{L}(T) \), though not necessarily vice versa. To simplify things, we assume that \( \mathcal{L}(T) \) always contains only finitely many non-logical symbols. The size of a term, a formula or a proof is the number of symbols in it. All variable, constant, function and relation symbols count as one symbol. We use vertical lines, \(| \cdot |\), to denote the size of a given syntactical object.

To measure the sizes of proofs precisely, one needs to fix a proof system. It is known that there are polynomial-time translations between the usual Hilbert-style systems and the sequent calculus with the cut rule [12], and between their tree-like or sequence-like versions [25, Theorem 4.1]. Therefore, in principle, it does not matter which proof system we choose here. In our arguments, we have in mind a Hilbert-style system like the one in Enderton [14, Section 2.4], except that instead of allowing all (universal closures of) propositional tautologies as axioms, we derive them from a finite number of propositional axiom schemes. We assume that, like in Enderton’s system, the only official connectives are \( \neg, \rightarrow \) and the only quantifier is \( \forall \). Recall [14, page 113] that a term \( t \) is substitutable for a variable symbol \( v \) in a formula \( \varphi \) if no free occurrence of \( v \) in \( \varphi \) is in the scope of a quantifier binding a variable that appears in \( t \).

**Definition 0.1.** Let \( T, T' \) be theories and let \( \Gamma \) be a set of sentences in the language \( \mathcal{L}(T) \cap \mathcal{L}(T') \). Then \( T \) polynomially simulates \( T' \) with respect to \( \Gamma \) if there exists a polynomial-time procedure that, given a proof of \( \gamma \in \Gamma \) in \( T' \) as input, outputs a proof of \( \gamma \) in \( T' \). (In particular, this implies that for every proof of \( \gamma \in \Gamma' \) in \( T' \), there is a proof of \( \gamma \) in \( T \) of at most polynomially larger size.)

\( T' \) has non-elementary speedup over \( T \) with respect to \( \Gamma \) if for each elementary recursive function \( f \), there exist \( \gamma \in \Gamma \) and a proof \( \pi \) of \( \gamma \) in \( T' \) such that no proof of \( \gamma \) in \( T \) has size \( \leq f(|\pi|) \).

## 1 Forcing interpretations

The technique we use for proving polynomial simulations was developed in Avigad [1] and relies on a certain kind of formalized forcing argument. In essence, the idea is to define a more general notion of interpretation: while a traditional interpretation gives a uniform way of defining a model \( M' \models T' \) inside a model \( M \models T \), an Avigad-style interpretation gives a way of describing a generic model \( M[G] \models T' \) which is not fully specified until the generic filter \( G \) is fixed. The point is that, even though no such \( G \) might be definable, the properties of \( M[G] \) that we care about in the context at hand do not depend on \( G \).

In this section, we discuss this idea, and its connection to questions of proof size, in some generality. We begin with an annoying technical issue.

### 1.1 Simplification

When defining interpretations, formulas with more than one function symbol in an atomic subformula can cause ambiguities. In this subsection, we explain how to get around the problem by avoiding such formulas as much as possible. This is quite similar in spirit to translating all formulas to a relational language.
**Definition 1.1.** A simple term is one of the form $v$ or $f(\bar{w})$, where $v, \bar{w}$ are variables. A simple formula is a formula in which every atomic subformula is either

- a formula with no function symbol; or
- an equation with exactly one function symbol.

**Definition 1.2.** For a formula $\theta$, its simple translation $\theta^*$ is defined in the following way.

First, for any given term $t$ and some canonically chosen variable $x$ not appearing in $t$, we translate $t$ into a formula $t[x]$ by induction on the subterms of $t$. For a variable $y$, the formula $y[x]$ is $x = y$. For $t$ of the form $f(t_1, \ldots, t_n)$, the formula $t[x]$ is $\forall x_1 \ldots \forall x_n (t_1[x_1] \rightarrow (t_2[x_2] \rightarrow \ldots \rightarrow (t_n[x_n] \rightarrow x = f(x_1, \ldots, x_n)) \ldots))$.

The formula $\theta^*$ equals $\theta$ for $\theta$ simple atomic. If $\theta$ is $R(t_1, \ldots, t_n)$ and is not simple, then $\theta^*$ is

$$\forall x_1 \ldots \forall x_n (t_1[x_1] \rightarrow (t_2[x_2] \rightarrow \ldots \rightarrow (t_n[x_n] \rightarrow R(x_1, \ldots, x_n)) \ldots))$$

for appropriately chosen variables $x_1, \ldots, x_n$. The translation commutes with $\neg$, $\rightarrow$, and $\forall$.

The following two lemmas mean that when studying questions of proof size, we can largely restrict attention to simple formulas. The lemmas can essentially be read off Visser’s exposition in [29, Section 7.3].

**Lemma 1.3.** There exists a polynomial-time procedure that, given a formula $\theta$, outputs a first-order logic proof of $\theta \leftrightarrow \theta^*$.

**Proof.** This is essentially Theorem 7.3.6 in Visser [29], with the following differences: (i) Visser defines the translation $\theta \mapsto \theta^*$ (where $\theta^*$ would be $\theta^* \circ$ in his notation) using $\exists$ and $\land$, while we use $\forall$ and $\rightarrow$ to match the official connectives of our proof system, and (ii) Visser only needs a formal proof of $\theta \leftrightarrow \theta^*$ of size polynomial in $|\theta|$, so he does not discuss the time required to construct it. However, these differences have no bearing on the proof, and the reader can verify that all of Visser’s constructions can be carried out in polynomial time.

**Lemma 1.4.** There is a polynomial-time procedure that, given a proof $\pi$ in first-order logic, outputs a proof $\pi^*$ with the following properties:

- $\pi^*$ is a proof in the same language as $\pi$;
- every formula in $\pi^*$ is simple;
- if $\theta_1, \theta_2, \ldots, \theta_\ell$ are the non-logical axioms in $\pi$, then $\theta_1^*, \theta_2^*, \ldots, \theta_\ell^*$ are the non-logical axioms in $\pi^*$;
- if $\eta$ is the conclusion of $\pi$, then $\eta^*$ is the conclusion of $\pi^*$.

**Proof.** See Theorem 7.3.3 and Theorem 7.3.4 in Visser [29], with the same caveats as in Lemma 1.3.
1.2 Forcing translations and interpretations

A traditionally understood interpretation of a theory $T'$ in a theory $T$ is essentially (i) a translation of $L(T')$ into $L(T)$ such that (ii) $T$ proves the translations of axioms of $T'$. Forcing interpretations of the sort we need here have a similar two-layered structure of “translation” and “interpretation proper”, except that now instead of translating an $L(T')$ formula $\varphi$ into $L(T)$, we have to translate “$\varphi$ is forced” into $L(T)$. Then we have to verify in $T$ that the axioms of $T'$ are forced.

**Definition 1.5.** A 
forcing translation

\[ \tau \]

from a language $L'$ to a language $L$ consists of $L$ formulas

\[ s \in \text{Cond}_\tau, \quad s' \trianglelefteq_s s, \quad s \vDash_{\tau} v \downarrow, \quad s \vDash_{\tau} \alpha(v_1, v_2, \ldots, v_\ell) \]

for every simple atomic $L'$ formula $\alpha(v_1, v_2, \ldots, v_\ell)$ such that

(FT1) $s' \trianglelefteq_s s$ contains $s' \in \text{Cond}_\tau$ as a conjunct;

(FT2) $s \vDash_{\tau} v \downarrow$ contains $s \in \text{Cond}_\tau$ as a conjunct;

(FT3) $s \vDash_{\tau} \alpha(v_1, v_2, \ldots, v_\ell)$ contains $\bigwedge_{i=1}^\ell s \vDash_{\tau} v_i \downarrow$ as a conjunct, whenever $\alpha(v_1, v_2, \ldots, v_\ell)$ is a simple atomic $L'$ formula; and

(FT4) if $\alpha(\bar{u}, v)$ is a simple atomic $L'$ formula and $w$ is a variable, then

\[ (s \vDash_{\tau} \alpha(\bar{u}, v))[v/w] \quad \text{and} \quad s \vDash_{\tau} (\alpha(\bar{u}, v)[v/w]) \]

are the same.

All formulas above have exactly the free variables shown. When there is no risk of ambiguity, we will often omit the subscript $\tau$ in the notation. We read ‘$s \vDash \ldots$’ as ‘$s$ forces $\ldots$’. By convention, the variable symbols $s, s', s'', \ldots$ are always distinct from $u, v, w, z, \ldots$. In accordance with the usual terminological customs related to forcing, we refer to elements $s$ satisfying $s \in \text{Cond}$ as forcing conditions, and we think of the objects denoted by the variables $u, v, \ldots$ as names (for the elements of the generic model being described).

**Remark.** In some contexts, it might make sense to allow forcing conditions to be tuples of elements instead of single elements. However, in this paper we deal exclusively with theories that have a definable pairing function, so to avoid complicating the notation we eschew that kind of generality and continue to write $s$ rather than $\bar{s}$.

The clauses we use to extend a forcing relation to arbitrary simple formulas follow those of a “good strong forcing notion” in the sense Avigad [1, Definition 4.2]. We deal with non-simple formulas using the $(\cdot)^*$ translation.

**Definition 1.6.** Let $\tau$ be a forcing translation from a language $L'$ to a language $L$. We define an $L$ formula $s \vDash_{\tau} \theta(v_1, v_2, \ldots, v_\ell)$ for each $L'$ formula $\theta(v_1, v_2, \ldots, v_\ell)$ by recursion on $\theta$ as follows.

(FT5) If $\theta(v_1, v_2, \ldots, v_\ell)$ is a simple $L'$ formula, then

\[ s \vDash_{\tau} \sim \theta(v_1, v_2, \ldots, v_\ell) \]
is defined to be
\[ \bigwedge_{i=1}^{\ell} "s \models \tau v_i \downarrow" \land \forall s' \leq \tau s \land s' \not\models \theta(\bar{v}). \]

(FT6) If \( \theta(u_1, u_2, \ldots, u_k, w_1, w_2, \ldots, w_m) \) and \( \eta(v_1, v_2, \ldots, v_\ell, w_1, w_2, \ldots, w_m) \) are simple \( L' \) formulas, then
\[ s \models \tau \theta(u_1, u_2, \ldots, u_k, w_1, w_2, \ldots, w_m) \rightarrow \eta(v_1, v_2, \ldots, v_\ell, w_1, w_2, \ldots, w_m) \]
is defined to be
\[ \bigwedge_{i=1}^{k} "s \models \tau u_i \downarrow" \land \bigwedge_{i=1}^{\ell} "s \models \tau v_i \downarrow" \land \bigwedge_{i=1}^{m} "s \models \tau w_i \downarrow" \land \forall s' \leq \tau s, \exists s'' \leq \tau s' ("s'' \models \tau \theta(\bar{u}, \bar{w})" \rightarrow "s'' \models \tau \eta(\bar{v}, \bar{w})"). \]

(FT7) If \( \theta(v_1, v_2, \ldots, v_\ell, w) \) is a simple \( L' \) formula, then
\[ s \models \tau \forall w \theta(v_1, v_2, \ldots, v_\ell, w) \]
is defined to be
\[ \bigwedge_{i=1}^{\ell} "s \models \tau v_i \downarrow" \land \forall w \forall s' \leq \tau s, \exists s'' \leq \tau s' ("s'' \models \tau w \downarrow" \rightarrow "s'' \models \tau \theta(\bar{v}, \bar{w})"). \]

(FT8) If \( \theta(\bar{v}) \) is an \( L' \) formula that is not simple, then \( s \models \tau \theta(\bar{v}) \) is defined to be \( s \models \tau \theta^*(\bar{v}) \), where \( \theta^*(\bar{v}) \) is the formula given by Lemma 1.3.

(FT9) \( \land, \lor, \rightarrow, \exists \) are defined in terms of \( \neg, \land, \lor \) in the usual way.

Here the quotation marks “...” are simply another type of brackets to enhance the readability of formulas. Our convention is that \( \models \) has lower precedence than all the logical connectives in first-order logic. All formulas above have exactly the free variables shown. We often abbreviate \( \bigwedge_{i=1}^{\ell} (s \models \tau v_i \downarrow) \) as \( s \models \tau \bar{v} \downarrow \).

The definitions above are designed to make (FT3)Item.3 and (FT4)Item.4 hold for all formulas.

**Lemma 1.7.** Let \( \tau \) be a forcing translation from a language \( L' \) to a language \( L \). Then:

(i) \( s \models \tau \theta(\bar{v}) \) contains \( s \models \bar{v} \downarrow \) as a conjunct for every \( L' \) formula \( \theta(\bar{v}) \),

(ii) if \( \theta(\bar{a}, v) \) is a \( L' \) formula and \( w \) is a variable, then
\[ (s \models \tau \theta(\bar{a}, v))[v/w] \land s \models \tau \theta(\bar{a}, v)[v/w] \]

are the same, and the variable \( w \) is substitutable for \( v \) in \( \theta(\bar{a}, v) \) exactly when it is so in \( s \models \tau \theta(\bar{a}, v) \).

**Proof.** Part (i) is true by construction. Part (ii), which is a special case of Lemma 4.3 in Avigad [1], can be proved by an easy induction on \( \theta \) using the definitions. For non-simple formulas \( \theta(\bar{a}, v) \), we assume the simplification procedure employs a canonical choice of variables.
Definition 1.8. A forcing interpretation of a theory $T'$ in a theory $T$ is a forcing translation $\tau$ from $\mathcal{L}(T')$ to $\mathcal{L}(T)$ such that $T$ proves

(F10) $\exists s \ (s \in \text{Cond}_\tau)$;

(F11) $\forall s \in \text{Cond}_\tau \ s \trianglelefteq_{\tau} s$;

(F12) $\forall s, s', s'' \in \text{Cond}_\tau \ (s'' \trianglelefteq_{\tau} s' \land s' \trianglelefteq_{\tau} s \rightarrow s'' \trianglelefteq_{\tau} s)$;

(F13) $\forall s \in \text{Cond}_\tau \ \exists s' \trianglelefteq_{\tau} s \ \exists v \ s' \triangleright_{\tau} v$;

(F14) $\forall s, s' \in \text{Cond}_\tau \ \forall v \ (s' \trianglelefteq_{\tau} s \land s \triangleright_{\tau} v \rightarrow s' \triangleright_{\tau} v')$;

(F15) for all simple atomic $\mathcal{L}(T')$ formulas $\alpha(\bar{v})$,

$$\forall s, s' \in \text{Cond}_\tau \ \forall \bar{v} \ (s' \trianglelefteq_{\tau} s \land s \triangleright_{\tau} \alpha(\bar{v}) \rightarrow s' \triangleright_{\tau} \alpha(\bar{v}))$$;

(F16) $\forall s \in \text{Cond}_\tau \ \forall v \ (s \triangleright_{\tau} v \rightarrow s \triangleright_{\tau} v = v)$;

(F17) $\forall s \in \text{Cond}_\tau \ \forall u, v \ (s \triangleright_{\tau} u = v \rightarrow s \triangleright_{\tau} v = u)$;

(F18) $\forall s \in \text{Cond}_\tau \ \forall u, v, w \ (s \triangleright_{\tau} u = v \land s \triangleright_{\tau} v = w \rightarrow s \triangleright_{\tau} u = w)$;

(F19) for all $k \in \mathbb{N}$ and all $k$-ary function symbols $f$ in $\mathcal{L}(T')$,

$$\forall s \in \text{Cond}_\tau \ \forall v_1, v_2, \ldots, v_k \left( \bigwedge_{i=1}^{k} "s \triangleright_{\tau} v_i" \rightarrow \forall s' \trianglelefteq_{\tau} s \ \exists s'' \trianglelefteq_{\tau} s' \ \exists w "s'' \triangleright_{\tau} w = f(\bar{v})" \right)$$

$$\land s \triangleright_{\tau} \forall w, w' \ (w = f(\bar{v}) \rightarrow (w' = f(\bar{v}) \rightarrow w = w'))$$;

(F10) for all simple $\mathcal{L}(T')$ terms $t(v_1, v_2, \ldots, v_k)$ with exactly the free variables

shown and all simple atomic $\mathcal{L}(T')$ formulas $\alpha(\bar{u}, \bar{w})$ such that $\alpha(\bar{u}, t(\bar{v}))$ is simple,

$$\forall s \in \text{Cond}_\tau \ \forall u, v_1, v_2, \ldots, v_k, w \ (s \triangleright_{\tau} w = t(\bar{v}) \rightarrow (s \triangleright_{\tau} \alpha(\bar{u}, w) \leftrightarrow s \triangleright_{\tau} \alpha(\bar{u}, t(\bar{v}))))$$;

(F11) $\forall s \in \text{Cond}_\tau \ \forall v \ (\forall s' \trianglelefteq_{\tau} s \ \exists s'' \trianglelefteq_{\tau} s' "s'' \triangleright_{\tau} v \rightarrow s \triangleright_{\tau} v")$;

(F12) for all simple atomic $\mathcal{L}(T')$ formulas $\alpha(\bar{v})$,

$$\forall s \in \text{Cond}_\tau \ \forall \bar{v} \ (\forall s' \trianglelefteq_{\tau} s \ \exists s'' \trianglelefteq_{\tau} s' "s'' \triangleright_{\tau} \alpha(\bar{v})" \rightarrow s \triangleright_{\tau} \alpha(\bar{v}))$$;

(F13) $\forall s \in \text{Cond}_\tau \ s \triangleright_{\tau} \sigma \ \ \text{for all} \ \sigma \in T'$.

We refer to a forcing translation satisfying [F10]Item.12 [F11]Item.24 (that is, to a forcing interpretation of pure logic) as a forcing interpretation of $\mathcal{L}(T')$ in $T$.\[8\]
Conditions (FI0) Item.12 and (FI3) Item.15 are technical. Conditions (FI1) Item.13–(FI2) Item.14 state that $\trianglelefteq$ is a preorder. On the other hand, we require neither antisymmetry nor the existence of a maximal element. (FI4) Item.16–(FI5) Item.17 form the base case of the usual requirement that forcing be closed downwards under $\trianglelefteq$. (FI6) Item.18–(FI10) Item.22 say roughly that equality axioms are forced. The variable symbol $w$ in (FI10) Item.22 may be syntactically equal to one of $\bar{u}$, but by convention $w_0$ cannot be. (FI11) Item.23–(FI12) Item.24 express the base case of the connection between forcing $\neg\neg \varphi$ and forcing $\varphi$. Finally, the crucial condition (FI13) Item.25 is what makes a forcing interpretation of pure logic be an interpretation of a theory $T'$: all axioms of $T'$ are forced.

Forcing interpretations are clearly closed under composition and definition by cases. As shown by the following example, our notion of forcing interpretations generalizes the usual notion of interpretation.

Example 1.9. Every interpretation (more precisely, every parameter-free one-dimensional global relative interpretation) $\tau$ of a theory $T'$ in a theory $T$ gives rise to a forcing interpretation of $T'$ in $T$ as follows.

- Define $s \in \text{Cond}_\tau$ and $s' \trianglelefteq \tau s$ to be respectively $s = s$ and $s' = s' \wedge s = s$.
- Define $s \models_\tau \nu \downarrow$ to be $s \in \text{Cond}_\tau \wedge \delta(v)$, where $\delta(v)$ is the defining formula for the domain of $\tau$.
- Define $s \models_\tau \alpha(\bar{v})$, where $\alpha(\bar{v})$ is a simple atomic $L(T')$ formula, according to the interpretation $\tau$.

1.3 Forcing interpretations and polynomial simulation

We now turn to the question of what is required of a forcing interpretation if it is to imply a polynomial simulation between theories.

We begin by verifying that proofs of some basic facts, including generalizations of (FI5) Item.17, (FI10) Item.22, and (FI12) Item.24 to all formulas, can be found in polynomial time for any forcing interpretation. Recall our convention that the language of a theory is always finite, which is implicitly used in the proofs of the lemmas below.

Lemma 1.10. Let $\tau$ be a forcing interpretation of a theory $T'$ in a theory $T$. Then there is a polynomial-time procedure which, given a $L(T')$ formula $\theta(\bar{v})$, outputs a proof in $T$ of

$$\forall s \in \text{Cond}_\tau \forall \bar{v} \neg ("s \models_\tau \theta(\bar{v})" \land "s \models_\tau \neg \theta(\bar{v})"").$$

Proof. Apply (FT5) Item.5.

Lemma 1.11. Let $\tau$ be a forcing interpretation of a theory $T'$ in a theory $T$. Then there is a polynomial-time procedure which, given a $L(T')$ formula $\theta(\bar{v})$, outputs proofs in $T$ of:

1. $\forall s, s' \in \text{Cond}_\tau \forall \bar{v} (s' \trianglelefteq \tau s \land "s \models_\tau \theta(\bar{v})" \rightarrow "s' \models_\tau \theta(\bar{v})")$; and
2. $\forall s \in \text{Cond}_\tau \forall \bar{v} (\forall s' \trianglelefteq \tau s \exists s'' \trianglelefteq \tau s' "s'' \models_\tau \theta(\bar{v})" \rightarrow s \models_\tau \theta(\bar{v})")$. 

9
Proof. These are part of Lemma 4.3 and Lemma 4.6 in Avigad [1], except that we additionally need to pay attention to the computational complexity of the proof constructions. The proofs are built using induction on the structure of $\theta$, with the inductive step split into cases depending on the outermost logical connective in $\theta$. In each case, the construction is straightforward.

Lemma 1.12. Let $\tau$ be a forcing interpretation of a theory $T'$ in a theory $T$. There is a polynomial-time procedure which – given a $L(T')$ term $t(\bar{v})$ with exactly the free variables shown and a simple $L(T')$ formula $\theta(\bar{u}, u_0)$ such that $\theta(\bar{u}, t(\bar{v}))$ is also simple, and both $w$ and $t(\bar{v})$ are substitutable for $u_0$ in $\theta(\bar{u}, u_0)$ – outputs a proof in $T$ of:

$$\forall s \in \text{Cond}_T \forall \bar{u}, \bar{v}, w \left(\left(\text{``}s \models_T w = t(\bar{v})\text{''} \rightarrow \left(\text{``}s \models_T \theta(\bar{u}, w)\text{''} \Leftrightarrow \text{``}s \models_T \theta(\bar{u}, t(\bar{v}))\text{''}\right)\right)\right).$$

Proof. This is part of Lemma 4.3 in Avigad [1], again with attention paid to the computational complexity of the construction. As previously, one proceeds by induction on the structure of $\theta$.

Definition 1.13. A forcing interpretation $\tau$ of a theory $T'$ in a theory $T$ is polynomial if there is polynomial-time procedure which, given any $\sigma \in T'$, outputs a proof in $T$ of the statement $\forall s \in \text{Cond}_T, s \models_T \sigma$.

Note that if $T'$ is finitely axiomatized, then any forcing interpretation of $T'$ is automatically a polynomial forcing interpretation.

Proposition 1.14. Let $\tau$ be a forcing interpretation of a theory $T'$ in a theory $T$. For all $L(T')$ formulas $\varphi(\bar{v}), \psi(\bar{v})$, if $T' + \varphi(\bar{v}) \vdash \psi(\bar{v})$, then $T$ proves $\forall s \in \text{Cond}_T \forall \bar{v} \left(\left(\text{``}s \models_T \varphi(\bar{v})\text{''} \rightarrow s \models_T \psi(\bar{v})\right)\right)$. (1)

If $\tau$ is polynomial, then a proof of (1) in $T'$ can be found in polynomial time given $\varphi, \psi$, and a proof $\pi$ of $\psi(\bar{v})$ from $T' + \varphi(\bar{v})$.

Proof. Let $\pi^*$ be the proof obtained from $\pi$ according to Lemma 1.14. By (FT8)Item.8 and (FI13)Item.25 for each non-logical axiom $\gamma$ used in $\pi^*$ we can find a proof in $T$ of $\forall s \in \text{Cond}_T (s \models_T \varphi(\bar{v}) \rightarrow s \models_T \gamma)$. If $\tau$ is a polynomial forcing interpretation of $T'$ in $T$, these proofs can be found in polynomial time.

We then construct analogous proofs for each line $\gamma$ in $\pi^*$, by a routine induction on the structure of $\pi^*$; cf. Proposition 4.8 in Avigad [1]. By (FT8)Item.8 in the case of the last line this is a proof of $\forall s \in \text{Cond}_T (s \models_T \varphi(\bar{v}) \rightarrow s \models_T \psi(\bar{v}))$.

We note the following important special case of Proposition 1.14.

Corollary 1.15. Let $\tau$ be a forcing interpretation of $T'$ in $T$. Then for each $L(T')$ sentence $\sigma$, if $T'$ proves $\sigma$, then $T$ proves $\forall s \in \text{Cond}_T (s \models_T \sigma)$. Moreover, if $\tau$ is polynomial, then a proof of $\forall s \in \text{Cond}_T (s \models_T \sigma)$ in $T$ can be found in polynomial time given a proof of $\sigma$ in $T'$.

Definition 1.16. Let $T, T'$ be theories and $\Gamma$ be a set of sentences in the language $L(T) \cap L(T')$. A forcing interpretation $\tau$ of $T'$ in $T$ is said to be $\Gamma$-reflecting if, for all $\gamma \in \Gamma$, $T$ proves $\forall s \in \text{Cond}_T (s \models_T \gamma) \rightarrow \gamma$. (2)

The interpretation is polynomially $\Gamma$-reflecting if a $T$-proof of (2) can be found in polynomial time on input $\gamma \in \Gamma$.  

10
The combinatorial statement is formulated using the concept of \( \alpha \)-largeness (originally introduced in [20] as a tool to study the unprovability of the Paris–Harrington theorem), which provides a framework for measuring the size of finite subsets of \( \mathbb{N} \) by means of countable ordinals rather than just natural numbers.

We now recall some basic notions related to that framework.

For a given \( \alpha < \omega \) and \( m \in \mathbb{N} \), define \( 0[m] = 0 \), \( \alpha[m] = \beta \) if \( \alpha = \beta + 1 \), and \( \alpha[m] = \beta + \omega^{n-1} \cdot m \) if \( \alpha = \beta + \omega^n \) for some \( n \geq 1 \).

Definition 2.3 (I(\( \Delta_0 \) + exp)). Let \( \alpha < \omega^\omega \). A set \( X = \{x_0 < \cdots < x_{\ell-1}\} \subseteq_{\text{fin}} \mathbb{N} \) is said to be \( \alpha \)-large if \( \alpha[x_0] \cdots [x_{\ell-1}] = 0 \). In other words, any finite set is 0-large, and \( X \) is \( \alpha \)-large when

- \( X \setminus \{\min X\} \) is \( \beta \)-large if \( \alpha = \beta + 1 \),
- \( X \setminus \{\min X\} \) is \( (\beta + \omega^{n-1} \cdot \min X) \)-large if \( \alpha = \beta + \omega^n \) for \( n \geq 1 \).

The proposition below lists some well-known and simple but important properties of \( \alpha \)-largeness. Their provability in I(\( \Delta_0 \) + exp) was pointed out e.g. in [21].

**Proposition 2.4.** The following are provable in I(\( \Delta_0 \) + exp):

(i) If a set \( X \) is \( \alpha \)-large and \( X \subseteq Y \subseteq_{\text{fin}} \mathbb{N} \), then \( Y \) is \( \alpha \)-large.

(ii) If \( \alpha = \omega^{n_k} + \cdots + \omega^{n_0} \) where \( n_k \geq \cdots \geq n_0 \), then a set \( X \) is \( \alpha \)-large if and only if there exist sets \( X_0, \ldots, X_k \) such that \( X = X_0 \sqcup \cdots \sqcup X_k \), each \( X_i \) is \( \omega^{n_i} \)-large, and for each \( i < k \), max \( X_i < \min X_{i+1} \).

As a result, if \( X \) is \( \omega^n \)-large and \( k = \min X \), then \( X \setminus \{k\} = X_0 \sqcup \cdots \sqcup X_{k-1} \) where each \( X_i \) is \( \omega^{n_i-1} \)-large, and for each \( i < k - 1 \), max \( X_i < \min X_{i+1} \).
(iii) If \( X \) is \( \omega^n \cdot 2 \)-large and \( X = X_0 \cup X_1 \), then at least one of \( X_0, X_1 \) is \( \omega^n \)-large.

The following result is crucial in the proof of the main theorem of [24].

**Proposition 2.5.** [24 Proposition 7.7] For every natural number \( n \) there exists a natural number \( m \) such that \( I\Sigma_1 \) proves: for every \( X \subseteq \text{fin} \mathbb{N} \) with \( \min X \geq 3 \), if \( X \) is \( \omega^m \)-large, then every colouring \( f : [X]^2 \to 2 \) has an \( \omega^n \)-large homogeneous set.

As the analysis in the remainder of this section will reveal, proving Theorem 2.1 involves obtaining an upper bound on \( m \) in terms of \( n \). The statement of Proposition 2.5 gives no such upper bound, and any bound that could be derived from a simple-minded analysis of the proof in [24] would be very weak — certainly not even elementary recursive. On the other hand, proving Theorem 2.1 requires a bound significantly better than exponential. We will use the bound obtained in [21].

**Theorem 2.6.** [21 Theorem 1.6] Provably in \( I\Sigma_1 \), if \( X \subseteq \text{fin} \mathbb{N} \) is \( \omega^{300x} \)-large, then every colouring \( f : [X]^2 \to 2 \) has an \( \omega^x \)-large homogeneous set.

**Remark.** Note that in [21], the theorem includes the additional assumption that \( \min X \geq 3 \). However, using the proof presented in [21] and some very straightforward combinatorics of \( \alpha \)-large sets, it is easy to verify that the assumption is in fact unnecessary. To avoid annoying technicalities, we nevertheless assume that any set \( X \) mentioned in the context of \( \alpha \)-largeness satisfies \( \min X \geq 1 \).

Our proof of Theorem 2.1 will involve an argument by cases, in which the case distinction is based on how much induction is available. We introduce an auxiliary theory, originally motivated by the case in which a certain amount of induction fails, which turns out to be useful in the other case as well. The theories and polynomial simulation relationships we consider in the proof are summarized in Figure 1.

**Definition 2.7.** \( \text{RCA}_0 + \mathbb{I} \) is a theory in the language of second-order arithmetic extended by a new unary predicate \( \mathbb{I} \) over the first-order sort.

The axioms of \( \text{RCA}_0 + \mathbb{I} \) are those of \( \text{RCA}_0 \) plus the following statements:

- \( (I) \) \( \mathbb{I} \) is a nonempty proper cut in the first-order universe,
- \( (II) \) \( \mathbb{I} \) is closed under addition,
- \( (III) \) for every infinite set \( S \), there exists finite \( X \subseteq S \) which is \( \omega^x \)-large for some \( x > \mathbb{I} \).

For our purposes, the most important property of the theory \( \text{RCA}_0 + \mathbb{I} \) will be that it polynomially simulates \( \text{WKL}_0 + \text{RT}^2_2 \) with respect to proofs of \( \forall \Sigma^0_2 \) sentences. Before proving that fact, which is Lemma 2.10 below, we verify some more basic properties of \( \text{RCA}_0 + \mathbb{I} \).

**Lemma 2.8.** \( \text{RCA}_0 + (I1) \) polynomially simulates \( \text{RCA}_0 + \mathbb{I} \) with respect to \( \mathcal{L}_2 \) sentences.
Figure 1: Polynomial simulations between the various theories in Section 2

Proof. It is enough to show that there is an interpretation of RCA\(_0 + \mathbb{I}\) in RCA\(_0 + (\Pi)\) that is the identity interpretation with respect to all symbols of L\(_2\).

So, let J consist of those numbers x for which every infinite set contains an \(\omega^x\)-large finite subset; this is easily seen to be a definable cut in RCA\(_0\).

Thus, I \cap J is, provably in RCA\(_0 + (\Pi)\), a proper definable cut with the property that every infinite set X actually contains an \(\omega^x\)-large subset for each x \(\in\) I \cap J. It follows that every infinite set X actually contains an \(\omega^x\)-large subset for some x \(>\) I \cap J, because the formula “X contains an \(\omega^x\)-large subset” is \(\Sigma^0_1\), so by I\(\Sigma^0_2\) it cannot define exactly the cut I \cap J. Let K be a cut closed under addition obtained from I \cap J by applying Solovay’s technique of shortening cuts: in other words, put K = \(\{ a \in I \cap J : \forall x \in I \cap J (a + x \in I \cap J) \}\). Then RCA\(_0 + (\Pi)\) proves the axioms of RCA\(_0 + \mathbb{I}\) with K substituted for I. \(\square\)

Lemma 2.9. RCA\(_0 + (\Pi)\) is polynomially simulated by:

(a) RCA\(_0 + \Sigma^0_2\) with respect to \(\forall \Sigma^0_3\) sentences,

(b) RCA\(_0 + \neg I\Sigma^0_2\) with respect to L\(_2\) sentences.

Proof. The proof of (b) is essentially immediate: by interpreting I as the intersection of all \(\Sigma^0_2\)-definable cuts, which is a definable proper cut in RCA\(_0 + \neg I\Sigma^0_2\), we obtain an interpretation of RCA\(_0 + (\Pi)\) in RCA\(_0 + \neg I\Sigma^0_2\) that is the identity with respect to L\(_2\).

In the case of (a) let us first explain the underlying model-theoretic argument, which shows how to construct a model of RCA\(_0 + (\Pi)\) from a model of RCA\(_0 + \Sigma^0_2\) while preserving the truth of a fixed \(\exists \Pi^0_3\) sentence. We then give some hints as to how this is reflected on the syntactic level.

Let \((M, \mathcal{X}) \models \text{RCA}_0 + I\Sigma^0_2 + \exists X \delta(X)\), where \(\delta \in \Sigma^0_4\). Take \(A \in \mathcal{X}\) such that \((M, \mathcal{X}) \models \delta(A)\). From Beklemishev [2] Theorem 5.1, we know that I\(\Sigma^0_2\) proves
the uniform \( \Pi^1 \) reflection principle for \( \Sigma^1 \). A relativization of this implies \( (M', \mathcal{P}') \models \text{Con}(\Sigma^1_0 + \delta(P)) \), where \( P \) is a new predicate. An application of the Low Arithmetized Completeness Theorem \cite[Theorem I.4.27]{Kechris} then gives an end-extension \( M' \supseteq M \) and \( A' \subseteq M' \) such that \( (M', A') \models \Sigma^1_0 + \delta(P) \).

Thus \( (M', \Delta_1(P)-\text{Def}(M', A'), M) \models \text{RCA}_0 + (\Pi_1 + \exists X \delta(X)) \). Moreover, the universe and operations in \( (M', A') \) are definable in \( M \), as is the range of a bijection between \( M \) and an initial segment of \( M' \) closed under \( + M' \).

Crucially, the formulas used to define \( (M', A') \) are not only independent of \( M \models \text{Con}(\Sigma^0_0 + \delta(P)) \), but they are also polynomial-time uniform in \( \delta \), in the sense that they can be constructed in polynomial time given \( \delta \) as input.

(Each of the formulas arises essentially by substituting the Gödel number of \( \exists X \delta(X) \) into a \( \Sigma^2 \) formula \( \gamma(x) \) describing a low branch of a binary tree that is \( \Delta^1 \)-definable with first-order parameter \( x \) and is infinite as long as \( x \) is the Gödel number of a sentence consistent with \( \Sigma^0_0 \).) So, if \( \pi \) is a proof of \( \perp \) from \( \text{RCA}_0 + (\Pi_1 + \exists X \delta(X)) \), we can build a proof of \( \perp \) from \( \Sigma^2_2(P) + \delta(P) \) by taking \( \pi^* \) and “relativizing it to \( (M', \Delta_1(P)-\text{Def}(M', P), M) \)”. That is, we relativize first-order quantifiers to the formula defining the universe of \( M' \) and replace first-order atoms by the formulas defining operations in \( M' \); replace second-order quantifiers by quantification over pairs of a \( \Sigma^1_1(P) \) formula and an equivalent \( \Pi_1(P) \) formula, changing atoms \( x \in X \) (for \( X \) different from \( P \)) into appropriate instances of the \( \Sigma^1_1(P) \)-universal formula; replace \( \perp \) by the formula defining the range of the inclusion \( M \hookrightarrow M' \). Additionally, it is necessary to add polynomially many new proof lines in order to derive the translated axioms of \( \text{RCA}_0 + (\Pi_1 + \exists X \delta(X)) \) from \( \Sigma^2_2(P) + \delta(P) \), and to deriving translations of conclusions from translations of premises for each inference in \( \pi^* \). The details of the latter task are quite standard and somewhat similar to those in the proof that \( \Sigma^1 \) polynomially simulates \( \text{RCA}_0 \) on the usual interpretation of \( \text{RCA}_0 \) in \( \Sigma^1 \), as described in e.g. \cite[Chapter 3]{Koer}. \( \square \)

**Lemma 2.10.** \( \text{RCA}_0 + I \) polynomially simulates \( \text{WKL}_0 + \text{RT}_2^2 \) with respect to \( \forall \Sigma^0_2 \) sentences.

To prove Lemma 2.10, we now define a particular forcing notion in \( \text{RCA}_0 + I \). We will then use a series of lemmas to show that this provides a polynomial forcing interpretation of \( \text{WKL}_0 + \text{RT}_2^2 \) in the sense of Section \ref{forcing} and, moreover, that the forcing interpretation is polynomially \( \forall \Sigma^0_2 \)-reflecting in the sense of Definition \ref{reflect}. \( \square \)

**Definition 2.11.** We let a \emph{finite} set \( s \) be a forcing condition \((s \in \text{Cond})\) if and only if \( s \) is \( \omega^x \)-large for some \( x > I \). The relation \( \subseteq \) is defined simply as \( \subseteq \). Note that there is no \( \subseteq \)-largest condition.

There are two sorts of names, a first- and a second-order sort. A name of the first-order sort is simply a natural number. A name of the second-sort is also simply a natural number, this time viewed as coding a finite set according to the Ackermann interpretation. (Formally, one may think of the first-order names being \((0, v)\) and second-order names being \((1, v)\) for various numbers \( v \).) To avoid notational confusion, we will write \( v \) for the first-order and \( V \) for the second-order names.

We say that \( s \vDash v \downarrow \) if \( s \cap [1, v] \) is not a condition, that is, if it is not \( \omega^x \)-large for any \( x > I \). On the other hand, \( s \vDash V \downarrow \) holds always.
For each simple atomic formula $\varphi(\bar{v})$, we define $s \vdash \varphi$ to be $(s \vdash v_1) \land \varphi(\bar{v})$. (In the special case of the simple atomic formula $v \in V$, this would be more precisely stated as $(s \vdash v \downarrow) \land v \in_{\text{Ack}} V$, where $\in_{\text{Ack}}$ is the standard Ackermann interpretation of $\in$ in arithmetic. We shall ignore this detail from now on.)

Before showing that this forcing notion provides a polynomial forcing interpretation of WKL$_0 + \text{RT}^2_{\omega}$, we first discuss the model-theoretic intuition behind it. Let $(\mathcal{M}, \mathcal{P}, I)$ be a countable model of RCA$_0 + \mathbb{I}$. The set of forcing conditions Cond and the relation $\trianglelefteq$ of Definition 2.7 are both definable in $(\mathcal{M}, \mathcal{P}, I)$. Take an $(\mathcal{M}, \mathcal{P}, I)$-generic filter $G$ of $(\text{Cond}, \trianglelefteq)$, and put $I_G = \sup_{\mathcal{M}} \{ \min s : s \in G \}$. One may then check that $(G, \text{Cod}(M/I_G)) \models \text{WKL}_0 + \text{RT}^2_{\omega}$ by simulating the inductive construction of a cut in the proof of [21, Theorem 3.3], where various choices made in the inductive steps are now replaced by the genericity of $G$.

One may also check that for any $v \in \mathcal{M}$, $v \in I_G$ if and only if $s \cap [1, v] \notin \text{Cond}$ for some $s \in \mathcal{G}$, and that $s \cap I_G$ is cofinal in $I_G$ for any $s \in \mathcal{G}$.

Thus, in Definition 2.11, $s \vdash v \downarrow$ is intended to hold only if $s \in G$ guarantees that $v$ will be in $I_G$. On the other hand, the name $V$ is intended to refer to $V \cap I_G$, which will always exist. The structure of $\vdash$, $\trianglelefteq$ on $I_G$ is inherited from the ground model, which motivates the trivial definition of forcing atomic formulas.

Remark. By Proposition 2.4(iii), whenever an $\omega^x$-large set is split into two subsets, one of them is $\omega^{x-1}$-large. As a consequence, if $s \vdash v \downarrow$, then $s \setminus [1, v]$ is a condition.

**Lemma 2.12.** The set of conditions Cond and the relations $\trianglelefteq$, $\vdash$ of Definition 2.11 determine a forcing interpretation of $\mathcal{L}_2$ in the theory RCA$_0 + \mathbb{I}$.

**Proof.** Checking that Cond, $\trianglelefteq$, $\vdash$ determine a forcing translation from $\mathcal{L}_2$ into $\mathcal{L}_2 \cup \{ \mathbb{I} \}$ is unproblematic, clause (FI13)Item.25 of the definition of forcing interpretation trivializes, and most of the other clauses are also very easy to verify in RCA$_0 + \mathbb{I}$. We discuss (FI9)Item.21 and (FI11)Item.23, which are perhaps less obvious than the rest.

To verify (FI9)Item.21, note that if the set $(x, y)$ is $\omega$-large, then $y \geq 2x$, and if it is $\omega^x$-large, then $y > x \omega^x \geq x^2$. Assume that $s$ is $\omega^x$-large for some $x > 1$ and that $s \vdash v_1 \downarrow$, $s \vdash w_1 \downarrow$ with say $v \leq w$. Since $s \cap [1, w]$ is not a condition and neither is $s \cap [w, w^2]$ (being a subset of the set $(w, w^2)$, which is not even $\omega^2$-large), it follows from Proposition 2.4(iii) that $s \cap [1, w^2]$ is not a condition. Thus, $s \vdash (w^2)\downarrow$, and a fortiori, $s$ forces both $(v + w)\downarrow$ and $(vw)\downarrow$. This proves the “existence” part of (FI9)Item.21. The “uniqueness” part follows easily from (FT7)Item.7 and (FT6)Item.6 and the definition of $\vdash$.

To show (FI11)Item.23 assume $s \not\vdash v \downarrow$. Then $s' := s \cap [1, v]$ is a condition with $s' \leq s$. However, for any condition $s'' \leq s'$, we have $s'' \cap [1, v] = s''$, so certainly $s'' \not\vdash v \downarrow$.

We can use the proof of Lemma 2.12 to show that, actually, a slightly stronger version of (FI9)Item.21 holds.

**Lemma 2.13.** There is a polynomial-time procedure which, given a term $t(\bar{v})$, outputs an RCA$_0 + \mathbb{I}$ proof of the statement that, for any condition $s$, if $s \vdash \bar{v}_1 \downarrow$, then $s \vdash t \downarrow$.  

15
Proof. Using a standard argument, we can prove in RCA$_0$ that $t < (\max \bar{v})^n$, for some $n \in \omega$ which is bounded by the number of symbols in $t$. By the proof of Lemma 2.12, we can also prove in RCA$_0 + \forall$ that for any $s$ and $w$, if $s \Vdash w \downarrow$, then $s \Vdash (w^2) \downarrow$. Iterating this reasoning $\log n$ times, we prove that if $s \Vdash w \downarrow$, then $s \Vdash (w^n) \downarrow$. We apply this to $w := \max \bar{v}$ in order to deduce $s \Vdash (\max \bar{v})^n \downarrow$ and hence $s \Vdash t \downarrow$. \hfill $\square$

Lemma 2.14. There is a polynomial-time procedure which, given a $\Delta_0$ formula $\theta(\bar{v}, \bar{V})$, outputs an RCA$_0 + \forall$ proof of

$$\forall s \forall \bar{v} \forall \bar{V} [s \Vdash \bar{v} \downarrow \rightarrow (\theta(\bar{v}, \bar{V}) \leftrightarrow (s \Vdash \theta(\bar{v}, \bar{V})))].$$

Proof. The construction of the proof proceeds by induction on the structure of $\theta$. As the reader will be able to check, it can be carried out in polynomial-time. We give an informal description of how the proof is built for the atomic step and for each case of the inductive step.

Suppose $\theta$ is atomic. If $\theta$ is a simple formula, then the equivalence between $\theta$ and $s \Vdash \theta$ follows immediately from the definition of our forcing relation. Otherwise, if $\theta$ is say $t_1(\bar{v}) \leq t_2(\bar{v})$ (the other cases are similar or easier), then $\theta^*$ is defined to be

$$\forall x_1 \forall x_2 (t_1(x_1) \rightarrow (t_2(x_2) \rightarrow x_1 \leq x_2)),$$

(cf. Definition 1.2) and by (FT8) Item.8, we have to prove that a condition forces $\theta^*$ exactly if it forces $\bar{v} \downarrow$ and $\bar{v}$ in fact holds. To achieve this, one first shows by induction on subterms $t'(\bar{u})$ of $t_1$ and $t_2$ (where $\bar{u}$ is a subtuple of $\bar{v}$) that, for the appropriate variable $x'$ and any $s \in \text{Cond}$,

$$s \Vdash t'(x') \text{ iff } (s \Vdash \bar{u} \downarrow \land x' = t(\bar{u})).$$  \hfill (3)

The inductive step is proved using (FT7) Item.7 and (FT6) Item.6, the definition of $\Vdash$, and Lemma 2.13. Once (3) is proved for $t'$ equal to $t_1$ and $t_2$, one can prove the equivalence of $s \Vdash \theta^*$ with $(s \Vdash \bar{u} \downarrow \land t_1 \leq t_2)$ by one more series of appeals to (FT7) Item.7 and (FT6) Item.6, the definition of $\Vdash$, and Lemma 2.13.

Suppose $\theta$ is $\neg \eta$. The inductive assumption gives us a proof that whenever $s \Vdash \bar{u} \downarrow$, then $s \Vdash \eta(\bar{v}, \bar{V})$ is equivalent to $\eta(\bar{v}, \bar{V})$. This equivalence generalizes to all $s' \leq s$, since for each such $s'$ we have $s' \Vdash \bar{u} \downarrow$ as well.

So, let $s \Vdash \neg \eta$. Assuming $\neg \eta$, we have $s' \not\Vdash \eta$ for any $s' \leq s$, and thus $s \Vdash \neg \eta$. On the other hand, if $s \Vdash \neg \eta$, then by the definition of forcing $s \not\Vdash \eta$ and hence $\neg \eta$.

Suppose $\theta$ is $\eta(\bar{v}, \bar{V}) \rightarrow \xi(\bar{v}, \bar{V})$. The inductive assumption gives us a proof that whenever $s \Vdash \bar{v} \downarrow$ and $s' \leq s$, then $s \Vdash \eta(\bar{v}, \bar{V})$ is equivalent to $\eta(\bar{v}, \bar{V})$ and likewise for $\xi$.

Let $s \Vdash \bar{v} \downarrow$. Assume $\eta \rightarrow \xi$. If $s' \leq s$ forces $\eta$, then $\eta$ holds and by our assumption so does $\xi$, which implies $s' \Vdash \xi$. This shows $s \Vdash \eta \rightarrow \xi$. Conversely, assume $\eta \land \neg \xi$. Then $s \Vdash \eta$ but for each $s' \leq s$ we have $s' \not\Vdash \xi$. So, $s \not\Vdash \eta \rightarrow \xi$.

16
Suppose \( \theta(\bar{v}, V) \) is \( \forall w \leq t \eta(w, \bar{v}, V) \). Recall that \( \forall w \leq t \ldots \) is shorthand for \( \forall w (w \leq t \rightarrow \ldots) \). (Here \( t = t(\bar{v}) \) is a term in the variables \( \bar{v} \).) The atomic step and the inductive assumption, respectively, give us proofs that if \( s \vDash \bar{v}_\downarrow \) and \( s \vDash w_\downarrow \), then

\[
s \vDash w \leq t \text{ is equivalent to } w \leq t,
\]
\[
s \vDash \eta(w, \ldots) \text{ is equivalent to } \eta(w, \ldots).
\]

As in the previous cases, the equivalences generalize to all \( s' \subseteq s \).

Moreover, by Lemma 2.13, we have a polynomial-time constructible proof that if \( s \vDash \bar{v}_\downarrow \), then \( s' \subseteq s \) implies \( s' \vDash t(\bar{v})_\downarrow \) and thus \( s' \vDash w_\downarrow \) for any \( w \leq t(\bar{v}) \).

Let \( s \vDash \bar{v}_\downarrow \) and assume \( \forall w \leq t \eta(w, \ldots) \). Take \( s' \subseteq s \) and \( w \) such that \( s' \vDash w_\downarrow \) and \( s' \vDash w \leq t \). Then the atomic step gives \( w \leq t \), which, by our assumption, implies \( \eta(w, \ldots) \). Thus \( s' \vDash \eta(w, \ldots) \), which is what we want.

Conversely, assume \( w \leq t \) is such that \( \neg \eta(w, \ldots) \). Then \( s \vDash w \leq t \). Also, for each \( s' \subseteq s \) we have \( s' \vDash w_\downarrow \), so \( s' \nvdash \eta(w, \ldots) \). Thus, \( s \nvdash \forall w \leq t \eta(w, \ldots) \). \( \square \)

**Lemma 2.15.** The relations \( \subseteq, \vdash \) of Definition 2.11 determine a polynomial forcing interpretation of \( \text{WKL}_0 + \text{RT}_2^2 \) in \( \text{RCA}_0 + \mathbb{I} \).

**Proof.** We have already shown in Lemma 2.12 that \( \subseteq, \vdash \) determine a forcing interpretation of \( L_2 \) in \( \text{RCA}_0 + \mathbb{I} \). So if we want to argue that this is actually a forcing interpretation of \( \text{WKL}_0 + \text{RT}_2^2 \), we only have to check the \( \text{PP} \)Item 25. In other words, we must show that any condition \( s \) forces each of the axioms of \( \text{WKL}_0 + \text{RT}_2^2 \). This is immediate for the axiomatization of non-negative parts of discrete ordered rings, so it remains to deal with \( \Delta^0_1 \)-comprehension, \( \Sigma^0_1 \)-induction, Weak König’s Lemma, and \( \text{RT}_2^2 \).

We reason in \( \text{RCA}_0 + \mathbb{I} \), sometimes making tacit use of the following simple observation: for any condition \( s \) and any \( k \), a condition \( s' \subseteq s \) with \( s' \vdash k_\downarrow \) exists exactly if \( s \setminus \{1, k\} \) is a condition.

We first deal with the \( \Sigma^0_1 \)-separation scheme, which implies both \( \Delta^0_1 \)-comprehension and Weak König’s Lemma. Consider a \( \Sigma^0_1 \)-separation axiom:

\[
\forall V \forall \bar{v} \exists k (\sigma_1(k, \bar{v}, V) \rightarrow \neg \sigma_2(k, \bar{v}, V)) \rightarrow \exists W \forall k ((\sigma_1(k, \bar{v}, V) \rightarrow k \in W) \land (k \in W \rightarrow \neg \sigma_2(k, \bar{v}, V))),
\]

where \( \sigma_1 := \exists \ell \eta_1(\ell, k, \ldots) \) and \( \sigma_2 := \exists \ell \eta_2(\ell, k, \ldots) \) are \( \Sigma^0_1 \). Let \( s \) be a condition and assume that \( s \) forces \( V_\downarrow, \bar{v}_\downarrow \) and that

\[
s \vdash \forall k (\sigma_1(k, \bar{v}, V) \rightarrow \neg \sigma_2(k, \bar{v}, V)). \tag{4}
\]

By the definition of forcing and Lemma 2.13, (4) implies the following: for every \( k, \ell_1 \leq \max s \) such that \( s \setminus \{1, \max(k, \ell_1)\} \) is a condition, if \( \eta_1(\ell_1, k, \ldots) \) holds, then there is no \( \ell_2 \leq \max s \) such that \( \eta_2(\ell_2, k, \ldots) \) and \( s \setminus \{1, \ell_2\} \) is a condition. In particular, it cannot be the case that \( \ell_1 \geq \ell_2 \). Thus, taking

\[
W_0 := \{ k \leq \max s : \exists \ell_1 \leq \max s (\eta_1(\ell_1, k, \ldots) \land \forall \ell_2 < \ell_1 \neg \eta_2(\ell_2, k, \ldots)) \},
\]

we conclude that \( s \vdash \forall k ( (\sigma_1(k, \ldots) \rightarrow k \in W_0) \land (k \in W_0 \rightarrow \neg \sigma_2(k, \ldots)) ). \)

We now turn to \( \Sigma^0_1 \)-induction. Consider the induction axiom for the \( \Sigma^0_1 \) formula \( \sigma := \exists \ell \eta(\ell, k, \bar{v}, V) \) with respect to the variable \( k \). Let \( s \) be a condition
forcing \( \dot{v}, \dot{V} \) and let \( j \) be such that \( s \setminus [1, j] \) is a condition. Let \( m = \min(s \setminus [1, j]) \) and let \( x > 1 \) be such that \( s \setminus [1, j] \) is \( \omega^x \)-large. It follows that \( s \setminus [1, m] \) is \( \omega^{x+1} \)-large, which by Proposition 22.1([ii]) means that we can write \( s \setminus [1, m] \) as a disjoint union

\[ s \setminus [1, m] = s_0 \sqcup \cdots \sqcup s_{m-1}, \]

where each set \( s_i \) is \( \omega^{x-1} \)-large (in particular, it is a condition) and, for each \( i < m-1 \), \( \max s_i < \min s_{i+1} \). Define:

\[ W_0 := \{ \ell : \exists k < j (\eta(\ell, k, \ldots) \land \forall \ell' < \ell \lnot \eta(\ell', k, \ldots)) \}. \]

Since \( W_0 \) has at most \( j \) elements and \( j < m \), the finite pigeonhole principle implies that \( W_0 \cap [\min s_{i_0}, \max s_{i_0}] = \emptyset \) for some \( i_0 < m \). Note that for every \( k < j \) we have

\[ \exists \ell < \min s_{i_0} \eta(\ell, k, \ldots) \text{ iff } \exists \ell \leq \max s_{i_0} \eta(\ell, k, \ldots). \quad (5) \]

By bounded induction on \( k \), either \( \lnot \exists \ell \leq \max s_{i_0} \eta(\ell, 0, \ldots) \) or there is a maximal \( k < j \) such that \( \exists \ell \leq \max s_{i_0} \eta(\ell, k, \ldots) \). By the property of \( s_{i_0} \) stated in ([5]), it follows that \( s_{i_0} \) forces one of the three statements \( \lnot \sigma(0, \ldots), \sigma(j-1, \ldots), \text{ or } \exists k < j-1 (\sigma(k, \ldots) \land \lnot \sigma(k+1, \ldots)) \). In each case, the existence of \( s_{i_0} \) implies that \( s \) forces induction for \( \sigma \) below \( j \).

Finally, we deal with \( RT^2_2 \). Assume that \( s \) is a condition such that \( s \vdash f \downarrow \) and \( s \vdash f : [\mathbb{N}]^2 \to 2 \). This means that \( f \) is defined on each pair \( (i, j) \) such that \( i < j < \max s \) and \( s \setminus [1, j] \) is a condition. Writing \( s = s_0 \sqcup s_1 \) where \( \max s_0 < \min s_1 \) and \( s_0, s_1 \) are both conditions, we can conclude that \( f \) is defined on all arguments below \( \max s_0 \).

Let \( x > 1 \) be such that \( s_0 \) is \( \omega^x \)-large. By Theorem 2.6 there exists an \( \omega^{x/300} \)-large set \( s' \subseteq s_0 \) which is homogeneous w.r.t \( f \). By ([2]), \( I \) is closed under addition, so \( x/300 > II \), and hence \( s' \) is a forcing condition.

It is easily verified that \( s' \) forces “\( s' \) is homogeneous w.r.t \( f \)”. However, we also have:

\[ s' \vdash \forall k \exists \ell (\ell > k \land \ell \in s'). \]

To see this, take a condition \( s'' \subseteq s' \) and \( k \) such that \( s'' \vdash k \downarrow \). Then \( s'' \setminus [1, k] \) is a condition, and so is \( s'' \setminus [1, \ell] \) for \( \ell := \min(s'' \setminus [1, k]) \). However, \( s'' \setminus [1, \ell] \vdash \ell \downarrow \), and since \( \ell > k \land \ell \in s' \) is a true \( \Delta_0 \) statement, Lemma 2.14 implies that it is forced by \( s'' \setminus [1, k] \) as well.

This completes the proof that \( \text{Cond, } \preceq, \vdash \text{ give a forcing interpretation of } WKL_0 + \text{RT}_2^2 \text{ in } RCA_0 + I \). Polynomiality of the interpretation is immediate if we assume that \( WKL_0 + \text{RT}_2^2 \) is finitely axiomatized. We may make this assumption w.l.o.g. because both \( WKL \) and \( \text{RT}_2^2 \) are single axioms, while \( RCA_0 \) can be axiomatized by using a finite number of instances of \( \Delta^0_1 \)-comprehension and \( \Sigma^0_1 \)-induction in such a way that a proof of any of the other instances can be constructed in polynomial time.

\[ \square \]

Lemma 2.16. The forcing interpretation of \( WKL_0 + \text{RT}_2^2 \) in \( RCA_0 + I \) given by \( \text{Cond, } \preceq, \vdash \) of Definition 2.11 is polynomially \( \forall \Sigma^0_2 \)-reflecting.

**Proof.** Let \( \varphi := \exists W \exists w \exists y \exists z \theta(W, w, y, z) \) be an \( \exists \Sigma^0_2 \) sentence. We sketch a proof in \( RCA_0 + I \) that, assuming \( \varphi \), there is a condition \( s \) such that \( s \not\vdash \forall W \forall w \exists y \exists z \lnot \theta \). On the basis of the sketch, it will be routine to verify that the proof can be constructed in polynomial time on input \( \varphi \).
Let $A,k$ be such that $\forall y \exists z \theta(A,k,y,z)$. By $\Sigma^0_1$ collection, for each $\ell$ there exists $m$ such that $\forall y \leq \ell \exists z \leq m \theta(A,k,y,z)$. Use primitive recursion to define a sequence of numbers by:

$$k_0 := k,$$
$$k_{n+1} := \text{smallest } m > k_n \text{ such that } \forall y \leq k_n \exists z \leq m \theta(A,k,y,z).$$

The axioms of $\text{RCA}_0$ imply the existence of the set $S = \{ k_n : n \in \mathbb{N} \}$, which is clearly infinite. Use the axiom (I3) of $\text{RCA} + I$ to obtain a finite set $s \subseteq S$ which is a forcing condition. Below, we abuse notation and write $A$ for the finite set $A |_{\max s}$.

We claim that $s \models \forall y \exists z \theta(A,k,y,z)$, which is enough to imply $s \not\Vdash \forall W \forall w \exists y \forall z \neg \theta$. Take $s' \triangleq s$ and $\ell$ such that $s' \models \ell \downarrow$. Since $s' \cap [1, \ell]$ is not a condition, $s' \setminus [1, \ell]$ must be one. Let $m_1 < m_2$ be the two smallest elements of $s' \setminus [1, \ell]$. By Proposition 2.4(iii) $s' \setminus [1, m_2]$ is also a condition and $(s' \setminus [1, m_2]) \Vdash m_2 \downarrow$. It follows from the definition of $s$ that $\forall y \leq m_1 \exists z \leq m_2 \theta(A,k,y,z)$. In particular, there is some $m \leq m_2$ such that $(s' \setminus [1, m_2]) \Vdash m \downarrow$ and $\theta(A,k,\ell,m)$. By Lemma 2.14 we get $(s' \setminus [1, m_2]) \Vdash \exists z \theta(A,k,\ell,m)$. This implies $(s' \setminus [1, m_2]) \Vdash \exists z \theta(A,k,\ell,z)$, which is what we wanted.

**Proof of Lemma 2.10** The statement of the Lemma follows directly from Lemma 2.15 Lemma 2.16 and Theorem 1.17.

**Proof of Theorem 2.1** By Lemma 2.10 there is a polynomial-time procedure which, given a $\text{WKL}_0 + \text{RT}_2^1$ proof $\pi$ of a $\forall \Sigma^0_2$ sentence $\psi$, outputs an $\text{RCA}_0 + I$ proof $\pi'$ of $\psi$. By Lemma 2.16 a further polynomial-time procedure outputs a proof $\pi''$ of $\psi$ in $\text{RCA}_0 + \Sigma^0_2$ and a proof $\pi'''$ of $\psi$ in $\text{RCA}_0 + \neg \Sigma^0_2$. Combine $\pi''$, $\pi'''$, and a case distinction to obtain a proof of $\psi$ in $\text{RCA}_0$.

**3 Ramsey for pairs: speedup over $\text{RCA}_0^*$**

We will now prove that Theorem 2.1 breaks down completely in the absence of $\Sigma^0_1$ induction, that is, if the base theory $\text{RCA}_0$ is replaced by $\text{RCA}_0^*$, although $\text{WKL}_0 + \text{RT}_2^1$ is also $\forall \Sigma^0_2$-conservative over $\text{RCA}_0^*$ [39].

**Theorem 3.1.** $\text{RT}_2^2$ has non-elementary speed-up over $\text{RCA}_0^*$ with respect to $\Sigma_1$ sentences.

We note that the speedup of $\text{RT}_2^2$ over $\text{RCA}_0^*$ can in fact be witnessed by $\Delta_0(\exp)$ sentences, though the proof involves a slightly larger amount of background than for the $\Sigma^0_1$ case. The idea of the argument is explained in the remark after the proof of Theorem 3.1.

The basic reason why Theorem 3.1 holds is expressed in the following lemma. Once the lemma is proved, the upper and lower bounds used to derive the theorem are obtained by more or less standard techniques described e.g. in the survey [25].

**Lemma 3.2.** $\text{RCA}_0^* + \text{RT}_2^2$ proves the following statement: “for every $k$, if each infinite set has a finite subset with $k$ elements, then each infinite set has a finite subset with $2^k$ elements.”
Proof. Clearly, it is enough to prove the statement with \(2^{k/2}\) substituted for \(2^k\). Working in \(\text{RCA}_0^* + \text{RT}_2^2\), let \(k\) be such that the infinite set \(A\) does not have a subset with \(2^{k/2}\) elements. W.l.o.g., we may assume that \(0 \in A\). We will use the fact that there is a 2-colouring of \([2^{k/2}]^2\) with no homogeneous set of size \(k\), which has a well-known probabilistic proof that easily formalizes in \(\text{RCA}_0^*\), to define a 2-colouring of \([n]^2\) such that every infinite homogeneous sets for the new colouring fails to have a \(k\)-element subset.

So, let \(f: [2^{k/2}]^2 \to 2\) have no homogeneous set of size \(k\). Let \(\{a_i: i \in I\}\), for some cut \(I\), be an increasing enumeration of \(A\). Note that \(2^{k/2} > I\). Define \(g: [n]^2 \to 2\) as follows:

\[
g(x, y) = \begin{cases} 
0 & \text{if there is } i < j \text{ s.t. } x \in [a_i, a_{i+1}), y \in [a_j, a_{j+1}), f(i, j) = 0, \\
1 & \text{if there is } i \text{ s.t. } x, y \in [a_i, a_{i+1}).
\end{cases}
\]

Let \(H\) be an infinite \(g\)-homogeneous set. By possibly thinning out \(H\), we may assume w.l.o.g. that for each \(i \in I\), there is at most one \(x \in H \cap [a_i, a_{i+1})\). This means that for any two \(x, y \in H\), the value \(g(x, y)\) is determined by the first clause of the definition of \(g\).

Let \(x_1, \ldots, x_\ell\) be the first \(\ell\) elements of \(H\), and let \(i_1, \ldots, i_\ell\) be such that \(x_j \in [a_{i_j}, a_{i_j+1})\). By our choice of \(H\) and the definition of \(g\), we see that \(f(i_j, i_{j'}) = g(x_j, x_{j'})\), so the set \(\{i_1, \ldots, i_\ell\}\) is homogeneous for \(f\). So \(\ell < k\) by our choice of \(f\). \(\square\)

Consider the simple variant of the iterated exponential function defined by \(2_0 = 1\) and \(2_{n+1} = 2^{2^n}\) for \(n \in \mathbb{N}\). To prove Theorem 3.1, we will make use of a family of sentences stating the existence of \(2_{2^n}\), for \(n\) a natural number. The relation \(y = 2_n\) has a \(\Delta_0\) definition which is well-behaved in \(= \Delta_0 - \text{see [11] \S 3.1}\), or see [17] Chapt. V.3(c)] for a similar idea applied to the more difficult case of \(y = 2^n\). Thus, \(\exists y(y = 2_{2^n})\) can be expressed as a \(\Sigma_1\) sentence of size \(O(\log n)\), for instance, by substituting the numeral for \(n\) into the fixed formula \(\exists y(y = 2_{2^n})\).

**Lemma 3.3.** In \(\text{RCA}_0^* + \text{RT}_2^2\), the sentence \(\exists y(y = 2_{2^n})\) has a proof of size polynomial in \(n\).

**Proof.** Consider the definable set

\[I = \{k: \text{every infinite set has a finite subset with } 2_k \text{ elements}\}. \]

In the absence of \(\Sigma_0^0\) induction, this is a proper initial segment of \(\mathbb{N}\) and, in general, not a set (i.e. a second-order object). However, Lemma 3.2 implies that \(I\) is a definable cut provably in \(\text{RCA}_0^* + \text{RT}_2^2\).

Consider also the set

\[A = \{\ell: \exists k (\ell = 2_k)\}. \]

Provably in \(\text{RCA}_0^*\), this is indeed a set, because the \(\exists k\) quantifier can be bounded by \(\ell\). By the totality of the exponential function, \(A\) is infinite. Moreover, since \(k \mapsto 2_k\) is an increasing monotone operation with a \(\Delta_0\)-definable graph, \(A\) has a subset with \(k\) elements exactly for those \(k\) for which \(2_k\) exists. It follows that for each \(k\), if \(I(k)\), then \(2_{2k}\) exists.
Now fix a natural number \( n \). By \[25\] Theorem 3.4.1, there are formulas \( I_0, I_1, \ldots, I_n \) such that \( I_0 \) is \( I \) and for each \( j < n \), the theory \( \text{RCA}_0^* + RT_2^j \) has a \( \text{poly}(n) \)-size proof that \( I_{j+1} \) is a cut and that \( \forall k \) \( (I_{j+1}(k) \rightarrow I_j(2^k)) \). By putting these proofs together, we get a \( \text{poly}(n) \)-size proof of \( I(2^n) \) and hence of the statement that \( 2_{2^n} \) exists.

\[ \square \]

**Lemma 3.4.** The size of the smallest \( \text{RCA}_0^* \) proof of \( \exists y(y = 2_{2^n}) \) grows faster than any elementary recursive function of \( n \).

**Proof.** In view of Theorem [16] in the next section, it suffices to show that the size of the shortest \( I\Delta_0 + \text{exp} \) proof of \( \exists y(y = 2_{2^n}) \) grows nonelementarily in \( n \). To this end, it is clearly enough to show the existence of a polynomial \( p \) such that, if the sentence \( \exists z \varphi(z) \) for a bounded formula \( \varphi \) has a proof in \( I\Delta_0 + \text{exp} \) of size \( n \), then \( \exists z \leq 2_{p(n)} \varphi(z) \) holds in \( \omega \).

So, assume that \( I\Delta_0 + \text{exp} \) has a size-\( n \) proof of \( \exists z \varphi(z) \) for \( \varphi \) bounded. We may assume that \( I\Delta_0 + \text{exp} \) is finitely axiomatized, because it has a finite axiomatization that polynomially simulates the more usual one in which \( \Delta_0 \) induction is a scheme. By standard upper bounds on the cut elimination procedure and its connection to Herbrand’s theorem \[25\] Section 5], there is a fixed polynomial \( p \) such that some Herbrand disjunction for \( \neg((I\Delta_0 + \text{exp}) \lor \exists z \varphi(z)) \) of size at most \( 2_{p(n)} \) is logically valid.

In more detail, what this means is as follows. Consider the conjunction of:

- \( \forall v_1 \ldots \forall v_k \delta(v_1, \ldots, v_k) \), a statement that induction holds for finitely many fixed \( \Delta_0 \) formulas with parameters among the \( v \); we may assume w.l.o.g. that all the quantifiers in \( \delta \) are bounded by one of the \( v_i \),

- \( \forall x \exists y(y = 2^x) \), with \( y = 2^x \) a \( \Delta_0 \) formula in the language of first-order arithmetic,

- \( \forall z \neg \varphi(z) \).

Put this conjunction in prenex normal form and skolemize it, obtaining the sentence \( \forall v \forall x \forall z \forall \bar{w} \exists \bar{v} \varphi(\bar{v}, x, z, \bar{w}) \), where the \( \forall \bar{v} \forall \bar{w} \) quantifiers are bounded and \( \varphi \) is quantifier-free (but contains the Skolem function symbols corresponding to the original existential quantifiers). Then, for some \( s \), there is a list of closed terms

\[
\begin{align*}
&l_{v_1, 1}, \ldots, l_{v_k, 1}, l_{x, 1}, l_{w_1, 1}, \ldots, l_{w_r, 1}, \ldots, l_{v_1, s}, \ldots, l_{v_k, s}, l_{x, s}, l_{z, s}, l_{w_1, s}, \ldots, l_{w_r, s}
\end{align*}
\]

(of the arithmetical language extended by the Skolem functions) such that the conjunction

\[
\Xi := \bigwedge_{i=1}^{s} \varphi(l_{v_1, i}, l_{x, i}, l_{z, i}, l_{w_i, i})
\]

has size at most \( 2_{p(n)} \) and is unsatisfiable.

We now interpret the function symbols appearing in \( \Xi \) as operations on \( \omega \). The symbols \( +, \cdot, 0, 1 \) are interpreted in the usual way. If \( f(\ldots, x, \ldots) \) is the Skolem function symbol corresponding to the \( \exists y \) quantifier \( f \) may have more arguments as a result of the move to prenex normal form), then we put \( f(\ldots, m, \ldots) = 2^m \). If \( f \) is the Skolem function symbol corresponding to a bounded existential quantifier \( \exists u \) in front of a subformula \( \psi \) of \( \delta \), \( \neg \varphi \), or \( y = 2^x \),
we let $f$ pick out the smallest witness to $\exists u \psi$ on those arguments for which there is such a witness; otherwise, $f$ returns 0.

Since each function interpreting a symbol in $\Xi$ increases its arguments at most exponentially, and $\Xi$ has size at most $2^{p(n)}$, it follows that under our interpretation each term in $\Xi$ has a value bounded by $2^{2^{p(n)}}$. Since we interpreted the Skolem function symbols as actual Skolem functions, and $I\Delta_0 + \text{exp}$ is a true theory, each substitution instance of the skolemizations of $\forall v_1 \ldots \forall v_k \delta(v_1, \ldots, v_k)$ and $\forall x \exists y (y = 2^x)$ appearing in $\Xi$ is true under our interpretation. However, $\Xi$ is unsatisfiable, which implies that some substitution instance of the skolemization of $\forall z \neg \varphi(z)$ must be false. Due to the way we defined our interpretation, this means that $\exists z \leq 2^{2^{p(n)}} \varphi(z)$ holds.

**Proof of Theorem 3.1.** Immediate from Lemmas 3.3 and 3.4.

**Remark.** As already mentioned, it is possible to witness the speedup of $\text{RCA}_0^i + \text{RT}^2_2$ over $\text{RCA}_0^i$ by a family of $\Delta_0(\text{exp})$ sentences. The sentences in question take the form $\text{Con}_{\Delta_0^i}((\text{I} \Delta_0 + \text{exp}))$, where $\text{Con}_\Sigma(T)$ is a formula stating in a natural way that there is no proof of contradiction in the theory $T$ of size at most $x$. Clearly, $\text{Con}_{\Sigma_2}(\Delta_0 + \text{exp})$ can be expressed by a $\Delta_0(\text{exp})$ sentence of size polynomial in $n$.

The so-called finitistic Gödel theorem, proved independently by Friedman and Pudlák (see [25, Theorem 6.3.2] and the references therein), implies that for a sufficiently strong finitely axiomatized $T$, the size of the smallest proof of $\text{Con}_\Delta(T)$ in $T$ is $n^{O(1)}$. As a consequence, the size of the smallest proof of $\text{Con}_{\Delta_2}(\Delta_0 + \text{exp})$ in $\Delta_0 + \text{exp}$ is $(2^n)^{O(1)}$. By Theorem 4.6 a proof of this sentence in $\text{RCA}_0^i$ also requires size $(2^n)^{O(1)}$.

On the other hand, essentially by formalizing the proof of Lemma 3.4 but with $\forall z \neg \varphi(z)$ replaced by $0 = 0$, one can prove in $\Delta_0 + \text{exp}$ that, for a certain fixed polynomial $p$, the existence of $2^{2^{p(n)}}$ implies $\text{Con}_{\Sigma_2}(\Delta_0 + \text{exp})$. By Lemma 3.3 $\text{RCA}_0^i + \text{RT}^2_2$ can prove the existence of $2^{2^{p(n)}}$, a number greater than $2^{2^{2^{p(n)}}}$, in size poly($n$). Thus, it can also prove $\text{Con}_{\Sigma_2}(\Delta_0 + \text{exp})$ in size poly($n$).

## 4 Induction versus collection

Since $\text{RT}^2_2$ proves the collection principle $B\Sigma_2$, Corollary 2.2 implies that $\Sigma_1$ polynomially simulates $B\Sigma_2$ w.r.t. proofs of $\Pi_2$ sentences. In this context, it is natural to ask whether the $\Pi_{n+1}$-conservativity of $B\Sigma_{n+1}$ over $\Sigma_n$ is also witnessed by a polynomial simulation for $n \neq 1$. A related problem was raised by Clote et al. [8, page 216]: is the conservativity provable in bounded arithmetic? Note that if a $\Pi_2$ sentence of the form $\forall x \exists y \varphi(x, y)$, such as a conservativity statement, is provable in bounded arithmetic, then the least witness for the $\exists y$ quantifier can be bounded by a polynomial-time function of $x$.

In this section, we present, for each $n \in \mathbb{N}$, a polynomial forcing interpretation $\tau(n)$ of $B\Sigma_{n+1} + \text{exp}$ in $\Sigma_n + \text{exp}$ that is polynomially $\Pi_{n+2}$-reflecting. The existence of these forcing interpretations, of course, implies that $\Sigma_n$ polynomially simulates $B\Sigma_{n+1}$ w.r.t. $\Pi_{n+2}$ sentences whenever $n \geq 1$; for $n = 0$, we get a polynomial simulation of $B\Sigma_1 + \text{exp}$ by $\Delta_0 + \text{exp}$. Our interpretations
can be formalized in (a fragment of) bounded arithmetic, which partially solves the problem from [8] as well: completely for \( n \geq 1 \), over exp for \( n = 0 \).

The only role of exp is to provide us with the usual universal \( \Sigma_{n+1} \) formula \( \Sigma_{n+1} + \text{Sat}(v, x) \). Recall that \( \Delta_0 + \text{exp} \) proves \( \theta(x) \leftrightarrow \Sigma_{n+1} - \text{Sat}(\bar{\theta}, [x]) \) for \( \theta \in \Sigma_{n+1} \), as well as the usual recursive conditions for a satisfaction relation restricted to \( \Sigma_{n+1} \) formulas; cf. [17] Chaps. I.1(d), V.5(b)].

**Remark.** The question from [8] as stated concerned provability in \( \Delta_0 + \Omega_1 \), which coincides with the \( L_1 \)-consequences of Buss’ theory \( S_2 \). The related Question 35 from the Clote-Krajíček list of open questions [10] – whether the stronger theory \( \Delta_0 + \text{exp} \) proves the mere equiconsistency of \( B \Sigma_{n+1} \) and \( \Sigma_n \) – was already answered positively in Hájek’s paper [16, Section 2] in the same volume where the list appeared.

**Remark.** While we were writing up our paper, we learnt that Fedor Pakhomov [private communication] independently proved that \( \Sigma_n \) polynomially simulates \( B \Sigma_{n+1} \) w.r.t. \( \Pi_{n+2} \) (and that bounded arithmetic proves this) by a different argument that avoids the need for exp in the case \( n = 0 \). Moreover, in an email discussion with us, Pakhomov found a parametric interpretation of \( B \Sigma_{n+1} + \text{exp} + \sigma \) in \( \Sigma_n + \text{exp} + \neg \sigma \) for each \( \Pi_{n+2} \) sentence \( \sigma \), at least for \( n > 0 \). It would be interesting to know whether his interpretation can be made independent of the \( \Pi_{n+2} \) sentence in question, and whether the use of parameters can be avoided.

The model-theoretic idea behind our forcing interpretations is simple and appeared already in [8] Lemma 12: given \( M \models \Sigma_n + \text{exp} \), if we construct a non-cofinal \( K \succ \Sigma_{n+1} M \), then \( M \models \Sigma_{n+1} \sup_K M \models B \Sigma_{n+1} + \text{exp} \). If the aim is only to show the \( \Pi_{n+2} \)-conservativity of \( B \Sigma_{n+1} + \text{exp} \) over \( \Sigma_n + \text{exp} \), one can build the extension \( K \) by any method, but to get a forcing interpretation a forcing construction is needed. The forcing construction we choose is that of a generic \( \Sigma_{n+1} \) ultrapower [22] pages 181ff.]; cf. the proof of Theorem B in [23]. In particular, the construction employs an ultrafilter \( \mathcal{U} \) on the Boolean algebra generated by the \( \Sigma_{n+1} \)-definable subsets of \( M \). Elements of the ultrapower \( K \) are represented by \( \Sigma_{n+1} \)-definable partial functions \( M \rightarrow M \) whose domain is an element of \( \mathcal{U} \). To ensure that the ultrapower is a non-cofinal extension, we require every element of \( \mathcal{U} \) to be cofinal in \( M \). It would be possible to first define a forcing relation \( \Vdash_{\tau(n)} \) for the ultrapower \( K \) and then define a second forcing relation \( \Vdash_{\tau'(n)} \) for the truncation \( \sup_K M \) in terms of \( \Vdash_{\tau(n)} \). To save some notation, we introduce the second one directly.

An alternative forcing interpretation, much closer in spirit to the one used to prove Theorem 23 and therefore perhaps easier to follow for the reader acquainted with Section 2, is mentioned in a remark after the proof of Theorem 4.3. However, that interpretation only works for \( n > 0 \).

In the remainder of this section, \( \exists^\infty x \) (“for infinitely many \( x \)” or “for cofinally many \( x \)” ) stands for \( \forall y \exists x \geq y \) and \( \forall^\infty x \) (“for almost all \( x \)” ) stands for \( \exists y \forall x \geq y \).

**Definition 4.1.** Let \( n \in \mathbb{N} \). Define a forcing translation \( \tau(n) \) from the language of first-order arithmetic to itself as follows.

Letting \( x \in \text{Ext}_{n+1}(s) \) stand for \( \Sigma_{n+1} \)-Sat\((s, [x])\), a \( \Sigma_{n+1} \) formula \( s \) will be in Cond\(_{(\tau(n))}\) if \( \exists^\infty x (x \in \text{Ext}_{n+1}(s)) \). For two conditions \( s, s' \), we let \( s' \subseteq_{\tau(n)} s \) if \( \text{Ext}_{n+1}(s') \subseteq \text{Ext}_{n+1}(s) \).
Names are also $\Sigma_{n+1}$ formulas. For a condition $s$ and name $v$, we say that $s \models_\tau(n) \downarrow_v$ if
\[
\forall x \in \text{Ext}_{n+1}(s) \exists y \Sigma_{n+1}\text{-Sat}(v, [x, y]) \\
\wedge \exists d \forall x \in \text{Ext}_{n+1}(s) \exists y \leq d \Sigma_{n+1}\text{-Sat}(v, [x, y]).
\]
For a simple atomic formula $\alpha(v_1, \ldots, v_\ell)$, we say that $s \models_\tau(n) \alpha(\bar{v})$ if $s \models_\tau(n) \downarrow_{\bar{v}}$ and
\[
\forall x \in \text{Ext}_{n+1}(s) \exists y_1, y_2, \ldots, y_\ell \left( \bigwedge_{i=1}^\ell \Sigma_{n+1}\text{-Sat}(v_i, [x, y_i]) \wedge \alpha(\bar{y}) \right).
\]
When there is no risk of ambiguity, we will often omit the subscript $\tau(n)$.

In other words, conditions are infinite $\Sigma_{n+1}$-definable sets and names are (to be viewed as) $\Sigma_{n+1}$-definable partial functions. A condition $s$ forces $v \downarrow$ if $v$ has a unique value on almost all elements of $s$ and if there is a common upper bound on almost all of those values.

It is known that $\Sigma_{n+1}$ ultrapowers of models of $\Sigma_n$ satisfy L"os’s Theorem for $\Sigma_{n+1}$ and $\Pi_{n+1}$ formulas [22, Lemma 5.13]. This ensures $\Sigma_{n+2}$-elementarity between the base model $M$ and the ultrapower $K$, although, as remarked earlier, $\Sigma_{n+1}$-elementarity is already enough to run our argument.

We transform these ideas into a syntactic proof in the following.

**Proposition 4.2.** There is a polynomial-time procedure which, given $n \in \mathbb{N}$ in unary and a formula $\theta(y_1, y_2, \ldots, y_\ell)$ in $\Sigma_n \cup \Pi_n$, outputs a proof of
\[
\forall s \in \text{Cond}_\tau(n) \forall v_1, v_2, \ldots, v_\ell \left( \begin{array}{l}
\text{“} s \models_\tau(n) \theta(\bar{v}) \text{”} \\
\leftrightarrow s \models_\tau(n) \downarrow_{\bar{v}} \wedge \forall x \in \text{Ext}_{n+1}(s) \exists y \left( \bigwedge_{i=1}^\ell \Sigma_{n+1}\text{-Sat}(v_i, [x, y_i]) \wedge \theta(\bar{y}) \right)
\end{array} \right)
\]
in $\Sigma_n + \exp$.

**Proof.** Just like the usual proof of L"os’s Theorem for $\Sigma_{n+1}$ ultrapowers [22, Lemma 5.13], the construction splits into two steps.

(1) For $\Delta_0$ formulas $\theta$, we build the required proof by induction on the structure of $\theta$.

(2) Beyond that, we build the proof using induction on the number of unbounded quantifier blocks in $\theta$.

Of course, the step for a single quantifier block in (2) itself splits into smaller steps corresponding to individual quantifiers.

In the remainder of the argument, we think of $n \in \mathbb{N}$ as fixed. We leave it to the reader to verify that the proofs in $\Sigma_n + \exp$ described below can be found in time polynomial in the given parameters. The polynomial dependence on $n$ rather than just on $|\theta|$ cannot be avoided: it comes for instance from the

---

1To get the required amount of elementarity, we follow Enayat–Mohsenipour [13, Theorem 2.1] instead of Clote et al. [8, Lemma 11].
universal \( \Sigma_n+1 \) formula, which is already used in the definition of \( \text{Ext}_{n+1} \), and thus of \( \text{Cond}_\tau(s) \) and \( \models \tau(s) \).

In the construction for (1), the base step is for an atomic formula \( \theta(\bar{v}) \). If \( \theta \) is simple atomic, the required equivalence follows immediately from the definition of \( \models \tau(s) \). For a non-simple atomic \( \theta \), the argument is similar to the one in Lemma 24.10. By induction on the structure of a term \( t(\bar{u}) \) (a subtuple of \( \bar{v} \)) appearing in \( \theta \), we build a proof that, for an appropriate fresh name \( w \), \( s \models t[w] \) if and only if the following happens: \( s \models \bar{u}_\bullet \), \( s \not\models w_\bullet \), and for almost all \( x \in \text{Ext}_{n+1}(s) \) the values \( \bar{y}, \bar{z} \) of \( u, w \) at \( x \) exist and satisfy \( t(\bar{y}) = \bar{z} \). Once this is proved for each term in \( \theta \), the required equivalence is proved using (FT7)Item.7 (FT6)Item.6 and the definition of \( \models \tau(s) \).

The inductive steps of the construction for (1) are more or less based on the usual proof of Łoś’s Theorem with the satisfaction relation for the ultrapower replaced by the forcing relation. As a demonstration, we show the case for negation.

Suppose \( \theta \) is \( \neg \eta \) for \( \eta \in \Delta_0 \). The inductive assumption gives us a proof that if \( s \) and \( \bar{v} \) are such that \( s \models \bar{v}_\bullet \), then \( s \models \eta(\bar{v}) \) exactly if for almost all \( x \in \text{Ext}_{n+1}(s) \) the values of the \( v_i \)'s at \( x \) exist and satisfy \( \eta \).

Take any \( s \in \text{Cond} \) and any \( \bar{e} \) such that \( s \models \bar{e}_\bullet \). If \( s \not\models \neg \eta(\bar{e}) \), then (FT5)Item.5 gives us \( s' \not\models s \) such that \( s' \models \eta(\bar{e}) \), and so, by the proof from the induction hypothesis,

\[
\forall^{\infty} x \in \text{Ext}_{n+1}(s') \exists \bar{y} \left( \bigwedge_{i=1}^{\ell} \Sigma_{n+1}\text{-Sat}(e_i, [x, y_i]) \land \eta(\bar{y}) \right),
\]

hence also

\[
\exists^{\infty} x \in \text{Ext}_{n+1}(s) \exists \bar{y} \left( \bigwedge_{i=1}^{\ell} \Sigma_{n+1}\text{-Sat}(e_i, [x, y_i]) \land \eta(\bar{y}) \right).
\]

Conversely, suppose \( s \not\models \neg \eta(\bar{e}) \). By (FT5)Item.5 and the proof from the induction hypothesis,

\[
\forall s' \preceq s \exists^{\infty} x \in \text{Ext}_{n+1}(s') \forall \bar{y} \left( \bigwedge_{i=1}^{\ell} \Sigma_{n+1}\text{-Sat}(e_i, [x, y_i]) \rightarrow \neg \eta(\bar{y}) \right).
\]

Since the conjunction of the formula \( s \) with \( \exists \bar{y} \left( M^{\ell}_{i=1} \Sigma_{n+1}\text{-Sat}(e_i, [x, y_i]) \land \eta(\bar{y}) \right) \) is equivalent to some \( \Sigma_{n+1} \) formula \( s' \), this implies

\[
\neg \exists^{\infty} x \in \text{Ext}_{n+1}(s) \exists \bar{y} \left( \bigwedge_{i=1}^{\ell} \Sigma_{n+1}\text{-Sat}(e_i, [x, y_i]) \land \eta(\bar{y}) \right),
\]

so, given that \( s \models \bar{e}_\bullet \),

\[
\forall^{\infty} x \in \text{Ext}_{n+1}(s) \exists \bar{y} \left( \bigwedge_{i=1}^{\ell} \Sigma_{n+1}\text{-Sat}(e_i, [x, y_i]) \land \neg \eta(\bar{y}) \right).
\]

If \( n = 0 \), then (1) already suffices. So suppose \( n > 0 \). In the construction for (2), we also imitate the usual proof of Łoś’s Theorem, but additionally we
need to take care of the fact that our anticipated ultrapower is truncated. The bounds involved will be provided by strong \( \Sigma_n \) collection, which is the principle that gives a \( \Sigma_n \) formula \( \varphi(x, y) \) and a bound \( a \), there is a bound \( b \) such that for all \( x < a \), if any \( y \) satisfying \( \varphi(x, y) \) exists, then some such \( y \) may be found below \( b \). It is well-known that for \( n > 0 \) this is equivalent to \( \varphi(x, y) \) if and only if for almost all \( x \in \text{Ext}_{n+1}(s) \) the values of the \( v_i \)'s and \( w_j \)'s at \( x \) exist and satisfy \( \eta \).

Take any \( s \in \text{Cond} \) and any \( \bar{e} \) such that \( s \models \bar{e} \). Assume

\[
\exists^\infty x \in \text{Ext}_{n+1}(s) \, \forall \bar{y} \, \left( \bigwedge_{i=1}^{\ell} \text{\texttt{Sat}}(e_i, [x, y_i]) \rightarrow \neg \forall \bar{w} \, \eta(\bar{y}, \bar{w}) \right).
\]

Let \( s' \) be a canonical \( \Sigma_{n+1} \) formula equivalent to the conjunction of \( s \) with

\[
\exists \bar{y} \left( \bigwedge_{i=1}^{\ell} \text{\texttt{Sat}}(e_i, [x, y_i]) \wedge \exists \bar{w} \, \neg \eta(\bar{y}, \bar{w}) \right).
\]

Then \( s' \in \text{Cond} \) by assumption. Recall that \( s \models \bar{e} \). Find \( d \) such that

\[
\bigwedge_{i=1}^{\ell} \forall^\infty x \in \text{Ext}_{n+1}(s) \, \exists y \leq d \, \Sigma_{n+1}\text{-Sat}(e_i, [x, y]).
\]

Suppose \( \bar{w} = w_1, w_2, \ldots, w_{\ell'} \) where \( \ell' \geq 1 \). For each \( j \in \{1, 2, \ldots, \ell'\} \), find a canonical \( \Sigma_{n+1} \) formula with free variables \( x, z \) that is equivalent to

\[
\exists \bar{y} \exists \bar{w} \left( \bigwedge_{i=1}^{\ell} \Sigma_{n+1}\text{-Sat}(e_i, [x, y_i]) \wedge \neg \eta(\bar{y}, \bar{w}) \wedge \forall \langle \bar{w}' \rangle \eta(\bar{y}, \bar{w}') \wedge z = w_j \right).
\]

Let this \( \Sigma_{n+1} \) formula be (coded by the number) \( e_j' \). Since \( \text{I} \Sigma_k \) holds, we know

\[
\bigwedge_{j=1}^{\ell'} \forall^\infty x \in \text{Ext}_{n+1}(s) \, \exists z \, \Sigma_{n+1}\text{-Sat}(e_j', [x, z]).
\]

Use strong \( \Pi_k \) collection to obtain \( d' \) such that

\[
\forall \bar{y} \leq d \left( \exists \bar{w} \, \neg \eta(\bar{y}, \bar{w}) \rightarrow \exists \langle \bar{w}' \rangle \leq d' \, \neg \eta(\bar{y}, \bar{w}) \right).
\]

As the reader can verify, this ensures

\[
\bigwedge_{j=1}^{\ell'} \forall^\infty x \in \text{Ext}_{n+1}(s') \, \exists z \leq d' \, \Sigma_{n+1}\text{-Sat}(e_j', [x, z]).
\]

Thus \( s' \models \bar{e}' \). If \( s'' \not\models s' \), then the proof from the induction hypothesis gives us \( s'' \models \neg \eta(\bar{e}, \bar{e}') \) and so \( s'' \not\models \eta(\bar{e}, \bar{e}') \) by Lemma 10. Hence (FT7)Item.7 tells us \( s \not\models \forall \bar{w} \, \eta(\bar{e}, \bar{w}) \).

\[]

**Remark.** For \( n > 0 \), the argument in the proof above relies on induction axioms in two places. Speaking in model-theoretic terms, first we use at most \( \text{I} \Sigma_{n-1} \)
to guarantee the existence of sufficiently many Skolem functions when proving
Loš’s Theorem for the ultrapower $K$, and then we use strong $\Sigma_n$ collection to
pass on the elementarity from $K$ to $\sup_K M$. For $n = 0$, the latter application
is not needed (and we do not have strong collection anyway), but we use full
$\Sigma_n$ to get the Loš Theorem we want.

In the usual argument proving $\Pi_{n+1}$ conservativity of $\BS_{n+1}$ of $\Sigma_n$ [17]
page 229 it is easy to deduce from the analogue of Proposition [12] that the truncated
ultrapower $\sup_K M$ is a $\Sigma_{n+1}$-elementary extension of the base model $M$.
In the model-theoretic setting, this directly implies $\sup_K M \models \Delta_0$. When
formulated in terms of forcing interpretations, elementarity becomes reflection, but
being reflecting in the sense of Definition [14] is not enough to ensure that every
condition forces $\Delta_0$, which is required by (FT13)Item.25. So we strengthen the
notion of reflection in the statement below. From this, one quickly derives exp
using the cofinality of $M$ inside $\sup_K M$ and the $\Delta_0$-provable monotonicity of the
exponential function. The model-theoretic idea behind the interpretation of $\BS_{n+1}$ is a standard underspill argument used to show that a $\Sigma_n$-elementary proper initial segment of $K \models \Sigma_n$ satisfies $\BS_{n+1}$. When $n > 0$, it is actually enough to assume $K \models \Sigma_{n-1} [7]$ Proposition 3], but we do not need this for our proof.

**Theorem 4.3.** For every $n \in \mathbb{N}$, the forcing translation $\tau(n)$ is a polynomial
forcing interpretation of $\BS_{n+1} + \exp$ in $\Sigma_n + \exp$ that is polynomially $\Pi_{n+2}$
reflecting.

In fact the $\tau(n)$ are polynomially $\Pi_{n+2}$-reflecting in the following strengthened and uniform sense: there is a single polynomial-time procedure which, given $n \in \mathbb{N}$ in unary and a $\Pi_{n+2}$ sentence $\gamma$, outputs a proof of

$$\exists s \in \Cond_{\tau(n)}(s \models \tau(n) \gamma) \rightarrow \gamma$$

in $\Sigma_n + \exp$.

**Proof.** As in the proof of Proposition [12] we think of $n \in \mathbb{N}$ as fixed and leave it to the reader to check that the running time of a procedure constructing the proofs in $\Sigma_n + \exp$ described below can be bounded by a polynomial in $n$ and $|\gamma|$.

**Strengthened reflection.** Let $\gamma$ be a $\Pi_{n+2}$ sentence of the form $\forall \bar{u} \exists \bar{v} \theta(\bar{u}, \bar{v})$, where $\theta(u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_{r})$ is $\Pi_n$. Let $s$ be such that $s \models \forall \bar{u} \exists \bar{v} \theta(\bar{u}, \bar{v})$. Take any $a_1, a_2, \ldots, a_k$. Then, as one can verify using [FT17]Item.7 and the
definition of $\Vdash_{\tau(n)}$, we have $s \models \exists \bar{v} \theta(\bar{a}, \bar{v})$, where each $a_i$ is the (code of the) the $\Delta_0$ formula $x = x \land y = a_i$. From this, we deduce that $\forall s' \leq s \exists s'' \leq s' \exists \bar{v} (s'' \models \theta(\bar{a}, \bar{v}))$.

We then apply Proposition [12] to get

$$\forall s' \leq s \exists s'' \leq s' \exists \bar{v} \forall \bar{x} \in \Ext_{n+1}(s'') \exists \gamma \left( \bigwedge_{j=1}^{r} \Sigma_{n+1} - \Sat(v_j, [x, y_j]) \land \theta(\bar{a}, \bar{y}) \right).$$

In particular, we know $\exists \bar{v} \theta(\bar{a}, \bar{v})$ holds.

**Interpretation of $\Delta_0 + \exp$.** The $\Delta_0$ part follows easily from strengthened reflection for $\Sigma_1$ sentences. For exp, simply prove that for any $s \in \Cond$ and any $e$, if $s \models e'$, then $s \models e' = 2^e$, where $e'$ is the $\Sigma_{n+1}$ formula

$$\exists z (\Sigma_{n+1} - \Sat(\mathcal{G}, [x, y]) \land y = 2^z).$$
**Interpretation of $B\Sigma_{n+1}$.** By induction on $m \leq n$, we construct proofs of

$$\forall s \in \text{Cond } (s \models BI_m).$$

This is sufficient by Proposition [1.14] and the well-known equivalence between $BI_n$ and $B\Sigma_{n+1}$. Take any $m \leq n$. In the case when $m \neq 0$, suppose that we already have a proof of $\forall s \in \text{Cond } (s \models BI_{m-1})$. Consider a $\Pi_m$ formula $\theta(u, v, w)$. Take $e$ and some $s \in \text{Cond}$ which forces $\forall u \leq v \exists v \theta(u, v, e)$. Then any element $v^*$ in the ultrapower $K$ that lies above the truncated part $\text{sup}_K M$ can be used to bound the $\exists v$ quantifier, i.e.,

$$\forall v^* \left( \forall^\infty x \in \text{Ext}_{n+1}(s) \exists y \Sigma_{n+1}\text{-Sat}(v^*, [x, y]) \land \forall d \forall^\infty x \in \text{Ext}_{n+1}(s) \exists y > d \Sigma_{n+1}\text{-Sat}(v^*, [x, y]) \rightarrow \forall^\infty x \in \text{Ext}_{n+1}(s) \exists y^*, z \Sigma_{n+1}\text{-Sat}(v^*, [y^*]) \land \Sigma_{n+1}\text{-Sat}(e, [x, z]) \land \forall u \leq v \exists v \leq y^* \theta(u, v, z) \right). \quad (6)$$

To see why this holds formally, first one seeks a $\models$-free description of what $s \models \forall u \leq v \exists v \theta(u, v, \cdot)$ means, using Proposition [1.12] ([FT1]Item.7, [FT6]Item.6) and basic facts about forcing interpretations from Section 4 (note that we already know that $\tau(n)$ is a polynomial forcing interpretation of some theory, and all that remains is to verify [FT13]Item.25 for $B\Sigma_n$). Unlike in (6), this leaves the $\forall u$ and the $\exists v$ quantifiers outside the big brackets. To push the $\exists v$ quantifier inside, use the bound $d$ in the definition of $s \models v_j$. To push the $\forall u$ quantifier inside, run one step of the proof of Loś’s Theorem for the ultrapower $K$. Since this part is similar to what we did for Proposition [1.2] we leave the details to the reader.

Feeding, for instance, $v^*$ equal to $\gamma x = y^\gamma$ – in other words, the diagonal element of the ultrapower – into (6), we see that for almost all $x \in \text{Ext}_{n+1}(s)$, we have some $y^*$ such that $\forall u \leq z \exists v \leq y^* \theta(u, v, z)$, where $z$ is the unique number satisfying $\Sigma_{n+1}\text{-Sat}(e, [x, z])$. By $\Sigma_m$-induction, we can find the smallest such $y^*$, which shows

$$\forall^\infty x \in \text{Ext}_{n+1}(s) \exists y \Sigma_{n+1}\text{-Sat}(e', [x, y]),$$

where $e'$ is a canonical $\Sigma_{n+1}$ formula equivalent over $I\Sigma_n$ to

$$\exists z \left( \Sigma_{n+1}\text{-Sat}(e, [x, z]) \land \forall u \leq z \exists v \leq y \theta(u, v, z) \land \forall y' < y^- \forall u \leq z \exists v \leq y^* \theta(u, v, z) \right).$$

In view of ([FT5]Item.5) and ([FT7]Item.7) it suffices to find $s' \preceq s$ which satisfies

$$\exists d \forall^\infty x \in \text{Ext}_{n+1}(s') \exists y \leq d \Sigma_{n+1}\text{-Sat}(e', [x, y]),$$

because this will imply $s' \models e' \downarrow$ and thus $s' \models \forall u \leq v \exists v \leq e' \theta(u, v, e)$ by Proposition [1.2]. Note that Proposition [1.2] applies only to $\Sigma_n$ or $\Pi_n$ formulas, but we can treat $\forall u \leq z \exists v \leq y \theta(u, v, z)$ as a $\Pi_m$ formula even when $m \geq 1$, thanks to Proposition [1.14] and the induction hypothesis that $s \models BI_{m-1}$.

We obtain the requisite $d$ by an underspill argument formalized using [6]. If there are infinitely many $x \in \text{Ext}_{n+1}(s)$ such that $\Sigma_{n+1}\text{-Sat}(e', [x, 0])$, then we are easily done. So suppose not. Let $e''$ be the $\Sigma_{n+1}$ formula

$$\Sigma_{n+1}\text{-Sat}(e'' [x, y + 1]).$$
Then by our supposition and the fact that \( \forall^\infty x \in \text{Ext}_{n+1}(s) \exists y \Sigma_{n+1}-\text{Sat}(e', [x, y]) \), we have
\[
\forall^\infty x \in \text{Ext}_{n+1}(s) \exists y \Sigma_{n+1}-\text{Sat}(e'', [x, y]).
\]
Moreover, the minimality condition in the formula \( e' \) implies
\[
\forall^\infty x \in \text{Ext}_{n+1}(s) \forall y, z \left( \Sigma_{n+1}-\text{Sat}(e'', [x, y]) \land \Sigma_{n+1}-\text{Sat}(e', [x, z]) \rightarrow \neg \forall u \leq z \exists v \leq y \theta(u, v, z) \right).
\]
So, by (1) applied to \( v^* := e'' \), there must be a \( d \) such that for infinitely many \( x \in \text{Ext}_{n+1}(s) \) we have \( \forall y > d \neg \Sigma_{n+1}-\text{Sat}(e'', [x, y]) \). Unravelling the definitions, this tells us
\[
\exists d \exists^\infty x \in \text{Ext}_{n+1}(s) \forall y > d \neg \Sigma_{n+1}-\text{Sat}(e', [x, y]).
\]
Fix such \( d \). Since \( \forall^\infty x \in \text{Ext}_{n+1}(s) \exists y \Sigma_{n+1}-\text{Sat}(e', [x, y]) \), we know
\[
s' \leq s \land \forall^\infty x \in \text{Ext}_{n+1}(s') \exists y \leq d \Sigma_{n+1}-\text{Sat}(e', [x, y]),
\]
where \( s' \) is some canonical \( \Sigma_{n+1} \) formula equivalent to
\[
x \in \text{Ext}_{n+1}(s) \land \exists y \leq d \Sigma_{n+1}-\text{Sat}(e', [x, y]).
\]

Remark. For \( n > 0 \), there is an alternative polynomially \( \Pi_{n+2} \)-reflecting forcing interpretation based on a generic cut construction similar to the one from Section 2. We work in a version \( \Delta_n^{\Pi} \) in which the axiom (3) is changed to “every infinite \( \Sigma_{n-1} \)-definable set contains an \( \omega^2 \)-large subset for some \( x > 1 \).” (The axioms (1)–(3) can then be eliminated by some variant of Lemma 2.4) as in Section 2. The role of infinite sets is now played by infinite \( \Sigma_{n-1} \)-definable sets. The forcing conditions are finite sets \( s \) (which now include e.g. all bounded \( \Delta_n \)-definable sets) that are \( \omega^2 \)-large for some \( x > 1 \) and, for \( n \geq 2 \), have the property that if \( \ell_1 < \ell_2 \) are elements of \( s \), then the Skolem function for the first \( 3 \) quantifier of the \( \Sigma_{n-1} \) universal formula takes values below \( \ell_2 \) on inputs below \( \ell_1 \). The rest of the argument is along the lines of Section 2, except that the step for \( \Pi_3 \) in the proof of Lemma 2.16 is replaced by the considerably easier step for the infinite \( \Delta_0 \) pigeonhole principle (a.k.a. \( \mathcal{B}_2 \)) relative to \( 0^{(n-1)} \).

In an analogue of this argument for \( n = 0 \), the conditions would be finite sets with more than \( 1 \) elements that are at least exponentially far apart from each other. However, the proof of Lemma 2.10 (reflection) no longer works. The proof of the reflection lemma makes use of the fact that for any element \( k \) there is a condition \( s \) such that \( \min s \geq k \). But in a model of \( \Pi_0 + \text{exp} \) there might be an element above which the exponential function may be iterated only a standard \( (a \text{ fortiori}, \text{ no greater than } \uparrow) \) number of times.

For each fixed \( n \in \mathbb{N} \), the proof of Theorem 1.3 can be formalized in the theory \( \text{PV} \), a fragment of bounded arithmetic corresponding to polynomial-time computation. As already mentioned, this provides a solution to a problem of Clote et al. [8] for \( n \geq 1 \) and a partial solution for \( n = 0 \).

Corollary 4.4. For each \( n \in \mathbb{N} \), the theory \( \Delta_n^{\Pi} + \text{exp} \) polynomially simulates \( B\Sigma_{n+1}^{\text{exp}} \) with respect to \( \Pi_{n+2} \) sentences. Moreover, the \( \Pi_{n+2} \)-conservativity of \( B\Sigma_{n+1}^{\text{exp}} \) over \( \Sigma_n^{\text{exp}} \) is provable in \( \text{PV} \).
Proof. The first part follows directly from Theorem 4.3 and Theorem 1.17. The second part is obtained by formalizing the proof of Theorem 4.3 for a fixed $n \in \mathbb{N}$ in PV.

In fact, the proof of Theorem 4.3 is witnessed by a single polynomial-time algorithm that takes $n$ (in unary) as input, and this algorithm can also be formalized in PV. As a result, we can conclude that the $\Pi_{n+2}$-conservativity of $B\Sigma_{n+1} + \text{exp}$ over $I\Sigma_n + \text{exp}$ holds provably in PV for any $n$ for which the axioms of $B\Sigma_{n+1} + \text{exp}$ over $I\Sigma_n + \text{exp}$ actually exist as finite strings – that is, for any $n$ in the definable cut Log, the domain of the exponential function.

Corollary 4.5. PV proves that for every $n \in \text{Log}$, the theory $B\Sigma_{n+1} + \text{exp}$ is $\Pi_{n+2}$-conservative over $I\Sigma_n + \text{exp}$.

As the reader may have noticed, one can actually extract from the truncated generic ultrapower construction (alternatively, from the generic cut construction for $n > 0$) a forcing interpretation of $\text{WKL}_0^* + B\Sigma^0_{n+1}$ in $I\Sigma_n + \text{exp}$ that is polynomially $\Pi_{n+2}$-reflecting. In fact, one can interpret more, as the following theorem shows.

Theorem 4.6. For each $n \in \mathbb{N}$, the theory $I\Sigma_{n} + \text{exp}$ polynomially simulates $\text{WKL}_0^* + B\Sigma^0_{n+1}$ with respect to $\Pi_{n+2}$ sentences.

Proof. First, notice there exists a polynomial forcing interpretation of $\text{WKL}_0^* + B\Sigma^0_{n+1}$ in $I\Sigma_n + \text{exp}$ that polynomially reflects $L_1$ sentences. For $n \geq 1$, this follows from Theorem 3.13(a) in Hájek [10]. If $n = 0$, then such a forcing interpretation can be extracted from Sections 5–8 in Avigad [1]. Then compose with $\tau(n)$ and invoke Theorem 1.17.

References

[1] Jeremy Avigad. Formalizing forcing arguments in subsystems of second-order arithmetic. *Annals of Pure and Applied Logic*, 82(2):165–191, 1996.

[2] Lev D. Beklemishev. A proof-theoretic analysis of collection. *Archive for Mathematical Logic*, 37(5–6):275–296, July 1998.

[3] Patrick Caldon and Aleksandar Ignjatović. On mathematical instrumentalism. *J. Symbolic Logic*, 70(3):778–794, 2005.

[4] Peter A. Cholak, Carl G. Jockusch, Jr., and Theodore A. Slaman. On the strength of Ramsey’s theorem for pairs. *J. Symbolic Logic*, 66(1):1–55, 2001.

[5] C. T. Chong, Theodore A. Slaman, and Yue Yang. The metamathematics of Stable Ramsey’s Theorem for Pairs. *J. Amer. Math. Soc.*, 27(3):863–892, 2014.

[6] C. T. Chong, Theodore A. Slaman, and Yue Yang. The inductive strength of Ramsey’s Theorem for Pairs. *Adv. Math.*, 308:121–141, 2017.

[7] Peter G. Clote. Partition relations in arithmetic. In Carlos Augusto Di Prisco, editor, *Methods in Mathematical Logic*, volume 1130 of *Lecture Notes in Mathematics*, pages 32–68, Berlin, 1985. Springer-Verlag.
[8] Peter G. Clote, Petr Hájek, and Jeff B. Paris. On some formalized conservation results in arithmetic. *Archive for Mathematical Logic*, 30(4):201–218, 1990.

[9] Peter G. Clote and Jan Krajíček, editors. *Arithmetic, Proof Theory, and Computational Complexity*, volume 23 of *Oxford Logic Guides*. Clarendon Press, Oxford, 1993. Papers from the conference held in Prague, July 2–5, 1991.

[10] Peter G. Clote and Jan Krajíček. Open problems. In *Arithmetic, Proof Theory, and Computational Complexity* [9], pages 1–19.

[11] Paola D’Aquino. A sharpened version of McAloon’s theorem on initial segments of models of I∆₀. *Annals of Pure and Applied Logic*, 61(1–2):49–62, May 1993.

[12] Elmar Eder. *Relative complexities of first order calculi*. Artificial Intelligence. Vieweg & Sohn, Braunschweig, 1992.

[13] Ali Enayat and Shahram Mohsenipour. Model theory of the regularity and reflection schemes. *Archive for Mathematical Logic*, 47(5):447–464, 2008.

[14] Herbert B. Enderton. *A Mathematical Introduction to Logic*. Harcourt/Academic Press, San Diego, second edition, 2001.

[15] Kurt Gödel. Über die Länge von Beweisen. In *Ergebnisse eines mathematischen Kolloquiums, Heft 7*, pages 23–24. Reprinted with English translation in Kurt Gödel, *Collected Works*, Vol. 1, pp. 396-399, Oxford University Press, 1986.

[16] Petr Hájek. Interpretability and fragments of arithmetic. In Clote and Krajíček [9], pages 185–196.

[17] Petr Hájek and Pavel Pudlák. *Metamathematics of First-Order Arithmetic*. Perspectives in Mathematical Logic. Springer-Verlag, Berlin, 1993.

[18] Denis R. Hirschfeldt. *Slicing the truth. On the computable and reverse mathematics of combinatorial principles*, volume 28 of *Lecture Notes Series. Institute for Mathematical Sciences. National University of Singapore*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2015.

[19] Aleksandar Djordje Ignjatović. *Fragments of First and Second Order Arithmetic and Lengths of Proofs*. PhD thesis, University of California at Berkeley, 1990.

[20] Jussi Ketonen and Robert Solovay. Rapidly growing Ramsey functions. *Ann. of Math. (2)*, 113(2):267–314, 1981.

[21] Leszek Aleksander Kołodziejczyk and Keita Yokoyama. Some upper bounds on ordinal-valued Ramsey numbers for colourings of pairs, 2018. Submitted. Available at arxiv.org/abs/1807.00616.

[22] Thomas G. McLaughlin. Sub-arithmetic ultrapowers: a survey. *Annals of Pure and Applied Logic*, 49(2):143–191, October 1990.
[23] J. B. Paris and L. A. S. Kirby. $\Sigma_n$-collection schemas in arithmetic. In Logic Colloquium '77 (Proc. Conf., Wroclaw, 1977), volume 96 of Stud. Logic Foundations Math., pages 199–209, 1978.

[24] Ludovic Patey and Keita Yokoyama. The proof-theoretic strength of Ramsey’s theorem for pairs and two colors. Adv. Math., 330:1034–1070, 2018.

[25] Pavel Pudlák. The lengths of proofs. In Samuel R. Buss, editor, Handbook of Proof Theory, volume 137 of Studies in Logic and the Foundations of Mathematics, chapter VIII, pages 547–637. Elsevier, Amsterdam, 1998.

[26] David Seetapun and Theodore A. Slaman. On the strength of Ramsey’s theorem. Notre Dame Journal of Formal Logic, 36(4):570–582, 1995.

[27] Stephen G. Simpson. Subsystems of Second Order Arithmetic. Perspectives in Mathematical Logic. Springer-Verlag, 1999. XIV + 445 pages; Second Edition, Perspectives in Logic, Association for Symbolic Logic, Cambridge University Press, 2009, XVI+ 444 pages.

[28] Stephen G. Simpson and Rick L. Smith. Factorization of polynomials and $\Sigma^0_1$ induction. Annals of Pure and Applied Logic, 31:289–306, 1986.

[29] Albert Visser. An inside view of EXP; or, the closed fragment of the provability logic of $I\Delta_0 + \Omega_1$ with a propositional constant for EXP. The Journal of Symbolic Logic, 57(1):131–165, March 1991.

[30] Keita Yokoyama. On the strength of Ramsey’s theorem without $\Sigma_1$-induction. Math. Log. Q., 59(1-2):108–111, 2013.